Vol. 15 / N. 1 / January - December 2025

Official Journal of the Italian Society of Physiotherapy (SIF)

ARCHIVES OF PHYSIOTHERAPY

ABOUTSCIENCE

GENERAL INFORMATION

Aims and Scope
Archives of Physiotherapy (AoP) is a Diamond open access journal disseminating research covering all aspects of physiotherapy. Featured topics include clinical trials across various disability domains (musculoskeletal, neurological, cardio-pulmonary), diagnostic studies measurement-focused research, observational studies, prognostic investigations, qualitative studies, systematic reviews, meta-analyses, and case reports.

In addition to research articles, the AoP board also invites expert and eminent authors to contribute to the Masterclass series.

Abstracting and Indexing
CINAHL
CrossRef
DOAJ
Google Scholar
Jisc Open Policy finder
OCLC WorldCat
Opac-ACNP (Catalogo Italiano dei Periodici)
PubMed
ROAD (Directory of Open Access Scholarly Resources)
Scilit
Scimago
Scopus

Publication process
Peer review
Papers submitted to Archives of Physiotherapy are subject to a rigorous peer review process, to ensure that the research published is valuable for its readership. Archives of Physiotherapy applies a double-blind review process and does not disclose the identity of its reviewers.

Lead times
Submission to final decision: 6-8 weeks
Acceptance to publication: 4 weeks

Publication fees
Articles submitted to Archives of Physiotherapy are published after having successfully completed the peer-review process. The publication of accepted open access articles is supported by the Italian Society of Physiotherapy, in line with their scope of fostering integrated information to support the improvement and recognition of professions strictly connected to physiotherapy.

There are no submission or publication fees, no fees for rejected articles and there is no additional cost for manuscript length, colour items or supplementary files, including videos.

Open access and copyright
All articles are published and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

Full author guidelines are available online at
www.archivesofphysiotherapy.com

EDITORIAL BOARD

Editor-in-Chief
Marco Barbero, PT, PhD
University of Applied Sciences and Arts of Southern Switzerland, Manno, Switzerland

Section Editors
Lara Allet, PT, PhD
University of Applied Sciences, School of Health, HES-SO Valais-Wallis, Sion, Switzerland

Davide Corbetta, PT
Vita-Salute San Raffaele University, Milan, Italy

Roma Forbes, PT, PhD
School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Australia

Andrew A. Guccione, PT, PhD, DPT
George Mason University, Fairfax Station, USA

Karen McCreesh, PT, PhD
University of Limerick, Limerick, Ireland

Anne W.S. Rutjes, PhD
UniCamillus, Rome, Italy

Stefan Schmid, PT, PhD
Bern University of Applied Sciences, Bern, Switzerland

Andrea Turolla, PT, PhD
Alma Mater Studiorum Università di Bologna, Bologna, Italy

Associate Editors
Ivan Bautmans, PT, PhD, *Brussels, Belgium*
Chad Cook, PT, PhD, *Durham, USA*
Patricia Almeida, *Cascais, Portugal*
Frida Eek, PT, PhD, *Lund, Sweden*
Roger Hilfiker, PT, PhD, *Lausanne, Switzerland*
Duncan Reid, PT, PhD, *Auckland, Australia*
Elisa Pelosin, PT, PhD, *Genoa, Italy*
Leonardo Pellicciari, PT, PhD, *Bologna, Italy*
Matteo Paci, PT, MSc, *Firenze, Italy*
Slavko Rogan, PT, PhD, *Bern, Switzerland*
Sandeep Subramanian, PT, MSc, PhD, *San Antonio, USA*
Stefano Vercelli, PT, MSc, PhD, *Manno, Switzerland*

Social Media Committee
Editor: Guillaume Christe, PT, MSc, *Lausanne, Switzerland*
Graphic designer: Laura Bassi, PT, *Ferrara, Italy*

ABOUTSCIENCE

Aboutscience Srl
Piazza Duca d'Aosta, 12 - 20124 Milano (Italy)

Disclaimer
The statements, opinions and data contained in this publication are solely those of the individual authors and contributors and do not reflect the opinion of the Editors or the Publisher. The Editors and the Publisher disclaim responsibility for any injury to persons or property resulting from any ideas or products referred to in the articles or advertisements. The use of registered names and trademarks in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Editorial and production enquiries
editorial@aboutscience.eu

Supplements, reprints and commercial enquiries
Lucia Steele - email: lucia.steele@aboutscience.eu

Publication data
eISSN: **2057-0082**
Continuous publication
Vol. 15 is published in 2025

ARCHIVES OF PHYSIOTHERAPY

Contents Volume 15 Number 1 January-December 2025

1 Coronavirus Disease 2019 (COVID-19) and pelvic floor signs and symptoms: a scoping review of the literature
Valentina Di Tommaso, Marta Rossi, Silvia Gianola, Greta Castellini, Silvia Bargeri, Giacomo Rossettini, Arianna Bortolami

9 Discriminative ability, responsiveness, and interpretability of smoothness index of gait in people with multiple sclerosis
Stefano Filippo Castiglia, Fulvio Dal Farra, Dante Trabassi, Andrea Turolla, Mariano Serrao, Ugo Nocentini, Paolo Brasiliano, Elena Bergamini, Marco Tramontano

19 Perceived bioethical issues in cancer rehabilitation: a qualitative study among Italian physiotherapists
Gianluca Bertoni, Simone Battista, Valentina Conti, Marco Testa, Sara Patuzzo Manzati

32 Comparisons of the effects of psychologically-informed and usual physiotherapy on pain sensitivity in chronic low back pain: an exploratory randomized controlled trial
Hugo Massé-Alarie, Amélie Desgagnés, Claudia Côté-Picard, Olivier Liberty, Pierre Langevin, Mathieu Piché, Yannick Tousignant-Laflamme

42 Scaling new heights: a pilot study of the impact of climbing on balance, agility, and dexterity in individuals with Parkinson's disease
Julie Ries, Molly Cupka, Andrew Guccione

49 The Use of Digital Technologies in Physiotherapy Higher Education: a Mixed-Methods Study
Francesca Trentini, Chiara Fante, Flavio Manganello, Marco Testa, Simone Battista

59 Mim-pong: a serious game for assessment and treatment of the lower limb in hemiparetic stroke patients
Fernando L. F. Eichinger, Fabrício Noveletto, Susana C. Domenech, Thierry Moulin, Yoshimasa Sagawa, Antonio Vinicius Soares

69 Entry-level physical therapist curricula in geriatric care: an Italian national survey study
Francesco Bucchi, Sara Dal Ben, Antonio Culcasi, Stefano Lecora, Roberto Tedeschi, Marco Da Roit

77 Effectiveness of physical therapy interventions for coccydynia: a systematic review with a narrative synthesis
Mohammad Sidiq, Hariharasudhan Ravichandran, Balamurugan Janakiraman, Aksh Chahal, Richa Hirendra Rai, Ahmed Hamoud Alotaibi, Ali Hamoud Alotaibi, Abdullah Saad Alotaibi, Aminu Alhassan Ibrahim, Eid Abed Alharbi, Faizan Zaffar Kashoo, Hemlata Vats

90 Intra- and inter-rater reliability of goniometric finger range of motion using a written protocol
Takuya Nakai, Satoru Amano, Chikako Murao, Haruki Taguchi, Kayoko Takahashi

97 Differences in cervical mobility and pain sensitivity between patients with rotator cuff-related shoulder pain and asymptomatic subjects: a cross-sectional study
Néstor Requejo-Salinas, Rubén Fernández Matías, Roy La Touche Arbizu, Javier Gordo Herrera, Pablo Conesa López, Enrique Lluch Girbés

110 Association of physiotherapy dose with motor recovery in early subacute phase after stroke: results from a multicenter Italian study
Silvia Salvalaggio, Nicola Valè, Thomas Bowman, Davide Cattaneo, Cristina Allera Longo, Serena Bocini, Viviana Bonci, Michele Gennuso, Francesca Marazzini, Francesco Giuseppe Materazzi, Elisa Pelosin, Martina Putzolu, Andrea Turolla, Susanna Mezzarobba, Fabiola Giovanna Mestanza Mattos, Simone Gambazza

118 Clinical relevance of the effects of robotic rehabilitation for upper limb recovery after stroke in randomized studies: a systematic review with meta-analysis
Sofia Verola, Alessandro Ugolini, Leonardo Pellicciari, Mauro Di Bari, Matteo Paci

AoP ARCHIVES OF PHYSIOTHERAPY

Contents

131 Application and adaptation of the Virtual Eggs Test for assessing hand dexterity in subjects with stroke: a pilot study
Michele Piazzini, Lucia Angelini, Tommaso Ciapetti, Stefano Doronzio, Marco Baccini, Robinson Guachi, Paola Mucci, Marco Controzzi, Francesca Cecchi

138 A new test protocol and device for measuring the cranio-cervical flexion test in participants with bruxism
Mareile Jung, Annika Zind, Markus Ernst

148 Validation of a Novel device for Assessing Neck Muscle Strength
Michail Arvanitidis, Hon Hin Ken Mak, Eduardo Martinez-Valdes, Marco Barbero, Deborah Falla

158 Head Pain Reproduction and Resolution Behavior in Response to Sustained Mobilization of the Upper Cervical Spine: A Case Series
Damien C Cummins, Lucy C Thomas, Peter Osmotherly

165 Evaluation and synthesis of physiotherapy protocols for Femoroacetabular Impingement Syndrome (FAIS): a scoping review
Céline Berguerand, Noemie Rossi, Rahel Caliesch

174 Motor and cognitive changes across 2 years: a longitudinal study in early Multiple Sclerosis
Elisa Gervasoni, Claudio Solaro, Denise Anastasi, Rachele Di Giovanni, Marco Rovaris, Giampaolo Brichetto, Paolo Confalonieri, Andrea Tacchino, Ilaria Carpinella, Davide Cattaneo

184 Primary care rehabilitation after Knee Replacement – a cross sectional study
Elin Östlind, Marcus Ljung, Caroline Ståhl, Anna Cronström, Thérèse Jönsson

192 Comments on: Impact of direct access on the quality of primary care musculoskeletal physiotherapy: a scoping review from a patient, provider, and societal perspective
Jean-François Kaux, Marc Schiltz, Patrick Linden, Geoffrey Brands, Bruno Schoonejans, Thierry Lejeune

194 Authors' reply to: "Comments on: Impact of direct access on the quality of primary care musculoskeletal physiotherapy: a scoping review from a patient, provider, and societal perspective"
Erik Cattrysse, Jona Van Den Broeck, Robin Petroons, Amber Teugels, Aldo Scafoglieri, Emiel van Trijffel

195 What do people search online about sciatica? What answers do they get from Dr. Google? A mixed analysis of Italian web-based data
Massimo Esposto, Matteo Cioeta, Giuseppe Giovannico, Valerio Barbari, Alessandra Arca

206 Biomechanical of bilateral heel rise, and its association with balance, functional mobility, and walking speed in older adults
Cristian Caparrós-Manosalva, Rodrigo Guzmán-Venegas, Francisca Gajardo-Garrido, Marion González-Fuentes, Víctor Pino-Domínguez, Jessica Espinoza-Araneda, Ivan Palomo, Nacim Molina

214 What are the evidence-based medical management approaches for the concussed youth athlete? A scoping review
Tom McKeever, Micheal Leavitt, Stephanie Valentin, David Hamilton

239 Biochemical and clinical effects of McKenzie therapy versus muscle endurance exercises in chronic low-back pain
Mistura Iyabo Olaoye, Raphael Okonji, Adekola Ademoyegun, Tadesse Gebrye, Gillian Yeowell, Francis Fatoye, Chidozie Mbada

AoP ARCHIVES OF PHYSIOTHERAPY

Contents

250 Metric properties of the Tonic Stretch Reflex Threshold (TSRT) as a measure of spasticity: a systematic review with meta-analysis
Diego Longo, Marco Baccini, Giulio Cherubini, Aurora Caporaso, Mario De Marco, Angela Politi, Guido Santini, Francesca Cecchi, Maria Angela Bagni

265 Usability of Myosuit exosuit and effects of device-mediated rehabilitation on chronic stroke survivors: a non-randomized pilot study
Stefano Doronzio, Silvia Campagnini, Andrea Mannini, Chiara Basla, Tommaso Ciapetti, Michele Piazzini, Angela Politi, Chiara Pedrini, Erika Guolo, Serena Malloggi, Teresa Barretta, Robert Riener, Francesca Cecchi

277 Development and feasibility testing of a video-based exercise program for improving upper limb motor function among stroke survivors in a low-income setting
Hafsat Maina Ali, Surajo Kamilu Sulaiman, Jibrin Sammani Usman, Muhammad Aliyu Abba, Aminu Alhassan Ibrahim, Umar Muhammad Bello, Dauda Salihu, Auwal Abdullahi

286 Physical therapists' perspectives and practices on weight management for chronic pain patients with obesity: a cross-sectional survey
Melanie Liechti, Massimo Menegon, Alexander Philipp Schurz, Matteo Vanroose, Ron Clijsen, Anneleen Malfliet, David Faeh, Heiner Baur, Jan Taeymans, Nathanael Lutz

297 Is resistance training with external loads superior to unloaded exercise in the management of chronic low back pain? A systematic review and meta-analysis
Marco Ranzani, Andrea Pozzi, Daniele Fornasari, Diego Ristori, Marco Testa

314 Postoperative education for patients after hip/knee arthroplasty: a systematic review with meta-analysis of randomized controlled trials
Leonardo Piano, Chiara Brossa, Leandro Francesco Gioia, Adele Occhionero, Andrea Dell'Isola, Marco Trucco

326 Erratum in: Evaluation and synthesis of physiotherapy protocols for Femoroacetabular Impingement Syndrome (FAIS): a scoping review
Céline Berguerand, Rahel Caliesch, Noemie Rossi

327 Mastering the use of hand-held dynamometry in clinical practice
Claire Minshull

Arch Physioter 2025; 15: 1-8
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3188
REVIEW

Coronavirus Disease 2019 (COVID-19) and pelvic floor signs and symptoms: a scoping review of the literature

Valentina Di Tommaso[1], Marta Rossi[2], Silvia Gianola[3], Greta Castellini[3], Silvia Bargeri[3], Giacomo Rossettini[4], Arianna Bortolami[5]

[1]Studio di fisioterapia PelviCare, Teramo - Italy
[2]Ospedale Papa Giovanni XXIII, Bergamo - Italy
[3]Unit of Clinical Epidemiology, IRCCS Istituto Ortopedico Galeazzi, Milan - Italy
[4]Department of Physiotherapy, Faculty of Sport Sciences, Universidad Europea de Madrid, Madrid - Spain
[5]Pelvic Floor Studio, Padua - Italy

ABSTRACT

Introduction: The Coronavirus Disease 2019 (COVID-19) pandemic presents a substantial global health challenge. While the disease is known to impact multiple systems, leading to long-term consequences that require monitoring and rehabilitation, its effects on the pelvic floor remain unclear.
This study aims to explore COVID-19-related signs and symptoms affecting pelvic floor functions through a scoping review.
Methods: We conducted a scoping review following the Arksey and O'Malley framework. A systematic search was performed in PubMed, CINAHL, and Embase databases up to March 19, 2024, to identify studies examining pelvic floor-related signs and symptoms in COVID-19 patients. Two independent reviewers extracted the study and participant characteristics, areas involved (e.g., anorectal, sexual, urogenital), and signs and symptoms using an ad-hoc data extraction form. Signs and symptoms were classified as direct (e.g., directly impacting the pelvic floor) or indirect (e.g., indirectly affecting the pelvic floor with potential long-term consequences).
Results: We included 104 studies, primarily a systematic review (N = 40; 38.46%) and focused on adult populations (N = 80; 76.92%), investigating 140 signs and symptoms from various regions worldwide. Most (N = 124; 88.57%) were indirect signs and symptoms, while the minority were direct (N = 16; 11.43%). The most prevalent indirect symptom was diarrhea (n = 81; 70,43%) in the anorectal bowel area (n = 115). The most prevalent direct were lower urinary tract symptoms (LUTS) (n = 16; 84.21%) in the urogenital area (n = 19).
Conclusion: This study highlights the significant prevalence of both direct and indirect pelvic floor symptoms in COVID-19 patients. Clinicians should be aware of the potential link between COVID-19 and pelvic floor dysfunction.

Keywords: Anorectal bowel symptoms, COVID-19, Lower urinary tract symptoms, Pelvic floor, Sexual dysfunction

What is already known about this topic:

- *The COVID-19 pandemic has resulted in death or temporary, permanent, or progressively worsening consequences for many people. Although the disease is known to affect many systems, its effects on the pelvic floor remain unclear.*

What does the study add:

- *This review maps the literature on the prevalence of COVID-19 symptoms affecting pelvic floor functions. Clinicians should be aware of the potential link between COVID-19 and pelvic floor dysfunction.*

Received: July 1, 2024
Accepted: December 20, 2024
Published online: January 20, 2025

This article includes supplementary material

Corresponding author:
Arianna Bortolami
email: arianna.bortolami@gmail.com

Introduction

The Coronavirus Disease 2019 (COVID-19) pandemic has determined, in addition to the death of many people, temporary, permanent, or progressively worsening consequences (1,2). An abundance of literature exists on respiratory system affections after COVID-19, but more information on other physiologic complications has yet to be published so far. It has been reported that the COVID-19 virus uses the

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

angiotensin-converting enzyme 2 (ACE2) as a host cell receptor, impacting not only the respiratory system but also the digestive system and the bladder (3). The respiratory symptoms of COVID-19, such as coughing and shortness of breath, cause an increase in intra-abdominal pressure that might develop pelvic floor under activity, contributing to the onset or worsening of urinary or fecal incontinence and/or pelvic organ prolapse. Furthermore, surviving people after a long-term hospitalization due to COVID-19 with a post-intensive care syndrome might have consequences on sexual, bowel, and bladder functioning (4). Thus, there could be signs and symptoms directly or indirectly related to the pelvic floor: the muscular and connective part that coincides with the genito-urinary-anal area, such as urinary and fecal incontinence and urgency, difficulty emptying of bladder and rectum, pelvic organ prolapse, sexual dysfunction and chronic pelvic pain (5,6).

The pathophysiological hypothesis for DIRECT symptoms correlates to a condition of systemic inflammation characterized by a cytokine cascade that amplifies the flogosis of pelvic structures (7). This inflammatory process can lead to specific symptoms of pelvic floor dysfunction, such as lower urinary tract symptoms (LUTS). These symptoms may manifest as difficulties in urination, urgency, and pain, highlighting the significant impact that the infection can have on urogenital health (8,9). Additionally, patients who have contracted COVID-19 may suffer from musculoskeletal or neuropathic pain or have had their chronic pain symptoms exacerbated by the virus. This can also include the pelvis, pelvic floor, and pudendal nerve (10).

For INDIRECT symptoms, the pathophysiology is related to: (a) cough and dyspnea of inducing incontinence, symptoms of the bladder and rectal filling phase (9), alterations of pelvic statics (12); (b) sexual dysfunctions (13); and (c) alterations of intestinal functionality, such as diarrhea that can generate anal incontinence and fecal urgency, and constipation with involvement of the pelvic floor. Additionally, the presence of inflammation and consequent pelvic pain, along with neuropathy of the pudendal nerve, contribute to these indirect symptoms.

However, health professionals often underestimate these signs and symptoms. Nevertheless, pelvic floor dysfunctions also have a significant role in worsening the quality of life of people, leading to sensations such as embarrassment, anxiety, depression, loneliness, alteration of the body pattern, lack of freedom due to the frequent need for the toilet, unsatisfactory sexuality, impossibility of procreation, impossibility of sitting position and fear of movement as in the case of pelvic pain chronic (14). Women and men affected by pelvic floor dysfunctions could represent an economic burden for the health system, both from direct financial costs derived from healthcare assistance, complications, psychological assistance, and the use of hygiene devices (e.g., pads), as well as indirect costs, derived from absenteeism and lost productivity (15,16,17).

During the COVID-19 pandemic, hospitals and clinics were particularly affected by a massive reduction of outpatient investigations, which reduced the access to care for patients suffering from pelvic floor dysfunctions (18). Therefore, signs and symptoms might not have been adequately reported due to the organizational limitations posed by the pandemic. Thus, given the insufficient literature on this topic, our scoping review aimed to fill the gap, mapping what was published on the topic and exploring possible direct and indirect signs correlated to the pelvic floor to investigate with research.

This information could benefit health professionals, particularly physiotherapists, specializing in the pelvic floor, encouraging them to investigate, during anamnesis and physical examination, the potential consequence of COVID-19 on pelvic floor functions in terms of direct and indirect symptoms in the genito-urinary-area. We hypothesize that this evidence could inform physiotherapists' clinical reasoning and decision-making process, guiding the adoption of appropriate treatments (e.g., manual therapy, therapeutic exercise, pelvic floor devices, and behavioral education) to properly address these issues.

Objectives

This scoping review aimed to investigate the prevalence of signs and symptoms of COVID-19 impacting directly or indirectly on pelvic floor function, to identify a population at risk of developing or worsening pelvic floor-related symptoms, and to inform clinicians of reasoning and decision-making.

Methods

Study design

The scoping review was conducted following the Joanna Bridge Institute methodology for scoping review (19) and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA ScR) Checklist for reporting (20).

The protocol information relating to this review is stored in the OSF repository at the following link: Online.

Defining the research question

This scoping review aimed to identify and describe the extent of the scientific evidence available that investigates whether the signs and symptoms of COVID-19 may have directly or indirectly led to the onset or worsening of pelvic floor symptoms related to different areas: urological, ano-rectal-bowel, urogynecological, and sexual.

Specific objectives were:

- to identify the extent of research available in relation to study characteristics;
- to map and summarize findings related to signs and symptoms of pelvic floor in COVID-19 patients;
- to discover and describe any gaps in the literature.

Eligibility criteria

We used the Population, Concept, Context (PCC) framework according to Joanna Bridge Institute methodology for scoping reviews to identify useful studies for our research (19). Specifically, we included studies published in any language focusing on:

(i) Population: pediatric, adolescent, and adult subjects (including elderly) or mixed population (e.g., presence of adults and children) regardless of sex. Pediatric subjects were included because it's important to identify as early as possible during life symptoms that may involve the intimate area of the person, thus also determining significant dysfunction in adulthood (21,22). Elderly subjects were involved in this study to analyze how their quality of life may have been changed in relation to the onset of the pelvic floor consequences caused by COVID-19 (23).
(ii) Concept: studies that describe signs and symptoms of the pelvic floor in the following categories: anorectal bowel (e.g., constipation), urogynecological (e.g., urinary incontinence, pelvic organ prolapse), sexual diseases (e.g., dyspareunia). In COVID-19 (including long-COVID);
(iii) Context: Any clinical context (patients who complain of urological, ano-rectal-bowel, gynecological, sexual, symptoms pelvic floor related) and/or organizational context during COVID-19;
(iv) Type of Evidence sources: systematic review, observational or experimental study published in English excluding other publication types such as letter, editorial, news, comment, case series, case report, and narrative review. Studies focusing on non-human populations and studies that referred to conditions related to the vaccine or with the title referring to the use of drugs or vaccine conditions were excluded. We have also decided to exclude all studies in which the manifestation of symptoms was linked to the contextual situation of the isolation induced by the pandemic as, for example, studies related to sexual activity limited by the forty: we have also decided to exclude all studies in which the manifestation of symptoms was linked to the contextual situation of the isolation induced by the pandemia.

Information sources

We used keywords and terminology referring to the international standards of the Scientific Societies and their reference Guidelines in this topic (6,24,25,26). We searched PubMed, Cumulated Index to Nursing and Allied Health Literature (CINAHL), and Embase by Mesh terms and free text words. Additional articles were handpicked from reference lists of included articles. We launched the search strategy on November 16, 2022. The full search strategy is reported in the **Supplementary File**.

Screening and study selection

After detecting the duplicates, two authors (VDT, MR) independently screened all publications by titles and abstract. A third author (AB) was consulted in case of disagreement. We used Rayyan software (27) to manage the study selection phase.

Charting the data

Two authors (VDT, MR) developed a data collection form to extract the characteristics of the included studies by piloting a small sample of ten records. The final data collection form was shared among review authors, reaching a consensus to collect and chart data. Finally, the two authors (VDT, MR) extracted data from the title and abstract of eligible studies and discussed any conflict.

We extracted the following data: name of the first author, year of publication, journal of publication, study design, aim (e.g., prevalence, intervention, prognosis), general characteristics of the population, including signs and symptoms directly or indirectly related to the pelvic-floor impacting on anorectal bowel, urogenital and sexual area. We defined direct signs or symptoms (as new onset or worsened) when they ranged from bladder or rectal emptying or filling symptoms, pelvic static alterations, chronic pelvic pain, and sexual dysfunction from pelvic floor dysfunction, as classified by International Scientific Societies of pelvic floor dysfunctions during COVID-19 (3,4). Indirect signs or symptoms are those symptoms developed during COVID-19 which can have consequences on the pelvic floor if continued over time, such as a persistent condition of bowel movement alteration, pelvic organ inflammation, pain and musculoskeletal dysfunction, sexual dysfunction, anxiety, depression, cough, dyspnea, neuropathies, and fatigue.

Summarizing and reporting results

The results were summarized in relation to the research questions with descriptive reporting. In addition, to identify the extent of research available in relation to study characteristics, we tabulated the main findings, whereas to map and summarize findings related to signs and symptoms of the pelvic floor in COVID-19 patients, we reported findings using figures. The symptomatology was categorized according to the anatomical area involved: urogenital area, sexual area, and anorectal bowel area.

Results

Selection of sources of evidence

After removing duplicates, we identified and screened 5323 references. A total of 577 references did not meet the inclusion criteria and have been excluded. Finally, 104 articles were included (Figure 1; references in Supplementary File).

Characteristics of sources of evidence

Among the 104 articles included, the main represented study designs were systematic reviews (N = 40; 38.46%) and retrospective studies (N = 39; 37.50%), followed by cross-sectional studies (N = 14; 13.46%) and prospective (N = 11; 10.58%) The majority (N = 80; 76.92%) were focused on the adult population, 13 (12.50%) on children and 11 on both (10.58%) (Table 1). Most studies had corresponding authors from the American region (N = 33; 31.73%), 29 from Northwestern Asia (27.88%), 20 from Europe (19.23%), and 22 from Southeast Asia (21.15%) (Figure 2). The graph was created with Excel to obtain data from the studies and create a table with the numbers of publications by countries of origin of the corresponding authors.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

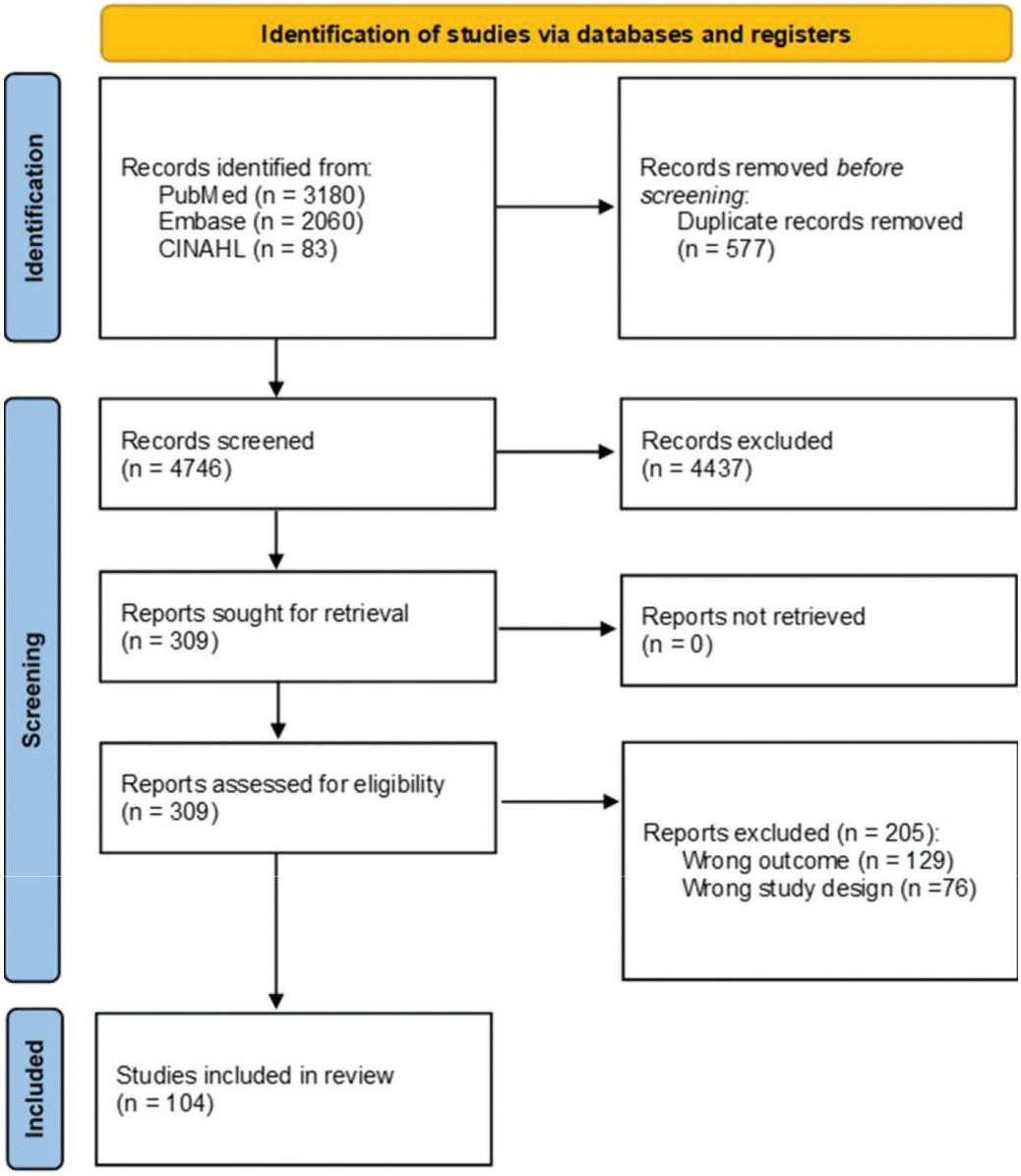


FIGURE 1 - Flow chart of study selection.
Identification of studies via databases and registers. After removing duplicates, 5323 references were identified and screened. A total of 3324 references did not meet the inclusion criteria and have been excluded. Finally, 104 articles were included.

TABLE 1 - Results of characteristics of studies

Characteristics	N (%)
Population	
Adults	76 (73.08)
*Adults women**	1 (0.96)
*Adults men**	3 (2.88)
Children	13 (12.50)
*Mixed^*	11 (10.58)
Study design	
Systematic review	40 (38.46)
Randomized controlled trial	0 (0.00)
Prospective	11 (10.58)
Retrospective	39 (37.50)
Cross-sectional	14 (13.46)

Characteristics	N (%)
Focus of study	
Prevalence	57 (54.81)
Prognosis	22 (21.15)
Intervention	0 (0.00)
Prevalence and Prognosis	25 (24.04)
Type of Covid-19	
Acute	85 (81.73)
Long covid	11 (10.57)
Post covid	4 (3.85)
Mixed$	4 (3.85)

Legend: *studies specifically focusing on women or men; ^both adults and children; $acute, long or post-COVID-19; N, number; % percentage. Percentages in brackets are calculated out of the 104 overall included studies.

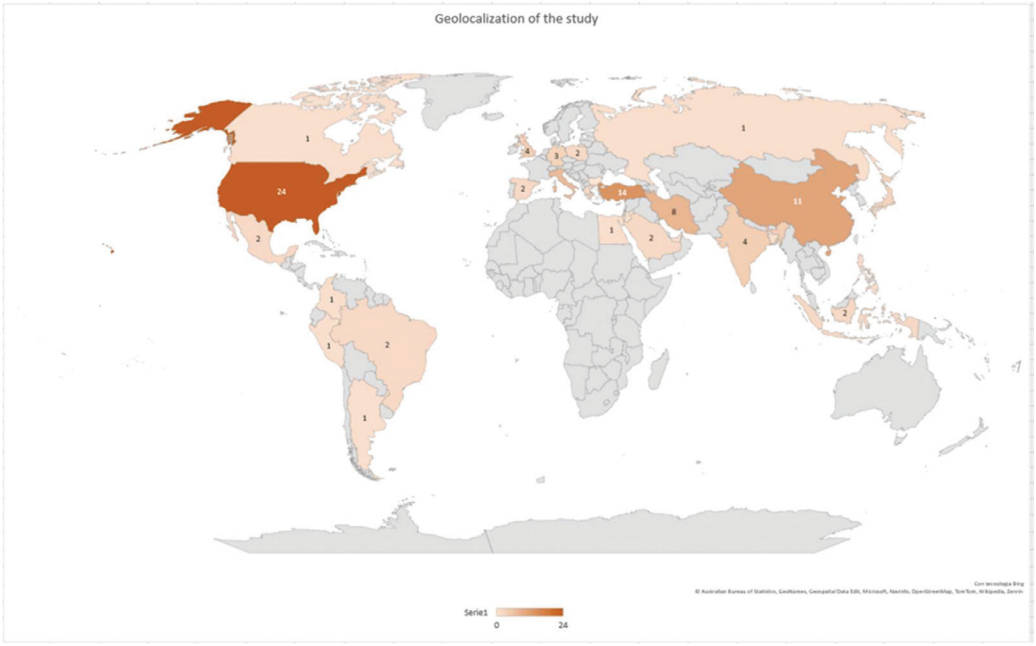


FIGURE 2 - Global prevalence of the studies analyzed.
The map represents the distribution of articles published in various countries. The intensity of the color progressively increases with the number of publications.

Symptoms classification

Overall, in 104 studies investigated, we analyzed 140 signs and symptoms. In particular, we found 16 (11.43%) signs and symptoms in 16 studies classified as direct and 124 (88.57%) signs and symptoms in 88 studies classified as indirect.

As reported in Table 2, the prevalent area was anorectal bowel (N = 115 signs and symptoms out of 140; 82,14%), followed by urogenital (N = 19 out of 140; 13,57%) and sexual area (N = 6 out of 140; 4,29%).

The prevalent symptoms in the urogenital area were LUTS (n = 16; 84.21%), while in the ano-rectal-bowel area was diarrhea (N = 81; 70.43%), and in the sexual area was erectile dysfunction (N = 4; 66.67%).

Figure 3 shows the signs and symptoms categorized based on whether their influence was direct or indirect on the symptomatic areas considered.

Among direct symptoms, the prevalent area was urogenital with LUTS, with the symptoms always presented (n = 16 out of 16; 100%). No studies reported direct symptoms for ano-rectal-bowel and sexual areas.

On the contrary, among the indirect symptoms, the prevalent reported signs and symptoms were anorectal bowel ones (n = 115; 92.74%), with diarrhea the most reported symptom (n = 81 out of 115; 71.7%)

Discussion

This scoping review, including 104 articles, covers anorectal, sexual, and urogenital signs and symptoms linked to the pelvic floor in all age populations resulting from COVID-19 disease. As a main finding, we found a high prevalence of both direct and indirect symptoms affecting pelvic floor function in COVID-19 patients. Indirect symptoms were more

TABLE 2 - Results of characteristic of symptoms

Signs and symptoms *area* *(n = 140)*	**N**	***Area***	***Symptoms***	**N (%)**
	115	***Ano-rectal-bowel***	*Diarrhea*	81 (70.43)
			Abdominal pain	25 (21.74)
			Constipation	9 (7.83)
	19	***Uro-genital***	*LUTS*	16 (84.21)
			Symptoms related to endometriosis	1 (5.26)
			Uro-genital inflammation and injections	2 (10.53)
	6	***Sexual***	*Erectile dysfunction*	4 (66.67)
			Sexual desire	2 (33.33)

Legend: LUTS, Lower Urinary Tract Symptoms; n, numbers; %, percentage. In brackets, we reported percentages of the overall number of symptoms in each category.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

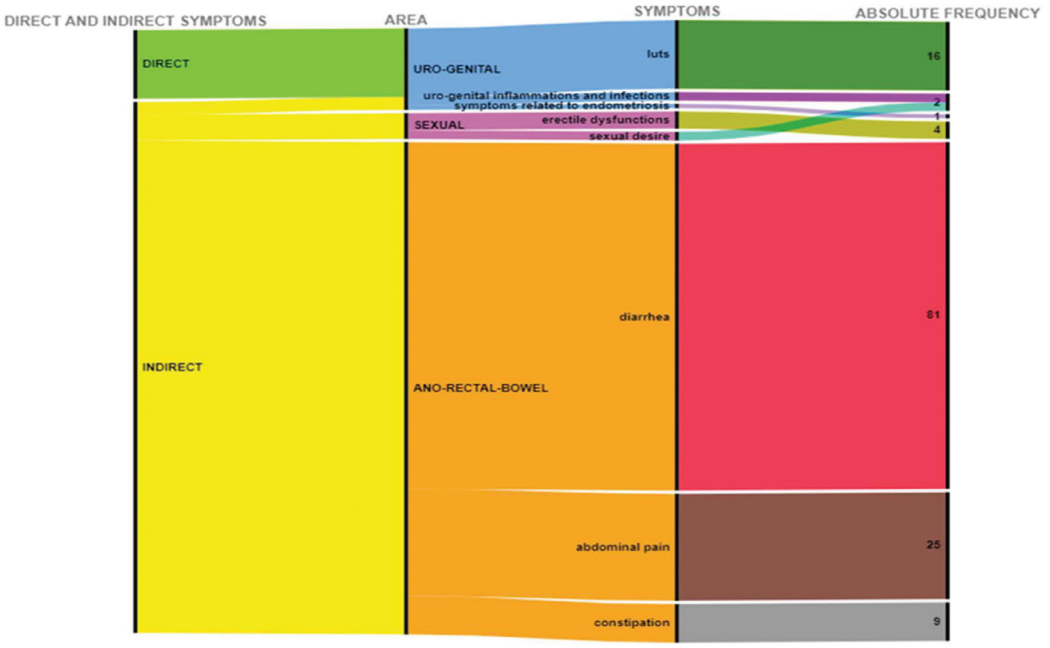


FIGURE 3 - Direct and indirect symptoms according to clinical area.
Direct and indirect symptoms-related to the pelvic floor caused or worsened by COVID-19 displayed according to the clinical area: the prevalent area among direct symptoms was urogenital with LUTS (upper part of the light blue color), and among indirect symptoms was anorectal bowel with diarrhea (upper part of orange color).

common, particularly in the anorectal and bowel area, with diarrhea being the main symptom. In fact, diarrhea could be a significant risk factor for fecal incontinence, suggesting that its presence, combined with SARS-CoV-2 infection, might have severe consequences for the pelvic floor (28,29).

Direct symptoms were primarily reported in the urogenital area, with LUTS being the most frequent issue. The severity of COVID-19 was found to correlate with an increase in LUTS. Sexual dysfunctions, though less common, were also reported, including erectile dysfunction and altered sexual desire, which can affect pelvic floor muscle function. In fact, male and female sexual dysfunction can cause and support symptoms involved in disability of pelvic floor muscles in their function (30,31).

Comparison with literature

To the best of our knowledge, this is the first scoping review investigating the reporting of direct and indirect signs and symptoms affecting urogenital, sexual, and anorectal bowel areas in COVID-19 patients, thereby limiting the comparison with existing evidence.

Another scoping review reported a relationship between gastrointestinal symptoms and COVID-19 infection in the pediatric population, with diarrhea being the most prevalent symptom (32). In fact, ACE1 and ACE2 are more expressed at the intestinal level in children than in adults, which explains why the pediatric population presents more evident gastrointestinal symptoms compared to adults (33).

Other scoping reviews focused on the effects of lockdowns, reporting a broad impact on sexual health (34) and, considering different biomedical fields, an increased risk of developing musculoskeletal (MSK) disorders due to working from and at non-ergonomic home environments, sedentariness, and the limited access to a healthcare facility (35).

Clinical Implications

The findings of this scoping review constitute a valuable resource to guide the clinical reasoning and decision-making of clinicians working with pelvic floor dysfunctions, emphasizing the need for comprehensive evaluation and management strategies in post-COVID care to enhance patient quality of life. Given the high prevalence of direct and indirect pelvic floor symptoms in COVID-19 patients, physiotherapists should also consider inquiring about COVID-19 history during anamnesis, and when detected by clinical examination (36), they should implement multimodal and multiprofessional treatment. Physiotherapists should educate the patient and, through the application of hands-on (e.g., manual therapy (37,38) and hands-off techniques (e.g., therapeutic exercises (39), address pelvic floor-related symptoms to improve the function of the pelvic-floor muscles. Additionally, physiotherapists should collaborate synergistically with other healthcare professionals, such as urologists, gastroenterologists, and internist physicians, to manage the systemic symptoms of COVID-19 with targeted treatments, implement preventive measures to mitigate virus-related damage and identify high-risk patients for chronification.

Research Implications

Our scoping review highlights the need for further research worldwide on the direct and indirect impacts of COVID-19 on the pelvic floor. Most studies have been conducted in the Americas and Asia. Adult populations (>18 years of age) have been predominantly studied. Retrospective studies were the most common, followed closely by systematic reviews, with fewer observational, prospective, and cross-sectional studies. Considering the limitations of the existing literature, future studies should explore under-researched areas such

as pelvic floor symptoms in pediatric populations and mixed-age groups. Moreover, research should involve patients from other regions (e.g., Africa, Europe, and Oceania). Qualitative studies (e.g., focus groups and interviews) could also inform clinicians about the lived experiences of patients with pelvic floor impairments related to COVID-19. Finally, longitudinal studies are required to understand the long-term impact of COVID-19 on pelvic health.

Strengths and Limitations

This review is the first to map the literature on the prevalence of COVID-19 symptoms affecting pelvic floor functions. It establishes a basis for hypothesizing the virus's impact beyond the respiratory system. However, some limitations must be acknowledged. Firstly, the relationship between exposure and disease development could not be established due to limited primary research. Secondly, confounding factors may have influenced symptom presentation. Thirdly, methodological quality and critical appraisal of included studies were not assessed, as the focus was on the breadth of evidence. Fourthly, symptom prevalence percentages were not extracted, aligning with the scoping review's purpose of mapping existing literature rather than detailed inspection (12). Lastly, only English-language articles were included. Additionally, the reduction in outpatient services during the pandemic may have led to underreporting and undertreatment of pelvic floor dysfunction, suggesting that current prevalence estimates might be underestimated.

Conclusions

COVID-19 is known to have multiple systemic implications, and our study demonstrates a significant prevalence of direct and indirect symptoms affecting pelvic floor functions, which can compromise patients' quality of life. Clinicians should be vigilant about the potential link between COVID-19 and pelvic floor symptoms to identy high-risk patients and develop effective treatment programs, enhancing clinical practice from initial assessment to comprehensive treatment.

Disclosures

Conflicts of Interest: The authors declare no conflicts of interest.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Author contributions: Conception and design: AB, SG, GR, SB, GC; Acquisition of data: VDT, MR; Analysis and interpretation of data: VDT, MR; Drafting the article: AB, SG, SB, GC, VDT, MR, GR; Review and editing the article: VDT, MR, SB; Final approval for the completed article: AB, SG, SB, GC, VDT, MR, GR; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data sharing is not applicable. No new data were created or analyzed in this study, or the data presented in this study are available as supplementary material to this article.

References

1. Khan M, Adil SF, Alkhathlan HZ, et al. COVID-19: a global challenge with old history, epidemiology and progress so far. Molecules. 2020;26(1):39. CrossRef PubMed
2. Leung TYM, Chan AYL, Chan EW, et al. Short- and potential long-term adverse health outcomes of COVID-19: a rapid review. Emerg Microbes Infect. 2020;9(1):2190-2199. CrossRef PubMed
3. Zheng L, Zhang L, Zheng Y, et al. Digestive system infection by SARS-CoV-2: entry mechanism, clinical symptoms and expression of major receptors (Review). Int J Mol Med. 2023;51(3):19. CrossRef PubMed
4. Siracusa C, Gray A. Pelvic floor considerations in COVID-19. J Womens Health Phys Therap. 2020;44(4):144-151. CrossRef PubMed
5. D'Ancona C, Haylen B, Oelke M, et al. Standardisation Steering Committee ICS and the ICS Working Group on terminology for male lower urinary tract & pelvic floor symptoms and dysfunction. The International Continence Society (ICS) report on the terminology for adult male lower urinary tract and pelvic floor symptoms and dysfunction. Neurourol Urodyn. 2019; 38(2):433-477. CrossRef PubMed
6. Haylen BT, de Ridder D, Freeman RM, et al. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for female pelvic floor dysfunction. Int Urogynecol J. 2010;21(1):5-26. CrossRef PubMed
7. Silva Andrade B, Siqueira S, de Assis Soares WR, et al. Long-COVID and post-COVID health complications: an up-to-date review on clinical conditions and their possible molecular mechanisms. Viruses. 2021;13(4):700. CrossRef PubMed
8. Zachariou A, Sapouna V, Kaltsas A, et al. Evaluation of overactive bladder symptoms in patients recovering from post-acute COVID-19 syndrome. J Multidiscip Healthc. 2022;15:2447-2452. CrossRef PubMed
9. Dhawan A, Ganduboina R, Dutta P, et al. COVID-associated cystitis: the culprit behind the bladder woes post-COVID infection? A review. Int Urol Nephrol. 2023;55(10):2367-2372. CrossRef PubMed
10. Lam CM, Sanderson M, Vu DT, et al. Musculoskeletal and neuropathic pain in COVID-19. Diagnostics (Basel). 2024;14(3):332. CrossRef PubMed
11. Thomaz RP, Colla C, Darski C, et al. Influence of pelvic floor muscle fatigue on stress urinary incontinence: a systematic review. Int Urogynecol J. 2018;29(2):197-204. CrossRef PubMed
12. Fitz FF, Bortolini MAT, Pereira GMV, et al. PEOPLE: Lifestyle and comorbidities as risk factors for pelvic organ prolapse-a systematic review and meta-analysis PEOPLE: Pelvic Organ Prolapse Lifestyle comorbidities. Int Urogynecol J. 2023;34(9):2007-2032. CrossRef PubMed
13. Yelin D, Daitch V, Kalfon T, et al. Long COVID sexual dysfunction among both genders: evaluation of a cohort of COVID-19 recoverees. Infect Dis Now. 2023;53(7):104750. CrossRef PubMed
14. Kalata U, Jarkiewicz MM, Barcz EM. Depression and anxiety in patients with pelvic floor disorders. Ginekol Pol. 2023; 94(9):748-751. CrossRef PubMed
15. Coyne KS, Wein AJ, Tubaro A, et al. The burden of lower urinary tract symptoms: evaluating the effect of LUTS on health-related quality of life, anxiety and depression: EpiLUTS. BJU Int. 2009;103(s3)(suppl 3):4-11. CrossRef PubMed
16. Felde G, Ebbesen MH, Hunskaar S. Anxiety and depression associated with urinary incontinence. A 10-year follow-up study from the Norwegian HUNT study (EPINCONT). Neurourol Urodyn. 2017;36(2):322-328. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

17. Irwin DE, Mungapen L, Milsom I, et al. The economic impact of overactive bladder syndrome in six Western countries. BJU Int. 2009;103(2):202-209. CrossRef PubMed
18. Sacco E, Gandi C, Li Marzi V, et al. Italian Society of Urodynamics (SIUD). Extensive impact of COVID-19 pandemic on pelvic floor dysfunctions care: a nationwide interdisciplinary survey. Neurourol Urodyn. 2021;40(2):695-704. CrossRef PubMed
19. Peters MDJGC, McInerney P, Munn Z, et al. Chapter 11: Scoping Reviews. Aromataris E, Munn Z (Editors). JBI Man. Evid. Synth. JBI, 2020. CrossRef.
20. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and Explanation. Ann Intern Med. 2018;169(7):467-473. CrossRef PubMed
21. Bolia R, Goel AD, Badkur M, et al. Gastrointestinal Manifestations of pediatric coronavirus disease and their relationship with a severe clinical course: a systematic review and meta-analysis. J Trop Pediatr. 2021;67(2):fmab051. CrossRef PubMed
22. Tiryaki S, Egil O, Birbilen AZ, et al. COVID-19 associated lower urinary tract symptoms in children. J Pediatr Urol. 2022;18(5): 680.e1-680.e7. CrossRef PubMed
23. Can O, Erkoç M, Ozer M, et al. The effect of COVID-19 on lower urinary tract symptoms in elderly men. Int J Clin Pract. 2021;75(6):e14110. CrossRef PubMed
24. Abrams P, Cardozo L, Fall M, et al. Standardisation Sub-Committee of the International Continence Society. The standardization of terminology in lower urinary tract function: report from the standardization sub-committee of the International Continence Society. Urology. 2003;61(1):37-49. CrossRef PubMed
25. Sultan AH, Monga A, Lee J, et al. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for female anorectal dysfunction. Int Urogynecol J. 2017;28(1):5-31. CrossRef PubMed
26. Basson R, Leiblum S, Brotto L, et al. Revised definitions of women's sexual dysfunction. J Sex Med. 2004;1(1):40-48. CrossRef PubMed
27. Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan-a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210. CrossRef PubMed
28. Whitehead WE, Borrud L, Goode PS, et al. Fecal incontinence in US adults: epidemiology and risk factors. Gastroenterology. 2009 Aug;137(2):512-517.e1-2. CrossRef
29. Bharucha AE, Knowles CH, Mack I, et al. Faecal incontinence in adults. Nat Rev Dis Primers. 2022;8(1):53. CrossRef PubMed
30. Seehuus M, Fertig M, Handy AB, et al. The impact of COVID-19 and long COVID on sexual function in cisgender women. J Sex Med. 2024;21(2):129-144. CrossRef PubMed
31. Salar R, Erbay G, Şanlı A, et al. Effect of COVID-19 infection on the erectile function. Andrology. 2023;11(1):10-16. CrossRef PubMed
32. Desconsi D, Araujo JP, Furtado MD, et al. Relationship between gastrointestinal symptoms and COVID-19 infection in the pediatric population: a scoping review. Rev Esc Enferm USP. 2024;58:e20230365. CrossRef PubMed
33. Canani RB, Comegna M, Paparo L, et al. Age-related differences in the expression of most relevant mediators of SARS-CoV-2 infection in human respiratory and gastrointestinal tract. Front Pediatr. 2021;9:697390. CrossRef PubMed
34. Toldam NE, Graugaard C, Meyer R, et al. Sexual health during COVID-19: a scoping review. Sex Med Rev. 2022;10(4):714-753. CrossRef
35. Gebrye T, Niyi-Odumosu F, Lawoe J, et al. The impact of COVID-19 related lockdown restrictions on musculoskeletal health: a systematic review. Rheumatol Int. 2023;43(11):2011-2019. CrossRef PubMed
36. Frawley H, Shelly B, Morin M, et al. An International Continence Society (ICS) report on the terminology for pelvic floor muscle assessment. Neurourol Urodyn. 2021;40(5):1217-1260. CrossRef PubMed
37. Farra FD, Aquino A, Tarantino AG, et al. Effectiveness of myofascial manual therapies in chronic pelvic pain syndrome: a systematic review and meta-analysis. Int Urogynecol J. 2022;33(11):2963-2976. CrossRef PubMed
38. Abaraogu UO, Igwe SE, Tabansi-Ochiogu CS, et al. A systematic review and meta-analysis of the efficacy of manipulative therapy in women with primary dysmenorrhea. Explore (NY). 2017;13(6):386-392. CrossRef PubMed
39. Alouini S, Memic S, Couillandre A. Pelvic floor muscle training for urinary incontinence with or without biofeedback or electrostimulation in women: a systematic review. Int J Environ Res Public Health. 2022;19(5):2789. CrossRef PubMed

Arch Physioter 2025; 15: 9-18
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3289
ORIGINAL RESEARCH ARTICLE

Discriminative ability, responsiveness, and interpretability of smoothness index of gait in people with multiple sclerosis

Stefano Filippo Castiglia[1,2], Fulvio Dal Farra[3], Dante Trabassi[1], Andrea Turolla[4,5], Mariano Serrao[1,6], Ugo Nocentini[7], Paolo Brasiliano[8], Elena Bergamini[9], Marco Tramontano[4,5]

[1]Department of Medico – Surgical Sciences and Biotechnologies, "Sapienza" University of Rome, Latina - Italy
[2]Department of Brain and Behavioral Sciences, University of Pavia, Pavia - Italy
[3]Department of Information Engineering, University of Brescia, Brescia - Italy
[4]Department of Biomedical and Neuromotor Sciences (DIBINEM), Alma Mater University of Bologna, Bologna - Italy
[5]Unit of Occupational Medicine, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna - Italy
[6]Gait Analysis LAB Policlinico Italia, Roma - Italy
[7]Department of Clinical Sciences and Translational Medicine, University of Rome Tor Vergata, Rome - Italy
[8]Department of Movement, Human and Health Sciences, University of Rome "Foro Italico", Rome - Italy
[9]Department of Management, Information and Production Engineering, University of Bergamo, Dalmine - Italy

ABSTRACT

Introduction: Gait impairments are common in People with Multiple Sclerosis (PwMS). Several studies have examined the clinometric properties of Inertial Measurement Units (IMUs), with LDLJa identified as a robust metric for gait smoothness. However, its responsiveness and interpretability have not been explored.
Methods: This cross-sectional study at IRCCS Santa Lucia Hospital enrolled 44 PwMS (age: 28-71; EDSS: 0-6) and 43 age- and gait-speed-matched healthy participants (HP). Two physiotherapists conducted assessments with five synchronized IMUs during a 10-meter walk at participants' preferred speed. Data were collected at baseline (T0) and after 4 weeks of training (T1).
Results: Significant differences in log dimensionless jerk (LDLJa) were found between PwMS and HP in the AP ($p < 0.001$, $d = 0.63$), ML ($p < 0.001$, $d = 1.08$), and CC ($p = 0.03$, $d = 0.68$) directions. PwMS had lower LDLJaAP values (< -4.88) and LDLJaML values (< -5.40) with probabilities of 63% and 76%, respectively. ΔLDLJaML demonstrated good responsiveness to rehabilitation (AUC ~0.80), with improvements >4.02% representing the optimal MCID for clinical improvement in MiniBesTest.
Conclusion: Lower LDLJa values in the AP and ML directions characterize gait smoothness impairment in PwMS. LDLJa in the ML direction is responsive to balance-focused rehabilitation, highlighting its potential for tracking gait disorders and rehabilitation progress.

Keywords: IMUs, Sensor, Multiple sclerosis, Responsiveness, Smoothness, Gait

Key messages:

- *What is already known about this topic: A series of trunk-acceleration-derived gait quality indexes are considered effective for assessing gait characteristics such as stability, symmetry, and smoothness in different populations, including patients with multiple sclerosis.*

What does the study add:

- *This study highlights the potential of smoothness index "LDLJa" to serve as a reliable and responsive measure for characterizing gait disorders and tracking rehabilitation progress in patients with multiple sclerosis.*

Received: September 4, 2024
Accepted: December 29, 2024
Published online: February 3, 2025

Study protocol number: received approval from the Local Independent Ethics Committee (CE/PROG.812)

Corresponding author:
Fulvio Dal Farra
email: fulvio.dalfarra@unibs.it

Introduction

Gait and balance impairments represent frequent issues in people with multiple sclerosis (PwMS), leading to limitations in daily activities and negatively impacting quality of life (1,2). In rehabilitation, gait training has proven to be effective in PwMS, improving gait speed, cadence, stride-to-stride variability, endurance, and postural stability. In this context, the quantitative assessment of the different gait-specific aspects is important and crucial (3,4).

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

In recent years, the use of wearable IMUs has become widespread in the clinical field due to their good reliability and validity in assessing human body movement (5), as well as their valuable properties in terms of practicability and cost-effectiveness (6). IMUs are motion-tracking devices consisting of an accelerometer, gyroscope, and (in some cases) magnetometer, which measure acceleration, angular velocity, and magnetic fields, allowing for the determination of motion and orientation.

In this context, several studies have investigated the clinometric properties of IMU sensors, obtaining satisfactory results in both healthy individuals and patients with neurological disorders (7,8).

A series of trunk-acceleration derived gait quality indexes can be retrieved from IMUs, that yield crucial information for evaluating motor abilities and monitoring recovery progress. Specifically, the normalized Root Mean Square, the short–term Lyapunov's exponent, the Harmonic ratio, the improved Harmonic Ratio, and the Log Dimensionless Jerk (LDLJ), obtained from trunk acceleration signals, are considered effective for assessing gait characteristics such as stability, symmetry, and smoothness in different populations (9-12), including PwMS (13,14). Recently, LDLJa (15), i.e., LDLJ calculated using the three-dimensional accelerations measured using a trunk-mounted IMU during a 10 Meter-Walk-Test (10MWT), has proven to be able to detect gait differences in terms of smoothness, in PwMS (16). Furthermore, LDLJa significantly improved after a rehabilitation program focusing on dynamic balance rehabilitation strategies (17). LDLJa has been described as the most robust metric for calculating gait smoothness (15,18). Indeed, it quantifies gait smoothness by calculating a normalized measure of the rate of change of acceleration that is dimensionless and scaled using the logarithm, thus removing the influence of movement duration and amplitude, allowing for comparisons across individuals and movement types (15). Gait smoothness has been described as an indicator of dynamic balance capabilities and fall risk in older adults and people with neurologic conditions (18-22). This relationship between gait smoothness and balance control is particularly relevant in people with neurological disorders, as smoothness reflects the efficiency of the neuromuscular system in coordinating motor tasks, which is closely linked to balance control (22). Additionally, a recent study revealed that balance-focused dynamic rehabilitation strategies can significantly modify gait smoothness (23). However, its discriminative ability, as well as its threshold values for identifying impairments in the gait smoothness of PwMS, has never been explored. Nevertheless, before considering LDLJa as a clinical outcome measure, its responsiveness should be assessed, which refers to its ability to change over time and align with clinical modifications after a rehabilitation intervention (24,25).

In this context, one of the most considered parameters is the Minimal Clinically Important Difference (MCID), also known as minimal important change, which is defined as the smallest change in a measure perceived as beneficial by individuals and potentially leading to different clinical decision-making (26). This definition implies that this change needs to be perceived by both clinicians and individuals. Studying the MCID could be useful for healthcare providers as it provides a benchmark to determine whether a patient's improvement or decline is clinically meaningful. This may help clinicians to make informed decisions about the effectiveness of interventions and to tailor rehabilitation programs more precisely to individual needs (27). There is no consensus on the best method to determine MCID, but the anchor-based approach is widely recognized. It uses an external criterion (the anchor), which is a well-interpretable measurement instrument, to establish what patients and clinicians consider important improvements or worsening. A commonly employed method in the anchor-based approach is the mean change, where the MCID is determined by calculating the average change in score on the assessed instrument within the subgroup of patients who have experienced minimal important change according to the anchor.

Our hypotheses were that LDLJa may accurately capture abnormalities in movement smoothness and reflect the overall postural and balance control of PwMS. We also hypothesized that LDLJa might be sensitive enough to detect clinically meaningful improvements in movement smoothness following rehabilitation. Therefore, the aims of this study were: i) to assess the ability of LDLJa to characterize gait abnormalities in PwMS in terms of movement smoothness, compared with HP; ii) to evaluate the sensitivity of LDLJa in detecting significant improvements in movement smoothness in PwMS performing the 10 Meter-Walk-Test (10MWT) following rehabilitation.

Materials and methods

Study design and participants

This cross-sectional study was conducted at the IRCCS Santa Lucia Hospital (Institute for Research and Healthcare) and received approval from the Local Independent Ethics Committee under protocol number CE/PROG.812. All procedures adhered to national and institutional ethical standards for human experimentation, the World Medical Association Declaration of Helsinki, and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines. Written consent was obtained from all participants to publish results from their clinical examinations and instrumental tests. Eligibility was determined by a researcher not involved in the assessment sessions, based on specific inclusion and exclusion criteria. Participants with MS diagnosed according to the McDonald's Criteria were recruited and enrolled through consecutive sampling between November 2020 and October 2022.

The inclusion criteria were as follows: i) a diagnosis of Relapsing-Remitting (RR) or Secondary-Progressive (SP) MS by an experienced neurologist; ii) native Italian speakers; iii) age between 28 and 71 years; iv) an Expanded Disability Status Scale (EDSS) (28) score between 0 and 6; v) ability to walk independently for at least 50 meters. Exclusion criteria included: i) the presence of psychiatric or neurological disorders (other than MS) or other conditions that could affect cognitive or motor performance; ii) a clinical relapse within three months prior to enrollment; iii) steroid therapies within 30 days prior to enrollment; iv) lower extremity fractures within three months prior to enrollment.

Each participant underwent a clinical evaluation using the Mini-BESTest, which assessed dynamic balance, postural responses, anticipatory postural adjustments, sensory orientation, and gait adaptability to changing task demands (29). As a result, data from 44 PwMS were analyzed.

For comparison purposes, gait data from 98 healthy individuals walking at preferred and self-determined slow walking speeds underwent an optimal matching the procedure using propensity score difference method (30). Age and gait speed (31,32) were input as covariates into a logistic regression model to compute propensity scores. Following an anamnestic and clinical examination of joint pain levels and range of motion, individuals with gait-affecting diseases such as peripheral neuropathies, clinically defined osteoarthritis, or joint replacement were excluded from the control group. As a result, gait data from 43 age- and gait speed–matched HP were included in this study. The effectiveness of the matching procedure was tested through a Mann – Whitney test after verifying the normality and homoscedasticity of the distributions through the Shapiro – Wilk test (Age: $p = 0.09$; gait speed: $p = 0.29$; sex: $p = 0.36$).

Sample size

A priori analysis of sample size was conducted using G*Power Version 3.1.9.4 software. Considering two groups and two repeated assessments, an effect size of 0.6, a type I error probability of 0.05, and a power effect of 0.80, the minimum required group size was 24 PwMS, accounting for dropouts (10%).

Data Collection

To ensure high-quality data, raters underwent specialized training to administer clinical outcome measures and conduct kinematic assessments. IMUs were securely attached to relevant body segments with Velcro straps to minimize oscillations and reduce motion artifacts. Two physiotherapists experienced in gait analysis with IMUs conducted the instrumental assessment and remained close to each participant during tests to prevent falls and ensure proper execution.

IMU-derived indices were measured as participants walked at their preferred gait speed along a 10-meter path, with an additional 2 meters at each end for turning, resulting in a total walking distance of 14 meters. (Fig. 1). Each task was performed three times, of which the mean value between trials was subsequently calculated and used for the statistical analysis. Data were collected at baseline (T0) and after 4 weeks of training (T1).

Participants were equipped with five synchronized IMUs (128 Hz, Opal, APDM, Portland, OR, USA), which acquired samples of three-dimensional linear accelerations and angular velocities. IMUs were located on the occipital cranium bone, near the lambdoid suture of the head (H), at the center of the sternum (S). For calculating the LDLJa, data from a single IMU placed at the L4/L5 level, just above the pelvis (P), was utilized. To segment steps and strides, data from the IMUs positioned on each lower leg, just above the lateral malleoli, were considered (33). During the initial static phase of each trial, a reference system aligned with the gravity vector was obtained. The time-invariant rotational matrix between each IMU and the defined reference system was computed and applied to the dynamic phase of the trials. As a result, accelerometer and gyroscope data were represented in a reference system roughly aligned with the antero-posterior (AP), medio-lateral (ML), and cranio-caudal (CC) anatomical axes. This method removed the acceleration due to gravity from the CC component of the accelerometer data (34). The accelerometer and gyroscope data were filtered using a second-order Butterworth low-pass filter with cutoff frequencies of 10 Hz and 6 Hz, respectively. Data were processed in the MATLAB® environment (MATLAB R2022b, MathWorks) for the computation of the LDLJa.

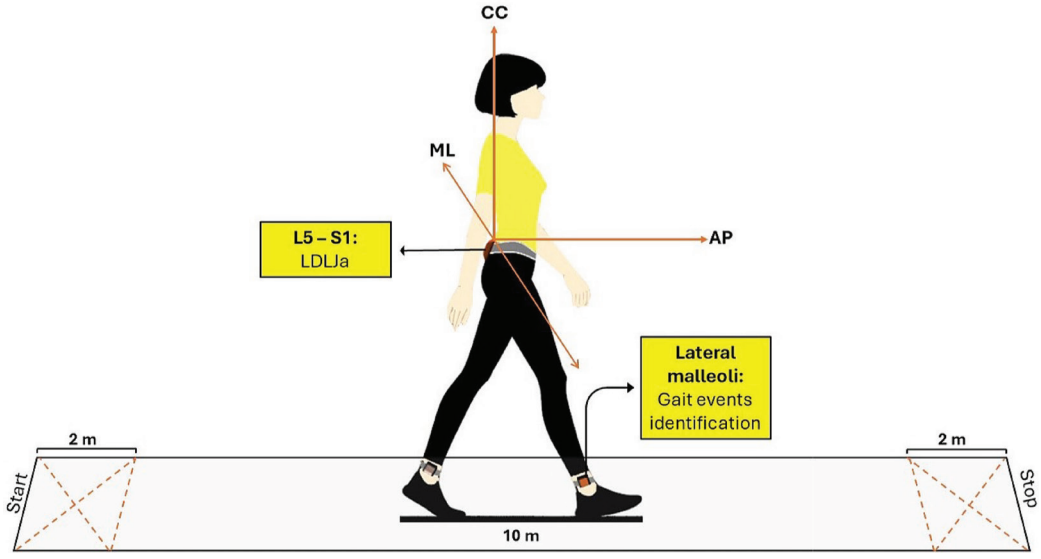


FIGURE 1 - Data collection procedure. IMU-derived indices recorded as participants walked at their gait speed along a 10-meter straight path, with an additional 2 meters at each end for turning, resulting in a total walking distance of 14 meters.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

The LDLJ was calculated at the pelvis level from the linear acceleration signals for each spatial direction according to the following formula (15):

$$\lambda_L^a(a) \triangleq -\ln\left(\frac{(t_2 - t_1)}{a_{peak}^2}\int_{t_1}^{t_2}\left|\frac{d^2}{dt^2}a(t)\right|\right);$$
$$a_{peak} \triangleq \max_{t\in[t_1,t_2]}|a(t)|$$

where $a(t)$ represents the time – domain acceleration of a movement, a normalization factor based on the peak of the acceleration profile a_{peak}, and t_1, t_2 are the start and stop times of the movement, respectively.

Interventions

The enrolled patients underwent 12 individual sessions of either conventional or cognitive-motor training, scheduled 3 days per week over a period of 4 weeks. Each session lasted 50 minutes. The conventional neuromotor rehabilitation program included 30 minutes of muscle stretching, active-assisted mobilizations, neuromuscular facilitation, gait training, and balance exercises using swinging platforms. Additionally, 20 minutes of dynamic postural stability training involved marching on unstable surfaces and walking on the treadmill with both open and closed eyes. For the cognitive-motor intervention, 20 minutes of treatment was incorporated into the conventional therapy. This involved a dual-task paradigm where each patient was instructed to walk continuously and respond to auditory stimuli by turning their head toward the sound and identifying a visual target. This dual task was performed while marching on an unstable surface and walking on the treadmill at varying velocities (23). The interventions were performed by physiotherapists not involved in the assessment.

Data analysis

After verifying the normality of the distributions through the Shapiro – Wilk test an independent sample t-test was used to identify significant differences between PwMS and HP at T0, and a paired samples t-test or Wilcoxon test was used to identify significant changes a T1, as well as Cohen's d with Hedge's correction was calculated to assess effect size.

Changes in gait variables and MiniBesTest scores at T1 were expressed as delta (Δ) values according to the following formula:

$$\Delta = (T1 - T0)/T0 * 100$$

Partial Spearman's correlation coefficient (ρ) accounting for tied scores and excluding the effects of EDSS values and gait speed were calculated to assess the correlation between the LDLJa and MiniBesTest at T0 as well as between the improvements in LDLJa and MiniBesTest.

Two logistic regression models were trained to predict binary outcomes using LDLJa scores at T0 and ΔLDLJa levels as the predictor variable with the following aims: i) to assess the ability of LDLJa to characterize the gait abnormalities of PwMS at T0 compared with HP as well as determining the related optimal cutoff value to characterize lower trunk smoothness abnormalities; ii) to assess the external responsiveness of LDLJa by using an improvement in MiniBesTest based on scores > 4 at T1 as anchor in PwMS who reported LDLJa abnormal values at T0.

To optimize and assess model performance, stratified k-fold cross-validation was implemented. This method ensured that the proportion of the two classes was maintained within each fold (35). For each iteration of cross-validation, the sample was partitioned into 5 folds. In each iteration, one fold was used as the validation set, and the remaining k-1 folds were used to train the model. Each sample was used once in the validation set and k-1 times in the training set across the k iterations. Therefore, a subject included in the validation fold in any given iteration was not included in the training folds during that same iteration (36). As a result, in the first model, the cross-validation procedure included 69 participants in the training sets and 17 participants in the validation sets for each iteration. The second model used 34 pwMS for training and 10 pwMS for validating each iteration.

To address the class imbalance between pwMS and HP, as well as between improved and non-improved PwMS, the Synthetic Minority Over-sampling Technique (SMOTE) was applied exclusively to the training data in each fold (6,37). Receiver Operating Characteristic (ROC) curves were plotted for each fold, and bootstrapping was used to estimate the 95% confidence intervals (38). The optimal cut-off point (OCP) for classifying PwMS from HP, as well as the MCID score reflecting an improvement ≥ 4 in the MiniBEsTest score (29), were determined as the values that maximized the sum of sensitivity and specificity. After cross-validation, the OCPs and MCIDs obtained from the training fold were then applied to the validation fold set to evaluate the model's performance on unseen data. Sensitivity (Se), specificity (Sp), AUC, F1 score, precision, and positive and negative likelihood ratios (LR+ and LR-, respectively), were calculated for both the training and validation sets within each fold and for the test set at the optimal thresholds. Likelihood ratios were transformed into positive and negative likelihood ratios (PTP+ and PTP-, respectively) through a Fagan's nomogram. The performance metrics were averaged across all folds to provide an overall assessment of the model (39). Additionally, the final optimal threshold was calculated as the mean of the optimal thresholds determined in each fold.

Results

IMU-measures were obtained from forty-four PwMS. Twenty-one participants were allocated to the conventional group, and 23 to the cognitive-motor group. The two groups were homogeneous with respect to demographic and anthropometric characteristics, and clinical status. Clinical and demographic characteristics of the 44 participants are described in Table 1. Two participants of the cognitive-motor group dropped out for reasons not related to the study, and data from 21 participants for each group were analyzed at T1.

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

TABLE 1 - Sample characteristics

		PwMS (n = 44)	HP (n = 43)
Disease duration, years [mean (SD)]		12.93 (9.73)	
EDSS [mean (SD)]		4.19 (1.58)	
MiniBesTest [mean (SD)]	T0	18.73 (6.75)	
	T1	20.78 (6.06)	
Age, years [mean (SD)]		49.76 (11.08)	57.57 (10.69)
Sex [n (%)]	F	12 (30 %)	17 (39.53 %)
	M	28 (70 %)	26 (60.46 %)
Gait speed (m/s)	T0	0.94 (0.37)	1.06 (0.23)
	T1	1.10 (0.39)	
LDLJaAP [mean (SD)]	T0	–5.22 (0.39)	–4.92 (0.46)
	T1	–5.17 (0.33)	
LDLJaML [mean (SD)]	T0	–5.48 (0.31)	–5.21 (0.24)
	T1	–5.29 (0.28)	
LDLJaCC [mean (SD)]	T0	–5.06 (0.43)	–5.02 (0.38)
	T1	–5.02 (0.33)	

PwMS, people with multiple sclerosis; HP, age and gait speed – matched healthy participants; EDSS, expanded disability status scale; SD, standard deviation; T0, assessment at baseline; T1, assessment following the rehabilitation period; LDLJa, log-dimensionless jerk score of acceleration signals; AP, antero-posterior direction of the acceleration signal; ML, medio-lateral direction of the acceleration signal; CC, cranio – caudal direction of the acceleration signal.

Ability to characterize gait smoothness differences of PwMS compared with HP

Significant differences between PwMS and HP were found in LDLJaAP ($p < 0.001$, $d = 0.63$), LDLJaML ($p < 0.001$; $d = 1.08$), and LDLJaCC ($p = 0.03$, $d = 0.68$). LDLJaAP and LDLJaML showed sufficient – to – moderate ability to identify the gait smoothness differences of PwMS compared with HP (Table 2). Lower LDLJaAP than –4.88 and lower LDLJaML values than –5.40 characterize PwMS with about 63% and 76% probability, respectively. The mean ROC curves, along with the confidence intervals and individual fold ROC curves, are shown in Fig. 2. The performance metrics are reported in Table 2. LDLJaCC did not reveal satisfactory discriminative ability, as shown by the AUC values around 0.50. Therefore, performance metrics and cutoff analysis were not conducted.

LDLJaML correlated with MiniBesTest values at T0, regardless of EDSS and gait speed values ($\rho = 0.35$, $p = 0.03$). LDLJaAP ($\rho = 0.23$, $p = 0.10$) and LDLJaCC (($\rho = 0.03$, $p = 0.79$) did not reveal a significant correlation with MiniBesTest values.

Responsiveness findings

Significant differences between T0 and T1 were found in LDLJaML (Fig. 3) and MiniBesTest ($p < 0.001$; $d = 0.68$), with LDLJaML values approaching HP values at T1. No significant differences between T0 and T1 were found in LDLJaAP ($p = 0.44$, $d = 0.12$). ΔLDLJaML correlated with ΔMiniBesTest ($\rho = 0.54$, $p < 0.00$). Fifteen individuals (26%) improved their MiniBesTest scores ≥ 4 points after rehabilitation. ΔLDLJaML showed good external responsiveness to rehabilitation (AUC = around 0.80, Fig. 4) with improvements >4.02 % representing the optimal MCID to reflect a clinical improvement in MiniBesTest, as reported in Table 3.

TABLE 2 - Performance metrics of LDLJaAP and LDLJaML in classifying PwMS and HP

		AUC (95% CI)	OCP	Se (95% CI)	Sp (95% CI)	F1 score (95% CI)	LR+ (95% CI)	LR-(95% CI)	PTP+	PTP-
LDLJaAP	Training folds	0.67 (0.64–0.69)	–4.88	0.87 (0.79–0.95)	0.49 (0.43–0.56)	0.73 (0.71–0.76)	1.73 (1.66–1.81)	0.25 (0.15–0.34)	63%	20%
	Validation folds	0.67 (0.56–0.78)		0.86 (0.73–0.98)	0.37 (0.24–0.51)	0.68 (0.62–0.74)	1.40 (1.26–1.54)	0.41 (0.26–0.52)	58%	29%
LDLJaML	Training folds	0.74 (0.70–0.78)	–5.40	0.65 (0.55–0.76)	0.75 (0.63–0.89)	0.69 (0.65–0.72)	3.20 (2.20–4.75)	0.46 (0.42–0.50)	76%	32%
	Validation folds	0.73 (0.54–0.91)		0.53 (0.23–0.84)	0.74 (0.41–0.99)	0.56 (0.32–0.81)	2.17 (1.47–3.05)	0.67 (0.55–0.82)	68%	40%
LDLJaCC	Training folds	0.50 (0.42–0.58)	n.a	n.a	n.a	n.a	n.a	n.a	n.a	n.a
	Validation folds	0.39 (0.22–0.55)								

AUC,area under the receiver operating characteristics curve; OCP, optimal cutoff point; Se, sensitivity; Sp, specificity; CI, confidence intervals; LR+, positive likelihood ratio; LR-, negative likelihood ratio; PTP+, positive post- test probability; PTP-, negative post – test probability; n.a., cutoff analysis not applicable because of the insufficient discriminative ability.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

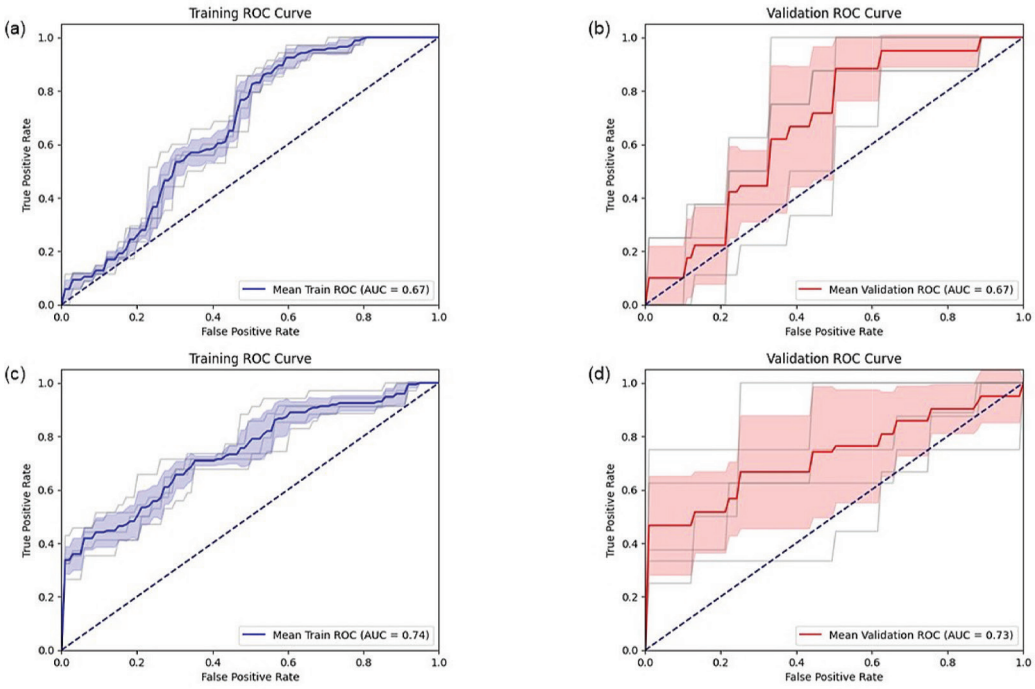


FIGURE 2 - ROC curves for training (a) and validation (b) folds of LDLJaAP, and for training (c) and validation (d) folds of LDLJaML.

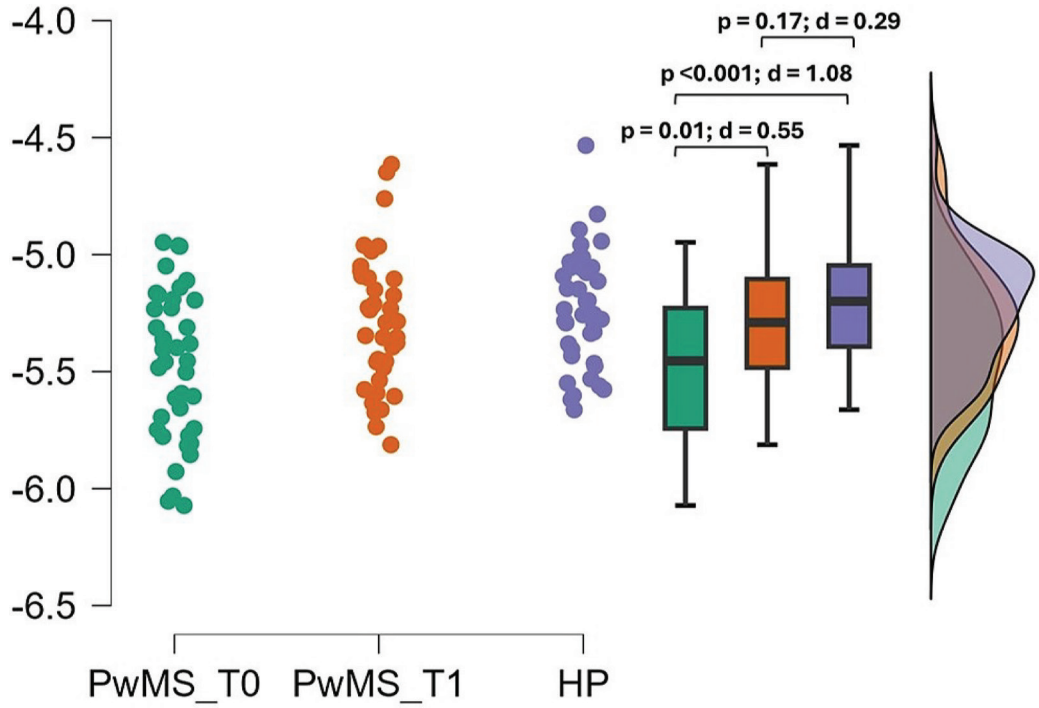


FIGURE 3 - Improvements in LDLJaML between T0 and T1.

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

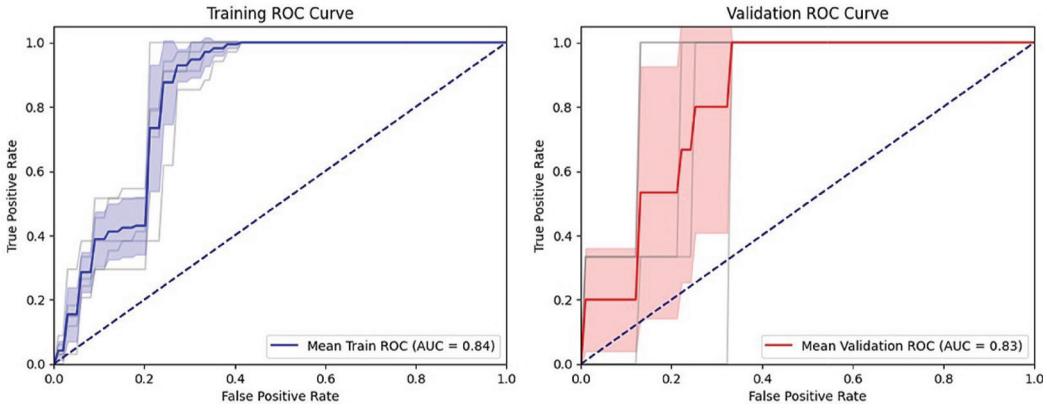


FIGURE 4 - Responsiveness of ΔLDLJaML.

TABLE 3 - Performance metrics for responsiveness of ΔLDLJaML in reflecting clinical improvement

	AUC (95% CI)	MCID	Se (95% CI)	Sp (95% CI)	F1 score (95% CI)	LR+ (95% CI)	LR- (95% CI)	PTP+ (95% CI)	PTP- (95% CI)
Training folds	0.84 (0.81–0.87)	> 4.02%	0.95 (0.87–0.99)	0.74 (0.87–0.99)	0.86 (0.81 – 0.90)	3.76 (3.34 – 4.26)	0.07 (0.01–0.14)	57% (54%–60%)	2% (0%–5%)
Validation folds	0.83 (0.70–0.96)		0.80 (0.60–0.99)	0.65 (0.49–0.87)	0.55 (0.14 – 0.96)	3.16 (1.12–5.60)	0.30 (0.00–0.90)	53% (28%–66%)	10% (0%–24%)

Discussion

The purpose of this study was to investigate the ability of the LDLJa to characterize gait differences in PwMS compared to HP, regardless of age or gait speed, and to assess its responsiveness to rehabilitation in terms of its ability to detect clinically meaningful improvements in movement smoothness following rehabilitation. We found significant differences in LDLJa values in the three anatomical directions (antero-posterior, medio-lateral and cranio-caudal) between PwMS and HP, regardless of age and gait speed, with lower LDLJa values indicating gait smoothness reductions in PwMS (Table 2). However, only LDLJaAP and LDLJaML showed sufficient – to – moderate ability to characterize the gait abnormalities of PwMS, as resulted from AUCs nearing 0.70. The OCP for classifying PwMS was found to be less than -4.88 for the AP direction and -5.40 for the ML direction, with a probability to correctly identify persons with MS at the OCP of around 63% to 76%, respectively. These cutoff values discriminate well also in unknown samples, as confirmed by the results in the validation folds set in the cross-validation model, thus increasing the external validity of the results (Table 2). Performance metrics were consistent across the training and validation samples, thus confirming the robustness of the model. As a result, LDLJaAP and LDLJaML can be considered as accurate metrics for identifying gait abnormalities in PwMS, consistent with previous studies highlighting the importance of gait smoothness in this population also in the early stages of the disease (16). Furthermore, LDLJaML positively correlated with MiniBesTest scores, with higher smoothness in medio–lateral direction reflecting higher dynamic postural control. Overall, these results highlight that PwMS experience an abruption of trunk smoothness during gait, and that lower smoothness in medio–lateral direction correlates with balance issues. This is consistent with the reported increased ML sway in the PwMS group, reflecting a reduced ability to control side-to-side motion when compared to controlling antero-posterior sway (40-42).

Nevertheless, only LDLJaML showed significant improvements following rehabilitation, and ΔLDLJaML moderately correlated with ΔMini-BESTest (ρ = 0.54). A greater improvement than 4.02% in LDLJaML was identified as the MCID for reflecting clinically meaningful improvements in smoothness. The AUC for responsiveness was approximately 0.80, reinforcing the utility of LDLJaML as a responsive measure for tracking rehabilitation outcomes, with consistent results between training and test performance metrics. Conversely, Gulde and colleagues (43) reported gait speed, but not smoothness to change following rehabilitation in a similar PwMS sample. However, the discrepancies in smoothness results can rely on the different metrics used for quantifying gait smoothness, namely the spectral arc length and the signal–to–noise ratio instead of LDLJ, which better discriminates smoothness of gait compared with other metrics (24). Moreover, Gulde and colleagues employed the 10- and 6-minute walking tests, instructing participants to walk at their maximal gait speed, which complicates the comparison with our results. Noteworthy, participants in the study by Gulde and colleagues (44) underwent a scheduled multidisciplinary rehabilitation treatment, whereas, in our study, PwMS performed a rehabilitation program specifically focused on dynamic postural stability, which could partially explain the contradictory results between the two studies. Smoothness during comfortable and fast walking speeds may engage distinct neuromuscular

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

strategies. Comfortable walking prioritizes efficiency and balance, which aligns with improvements captured by the Mini-BEST, a tool focused on dynamic balance. In contrast, fast walking emphasizes propulsion and endurance, potentially overshadowing improvements in smoothness. While the 10-meter walk test provides a snapshot of walking patterns over short distances relevant to daily activities (45,46), it may not fully encompass the broader ecological context of community ambulation, particularly in populations with greater disability levels (e.g., EDSS > 4). Tests like the 2-minute or 6-minute walk test are better suited for evaluating endurance and community mobility. Future studies could integrate longer walking tests to better capture endurance-related parameters and their interaction with smoothness in a broader ecological context. The findings of our study have significant clinical implications. The established MCID for LDLJaML provides clinicians with a quantifiable target for rehabilitation interventions aimed at improving gait smoothness in PwMS. Anchoring the Mini-BESTest, a widely used clinical scale for balance assessment, bridges the gap between quantitative gait analysis and clinical practice. This enhances the interpretability and clinical relevance of LDLJaML, facilitating its integration into routine clinical assessments and personalized rehabilitation programs. To our knowledge, no clinical scales specifically assess gait smoothness, a dynamic motor ability reflecting the efficiency of neuromuscular control during movement. The Mini-BEST was chosen as an anchor due to its focus on dynamic balance, which shares underlying mechanisms with smoothness, such as postural control and motor coordination. However, this approach has limitations, as the Mini-BEST primarily evaluates dynamic balance. Furthermore, the robustness of our findings is supported by the applied statistical procedure, including the use of stratified k-fold cross-validation and data balancing techniques to address class imbalance (6). The consistency of performance metrics across training and test sets indicates the reliability and generalizability of our logistic regression models, thus supporting the external validity of the results. The relatively small sample size led to an undersized test set, which contributed to some variability in the validation fold results. Nevertheless, the consistency of performance metrics across training and validation sets supports the reliability and generalizability of our logistic regression models, reinforcing the external validity of the findings. To further validate and refine the MCID thresholds for gait smoothness, larger pragmatic multicenter studies are needed (14). Despite the promising results, the following limitations should be considered. The sample size, though adequate for initial validation, limits the generalizability of the findings. Furthermore, although the abnormal gait pattern in PwMS is independent of the disease phenotype (47), this study did not account for heterogeneity in the sample's walking ability, as well as MS subtypes and impairment levels.

Additionally, the study focused on a specific cohort of PwMS undergoing a new dynamic rehabilitative approach, which may not reflect conventional treatments. Future research should explore the longitudinal stability of LDLJaML measurements and their predictive value for long-term clinical outcomes. Integrating additional sensor-based gait parameters could provide a more comprehensive assessment of gait abnormalities and their response to rehabilitation. Moreover, utilizing multiple inertial sensors can enhance the collection of kinematic data, offering a more detailed analysis of movement patterns.

Conclusions

Lower LDLJa values in the AP and ML directions characterize the impairment in smoothness of gait in PwMS, compared to HP, regardless of age and gait speed. These differences indicate reduced smoothness, which is associated with compromised balance control. Furthermore, LDLJa in the medio–lateral direction can be considered as a responsive metric to balance–focused rehabilitation interventions, thus highlighting the potential of LDLJa to serve as a reliable and responsive measure for characterizing gait disorders and tracking rehabilitation progress in PwMS. These insights contribute to the development of targeted, data-driven rehabilitation strategies aimed at enhancing balance and gait in PwMS.

Disclosures

Conflict of interest: The authors declare no conflict of interest.

Financial support: This research received no external fundings.

Author's contributors role: Conceptualization: MT, SFC, FDF. Methodology: MT, SFC, FDF. Data curation: SFC, DT, MS, MT. Formal analysis: SFC, DT, MS, MT. Investigation: SFC, FDF, MT. Writing-original draft: SFC, FDF, MT. Writing-review and editing: all the authors. Supervision: AT, EB, MT.

Data availability statement: The data presented in this study are available upon reasonable request to the corresponding author.

References

1. Comber L, Galvin R, Coote S. Gait deficits in people with multiple sclerosis: A systematic review and meta-analysis. *Gait Posture*. 2017;51:25-35. CrossRef PubMed
2. Gunn H, Creanor S, Haas B, Marsden J, Freeman J. Risk factors for falls in multiple sclerosis: an observational study. *Mult Scler*. 2013;19(14):1913-1922. CrossRef PubMed
3. Galperin I, Mirelman A, Schmitz-Hübsch T, et al. Treadmill training with virtual reality to enhance gait and cognitive function among people with multiple sclerosis: a randomized controlled trial. *J Neurol*. 2023;270(3):1388-1401. CrossRef PubMed
4. Corrini C, Gervasoni E, Perini G, et al. Mobility and balance rehabilitation in multiple sclerosis: A systematic review and dose-response meta-analysis. *Mult Scler Relat Disord*. 2023; 69:104424. CrossRef PubMed
5. Gordt K, Gerhardy T, Najafi B, Schwenk M. Effects of Wearable Sensor-Based Balance and Gait Training on Balance, Gait, and Functional Performance in Healthy and Patient Populations: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Gerontology*. 2018;64(1):74-89. CrossRef PubMed
6. Trabassi D, Castiglia SF, Bini F, et al. Optimizing Rare Disease Gait Classification through Data Balancing and Generative AI: Insights from Hereditary Cerebellar Ataxia. *Sensors (Basel)*. 2024;24(11):3613. CrossRef
7. Belluscio V, Bergamini E, Tramontano M, et al. Gait Quality Assessment in Survivors from Severe Traumatic Brain Injury: An Instrumented Approach Based on Inertial Sensors. *Sensors*

(Basel). 2019;19(23):5315. CrossRef PubMed
8. Tramontano M, Orejel Bustos AS, Montemurro R, et al. Dynamic Stability, Symmetry, and Smoothness of Gait in People with Neurological Health Conditions. *Sensors (Basel)*. 2024;24(8): 2451. CrossRef PubMed
9. Buckley C, Galna B, Rochester L, Mazzà C. Upper body accelerations as a biomarker of gait impairment in the early stages of Parkinson's disease. *Gait Posture*. 2019;71:289-295. CrossRef PubMed
10. Castiglia SF, Trabassi D, Tatarelli A, et al. Identification of Gait Unbalance and Fallers Among Subjects with Cerebellar Ataxia by a Set of Trunk Acceleration-Derived Indices of Gait. *Cerebellum*. 2023;22(1):46-58. CrossRef PubMed
11. Castiglia SF, Tatarelli A, Trabassi D, et al. Ability of a Set of Trunk Inertial Indexes of Gait to Identify Gait Instability and Recurrent Fallers in Parkinson's Disease. *Sensors (Basel)*. 2021; 21(10):3449. CrossRef
12. Tramontano M, Manzari L, Bustos ASO, et al. Instrumental assessment of dynamic postural stability in patients with unilateral vestibular hypofunction during straight, curved, and blindfolded gait. *Eur Arch Otorhinolaryngol*. 2024;281(1):83-94. CrossRef PubMed
13. Spain RI, St George RJ, Salarian A, et al. Body-worn motion sensors detect balance and gait deficits in people with multiple sclerosis who have normal walking speed. *Gait Posture*. 2012;35(4):573-578. CrossRef PubMed
14. Angelini L, Carpinella I, Cattaneo D, et al. Is a Wearable Sensor-Based Characterisation of Gait Robust Enough to Overcome Differences Between Measurement Protocols? A Multi-Centric Pragmatic Study in Patients with Multiple Sclerosis. *Sensors (Basel)*. 2019;20(1):79. CrossRef
15. Melendez-Calderon A, Shirota C, Balasubramanian S. Estimating Movement Smoothness From Inertial Measurement Units. *Front Bioeng Biotechnol*. 2021;8:558771. CrossRef
16. Pau M, Mandaresu S, Pilloni G, et al. Smoothness of gait detects early alterations of walking in persons with multiple sclerosis without disability. *Gait Posture*. 2017;58:307-309. CrossRef PubMed
17. Tramontano M, Belluscio V, Bergamini E, et al. Vestibular Rehabilitation Improves Gait Quality and Activities of Daily Living in People with Severe Traumatic Brain Injury: A Randomized Clinical Trial. *Sensors (Basel)*. 2022;22(21):8553. CrossRef
18. Germanotta M, Iacovelli C, Aprile I. Evaluation of Gait Smoothness in Patients with Stroke Undergoing Rehabilitation: Comparison between Two Metrics. *Int J Environ Res Public Health*. 2022;19(20):13440. CrossRef
19. Miller HL, Templin TN, Fears NE, Sherrod GM, Patterson RM, Bugnariu NL. Movement smoothness during dynamic postural control to a static target differs between autistic and neurotypical children. *Gait Posture*. 2023;99:76-82. CrossRef PubMed
20. Dixon PC, Stirling L, Xu X, Chang CC, Dennerlein JT, Schiffman JM. Aging may negatively impact movement smoothness during stair negotiation. *Hum Mov Sci*. 2018;60:78-86. CrossRef PubMed
21. Suri A, Rosso AL, VanSwearingen J, et al. Mobility of Older Adults: Gait Quality Measures Are Associated With Life-Space Assessment Scores. *J Gerontol A Biol Sci Med Sci*. 2021;76(10): e299-e306. CrossRef PubMed
22. Garcia FDV, da Cunha MJ, Schuch CP, Schifino GP, Balbinot G, Pagnussat AS. Movement smoothness in chronic post-stroke individuals walking in an outdoor environment – A cross-sectional study using IMU sensors. *PLoS One*. 2021;16(4): e0250100. CrossRef PubMed
23. Tramontano M, Argento O, Orejel Bustos AS, et al. Cognitive-motor dual-task training improves dynamic stability during straight and curved gait in patients with multiple sclerosis: a randomized controlled trial. *Eur J Phys Rehabil Med*. 2024; 60(1):27-36. CrossRef PubMed
24. Antonelli M, Caselli E, Gastaldi L. Comparison of Gait Smoothness Metrics in Healthy Elderly and Young People. *Appl Sci (Basel)*. 2024;14(2):911. CrossRef
25. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol*. 2000;53(5):459-468. CrossRef PubMed
26. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol*. 2003;56(5):395-407. CrossRef PubMed
27. Baert I, Freeman J, Smedal T, et al. Responsiveness and clinically meaningful improvement, according to disability level, of five walking measures after rehabilitation in multiple sclerosis: a European multicenter study. *Neurorehabil Neural Repair*. 2014;28(7):621-631. CrossRef PubMed
28. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*. 1983;33(11):1444-1452. CrossRef PubMed
29. Franchignoni F, Horak F, Godi M, Nardone A, Giordano A. Using psychometric techniques to improve the Balance Evaluation Systems Test: the mini-BESTest. *J Rehabil Med*. 2010;42(4): 323-331. CrossRef PubMed
30. Godi M, Franchignoni F, Caligari M, Giordano A, Turcato AM, Nardone A. Comparison of reliability, validity, and responsiveness of the mini-BESTest and Berg Balance Scale in patients with balance disorders. *Phys Ther*. 2013;93(2): 158-167. CrossRef PubMed
31. Yao XI, Wang X, Speicher PJ, et al. Reporting and Guidelines in Propensity Score Analysis: A Systematic Review of Cancer and Cancer Surgical Studies. *J Natl Cancer Inst*. 2017;109(8):djw323. CrossRef PubMed
32. Huijben B, van Schooten KS, van Dieën JH, Pijnappels M. The effect of walking speed on quality of gait in older adults. *Gait Posture*. 2018;65:112-116. CrossRef PubMed
33. Lowry KA, Vanswearingen JM, Perera S, Studenski SA, Brach JS. Walking smoothness is associated with self-reported function after accounting for gait speed. *J Gerontol A Biol Sci Med Sci*. 2013;68(10):1286-1290. CrossRef PubMed
34. Trojaniello D, Cereatti A, Pelosin E, et al. Estimation of step-by-step spatio-temporal parameters of normal and impaired gait using shank-mounted magneto-inertial sensors: application to elderly, hemiparetic, parkinsonian and choreic gait. *J Neuroeng Rehabil*. 2014;11:152. CrossRef
35. Bergamini E, Ligorio G, Summa A, Vannozzi G, Cappozzo A, Sabatini AM. Estimating orientation using magnetic and inertial sensors and different sensor fusion approaches: accuracy assessment in manual and locomotion tasks. *Sensors (Basel)*. 2014;14(10):18625-18649. CrossRef
36. Jiang D, Huang J, Zhang Y. The cross-validated AUC for MCP-logistic regression with high-dimensional data. *Stat Methods Med Res*. 2013;22(5):505-518. CrossRef PubMed
37. Eertink JJ, Heymans MW, Zwezerijnen GJC, Zijlstra JM, de Vet HCW, Boellaard R. External validation: a simulation study to compare cross-validation versus holdout or external testing to assess the performance of clinical prediction models using PET data from DLBCL patients. *EJNMMI Res*. 2022;12(1):58. CrossRef
38. Szeghalmy S, Fazekas A. A Comparative Study of the Use of Stratified Cross-Validation and Distribution-Balanced Stratified Cross-Validation in Imbalanced Learning. *Sensors (Basel)*. 2023;23(4):2333. CrossRef
39. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection; Proceedings of the 14th International Joint Conference on Artificial Intelligence

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

(IJCAI'95); Montreal, QC, Canada. 20–25 August 1995; pp. 1137-1145. CrossRef
40. Forman G, Scholz M. Apples-to-apples in cross-validation studies: pitfalls in classifier performance measurement. *SIGKDD Explor*. 2010;12(1):49-57. CrossRef
41. Morrison S, Rynders CA, Sosnoff JJ. Deficits in medio-lateral balance control and the implications for falls in individuals with multiple sclerosis. *Gait Posture*. 2016;49:148-154. CrossRef PubMed
42. Comber L, Sosnoff JJ, Galvin R, Coote S. Postural control deficits in people with Multiple Sclerosis: A systematic review and meta-analysis. *Gait Posture*. 2018;61:445-452. CrossRef PubMed
43. Lencioni T, Anastasi D, Carpinella I, et al. Strategies for maintaining dynamic balance in persons with neurological disorders during overground walking. *Proc Inst Mech Eng H*. 2021;235(9):1079-1087. CrossRef PubMed
44. Gulde P, Hermsdörfer J, Rieckmann P. Speed but Not Smoothness of Gait Reacts to Rehabilitation in Multiple Sclerosis. *Mult Scler Int*. 2021;2021:5589562. Published 2021 Jun 3. CrossRef
45. Paltamaa J, Sarasoja T, Leskinen E, Wikström J, Mälkiä E. Measures of physical functioning predict self-reported performance in self-care, mobility, and domestic life in ambulatory persons with multiple sclerosis. *Arch Phys Med Rehabil*. 2007;88(12):1649-1657. CrossRef PubMed
46. Kieseier, B. C., & Pozzilli, C. Assessing walking disability in multiple sclerosis. *Multiple sclerosis* 2012 (Houndmills, Basingstoke, England), 18(7), 914-924. CrossRef
47. Dujmovic I, Radovanovic S, Martinovic V, et al. Gait pattern in patients with different multiple sclerosis phenotypes. *Mult Scler Relat Disord*. 2017;13:13-20. CrossRef PubMed

Arch Physioter 2025; 15: 19-31
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3299
ORIGINAL RESEARCH ARTICLE

Perceived bioethical issues in cancer rehabilitation: a qualitative study among Italian physiotherapists

Gianluca Bertoni[1-3], **Simone Battista**[4], **Valentina Conti**[5], **Marco Testa**[1], **Sara Patuzzo**[6]

[1]Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genoa, Campus of Savona - Italy
[2]Department of Clinical and Experimental Sciences, University of Brescia, Brescia - Italy
[3]Training Unit, Azienda Socio-Sociosanitaria Territoriale di Cremona, Cremona - Italy
[4]School of Health & Society, Centre for Human Movement and Rehabilitation, Salford, Greater Manchester - UK
[5]School of Medicine and Surgery, University of Milan-Bicocca, Milan - Italy
[6]Department of Surgery, Dentistry, Paediatrics and Gynaecology, University of Verona, Verona - Italy

ABSTRACT

Introduction: Literature on bioethics in physiotherapy, particularly in cancer management, is limited. This study explores the perceived bioethical issues in cancer rehabilitation by Italian physiotherapists.
Participants: Thirty-one physiotherapists (Age: 42 ± 10.5 years; 20 women, 11 men) with expertise in cancer rehabilitation were purposefully selected.
Data Collection: Six online focus groups were conducted, guided by a focus group guide based on existing literature and refined by experts in cancer rehabilitation and bioethics.
Data Analysis: Sessions were recorded, transcribed, and analyzed using Braun and Clarke's 'Reflexive Thematic Analysis'.
Results: Four primary themes emerged: 1) *Challenges of (Non)-Disclosure in Diagnosis and Prognosis* – ethical difficulties around withholding diagnosis or prognosis information; 2) *Balancing Hope and Realism in Patient and Caregiver Expectations* – navigating hope versus realistic rehabilitation goals; 3) *Weighing Efficacy and Safety in Cancer Rehabilitation* – balancing treatment outcomes with patient safety; 4) *Decisions on Withdrawing Treatment* – ethical considerations in discontinuing treatment.
Discussion: These themes highlight common ethical dilemmas faced by physiotherapists in cancer rehabilitation, mirroring broader healthcare challenges. Addressing them requires a nuanced understanding of ethical principles within the cancer rehabilitation context.
Conclusions: The study provides insights into the bioethical issues in cancer rehabilitation, stressing the need for a patient-centered approach to navigate these challenges effectively.

Keywords: Bioethics, Neoplasm, Physiotherapy, Qualitative Research, Rehabilitation

What is already known about this topic:

- *Physiotherapists face ethical challenges in cancer rehabilitation, balancing treatment benefits with patient autonomy and well-being. However, bioethical dimensions in this context remain underexplored.*

What the study adds:

- *This study identifies specific bioethical dilemmas in Italian physiotherapists' cancer care practices, highlighting the need to explore this topic further and enhance ethical training and support systems in physiotherapy.*

Received: September 11, 2024
Accepted: January 20, 2025
Published online: February 13, 2025

This article includes supplementary material

Corresponding author:
Gianluca Bertoni
email: gianluca.bertoni@edu.unige.it

Introduction

Thanks to significant clinical and pharmacological innovations, today, we can treat oncological conditions and prolong the lives of many people with cancer who once had no hope of survival (1,2). However, the limitations of medicine still prevent us from always defeating the disease, creating difficulties in decision-making and the potential for prolonged suffering in some instances (3–5).

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

This context raises the interest of the bioethical discipline, which questions the moral implications of biomedical and biotechnological progress (6). Are the new medical opportunities to address oncological conditions always beneficial for the patient? When can intervention be classified as therapeutic obstinacy, as futile care that causes harm rather than benefit the patient? Should we support the duration of biological life even if it is at the expense of biographical life? (7). These and other similar questions fuel the bioethical debate on end-of-life issues, involving other fields of study, such as medical ethics (8) and medical deontology (9), which respectively identify ethical principles and their translation into rules of conduct aimed at guiding physicians and healthcare professionals in their profession. Does the duty to intervene to cure (ethical principle of beneficence) have boundaries (10)? How does it relate to the prohibition of harming the patient (ethical principle of non-maleficence) given that every treatment entails serious side effects (10,11)? How does it intersect with subjective concepts such as well-being and quality of life, and when is it appropriate to shift the focus from therapy to palliation (12,13)? How should communication with the patient be managed to ensure that their consent or dissent is fully informed (ethical principle of autonomy) (10,14)?

The literature has extensively investigated these questions, considering the physicians' perspectives (15–17). However, cancer management is multidisciplinary, and different health professionals work with patients. In particular, the physiotherapist's role has become an integral part of the medical team in oncology, as rehabilitation is fundamental to increasing the patient's quality of life while reducing disability levels (18,19). Consequently, ethical perspectives on this context should include physiotherapists. Nevertheless, the literature has so far focused on investigating the link between ethics and physiotherapy in general (20–33) or more partially on end-of-life issues (34–37), with little to no studies in oncology (38,39).

Hence, this qualitative focus group study aimed to identify and explore the bioethical issues in oncological rehabilitation as perceived by a group of Italian physiotherapists experts in this field. Specifically, it seeks to serve as a tool for descriptive ethics, helpful in capturing an existing reality about what physiotherapists perceive or interpret as ethical dilemmas in their professional practice in oncology.

Conducting this study in Italy represents a unique opportunity due to the country's socio-cultural context that can significantly shape bioethical considerations in cancer rehabilitation. Italian society, with its predominantly conservative values, deep-rooted Catholic traditions, and family-centered decision-making processes, creates a distinctive environment in which these ethical themes unfold (40,41). The Vatican further reinforces moral considerations aligned closely with religious teachings, which can influence healthcare providers' approaches to ethical challenges (42,43). For instance, decisions regarding the disclosure of diagnoses often involve requests from family members to shield patients from distressing information, reflecting a paternalistic view of care (40,41). Moreover, this cultural backdrop shapes the balancing act between fostering hope and maintaining realistic expectations (31,44). Due to their frequent and close contact with patients, physiotherapists often become confidants, sometimes feeling obliged to support patient optimism even when transparency might be compromised (45). Finally, these socio-cultural factors influence decisions on treatment continuation or withdrawal, particularly in terminal cases. In such instances, the wishes of family members can take precedence over other considerations. This culturally embedded context highlights the importance of considering Italy's distinctive social and ethical influences when examining bioethical practices in cancer rehabilitation.

Methods

Study Design

The authors conducted a qualitative focus group study. Qualitative research is the most effective method for gathering experts' opinions (46). A focus group is the ideal methodological tool to foster the development of peer support, as the group can help explore and clarify the views of a group of individuals (47). The study was performed per the Declaration of Helsinki and reported following the Consolidated Criteria for Reporting Qualitative Research (COREQ (48)). The COREQ summary sheet can be found in Supplementary File 1. Ethical approval was obtained from the Ethics Committee for University Research, University of Genova (Approval date: 27/06/2023; Genova), and informed consent was collected.

Participants

Study participants were recruited through purposive sampling (49). Specifically, participants in this study were carefully selected to include a range of perspectives on cancer rehabilitation. We had Italian physiotherapists with clinical and academic expertise in the field. To be considered experts and participate in the focus groups, the participants must have had at least five years of continuous experience in cancer rehabilitation or possess advanced education degrees and training in cancer rehabilitation. Participants were contacted via email and sourced through universities, oncology facilities, personal networks, and snowball sampling, with eligibility determined by analyzing their professional backgrounds. Then, participants were selected primarily through the analysis of their curricula vitae. GB conducted the curricula vitae analysis. Once GB identified the eligible participants, he recruited them by email. The email reported the purpose of the study, how the research would be conducted (e.g., through focus groups), and the confidentiality and anonymity of the data. The informed consent form was sent as an attachment, which participants were required to complete, sign, and return via email. Each focus group consisted of three to seven participants. The disparity in participant numbers across different focus groups can be attributed to the participants' practical constraints related to work and family commitments, affecting their ability to attend the scheduled online meetings. Each focus group lasted between one to two hours.

Data Collection

An open-question-based focus group guide (Table 1) was constructed, based on existing literature on cancer rehabilitation (2,50), by a physiotherapist experienced in cancer rehabilitation (GB), a physiotherapist experienced in qualitative research (SB) and a bioethicist (SP). Relevant information about Focus group facilitators and researchers' profiles can be found in Supplementary File 2 (Focus group facilitators and researchers' profiles).

TABLE 1 - Steps of the focus group and questions

	Introduction
1.	Introduction to the project and presentation of the moderators
2.	Presentation of the participants in the focus group
3.	The researchers provided a definition of ethical dilemma to the participants to have a shared language. Specifically, we reported that an ethical dilemma is a complex situation that raises moral questions and prompts reflections on what is right and what is wrong. It often involves conflicts between values, duties, or interests, challenging the morality of the actions or decisions involved. It can stem from specific circumstances, such as difficult medical decisions or ethical issues in healthcare, and requires a weighted assessment of the various factors involved to find the best possible solution.
	Questions
1.	Cancer is still a widespread condition that limits patients' quality of life. Physiotherapy is certainly useful and important for individuals with cancer. Does bioethics play a role in cancer rehabilitation?
2.	Could you tell us about the ethical dilemmas you have faced during your clinical practice in cancer rehabilitation?
3.	Are there any ethical dilemmas that we have not mentioned that are nonetheless relevant or important to you?
4.	How have you addressed these ethical dilemmas? (with a colleague/coordinator/ethics committee)
5.	Where does your sensitivity related to these bioethical issues come from? (e.g., from a course? From religion? From university training?)
6.	Is there any topic we have not touched on that you feel is important to highlight or elaborate on?

For each explored thematic area, we formulated stimulus questions to encourage dialogue and discussion among participants to answer our research question. The guide was also reviewed by two patients who underwent cancer rehabilitation to grant patients perspective in our research (51,52). Finally, a pilot interview was conducted with a lecturer in cancer rehabilitation to test the guide's relevance and understandability. The individual involved in the pilot interview is a male physiotherapist who has been working in oncological and palliative rehabilitation for 15 years. Additionally, he has taught 'Rehabilitation in Oncological and Palliative Care' for the past five academic years in a BSc in Physiotherapy at the University of Brescia. The focus groups were conducted online with only the moderators and participants. The software used for the focus groups was Microsoft Teams. Three moderators (GB, SB, and SP) were present during the focus groups. No close relationships were established before the study between the focus group moderators and the participants. No follow-up focus groups were performed. The focus groups were recorded and transcribed *verbatim*. The transcription was obtained through the software's automatic transcription feature and checked for precision and accuracy by GB and VC by comparing the transcription to the audio recording.

The recordings were preserved in a secure database and deleted after data transcription. While conducting the interview, GB anonymized the participants as 'Participant 1', 'Participant 2', etc., according to the chronological order of the interviews. This label is the only information shared with the rest of the group. The analysis of the collected data was carried out after focus group transcription.

Data Analysis

We collected descriptive data related to participants' gender, age, geographic origin, and professional role. Data analysis was performed according to the principles of Braun's and Clark's 'Reflexive Thematic Analysis' (RTA) (53). This choice was made because the research aims to identify patterns of meaning – and consequently themes – relating to the role of physiotherapists in oncology, focusing on bioethical issues in cancer rehabilitation as perceived by physiotherapists (54). More details on the characteristics of the authors to understand their standpoint in the reflective process can be found in Supplementary File 2 (Focus group facilitators and researchers' profiles). More details on the analysis process through Reflexive Thematic Analysis can be found in Table 2 (six steps of RTA) & in Supplementary File 3 (Theoretical standpoint).

Results

Six focus groups were conducted in July and September 2022 with thirty-one participants (Age: 42 ± 10,5; 32% Men N = 10; 68% Women N = 21, Table 3). Of the participants, all were expert clinicians, six were clinicians and lecturers, and one was a researcher in cancer rehabilitation who also possesses several years of clinical experience in this field. All the contacted participants accepted to partake in the study.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - Six steps of the RTA

Phases	Process	Authors' Involvement	Authors' Actions
1) Data familiarization	All authors read and reread several times the transcriptions of the focus groups. This process is fundamental to getting in contact with the data and taking notes of any insights.	All authors engaged in this phase, and they met to reflect upon their first insights	- Document theoretical and reflective thoughts: GB documented field notes ("Memos" and diary) during and after each focus group to promote reflexivity. - Keep records of all data field notes, transcripts, and reflexive diary - Prolong engagement with data and triangulate different data collection modes to increase the probability that the research findings and interpretations will be found credible: GB e VC read and reread the data (transcripts of the focus groups, memos, and reflexive diary)
2) Coding	In this phase, two researchers systematically coded the data through an open, evolving, and organic process.	GB and VC systematically coded the data. They adopted semantic data coding.	- Peer debriefing: memos were shared during research meetings for reflexive thoughts. - Audit trail of code generation: GB and VC coded data through the entire data set to identify interesting aspects in the data items that may form the basis of themes across the data set. - Documentation of all team meetings and peer debriefings to help researchers examine how their thoughts and ideas evolve as they engage more deeply with the data
3) Generating initial themes	The researchers generated initial themes from the codes, clustering similar or related codes.	GB and VC generated initial themes separately, clustering similar codes together.	- Diagramming to make sense of theme connections: GB and VC generated initial themes through deductive thematic analysis.
4) Reviewing and refining themes	The researcher reviewed the initial themes, reworking or discarding some until finding a final set of themes fitting the data.	All authors reviewed the coding and initial themes separately and then jointly and generated four themes that fit the data the most. GB and VC reviewed the agreed themes against the codes and the entire dataset.	- Themes vetted by team members: the research team frequently met to refine the themes and clearly show how each theme was derived from the data.
5) Defining and naming themes	The 'story' of each theme is developed by finalizing theme names and their definition.	All authors finalized the final themes and definitions to set the basis of the written report.	- Peer debriefing and team consensus on themes: the research team met until the final themes were reached. - Documentation of theme naming.
6) Producing the report	The authors produced the final report and refined them if necessary.	GB and VC selected the illustrative quotations from the interviews, and all authors reviewed and agreed. GB, SB, and SP led the writing of the paper, and all authors participated in this phase.	- Producing the report using direct quotes from participants. - Report on reasons for theoretical, methodological, and analytical choices throughout the entire study.

TABLE 3 - Descriptive statistics

Participant	Age	Gender	Educational level, Professional role	Y. of Expertise	Region
			1st Focus Group (November 2023)		
P1	53	Woman	BSc, Clinical expert	13	North
P2	54	Man	BSc, Clinical expert, and lecturer	30	North
P3	38	Woman	MSc, Clinical expert	40	North
			2nd Focus Group (November 2023)		
P4	55	Woman	MSc, Clinical expert	29	North
P5	39	Woman	MSc, Clinical expert, and lecturer	13	South & Islands

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Participant	Age	Gender	Educational level, Professional role	Y. of Expertise	Region
P6	28	Woman	BSc, Clinical expert	6	North
P7	41	Woman	BSc, Clinical expert	19	North
P8	43	Woman	BSc, Clinical expert	21	North
3rd Focus Group (November 2023)					
P9	35	Man	MSc, Clinical expert	11	North
P10	39	Man	MSc, Clinical expert	16	North
P11	26	Woman	MSc, Clinical expert	5	North
P12	45	Man	MSc, Clinical expert	22	North
P13	45	Woman	MSc, Clinical expert	10	South & Islands
4th Focus Group (November 2023)					
P14	60	Man	MSc, Clinical expert, and lecturer	30	North
P15	26	Man	BSc, Clinical expert	5	North
P16	37	Man	BSc, Clinical expert	8	South & Islands
P17	28	Woman	MSc, Clinical expert and lecturer	7	North
5th Focus Group (November 2023)					
P18	42	Woman	BSc, Clinical expert	19	North
P19	27	Man	MSc, Clinical expert	5	North
P20	34	Woman	MSc, Clinical expert	13	North
P21	49	Woman	BSc, Clinical expert	25	North
P22	41	Woman	MSc, Clinical expert	19	North
P23	61	Woman	MSc, Clinical expert, and lecturer	39	North
P24	54	Man	MSc, Clinical expert	28	North
6th Focus Group (November 2023)					
P25	36	Man	MSc, Clinical expert	13	North
P26	52	Woman	PhD, Clinical expert, and Researcher	27	North
P27	59	Woman	MSc, Clinical expert	35	North
P28	51	Man	MSc, Clinical expert	24	North
P29	48	Woman	MSc, Clinical expert, and lecturer	26	North
P30	30	Woman	MSc, Clinical expert	8	North
P31	42	Woman	MSc, Clinical expert	18	North

Legend: P, participant
Y, Years
BSc, Bachelor of Science
MSc, Master of Science
PhD, Doctor of Philosophy

From the analysis of the focus groups, four themes were generated (see Table 4 for an example of the coding process and relevant quotations and Supplementary File 4 for further quotations). According to our participants, these issues encapsulated the primary bioethical challenges encountered by physiotherapists in the clinical care of people with cancer: 1. 'Challenges of (Non)-Disclosure in Diagnosis and Prognosis'; 2. 'Balancing Hope and Realism in Patient and Caregiver Expectations'; 3. 'Weighing Efficacy and Safety in Cancer Rehabilitation'; 4. 'Decisions on Withdrawing Treatment'.

Theme 1: 'Challenges of (Non)-Disclosure in Diagnosis and Prognosis'

The main bioethical issue perceived by physiotherapists in cancer rehabilitation centers on the ethical dilemma of (non)-disclosure, particularly when this leads to communication challenges regarding patients' diagnoses and prognoses. This dilemma prompted the authors to generate Theme 1, which addresses the complexities physiotherapists encounter when managing situations where essential information is withheld, impacting patient understanding and informed consent. Participants reported that

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 4 - Defined codes for the generated themes with example of quotes

Theme 1: Challenges of (Non)-Disclosure in Diagnosis and Prognosis	
Codes defined by the researchers	**Example of quotes extracted from the focus groups**
Navigating Diagnosis/ Prognosis Disclosure: Implications on Informed Consent/Disagreement	'Very often we find ourselves in a situation where the patient is sent for rehabilitation without being informed about the prognosis, maintaining ignorance about the severity of their condition. This puts us face to face with the main dilemma: should we still communicate with a patient who is unaware of their situation? In the field of oncology, patients are constantly misled with experimental therapies and special protocols, thus they continue on their path without knowing the real severity of the situation. However, the deterioration of their health becomes inevitable, their body speaks to them, and we as healthcare professionals find ourselves having to confront this reality' (P1 – Woman – 53)
Bridging the Gap: Ethical and Legal Communication with Caregivers	'There are multiple motivations that drive family members or even the patient themselves to take certain positions within a family or caregiving context. We often focus on the patient's perspective, but it's important to address the ethical issue when the patient expresses a desire not to inform their family, such as their spouse or child. On the other hand, there are situations where there are constant requests for information from family members eager to be close to the sick person. This raises further ethical questions and could open new chapters of reflection on this complex topic' (P14 – Man – 60)
Addressing Uncomfortable Questions: Managing Patient Discomfort	'Questions about death are often directed at us physiotherapists, mainly for a quantitative reason - we are the profession that spends the most time directly in close contact with the patient. We are the ones who, in terms of minutes per week, spend the most time with them, and you find yourself being asked questions like: What is life? What is death? Why illness? Why me specifically? How should I face it? And well, it's not easy at all' (P9 – Man – 35)
	Theme 2: Balancing Hope and Realism in Patient and Caregiver Expectations
Codes defined by the researchers	**Example of quotes extracted from the focus groups**
Aligning Patient Expectations with Physiotherapist Goals	'For us too, often the lack of awareness of the diagnosis, when patients come to us, translates into the fact that they are referred to you. They are told, "go there, get rehabilitated, then come back to oncology and you'll do the next cycle of specific therapy." So, there's also, pardon the term, I'm being a bit blunt, but somewhat misleading communication, right? Toward the patient, in the sense that one thing is not explaining well what condition you have? And still fostering rehabilitative expectations that, unfortunately, we find ourselves in the position of having to somewhat downplay. Downplay, however, in a way, being very careful because if we go in too harshly, clearly, we devastate a person, I mean, our, maybe not taking care of the patient indirectly becomes communication, right? I mean, we don't say things, but if we don't then do them, we're saying things, I won't treat you. So, if you don't treat me, it means there's nothing to be done for me, so it becomes a very slippery slope on which we must work with the entire team' (P4 – Woman – 55)
Harmonizing Caregiver/ Family Expectations with Physiotherapist/Team Objectives	'The matter of rehabilitative treatment extending to the very end, even now of passing, is a complex and sensitive one. Typically, I assess each situation individually to determine whether to continue rehabilitation, but the decision isn't always mine alone. Sometimes, we may opt to continue passive mobilization even if the patient is in a coma, simply because the family wishes it, based on the patient's past enjoyment. The family's request to continue passive mobilization may stem from a desire to provide comfort, prevent pain, avoid stiffness, and alleviate discomfort from prolonged pressure on the anti-decubitus mattress. Even if clinical conditions suggest that rehabilitative treatment no longer offers direct rehabilitative benefits, as a physiotherapist, it can be challenging to refuse, considering the potential improvement in the patient's comfort and perceived quality of life. In such situations, our practice extends beyond traditional rehabilitation goals, addressing the emotional and relational needs of both the patient and their family. Thus, deciding whether to continue or stop rehabilitative treatment becomes a thoughtful consideration of how to compassionately meet these needs, even when direct rehabilitative benefits may be limited' (P2 – Man – 54)
	Theme 3: Weighing Efficacy and Safety in Cancer Rehabilitation
Codes defined by the researchers	**Example of quotes extracted from the focus groups**
Rehabilitation: Handling Clinical Outcomes and Risk Management	'For me, one issue is the presence of lytic bone metastases in oncology patients, because often the approach is not consistent for everyone. Let's consider a patient with a vertebral metastasis. In some cases, they'll tell you they can move and walk, while in others, they'll say no, they need a brace before they can walk. So, the approach is always a bit inconsistent, and I find the same thing among my physiotherapist colleagues. Some keep the patient in bed or barely seated, while others, like me, try until the very end. Some are more cautious, while others take some risks. My focus is on recovery, whether it's going to the bathroom or taking a few steps in the room. But at least personally, I always try to accommodate the patient's needs, compatibly with the pathology, of course. Others choose to never take risk' (P20 – Woman – 34)

Advancing Research in Cancer Rehabilitation	'I'm unsure whether my research will truly help the patient or, conversely, just wear them out even more. Quite often, we ask the patient for additional appointments and to fill out long questionnaires. It leaves me grappling with the dilemma of how much the patient is truly engaging, consciously, in a research project. I always question whether, deep down, I'm really doing them any good. Then, there's already plenty of evidence on the effectiveness of physical exercise in all stages of oncological disease, from diagnosis to the terminal phase. But this evidence often doesn't translate into clinical practice. It's a dilemma that nags at me because I keep on with my research, yet there's a lack of resources to actually apply these findings. Many times, it feels like I'm researching just to advance my career, rather than genuinely for the patients' well-being' (P26 – Woman – 2)
	Theme 4: Decisions on Withdrawing Treatment
Codes defined by the researchers	**Example of quotes extracted from the focus groups**
Identifying Therapeutic Futility: Determining When Treatment Becomes Ineffective	'Often, we find ourselves facing numerous oncology patients, and the oncologist, perhaps out of a lack of courage or to avoid admitting failure, continues to propose extreme and unrealistic treatments. We wonder why they persist in offering such unrealistic proposals, especially when the patient is exhausted and can no longer bear further treatments, radiotherapies, or oncological therapies. This amounts to a case of therapeutic obstinacy, which becomes even more apparent in the pediatric context. In these cases, children are encouraged to play every possible card to win their battle, without realizing that they are sometimes overwhelmed with numerous treatments, including physiotherapy and often orthoses like braces or similar devices. We wonder: what is a child supposed to do when, in the end, they find themselves saying "enough, I can't take it anymore" after being subjected to so many treatments?' (P1 – Woman – 53)
Emotional Management in Treatment Withdrawal	'In this process, we often do well with some patients, while with others it's a bit more challenging. This might happen because they're young or they have high expectations, especially regarding physiotherapy itself. Maybe they've had positive past experiences and they're trying to hold onto those. When a strong bond is formed with the patient, it becomes difficult to halt the treatment, even if it might be necessary for the patient's benefit and in consultation with the entire team. There are sometimes obstacles in stopping the treatment, perhaps because I also need to gain more experience, so I find it hard to stop at the exact moment it would be right' (P19 – Man – 27)

non-transparent communication on these issues creates significant problems regarding patients' informed consent/disagreement and the legitimacy of therapies administered by professionals. Additionally, the interviewees noted that these issues became even more significant when family members or caregivers were involved, especially if diagnoses/prognoses were disclosed to them instead of the patient. In such instances, relatives might request to keep the patient unaware of this information to "protect them". However, this process adds complexity and difficulty to the physiotherapist's work, as they must continuously interact with an uninformed patient who might also inquire about their health condition. The interviewees reported that negotiating this delicate balance of what is said and left unsaid complicated the professional's relationship with the patient. As outlined in the subsequent theme, unclear and ineffective communication could also lead to issues and misunderstandings concerning patient and family expectations.

Theme 2: 'Balancing Hope and Realism in Patient and Caregiver Expectations'

The second theme addressed in this study revolves around managing patient expectations and navigating between hope and realism. Participants in the focus groups highlighted two critical issues: collaboratively shaping rehabilitation goals with patients to align with their expectations and supporting family members and caregivers in understanding achievable rehabilitation objectives. Throughout the focus groups, it became apparent that managing patients' expectations poses a complex challenge for physiotherapists. Participants emphasized that establishing clear and realistic communication with patients is necessary to ensure that their expectations align with the predefined therapeutic goals. The failure to create realistic expectations was perceived to lead to dissatisfaction, disappointment, and disillusionment. Nevertheless, it is equally crucial to synchronize caregiver and family expectations with the objectives of the physiotherapy team, fostering meaningful involvement and a shared comprehension of the patient's rehabilitation trajectory. The absence of such mutual understanding seemed to foster unrealistic expectations, potentially escalating into conflicts and challenges. This dual responsibility mandates meticulous management and a delicate equilibrium between hope and realism to optimize treatment efficacy and enhance overall patient welfare. This parallels the imperative to balance clinical efficacy and safety, a concept to be further elucidated in the subsequent theme.

Theme 3: 'Weighing Efficacy and Safety in Cancer Rehabilitation'

The third theme was created based on focus group participants' opinions concerning the challenge of balancing clinical effectiveness and safety in oncological rehabilitation. According to the participants, this issue significantly impacts rehabilitation practice and rehabilitative outcomes based on the chosen approach. Some participants noted that opting for a more cautious approach may be viewed as reducing risks for the patient, albeit potentially leading to reduced outcomes. Conversely, others stressed the importance of a

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

more aggressive approach to achieve better results, even at the expense of increased risks for the patient. The risks the participants referred to included fractures in the presence of bone metastases or falls in patients with balance issues. Additionally, participants observed that these considerations extend to cancer rehabilitation research, where practitioners carefully assess the extent of risk-taking with patients and the degree to which a cautious approach should be adopted. The discussion on balancing clinical efficacy and safety in oncological rehabilitation naturally transitions to the complex issue of therapeutic relentlessness and treatment discontinuation. As professionals strive to maximize treatment outcomes for oncology patients, they are confronted with crucial ethical decisions regarding the continuation of therapies. The balance between seeking optimal results and ensuring patient safety thus becomes central in the context of the decision to continue or discontinue treatments. This sets the stage for the generation of the next fourth theme.

Theme 4: 'Decisions on Withdrawing Treatment'

The fourth theme generated in this study revolves around therapeutic futility and the emotional management of treatment withdrawal. This theme brings to the forefront critical considerations regarding treatment persistence and the challenging decisions associated with withdrawing care. During the focus groups, participants highlighted the hurdles related to therapeutic futility, stressing the importance of identifying signs indicating ineffective treatment and addressing the emotional repercussions linked to treatment withdrawal. Reflections on therapeutic futility and the perceived inefficacy of treatments underscore the need to balance pursuing therapeutic objectives and upholding the patient's dignity and quality of life. This consideration entailed addressing the ethical and emotional complexities that may arise during this process. Specifically, participants discussed physiotherapists' concerns regarding the management of pain, suffering, and disappointment in patients and their caregivers when the decision to withdraw treatments becomes necessary. This theme sheds light on the intricacies of clinical decision-making and underscores the importance of adopting an empathetic, patient-centered approach in navigating the conclusion of rehabilitative treatment.

Discussion

This study, situated within the framework of empirical or descriptive bioethics, investigates the bioethical dilemmas perceived by a group of Italian physiotherapists in the context of oncological rehabilitation. Through focus groups, we aimed to capture the bioethical issues recognized by these healthcare professionals. Our findings generated four themes: 'Challenges of (Non)-Disclosure in Diagnosis and Prognosis'; 'Balancing Hope and Realism in Patient and Caregiver Expectations'; 'Weighing Efficacy and Safety in Cancer Rehabilitation'; and 'Decisions on Withdrawing Treatment'. These themes reflect the nuanced ethical considerations that physiotherapists encounter, many of which are influenced by the unique socio-cultural context in Italy.

These themes align closely with the essential phases of the care relationship as outlined by Italian Law 219/2017: *treatment selection* (clinical appropriateness), *therapeutic proposal* (information-communication and consent or dissent), and *implementation of the intervention* (withholding or withdrawing) (see Table 5) (55,56). The care process in cancer rehabilitation, encompassing these phases, functions as a continuum where each stage intertwines with the next, presenting healthcare professionals with distinct bioethical challenges.

TABLE 5 - Tracing back the themes identified in the research to the fundamental components of the care relationship

Care Relationship	**1**	**2**			**3**	
	Treatment Selection	**Therapeutic Proposal (Information-Communication *and Consent or Dissent*)**			**Implementation of the Intervention (Withholding or Withdrawing)**	
	Theme 3	*Theme 1*		*Theme 2*	*Theme 3*	*Theme 4*
Generated Themes	Weighing Efficacy and Safety in Cancer Rehabilitation	Challenges of (Non)-Disclosure in Diagnosis and Prognosis		Balancing Hope and Realism in Patient and Caregiver Expectations	Weighing Efficacy and Safety in Cancer Rehabilitation	Decisions on Withdrawing Treatment
		Communicating the truth to the patient	The patient's questions of meaning			

In the Italian context, this regulation protects patient self-determination and the duty of healthcare professionals to provide clear information (55,56).

Healthcare professionals are required to inform patients about proposed treatments and obtain informed consent before proceeding. This legislation is particularly relevant in end-of-life or critical treatment contexts, guiding ethical practices aligned with patient self-determination. While the socio-cultural environment in Italy, with its conservative values, family-centered ethos, and Catholic influence, may shape the ethical challenges faced by physiotherapists, our study does not explore these specific influences (40,41). While this analysis suggests potential religious and cultural influences on bioethical decisions in cancer rehabilitation, we do not examine these specific influences, which could be explored in future studies. However, these cultural factors seem to particularly impact the second stage of the care relationship (Therapeutic Proposal), as detailed below, while

having less influence on the first and third phases (Treatment Selection and Implementation of the Intervention).

Treatment Selection (Clinical Appropriateness)

Theme 3, 'Weighing Efficacy and Safety in Cancer Rehabilitation'

According to the ethical principle of beneficence, a good healthcare professional pursues the patient's well-being, which is understood as clinical good and encompasses the traditional goals of medicine and related professions: safeguarding health and preserving individuals' biological lives. The evaluation of the clinical case by the healthcare professional is historically based on a conscience that guides technical-professional skills toward the most effective treatment to counteract the pathology (20). Since every medical intervention entails risks and potential adverse effects, aspects of prudence also contribute to this assessment, aiming not to violate the ethical principle of non-maleficence, understood as the prohibition of causing harm to the patient. At the moment of decision-making, when selecting the treatments to propose to the patient, the healthcare professional must, therefore, identify the effective treatment within a framework where the clinical benefits outweigh the risks, recognizing the uncertainty of medicine and the individual response to the proposed intervention. This decision-making process becomes more complex when the patient is affected by oncological pathology since the balance between beneficence and non-maleficence can be dichotomous: on the one hand, the appropriate intervention may be the most prudent, given the vulnerability of the oncological patient; on the other hand, precisely because of the presence of cancer, it may be necessary to push as much as possible, considering that there is nothing worse than the progression of the disease and its outcome. What is interesting to note is that in this hermeneutical dilemma about clinically appropriate care, the examination of its ethical proportionality is fully relevant, demonstrating how clinical decisions are never devoid of moral significance: clinical reasoning and ethical reasoning are intertwined (28). Physiotherapists, by expressing this dilemma in the present study, demonstrated that the selection of treatments, especially in the oncological field, is not a matter solely reserved for physicians, involving rehabilitation professionals as well concerning their specific competencies.

Therapeutic Proposal (Information-Communication and Consent or Dissent)

Theme 1, 'Challenges of (Non)-Disclosure in Diagnosis and Prognosis' (Communicating the truth to the patient)

The transition from paternalism to informed consent is widely recognized and legally endorsed, yet often does not effectively reflect in clinical practice, where behaviors persist that, for the "patient's good", circumvent the principle of information disclosure. It is possible that a patient, in the healthcare professional's perspective, appears to waive this right, when this might not be the case, or the patient might be shielded by the healthcare professional, often in agreement with family members, to spare them unnecessary distress and the awareness of a condition that no one would ever want to confront.

For instance, the theme of (non)-disclosure in diagnosis and prognosis underscores the tendency among family members to request that healthcare providers withhold distressing information from patients (40). While intended to "protect" patients, this paternalistic approach can place physiotherapists in ethically complex situations where they must navigate between respecting patient autonomy and complying with family wishes—a practice deeply rooted in Italian tradition (40). These instances highlight the importance of culturally sensitive communication strategies that honor individual rights and collective family ethical principles.

If physiotherapists raised this issue in this theme, it is to highlight its prevalence and impact in clinical routine, especially within oncological healthcare settings. The ethical dilemma is real, as it raises the question of whether or not to inform the person in whose care (the healthcare provider, the family members) they are placed. However, professionals also expressed concern about proceeding without the assurance that the patient is fully aware of their health condition, fearing they might create a distorted representation of reality, thereby invalidating treatment consent (29). The ethical problem intensifies in the consideration that someone might actually be substituting for the patient and their will, as well as in the fear of violating legal principles and norms requiring valid informed consent before any intervention (57). From this latter perspective, healthcare professionals' defensive stance is understandable, caught between the duty to recognize the ethical principle (and related legal norms) of patient autonomy and the relationship with the strong presence of family members.

Theme 2, 'Balancing Hope and Realism in Patient and Caregiver Expectations'

When a physiotherapist communicates with a patient regarding an oncological diagnosis and its prognosis, this communication should occur within a context of full understanding and transparency, in alignment with the information already provided by the physician. However, the physiotherapist cannot avoid confronting complex moral issues even in this scenario, as highlighted in this theme. The truth may sometimes be at odds with the patient's expectations, potentially leading to disappointment. This discrepancy can arise from an uncertain prognosis or a desire to offer moral support (30). Similarly, the challenge of balancing hope and realism (Theme 2) also reveals cultural nuances within the Italian healthcare setting. Physiotherapists frequently take on a dual role, supporting patient optimism while managing expectations around treatment outcomes (31). In Italy, preserving patient's optimism is paramount, often requiring physiotherapists to carefully balance transparent communication with a compassionate approach that aligns with ethical principles (45). This dual responsibility reflects a broader societal preference to maintain hope and protect patient morale, which may sometimes contrast with more transparent healthcare models in other contexts (31,45). By addressing these ethical dilemmas, our study underscores the need for a patient-centered approach that integrates ethical

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

sensitivity into clinical decision-making and fosters collaboration between patients, families, and healthcare professionals. Physiotherapists are guided to make ethically sound decisions that integrate professional standards and the socio-cultural context in which they operate (22,24,25,28). Such considerations highlight the importance of ethics training that includes ethics, especially in cancer rehabilitation settings where patients and families face challenging choices (21).

Like other healthcare professionals, physiotherapists found themselves balancing the conveyance of hope with a realistic view of the situation. However, a critical question emerges among physiotherapists: where does the boundary lie between authentic hope and illusion?

Implementation of the Intervention (Withholding or Withdrawing)

Theme 3, 'Weighing Efficacy and Safety in Cancer Rehabilitation'

In cancer rehabilitation, balancing efficacy and safety requires careful assessment of the patient's clinical condition and potential complications (58). Rehabilitation approaches range from more cautious strategies, which minimize risks but may limit benefits, to more aggressive interventions that, while potentially improving functional outcomes, expose patients to greater dangers, such as fractures or falls in cases of bone metastases or balance deficits (58). This dynamic between clinical benefit and potential harm (principle of beneficence vs. non-maleficence) confronts physiotherapists with complex ethical decisions.

A critical element is the oncological context, where safety is paramount but must be considered in light of quality of life and patient expectations (59–61). In particular, physiotherapists find themselves managing the delicate balance of providing treatments that are not only physically safe but also psychologically beneficial, respecting patient preferences. Consequently, the role of the physiotherapist extends beyond rehabilitation to incorporate broader considerations of what efficacy and safety mean for both the patient and their family, who may often desire intensive treatment even when the benefits are uncertain (58).

Furthermore, this discussion extends to research in the field of cancer rehabilitation, where physiotherapists and researchers must grapple with the dilemma of risk in clinical experimentation (62). On the one hand, research aims to optimize rehabilitation approaches, but on the other, it may be challenging to balance this objective with respect for patients' conditions and vulnerabilities (62). The lack of resources or specific protocols for cancer rehabilitation can further limit achieving optimal clinical outcomes without compromising safety (58,62).

Finally, in deciding whether to continue or withdraw a rehabilitative intervention, the continuous evaluation of efficacy and safety is essential (60). Physiotherapists may be reluctant to withdraw from treatment, especially with patients who respond positively or express high expectations for improvement. However, persisting in treatment without concrete benefits can turn intervention into therapeutic obstinacy, moving away from a patient-centered approach (37). In this context, open dialogue with the patient and family is essential to share an understanding of the treatment's risks and limitations, thereby ensuring an informed choice that respects the patient's dignity.

Theme 4, 'Decisions on Withdrawing Treatment'

Physiotherapy stands out from many other areas of medicine for the lack of a clear endpoint (21). It is a territory where the line between the necessary continuation of treatment and its conclusion becomes blurred. Patients might perceive further therapeutic sessions as beneficial, and the physiotherapist might share this belief (21). Although discontinuing rehabilitative intervention does not equate, in terms of fatal consequences, to the decision of Withdrawing Life-Sustaining Treatment, physiotherapists still consider it ethically problematic. The central issue concerns the actual utility of rehabilitation: how can we determine whether it is providing significant benefits to the patient or, conversely, if it is futile, unnecessary, and potentially harmful? In some cases, there are older or terminal patients who might not be fully aware and upon whom physiotherapists perform rehabilitative interventions, even though they know such interventions will not lead to the recovery of their motor functions and might even be perceived as bothersome by the individual. However, the general belief persists that rehabilitation cannot harm, primarily if its intent extends beyond restoring physical functions and encompasses a moral aspect (37). This concept raises questions for physiotherapists about ensuring that their practice is always genuinely beneficial to themselves and society. They are also concerned with the fear of abandoning the patient, especially if their training does not include palliative approaches or a broader view of rehabilitation that considers care's psychological, social, philosophical, and spiritual aspects (30,32,63).

Strengths and Limitations

This study offers valuable insights into the unique bioethical challenges faced by physiotherapists in cancer rehabilitation, an area that has received limited attention in the bioethics literature. By focusing on this specific professional group, our study contributes to a better understanding of the ethical landscape within multidisciplinary cancer care, highlighting complex issues related to the ethical principles of self-determination, beneficence, non-maleficence, and justice and equity in the rehabilitation context.

However, several limitations should be acknowledged. Firstly, the study is limited to a single country, Italy, where the specific socio-cultural and religious context—characterized by predominantly conservative values, a strong family-centered decision-making ethos, and Catholic influences—may shape the ethical perspectives encountered. This cultural specificity could limit the transferability of our findings to other settings; however, given similar social and cultural norms in various European countries, the results may still offer relevant insights for comparable contexts.

Additionally, despite efforts to recruit a diverse participant pool, the study includes only white physiotherapists

from northern Italy, potentially overlooking the perspectives of professionals from different ethnic backgrounds, regions, or underrepresented communities within Italy. This demographic limitation points to a need for future studies that encompass a broader range of backgrounds to capture the full spectrum of ethical concerns within cancer rehabilitation.

Furthermore, our study exclusively reflects the perspectives of physiotherapists, and we recognize that these may not fully align with those of patients, caregivers, or other healthcare professionals. Given that patient views on ethical issues could provide a more rounded understanding of these challenges, future research incorporating patient perspectives would be valuable.

Methodologically, the use of online focus groups enabled wider geographic participation and the inclusion of diverse viewpoints; however, this approach has inherent limitations. Online discussions may restrict the depth of non-verbal communication and subtly impact group dynamics, as participants might feel more reserved in discussing sensitive topics. Future studies might consider combining online and in-person methods to enhance the richness of participant interaction. In summary, while these limitations suggest caution in transferring our findings, this study started to fill a notable gap in existing research. By centering on physiotherapists' experiences, it sheds some light on bioethical issues specific to rehabilitation in oncology, which could inform both practice and ethical training in the profession.

Conclusion

This study explored the perceived bioethical concerns in cancer rehabilitation as expressed by Italian physiotherapists, highlighting key issues such as the selection of treatment approaches, patient self-determination, and informed consent or dissent, the balance between providing hope and maintaining realistic expectations, and the complexities around discontinuing care. These bioethical considerations echo the broader literature on ethical challenges in healthcare, suggesting that physiotherapists, like other health professionals, face nuanced ethical dilemmas in their practice (27,33). These findings underscore the need for ethical sensitivity in clinical reasoning, affirming that decision-making in physiotherapy extends beyond technical considerations to involve moral and ethical reflection.

Given the ethical complexities identified in this study, there are clear implications for education and training within the field. Integrating bioethical principles, effective communication techniques, and shared decision-making into physiotherapy training could enhance practitioners' preparedness for the ethical challenges of oncological rehabilitation. Moreover, adopting an interdisciplinary approach within training programs would promote collaborative, ethically sound decision-making in practice. This enhanced focus on ethics within education would support physiotherapists in balancing empathy with clinical transparency and safety, ultimately strengthening their ability to engage patients and families in a respectful, patient-centered manner. Although the cultural context of Italy could shape our findings, they offer valuable insights that may inform bioethical practices in similar healthcare settings. Future research should consider patient and caregiver perspectives to further illuminate these issues, as well as cross-cultural studies to assess the transferability of our findings. Expanding this research could contribute to ethically informed, patient-centered approaches to cancer rehabilitation worldwide.

Acknowledgment

This work was developed within the DINOGMI Department of Excellence framework of MIUR 2018-2022 (Legge 232 del 2016). The authors would like to thank all the participants to the focus group for bringing their expertise, insights into our research on top of the time they dedicated to us.

Disclosures

Conflicts of interest: The authors declare that there is no conflict of interest regarding the publication of this article.

Funding statements: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Authors contributor role: GB: Conceptualization – Data Curation – Formal Analysis – Investigation – Writing original draft – Writing review & Editing; SB: Conceptualization – Investigation – Methodology – Supervision – Writing review & Editing; VC: Investigation – Writing review & Editing; MT: Conceptualization – Project administration – Supervision – Writing review & Editing; SP: Conceptualization – Project administration – Supervision – Writing review & Editing.

Data availability statement: The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Ethics approval: The study was performed in respect of the Declaration of Helsinki and reported following the Consolidated Criteria for Reporting Qualitative Research (COREQ). Ethical approval was obtained from the Ethics Committee for University Research (CERA: Comitato Etico per la Ricerca di Ateneo), University of Genova (Approval date: 13/07/2023; CERA 2023.50).

References

1. Smith SR, Zheng JY, Silver J, Haig AJ, Cheville A. Cancer rehabilitation as an essential component of quality care and survivorship from an international perspective. Disabil Rehabil. 2020;42(1):8-13. PubMed CrossRef
2. Schieroni MP. La riabilitazione in oncologia. La presa in carico multidisciplinare e i percorsi riabilitativi diagnostico-terapeutici dei pazienti affetti da tumore. Edizioni Medico-Scientifiche; 2017.
3. Shih YA, Wang C, Jin S, Feng W, Lu Q. Decision Making of Artificial Nutrition and Hydration for Cancer Patients at Terminal Stage-A Systematic Review of the Views From Patients, Families, and Healthcare Professionals. J Pain Symptom Manage. 2021;62(5):1065-1078. PubMed CrossRef
4. Ahn E, Shin DW, Choi JY, et al. The impact of awareness of terminal illness on quality of death and care decision making: a prospective nationwide survey of bereaved family members of advanced cancer patients. Psychooncology. 2013;22(12):2771-2778. PubMed CrossRef
5. Luna-Meza A, Godoy-Casasbuenas N, Calvache JA, et al. Decision making in the end-of-life care of patients who are terminally ill with cancer - a qualitative descriptive study

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

with a phenomenological approach from the experience of healthcare workers. BMC Palliat Care. 2021;20(1):76. PubMed CrossRef
6. Potter VR. Bioethics: Bridge to the Future. Prentice-Hall; 1971.
7. Casella C, Graziano V, Lorenzo PD, Capasso E, Niola M. Unreasonable obstinacy: Ethical, deontological and forensic medical problems. J Public Health Res. 2018;7(3):1460. PubMed CrossRef
8. Percival T. Medical Ethics. London: Russell for J. Johnson, St. Paul's Church Yard and R. Bickerstaff, Strand; 1803.
9. Simon M. Déontologie Médicale. J.B. Bailliere; 1845.
10. Beauchamp TL, Childress JF. Principles of Biomedical Ethics. Oxford University Press; 1977.
11. Morrison RS, Meier DE. Clinical practice. Palliative care. N Engl J Med. 2004;350(25):2582-2590. PubMed CrossRef
12. Singer PA, Martin DK, Kelner M. Quality end-of-life care: patients' perspectives. JAMA. 1999;281(2):163-168. PubMed CrossRef
13. Hofmann B. On the triad disease, illness and sickness. J Med Philos. 2002;27(6):651-673. PubMed CrossRef
14. Quill TE, Brody H. Physician recommendations and patient autonomy: finding a balance between physician power and patient choice. Ann Intern Med. 1996;125(9):763-769. PubMed CrossRef
15. Hofmann B. Ethical issues with colorectal cancer screening-a systematic review. J Eval Clin Pract. 2017;23(3):631-641. PubMed CrossRef
16. Morgan G. Issues and Ethical Considerations in Pharmaco-oncogenomics. Adv Exp Med Biol. 2019;1168:91-101. PubMed CrossRef
17. Caplan AL. Cancer and bioethics: caring and consensus. Cancer. 2008;113(7)(suppl):1801-1806. PubMed CrossRef
18. Tan GA, Peiris CL, Dennett AM. Cancer survivors maintain health benefits 6 to 12 months after exercise-based rehabilitation: a systematic review and meta-analysis. J Cancer Surviv. 2022;1-22. PubMed Online
19. Rogers LQ, Courneya KS, Oster RA,et al. Physical activity intervention benefits persist months post-intervention: randomized trial in breast cancer survivors. J Cancer Surviv. 2023 Dec;17(6):1834-1846. CrossRef PubMed
20. Praestegaard J, Gard G. Ethical issues in physiotherapy-reflected from the perspective of physiotherapists in private practice. Physiother Theory Pract. 2013;29(2):96-112. PubMed CrossRef
21. Poulis I. Bioethics and physiotherapy. J Med Ethics. 2007;33(8):435-436. PubMed CrossRef Online
22. Kulju K, Suhonen R, Leino-Kilpi H. Ethical problems and moral sensitivity in physiotherapy: a descriptive study. Nurs Ethics. 2013;20(5):568-577. PubMed CrossRef
23. Richardson RW. Ethical issues in physical therapy. Curr Rev Musculoskelet Med. 2015 Jun;8(2):118-21. PubMed CrossRef
24. Delany CM, Edwards I, Jensen GM, Skinner E. Closing the gap between ethics knowledge and practice through active engagement: an applied model of physical therapy ethics. Phys Ther. 2010;90(7):1068-1078. PubMed CrossRef
25. Ladeira TL, Koifman L. Bioethics in the physical therapist's clinical practice: conceptions by faculty members and students. Rev bioét (Impr). 2021;29(3):588. CrossRef
26. Aguilar-Rodríguez M, Marques-Sule E, Serra-Añó P, Espí-López GV, Dueñas-Moscardó L, Pérez-Alenda S. A blended-learning programme regarding professional ethics in physiotherapy students. Nurs Ethics. 2019;26(5):1410-1423. PubMed CrossRef
27. Finch E, Geddes EL, Larin H. Ethically-based clinical decision-making in physical therapy: process and issues. Physiother Theory Pract. 2005;21(3):147-162. PubMed CrossRef
28. Edwards I, Braunack-Mayer A, Jones M. Ethical reasoning as a clinical-reasoning strategy in physiotherapy. Physiotherapy. 2005;91(4):229-236. CrossRef
29. Guccione AA. Ethical issues in physical therapy practice. A survey of physical therapists in New England. Phys Ther. 1980;60(10):1264-1272. PubMed CrossRef
30. Presidenza del Consiglio dei Ministri, Comitato Nazionale per la Bioetica. Bioetica e riabilitazione. 2006 Mar. Online (Accessed September 2024)
31. Pace P, Cisari C. Riflessioni bioetiche in medicina riabilitativa. Italian Journal of Rehabilitation Medicine. 2008;22(3):219-224. Online
32. Hudon A, Laliberté M, Hunt M, et al. What place for ethics? An overview of ethics teaching in occupational therapy and physiotherapy programs in Canada. Disabil Rehabil. 2014;36(9):775-780. PubMed CrossRef
33. Swisher LL. A retrospective analysis of ethics knowledge in physical therapy (1970-2000). Phys Ther. 2002;82(7):692-706. PubMed CrossRef
34. Michel TH. Do physiotherapists have a role in palliative care? Physiother Res Int. 2001;6(1):v-vi. PubMed CrossRef
35. Priscila Costa B, Duarte LA. Bioethical reflections about the finitude of life, palliative care and physical therapy. Rev bioét (Impr). 2019;27(3):510-515. CrossRef
36. Chigbo NN, Ezeome ER, Onyeka TC, Amah CC. Ethics of physiotherapy practice in terminally ill patients in a developing country, Nigeria. Niger J Clin Pract. 2015;18(7)(suppl):S40-S45. PubMed CrossRef Online
37. Spill GR, Vente T, Frader J, et al. Futility in Rehabilitation. PM R. 2019;11(4):420-428. PubMed CrossRef
38. Jedličková A. [Ethical aspects in oncology]. Vnitr Lek. 2023;69 (E-1):3-10. PubMed
39. Culyer AJ. Ethics, priorities and cancer. J Cancer Policy. 2017;11:6-11. CrossRef
40. Luciano M, Sampogna G, del Vecchio V, et al. The family in Italy: cultural changes and implications for treatment. Int Rev Psychiatry. 2012;24(2):149-156. PubMed CrossRef
41. Sampogna G, Luciano M, Del Vecchio V, et al. The effects of psychoeducational family intervention on coping strategies of relatives of patients with bipolar I disorder: results from a controlled, real-world, multicentric study. Neuropsychiatr Dis Treat. 2018;14:977-989. PubMed CrossRef
42. Garelli F. The church and Catholicism in contemporary Italy. J Mod Ital Stud. 2007;12(1):2-7. CrossRef
43. Callahan I, Loscocco K. The Prevalence and Persistence of Homophobia in Italy. J Homosex. 2023;70(2):228-249. PubMed CrossRef
44. Corbellini G. Scientists, bioethics and democracy: the Italian case and its meanings. J Med Ethics. 2007;33(6):349-352. PubMed CrossRef
45. Monaco S, Renzi A, Galluzzi B, et al. The Relationship between Physiotherapist and Patient: A Qualitative Study on Physiotherapists' Representations on This Theme. Healthcare 2022. 2022 Oct 25;10(11):2123. CrossRef
46. Moser A, Korstjens I. Series: Practical guidance to qualitative research. Part 1: Introduction. Eur J Gen Pract. 2017;23(1): 271-273. PubMed CrossRef
47. Kitzinger J. Focus Groups. Qualitative Research in Health Care: Third Edition [Internet]. 2007 Nov 26; 21–31. Online (Accessed September 2024)
48. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care. 2007;19(6):349-357. PubMed CrossRef

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

49. Campbell S, Greenwood M, Prior S, et al. Purposive sampling: complex or simple? Research case examples. J Res Nurs. 2020;25(8):652-661. PubMed CrossRef
50. Mayer RS, Engle J. Rehabilitation of Individuals With Cancer. Ann Rehabil Med. 2022;46(2):60. CrossRef
51. Rolfe DE, Ramsden VR, Banner D, Graham ID. Using qualitative Health Research methods to improve patient and public involvement and engagement in research. Res Involv Engagem. 2018;4(1):49. PubMed CrossRef
52. Battista S, Manoni M, Dell'Isola A, et al. Giving an account of patients' experience: A qualitative study on the care process of hip and knee osteoarthritis. Health Expectations. 2022 Mar 9; CrossRef
53. Braun V, Clarke V. Thematic Analysis: A Practical Guide [Internet]. SAGE Publications; 2021. Online (Accessed September 2024)
54. Ayre J, McCaffery KJ. Research Note: thematic analysis in qualitative research. J Physiother. 2022;68(1):76-79. PubMed CrossRef
55. Gazzetta Ufficiale della Repubblica Italiana. Law 219/2017 'Provisions for informed consent and advance treatment directives.'. 2018 Jan 16;12. Online (Accessed September 2024)
56. Di Paolo M, Gori F, Papi L, Turillazzi E. A review and analysis of new Italian law 219/2017: 'provisions for informed consent and advance directives treatment'. BMC Med Ethics. 2019;20(1):17. PubMed CrossRef
57. Triezenberg HL, Purtilo R. The identification of ethical issues in physical therapy practice. Phys Ther. 1996;76(10):1097-1107. PubMed CrossRef
58. Maltser S, Cristian A, Silver JK, et al. A Focused Review of Safety Considerations in Cancer Rehabilitation. PM R. 2017 Sep 1; 9(9S2):S415-28. PubMed CrossRef
59. Lee RT, Moore JA. Integrative Oncology and Ethics. Ethical Challenges in Oncology: Patient Care, Research. Educ Econ. 2017;101-119. CrossRef
60. Rodriguez MA, Lewis-Patterson P, Palos GR. Ethical Issues in Cancer Survivorship. Ethical Challenges in Oncology: Patient Care, Research. Educ Econ. 2017;85-100. CrossRef
61. Benjamin RS, Benjamin NB. Clinician Experience: Some Current Ethical Considerations. Ethical Challenges in Oncology: Patient Care, Research. Educ Econ. 2017;121-126. CrossRef
62. Hoover TC, Buzdar A. Ethical Considerations in Human Subjects Research: Emerging Issues in Cancer Research. Ethical Challenges in Oncology: Patient Care, Research. Educ Econ. 2017;145-157. CrossRef
63. Bioética R, Ladeira TL, Koifman L. Bioethics in the physical therapist's clinical practice: conceptions by faculty members and students. Rev bioét (Impr). 2021;29(3):588. CrossRef

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 32-41
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3323
ORIGINAL RESEARCH ARTICLE

Comparisons of the effects of psychologically-informed and usual physiotherapy on pain sensitivity in chronic low back pain: an exploratory randomized controlled trial

Hugo Massé-Alarie[1], Amélie Desgagnés[1], Claudia Côté-Picard[1], Olivier Liberty[1], Pierre Langevin[1], Mathieu Piché[2], Yannick Tousignant-Laflamme[3]

[1]Cirris, School of Rehabilitation Sciences, Université Laval, Quebec City - Canada
[2]Chaire de Recherche Internationale en Santé Neuromusculosquelettique, Université du Québec à Trois-Rivières, Trois-Rivières - Canada
[3]Centre de recherche du CHUS, Université de Sherbrooke, Sherbrooke - Canada

ABSTRACT

Introduction: The presence of altered central pain processing and modulation, as well as negative psychological factors, have been suggested to impede recovery in chronic low back pain (CLBP). Psychologically-informed physiotherapy (PiP) aims to specifically address the latter factors—in addition to physical factors—to improve treatment effects. This study aims to determine if the effect of PiP is superior to usual physiotherapy (UP) on pain sensitivity and modulation in participants with CLBP and if changes in these variables were associated with changes in clinical outcomes.

Methods: Forty participants with CLBP were randomly allocated to PiP or UP. Seven physiotherapy sessions over 6 weeks plus a booster session at an 11-week follow-up were delivered. Pressure pain threshold (PPT), temporal summation of pain (TSP), and exercise-induced hypoalgesia were assessed on lumbar, upper, and lower limb sites at baseline and after 6 weeks. Linear mixed models tested if PiP was superior to UP on pain sensitivity/modulation. Linear regressions tested if pain sensitivity/modulation changes were associated with changes in clinical outcomes (pain intensity, physical functioning, symptoms of central sensitization).

Results: PiP was not superior to UP to modulate pain sensitivity/modulation variables. All PPTs increased after 6 weeks regardless of the approach. Lumbar PPT and lumbar and lower limb TSP changes were associated with physical functioning changes.

Conclusion: Although our study suggests that neither approach has a superiority to impact on pain sensitivity, both approaches elicited widespread hypoalgesia. Future powered trials should verify if pain sensitivity can be a mediator of physical functioning improvement, as suggested by our results.

Keywords: Exercise-induced hypoalgesia, Pressure pain threshold, Psychologically informed physiotherapy, Quantitative sensory testing, Temporal summation of pain, Usual physiotherapy

What is already known about this topic?

- *Psychologically-informed Physiotherapy (PiP) consists of identifying and addressing both physical (including sensitization) and psychological factors contributing to pain. This approach is deemed superior to usual physiotherapy (UP) to manage chronic low back pain (CLBP).*

What does the study add?

- *PiP has been claimed to influence pain sensitivity, modulation, and psychological factors. In this exploratory clinical trial, we observed that both physiotherapy approaches reduced pressure pain sensitivity.*

Received: September 30, 2024
Accepted: January 27, 2025
Published online: February 17, 2025

Clinical Trial Protocol Number
NCT04979403

Corresponding author:
Hugo Massé-Alarie
email: hugo.masse-alarie@fmed.ulaval.ca

Introduction

Chronic low back pain (CLBP) is highly prevalent in the general population (1) and represents the leading cause of disability worldwide (2). Considering that the specific cause of CLBP is elusive in most patients (90%), the term non-specific CLBP is commonly used (3). This may be explained by the multidimensional nature of CLBP. Indeed, several factors such as psychological, social, biophysical, genetics, and comorbidities may modulate both central pain processing and nociceptive inputs, giving rise to an individual's pain experience (4). For example,

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

it has been demonstrated that participants with CLBP had reduced exercise-induced hypoalgesia (EIH) and increased pain during a muscle contraction remote from the lower back area, suggesting an alteration in central mechanisms for pain control (5). This is coherent with meta-analyses reporting higher pain sensitivity in CLBP compared to controls, although pain modulation (frequently measured using conditioned pain modulation paradigm) was altered only in some studies (6). This could reflect the presence of nociplastic pain mechanisms (7-9), which is suggested to contribute to chronicity and long-term disability. Some authors proposed, although not all agree (8-10), that patients with predominant nociplastic CLBP phenotype often present with negative psychological factors (e.g., pain catastrophizing, fear) that are predictors of poor prognosis (11). Consequently, most contemporary treatments to manage CLBP consider the multiple factors impacting the pain experience, for example, by targeting (i) nociplastic pain mechanisms and (ii) psychological factors associated with poor prognosis through psychologically-informed approaches (12-14).

Psychologically informed approaches integrate physical, behavioral, and psychological interventions to improve a patient's quality of life (15). When delivered by physiotherapists, it is referred to as psychologically-informed physiotherapy (PiP). PiP includes usual physiotherapy (UP) interventions such as manual therapy and exercises along with strategies based on cognitive-behavioral principles and pain neuroscience education (12). Although a meta-analysis demonstrated that interventions similar to PiP (i.e., combining physical and psychological interventions) are more effective in improving pain and physical functioning than usual care (16), the impact of PiP on central pain mechanisms (e.g., pain sensitivity and modulation alteration) has been scarcely studied. PiP may have an impact on pain sensitivity and modulation via different mechanisms. First, it has been suggested that targeting negative emotional and cognitive factors could influence the pain modulation system (17) through the corticolimbic system (18). For example, studies have demonstrated associations between pain catastrophizing and pain sensitivity (19). Second, pain neuroscience education—which also targets psychological factors and pain conceptualization—has been suggested to "desensitize" the central nervous system, thus influencing pain sensitivity and modulation variables and, ultimately, pain-related outcomes (20,21). Emerging evidence may support these hypotheses. A randomized controlled trial demonstrated that the combination of pain neuroscience education and exercises reduced pain sensitivity of the lower back area more than exercise alone (21). Nonetheless, interventions commonly used by physiotherapists may also influence pain sensitivity. Specifically, meta-analyses reported a reduction in pain sensitivity following manipulation (22) and exercises (23). It remains unclear whether PiP has a greater impact on pain sensitivity and modulation variables than UP. Moreover, despite the fact that some groups proposed that PiP may "desensitize" the central nervous system (20,21) and consequently affect pain-related outcomes, it is not known whether changes in pain sensitivity and modulation are associated with changes in pain and disability.

The objectives of this exploratory study are (i) to compare the effects of PiP and UP on pain sensitivity and modulation variables and (ii) to determine if changes in pain sensitivity and modulation are associated with changes in clinical outcomes. We hypothesized that (i) PiP will have a greater influence on pain sensitivity and modulation variables than UP, and (ii) changes in pain sensitivity/modulation variables will be associated with changes in patient-reported outcomes.

Methods

Study design

The complete methodological details and results (feasibility and clinical effect of physiotherapy) of this pilot and exploratory randomized controlled trial were published (24), and the protocol was prospectively registered in ClinicalTrials (NCT04979403). A brief overview of the methods is provided in the following sections.

A two-arm parallel pilot randomized clinical trial (RCT) with a 6-month follow-up was conducted. Quantitative sensory testing (QST) measures were collected at Cirris (research center) at baseline and at a 6-week follow-up. Patient-reported outcome measures (PROM) were assessed at baseline, 6 weeks, 12 weeks, and 24 weeks after randomization using Research Electronic Data Capture (REDCap) (25). Considering the PROMs have been published, only PROMs collected at baseline and 6-week follow-up are reported in this study to fulfill the second objective of the study.

All participants were randomly allocated to one of the two groups (PiP, UP) and underwent their assigned interventions. Each intervention consisted of eight sessions spread over 11 weeks. Precisely, a 60-minute assessment meeting and a follow-up session were delivered during the 1st week, followed by weekly sessions until the 6th week, and a final follow-up session (i.e., "booster" session) in the 11th week (between 30-45 mins per follow-up session).

Participants

Forty participants with CLBP were recruited according to the following inclusion criteria: (i) adults between 18 and 65 years, (ii) non-specific CLBP (>3 months), and (iii) classified at high risk of poor prognosis based on the STarT Back Screening Tool (i.e., psychosocial subscore ≥4/5) (26). Exclusion criteria were any cause of specific LBP (e.g., fracture, cancer) (3), litigation, and neuropathic pain [>4 on the *Douleur Neuropathique 4* (DN4) questionnaire (27)]. The presence of leg pain *per se* was not an exclusion criterion in the absence of signs and symptoms of neuropathic pain. Considering this study was planned as a feasibility and pilot RCT, no *a priori* sample size was calculated as recommended for this type of study (28). Nonetheless, considering $\alpha = 0.05$ and $\beta = 0.20$, an effect size of $d = 0.91$ (objective 1) and a coefficient of correlation of $r = 0.41$ (objective 2) can be detected based on the study's sample size.

The study was approved by the Ethics Research Committee of the *Centre Intégré Universitaire de Santé et de Services Sociaux de la Capitale-Nationale* (Project: # 2021-2227) in accordance with the Declaration of Helsinki. All participants provided their written informed consent prior to their inclusion in the clinical trial.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Randomization/Blinding

A random allocation sequence (1:1) was generated with the random function on an Excel file, stratified by sex and physical functioning (Oswestry Disability Index [ODI] cutoff = 20) in blocks of 4. A researcher not involved in the assessment and interventions (CCP) carried out concealed allocation, assigned participants to an intervention group, and contacted each participant after the baseline evaluation to inform them of the location of the clinics to receive treatment. The evaluator of QST measures was blinded to participants' allocation (AD). Participants were instructed not to discuss with the evaluator about the intervention received, and the physiotherapist and the clinic consulted.

Interventions

Physiotherapists in each arm worked at two different sites of a private physiotherapy clinic (PhysioInteractive group, Quebec City, Canada—urban setting) to limit contamination. Physiotherapists from one site provided UP, while physiotherapists from the other site provided PiP. UP included all interventions that a physiotherapist can provide in the province of Quebec (e.g., education, advice to stay active, exercises, manual therapy). PiP included UP along with a psychologically informed approach (12). Management strategies such as establishing common goals with the patient and therapeutic alliance, behavior change model strategies, motivational interviews, education on pain neurophysiology, gradual exposure, and stress management techniques were used in PiP to mitigate psychosocial factors/barriers. Interventions are fully detailed elsewhere (24).

Outcomes

Quantitative sensory testing (QST)

Static and dynamic pain measurements were tested using standardized verbal instructions according to the recommendations from the German Research Network on Neuropathic Pain (29). Static pain sensitivity was measured using the pressure pain threshold (PPT) (30). PPTs were measured ipsilateral to the side of worst pain for the lower back and upper limb but contralateral to pain for the lower limb. If pain was central, it was considered as a right LBP. Three sites were tested: lumbar *erector spinae* (lumbar) muscles on the muscle bulk at 2-3 cm lateral to the L4-L5 interspinous space, *tibialis anterior* (lower limb) at one-third of the distance between the upper part of the head of the fibula and the lower part of the medial malleolus, and wrist flexors (upper limb) at one-third of the distance between the medial epicondyle and the styloid process of the radius. Participants were positioned supine for upper and lower limb measurements and prone for lumbar measurements. All anatomical landmarks were marked before measurement. The evaluator used a handheld digital algometer (1-cm^2 probe—FPIX, Wagner Instruments, Greenwich, CT, USA) to apply pressure at a rate of ~0.5 $kg/cm^2/s$. A one-minute break was taken between trials. The testing order was randomized. An average of three trials was used. The algometer can measure up to 11 kg/cm^2. Out of 720 measurements, 43 (5.97%) were above this limit and considered as 11 kg/cm^2.

Dynamic pain sensitivity was measured using temporal summation of pain (TSP) (31) and EIH (5, 32). TSP is considered a proxy of the wind-up phenomenon and would represent the change in excitability of the spinal neurons (29). For TSP, three sites were tested, namely the L4-L5 interspinous space (lumbar), the dorsal part of the first cuneiform (lower limb), and the dorsal part of the capitate (upper limb). Participants were positioned in crooked lying for wrist and leg measurements and prone for lumbar measurements. The evaluator used a pinprick stimulator (256 mN, MRC Systems GmbH, Heidelberg, Germany) for a series of ten punctuate stimuli at 1 Hz. The evaluator used a visual metronome out of the participant's sight to respect the frequency of 1 Hz. A one-minute break was taken between the three trials. The average of the three trials was used. TSP was the difference between the pain rated on the numerical pain rating scale (NPRS—from 0 [no pain] to 10 [worst imaginable pain]) after ten stimuli and that after a single stimulus. TSP and PPT are reliable measurements (33,34).

EIH is characterized by a reduction in pain sensitivity following an acute bout of exercise (35,36) and is mediated by opioids and non-opioid mechanisms (e.g., endocannabinoid, serotoninergic) (36,37). To measure EIH, all PPT measurements were repeated after the performance of an isometric wrist contraction. Electromyography was used to standardize the level of muscle activation during wrist contraction by providing feedback to participants to ensure the consistency of contraction. An electromyographic wireless sensor (TrignoTM Wireless EMG System, Delsys, USA) was placed on the PPT landmark of the wrist flexors following SENIAM recommendations (38). The participant was seated, elbow at 90°, the forearm was supinated and the wrist placed under the edge of a table to perform an isometric wrist flexion. Maximal voluntary contractions (MVC) of wrist flexion were measured three times with one-minute breaks between each trial. The highest trial, measured as peak-to-peak, was retained as MVC. Then, an isometric wrist flexion contraction was maintained for 4 minutes at 25% MVC, and visual feedback was provided on a monitor. Two 5-s breaks were allowed for all participants if pain or fatigue was present (32,39). The percentage of change from post- to pre-contraction PPT measurements was calculated and represent EIH.

Patient-reported outcome measures

As reported in the published article (24), many PROMs (e.g., pain intensity, physical functioning, pain catastrophizing) were measured to explore the effects of the intervention. Pain intensity and physical functioning, which are the most widely used clinical outcomes, were used to represent clinical changes of participants with CLBP using the average pain intensity as measured with the NPRS in the last week and the Oswestry Disability Index (ODI) (40). For objective 2, the two latter measures were considered the primary outcomes. We also tested the association between symptoms of central sensitization [measured using Central Sensitization Inventory [CSI] (41)] and QST measures, considering they have been hypothesized to represent overlapping, although different, constructs (42). Other PROMs were used to characterize the study sample and as covariates/confounders in statistical analyses.

Statistical analysis

Sociodemographic data were compared with independent samples t-tests for interval variables, with Mann-Whitney

tests for ordinal variables, and with Chi-square tests for proportion variables.

For objective 1, we used linear mixed models using Group (PiP, UP), Time (baseline, 6 weeks), and Group × Time as fixed factors, participants' intercept as a random factor, and age and sex as covariates to explore the effect of the PiP compared to UP on QST outcomes (PPT, EIH, TSP). A compound symmetry covariance matrix was used, and Sidak's corrections for multiple comparisons were applied. Intention-to-treat analyses were used, i.e., each participant was analyzed in their allocated group.

Because analyses of the PROMs were published (24), no further statistical analyses on these outcomes are reported in this paper. Means and mean differences between 6-week follow-up and baseline were reported for pain intensity, physical functioning, and symptoms of central sensitization.

For objective 2, linear regressions were computed to determine if differences in QST were associated with differences in clinical profile by pooling the results from both groups. Differences between the 6-week follow-up and baseline data were measured for QST (ΔPPT, ΔEIH, ΔTSP) and clinical variables (ΔPain, ΔODI, ΔCSI). Outcomes (dependent variables) of linear regressions were ΔPain, ΔODI, ΔCSI. Independent variables were ΔPPT, ΔEIH, and ΔTSP at the different sites. Linear regressions were adjusted for potential confounders as self-reported age, sex, baseline pain catastrophizing scale and baseline symptoms of central sensitization. Potential confounders were selected based on potential causal influences with dependent and independent variables (43).

The mean and standard deviation are presented in the manuscript unless otherwise specified. A value of $p < 0.05$ was considered statistically significant.

Results

Figure 1 depicts the CONSORT flow diagram of the study. Of the 291 potential participants assessed for eligibility,

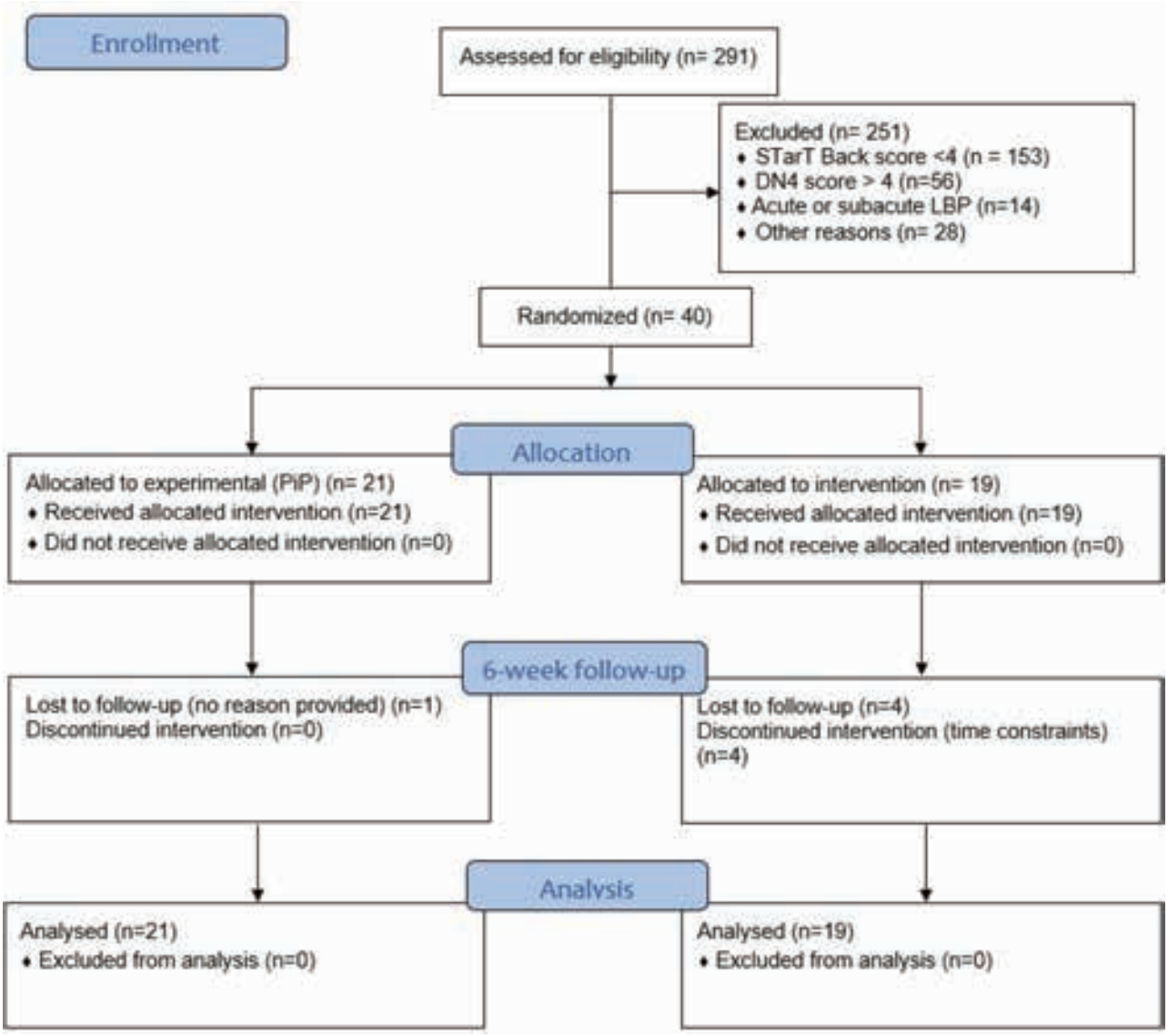


FIGURE 1 - CONSORT flowchart depicting the pilot randomized trial.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

40 were randomly allocated to one of the two groups (PiP or UP). In addition to the five participants who were lost at the 6-week follow-up, baseline, follow-up data for EIH (low back and lower limb) for one participant was not available since it reached the limit of the algometer (11 kg²/cm) at each trial, impeding the measurement of PPT change.

Sociodemographic and baseline data of participants

Sociodemographic variables and scores measured by self-reported questionnaires at baseline were compared between groups and presented in Table 1. Both groups were comparable on all variables ($p < 0.05$), except for pain catastrophizing, which was significantly higher for UP.

Quantitative sensory testing

Table 2 reports the estimated means and 95% confidence intervals extracted from the linear mixed models for all quantitative sensory testing outcomes, in addition to the within-sample difference (p-values of the effect of time) and Group × Time interaction (p-values). Specifically, a Group × Time interaction was observed for TSP of the wrist flexors ($F_{(1, 33.95)} = 4.84$; $p = 0.04$). However, pairwise comparisons did not detect differences between groups at baseline (0.06 [−0.89, 1.02]; $p = 0.90$) and at 6 weeks (−0.74 [−1.73; 0.25]; $p = 0.14$), and within-group for PiP (0.5 [−0.0, 1.0]; $p = 0.051$) or UP (−0.3 [−0.2, 0.9]; $p = 0.26$). No other interaction was detected ($p > 0.10$—Table 2). Main effects of time were observed for PPT of lumbar ($F_{(1, 33.731)} = 5.81$; $p = 0.02$), lower limb ($F_{(1, 34.61)} = 8.85$; $p = 0.01$), and upper limb ($F_{(1, 35.19)} = 6.36$; $p = 0.02$) sites, meaning that PPT at all sites increased over time (reduced pain sensitivity) regardless of the group. No other main effect of time was detected for EIH ($p > 0.48$) or TSP ($p > 0.47$). Figure 2 displays the individual changes between baseline and 6-week follow-up of all the QST outcomes.

Patient-reported outcomes measures

Table 3 describes changes in PROMs between baseline and 6-week follow-up for physical functioning, pain intensity, and symptoms of central sensitization. Although no additional analyses were undertaken in the current manuscript, significant effects of time were observed without significant Group × Time interaction; the detailed analyses are reported elsewhere (24).

Linear regressions between QST and pain-related outcomes

Table 4 reports the results of the linear regressions between changes in pain sensitivity/modulation and changes in pain-related outcomes. Out of the 27 analyses, three were statistically significant. Specifically, there was a significant negative association between the change in lumbar PPT and the change in physical functioning, meaning that each increase of 1 kg/cm² for lumbar PPT was associated with a reduction of 2.36% on the ODI score (β = −2.36 [−4.25, −0.46]; $p = 0.02$). Also, significant positive associations were detected between the lumbar and lower limb TSP and changes in physical functioning; a reduction of 1 point on TSP resulted in a reduction of 3.50% and 3.10% on ODI score, respectively (lumbar TSP: β = 3.50 [1.01, 6.00]; $p = 0.01$; foot TSP: β = 3.10 (0.38, 5.82); $p = 0.03$). No other significant associations were observed.

Discussion

The first objective of this study was to explore the effects of PiP and UP on pain sensitivity and modulation variables using PPT, EIH, and TSP. We observed an increase in PPTs at 6 weeks regardless of the interventions, suggesting a widespread induced hypoalgesia without superiority of a physiotherapy approach. These findings refute our hypothesis that PiP would have a larger influence on pain sensitivity and modulation variables. The second objective was to determine if changes in pain sensitivity and modulation were associated with changes in pain and physical functioning. Significant

TABLE 1 - Sociodemographic and baseline variables for both groups*

	UP (n = 19)	PiP (n = 21)	*p-value*
Age	33.1 (8.8)	35.8 (11.0)	0.41
Male, n (%)	10 (52.6)	12 (57.1)	1.00
LBP duration, n (%)			0.10
3-5 months	2 (11)	0 (0)	
6-11 months	4 (21)	4 (19)	
1-5 years	3 (16)	10 (48)	
More than 5 years	10 (53)	7 (33)	
Pain intensity (average last 7 days)	5.7 (2.0)	5.0 (1.9)	0.31
Physical functioning (ODI – range: 0-100)	27.1 (12.2)	23.9 (11.8)	0.41
Kinesiophobia (TSK-11 – range 11-44)	28.7 (8.1)	25.9 (7.4)	0.25
Pain catastrophizing (PCS – range: 0-52)	25.7 (8.5)	18.3 (10.8)	**0.02**
Symptoms of central sensitization (9-item CSI – 0-36)	20.7 (6.2)	17.6 (5.0)	0.09

*Data are presented as mean (SD) unless stated otherwise.
UP: Usual physiotherapy; PiP: psychologically-informed physiotherapy; LBP: low back pain; ODI; Oswestry Disability Index; TSK: Tampa Scale for Kinesiophobia; PCS: Pain Catastrophizing Scale; CSI: Central Sensitization Inventory.

TABLE 2 - Estimated means, estimated mean differences, 95% confidence interval, and p-values for quantitative sensory testing measures

	PiP			UP			ΔUP – ΔPiP*	Interaction (p-value)	Within-sample difference¥	Time effect (p-value)
	Baseline	6 weeks	ΔPiP	Baseline	6 weeks	ΔUP				
PPT (kg/cm²)										
Lumbar	5.0 [3.9, 6.1]	5.3 [4.2, 6.4]	0.3 [-0.5, 1.1]	3.6 [2.5, 4.8]	5.0 [3.9, 6.1]	1.3 [0.2, 2.0]	1.0	0.16	0.7 [0.1, 1.3]	**0.02**
Lower limb	4.8 [3.9, 5.7]	5.4 [4.5, 6.3]	0.6 [-0.1, 1.3]	4.2 [3.2, 5.1]	5.0 [4.0, 6.0]	0.9 [0.1, 1.6]	0.3	0.60	0.7 [0.2, 1.2]	**0.01**
Upper limb	3.2 [2.5, 4.0]	3.4 [2.7, 4.2]	0.2 [-0.4, 0.8]	2.5 [1.7, 3.3]	3.4 [2.6, 4.3]	0.9 [0.3, 1.6]	0.7	0.10	0.6 [0.1, 1.0]	**0.02**
EIH (% of change)										
Lumbar	3.8 [–7.9, 15.4]	1.6 [–10.3, 13.6]	–2.1 [–16.1, 11.9]	14.3 [2.3, 26.3]	15.3 [2.0, 28.6]	1.0 [–14.2, 16.3]	3.1	0.76	–0.5 [–10.9, 9.8]	0.92
Lower limb	7.4 [–3.0, 17.8]	–3.1 [–13.7, 7.6]	–10.5 [–24.2, 3.1]	4.6 [–6.1, 15.3]	8.1 [–3.9, 20.0]	3.5 [–11.3, 18.3]	14.0	0.17	–3.5 [–13.6, 6.6]	0.48
Upper limb	8.2 [–3.3, 19.7]	5.7 [–6.1, 17.4]	–2.5 [–18.3, 13.2]	9.2 [–2.8, 21.3]	5.4 [–8.2, 19.0]	–3.9 [–21.3, 13.6]	1.4	0.91	–3.2 [–15.0, 8.6]	0.59
TSP										
Lumbar	2.0 [1.4, 2.7]	2.3 [1.6, 2.9]	0.2 [–0.5, 0.8]	2.4 [1.4, 2.7]	1.9 [1.1, 2.6]	–0.6 [–1.3, 0.2]	–0.8	0.13	–0.2 [–0.7, 0.3]	0.47
Lower limb	2.5 [1.8, 3.2]	2.5 [1.8, 3.2]	0.0 [–0.7, 0.7]	2.0 [1.2, 2.7]	1.9 [1.1, 2.7]	–0.0 [–0.8, 0.7]	–0.1	0.90	–0.0 [–0.5, 0.5]	0.95
Upper limb	1.9 [1.2, 2.5]	2.4 [1.7, 3.0]	0.5 [–0.0, 1.0]	1.9 [1.2, 2.6]	1.6 [09, 2.3]	–0.3 [–0.2, 0.9]	–0.8	**0.04**	0.1 [–0.3, 0.5]	0.65

*ΔUP – ΔPiP: Between-group difference at 6 weeks; ¥ Within-group difference of the whole sample (Post – pre).
PiP: psychologically-informed physiotherapy; UP: usual physiotherapy; ΔPiP: Within-group difference (Post-pre) for PiP; ΔUP: Within-group difference (Post-pre) for PiP; PPT: Pain pressure threshold; EIH: Exercise-induced hypoalgesia; TSP: Temporal summation of pain.
Bold indicates significant p-values.

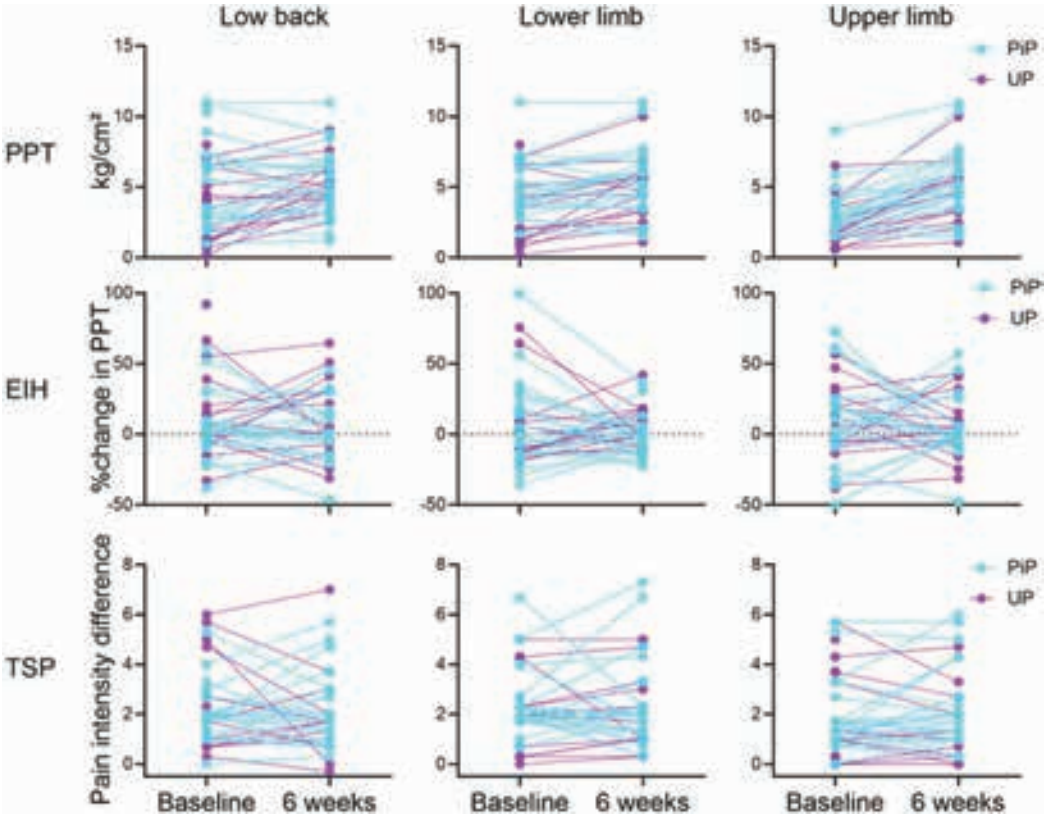


FIGURE 2 - Individual pain sensitivity (PPT) and modulation (TSP, EIH) variable changes (baseline to 6 weeks) at the low back, lower limb, and upper limb sites for both physiotherapy approaches. Note the increase in PPTs detected in both groups. PPT: pressure pain threshold; EIH: exercise-induced hypoalgesia; TSP: temporal pain summation; PiP: psychologically-informed physiotherapy; UP: usual physiotherapy.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 3 - Trajectories of physical functioning, pain, and symptoms of central sensitization scores at baseline and 6-week follow-up [mean (SD)]

	PiP			UP			Within sample difference
	Baseline	6 weeks	ΔPiP	Baseline	6 weeks	ΔUP	
Physical functioning	23.9 (11.8)	16.8 (10.4)	–7.1 (9.6)	27.1 (12.2)	21.9 (14.8)	–8.4 (9.4)	–6.3 (10)
Pain intensity	5.0 (1.9)	3.1 (2.0)	–2.1 (2.1)	5.7 (2.0)	3.7 (2.1)	–1.8 (2.2)	–2.0 (2.1)
Symptoms of central sensitization	17.6 (5.0)	15.2 (6.4)	–2.3 (5.7)	20.7 (6.2)	15.7 (7.2)	–4.3 (6.1)	–3.2 (5.9)

PiP: Psychologically-informed physiotherapy; UP: Usual physiotherapy; ΔPiP: Within-group difference (Post-pre) for PiP; ΔUP: Within-group difference (Post-pre) for PiP; SD: standard deviation.
Note that all scores were significantly different at 6 weeks compared to baseline (see Desgagnés et al., 2024).

TABLE 4 - Results of linear regressions (unstandardized β coefficient adjusted for age, sex, CSI baseline, and PCS baseline [95% confidence interval]) between pain sensitivity/modulation changes and clinical changes

	ΔPain	ΔODI	ΔCSI
ΔPPT			
Lumbar	β = –0.31 (–0.73, 0.11); p = 0.15	***β = –2.36 (–4.25, –0.46); p = 0.02***	β = –0.09 (–1.15, 1.32); p = 0.88
Lower limb	β = 0.06 (–0.48, 0.60); p = 0.83	β = –1.05 (–3.52, 1.43); p = 0.39	β = –0.39 (–1.85, 1.08); p = 0.59
Upper limb	β = –0.27 (–0.85, 0.31); p = 0.35	β = –2.46 (–5.08, 0.16); p = 0.07	β = –0.35 (–1.98, 1.29); p = 0.67
ΔEIH			
Lumbar	β = 0.01 (–0.020, 0.036); p = 0.58	β = –0.01 (–0.14, 0.12); p = 0.90	β = –0.02 (–0.09, 0.06); p = 0.69
Lower limb	β = 0.01 (–0.016, 0.037); p = 0.41	β = 0.05 (–0.08, 0.17); p = 0.44	β = 0.03 (–0.04, 0.11); p = 0.35
Upper limb	β = –0.01 (–0.030, 0.017); p = 0.56	β = 0.02 (–0.10, 0.13); p = 0.74	β = 0.002 (–0.07, 0.07); p = 0.95
ΔTSP			
Lumbar	β = 0.53 (–0.02, 1.07); p = 0.06	***β = 3.50 (1.01, 6.00); p = 0.01***	β = 1.15 (–0.45, 2.76); p = 0.15
Lower limb	β = 0.26 (–0.31, 0.83); p = 0.36	***β = 3.10 (0.38, 5.82); p = 0.03***	β = 0.05 (–1.70, 1.79); p = 0.96
Upper limb	β = –0.11 (–0.83, 0.62); p = 0.76	β = 0.01 (–3.62, 3.64); p = 1.00	β = –0.85 (–2.96, 1.25); p = 0.41

CSI: Central sensitization inventory; PCS: Pain catastrophizing scale; ODI: Oswestry disability index; PPT: Pressure pain threshold; EIH: Exercise-induced hypoalgesia; TSP: Temporal summation of pain; Δ: score at 6 weeks minus score at baseline.
Bold and italic: $p < 0.05$.

associations between differences in physical functioning and (i) PPT in the lumbar region, (ii) TSP in the lumbar region, and (iii) TSP at the lower limb were observed. However, no association with changes in pain intensity or symptoms of central sensitization was detected. These findings tend to dispute the hypothesis that changes in sensitivity/modulation variables are associated with clinical changes for most variables collected, with the exception of physical functioning. Nonetheless, this is important to consider that this is an exploratory study with a small sample size, and findings need to be interpreted with caution and replicated in future research.

The lack of PiP superiority over UP to influence pain sensitivity and modulation may appear surprising, considering that PiP—including pain neuroscience education—suggests a "desensitization" effect in chronic pain (17,20). Nonetheless, the literature specific to CLBP is scarce and only a few clinical trials reported the effects of PiP on pain sensitivity. Bodes Pardo et al., (2018) (2018) (21) reported an increase in PPT in favor of pain neuroscience education combined with therapeutic exercise compared to therapeutic exercise alone in participants with CLBP. Pain neuroscience education aims to modify beliefs through pain reconceptualization (44) and is commonly included in PiP. Furthermore, Cognitive Functional Therapy in patients with severe CLBP increased back PPT in a single-arm clinical trial (45). Cognitive Functional Therapy is an approach consistent with PiP [14]. However, due to the absence of a control group, it is not possible to determine if the increase was caused by the intervention (45). Another RCT, but in participants with knee osteoarthritis (46), reported no superiority of pain neuroscience education over biomedical education plus knee mobilization; both interventions increased PPTs and did not influence other QST measurements. Although the literature seems to support an increase in PPTs following PiP, other approaches also increase them. Indeed, both manual therapy and exercises, which were the core of the UP group, are known to reduce pain sensitivity. Meta-analyses observed that exercise [23] and spinal

manipulation (22) significantly increase PPTs. Thus, discrepancies between our results and those from Bodes Pardo et al., (2018) (21) could be explained by the presence of both manual and exercise therapies in UP. Different mechanisms may explain this intervention-induced hypoalgesia (i.e., an increase in PPTs). For example, exercises may (i) normalize the alteration in areas of the central nervous system contributing to pain processing via neuroplasticity, (ii) reduce nociception related to spine microtrauma via improved motor control, or (iii) indirectly modify psychological risk factors (e.g., reduction in fear of movement) (23). Similarly, spinal manipulation was suggested to activate pain modulation mechanisms at both spinal and supraspinal levels (22). Considering that both physiotherapy approaches comprised (i) manual and exercise therapies and (ii) may have influenced psychological factors directly (PiP) or indirectly (UP), it is possible that overlapping mechanisms may have contributed to the increase in PPTs. Overall, our results suggest that both physiotherapy approaches may reduce pressure pain sensitivity. Nonetheless, the exploratory nature of the study and the few numbers of studies published in the area preclude any clinical recommendations on the most appropriate intervention to influence pain sensitivity. If future powered RCTs confirm our findings, it may suggest that different physiotherapy approaches (exercise, manual therapy, pain education, etc.) may be used to reduce pain sensitivity if this is a specific therapeutic objective.

Despite the fact that PPTs were increased following both interventions, no effect was observed for TSP or EIH. Although PPT is the QST measurement that is most often used as a pain sensitivity proxy, some studies tested the effects of various interventions on other QST measurements. In contrast to our results, a systematic review suggested that physiotherapy interventions may influence various types of QST measurements (e.g., TSP, conditioned pain modulation) (47). For example, spinal manipulation reduced TSP more than stationary bicycle in patients with CLBP (48). These discrepancies could be explained by the different types of paradigms used. For example, most studies included in the systematic review used a suprathreshold heat response consisting of rating pain intensity following repeated heat stimuli (47). In contrast, we used a pinprick to measure TSP. Other reasons for discrepancies could be related to different clinical populations recruited or the time elapsed between measurements. EIH was not included as an outcome in the meta-analysis. In a recent study, we demonstrated that EIH was absent in participants with CLBP (5). In contrast, our sample of participants with CLBP seemed to present with "normal" EIH (i.e., hypoalgesia was present at baseline for most sites). This difference might be explained by the low reliability of EIH, which may, therefore, be less suitable for use as an outcome (49).

For the second objective, aiming to determine the associations between changes in PROMs and in pain sensitivity and modulation variables, we observed that changes in PPT and TSP were associated with changes in physical functioning but not in pain intensity. The absence of significant associations with pain remains intriguing. We could speculate that the reduction in pain sensitivity following both physiotherapy approaches allowed participants to re-engage in everyday life activity and then improve their physical functioning. For example, a study of participants with knee osteoarthritis observed significant correlations between TSP at the knee and sensitivity to physical activity (50). Although indirect evidence, it is possible that pain sensitivity variables such as TSP may be related to sensitivity during activity and influence physical functioning. To our knowledge, only a few studies tested the associations between changes in pain sensitivity and PROMs and did not identify significant associations. For instance, Nim et al. (2021) (50) reported no significant association between changes in QST variables and clinical improvement outcomes following spine manipulation in participants with LBP. Similarly, Palsson et al., 2021 (51) did not report any correlations between changes in PROMs and in QST measurements following a rehabilitation program, including exercises, even though pain and physical functioning both improved. Thus, high discrepancies are present, making it difficult to interpret our results. Our hypothesis explaining the influence of pain sensitivity on physical functioning underlies causal influence which cannot be ascertained with the current analyses. Also, the small proportion of significant associations between changes in clinical outcomes and QST can be explained by the large number of statistical analyses done (type I error). Further studies with larger sample sizes and additional analyses are necessary to confirm that change in pain sensitivity could be a mediator of the improvement in physical functioning.

Another interesting finding is the absence of a correlation between changes in pain sensitivity and modulation variables and changes in symptoms of "central sensitization." This highlights the contrast in the constructs between QST and CSI; sensitization to nociceptive/somatosensory stimuli as measured with QST does not grasp the whole construct measured by the CSI. While certain aspects may appear analogous, CSI encompasses a wider range of symptoms typically observed in patients with features of "central sensitization" associated with a general hypersensitivity to various sensory stimuli (e.g., sound, odor), general health (e.g., gastrointestinal, dermatology) and lifestyle behaviors (e.g., sleep). This is in line with a study demonstrating low or no association between CSI scores and PPT and conditioned pain modulation, respectively (42).

Limitations

The current study has limitations that must be considered when interpreting our results and their applicability to clinical settings. First, the small sample size limits the statistical power of our analyses, and our results should be interpreted cautiously. Second, the absence of a "placebo" or minimal intervention group precludes a definitive determination of whether the increased PPTs were attributable to the interventions received or the natural trajectory of pain sensitization in CLBP patients. For example, reliability analyses showed that PPT could increase at the second measurement within a single session (34), even though estimated PPT changes were superior to the minimal detectable changes (33).

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Conclusions

Usual and psychologically informed physiotherapy approaches elicited widespread hypoalgesia after 6 weeks of treatment as measured by PPTs, but not for the TSP or EIH. In addition, changes in pain sensitivity (e.g., PPT and TSP) were associated with changes in physical functioning but not in pain intensity. Contrary to initial expectations, our study suggests that neither approach demonstrated a significant advantage over the other to impact pain sensitivity. Future studies with a larger sample size and including a "minimal intervention" group may help to determine if the hypoalgesia induced by physiotherapy approaches is caused by physiotherapy interventions or if it is associated with the natural evolution of pain sensitivity over time.

Disclosures

Conflict of interest: None of the authors have potential conflicts of interest to be disclosed.

Financial support: The project was funded by a grant from REPAR (rehabilitation network of the Fonds de recherche du Québec). HMA is supported by a research scholar from Fonds de recherche du Québec – Santé (281961). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Meucci RD, Fassa AG, Faria NM. Prevalence of chronic low back pain: systematic review. *Rev Saude Publica*. 2015;49(0):1. CrossRef PubMed
2. James SL, Abate D, Abate KH, et al. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392(10159):1789-1858. CrossRef PubMed
3. Maher C, Underwood M, Buchbinder R. Non-specific low back pain. *Lancet*. 2017;389(10070):736-747. CrossRef PubMed
4. Hartvigsen J, Hancock MJ, Kongsted A, et al. Lancet Low Back Pain Series Working Group. What low back pain is and why we need to pay attention. *Lancet*. 2018;391(10137):2356-2367. CrossRef PubMed
5. Patricio P, Mailloux C, Wideman TH, et al. Assessment of exercise-induced hypoalgesia in chronic low back pain and potential associations with psychological factors and central sensitization symptoms: A case-control study. *Pain Pract*. 2023;23(3):264-276. CrossRef PubMed
6. den Bandt HL, Ickmans K, Leemans L, et al. Differences in quantitative sensory testing outcomes between patients with low back pain in primary care and pain-free controls. *Clin J Pain*. 2022;38(6):381-387. CrossRef PubMed
7. Nijs J, Kosek E, Chiarotto A, et al. Nociceptive, neuropathic, or nociplastic low back pain? The low back pain phenotyping (BACPAP) consortium's international and multidisciplinary consensus recommendations. *Lancet Rheumatol*. 2024;6(3):e178-e188. CrossRef PubMed
8. Shraim MA, Massé-Alarie H, Hodges PW. Methods to discriminate between mechanism-based categories of pain experienced in the musculoskeletal system: a systematic review. *Pain*. 2021;162(4):1007-1037. CrossRef PubMed
9. Shraim MA, Sluka KA, Sterling M, et al. Features and methods to discriminate between mechanism-based categories of pain experienced in the musculoskeletal system: a Delphi expert consensus study. *Pain*. 2022;163(9):1812-1828. CrossRef PubMed
10. Smart KM, Blake C, Staines A, et al. Clinical indicators of 'nociceptive,' 'peripheral neuropathic' and 'central' mechanisms of musculoskeletal pain. A Delphi survey of expert clinicians. *Man Ther*. 2010;15(1):80-87. CrossRef PubMed
11. da C Menezes Costa L, Maher CG, Hancock MJ, McAuley JH, Herbert RD, Costa LO. The prognosis of acute and persistent low-back pain: a meta-analysis. *CMAJ*. 2012;184(11):E613-E624. CrossRef PubMed
12. Ballengee LA, Zullig LL, George SZ. Implementation of psychologically informed physical therapy for low back pain: where do we stand, where do we go? *J Pain Res*. 2021;14:3747-3757. CrossRef PubMed
13. Hodges PW. Hybrid approach to treatment tailoring for low back pain: a proposed model of care. *J Orthop Sports Phys Ther*. 2019;49(6):453-463. CrossRef PubMed
14. O'Sullivan PB, Caneiro JP, O'Keeffe M, et al. Cognitive functional therapy: an integrated behavioral approach for the targeted management of disabling low back pain. *Phys Ther*. 2018;98(5):408-423. CrossRef PubMed
15. Main CJ, George SZ. Psychologically informed practice for management of low back pain: future directions in practice and research. *Phys Ther*. 2011;91(5):820-824. CrossRef PubMed
16. Ho EK, Chen L, Simic M, et al. Psychological interventions for chronic, non-specific low back pain: systematic review with network meta-analysis. *BMJ*. 2022;376:e067718. CrossRef PubMed
17. Van Oosterwijck J, Meeus M, Paul L, et al. Pain physiology education improves health status and endogenous pain inhibition in fibromyalgia: a double-blind randomized controlled trial. *Clin J Pain*. 2013;29(10):873-882. CrossRef PubMed
18. Tracey I, Mantyh PW. The cerebral signature for pain perception and its modulation. *Neuron*. 2007;55(3):377-391. CrossRef PubMed
19. Meints SM, Mawla I, Napadow V, et al. The relationship between catastrophizing and altered pain sensitivity in patients with chronic low-back pain. *Pain*. 2019;160(4):833-843. CrossRef PubMed
20. Van Oosterwijck J, Nijs J, Meeus M, et al. Pain neurophysiology education improves cognitions, pain thresholds, and movement performance in people with chronic whiplash: a pilot study. *J Rehabil Res Dev*. 2011;48(1):43-58. CrossRef PubMed
21. Bodes Pardo G, Lluch Girbés E, Roussel NA, et al. Pain neurophysiology education and therapeutic exercise for patients with chronic low back pain: a single-blind randomized controlled trial. *Arch Phys Med Rehabil*. 2018;99(2):338-347. CrossRef PubMed
22. Coronado RA, Gay CW, Bialosky JE, et al. Changes in pain sensitivity following spinal manipulation: a systematic review and meta-analysis. *J Electromyogr Kinesiol*. 2012;22(5):752-767. CrossRef PubMed
23. Belavy DL, Van Oosterwijck J, Clarkson M, et al. Pain sensitivity is reduced by exercise training: evidence from a systematic review and meta-analysis. *Neurosci Biobehav Rev*. 2021;120:100-108. CrossRef PubMed
24. Desgagnés A, Côté-Picard C, Gaumond A, et al. Efficacy of a psychologically-informed physiotherapy intervention in patients with chronic low back pain at high risk of poor prognosis: a pilot and feasibility randomized controlled trial. *Physiother Can*. 2024;76(2):163-174. CrossRef PubMed
25. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow

process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377-381. CrossRef PubMed
26. Hill JC, Dunn KM, Lewis M, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum*. 2008;59(5):632-641. CrossRef PubMed
27. Bouhassira D, Attal N, Alchaar H, et al. Comparison of pain syndromes associated with nervous or somatic lesions and development of a new neuropathic pain diagnostic questionnaire (DN4). *Pain*. 2005;114(1-2):29-36. CrossRef PubMed
28. Abbott JH. The distinction between randomized clinical trials (RCTs) and preliminary feasibility and pilot studies: what they are and are not. *J Orthop Sports Phys Ther*. 2014;44(8): 555-558. CrossRef PubMed
29. Rolke R, Baron R, Maier C, et al. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): standardized protocol and reference values. *Pain*. 2006;123(3):231-243. CrossRef PubMed
30. Coronado RA, Bialosky JE, Robinson ME, et al. Pain sensitivity subgroups in individuals with spine pain: potential relevance to short-term clinical outcome. *Phys Ther*. 2014;94(8): 1111-1122. CrossRef PubMed
31. McPhee ME, Vaegter HB, Graven-Nielsen T. Alterations in pronociceptive and antinociceptive mechanisms in patients with low back pain: a systematic review with meta-analysis. *Pain*. 2020;161(3):464-475. CrossRef PubMed
32. Mailloux C, Wideman TH, Massé-Alarie H. Wrist, but not back, isometric contraction induced widespread hypoalgesia in healthy participants. *Front Pain Res (Lausanne)*. 2021;2:701830. CrossRef PubMed
33. Mailloux C, Beaulieu LD, Wideman TH, et al. Within-session test-retest reliability of pressure pain threshold and mechanical temporal summation in healthy subjects. *PLoS One*. 2021;16(1) :e0245278. CrossRef PubMed
34. de Oliveira FCL, Cossette C, Mailloux C, et al. Within-session test-retest reliability of pressure pain threshold and mechanical temporal summation in chronic low back pain. *Clin J Pain*. 2023;39(5):217-225. CrossRef PubMed
35. Naugle KM, Fillingim RB, Riley JL III. A meta-analytic review of the hypoalgesic effects of exercise. *J Pain*. 2012;13(12): 1139-1150. CrossRef PubMed
36. Rice D, Nijs J, Kosek E, et al. Exercise-induced hypoalgesia in pain-free and chronic pain populations: state of the art and future directions. *J Pain*. 2019;20(11):1249-1266. CrossRef PubMed
37. Koltyn KF, Brellenthin AG, Cook DB, et al. Mechanisms of exercise-induced hypoalgesia. *J Pain*. 2014;15(12):1294-1304. CrossRef PubMed
38. Hermens HJ, Freriks B, Disselhorst-Klug C, et al. Development of recommendations for SEMG sensors and sensor placement procedures. *J Electromyogr Kinesiol*. Oct 2000;10(5):361-74. CrossRef
39. Chiarotto A, Boers M, Deyo RA, et al. Core outcome measurement instruments for clinical trials in non-specific low back pain. *Pain*. 2018;159(3):481-495. CrossRef PubMed
40. Mayer TG, Neblett R, Cohen H, et al. The development and psychometric validation of the central sensitization inventory. *Pain Pract*. 2012;12(4):276-285. CrossRef PubMed
41. Kregel J, Schumacher C, Dolphens M, et al. Convergent validity of the dutch central sensitization inventory: associations with psychophysical pain measures, quality of life, disability, and pain cognitions in patients with chronic spinal pain. *Pain Pract*. 2018;18(6):777-787. CrossRef PubMed
42. McNamee R. Confounding and confounders. *Occup Environ Med*. 2003;60(3):227-234. CrossRef PubMed
43. Nijs J, Meeus M, Cagnie B, et al. A modern neuroscience approach to chronic spinal pain: combining pain neuroscience education with cognition-targeted motor control training. *Phys Ther*. 2014;94(5):730-738. CrossRef PubMed
44. Vaegter HB, Ussing K, Johansen JV, et al. Improvements in clinical pain and experimental pain sensitivity after cognitive functional therapy in patients with severe persistent low back pain. *Pain Rep*. 2019;5(1):e802. CrossRef PubMed
45. Lluch E, Dueñas L, Falla D, et al. Preoperative pain neuroscience education combined with knee joint mobilization for knee osteoarthritis: a randomized controlled trial. *Clin J Pain*. 2018;34(1):44-52. CrossRef PubMed
46. Arribas-Romano A, Fernández-Carnero J, Molina-Rueda F, et al. efficacy of physical therapy on nociceptive pain processing alterations in patients with chronic musculoskeletal pain: a systematic review and meta-analysis. *Pain Med*. 2020;21(10):2502-2517. CrossRef PubMed
47. Bialosky JE, Bishop MD, Robinson ME, et al. Spinal manipulative therapy has an immediate effect on thermal pain sensitivity in people with low back pain: a randomized controlled trial. *Phys Ther*. 2009;89(12):1292-1303. CrossRef PubMed
48. Vaegter HB, Lyng KD, Yttereng FW, et al. Exercise-induced hypoalgesia after isometric wall squat exercise: a test-retest reliability study. *Pain Med*. 2019;20(1):129-137. CrossRef PubMed
49. Wideman TH, Finan PH, Edwards RR, et al. Increased sensitivity to physical activity among individuals with knee osteoarthritis: relation to pain outcomes, psychological factors, and responses to quantitative sensory testing. *Pain*. 2014;155(4):703-711. CrossRef PubMed
50. Nim CG, Kawchuk GN, Schiøttz-Christensen B, et al. Changes in pain sensitivity and spinal stiffness in relation to responder status following spinal manipulative therapy in chronic low Back pain: a secondary explorative analysis of a randomized trial. *BMC Musculoskelet Disord*. 2021;22(1):23. CrossRef PubMed
51. Palsson TS, Christensen SWM, De Martino E, et al. Pain and disability in low back pain can be reduced despite no significant improvements in mechanistic pain biomarkers. *Clin J Pain*. 2021;37(5):330-338. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 42-48
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3339
ORIGINAL RESEARCH ARTICLE

Scaling new heights: a pilot study of the impact of climbing on balance, agility, and dexterity in individuals with Parkinson's disease

Julie D Ries[1,2], **Molly Cupka**[3], **Andrew A Guccione**[4]

[1]College of Health and Education, School of Health Professions, Program in Physical Therapy, Marymount University, Arlington County, Virginia - United States of America
[2]Marymount University Center for Optimal Aging, Marymount University, Arlington County, Virginia - United States of America
[3]SportRock Climbing Centers, Alexandria, Virginia - United States of America
[4]George Mason University, Fairfax, Virginia - United States of America

ABSTRACT

Introduction: Exercise is known to be beneficial for individuals with Parkinson's disease (PD). Rock climbing contains exercise characteristics highlighted in published clinical guidelines for PD (e.g., aerobic, resistance, balance training, cued-movements, community-based) and also has unique somatosensory and visuospatial experiences that may facilitate motor learning. The purpose of this study was to investigate the effects of a climbing program on the physical function of individuals with PD.
Methods: This quasi-experimental observational pilot study used pre-to-post-test comparisons to assess participants with mild to moderate PD (Hoehn and Yahr 1-3) who walked independently. The intervention was 12 weeks of community-based, twice weekly top-rope climbing sessions under one-on-one supervision and tailored to skill level. Wall angles, hand/foot holds, and routes varied from climb to climb and became progressively more difficult as skills increased. The primary outcome measure was the Community Balance & Mobility Scale (CBMS); secondary measures were the Agility T-Test (ATT), 9-Hole Peg Test (9HPT), upper extremity reaction time using BlazePods (UE-React), and grip strength (GS).
Results: 28 participants completed the study: 8 women/20 men; mean age = 66.1 (sd = 7.4) years; average disease duration = 4.0 (sd = 3.6) years. *Paired t-tests* comparing pre- and post-test scores and effect sizes (ES) with 95% confidence intervals (CI) were calculated for statistically significant results using *Hedge's g*. Findings were: CBMS ($p < 0.001$; ES = 0.573, 95% CI = 0.178-0.960), ATT ($p < 0.001$; ES = 0.462, 95% CI = 0.078-0.838), 9HPT ($p < 0.001$; ES = 0.480, 95% CI = 0.094-0.858), UE-React ($p < 0.001$; ES = 0.329, 95% CI = −0.045-0.696); GS changes were non-significant.
Conclusions: Rock climbing demonstrated medium-size effects on mobility/balance and small-size effects on agility and dexterity that could impact functioning in everyday activities.

Keywords: motor learning, physical activity, Parkinson's disease, rehabilitation

What is already known about this topic?

- *Physical activity is known to be beneficial for individuals with PD, and certain components have been highlighted in clinical practice guidelines as key characteristics (e.g., aerobic, strength, and balance training, cued responses, and community-based). Rock climbing encompasses all of these traits.*

What does this study add?

- *Rock climbing contains multiple PD-recommended exercise components in one activity and offers unique motor learning experiences for individuals with PD. Vertical climbing training could lead to improved horizontal over-ground performance.*

Received: October 14, 2024
Accepted: February 18, 2025
Published online: February 26, 2025

Corresponding author:
Julie D Ries
email: jries@marymount.edu

Introduction

Parkinson's Disease (PD) is a major public health crisis in the making. It is the second most common neurological condition after Alzheimer's disease, affecting nearly 1 million people in the US (1) and causing a significant public health burden (2). While there are a few treatment options that may mitigate symptoms, there is currently no cure for this health condition.

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

Exercise is commonly prescribed for individuals with PD as both an intervention for and prevention of further disability (3,4). Ernst et al. support the beneficial effects of physical exercise, including gait/balance/functional exercise and multi-domain exercise, on motor signs in PD (4) Osborne et al. offer clinical practice guidelines that include multiple types of activity interventions (e.g., aerobic, resistance, balance, externally cued, community-based) (5). Performance-based interventions emphasizing aerobic intensity and resistance training in community settings have gained traction among patients and health professionals (5), including dancing (6), and boxing (7). Despite their increasing popularity, the evidence base for these programs is still developing and far from conclusive (8). More recently, rock climbing programs have been identified as providing potential therapeutic benefits for people with PD (9-11) and other health conditions (12,13). Rock climbing uniquely encompasses many of the recommended characteristics of activity interventions (4,5). Langer and colleagues (9-11) suggest that climbing might positively impact motor symptoms, posture, and gait speed in individuals with PD.

The purpose of this pilot study was to identify the primary physical effects associated with participating in rock climbing on functional performance for tasks of community mobility and upper extremity (UE) functioning. The specific aims of this study were to characterize the effects of rock climbing on a primary outcome of mobility and locomotion, as well as secondary outcomes of agility, dexterity, UE reaction time, and grip strength (GS). Based on anecdotal evidence and the limited research available, we hypothesized improvement in all outcome measures.

Methods

Study Design

This pilot study used a quasi-experimental observational design to follow a cohort of individuals with PD who had limited to no exposure to rock climbing. Participants completed physical performance tests twice, prior to the first climbing session and then again after completing 24 sessions for before and after comparisons.

Target Population

Consenting adults with mild-to-moderate PD (Hoehn and Yahr 1-3) who met the inclusion criteria were eligible to participate.

Recruitment

Participants were recruited from the greater Washington DC metro area by word of mouth, healthcare provider referral, support groups, social media, and posted fliers. Interested individuals were directed to a recruitment website that listed the specific inclusion and exclusion criteria to participate, allowing prospective participants to self-determine their eligibility before contacting the research team for screening. Prospective participants then contacted one member of the research team (MC) who confirmed eligibility, reviewing each of the inclusion and exclusion criteria. Data collection occurred in an exercise room at the climbing gym and spanned August 2023 to August 2024.

Inclusion criteria were: ≥18 years old; diagnosis of PD by a health professional; Hoehn and Yahr score 1-3; able to speak and read in English; and able to ambulate at least 10 meters with no assistive device or human assistance. Exclusion criteria were: Diagnosis of neurological disease other than PD; uncontrolled cardiovascular, pulmonary, neurological, or metabolic disease which might impact the ability to exercise or for which exercise is contraindicated; cognitive or psychiatric impairment precluding informed consent or ability to follow instructions; and pregnancy. Participants engaging in routine physical activities were included in the study, but participants did not begin any new structured exercise over the course of the study.

Ethics Approval

The pilot study proposal was approved by the Marymount University (Arlington VA) Institutional Review Board (IRB Protocol #831) and registered on Clinical Trials (NCT05919771). Consent was obtained by a single member of the research team (MC).

Intervention Description and Protocol

For the purposes of this pilot study, rock climbing was described as a vertical ascent of a surface specifically designed to elicit the skills that might be required to scale a naturally occurring rock formation. The rock climbing experiences offered by UpENDing Parkinson's (UEP), a non-profit organization that has provided free training in rock climbing for individuals with PD since 2012 (https://www.upending-parkinsons.org), served as the study intervention. Sessions were integrated into an existing group of climbers with PD and conducted under one-on-one supervision using a harness and rope to belay the climber (consistent with "top-rope" or "sport" climbing). Sessions were supervised by UEP volunteers with extensive climbing experience and were tailored to the skill level of the participant. The angles of the walls, the holds on which climbers grab or step, and the routes varied from climb to climb, and climbers progressed to more difficult courses as their skill levels increased. Participants had the opportunity to climb two times per week free of charge. Study participants could also climb outside of group climbing sessions if they received safety clearance from the UEP staff. Participants were encouraged to climb at least two times a week but no more than three times in a seven-day period, and attendance was monitored by UEP staff upon entering the climbing facility. Post-test data were collected after 24 climbing sessions, which generally took participants ~12 weeks to complete.

Outcome Measure Selection

Our study team relied on a research advisory group that included a climber with PD and the spouse of a climber to identify the array of concepts relevant to physical mobility and well-being. With this group's input, we selected candidate instruments for further review. Following the initial culling of potential tests and measures, the research team

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

gathered a small group of experienced climbers with PD for a trial administration of measures to identify floor and ceiling effects of the proposed instruments. Additionally, to finalize the battery of tests and measures used in this study, the same experienced climbers participated in a focus group to provide feedback on what outcomes were important and meaningful to them as individuals living with PD.

Primary Outcome Measure

The primary outcome measure for this observational study was a modified version of the Community Balance and Mobility Scale (14) (CBMS) consisting of eight separate upright mobility tasks, four of which are performed with both left and right lower extremity (LE). The eight tasks were selected from the original 13 based on feedback regarding redundancy across tests and subject fatigue provided by our test group of experienced climbers with PD. The CBMS was scored from 0-60 with higher scores indicating better performance.

Secondary Outcome Measures

Four other measures were used, including the Agility T-Test (ATT) (15), in which a participant is timed as they move quickly toward cone targets that they must tap with a hand, moving forward, laterally to the left, laterally to the right, back to the left, and then backward to their starting point. Faster scores denote better agility.

The Nine-Hole Peg Test (16) was used to assess UE dexterity. This test measures the time it takes to remove nine pegs (about the size of a golf tee) from a tray, place each peg in a hole, and then remove the pegs and place them back in the tray. Participants were seated at a table and chose which hand to test (stabilizing the peg board with the other hand) and repeated with the same hand on post-testing. Faster times indicated better dexterity.

Reaction time (quantifying the combined cognitive and physical response) was captured using the BlazePod®[a] system (17). A standard semicircle of four pods (pod cases were affixed to plywood base) emitted red light stimuli in a random sequence to trigger the participant's firm tap of the light source. Participants were seated with the pods on the table in front of them and could use either or both hands to respond to lit pods. Total reaction time was the system's recorded time response to 30 random flashing stimuli, with lower scores denoting better reaction time.

Grip strength (GS) (16) is a measure of force production that was assessed using a standard hand-held dynamometer (Jamar[b]), yielding the force output in pounds, with higher force representing greater strength. Participants sat in a chair with their arms at their side and elbow in ninety degrees of flexion. Participants chose which hand to test and repeated with the same hand on post-testing.

Data Collection Procedures

Data collection procedures were taught to volunteer climbers and physical therapy student research assistants by the lead researchers (two physical therapists and one climbing instructor) and practiced several times before the first participant was tested. Participant testing sessions occurred within two weeks of beginning/finishing the 12-week intervention. One author (AG) had minimal involvement in data collection and did not engage in the pre- or post-test of the same individual, in any case, to remain blinded to the results, and two authors (JR and AG) did not attend any of the climbing sessions. Data analysis was performed using a deidentified data set.

Data Analysis, Power, and Sample Size

Descriptive statistics were evaluated by calculating the mean, standard deviation, and range for ratio data and evaluating percentages for nominal data. Outcome data were assessed for normal distribution using the *Shapiro-Wilk Test*. Then, each outcome measure was analyzed using *paired t-tests* to compare pre- and post-test values. Because there are no studies of rock climbing using community mobility as the primary outcome, the research team proceeded very conservatively with power and sample size estimations based on a desired effect size (ES) of 0.5. A Cohen's *d* of 0.2 is generally considered to be "small," 0.5 to be "medium," and 0.8 to be "large." (18) In the context of a rehabilitation intervention study, a moderate ES is generally considered to be clinically meaningful. We computed our sample size based on our primary variable, CBMS, using a one-tailed paired t-test with a significance level of 0.05 and 80% power. Based on our calculations, our *a priori* required sample size was 27 individuals with PD.

Because ES may be inflated in small samples and the preliminary nature of a pilot requires statistical restraint, we chose to estimate ES for within-subject differences using IBM SPSS Statistics (Version 29.0.2.0)[c] in the most conservative (i.e., unbiased) way by calculating *Hedge's g* [with 95% confidence intervals (CIs)], which uses the square root of the average variance of measures in the denominator, plus a correction factor.

Results

A total of 32 participants consented to the study, of whom 28 completed 12 weeks of training and both pre- and post-testing sessions. The drop-out rate was 12.5%, far less than typically expected for an exercise study (19, 20). Of the four subjects who did not complete the study, two reported a lack of interest in rock climbing, one became too busy to climb regularly, and one experienced a shoulder injury unrelated to the study. Comparisons of participants who completed the study and those who did not, showed no appreciable differences with respect to age (*independent t-test*: $t = 0.161$, $p = 0.881$), time since PD diagnosis (*independent t-test*: $t = -1.468$, $p = 0.230$), or Hoehn and Yahr score (*chi-square test:* $\chi^2 = 2.245$; $p = 0.326$), but non-completers were different in sex distribution (*chi-square test:* $\chi^2 = 8.837$; $p = 0.014$), in that all four were female. There were no adverse events during the course of the study.

Table 1 presents the demographics of the participants who completed the study. In general, participants were mostly male older adults with PD for less than 5 years and

less severely impacted by the clinical signs and symptoms of PD, as evidenced by Hoehn and Yahr scores. All participants self-reported their health as either good or excellent, and most exercised routinely prior to entering the study.

TABLE 1 - Participant demographics

Participant Characteristic	Study Participants (n = 28)
Mean Age (years [sd], range)	66.1 [7.4] Range 45-78
Sex	28.6% Female (8) 71.4% Male (20)
Disease Duration (years [sd], range)	4.0 [3.6] Range 1-14
Hoehn & Yahr (H&Y) Score	64.3% H&Y 1 (18) 25.0% H&Y 2 (7) 10.7& H&Y (3)
Race/Ethnicity	85.7% White (26) 7.1% Hispanic or Latino (2) 7.1% Declined to Answer (2)
Education	64.3% Graduate Degree (18) 3.6% Some Grad School (1) 25.0% Bachelor's Degree (7) 3.6% Tech School Degree (1) 3.6% Some College (1)
Self-Health Rating	21.4 % Excellent (6) 78.6% Good (22)
Participate in Regular Exercise	92.9% Yes (26) 7.1% No (2)
History of Falls (Y/N)	25% Yes (7) 75% No (21)

Paired t-tests for the outcome measures demonstrated statistically significant improvement in performance on four of the five outcome measures (Table 2). There was no statistically significant change in GS, one of the secondary outcome measures. We calculated an ES for each of the four statistically significant comparisons to gauge the magnitude of the change. As displayed in Table 2 as point estimates and 95% CIs, we found medium ES for our primary outcome measure, CBMS. We found small ES for three of our four secondary outcome measures [ATT, 9-Hole Peg Test (9HPT), and Reaction Time as measured by BlazePods®]; however, the Reaction Time CI includes zero, impacting our confidence in the estimate of ES on this outcome.

Discussion

This observational pilot study of a 12-week climbing intervention for individuals with mild to moderate PD demonstrated participant improvement, with a medium effect on mobility and balance (CBMS), the primary outcome, and a small effect on agility (ATT) and dexterity (9HPT). There was negligible effect on UE reaction time and GS.

The growth of climbing as a therapeutic adjunct has followed the expansion in popularity of climbing in the fitness community and has a history of use in orthopedic conditions, depression, and multiple sclerosis (13). To our knowledge, there is one published study on top-rope climbing for people with PD, a randomized controlled trial that is the subject of three separate publications (9-11) Langer and colleagues demonstrated that a 12-week, once per week 90-minute top-rope climbing intervention for participants with mild to moderate PD resulted in improvements in motor symptoms (9), posture (10), and gait speed (11). Our positive results are consistent with the findings of that study.

Exercise is known to be therapeutic in the management of PD motor signs and symptoms (21,22), although a recent Cochrane review finds fault in the rigor of evidence for exercise in existing research (4). Physical therapy clinical practice guidelines (5,23) strongly support exercise that has aerobic and resistance training components, balance demands, external cueing, and takes place in a community-based format – all characteristics that are found in climbing. These traits are not unique to climbing, but climbing provides all of these in a single session.

Climbing also encompasses many routinely supported principles of exercise and activity interventions. The principle of progressive overload of biomechanical and bioenergetic demands is a well-supported framework in musculoskeletal and cardiopulmonary systems training, and an appropriate level of challenge is pivotal to driving neuroplastic changes. The difficulty level of routes and the amount of support provided by the belayer are dynamic variables contributing to overload and challenges that are modified over time. Climbing offers variability within and between climbs, and

TABLE 2 - Mean differences, standard deviations, significance, and ESs with 95% CIs for outcome measures

Outcome Measure	Mean Difference	Standard Deviation	*t statistic*	*p-value*	ES (*Cohen's d*) (95% CI)
CBMS	6.250	6.942	4.764	<0.001	0.573 (0.178-0.960)
ATT	2.419	3.417	3.746	<0.001	0.462 (0.078-0.838)
9HPT	2.148	3.125	3.637	<0.001	0.480 (0.094-0.858)
Reaction Time	0.740	0.998	3.922	<0.001	0.329 (–0.045-0.696)
GS	1.661	14.337	0.613	0.273	N/A*

CI = Confidence Interval; CBMS = Community Balance and Mobility Scale.
Mean Difference = Difference between pre- and post-test scores on each outcome measure.
*Effect size calculation not applicable (N/A) due to lack of statistically significant findings.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

practice variability has long been considered critical to creating broad and adaptable movement strategies.

Additionally, there are novel components to climbing that are not readily found in other therapeutic or activity interventions that foster motor adaptation. Table 3 identifies some of the unique characteristics of therapeutic rock climbing as related to our understanding of current motor control and motor learning theory. Of particular note are the advantages of rock climbing with respect to our emerging understanding of the optimization of motor learning in the context of social, cognitive, and affective domains. The Optimizing Performance Through Intrinsic Motivation and Attention for Learning (OPTIMAL) theory (24) emphasizes autonomy and enhanced expectations of the learner with an externally focused goal. Rock climbers are encouraged and empowered to solve the sequence of motor problems at hand: they "take on" the wall with the expectation of success. In summary, rock climbing as a physical activity mimics many of the characteristics of structured exercise programs, but it also offers the critical attributes of enhanced motor learning offered by the OPTIMAL theory.

Our findings suggest that participants adapted in response to the challenges of climbing. Participants did not just get better at climbing (task-specific improvements as evidenced in their progression of climbing skill level over the course of the study) but demonstrated improvement in functional movements relevant to upright mobility and activities of daily living. Motor learning is the result of experience-dependent neuroplasticity (25), and the capacity to benefit from exercise-induced neuroplasticity in individuals with PD has been documented (3,26). Individuals with PD are capable of motor learning, although the speed, scope, and generalizability may be less than that of healthy individuals (27). Our findings strongly suggest that rock climbing can support enhanced motor learning and promote movement capabilities beyond task-specific gains.

Climbing, a form of "vertical quadruped locomotion" (28) training, appears to have translated to horizontal bipedal mobility gains in this pilot data. Although we cannot speculate beyond the limits of this study, future research should explore the neuroanatomical underpinnings of motor learning in PD. There is a tendency in the rehabilitation literature

TABLE 3 - Unique motor control and motor learning characteristics of therapeutic rock climbing

Unique to Rock Climbing	Specific Characteristics or Demands	Motor Control and Motor Learning within the Context of Founding Theories
Vertical Orientation	• The vertical orientation of the movement task drives substantially different demands on posture and movement (vs. traditional movement in the horizontal plane as in over-ground activities) • Orientation to the wall counters flexion posture tendency and anterior (sagittal plane) instability	• Changes movement landscape: Removes climber from habitual, entrenched movement strategies (i.e., deep "attractor wells" in *Dynamic Systems Motor Control Theory)*, providing an opportunity for discovery of untapped movement potential (30)
	• Unique somatosensory experiences (e.g., the stable surface is anterior to the climber; position of body and extremities is constantly changing; different holds and wall contours force climber to adapt; feel of harness, shoes, ropes) • Unique visuospatial experiences (e.g., functioning at different elevations, locating and targeting movement to appropriate holds) used in perception-action coupling	• Ongoing assessment of relevant environmental and task cues within the perceptual-motor workspace in the context of goal-directed behavior (*Ecological Motor Learning Theory*)[31]
Full-body engagement	• Bilateral inter- and intra-limb coordination demands in multiplanar directions and varying amplitudes with repeated midline crossing • Open and closed chain functioning of all extremities requires varying mobility and stability skills	• Application of practice variability and movement complexity forces expansion of the behavioral repertoire (32, 33)
Cognitive demand	• Concurrent cognitive demands of determining the route and optimal moves: Constantly making decisions about how to orient/arrange one's body and extremities relative to the wall/holds for best success	• Dual-task (motor and cognitive) training is effective in improving single and dual-task performance in PD (34, 35)
Salutary stressors	• Poses therapeutic level of stress/anxiety (associated with novelty, heights, equipment, etc.) • Consistently increasing challenges by progressive overloading of biomechanical and bioenergetic demands	• Stressors "prime" the neuromotor system for adaptability (36, 37)
Task-oriented goal in a unique setting	• Integrates expectation of success, and autonomous control over movement and goal attainment	• Facilitates motivation and confidence of learner (*OPTIMAL Motor Learning Theory*) (24)

to delineate types of motor learning to specific brain regions: The pre-frontal cortex for strategy-based learning; basal ganglia for reward-based learning; the motor cortex and spinal cord for repetition-based learning; and cerebellum for error-based learning (25). This makes it tempting to conclude that those with PD (a primarily basal ganglia pathology) are unable to benefit from reward-based learning (basal ganglia mediated). It may be more accurate to acknowledge that motor learning is not the product of mutually exclusive functioning of different brain areas but rather the result of shared responsibility of multiple dynamic and adaptive systems and substrates (29).

Future research may also determine the best methodology for documenting outcomes related to climbing, as some of our secondary outcome measures (i.e., reaction time using BlazePods® and GS using hand-held dynamometer) did not show the results we anticipated.

Limitations

The generalizability of this pilot study is limited by the lack of a control group and a small sample despite the adequate power to detect change. Assessors were not blinded to the purpose of the study. Our findings are limited to novel climbers immediately following a 12-week climbing intervention. Furthermore, as in almost all exercise and physical activity studies, participants are a self-selected group whose interests match the study intervention. Our participants were generally healthy, routine exercisers with less severe PD for less than 5 years on average. Moreover, they were very well educated and white, social determinants of health that are known to positively influence health outcomes. Interpretation of the data is hampered by a lack of comparative studies that would allow a more complete assessment of the context and import of our findings.

Conclusion

Rock climbing appears to have modest positive effects on physical performance that may benefit individuals with PD in their everyday lives. The findings of this study may be a basis for future research, which may establish rock climbing as a viable adjunct to traditional medical and rehabilitation therapies to reduce the impact of this health condition and enhance physical well-being in this population.

Ethical Approval

This study was approved by Marymount University (Arlington VA USA) IRB as Protocol 831 on June 19, 2023. All participants gave written consent and signed Informed Consent documents, which are in the possession of the principal researcher.

Acknowledgments

We'd like to thank the members of our local PD climbing community who provided advice and guidance as we planned this research, as well as our volunteer research assistants who aided in data collection and data entry. Finally, a special thanks to our research participants.

Disclosures

Conflict of interest: The second author, Molly Cupka, is the President and Founder of UpENDing Parkinson's, which is the organization that provided climbing instruction and oversight for the participants of this study. The other authors, Julie Ries and Andrew Guccione, have no conflicts to report.

Financial support: The authors of this study received no financial support or grants for this study.

Author contributor roles: JR: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Supervision, Visualization, Writing (original draft), Writing (review & editing); MC: Conceptualization, Data Curation, Investigation, Methodology, Project Administration, Resources, Supervision, Visualization, Writing (review & editing); AG: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Visualization, Writing (original draft), Writing (review & editing).

Data availability statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available.

Clinical trial protocol number: This study was registered with clinicaltrials.gov with registration number NCT05919771 and title "Effects of Rock Climbing on Parkinson's Disease." (https://clinicaltrials.gov/study/NCT05919771)

References

1. Willis AW, Roberts E, Beck JC, et al. Parkinson's Foundation P4 Group. Incidence of Parkinson disease in North America. NPJ Parkinsons Dis. 2022;8(1):170. CrossRef PubMed
2. Dorsey ER, Elbaz A, Nichols E, et al. GBD 2016 Parkinson's Disease Collaborators. Global, regional, and national burden of Parkinson's disease, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol. 2018;17(11):939-953. CrossRef PubMed
3. Hirsch MA, Iyer SS, Sanjak M. Exercise-induced neuroplasticity in human Parkinson's disease: what is the evidence telling us? Parkinsonism Relat Disord. 2016;22(suppl 1):S78-S81. CrossRef PubMed
4. Ernst M, Folkerts AK, Gollan R, et al. Physical exercise for people with Parkinson's disease: a systematic review and network meta-analysis. Cochrane Database Syst Rev. 2023;1(1):CD013856. CrossRef PubMed
5. Osborne JA, Botkin R, Colon-Semenza C, et al. Physical therapist management of Parkinson disease: A clinical practice guideline from the American Physical Therapy Association. Phys Ther. 2022;102(4):pzab302. CrossRef PubMed
6. Carapellotti AM, Stevenson R, Doumas M. The efficacy of dance for improving motor impairments, non-motor symptoms, and quality of life in Parkinson's disease: a systematic review and meta-analysis. PLoS One. 2020;15(8):e0236820. CrossRef PubMed
7. Chrysagis N, Trompouki G, Petropaulis D, et al. Effect of boxing exercises on the functional ability and quality of life of individuals with Parkinson's disease: a systematic review. Eur J Investig Health Psychol Educ. 2024;14(5):1295-1310. CrossRef PubMed
8. Domingos J, Dean J, Fernandes JB, et al. Community exercise: a new tool for personalized Parkinson's care or just an addition to formal care? Front Syst Neurosci. 2022;16:916237. CrossRef PubMed
9. Langer A, Hasenauer S, Flotz A, et al. A randomized controlled trial on effectiveness and feasibility of sport climbing in Parkinson's disease. NPJ Parkinsons Dis. 2021;7(1):49. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

10. Langer A, Roth D, Santer A, et al. Climb up! Head up! Climbing improves posture in Parkinson's disease. A secondary analysis from a randomized controlled trial. Clin Rehabil. 2023;37(11):1492-1500. CrossRef PubMed
11. Langer A, Hansen C, Roth D, et al. Vertical locomotion improves horizontal locomotion: effects of climbing on gait and other mobility aspects in Parkinson's disease. A secondary analysis from a randomized controlled trial. J Neuroeng Rehabil. 2024;21(1):63. CrossRef PubMed
12. Gassner L, Dabnichki P, Langer A, et al. The therapeutic effects of climbing: a systematic review and meta-analysis. PM R. 2023;15(9):1194-1209. CrossRef PubMed
13. Liu S, Gong X, Li H, Li Y. The origin, application and mechanism of therapeutic climbing: a narrative review. Int J Environ Res Public Health. 2022;19(15):9696. CrossRef PubMed
14. Dsouza ZL, Rebello SR, Dsilva C. Correlation between community balance and mobility scale (CB&M) with a battery of outcome measures to assess balance in Parkinson's disease - a cross-sectional study. Arch Physiother. 2021;11(1):25. CrossRef PubMed
15. Pauole K, Madole K, Garhammer J, et al. Reliability and validity of the T-Test as a measure of agility, leg power, and leg speed in college-aged men and women. J Strength Cond Res. 2000;14(4):443. CrossRef
16. Alonso CCG, de Freitas PB, Pires RS, et al. Exploring the ability of strength and dexterity tests to detect hand function impairment in individuals with Parkinson's disease. Physiother Theory Pract. 2023;39(2):395-404. CrossRef PubMed
17. de-Oliveira LA, Matos MV, Fernandes IGS, et al. Test-retest reliability of a visual-cognitive technology (BlazePod™) to measure response time. J Sports Sci Med. 2021;179-180. CrossRef
18. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol. 2013;4:863. CrossRef
19. Allen NE, Sherrington C, Suriyarachchi GD, et al. Exercise and motor training in people with Parkinson's disease: a systematic review of participant characteristics, intervention delivery, retention rates, adherence, and adverse events in clinical trials. Parkinsons Dis. 2012;2012:854328. CrossRef PubMed
20. Collins KA, Huffman KM, Wolever RQ, et al. Determinants of drop-out from and variation in adherence to an exercise intervention: the STRRIDE randomized trials. Transl J Am Coll Sports Med. 2022;7(1):e000190. CrossRef PubMed
21. Álvarez-Bueno C, Deeks JJ, Cavero-Redondo I, et al. Effect of exercise on motor symptoms in patients with Parkinson's Disease: a network meta-analysis. J Geriatr Phys Ther. 2023;46(2):E87-E105. CrossRef PubMed
22. Lorenzo-García P, Cavero-Redondo I, Núñez de Arenas-Arroyo S, et al. Effects of physical exercise interventions on balance, postural stability and general mobility in Parkinson's disease: a network meta-analysis. J Rehabil Med. 2024;56:jrm10329. CrossRef PubMed
23. Keus S, Munneke M, Graziano M, et al. European Physiotherapy Guideline for Parkinson's disease. 2014. Online (Accessed October 2024)
24. Wulf G, Lewthwaite R. Optimizing performance through intrinsic motivation and attention for learning: the OPTIMAL theory of motor learning. Psychon Bull Rev. 2016;23(5):1382-1414. CrossRef PubMed
25. Leech KA, Roemmich RT, Gordon J, et al. Updates in motor learning: implications for physical therapist practice and education. Phys Ther. 2022;102(1):pzab250. CrossRef PubMed
26. Johansson H, Hagströmer M, Grooten WJA, et al. Exercise-induced neuroplasticity in Parkinson's disease: a metasynthesis of the literature. Neural Plast. 2020;2020:8961493. CrossRef PubMed
27. Olson M, Lockhart TE, Lieberman A. Motor learning deficits in Parkinson's disease (PD) and their effect on training response in gait and balance: a narrative review. Front Neurol. 2019;10:62. CrossRef PubMed
28. Di Paola M, Caltagirone C, Petrosini L. Prolonged rock climbing activity induces structural changes in cerebellum and parietal lobe. Hum Brain Mapp. 2013;34(10):2707-2714. CrossRef PubMed
29. Bostan AC, Strick PL. The basal ganglia and the cerebellum: nodes in an integrated network. Nat Rev Neurosci. 2018;19(6):338-350. CrossRef PubMed
30. Heino MTJ, Proverbio D, Marchand G, et al. Attractor landscapes: a unifying conceptual model for understanding behaviour change across scales of observation. Health Psychol Rev. 2023;17(4):655-672. CrossRef PubMed
31. Pacheco MM, Lafe CW, Newell KM. Search Strategies in the Perceptual-Motor Workspace and the Acquisition of Coordination, Control, and Skill. Front Psychol. 2019;10:1874. CrossRef PubMed
32. Ranganathan R, Newell KM. Changing up the routine: intervention-induced variability in motor learning. Exerc Sport Sci Rev. 2013;41(1):64-70. CrossRef PubMed
33. Dhawale AK, Smith MA, Ölveczky BP. The role of variability in motor learning. Annu Rev Neurosci. 2017;40(1):479-498. CrossRef PubMed
34. García-López H, de Los Ángeles Castillo-Pintor M, Castro-Sánchez AM, et al. Efficacy of dual-task training in patients with Parkinson's disease: a systematic review with meta-analysis. Mov Disord Clin Pract. 2023;10(9):1268-1284. CrossRef PubMed
35. Johansson H, Folkerts AK, Hammarström I, et al. Effects of motor-cognitive training on dual-task performance in people with Parkinson's disease: a systematic review and meta-analysis. J Neurol. 2023;270(6):2890-2907. CrossRef PubMed
36. Hordacre B, Immink MA, Ridding MC, et al. Perceptual-motor learning benefits from increased stress and anxiety. Hum Mov Sci. 2016;49:36-46. CrossRef PubMed
37. Mellow ML, Goldsworthy MR, Coussens S, et al. Acute aerobic exercise and neuroplasticity of the motor cortex: a systematic review. J Sci Med Sport. 2020;23(4):408-414. CrossRef PubMed

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Arch Physioter 2025; 15: 49-58
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3334
ORIGINAL RESEARCH ARTICLE

The Use of Digital Technologies in Physiotherapy Higher Education: a Mixed-Methods Study

Francesca Trentini[1], Chiara Fante[2], Flavio Manganello[2], Marco Testa[3], Simone Battista[4]

[1]Rehabilitation Department, Villa Rosa Rehabilitation Hospital, Provincial Health Services Company (APSS), Trento - Italy
[2]Institute for Educational Technologies, National Research Council, Genova - Italy
[3]Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genova, Campus of Savona - Italy
[4]School of Health and Society, Centre for Human Movement and Rehabilitation, University of Salford, Salford, Greater Manchester - United Kingdom

ABSTRACT

Background: Previous studies demonstrated that digital tools can be effectively integrated into physiotherapy higher education. However, their adoption remains limited. This study aimed to 1) evaluate the perceived knowledge, confidence, and frequency of digital technology use among Italian lecturers and 2) explore lecturers' experiences with digital technology in higher education.

Methods: We performed a convergent mixed-method study using an online survey instrument for Italian physiotherapy lectures. We employed a 5-point Likert to evaluate perceived knowledge, confidence, and frequency of digital technology use, with consensus defined as an agreement of ≥70% on a statement. An optional qualitative section explored lecturers' experiences with technology, which we analyzed using Reflexive Thematic Analysis.

Results: Between June and September 2023, 118 lecturers (mean age: 45 ± 11; 69% female, n = 81) completed the survey. Participants expressed confidence in utilizing digital tools such as videoconferencing (95%), online repositories (88%), and communication apps (78%). On average, 32% reported using technologies "often" or "always." In the qualitative section, completed by 77 participants, we generated three themes: 1) "Technology can promote a constructive educational approach"; 2) "Action of technologies on students' learning process," with mixed results on their impact; and 3) "Technology is not within everyone's reach," due to barriers to its implementation in didactics. Qualitative and quantitative findings confirmed one another, allowing for a deeper understanding of digital technologies among Italian physiotherapy lecturers.

Conclusions: Our findings indicate that digital tools are still underutilized among Italian physiotherapy lecturers. The main barriers include inadequate infrastructure and a lack of digital competencies.

Keywords: Allied health personnel, Digital technology, Education, Mixed methods, Physical therapy modalities, Physical therapy specialty, Public health professional

What is already known about this topic:

- *Digital technologies, including learning management systems and virtual reality, enhance physiotherapy education, improving student engagement, knowledge, and skills. However, barriers such as limited lecturer confidence, inconsistent use, and inadequate infrastructure hinder widespread adoption.*

What the study adds:

- *The study enhances current knowledge by analyzing digital technology use in Italian physiotherapy education, identifying faculty competency gaps and barriers, and providing insights applicable to global medical education, focusing on technology integration and curriculum innovation.*

Received: October 7, 2024
Accepted: February 14, 2025
Published online: March 10, 2025

This article includes supplementary materials

Corresponding author:
Simone Battista
email: s.battista@salford.ac.uk

Introduction

Digital education is an umbrella term encompassing various approaches, concepts, methods, and technologies (1). Digital Health Education Collaboration has defined digital education as "the act of teaching and learning by means of digital technologies" and digital technologies as "tools, systems and electronic resources that can generate, store or process data" (2). The most common digital technologies

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

in digital learning are learning management systems (3,4), online repositories (5,6), synchronous and asynchronous e-learning technologies (7), social media (8,9), communication apps and tools (10), forums and online learning groups (11), audience response systems (ARS) (12,13), virtual reality interfaces (14-16). In the last few years, several studies have been conducted to investigate the role of technology in education (17-20). The United Nations Educational, Scientific and Cultural Organization (UNESCO) stated that digital learning can enhance teaching practices and improve the quality and sustainability of higher education (21). Evidence showed that incorporating technology into allied health professionals' (e.g., physiotherapists) education is increasingly widespread with several benefits, including flexibility of access to quality content, support of life-long learning, improved communication networks, and personalization of learner experience (22).

For example, online learning modules and learning management systems have been shown to facilitate peer learning, interaction, and training flexibility (3). Multimedia resources and interactive tools such as podcasts and ARS have been proven effective in supporting active student engagement (13,23,24). Virtual reality and computer-assisted learning methods could improve learning outcomes and promote interprofessional collaboration (25). Online technologies in undergraduate physiotherapy education enhance practical performance, facilitate knowledge acquisition, stimulate reflective thinking, and increase awareness of the professional core values (22,26). Additionally, technology-supported interactive teaching models have been found effective in increasing motivation and engagement, improving self-management skills, providing flexibility and transparency in the learning process, and supporting the achievement of educational goals (22,26).

Despite the advances in learning digital technologies, their use remains underutilized and inconsistently implemented across higher education courses, limiting the potential of these resources (30,31). Successful adoption of digital technologies requires careful planning (32), collaborative efforts among educators (33), use of student-centered tools (29), clear instructions (23), flexible and accessible resources (11), effective feedback mechanisms (13), and integration of authentic virtual environments (25). In physiotherapy education, effective integration of digital tools relies on lecturers' technical skills, adaptability, and ongoing professional development (27). Beyond technical literacy, lecturers' belief in their ability to successfully implement technology is a crucial factor influencing its use, as widely reported in the literature (28,29). Additional challenges to adopting digital technologies in physiotherapy education include organizational issues such as large class sizes, limited budgets, and time constraints (32), technical problems, and unclear instructions (33). The lack of technical infrastructure and support was also a barrier (34).

Therefore, the Research Agenda of the American Physical Therapy Association (APTA) prioritized research on using technology to enhance student learning in Education and Professional Development (35). No previous study has investigated the use of educational technologies in Italian physiotherapy undergraduate courses. Therefore, this mixed-methods study aimed to explore Italian lecturers' perceived knowledge, confidence, and frequency of use of key educational technologies in physiotherapy undergraduate programs. The second purpose was to investigate faculty insights into the advantages, disadvantages, barriers, and facilitators of using digital resources, considering the importance of addressing obstacles in technology integration. The findings of this study could guide strategies to improve technology use in physiotherapy education, bridging traditional methods with the digital landscape, enhancing student engagement, and preparing graduates for modern clinical practice.

Materials and Methods

Study Design and Ethical Issues

This study used a convergent mixed-methods design to gather complementary quantitative and qualitative data (36). Data collection and analysis were conducted concurrently and independently. Results were compared and integrated for a more comprehensive understanding. This mixed-methods design was chosen because of its potential to gain deep knowledge of the use and experience with digital technologies in Italian physiotherapy higher education. We performed a web-based cross-sectional survey exploring the perceived level of knowledge, confidence, and frequency of use of the leading educational technologies by lecturers in Italian university physical therapy courses. It also deepened faculty perspectives on the potential of technological tools and possible facilitating and hindering factors in their integration into teaching practices.

Ethical approval was obtained from the Ethics Committee for University Research (CERA: Comitato Etico per la Ricerca di Ateneo), University of Genova (approval date: 22/05/2023; Code: 2023.31). The study was conducted according to the Declaration of Helsinki principles, ensuring ethical standards were upheld to protect the rights and welfare of participants, and reported following the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement (37), the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) (38), the Consolidated Criteria for Reporting Qualitative Research (COREQ) (39) and the Good Reporting of a Mixed Methods Study (GRAMMS) (40).

Survey Instrument Creation and Development

For the creation of the survey instrument, a preliminary literature review was performed to identify the key educational technologies adopted in physiotherapy and other health sciences education (see Supplementary 1 for further information). Nine key educational technologies were identified: learning management systems, online repositories, synchronous e-learning technologies, asynchronous e-learning technologies, social media, communication apps and tools, forums and online learning groups, ARS, and virtual reality interfaces.

Following the International Handbook of Survey Methodology, the authors developed the survey instrument based on the research question, the literature review results, and the expertise of the individual authors (36). The research

team members (FT, SB, CF, FM, and MT) defined the survey instrument's objective and design. FT (She/Her), MT (He/His), and SB (He/His) are all physiotherapists. FT has a Master's degree in Health Professional Rehabilitation Sciences. MT has a PhD in Rehabilitation Science and Physiotherapy. SB has a joint PhD in Neurosciences and Medical Science and is proficient in conducting mixed-methods studies. CF (She/Her) is a researcher with a PhD in clinical psychology, and FM (He/His) is a researcher with a PhD in Engineering Science with a specialization in e-learning. They both have a background in medical education. After two rounds of revision, the authors reached a consensus, and the final draft was compiled. Before being distributed, the questionnaire was tested on five physiotherapy lecturers who agreed on its relevance and clarity. The final online version of the survey instrument was delivered in Italian and lasted 10-20 minutes based on participants' willingness to reply to the qualitative parts. It consisted of 17 items ordered in four sections across four pages. Items were not randomized. A detailed description of the survey instrument and its translated version is reported in Supplementary 2.

Quantitative Section

In section 1 (questions 1 to 2), there were links to the participant information sheet and privacy policy about the study. Participants need to read and accept they had read the documents and agreed with them to proceed. In section 2 (questions 3 to 9), we identified the populations' characteristics, acting as eligibility criteria. Firstly, the participant had to declare to be a physiotherapist and lecturer in a BSc in Physiotherapy in Italy. Then, participants' demographic characteristics were investigated: age, gender, number of years as lecturers, Italian region of the University they were lecturing, and subject(s) they lectured.

In section 3 (questions 10 to 12), the participants had to agree with three statements about their perceived level of knowledge of digital technologies in education, their perceived confidence in using technologies, and their frequency of use of digital tools in the academic year 2022-2023. We chose not to investigate previous academic years because we believe that the COVID-19 pandemic particularly influenced the use of technology in education in those years. Questions in this section were based on the nine clusters of digital technologies identified through the preliminary literature review.

Qualitative Section

Section 4 (questions 13 to 17) was optional and included four open-ended questions to qualitatively investigate the advantages and disadvantages of digital resources and factors limiting and facilitating the use of technology in physiotherapy higher education.

Data Collection

The implemented survey instrument was disseminated online between June 15 and September 25, 2023, through *Microsoft 365 Forms*, a secure web application compliant with the European General Data Protection Regulations (EU Regulation 2016/679) designed to build and manage online surveys and databases while ensuring data encryption, controlled access, and privacy protection throughout the data collection process (41). Consent for participation was gained at the beginning of the survey after a brief cover letter outlining the study's aim and duration. It was necessary to proceed and complete the survey. Completion of the survey instrument was anonymous and entirely voluntary. Researchers' contact details were supplied to enable any questions or concerns to be answered before completing the online survey instrument. Respondents were able to review and change answers before submitting the questionnaire. To submit the questionnaire, all questions had to be answered.

Participants

This study included physiotherapists who had taught for at least one academic year before enrolling in the study. To ensure that participants had these characteristics, at the beginning of section one of the survey instrument, we asked if they had a degree in Physiotherapy and if they had taught for at least one academic year in Italy. People who answered "No" to at least one of these two questions could not continue the survey.

An open survey design was employed. Participants were reached by sending out the online version of the survey instrument link via e-mail through the course leaders of different 'Physiotherapy' Bachelor of Science (BSc) courses. To get in touch with them, we contacted the National Commission for Degree Courses in Physiotherapy, which shared our survey link with the course leaders. This entity coordinates the educational activities of physiotherapy degree programs in Italy and comprises course leaders from Physiotherapy BSc programs across Italy. No incentives were offered to participants to complete the survey instrument. Responses were solicited twice via an e-mail reminder during the data collection period.

Data Analysis

Quantitative Data Analysis

Demographic and Academic Characteristics of the Sample

Descriptive analysis was carried out to understand demographic and academic characteristics. The open answers to question 9 (the lecture's title) were categorized based on the type of subjects (i.e., musculoskeletal rehabilitation, neurological rehabilitation, cardiovascular and chest physiotherapy, urogynecological rehabilitation, oncological rehabilitation, geriatric rehabilitation, pediatric rehabilitation, professional laboratory, anatomy and kinesiology, professional ethics and management, research methodology, aids, and assistive technologies). The categories were identified based on physiotherapists' specific competence areas defined by the Italian professional profile and the deontological code (www.aifi.net).

Knowledge and Confidence in Using Digital Technologies

A 5-point Likert-type scale was used to measure the perceived knowledge (question 10) and confidence (question 11)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

in using digital tools. The scale ranged "completely false" (score 1), "partially false" (score 2), "neither true nor false" (score 3), "partially true" (score 4), and "completely true" (score 5) (42). Participants who considered the statement partially or completely true agreed with the statement, with 70% of the agreement considered consensus (43,44). The frequencies of the answers were reported in a table and graphically represented in a bar chart.

Frequency of Use of Digital Technologies

Participants indicate how frequently they used digital technologies in the last academic year (2022-2023) with a 5-point Likert scale ranging "never" (score 1), "rarely" (score 2), "occasionally" (score 3), "often" (score 4), "always" (score 5). Frequencies were calculated and reported in a table.

Qualitative Data Analysis

Thematic Analysis (TA) was used to analyze the open answers reported in Section 4 (questions 14-17). The analysis process followed the principles of Braun's and Clarke's "Reflexive Thematic Analysis" (RTA). RTA is an interpretive method for analyzing qualitative data that enables identifying and examining patterns or themes within a given dataset (45-48). See Supplementary 3 for further information and theoretical framework.

Integration

Integration was conducted to enhance validation and identify further insights through triangulation of quantitative and qualitative results. Comparisons between quantitative and qualitative findings were interpreted in terms of three possible outcomes (36): (1) confirmation (when quantitative results reinforced qualitative results or vice versa), (2) expansion (when quantitative and qualitative results diverged and helped expand insights by addressing different or complementary findings), and (3) discordance (when quantitative and qualitative findings contradicted or disagreed with one another). In this study, quantitative and qualitative data were analyzed independently using appropriate methods for each data type as described above. Quantitative data were analyzed first, followed by the independent analysis of the qualitative data. No priority was given to one kind of data over the other. No transformation processes (e.g., from quantitative to qualitative or vice versa) were applied. Instead, findings were integrated during the interpretative phase, where results from both datasets were systematically compared to identify convergences, divergences, and complementary insights (36).

Results

Quantitative Analysis

A total of 156 people accepted the invitation to participate in the study, and 118 (age: 45 ± 11, 69% female) completed the survey instrument between June and September 2023. Among them, 77 (age: 44 ±11, 68% female) also completed the optional (qualitative) section 4. Only questionnaires with at least sections 1 to 3 completed were analyzed, as Microsoft Forms 365 does not record submissions where mandatory fields are left blank, and the quantitative section was compulsory. Descriptive characteristics are reported in Table 1.

Perceived Knowledge of and Confidence in the Use of Digital Technologies

Regarding knowledge of the possibilities of using digital technologies in education, consensus was achieved for 5 (56%) out of 9 statements (Figure 1). Physiotherapists declared to know the potential of tools for video conferencing and remote synchronous lessons (95%), online repositories (88%) and learning management systems (82%), tools for blended and asynchronous learning (80%), apps for communication (78%). Regarding confidence in using digital technologies in

TABLE 1 - Participants' demographic characteristics

Descriptive Data	
Age [years (mean, SD)]	45 (11)
Gender [N, (%)]	
Female	81 (69)
Male	37 (31)
Other	0 (0)
Years as lecturers [N, (%)]	
Less than 1 year	4 (3)
1-2 years	18 (15)
3-4 years	23 (20)
5-10 years	32 (27)
More than 10 years	41 (35)
Area of Italy where they lecture [N, (%)]	
North	97 (82)
Center	16 (14)
South	5 (4)
Lectures' topic [N, (%)]	
Musculoskeletal rehabilitation	25 (18)
Neurological rehabilitation	21 (15)
Cardiovascular and chest physiotherapy	10 (7)
Urogynecological rehabilitation	6 (4)
Oncological rehabilitation	6 (4)
Geriatric rehabilitation	8 (6)
Pediatric rehabilitation	4 (3)
Professional laboratory	25 (18)
Anatomy and kinesiology	12 (9)
Professional ethics and management	7 (5)
Research methodology	10 (7)
Aids and assistive technologies	3 (2)

Legend: SD, standard deviation; N, number.

higher education, consensus was achieved for 3 (33%) out of 9 statements (Figure 2). Participants declared themselves to be capable of using apps for communication, tools for video conferencing and remote synchronous lessons, and online repositories and storage systems. Raw data are reported in Table 2 and Table 3.

Frequency of Use of Digital Technologies

Referring to the last academic year 2022–2023, most participants declared they never used virtual reality (81%), forums and online study groups (65%), social media (65%), audience response systems (62%), tools for blended or asynchronous learning (45%), apps for communication (37%) and learning management systems (36%). Tools for video conferencing and remote synchronous lessons were often used by lecturers (20%), as well as online repositories and storage systems. Raw data are reported in Table 4.

Qualitative Analysis: Lecturers' Experience Using Technology

A total of 77 physiotherapists (age: 44 ± 11, 68% female) agreed to participate in the optional section of the survey instrument and answered four open questions about their experience using digital technologies in physiotherapy higher education. Most of them had been teaching for more than ten years (31%), 30% had been teaching for 5-10 years, 19% for 3-4 years, 16% for one or two years and 4% for less than one year. Three themes were developed from the analysis of the lecturers' answers: 1) Technology can promote a constructive educational approach, 2) Action of technologies on students' learning process, and 3) Technology is not within everyone's reach. Supplementary 4 contains all the quotations and codes behind the themes.

Theme 1: Technology Can Promote a Constructive Educational Approach

Lecturers identified various advantages of integrating technology in education and promoting a constructive educational context, leading to the theme: "Technology Can Promote a Constructive Educational Approach" (see Supplementary 4 for quotations and codes). These advantages include the ability to easily store and share data and resources, greater accessibility, and flexibility in teaching, described as *close to the world where students come from*.

TABLE 2 - Perceived knowledge and confidence in using digital technologies

I am aware of the possibility of using the following tools [N, (%)]	Completely false	Partially false	Neither true nor false	Partially true	Completely true
Tools for video conferencing and remote synchronous lessons	1 (0)	1 (1)	5 (4)	19 (16)	93 (79)
Online repositories and storage systems	0 (0)	9 (8)	5 (4)	25 (21)	79 (67)
Learning Management Systems	12 (10)	5 (4)	4 (3)	40 (34)	57 (48)
Tools for blended learning and/or asynchronous e-learning	4 (3)	8 (7)	12 (10)	40 (34)	54 (46)
Apps for communication	6 (5)	5 (4)	15 (13)	26 (22)	66 (56)
Social media	10 (8)	9 (7)	21 (18)	35 (30)	44 (37)
Forums and online study groups	7 (5)	15 (13)	28 (24)	39 (33)	30 (25)
Audience response system	34 (29)	20 (17)	10 (8)	31 (26)	23 (19)
Virtual reality interfaces	32 (27)	24 (20)	28 (24)	22 (19)	12 (10)

Legend: N, number.

TABLE 3 - Perceived confidence in using digital technologies

I feel confident in using the following tools [N, (%)]	Completely false	Partially false	Neither true nor false	Partially true	Completely true
Tools for video conferencing and remote synchronous lessons	4 (3)	3 (3)	5 (4)	37 (31)	69 (58)
Online repositories and storage systems	3 (3)	6 (5)	10 (8)	42 (36)	57 (48)
Apps for communication	7 (6)	6 (5)	12 (10)	31 (26)	62 (53)
Learning Management Systems	18 (15)	11 (9)	13 (11)	45 (38)	31 (26)
Tools for blended learning and/or asynchronous e-learning	9 (8)	13 (11)	23 (19)	40 (34)	33 (28)
Social media	14 (12)	12 (10)	26 (22)	29 (25)	37 (31)
Forums and online study groups	17 (14)	20 (17)	33 (28)	27 (23)	21 (18)
Audience response system	44 (37)	20 (17)	18 (15)	19 (16)	17 (14)
Virtual reality interfaces	61 (52)	22 (19)	14 (12)	16 (14)	5 (4)

Legend: N, number.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

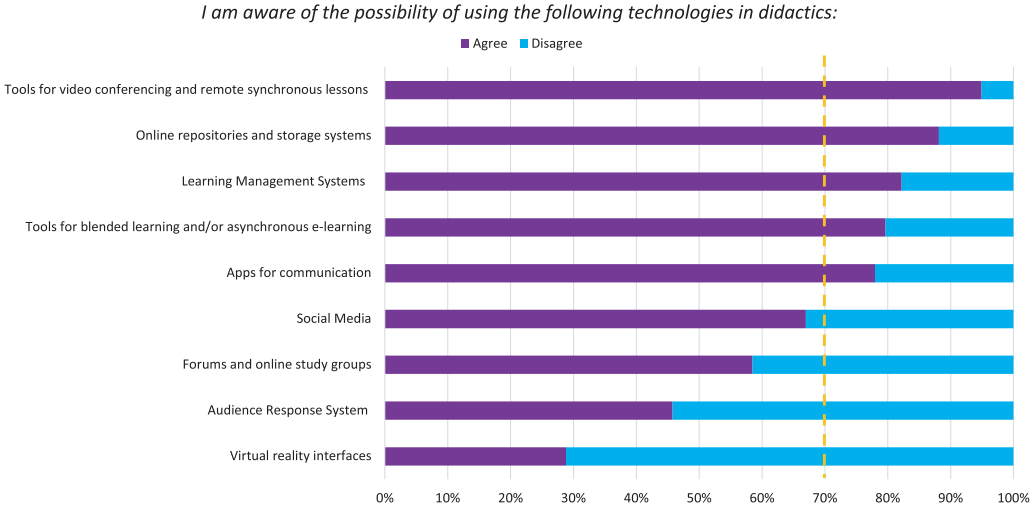


FIGURE 1 - Knowledge of digital technologies

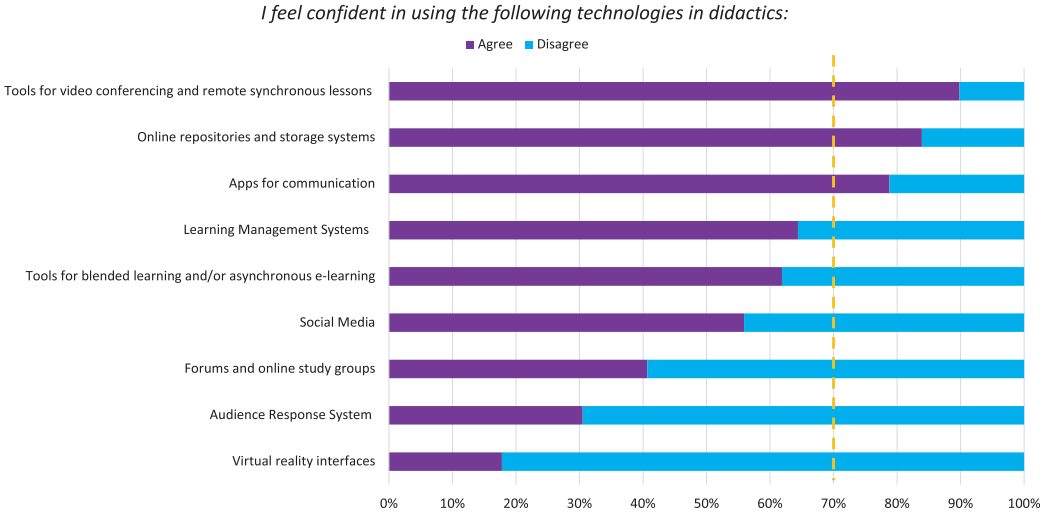


FIGURE 2 - Confidence in the use of digital technologies

They emphasized improved communication and inclusivity enabled by digital tools, which *encourages the participation of even the most timid students*. Digital technology is also a *great support during lessons and for students with difficulty*, allowing the inclusion of students who could not reach the University through distance learning methods. Regarding interactivity, opinions differed: some believed technology increased lesson interactivity, while others felt it did not facilitate interaction, especially in distance learning. Technology enhances lessons and diversifies teaching tools and methodologies, offering blended or asynchronous learning flexibility. Distance learning was viewed as inclusive, accessible, cost-effective, and able to *reduce climate-altering emissions and time lost* by eliminating travel issues. Some participants argued that digital technologies optimize time, though others noted the need for more preparation time and the risk of wasting time during lessons. Nevertheless, technology supports lecturers in organizing resources and lessons, being an *extra tool for the teacher to manage students and lessons* and allowing the provision and receipt of real-time feedback while easily tracking teaching progress. From the lecturers' perspective, using digital technologies completes the educational process and helps achieve educational objectives, increasing effectiveness. A key issue was in clinical practice, where participants recognized the usefulness of virtual reality and simulations but expressed concerns about a potential reduction in hands-on experience in natural settings.

Theme 2: Action of Technologies on Students' Learning Process

Digital technologies can impact each student's learning process, leading to the theme: "Action of Technologies on Students' Learning Process" (see Supplementary 4 for

TABLE 4 - Frequency of use of digital technologies in the last academic year

How often did I use digital tools in the last academic year? [N, (%)] N = 118	Never	Rarely	Occasionally	Often	Always
Learning management systems	42 (36)	6 (5)	16 (14)	32 (27)	22 (19)
Online repositories and storage systems	19 (16)	12 (10)	23 (19)	34 (29)	30 (25)
Tools for video conferencing and remote synchronous lessons	24 (20)	11 (9)	24 (20)	37 (31)	22 (19)
Tools for blended learning and/or asynchronous e-learning	53 (45)	13 (11)	20 (17)	24 (20)	8 (7)
Social media	77 (65)	10 (8)	10 (8)	16 (14)	5 (4)
Apps for communication	44 (37)	17 (14)	19 (16)	15 (13)	23 (19)
Forums and online study groups	77 (65)	12 (10)	10 (8)	12 (10)	7 (6)
Audience response system	73 (62)	13 (11)	13 (11)	10 (8)	9 (8)
Virtual reality interfaces	95 (81)	10 (8)	10 (8)	2 (2)	1 (1)

Legend: N, number.

quotations and codes). The analysis showed that digital tools positively affect students' motivation and engagement. Lecturers agreed that technology supports and accelerates learning, stimulates cooperative learning, and is *excellent for thesis writing or group work meetings*. It facilitates in-depth study and integrates diverse information sources. However, some noted that the rapid pace of digitalization might hinder reflective abilities: *not always everything has to be so fast*. Regarding attention, some felt digital tools enhanced participation and attentiveness, while others believed remote learning and pre-prepared materials might reduce focus during lectures. Regarding interpersonal relationships, lecturers observed that distance learning could diminish human interactions, raising concerns about depersonalization and student passivity, who *may never feel part of the academic journey*. This detachment might adversely affect the acquisition of professional relational skills indispensable for aspiring physiotherapists.

Theme 3: Technology is Not Within Everyone's Reach

Lecturers reported several challenges hindering the effective integration of digital technologies in physiotherapy higher education. This consideration led to the theme: "Technology is Not Within Everyone's Reach" (see Supplementary 4 for quotations and codes). Participants raised concerns about inadequate infrastructure, tool functionality, and university connectivity. They highlighted the need for financial resources, better equipment, and economic incentives for lecturers. Significant obstacles were the lack of university support, comprehensive information, and regulations on using digital technologies in higher education. Robust academic regulations and well-informed guidelines are needed to help lecturers adopt technology and promote blended learning. Finally, another limiting factor reported was "technological illiteracy." Digital tools are used inappropriately and unwittingly because of the lack of knowledge and specific technological skills. The difficulty in using technology was seen as influenced by the age of both lecturers and students. Participants emphasized the need for continuous training for lecturers and the importance of exchanging information among them. Finally, they stressed that lecturers' mentality, motivation, and willingness to change are key to innovation, along with *collaboration with and interest of the students*.

Integration

Integrating quantitative and qualitative findings revealed both confirmation and expansion of the results. Quantitative findings regarding knowledge of digital technologies demonstrated by participants were confirmed by qualitative insights, which emphasized the benefits of digital tools in enhancing accessibility, flexibility, and educational outcomes. Concerning the lack of confidence and the low frequency of use of digital technologies highlighted by the quantitative data, the qualitative findings expanded upon these insights. Specifically, qualitative responses detailed significant barriers and difficulties faced by lecturers in Italian universities, uncovering nuanced perspectives on the challenges of technology adoption, such as inadequate infrastructure, lack of training, and digital literacy gaps, which were not explicitly addressed in quantitative results.

Discussions

This study investigated the perceived knowledge, confidence, and frequency of use of digital technologies in the context of Italian physiotherapy higher education. Our results reflect the literature evidence revealing a general underutilization of digital technologies in physiotherapy education, limited knowledge, confidence, and use of these tools among lecturers, as well as persistent barriers to their integration.

Knowledge of and Confidence in Digital Technologies

Our findings show limited knowledge and confidence in digital tools for physiotherapy education, consistent with previous research on the lack of systematic adoption and reliance on intuitive use (29). Video conferencing and remote lesson tools were the most familiar, likely due to their rapid adoption during the COVID-19 pandemic. These tools' flexibility and cost-effectiveness align with the advantages

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

noted by Vaona et al. (2015) (7). However, many educators expressed concerns about reduced in-person interaction that could lead to a sense of detachment among students, supporting the view that digital tools should complement rather than replace traditional teaching (7,25,44). Aligning with previous research, there was also agreement on the importance of apps and communication tools for maintaining interaction and feedback exchange, which is beneficial for analyzing and adapting activities during lessons (13,30). Lecturers acknowledged that virtual reality might support practical activities in professional laboratories, which aligns with Lucena-Anton and colleagues' statements (2022) (16). However, concerns were raised about the potential devaluation of hands-on experience, which is essential in physiotherapy education. Our findings suggested that integrating various digital resources enables diverse didactic strategies and communication channels, enriching educational offerings. According to previous research, personalized learning experiences, supported by well-aligned technologies and clear learning objectives, enhance students' achievement of their goals (29,45).

Frequency of use

Our findings, consistent with previous studies, revealed that digital technologies were rarely used to support teaching (22,30,31). However, tools for videoconferencing, communication apps, and learning management systems saw high usage rates. These technologies help organize and share didactic materials and create innovative online environments that facilitate active and cooperative learning, promote interaction and collaborative learning, and foster critical thinking (5,44).

Barriers to Implementation

Despite our participants appreciating the benefits of digital technologies, their use remains constrained by several challenges and deficiencies. Technical issues such as inadequate infrastructure and outdated devices were major challenges, emphasizing the need for strong technical support and infrastructure for effective technology use, as reported elsewhere. Revising regulations to foster a blended approach that combines in-person and online learning may offer a more deliberate framework for effectively leveraging technology in higher education (7,25).

Participants underscored internal factors like "technological illiteracy," highlighting the need for better training, organizational support, and incentives from universities. This supports the views of Royal et al. (2016), who stated that being a good lecturer requires continuous knowledge updates, collaboration with education specialists, and considering and studying education as a subject of its own (51). Although Italian universities offer services through "Teaching and Learning Centers" to promote teaching quality and encourage lecturers' training, these services appear underutilized.

Time is crucial for participants, especially since many physiotherapists work simultaneously in University and clinical settings. Another challenge is the resistance to change reported by lecturers, who need help to innovate teaching methodologies. Successful introduction of innovative student-centered approaches relies on the motivation and willingness of lecturers to embrace change, together with students' interest and collaboration (52).

Implications for Education and Research

Our findings underscore the need for targeted interventions to enhance digital technology adoption in physiotherapy education. Universities should provide clearer guidelines and stronger institutional support to foster a blended learning approach that integrates digital and traditional methods effectively. Investments in infrastructure, comprehensive training programs, and incentives for lecturers are crucial for overcoming technical and organizational barriers. Educators must engage in continuous professional development to align their teaching practices with technological advancements. Future research should investigate strategies to increase lecturer and student engagement with digital tools, assessing their impact on learning outcomes and professional preparedness. Furthermore, examining the potential of virtual reality and other immersive technologies could reveal innovative applications for practical learning in physiotherapy.

Strengths and Limitations

Some limitations and strengths of these studies and possible biases need to be acknowledged. This study provides valuable insights into the challenges and opportunities of integrating digital technologies in Italian physiotherapy education. The mixed-methods approach allowed us to capture both quantitative trends and qualitative nuances, enriching the understanding of lecturers' experiences to explore the phenomenon's complexity better. However, our study had some limitations that should be acknowledged. Firstly, our findings were based on a survey instrument created ad hoc by the research team relying on the latest evidence and the individual experience. Calculating the sample size or response rate was impossible, as the total number of physiotherapists who serve as lecturers in Italian physiotherapy degree programs is undefined. Besides, most participants were women, and most of them lived in a similar geographical area (i.e., northern Italy), which could affect the generalizability of the results. We were aware of the risk of selection bias and attempted to mitigate it by distributing the survey across all regions and universities in Italy. We also ensured anonymity to minimize the potential for social desirability bias. Regarding the risk of recall bias, we narrowed the recall period to the previous academic year. We provided clear definitions and examples of the tools mentioned in the questionnaire to enhance the accuracy of participants' responses. For digital bias, we prioritized accessibility by providing clear instructions and user-friendly interfaces, ensuring compatibility across different devices (smartphones, tablets, or computers). To mitigate response bias, we assured anonymity and adopted neutral language to avoid influencing respondents' answers. To address time and availability bias, we offered flexible timing for participants and sent reminders to potential respondents. Finally, to obtain a more representative

sample of participants and ensure that all participants had an equal likelihood of being selected, we sampled from different institutions across Italy.

Conclusions

Despite the literature highlighting the importance of using digital technologies in higher education, our findings indicated that the frequency of use of digital technologies in the last academic year was generally low, and there was a lack of knowledge and confidence in the use of most of the investigated digital tools by the lecturers. However, the lecturers recognized the potential benefits of these capabilities. Among the main factors limiting the introduction of technologies in didactics are the lack of adequate infrastructure and organizational support in universities and the lack of lecturers' digital competencies. Further studies are needed to clarify barriers to the use of technologies better and improve lecturers' adherence to the use of digital tools.

Acknowledgments

The authors thank the Italian Commission for Degree Courses in Physiotherapy and all the course leaders of the different BScs in physiotherapy who helped with the survey distribution.

Disclosures

Conflict of interest: The authors declare no conflict of interest.

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethical approval: Ethical approval was obtained from the Ethics Committee for University Research (CERA: Comitato Etico per la Ricerca di Ateneo), University of Genova (approval date: 22/05/2023; Code: 2023.31).

Authors' contribution: All authors made substantial contributions to the conception or design of the work or the acquisition, analysis, or interpretation of data for the work. All authors drafted the work or revisited it critically for important intellectual content. All authors gave final approval of the version to be published. All authors agree to be accountable for all aspects of the work to ensure that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Ødegaard NB, Røe Y, Dahl-Michelsen T. "Learning is about being active, but the digital is not really active": physiotherapy teachers' attitudes toward and experiences with digital education. Physiother Theory Pract. 2024;40(3):494-504. Internet. CrossRef PubMed
2. Car J, Carlstedt-Duke J, Tudor Car L, et al. Digital Health Education Collaboration. Digital education in health professions: the need for overarching evidence synthesis. J Med Internet Res. 2019;21(2):e12913. CrossRef PubMed
3. Green RA, Farchione D, Hughes DL, et al. Participation in asynchronous online discussion forums does improve student learning of gross anatomy. Anat Sci Educ. 2014;7(1):71-76. CrossRef PubMed
4. Hammarlund CS, Nilsson MH, Gummesson C. External and internal factors influencing self-directed online learning of physiotherapy undergraduate students in Sweden: a qualitative study. J Educ Eval Health Prof. 2015;12:33. CrossRef PubMed
5. Rowe M, Bozalek V, Frantz J. Using Google Drive to facilitate a blended approach to authentic learning. Br J Educ Technol. 2013;44(4):594-606. CrossRef
6. Tilson JK, Loeb K, Barbosa S, et al. Use of tablet computers to promote physical therapy students' engagement in knowledge translation during clinical experiences. J Neurol Phys Ther. 2016;40(2):81-89. CrossRef PubMed
7. Vaona A, Rigon G, Banzi R, et al, et al. E-learning for health professionals. Cochrane Database of Systematic Reviews;2015. CrossRef
8. Gagnon K. Using Twitter in health professional education: a case study. J Allied Health. 2015;44(1):25-33. PubMed
9. Tunnecliff J, Weiner J, Gaida JE, et al. Translating evidence to practice in the health professions: a randomized trial of Twitter vs Facebook. J Am Med Inform Assoc. 2017;24(2):403-408. CrossRef PubMed
10. Frantz JM, Rowe M. Developing reflection and research skills through blogging in an evidence-based practice postgraduate physiotherapy module. Afr J Health Prof Educ. 2013;5(1):3. CrossRef
11. Phadtare A, Bahmani A, Shah A, et al. Scientific writing: a randomized controlled trial comparing standard and online instruction. BMC Med Educ. 2009;9(1):27. CrossRef PubMed
12. Annan-Coultas DL, Kazley AS, Seif GA. Effectiveness of audience response-enhanced case learning activities in graduate health professions education. J Allied Health. 2014;43(3):e53-e58. PubMed
13. Wait KR, Cloud BA, Forster LA, et al. Use of an audience response system during peer teaching among physical therapy students in human gross anatomy: perceptions of peer teachers and students. Anat Sci Educ. 2009;2(6):286-293. CrossRef PubMed
14. Kurul R, Ögün MN, Neriman Narin A, et al. An alternative method for anatomy training: immersive virtual reality. Anat Sci Educ. 2020;13(5):648-656. CrossRef PubMed
15. Liaw SY, Soh SL-H, Tan KK, et al. Design and evaluation of a 3D virtual environment for collaborative learning in interprofessional team care delivery. Nurse Educ Today. 2019;81:64-71. CrossRef PubMed
16. Lucena-Anton D, Fernandez-Lopez JC, Pacheco-Serrano AI, et al. Virtual and augmented reality versus traditional methods for teaching physiotherapy: a systematic review. Eur J Investig Health Psychol Educ. 2022;12(12):1780-1792. CrossRef PubMed
17. Biggs JB. Teaching for Quality Learning at University. (Internet) 1999. Online (Accessed October 2024)
18. Rowe M, Frantz J, Bozalek V. The role of blended learning in the clinical education of healthcare students: a systematic review. Med Teach. 2012;34(4):e216-e221. CrossRef PubMed
19. Hew KF, Lo CK. Flipped classroom improves student learning in health professions education: a meta-analysis. BMC Med Educ. 2018;18(1):38. CrossRef PubMed
20. Lilllejord S, Børte K, Nesje K, et al. Learning and teaching eith technology in higher education - a systematic review. DigiTrans View project. 2018. Online (Accessed October 2024)
21. UNESCO. Technology in education. 2018. CrossRef (Accessed October 2024)
22. Mącznik AK, Ribeiro DC, Baxter GD. Online technology use in physiotherapy teaching and learning: a systematic review of effectiveness and users' perceptions. BMC Med Educ. 2015;15(1):160. CrossRef PubMed
23. Meade O, Bowskill D, Lymn JS. Pharmacology podcasts: a qualitative study of non-medical prescribing students'

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

use, perceptions and impact on learning. BMC Med Educ. 2011;11(1):2. CrossRef PubMed
24. Bertoni G, Marchesini E, Zanchettin FE, et al. Use of audience response systems (ARS) in physiotherapists' training: a qualitative study. BMJ Open. 2023;13(8):e073025. CrossRef PubMed
25. Sabus C, Sabata D, Antonacci D. Use of a virtual environment to facilitate instruction of an interprofessional home assessment. J Allied Health. 2011;40(4):199-205. PubMed
26. Ødegaard NB, Myrhaug HT, Dahl-Michelsen T, et al. Digital learning designs in physiotherapy education: a systematic review and meta-analysis. BMC Med Educ. 2021;21(1):48. CrossRef PubMed
27. Thomas KJ, Denham BE, Dinolfo JD. Perceptions among occupational and physical therapy students of a nontraditional methodology for teaching laboratory gross anatomy. Anat Sci Educ. 2011;4(2):71-77. CrossRef PubMed
28. Abbitt JT. An Investigation of the Relationship between self-efficacy beliefs about Technology Integration and Technological Pedagogical Content Knowledge (TPACK) among preservice teachers. J Digit Learn Teach Educ. 2014;27(4):134-143. CrossRef
29. Zeng Y, Wang Y, Li S. The relationship between teachers' information technology integration self-efficacy and TPACK: A meta-analysis. Front Psychol. 2022;13:1091017. CrossRef PubMed
30. Unge J, Lundh P, Gummesson C, et al. Learning spaces for health sciences–what is the role of e-learning in physiotherapy and occupational therapy education? A literature review. Phys Ther Rev. 2018;23(1):50-60. CrossRef
31. Dolan EL, Collins JP. We must teach more effectively: here are four ways to get started. Mol Biol Cell. American Society for Cell Biology; 2015:2151-2155.
32. Brahler CJ, Quitadamo IJ, Johnson EC. Student critical thinking is enhanced by developing exercise prescriptions using online learning modules. Adv Physiol Educ. 2002;26(1-4):210-221. CrossRef PubMed
33. Juntunen A, Heikkinen E. Lessons from interprofessional e-learning: piloting a care of the elderly module. J Interprof Care. 2004;18(3):269-278. CrossRef PubMed
34. Carbonaro M, King S, Taylor E, et al. Integration of e-learning technologies in an interprofessional health science course. Med Teach. 2008;30(1):25-33. CrossRef PubMed
35. Stevens-Lapsley J, Hicks GE, Zimney K, et al. Research agenda for physical therapy from the American Physical Therapy Association. Phys Ther. 2023;103(10):1-7. CrossRef PubMed
36. Creswell JW, Clark VLP. Designing and conducting mixed methods research. SAGE Publications; 2017. ISBN 1483346986, 9781483346984
37. Von Elm E, Altman D, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med. 2007;147:573–8.
38. Eysenbach G. Improving the Quality of Web Surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. 2004;6:e34. Online (Accessed October 2024)
39. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care. 2007;19(6):349-357. CrossRef PubMed
40. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy. 2008;13(2):92-98. CrossRef PubMed
41. Union E. Regulation (EU) 2016/679 of the European Parliament and of the Council of April 27 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Da. Official Journal of the European Union. 2016;L119:1-88. Online (Accessed October 2024)
42. Joshi A, Kale S, Chandel S, et al. Likert scale: explored and explained. Br J Appl Sci Technol. 2015;7(4):396-403. CrossRef
43. Cooper ID, Johnson TP. How to use survey results. J Med Libr Assoc. 2016;104(2):174-177. CrossRef PubMed
44. Suárez-Álvarez J, Pedrosa I, Lozano LM, et al. Using reversed items in Likert scales: a questionable practice. Psicothema. 2018;30(2):149-158. CrossRef PubMed
45. Braun V, Clarke V. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. Couns Psychother Res. 2021;21(1):37-47. CrossRef
46. Braun V, Clarke V. Is thematic analysis used well in health psychology? A critical review of published research, with recommendations for quality practice and reporting. Health Psychol Rev. 2023;17(4):695-718.CrossRef PubMed
47. Braun V, Clarke V. Thematic analysis: a practical guide. SAGE Publications; 2021.
48. Byrne D. A worked example of Braun and Clarke's approach to reflexive thematic analysis. Qual Quant. 2022;56(3):1391-1412. CrossRef
49. Olivier B, Verdonck M, Caseleijn D. Digital technologies in undergraduate and postgraduate education in occupational therapy and physiotherapy: a scoping review. JBI Evid Synth. Lippincott Williams and Wilkins; 2020:863-892.
50. Priniski SJ, Hecht CA, Harackiewicz JM. Making learning personally meaningful: a new framework for relevance research. J Exp Educ. 2018;86(1):11-29. CrossRef PubMed
51. Royal KD, Rinaldo JCB. There's education, and then there's education in medicine. J Adv Med Educ Prof. 2016;4(3):150-154. PubMed
52. McCabe A, O'Connor U. Student-centred learning: the role and responsibility of the lecturer. Teach High Educ. 2014;19(4):350-359. CrossRef

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Arch Physioter 2025; 15: 59-68
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3301
ORIGINAL RESEARCH ARTICLE

Mim-pong: a serious game for assessment and treatment of the lower limb in hemiparetic stroke patients

Fernando LF Eichinger[1], Fabrício Noveletto[2], Susana C Domenech[3], Thierry Moulin[4], Yoshimasa Sagawa[4], Antonio V Soares[1,4]

[1]Faculty of Physiotherapy and Graduate Program in Health and Environment, University of Joinville Region, Joinville - Brazil
[2]Faculty of Electrical Engineering, State University of Santa Catarina, Joinville - Brazil
[3]Faculty of Design, State University of Santa Catarina, Florianopolis - Brazil
[4]Laboratory for Integrative Neuroscience and Cognitive Psychology (LINC) UMR 1322 INSERM, University of Franche-Comté, Besançon - France

ABSTRACT

Introduction: The motor impairment evidenced post-stroke results in limitations to performing activities of daily living (ADL), especially when it involves locomotion. The Serious Games (SG) are an interesting therapeutic option, as they allow the performance of exercises according to stroke treatment guidelines. However, there is little research exploring the evaluation potential of SG. This study aimed to evaluate the possible metric properties of the mim-pong SG in addition to the therapeutic effects.
Methods: Twenty-four hemiparetic stroke patients were divided into two non-randomized groups: the experimental group (EG) (n = 16) and the control group (CG) (n = 8). Participants were evaluated in terms of motor impairment (lower limb), muscle strength (MS), motor control, and functional mobility.
Results: The significant correlations observed between the score generated by the SG and clinical variables in both groups are highlighted, especially with MS (rho = 0.62-0.66; p = 0.000, and rho = 0.67-0.71; p = 0.002-0.005, for the experimental and CGs, respectively) and motor function of the lower limb for the EG (rho = 0.41, p = 0.018). In addition, the results indicated improvements in all variables in the EG, with superiority over the CG.
Conclusions: This study showed that the mim-pong serious game could be considered a potential resource for the assessment and treatment of hemiparetic stroke patients.

Keywords: Exercise therapy, Hemiparesis, Lower extremity, Rehabilitation, Serious games, Stroke

What is already known about this topic:

- *Serious Games represent a new technology that addresses the main requirements of rehabilitation: attention, motivation and training (repetition).*

What does the study add:

- *The study presents an SG developed for post-stroke rehabilitation. The results showed the therapeutic effects and potential metric properties of this technology.*

Introduction

Stroke is the second leading cause of death and the main cause of long-term disability worldwide (1,2). According to the World Health Organization (WHO), more than 15 million people suffer from a stroke annually (3). Among the clinical aspects, motor dysfunction is the most common and disabling complication of patients, present in approximately 80% of cases (4,5).

Hemiparesis is the typical post-stroke dysfunction, described as partial loss of muscle strength (MS) in the contralateral hemibody to the brain injury (6). Impairment in the lower limbs, especially in the quadriceps femoris (QF) and the hamstrings (HS) muscle groups, has a great impact on the functional independence of patients, as it affects locomotor skills and increases the risk of falls in most patients (7).

The motor impairment evidenced after stroke results in limitations in performing activities of daily living (ADLs) (8), especially when it involves the lower limb. Thus, there are often

Received: September 12, 2024
Accepted: February 25, 2025
Published online: March 24, 2025

Clinical Trial Protocol number: Brazilian Registry of Clinical Trials (RBR-2MF595).

Corresponding author:
Antonio V Soares
email: antonio.vinicius@univille

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

difficulties in performing basic and essential tasks such as walking, sitting, standing up, and going up and down stairs, which significantly reduces quality of life (9). In this context, stroke rehabilitation plays a key role, especially if started early (10).

Motor function recovery is a major rehabilitation target for patients with stroke (11), as it allows the optimization of residual capacities to improve functionality, reflecting greater participation in social activities (5,10). In this direction, the recovery of the lower limb can improve gait, which is one of the main objectives of these patients (12).

There are several therapeutic strategies aimed at lower limb rehabilitation, including conventional kinesiotherapy based on mobilization, stretching, and strengthening exercises (13). Although evidence indicates the beneficial effects of this type of therapy, there are important limitations that must be considered, especially those related to repetitiveness, boredom, and monotony, as they can result in difficulty adhering to treatment (14–16).

Seeking to provide a more engaging and motivating experience for patients and overcoming the limitations of conventional therapy, the use of games as a therapeutic resource for patients with stroke has increased in recent years (17). From this perspective, Serious Games (SG) stand out, which are games created in a virtual environment with specific objectives, allowing customization according to the patient's limitations while combining physical training and motor rehabilitation without losing the common motivational focus of an electronic game (18–20).

SG is an interesting therapeutic option because it allows for the performance of exercises in accordance with the guidelines for stroke treatment, which recommend task-oriented practice with intensity and repetitiveness, adding a higher level of attention and motivation during sessions (13,18). The low cost, accessibility, and portability of SG have made them invaluable for patient assessment and treatment (21).

The effectiveness of SG in stroke rehabilitation has already been proven, especially in approaches aimed at the upper limb (17,18). Regarding application to the lower limb, there are still few studies on effectiveness, and publications aimed at the potential for evaluating these resources are even more uncommon.

Due to the limited research exploring the potential of SG for assessment, the aim of this study was to verify whether the mim-Pong SG developed for the evaluation and treatment of lower limbs in stroke patients has metric properties.

Materials and methods

Design and context of the study

This Non-Randomized Controlled Clinical Trial involved patients with stroke. The study was carried out at the Center of Research in Neurorehabilitation at the Neurology Outpatient Clinic of the Guilherme Guimbala College, Joinville, Santa Catarina, Brazil.

Participants

The inclusion criteria were patients with subacute or chronic hemiparesis due to stroke (injury time ≥3 months), clinically stable, and able to walk independently. The exclusion criteria were hemiparesis due to other diseases, symmetric bilateral motor impairment, severe visual and/or auditory impairment, uncooperative patients and/or with severe cognitive deficits, and patients who were performing or who performed (in the last 3 months) any other type of rehabilitation for the trunk and/or lower limbs.

This study was approved by the Ethics Committee for Research with Humans (CAAE 56995816.6.0000.0118), by the Brazilian Registry of Clinical Trials (RBR-2MF595), and the procedures were in accordance with the Helsinki Declaration.

Evaluations

The evaluations and instruments used were chosen according to the International Classification of Functioning, Disability, and Health (ICF) domains, developed by the WHO in 2004.

- Anthropometric digital scale and stadiometer to measure body mass and height, respectively.
- Fugl-Meyer Assessment Scale (FMAS) to measure the level of motor impairment in the patients (22). Only the section for motor evaluation of the lower limb was used, which includes the analysis of reflex activity, synergic muscular action in flexion and extension, and movements with and without synergy.
- Mini-Mental State Examination (MMSE) to assess the cognitive level of patients. This instrument was used only for patient screening following the cut-off points related to schooling (23).
- *Mim-Pong* SG: for assessment of the performance during the game, MS and motor control (MC). *Mim-pong* SG was also used for the training of patients. This game operates by means of a system based on a load cell (IWM brand, single point load type, Brazil, capacity of 589 N, accuracy <0.017% – full scale).

The score for assessing the patient's performance during the training considers the aspects related to the gameplay and is obtained through the equation proposed previously by this study's research group (19). This SG has simple visual aspects that focus the patient on hitting the ball using a "racket" (Fig. 1). Rackets move simultaneously on the screen's vertical sides based on the strength applied to the load cell, which can be calibrated individually with the evaluation of the maximal voluntary isometric contraction (MVIC) of each muscle group in each session. However, when no strength is applied, the rackets stay still at the bottom of the screen. The upper limit of the rackets (at the top of the screen) is based on the MVIC of each subject. The horizontal walls bounce the ball back. To improve the challenge, game parameters can be adjusted: for example, the racket size, ball size, ball speed, and duration of the match.

Two muscle groups of the lower limb were evaluated: QF and HS. For operation, specific software and hardware were developed. The hardware used the signal obtained from the load cells as a dynamometry (Fig. 2).

For the evaluation of MS, the load cell was coupled to a device that allows the adjustment and positioning of the region to be worked, acting as a hand-held dynamometer for

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

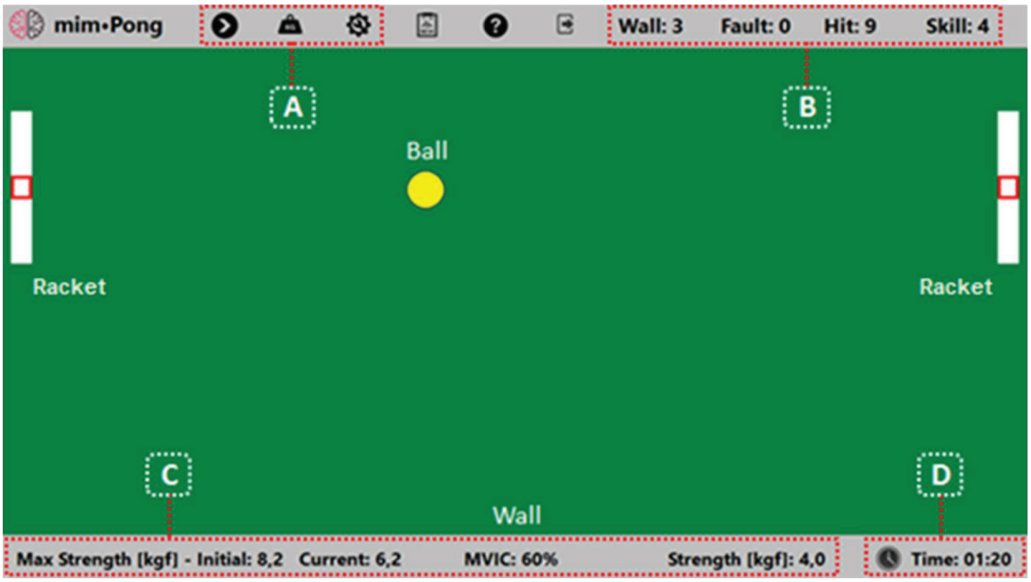


FIGURE 1 - The mim-pong SG main screen.
A) Game control and settings; B) Game score components; C) Muscular strength data; D) Game session time elapsed. The figure was created by (7).

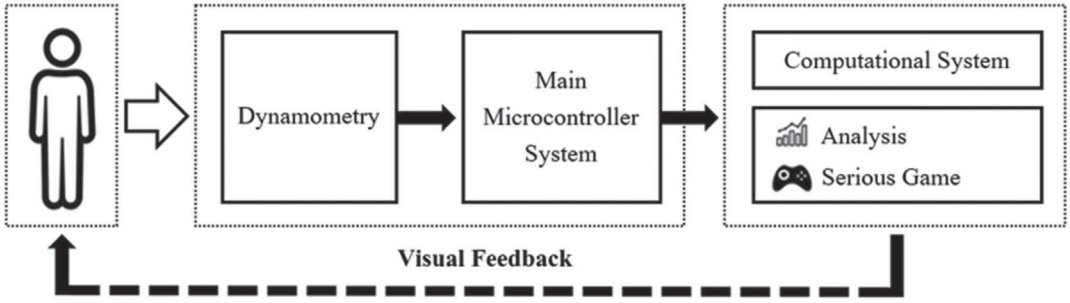


FIGURE 2 - General diagram illustrating the basic units of the system used for the evaluation and treatment of patients.

measuring any muscle group (19). The software enables the configuration of hardware-related parameters, calibration for acquisition initialization, real-time visualization of the captured signal, and recording of acquisition data.

For evaluation and training with the *mim-Pong* SG, a leg extensor apparatus was prepared and adapted with a load cell, which was coupled to the chair and allowed to assess the strength bilaterally of the muscle groups previously mentioned (QF and HS). Both the patient and device position depend on the muscular group analyzed as follows. For the QF, the patient sits in the chair with the trunk supported by the backrest, legs hanging, hip at 110º of flexion relative to the trunk, and knees flexed at 90º (7). The sensor is placed perpendicular to the distal third of the leg (just above the malleolar region) on the anterior face. For the HS test, the patient is positioned in the chair as in the QF test, but with knees flexed at 60º (24), and the sensor is placed perpendicular to the distal third of the leg on the posterior face (19).

For the evaluation of MS, three measures were performed bilaterally of each muscle group in the MVIC for 5 seconds (24), with an interval of 1 minute between each measurement. All evaluations (QF and HS) were conducted on alternate days, with a minimum interval of 24 hours, and performed by the same examiners. The patient was instructed to perform as much force as possible from a green signal projected on the screen, which indicated the beginning of the test, and this force should be maintained until the green color disappears, indicating the end of the test. After the MS measurement, the patient had a rest period of 2 minutes to then carry out the MC assessment protocol.

The assessment of patients' MC consisted of a task in which the goal was to hit a ball at 5 different levels, from 1 to 5 (Figure 3). These levels are associated with the MVIC obtained during the MS. At level 1, the patient should produce strength equivalent to 20% MVIC to hit the ball in the racquet's center. Levels 2 to 5 represent the percentages of 40, 60, 80, and 100% MVIC, respectively. The rationale for these different levels of strength is that most tasks performed in daily life do not require the use of maximum strength for their performance (25,26). The ball moved in a straight line from the left side of the screen to the right side, where the racket was located at different levels, representing the percentages of the MVIC. One ball leaves every 10 seconds, starting with the lower strength level (20% of the MVIC) up to the maximum level (100% of the MVIC). At the end of the test, the *mim-pong* SG provides an evaluation score that indicates the patient's MC. The maximum score in the evaluation

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

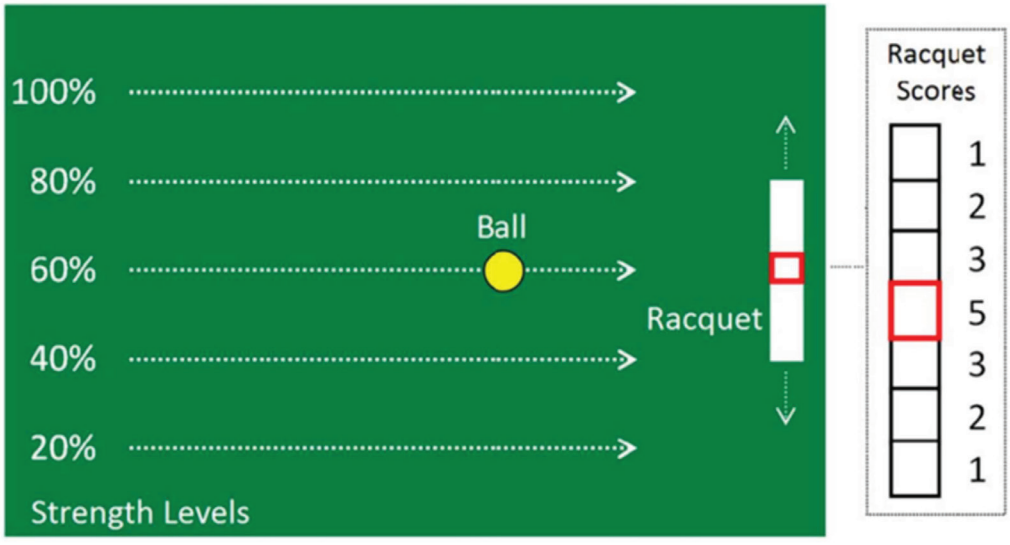


FIGURE 3 - Assessment protocol for motor control.
The hit score values on the racquet are shown zoomed on the right side. The closer to the racquet's center, the higher the score. The score is equal to zero if the ball does not hit the racquet. This figure was created by (19).

mode is 100. The patient's MC score was calculated according to the equation created by (19), which considers the position of hitting the racket as well as the effort required to hit the ball at each level of the assessment mode. Lower levels require less MS and, therefore, represent a smaller weight in the score's composition.

– Timed Up and Go (TUG) Test: to assess functional mobility by measuring the time taken to perform the proposed task (7).

Interventions

After the evaluation phase, the patients of both groups were submitted to the same treatment period, 10 consecutive weeks, with two sessions per week on alternate days for a total of 10 hours of training (each training session lasted 30 minutes). The participants were divided into two non-randomized groups: the experimental group (EG 16 patients) and the control group (CG 8 patients). Due to operational constraints related to the researchers' availability to conduct the experiment, the CG consisted of only 8 patients.

The EG received a rehabilitation program for the paretic lower limb based on exercises using the *mim-pong* SG. This exercise program (as described in assessing the patient's performance during the training) was divided into two phases with 10 sessions: phase 1 (using 60% of the MVIC) and phase 2 (using 80% of the MVIC). These percentage values fulfill the guidelines of resistance exercises for patients with stroke (27). The same configuration of the game (racket size, ball size, and ball speed) was standardized for all patients. In each session, the passive mobilization of the paretic side was performed for a period of 10 minutes, after which the training was started. An SG calibration was always performed by measuring the MVIC, and after obtaining this data, the exercise was started with the patient. Each session was composed of three sets of 2 minutes with a 1-minute break between each set for rest for each paretic muscle group (QF and HS). The sessions always started with the strengthening of the QF, and then the exercises for HS were performed after.

The CG received a rehabilitation program based on conventional kinesiotherapy (passive mobilization and stretching exercises). The sessions performed with this group also lasted 30 minutes. Initially, 10 minutes of passive mobilization were performed. This was followed by 20 minutes of strengthening exercises using manually resisted active movements conducted by the researcher, as load tolerance varied among patients. It is important to note that, as in the EG, training was performed exclusively on the paretic side.

Evaluations

The evaluations were carried out three times in the week prior to initiating the rehabilitation and three times in the week following the end of the program. The arithmetic mean resulting from these three evaluations was used for further analysis.

Statistical analysis

The data was analyzed using IBM SPSS for Windows, version 20.0. A per-protocol analysis was performed. Descriptive statistics (mean, median, standard deviation, and frequency distribution with absolute and percentage values) were used for participant characterization. Data normality was verified by the Shapiro-Wilk test and the homogeneity of variances by the Levene test. The one-way ANOVA and Tukey's multiple comparison tests were used to verify the effect of the group (EG *vs* CG) and the effect of the condition (pre- and post-intervention) on the studied variables. The variables that did not meet the assumptions for the ANOVA underwent two-group comparison tests: the Mann-Whitney U test when comparing the EG and CG or the Wilcoxon rank-sum test when comparing pre- and post-intervention conditions. Effect Size (ES) was calculated to evaluate whether the observed differences corresponded to clinically significant effects. For parametric data, ES values <0.2, 0.2-0.8, and ≥0.8 were considered small, medium, and large, respectively. For non-parametric data, ES values were <0.1, 0.2-0.5, and ≥0.5. The Spearman correlation test was used to evaluate the potential

metric properties of the *mim-pong* SG. The correlation values <0.3, 0.2-0.5, and >0.5 were considered small, medium, and large, respectively. A significance level of 0.05 was used for all tests.

Results

Table 1 presents the sociodemographic, clinical, and anthropometric characterization of the participants of the EG and CG. Twenty-four patients (12 males and 12 females; mean (SD) age of 58 (10) years old; mean (SD) time since stroke of 16.8 (19.6) months) participated in this study. In both groups, a predominance of ischemic stroke was observed (EG = 80% and CG = 62.5%), and there was a prevalence of hemiparesis on the left side of the body. All patients completed all stages of the experiment. A comparison of the sociodemographic, clinical, and anthropometric variables between the EG and the CG in the pre-intervention period revealed that there were no statistically significant differences between groups in any of the variables studied.

Table 2 shows the results of the correlation analyses between the scores obtained with mim-pong SG and the other variables studied, which were used to examine the potential metric properties of the SG. The score provided by the SG was significantly correlated with MS in both the QF and HS (large correlation) in both groups. A moderate and significant correlation between the HS score and the TUG was observed only in the CG, while a medium and significant correlation between the QF score and the FMAS was found only in the EG.

Regarding the comparison between groups (Table 3), in the pre-intervention condition, no statistically significant differences were observed in any variable.

Regarding the comparison between groups (Table 3), in the post-intervention condition, statistically significant differences were observed for QF MC of the paretic (p = 0.001) and nonparetic (p = 0.029) sides and for HS MC of the nonparetic sides (p = 0.029).

Comparisons between pre- and post-intervention conditions in the EG showed statistically significant differences for all the variables studied except the QF MS of the non-paretic side (p = 0.093).

Comparisons between pre- and post-intervention conditions in the CG showed statistically significant differences only for QF MC (p = 0.003) and HS MC (p = 0.006).

TABLE 1 - Characterization of people with stroke

Characteristics	EG (n = 16)	CG (n = 8)	*p-value*
Sex (female; f, %)	8.50%	4.50%	–
Age (years)	56.8 (10.8)	59.8 (9.8)	0.515[a]
BMI (kg/m^2)	29.1 (7.8)	28.3 (5.8)	0.796[a]
Hemiparesis (left; f, %)	9.56%	4.50%	–
Stroke time (months)	19.3 (23.1)	13.8 (12.3)	0,666[b]
Phase of the stroke (chronic; f, %)	11.67%	5.62%	–
Type of the stroke (Ischemic; f, %)	13.81%	5.62%	–

Significance level p < 0,05; p-value: probability of significance obtained by the independent t-test ([a]) or by the Mann-Whitney U test ([b]). EG: experimental group; CG: control group; BMI: body mass index; n: total sample; f: absolute frequency.

Discussion

There are already some studies indicating the potential benefits of SG for the rehabilitation of patients with stroke. Our study aimed to explore the possible metric properties of the *mim-pong* SG, as well as to evaluate the therapeutic effects derived from an exercise program using this resource.

The analysis of the metric properties of the game (Table 2) indicated that the *mim-pong* SG can be used not only for treatment but also as an assessment tool for patients. The highest correlations observed (large and significant) were between the scores generated by the SG and MS variables in the QF and HS in both groups. One hypothesis to explain this finding is related to the elements that compose the formula for calculating the game score, with MS being one of the clinical elements considered. Therefore, as patients show improvement in their MS levels, the score also tends to be higher.

TABLE 2 - Correlation coefficients of the score obtained through the mim-pong SG and the other variables of the QF, hamstring muscles, Fugl-Meyer scale, and timed up-and-go test in both groups

	Mim-pong scores							
	EG				CG			
Variables	Score QF		Score HS		Score QF		Score HS	
	rho	*p*	rho	*p*	rho	*p*	rho	*p*
QF strength*	0.62	0.000	0.29	0.109	0.67	0.005	0.35	0.184
Hamstrings strength*	0.66	0.000	0.41	0.019	0.71	0.002	0.74	0.001
FMAS	0.41	0.018	0.34	0.061	0.35	0.185	0.45	0.082
Timed up and go	−0.29	0.105	−0.33	0.065	−0.39	0.138	−0.54	0.032

rho, Spearman correlation; *p*, Significance level <0.05; *only paretic side; Numbers in red indicate statistical significance.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 3 - Results of the group's effect analysis and the condition (pre- and post-intervention) on the studied variables

	Group	Pre	Post	p-value (pre vs post)	ES	p-value (EG vs CG)*
Paretic side						
QF MS (kgf)b	Exp.	14.8 (6.6)	21.2 (11.6)	0.000	0.6	0.268
	Cont.	14.7 (11.9)	15.4 (12.0)	0.185	0.1	
Hamstrings MS (kgf)[a]	Exp.	5.4 (3.0)	10.2 (4.3)	0.000	1.3	0.322
	Cont.	7.1 (6.3)	8.1 (6.0)	0.069	0.5	
QF motor control (score)[b]	Exp.	67.8 (18.7)	87.8 (12.8)	0.000	0.6	0.001
	Cont.	52.4 (19.9)	65.0 (17.1)	0.003	0.7	
Hamstrings motor control (score)[a]	Exp.	54.3 (21.8)	78.9 (16.6)	0.000	1.3	0.220
	Cont.	51.3 (27.7)	64.8 (25.2)	0.006	0.5	
Nonparetic side						
QF MS (kgf)[b]	Exp.	24.3 (7.3)	27.5 (10.9)	0.093	0.3	0.462
	Cont.	23.9 (8.8)	24.2 (10.0)	0.761	0.1	
Hamstrings MS (kgf)[a]	Exp.	11.7 (4.2)	15.2 (5.2)	0.000	0.8	0.342
	Cont.	12.8 (7.0)	13.1 (4.4)	0.744	0.1	
QF motor control (score)[b]	Exp.	78.8 (11.7)	87.6 (4.7)	0.004	0.5	0.029
	Cont.	65.6 (16.1)	74.3 (13.7)	0.142	0.5	
Hamstrings motor control (score)[a]	Exp.	69.9 (14.2)	82.3 (12.6)	0.001	0.8	0.029
	Cont.	64.8 (20.9)	69.3 (13.4)	0.407	0.3	
FMAS[b]	Exp.	21.1 (4.8)	24.3 (3.3)	0.000	0.6	0.877
	Cont.	22.4 (6.3)	23.9 (4.6)	0.068	0.5	
TUG test (s)[b]	Exp.	22.5 (20.4)	18.4 (15.7)	0,001	0.6	0.668
	Cont.	33.3 (28.4)	28.9 (22.0)	0.247	0.4	

Significance level $p < 0.05$; p-value: probability of significance obtained by the paired t-test ([a]) or by the Wilcoxon signed-rank test ([b]); ES: effect size; EG: experimental group; CG: control group. *comparison post-intervention

Among the important variables in resistance training are intensity and the number of repetitions performed during an exercise (28). The utilized SG provided an intense training session with a high number of exercise repetitions, which contributed to the observed improvements. In this sense, training with adequate intensity and repeatability can improve MC in the lower limbs, in addition to increasing MS (29).

In the EG, a moderate and significant correlation was also found between the score and the FMAS, which evaluates motor impairment in patients. This scale includes the analysis of various factors related to motor function, including synergistic muscle action in flexion and extension (22), which may explain the findings. In the CG, a moderate and significant correlation was observed between the score and functional mobility. This is supported by a study involving 21 stroke patients. It was found that MC deficits rather than MS deficits predict functional impairments in mobility (30). Therefore, it is recommended to prioritize rehabilitation interventions aimed at improving MC as they may yield better functional outcomes in these patients. In fact, most daily activities do not require maximum levels of MS (25), which makes the *mim-pong* SG even more valuable as it takes this factor into account by requiring patients to adjust to various levels of contraction, thus promoting improvements in MC.

In commercial games frequently used in rehabilitation, there is no concern while designing game scores toward the patient's clinical monitoring (16). Thus, as already mentioned, the SG scores were related to important clinical variables (MS, motor impairment, and functional mobility), indicating that the *mim-pong* SG can contribute not only to the treatment but also to the assessment of patients, being a measure of therapeutic progress.

The muscle strengthening programs based on SG should include elements described as essential in rehabilitation, for example, intensity, repetitiveness, and task-oriented training (19). Thus, they become effective for the rehabilitation of these patients and contribute to a more attractive and motivational therapeutic process, overcoming the limitations often evidenced in conventional therapies (13,18,19). It is believed that motivated behaviors lead to improved performance, and thus, motivation has been shown to be an important element

in the association between cognition and motor performance, playing a crucial role in rehabilitation outcomes (31).

In this study, the exercise program with the *mim-pong* SG resulted in significant improvements in all variables analyzed in the EG. In the group receiving conventional treatment, significant improvements were observed only in the MC of the QF and the HS in the paretic limb.

The hemiparesis present after a stroke is related to high rates of functional disability, evidenced by the difficulty in performing ADL (8,11,32). Muscle weakness in the lower limbs limits fundamental activities such as walking, sitting, standing, and climbing stairs, which significantly impacts quality of life (6,33). In general, the participants of this study had an important hemiparesis in both muscle groups, but more evident in the HS. In the EG, 39.3% lower QF strength and 53.0% lower HS strength were observed than in the non-paretic side. In the CG, much lower strength levels were also observed on the paretic side, 38.5% in the QF and 44.5% in the HS.

An important clinical result of this study was the MS gain. In the EG, significant gains were obtained on the paretic side with a large ES for the HS (p = 0.000, ES = 1.3) and moderate ES for the QF (p = 0.003, ES = 0.6). The EG also showed an improvement in the non-paretic side's MS, with statistically significant results and a large ES for the QF (p = 0.004, ES = 0.5) and the HS (p = 0.001, ES = 0.8). In the CG, there were no significant improvements, and the ES was small for the QF and moderate for the HS on the paretic side. Although for MS, no statistically significant differences were found when comparing between groups in the post-experimental period, it is possible to see that the effects were greater in the EG. In recent years, there has been an increase in studies attempting to detect the so-called minimum clinically important difference (MCID), defined by Jaeschke et al. (34) as the "smallest difference in scores in the domain of interest that patients perceive as beneficial and that would require, in the absence of troublesome side effects and excessive costs, a change in patient management." In this sense, the literature indicates that for MS in the QF in individuals with limited physical activity, an MCID of 3.0 kgf is considered (35). Thus, comparing the means between the groups, indeed, the EG obtained greater improvement (difference of 5.7), while the CG did not reach the MCID. In the MS of the HS, a greater difference was also noted in the EG compared to the CG (difference of 3.8).

Systematic reviews show that resistance training has been beneficial in patients with stroke for improving lower limb MS (28, 36). In this sense, the *mim-pong* SG was effective in recovering the patients' MS, with greater improvements than those obtained with conventional therapy, which involved mobilization, free and resisted active exercises, and stretching. Other research supports these findings (37-39).

Post-stroke muscle dysfunction is a multifactorial phenomenon (40), so there is already evidence that hemiparesis involves factors that go beyond the weakness (40). Thus, therapeutic modalities that provide improvements in MC, in addition to MS, are extremely important. The exercise program using the *mim-pong* SG is an example of an intervention that has this purpose. In this SG, the score of the game evaluates the MC, where the closer to the center of the racket the patient hits the ball, the higher the score. In this way, the SG provides increased MS using different muscle activation strategies.

After the experimental treatment, there were important improvements in MC only in the patients of the EG. In this group, a significant improvement with a large ES was observed in the QF and HS in the paretic limb (p = 0.000, ES = 0.6/p = 0.000, ES = 1.3) and in the non-paretic limb (p = 0.004, ES = 0.5/p = 0.001, ES = 0.8). The CG only showed significant improvements in the paretic limb, but it was moderate ES (p = 0.003, ES = 0.7/p = 0.006, ES = 0.5). The comparison of these findings is hampered by the fact that there were no studies that analyzed this variable in a similar way.

Although exercises were only performed with the paretic limb, hypotheses can be raised to justify the improvements observed in MS and MC on the contralateral side. A justification for this is related to learning, where the individual presents a greater facility in performing activities in which they already have some experience (11). Another factor that may have resulted in improved MC in the non-paretic limb is the so-called "cross-effect" of training. From the neurological point of view, this can be explained since there are not only intrahemispheric communication pathways but also interhemispheric pathways. The main communication structure between the cerebral hemispheres is the corpus callosum, which consists of the largest set of nerve fibers that perform this function (41), causing the transcortical flow of information and playing an important role in the control of movement (42). Indeed, studies suggest that the integrity of the sensorimotor regions of the corpus callosum correlates with motor function after stroke (43).

In a randomized clinical trial involving 139 patients with stroke, it was found that strength training on the non-hemiparetic side promotes motor function recovery (11). Similarly, a study investigating the effects of non-paretic limb training for improvements in the paretic limb of chronic stroke patients has found that this type of intervention can result in gains related to MS and muscle activation (45). Our research indicated significant improvements with a large ES in MS variable after intervention using the exercise program from the SG in the EG (p = 0.001, ES = 0.6). In the CG, which received treatment based on conventional kinesiotherapy, the improvements were not significant, and the ES was medium (p = 0.247, ES = 0.4). Some studies have used SG, virtual reality, or video games to improve lower limb motor functions in patients with stroke, especially with regard to mobility (46-49). Our findings align with the results of these prior studies. A meta-analysis involving 22 studies found that Wii Fit-based rehabilitation accompanied by conventional therapy results in improvements in functional mobility evaluated with the TUG (46). However, another meta-analysis indicates that there is still a lack of high-quality evidence for the use of virtual reality rehabilitation in stroke treatment (49). It is known that a difference of 2.9 seconds in TUG is the MCID for patients with stroke (50). Thus, in the present study, superiority in the results of the EG in relation to the CG.

Regarding the variable motor impairment, significant improvements were also observed, with large ES in the EG

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

(FMAS p = 0.000, ES = 0.6). Thus, the *mim-pong* SG can be considered a beneficial resource for motor rehabilitation for patients with stroke. In fact, SG therapy can enhance the effects of conventional rehabilitation when they are used in combination (51).

Patients with stroke commonly face challenges in understanding data indicating the clinical improvements that occur periodically during treatment. In this context, the scores generated by the *mim-pong* SG can contribute to a clearer perception of the changes observed from one session to another, establishing a more easily comprehensible parameter for comparison. Collectively, the results (MS, MC, FMAS) corroborate with patients' perceptions as when they were asked about their perception of improvement throughout the clinical trial, some participants in the present study responded that they were getting better because their scores were increasing day by day.

Finally, the present study involved two groups of patients with similar general characteristics: gender, age, body mass index (BMI), laterality, hemiparesis, stroke time, and phase of stroke. This allows a more reliable comparison between the effects of the different treatment proposals used with the groups. It is important to emphasize that most patients from both groups were in the chronic phase, a period characterized by greater difficulty in obtaining improvements when compared to the subacute or acute phase (52). These characteristics of the participants make the results even more valuable since they indicate that the intervention using the *mim-pong* SG can result in important improvements, even in chronic patients who are often already disillusioned with rehabilitation. Improvements in the chronic phase of stroke are more challenging to achieve compared to the acute or subacute phases. The chronic phase, defined as six months or more after stroke, is characterized by a reduced neuroplastic capacity due to factors such as alterations in cortical activity levels, changes and decreased responsiveness of neuroplastic mechanisms, as well as a reduction in neurotrophic factors such as Brain-Derived Neurotrophic Factor (53,54,55). Thus, since the Serious Game involves specific and high-intensity exercises, it can enhance cortical stimulation, promoting greater neuroplastic activity.

MS training promotes improvements and maintenance of this aspect for up to 4 years post-stroke (50). The present research corroborates with these authors about the existence of improvements in the chronic phase. However, significant improvements were also observed in patients with lesions received approximately eight years ago. This finding may justify the functional improvements that are often achieved with rehabilitation many years after the event (36).

The limitations of our current study include the lack of randomization and a small participant pool, as well as the absence of tools to assess patients' fatigue and motivation.

Conclusions

In conclusion, our study showed that the exercise program conducted with SG resulted in improvements across all analyzed variables, surpassing the conventional therapy used. Furthermore, correlations were observed between the SG-generated scores and clinical variables, particularly MS and lower limb motor function. Thus, the *mim-pong* SG can be considered a potential resource for the assessment and treatment of patients with stroke.

For future research, it is recommended to conduct a randomized clinical trial with a larger sample size and to explore the use of the *mim-Pong* SG with different populations and objectives.

Acknowledgment

The authors would like to thank the institutional support of the University of Joinville Region, the State University of Santa Catarina (UDESC), and Guilherme Guimbala College. The authors would like to thank Alexander Hansen for his editorial assistance.

Disclosures

Conflicts of interest: There is no conflict of interest regarding the present study.

Author contributions: All authors made a significant intellectual contribution to the development of this article, participating in all stages (design, collection, analysis, data interpretation, writing, and reviewing the text). Everyone approved the final version of the article.

References

1. Lavis H, Van Vliet P, Tavener M. Stroke survivor, caregiver and therapist experiences of home-based stroke rehabilitation: a thematic synthesis of qualitative studies. Phys Ther Rev. 2023;28(2):157-173. CrossRef
2. Campagnini S, Liuzzi P, Mannini A, et al. Effects of control strategies on gait in robot-assisted post-stroke lower limb rehabilitation: a systematic review. J Neuroeng Rehabil. 2022;19(1):52. CrossRef PubMed
3. Ahmed T, Islam MR, Brahmi B, et al. Robustness and tracking performance evaluation of PID motion control of 7 DoF anthropomorphic exoskeleton robot assisted upper limb rehabilitation. Sensors (Basel). 2022;22(10):3747. CrossRef PubMed
4. Park S, Tang A, Pollock C, et al. Telerehabilitation for lower extremity recovery poststroke: a systematic review and meta-analysis protocol. BMJ Open. 2022;12(3):e055527. CrossRef PubMed
5. Xiong F, Liao X, Xiao J, et al. Emerging limb rehabilitation therapy after post-stroke motor recovery. Front Aging Neurosci. 2022;14:863379. CrossRef PubMed
6. Wang CY, Olson SL, Protas EJ. Test-retest strength reliability: hand-held dynamometry in community-dwelling elderly fallers. Arch Phys Med Rehabil. 2002;83(6):811-815. CrossRef PubMed
7. Eichinger FLF, Soares AV, Noveletto F, et al. Serious game for locomotor rehabilitation of hemiparetic stroke patients. Fisioter Mov. 2020;33:e003316. CrossRef
8. Schwarz A, Al-Haj Husain A, Einaudi L, et al. Reliability and validity of a wearable sensing system and online gait analysis report in persons after stroke. Sensors (Basel). 2023;23(2):624. CrossRef PubMed
9. Wang YF, Chen WY, Lee CT, et al. Combinations of scalp acupuncture location for the treatment of post-stroke hemiparesis: a

systematic review and Apriori algorithm-based association rule analysis. Front Neurosci. 2022;16:956854. CrossRef PubMed
10. Demeco A, Zola L, Frizziero A, et al. Immersive virtual reality in post-stroke rehabilitation: a systematic review. Sensors (Basel). 2023;23(3):1712. CrossRef PubMed
11. Shao C, Wang Y, Gou H, et al. Strength training of the nonhemiplegic side promotes motor function recovery in patients with stroke: a randomized controlled trial. Arch Phys Med Rehabil. 2023;104(2):188-194. CrossRef PubMed
12. Felius RAW, Geerars M, Bruijn SM, et al. Reliability of IMU-based gait assessment in clinical stroke rehabilitation. Sensors (Basel). 2022;22(3):908. CrossRef PubMed
13. Winstein CJ, Stein J, Arena R, et al. American Heart Association Stroke Council, Council on Cardiovascular and Stroke Nursing, Council on Clinical Cardiology, and Council on Quality of Care and Outcomes Research. Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2016;47(6):e98-e169. CrossRef PubMed
14. Khokale R, S Mathew G, Ahmed S, et al. Virtual and augmented reality in post-stroke rehabilitation: a narrative review. Cureus. 2023;15(4):e37559. CrossRef PubMed
15. Phan HL, Le TH, Lim JM, et al. Effectiveness of augmented reality in stroke rehabilitation: a meta-analysis. Appl Sci (Basel). 2022;12(4):1848. CrossRef
16. Noveletto F, Soares AV, Mello BA, et al. Biomedical serious game system for balance rehabilitation of hemiparetic stroke patients. IEEE Trans Neural Syst Rehabil Eng. 2018;26(11): 2179-2188. CrossRef PubMed
17. Mubin O, Alnajjar F, Al Mahmud A, et al. Exploring serious games for stroke rehabilitation: a scoping review. Disabil Rehabil Assist Technol. 2022;17(2):159-165. CrossRef PubMed
18. Doumas I, Everard G, Dehem S, et al. Serious games for upper limb rehabilitation after stroke: a meta-analysis. J Neuroeng Rehabil. 2021;18(1):100. CrossRef PubMed
19. Noveletto F, Soares AV, Eichinger FLF, et al. Biomedical serious game system for lower limb motor rehabilitation of hemiparetic stroke patients. IEEE Trans Neural Syst Rehabil Eng. 2020;28(6):1481-1487. CrossRef PubMed
20. Perrochon A, Borel B, Istrate D, et al. Exercise-based games interventions at home in individuals with a neurological disease: a systematic review and meta-analysis. Ann Phys Rehabil Med. 2019;62(5):366-378. CrossRef PubMed
21. Pereira VA, Eldebrando L, Silva HED, et al. Biomedical system to evaluate pulmonary function in patients with post-stroke hemiparesis. Fisioter Mov. 2020;33:e003338. CrossRef
22. Fugl-Meyer AR. Post-stroke hemiplegia assessment of physical properties. Scand J Rehabil Med Suppl. 1980;7:85-93. PubMed
23. Appelros P. Characteristics of mini-mental state examination 1 year after stroke. Acta Neurol Scand. 2005;112(2):88-92. CrossRef PubMed
24. Souza LA, Martins JC, Teixeira-Salmela LF, et al. Validity and reliability of the modified sphygmomanometer test to assess strength of the lower limbs and trunk muscles after stroke. J Rehabil Med. 2014;46(7):620-628. CrossRef PubMed
25. Miller W, Jeon S, Ye X. An examination of acute cross-over effects following unilateral low intensity concentric and eccentric exercise. Sports Med Health Sci. 2020;2(3):141-152. CrossRef PubMed
26. Cunha BP, Freitas SMSFD, Menezes VVDBD, et al. Ipsilesional upper limb performance in stroke individuals: relationship among outcomes of different tests used to assess hand function. Fisioter Mov. 2016;29(3):561-568. CrossRef
27. Billinger SA, Arena R, Bernhardt J, et al. American Heart Association Stroke Council; Council on Cardiovascular and Stroke Nursing; Council on Lifestyle and Cardiometabolic Health; Council on Epidemiology and Prevention; Council on Clinical Cardiology. Physical activity and exercise recommendations for stroke survivors: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2014;45(8):2532-2553. CrossRef PubMed
28. Veldema J, Jansen P. Resistance training in stroke rehabilitation: systematic review and meta-analysis. Clin Rehabil. 2020;34(9):1173-1197. CrossRef PubMed
29. Chen SC, Kang JH, Peng CW, et al. Adjustable parameters and the effectiveness of adjunct robot-assisted gait training in individuals with chronic stroke. Int J Environ Res Public Health. 2022;19(13):8186. CrossRef PubMed
30. Lodha N, Patel P, Casamento-Moran A, et al. Strength or motor control: what matters in high-functioning stroke? Front Neurol. 2019;9:1160. CrossRef PubMed
31. Verrienti G, Raccagni C, Lombardozzi G, et al. Motivation as a measurable outcome in stroke rehabilitation: a systematic review of the literature. Int J Environ Res Public Health. 2023;20(5):4187. CrossRef PubMed
32. Dorsch S, Ada L, Sorial T, et al. The relationship between strength of the affected leg and walking speed after stroke varies according to the level of walking disability: a systematic review. Phys Ther. 2021;101(12):pzab233. CrossRef PubMed
33. Harjpal P, Qureshi MI, Kovela RK, et al. Efficacy of bilateral lower-limb training over unilateral lower-limb training to reeducate balance and walking in post-stroke survivors: a randomized clinical trial. Cureus. 2022;14(10):e30748. CrossRef PubMed
34. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989;10(4):407-415. CrossRef PubMed
35. Lee MY, Sung KS, Ham H, et al. Knee extensor strength measurement in patients with limited physical activity using a supine dynamometer anchoring frame. Ann Rehabil Med. 2020;44(6):502-509. CrossRef PubMed
36. Veerbeek JM, van Wegen E, van Peppen R, et al. What is the evidence for physical therapy poststroke? A systematic review and meta-analysis. PLoS One. 2014;9(2):e87987. CrossRef PubMed
37. Viana RB, de Oliveira VN, Dankel SJ, et al. The effects of exergames on muscle strength: a systematic review and meta-analysis. Scand J Med Sci Sports. 2021;31(8):1592-1611. CrossRef PubMed
38. Park JS, Lee G, Choi JB, et al. Game-based hand resistance exercise versus traditional manual hand exercises for improving hand strength, motor function, and compliance in stroke patients: a multi-center randomized controlled study. NeuroRehabilitation. 2019;45(2):221-227. CrossRef PubMed
39. Tăut D, Pintea S, Roovers JWR, et al. Play seriously: effectiveness of serious games and their features in motor rehabilitation. A meta-analysis. NeuroRehabilitation. 2017;41(1):105-118. CrossRef PubMed
40. Hunnicutt JL, Gregory CM. Skeletal muscle changes following stroke: a systematic review and comparison to healthy individuals. Top Stroke Rehabil. 2017;24(6):463-471. CrossRef PubMed
41. Chang SH, Francisco GE, Zhou P, et al. Spasticity, weakness, force variability, and sustained spontaneous motor unit discharges of resting spastic-paretic biceps brachii muscles in chronic stroke. Muscle Nerve. 2013;48(1):85-92. CrossRef PubMed
42. Ocklenburg S, Gerding WM, Arning L, et al. Myelin genes and the corpus callosum: proteolipid protein 1 (PLP1) and contactin

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

1 (CNTN1) gene variation modulates interhemispheric integration. Mol Neurobiol. 2017;54(10):7908-7916. CrossRef PubMed
43. Stewart JC, Dewanjee P, Tran G, et al. Role of corpus callosum integrity in arm function differs based on motor severity after stroke. Neuroimage Clin. 2017;14:641-647. CrossRef PubMed
44. Li Y, Wu P, Liang F, et al. The microstructural status of the corpus callosum is associated with the degree of motor function and neurological deficit in stroke patients. PLoS One. 2015;10(4):e0122615. CrossRef PubMed
45. Dragert K, Zehr EP. High-intensity unilateral dorsiflexor resistance training results in bilateral neuromuscular plasticity after stroke. Exp Brain Res. 2013;225(1):93-104. CrossRef PubMed
46. Ghazavi Dozin SM, Mohammad Rahimi N, Aminzadeh R. Wii fit-based biofeedback rehabilitation among post-stroke patients: a systematic review and meta-analysis of randomized controlled trial. Biol Res Nurs. 2024;26(1):5-20. CrossRef PubMed
47. Spina S, Facciorusso S, D'Ascanio MC, et al. Sensor based assessment of turning during instrumented Timed Up and Go Test for quantifying mobility in chronic stroke patients. Eur J Phys Rehabil Med. 2023;59(1):6-13. CrossRef PubMed
48. Lenkulkul F, Franco G, Soares AV, et al. Jogo Sério para reabilitação de hemiparéticos por acidente vascular cerebral: impacto na qualidade de vida. RSD. 2022;11(14):e87111435976. CrossRef
49. Peng QC, Yin L, Cao Y. Effectiveness of virtual reality in the rehabilitation of motor function of patients with subacute stroke: a meta-analysis. Front Neurol. 2021;12:639535. CrossRef PubMed
50. Flansbjer UB, Holmbäck AM, Downham D, et al. Reliability of gait performance tests in men and women with hemiparesis after stroke. J Rehabil Med. 2005;37(2):75-82. CrossRef PubMed
51. Soares AV, Borges Júnior NG, Hounsell MS, et al. A serious game developed for physical rehabilitation of frail elderly. European Research in Telemedicine/La Recherche Européenne En Télémédecine. 2016;5:45-53. CrossRef
52. Van Hoornweder S, Vanderzande L, Bloemers E, et al. The effects of transcranial direct current stimulation on upper-limb function post-stroke: a meta-analysis of multiple-session studies. Clin Neurophysiol. 2021;132(8):1897-1918. CrossRef PubMed
53. Dromerick AW, Geed S, Barth J, et al. Critical Period After Stroke Study (CPASS): a phase II clinical trial testing an optimal time for motor recovery after stroke in humans. Proc Natl Acad Sci USA. 2021;118(39):e2026676118. CrossRef PubMed
54. Mojtabavi H, Shaka Z, Momtazmanesh S, et al. Circulating brain-derived neurotrophic factor as a potential biomarker in stroke: a systematic review and meta-analysis. J Transl Med. 2022;20(1):126. CrossRef PubMed
55. Murphy TH, Corbett D. Plasticity during stroke recovery: from synapse to behaviour. Nat Rev Neurosci. 2009;10(12):861-872. CrossRef PubMed

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Arch Physioter 2025; 15: 69-76
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3418
ORIGINAL RESEARCH ARTICLE

Entry-level physical therapist curricula in geriatric care: an Italian national survey study

Francesco Bucchi[1], Sara Dal Ben[2], Antonio Culcasi[3], Stefano Lecora[4], Roberto Tedeschi[5], Marco Da Roit[6]

[1]CIDAS Social Cooperative, Ferrara - Italy; Geriatric Physiotherapy Specialist Interest Group, Italian Physiotherapy Association
[2]Istituto AltaVita-IRA, Padua - Italy; School of Physiotherapy, University of Padua, Padua - Italy; Geriatric Physiotherapy Specialist Interest Group, Italian Physiotherapy Association
[3]Nursing, Technical and Rehabilitative Assistance Service, IRCCS Istituto Ortopedico Rizzoli, Bologna - Italy; School of Physiotherapy, Alma Mater Studiorum, University of Bologna, Bologna - Italy; Geriatric Physiotherapy Specialist Interest Group, Italian Physiotherapy Association
[4]Private Practitioner, Naples - Italy; Geriatric Physiotherapy Specialist Interest Group, Italian Physiotherapy Association
[5]Independent Researcher, Bologna - Italy
[6]Quality Service, AULSS1 Dolomiti, Belluno - Italy, AULSS1, Dolomiti - Italy; Geriatric Physiotherapy Specialist Interest Group, Italian Physiotherapy Association

ABSTRACT

Introduction: To address the health needs of the aging population, it is necessary to map entry-level curricula for health professionals. This survey investigated geriatric-related content in entry-level physical therapist (PT) curricula offered by Italian universities.

Methods: A cross-sectional observational study was conducted using the CROSS checklist. A 66 questions survey was developed and sent via e-mail to all entry-level PT program directors of the Italian universities.

Results: A total of 34 physical therapist undergraduate course directors out of 62 completed the survey, yielding a 54.8% response rate. These results highlight the need for greater emphasis on health promotion and prevention. Although essential competencies such as the promotion of an active lifestyle and fall prevention appear to be well covered, other aspects remain underrepresented. Relevant gaps were also noted in the care and rehabilitation of some common chronic conditions (e.g., constipation, depression, and diabetes), as well as in therapeutic education. Greater attention to these topics could help align training with the emerging needs of the healthcare system. However, 11.76% of Italian PT programs do not include specific modules or courses dedicated to geriatric rehabilitation. Overall, there is considerable variability in teaching hours, topic coverage, and depth.

Conclusions: This study provides meaningful insights for updating the current PT curricula in geriatric care. This survey could represent a tool for future longitudinal research on mapping curricula over time in response to the aging population.

Keywords: Geriatrics, Competence, Curricula, Entry-level education, Physical therapist

What is already known about this topic:

- *There is a paucity of studies mapping the content of health profession entry-level curricula in geriatric care, with none at European or Italian levels.*

What does the study add:

- *This survey maps entry-level physical therapy curricula in the geriatric field at all Italian universities, in-depth exploration of core geriatric competencies, teaching and assessment strategies, and course directors' perceptions.*

Received: December 4, 2024
Accepted: March 3, 2025
Published online: April 25, 2025

This article includes supplementary materials.

Corresponding author:
Marco Da Roit
email: marco.daroit@aulss1.veneto.it

Introduction

Globally, all regions are experiencing an aging population. Developed countries are expected to move to a more advanced stage of population aging, with the proportion of older individuals rising from 20% in 2023 to 28% by 2050 [1]. Given the growing population of frail and older individuals, there is a need to educate students and professionals in health and social care, giving them specific competencies necessary to assess, treat, and work with older people [2-4].

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

The International Association of Physiotherapists Working with Older People (IPTOP) published the "Standards of Clinical Practice" in 2021 [5]. This document provides guidance for physical therapists and represents an educational resource for students to learn about geriatric care. However, the document is not intended to establish educational standards or provide a framework for acceptable professional behavior or knowledge.

Previous studies have reported that integrating geriatric content into didactic coursework and clinical placement can enhance physical therapy students' competencies in geriatrics [6]. However, known barriers to improving geriatric training include limited time within packed curricula and the number of educators with geriatric expertise [7]. Moreover, ageism and the complexity of managing multiple comorbidities have been identified as factors that decrease the willingness and positive attitudes of physical therapy students toward working with older adults [6, 7]. These challenges can be mitigated by enhancing students' knowledge of aging and promoting contact with older adults during entry-level education [8]. To address this challenge, changes in physical therapists' practice and educational pathways are required [9].

Recognizing the need, the Academy of Geriatric Physical Therapy (American Physical Therapy Association, APTA) published the document "Essential Competencies in the Care of Older Adults at the Completion of the Entry-level Physical Therapist Professional Program of Study" in 2017 [10]. This work describes the competencies in the geriatrics field that students of physical therapy entry-level degree programs should acquire in the United States. Specifically, the document states the core skills required to provide competent care to older people across the following domains: (i) health promotion and safety; (ii) evaluation and assessment; (iii) care planning and coordination across the care spectrum (including end-of-life care); (iv) interdisciplinary and team-based care; (v) caregiver support; and (vi) healthcare systems and benefits. To our knowledge, there are few similar documents elsewhere, but none describe physical therapist competence in depth [2,11]. Nevertheless, such a document could serve as a reference point for similar projects to be contextualized in other nations.

At the national level, a document defining the core curriculum and competencies for entry-level physical therapists was published in 2012 [12]. However, the national framework lacks the field-specific focus required to address the evolving needs of the patient population.

To our knowledge, there is a paucity of studies mapping the content of health profession entry-level curricula in geriatric care. Few studies with different methodological approaches are available for medical doctors, nurses, occupational therapists, physical therapist assistants, and other education programs [13-17]. In particular, three surveys were conducted in the United States over a 14-year period and recently, a survey conducted in the Philippines assessed entry-level physical therapist curricula [18-21]. We are not aware of any studies that have examined the content of physical therapy degree courses in geriatric physical therapy in Europe or Italy.

This survey investigated geriatric-related content in entry-level physical therapy degree programs offered by Italian universities. The aim was to identify weaknesses or gaps in the current curriculum.

Methods

Study design

We conducted a cross-sectional observational study following the CROSS checklist for survey-based research design [22].

Survey development

Two investigators (AC and MDR) independently reviewed the literature in February 2023. No validated survey was found in the literature. Potential questions were derived from the relevant studies and then refined several times. Each interaction combined deductive reasoning and inductive insights. This process involved two investigators (FB and SDB) and led to a survey draft consisting of 63 questions. The preliminary survey was pre-tested by five geriatric physical therapists with teaching experience in entry-level courses and at least ten years of clinical expertise with older people in different settings (e.g., nursing homes, hospitals, and home-care physical therapy). During the pre-test, the relevance and clarity of the questions were assessed in two rounds. Face and content validity were tested during this phase. Specifically, seven questions were added, and four questions were deleted based on expert feedback, as they were deemed essential or irrelevant to the research. The final survey consisted of 66 questions (Supplementary Material 1). The structure of the survey was based on geriatric physical therapists' competencies as defined by the American Physical Therapy Association Section on Geriatrics [10]. The survey was divided into sections with a variable number of questions, as follows: (i) health promotion (seven questions); (ii) prevention (eight questions); (iii) geriatric rehabilitation-specific teaching (four questions) (iv) care and rehabilitation (30 questions); (v) therapeutic education (six questions); (vi) communication and interpersonal relation (six questions); (vii) perception on importance and adequacy of geriatric-related physical therapist (PT) curricula (two questions); and (viii) methodology of teaching and students' judging (three questions). For each section, the questions used a four-point scale to assess both the content covered in each PT program and the directors' perceptions of the importance and suitability of the content (strongly yes/more yes than no/more no than yes/strongly no). Yes/no and multiple-choice questions were used to explore the geriatric rehabilitation teaching methodology and students' knowledge and competence assessment.

Sample characteristics

We aimed to recruit entry-level PT program directors from all Italian universities. The eligible population was identified through https://www.universitaly.it, the official portal of the Italian Ministry of University and Research. The contact information (names, e-mails, and phone numbers) of entry-level PT program directors was collected from publicly accessible university websites. Forty-two universities were identified, and 62 participants were invited to participate in this study.

Survey administration

An investigator (SDB) sent an e-mail invitation with a link to the participant information sheet and a consent form to all entry-level PT program directors identified as suitable to participate in the study. After providing consent, the participants were asked to complete the online survey (administered via Microsoft Form) within three months. A follow-up e-mail was sent after a three-week period to encourage participation. After three e-mail reminders, three investigators (SDB, FB, and SL) phoned entry-level PT program directors to further promote participation in the study. Appropriate Microsoft Form settings were used by each participant to prevent multiple entries.

Ethical considerations

The University of Bologna Bioethics Committee approved this protocol (approval number: 0335961), and informed consent was obtained from all participants. Participants' anonymity was ensured because Microsoft Forms did not collect e-mail addresses. The study protocol was registered in the Open Science Framework on January 7, 2024, prior to data collection (accessible at https://doi.org/10.17605/OSF.IO/MVEJX).

Statistical analysis

Data collected using Microsoft Forms were exported to a Microsoft Excel spreadsheet. Descriptive data analysis was performed, and the Shapiro–Wilk test was used to assess the normality of continuous variable distributions. In the case of distribution symmetry, the variables were represented as means and standard deviations, whereas the median value and interquartile range were used in the case of distributional non-normality. Categorical variables are reported as absolute and percentage values. Comparisons between groups were performed using the χ^2 or Fisher's exact test for categorical variables. Statistical significance was set at p-value < 0.05. All analyses were conducted using the STATA 11 software (StataCorp, USA).

Results

As described in the methods section, the survey was sent to all the identified course directors. Over the planned timeframe, a total of 34 physical therapy undergraduate course directors out of 62 completed the survey, yielding an overall 54.8% response rate. Demographic data were not collected from respondents. Given the anonymous nature of the survey, it was not possible to investigate the reasons for non-responses. Nevertheless, a range of strategies was implemented to promote and enhance participation in subsequent surveys. The results are presented following the same scheme to maintain consistency with the survey structure. Statistical tests were initially planned for comparison between the groups. However, the data distribution prevented the application of these tests because the groups were too small or unbalanced to yield valid comparisons. Therefore, we opted to present only the descriptive statistics.

Health promotion

The responses to questions on whether the curriculum addresses knowledge and skills in health promotion are shown in Fig. 1. In this section, a large proportion of responses were registered as "no rather than yes" or "yes rather than no." The question about competence in the role of physical activity was the only one with no "strongly disagree" response and, at the same time, with the highest proportion of "strongly agree" responses (67%). All the questions have some "strongly disagree" responses, ranging from active lifestyle promotion (3%) to the quality of sleep (15%).

Prevention

Responses to the questions about whether the curriculum addresses prevention knowledge and skills are shown in Fig. 2. In this section, the questions about fall risk prevention and assessment tools were the only two with no "absolutely no" and had the highest rate of "absolutely yes" responses, 20 (58%) and 18 (52%), respectively. For all other questions, most responses fell into the two central categories, with a slight predominance of "more yes than no" responses.

Care and rehabilitation

Responses to questions on whether the curriculum addresses knowledge and skills in the care and rehabilitation areas are shown in Fig. 3. Analysis of the responses in this

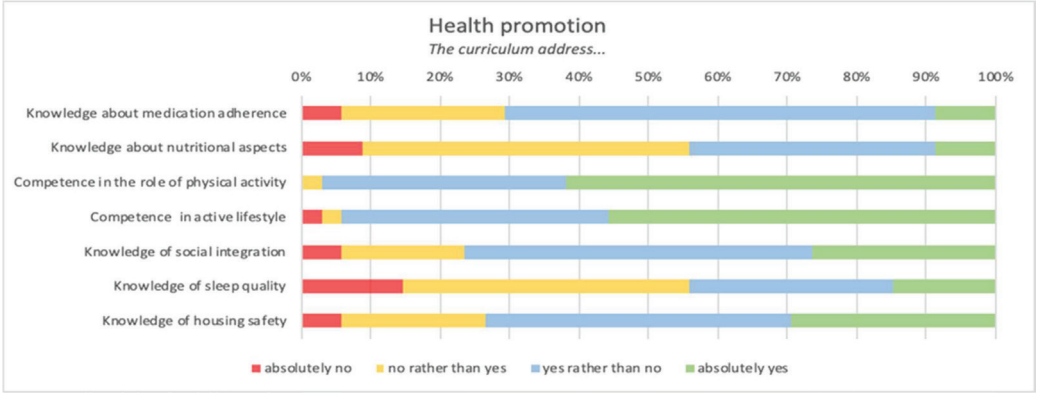


FIGURE 1 - Results of survey questions concerning health promotion.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

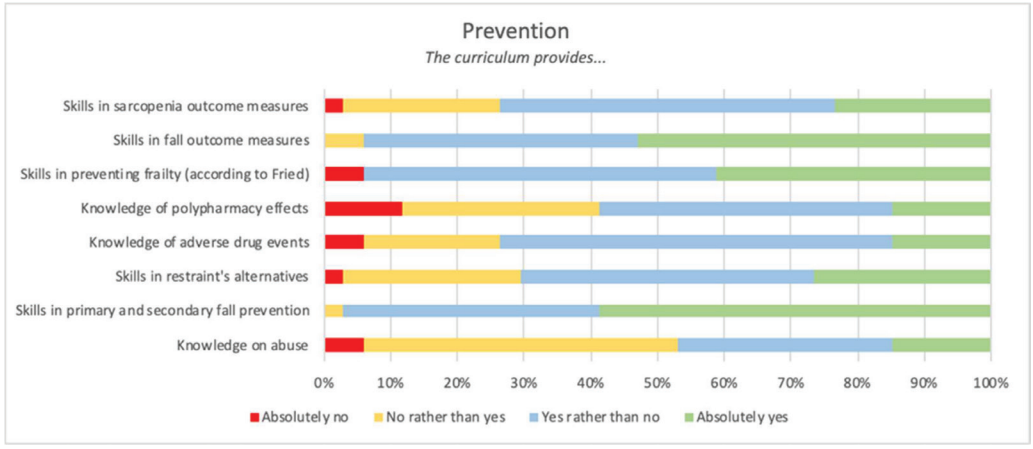


FIGURE 2 - Results of survey questions concerning prevention

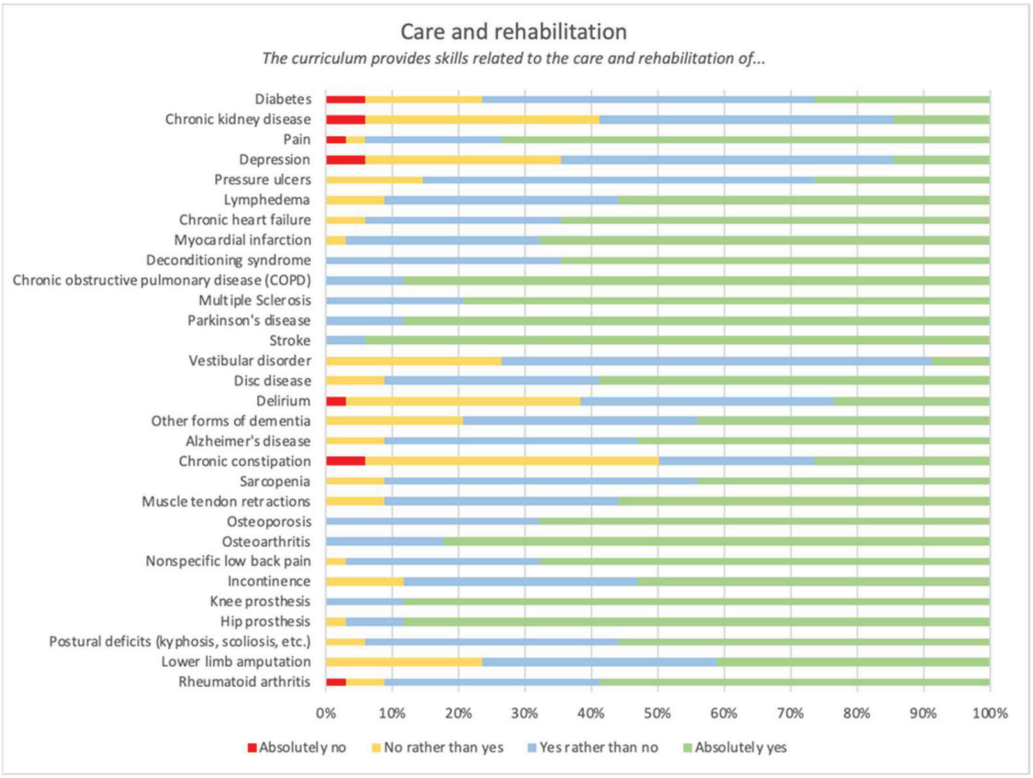


FIGURE 3 - Results of survey questions concerning care and rehabilitation

section revealed that the vast majority of questions did not yield any "absolutely no" answers. Only seven of the 30 questions elicited a small number of negative responses, namely, those pertaining to rheumatoid arthritis, chronic constipation, delirium, diabetes, depression, pain, and chronic kidney disease. In this area, the majority of registered responses were "yes rather than no" or "absolutely yes."

Therapeutic education

The responses to questions on whether the curriculum addresses skills in the area of therapeutic education are shown in Fig. 4. Here, we observed that three questions about home setting, red flags assessment, and aids appropriate use registered no "absolutely no" and, at the same time, all had more than 50% "absolutely yes" responses. The questions about the use of telemonitoring devices, adherence to medication therapy, and self-care programs registered some "absolutely no" responses (8%, 8%, and 2%, respectively), while the majority of responses fell into the two central categories, with a slight predominance of "more yes than no" responses.

Communication and interpersonal relation

Responses to questions on whether the curriculum addresses communication skills are shown in Fig. 5. In this area, we observe that "absolutely yes" and "rather yes than

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

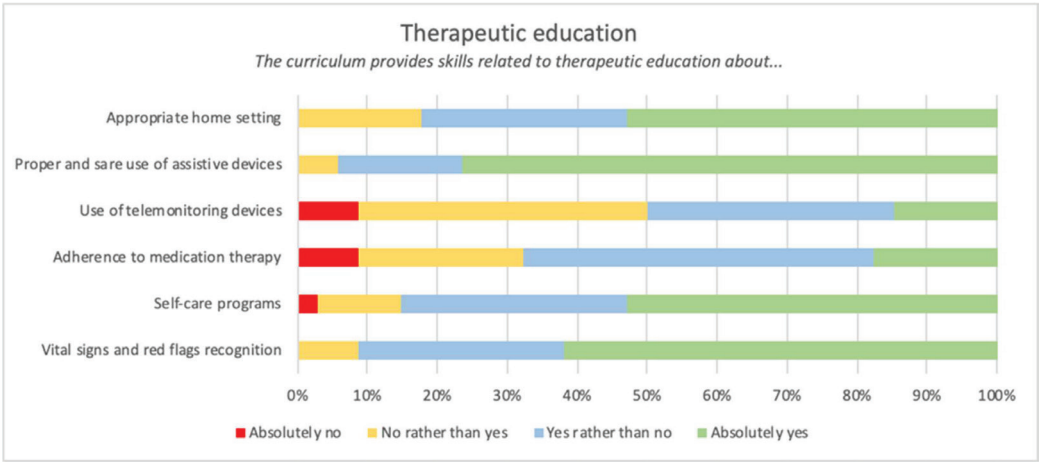


FIGURE 4 - Results of survey questions concerning therapeutic education

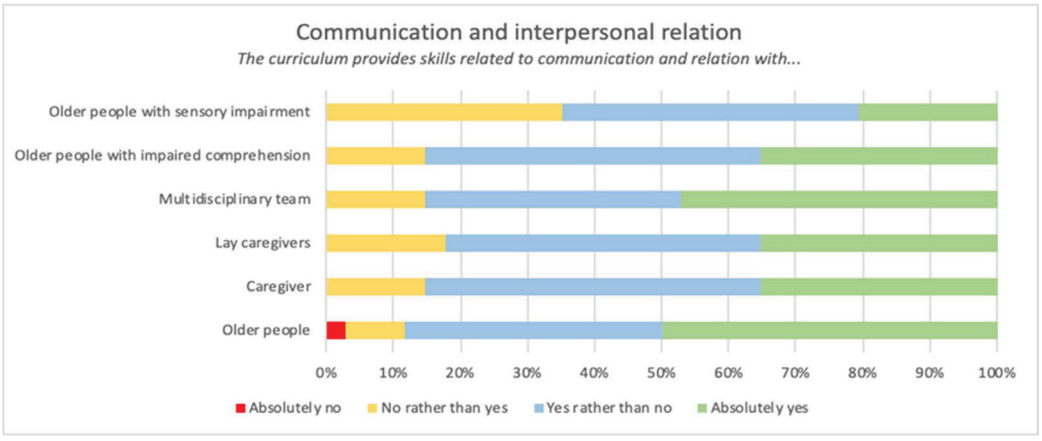


FIGURE 5 - Results of survey questions concerning communication and interpersonal relation

no" responses were the most frequent of all the questions. Only the question about skill in communication with older people registered some "absolutely no" responses (2%), and those about communication with older people with sensory deficits registered 35% of "rather no than yes" responses.

Course directors' perceptions

Our survey also assessed course directors' perceptions of the degree to which their course curriculum provides adequate competencies for managing older people and whether these competencies are developed during basic training. We observed that positive responses constituted the vast majority, with "rather yes than no" (61.76%) and "completely yes" (20.59%). Slightly negative responses, "rather no than yes," were observed in 17.75% of directors' responses, while no completely negative responses were recorded. The importance of providing these competencies in the basic training is indicated by almost all the academic program directors, with "absolutely yes" and "rather yes than no" responses respectively representing 70.59% and 26.47%, respectively.

Specific course description

With reference to the sample size that responded to the survey, 30 (88.24%) curricula included a specific module/course on geriatric rehabilitation, while in the remaining four (11.76%), such a course was not present. Regarding the 30 curricula with a specific course on geriatric rehabilitation, data were gathered about the extent of the teaching included in the curriculum, the professional profile of the teacher in charge of this course, and the teaching and evaluation methodologies used. The geriatric rehabilitation teaching module has, on average, 3.43 [(standard deviation (SD) 2.14] ECTS credits, with a range recorded between 1 and 8 ECTS. This corresponds to an average of 31.87 (SD 22.58) hours of teaching activity. The teacher was a physical therapist in all cases; in 40% of cases, they were the only professionals, while in the remaining 60%, they were joined by medical figures of different specialties (geriatrician, internist, physiatrist, neurologist, or orthopedist) and with different combinations. The most frequent combination was a physical therapist plus a geriatrician, which was recorded in 33% of the responses. The teaching and evaluation methodologies used in the courses are represented in Fig. 6. This is indicated by how many courses a specific methodology has been used, both individually and in combination with others.

In order to develop practical skills, we investigated the number of internship hours required to build competencies in this population. The aggregated median of 150 [interquartile range (IQR) 102.5] reflected responses from 28 curricula,

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

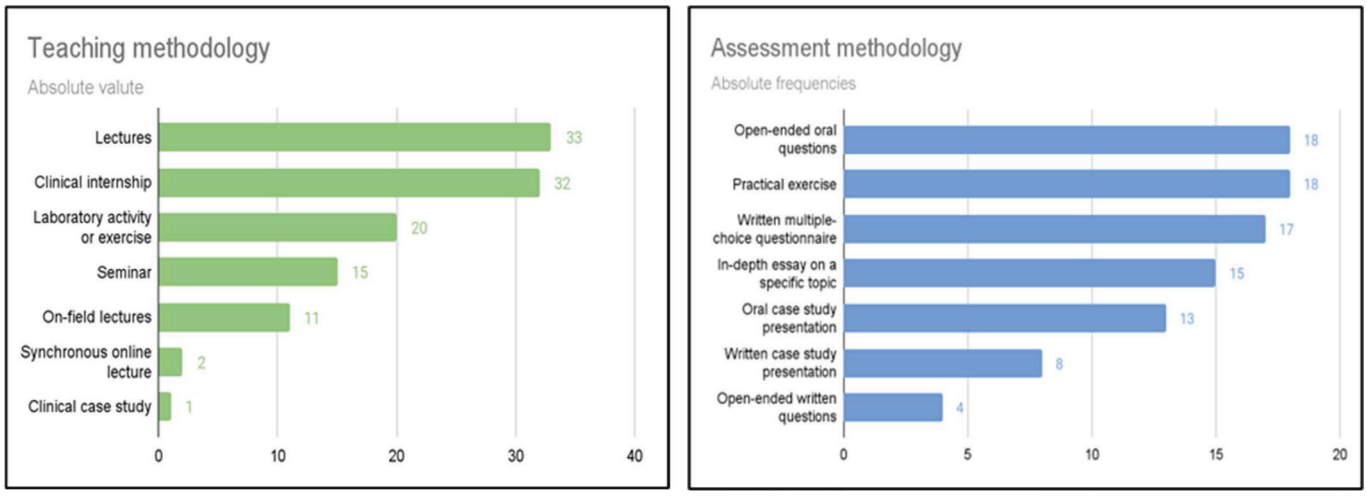


FIGURE 6 - Teaching and assessment methodology absolute frequencies

reported as the median and IQR due to the non-normal distribution of the data. For the remaining six curricula, one indicated 30% of the total, but the total number of hours was not specified, while for the other five, the internship hour data were missing.

Discussion

The findings of this study offer compelling insights into the integration of geriatric-related content into the curricula of physical therapy degree programs offered by Italian universities. Our survey achieved an overall response rate of 54.8%, which was comparable to or slightly higher than those reported in previous studies [17, 20].

Although 88.24% of the curricula included a specific course on geriatric rehabilitation, there was considerable variability in teaching hours, topic coverage, and depth. This variability may reflect insufficient standardization of geriatric training, an issue previously highlighted by Granick [18] and Wong [20].

Regarding the competencies defined by the Academy of Geriatric Physical Therapy [American Physical Therapy Association (APTA)] [10], our survey indicates that fall prevention is a particularly strong area within the curricula, with questions on this topic yielding highly positive responses. This result aligns with prior studies [20] and reflects the growing global recognition of the importance of fall prevention in older adults, especially considering that falls are one of the leading causes of morbidity and mortality in this age group, as emphasized by Bardach and Rowles [7].

Increasing awareness was also noted in the areas of care and rehabilitation, and our study demonstrated that the most relevant content was entry-level training, with some exceptions, such as rheumatoid arthritis, constipation, delirium, and diabetes. These gaps suggest potential deficiencies in the integration of training related to the management of common chronic conditions among older adults. Perez identified similar gaps in a study conducted in the Philippines, where geriatric-specific training was found to be insufficient, particularly in chronic disease management [21].

Other studies have also highlighted a lack of focus on multimorbidity and the complexity of conditions commonly observed in older adults [18, 20]. Similar deficiencies have been reported in medical education [17], indicating inadequate preparation for managing complex comorbidities and a lack of multidisciplinary skills required of older people. A revision or update of the current curricula should consider these elements to better address the health needs of the aging population [4, 7].

Another interesting aspect was the presence of dementia. Our study seems to indicate that dementia education, while present in physical therapist curricula, is not yet sufficiently integrated despite the growing global prevalence of dementia and its impact on motor and cognitive functions. Our findings align with previous reports from Ireland, the United Kingdom, and New Zealand, which revealed that while fundamental dementia management concepts are included, training remains partial and insufficient [23]. The same issues were raised by Quick in a study conducted in Australia and Canada, in which training primarily focused on the management of physical symptoms, neglecting an integrated approach that addressed the cognitive and behavioral challenges associated with dementia [24]. Both studies highlighted difficulties such as the lack of adequate educational resources, the need for greater involvement of dementia experts in training programs, and the need for a cultural shift in physical therapist schools to recognize dementia as a crucial competence for physical therapists [23, 24].

Greater emphasis should also be placed on health promotion, as our study showed that the inclusion of competencies related to health promotion within the core training of physical therapists still has significant room for improvement. These findings are consistent with those of other studies [18, 20, 21], which underscore the importance of this area but acknowledge that it remains underrepresented in educational programs. In particular, the promotion of an active lifestyle and the key role of physical activity in managing older adults can be reinforced, whereas other important aspects, such as sleep management and nutritional considerations, remain underrepresented.

The results of our study also show that therapeutic education seems to be satisfactorily covered in physical therapy curricula, with a majority of positive responses from the course directors. Nevertheless, it is interesting to note that some areas, such as pharmacological therapy and the use of telemonitoring devices, still present challenges. This latter aspect may indicate that, despite the growing importance of these technologies and approaches in modern healthcare, they have not yet been thoroughly addressed in all degree programs. Paying greater attention to new technologies and innovative health management approaches could help align the training more closely with the emerging needs of the healthcare system and better prepare future physical therapists to address clinical challenges.

Finally, our study suggests that communication and interpersonal relation skills seem to be adequately addressed in physical therapy curricula, with most course directors responding with either "absolutely yes" or "more yes than no." However, communication with older adults with sensory deficits elicited a higher percentage of negative responses. This finding suggests a potential gap in addressing the specific communication challenges that arise when working with frail older adults.

Despite this, there has been significant improvement over time in this area. In 2001, Wong noted that communication and interpersonal skills were not always well integrated or emphasized in the curricula [20].

The majority of course directors perceive their programs as well-equipped to provide the necessary competencies for managing older adults. This is an encouraging finding, as it reflects a growing awareness of the importance of training physical therapists with specific skills in the care of older people, in line with the increasing demand for healthcare services in this population. However, the fact that 60% of responses fall under the "more yes than no" category suggests there is room for improvement in fully meeting the health needs of this population.

A particularly concerning aspect revealed by our study is that, despite the growing importance of geriatrics in physical therapy, 11.76% of Italian physical therapy degree programs do not include any specific module or course dedicated to geriatric rehabilitation. This finding is consistent with previous international studies [18, 20, 21] that have shown that many physical therapy programs lack geriatric-specific content.

Another important issue is the time dedicated to training. In our study, an average of 31.87 (SD 22.58) hours were dedicated to teaching activities. Even if there is still room for improvement, this is an encouraging improvement compared with studies from the 1980s, which reported an average of 10-16 hours [13, 17, 18].

We also investigated the development of practical skills. The allocation of internship hours specific to working with the older population varied across the programs, with an average of 150 hours. However, in six programs, data on internship hours were either partial or missing, raising concerns about whether students were given sufficient opportunities to acquire practical experience in a geriatric setting. As noted in previous studies, direct interactions with older patients during training can reduce ageism and improve students' attitudes toward geriatric care [7, 8].

These findings highlight the need to expand teaching strategies, as previously suggested by Wong [20], by incorporating more clinical aspects and a deeper understanding of the multifactorial conditions associated with aging. Interprofessional collaboration should be encouraged, and instructors who practice in geriatric-specific settings should be identified. In our study, only 33% of the teaching teams for the module comprised both a geriatrician and a physical therapist.

Limitations

Our findings should be considered while acknowledging several limitations. Despite the various strategies implemented during the construction process, the instrument remains unvalidated. Furthermore, although our survey achieved a response rate of 54.8%, we could not assess whether the data were influenced by the responses offered by course directors who were more interested in this topic. Therefore, caution should be exercised when generalizing these results.

Conclusion

Future studies should consider conducting a Europe-wide survey to compare the geriatric physical therapy curricula across different countries. In addition, longitudinal research could help track the evolution of curricula over time, particularly in response to the aging population and increasing need for geriatric-specific competencies. It would also be valuable to explore the barriers perceived by instructors and course directors in creating and integrating specific geriatric modules, as well as to investigate students' perceptions of geriatrics and how these influence their motivation to work with older patients. Moreover, examining students' practical experiences during internships with older adults could help determine whether the acquired skills align with the actual needs of the clinical setting.

This survey could provide crucial insights for developing more targeted and effective educational strategies that address not only educational needs but also the perceptions and motivations of those training and working with the older population.

Acknowledgments

The authors acknowledge Dr. Foggetti AM, Dr. Martinotta S and the Geriatric Physiotherapy Special Interest Group of Italian Physiotherapy Association (Gruppo di Interesse Specialistico in Fisioterapia nell'Anziano e nell'Invecchiamento Attivo) for support.

Disclosures

Conflict of interest: The authors declare no conflict of interest.

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

The Clinical Trial Protocol number is not applicable, but the study protocol was registered on Open Science Framework on January 7, 2024, prior to the beginning of data collection CrossRef

References

1. Department of Economic and Social Affairs. *World Population Ageing 2023: Challenges and Opportunities of Population Ageing in the Least Developed Countries.* United Nations Research Institute for Social Development; 2024. Online
2. Dijkman B, Roodbol PF, Aho J, et al. European Core Competences Framework for Health and Social Care Professionals Working with Older People. 2016. Online
3. World Health Organization. Long-term care for older people: package for universal health coverage. Online (Accessed December 2024)
4. Wong R, Odom CJ, Barr JO. Building the physical therapy workforce for an aging America. *J Phys Ther Educ*. 2014;28(2):12-21. CrossRef
5. The International Association for Physiotherapists Working with Older People (IPTOP). 2021. Standards of Clinical Practice: Revised Edition. 2021. Online (Accessed December 2024)
6. Blackwood J, Sweet C. The influence of ageism, experience, and relationships with older adults on physical therapy students' perception of geriatrics. *Gerontol Geriatr Educ*. 2017;38(2):219-231. CrossRef PubMed
7. Bardach SH, Rowles GD. Geriatric education in the health professions: are we making progress? *Gerontologist*. 2012;52(5):607-618. CrossRef PubMed
8. Almarwani M. Does ageism exist among multidisciplinary rehabilitation students? Predictors of attitudes toward older adults. *J Multidiscip Healthc*. 2023;16:181-189. CrossRef PubMed
9. Guccione AA, Frost J, Barr JO. Forecasting health care delivery for older adults in the midst of change: challenges and opportunities for the physical therapy profession in an evolving environment. *J Phys Ther Educ*. 2014;28(2):7.
10. Geriatrics Section APTA. Essential Competencies in the Care of Older Adults at the Completion of the Entry-level Physical Therapist Professional Program of Study. 2017. Online (Accessed December 2024)
11. Eleazer P, McRae T, Knebl J, et al. The Education Committee Writing Group of the American Geriatrics Society. Core competencies for the care of older patients: recommendations of the American Geriatrics Society. *Acad Med*. 2000;75(3):252-255. PubMed
12. Bielli S, Bozzolan M, Cortini S, et al. *La formazione "core" del fisioterapista.* Scienza Riabilitativa; 2011:13.
13. Liddle J, Beattie E, Gannon B, Bennett S, Pachana NA. Ageing as part of the curriculum for healthcare professionals: A Queensland Survey. *Australas J Ageing*. 2020;39(3):e447-e453. CrossRef PubMed
14. Strasburg DM, Gingher MC. A review of entry level education in gerontology. *Am J Occup Ther*. 1986;40(8):557-560. CrossRef PubMed
15. Wedge F, Mendoza M, Reft J. Development of geriatric curricular content within a physical therapist assistant education program. *Journal of Physical Therapy Education*. 2014;28(2). CrossRef
16. Ironside PM, Tagliareni ME, McLaughlin B, et al. Fostering geriatrics in associate degree nursing education: an assessment of current curricula and clinical experiences. *J Nurs Educ*. 2010;49(5):246-252. CrossRef PubMed
17. Eleazer GP, Doshi R, Wieland D, et al. Geriatric content in medical school curricula: results of a national survey. *J Am Geriatr Soc*. 2005;53(1):136-140. CrossRef PubMed
18. Granick R, Simson S, Wilson LB. Survey of curriculum content related to geriatrics in physical therapy education programs. *Phys Ther*. 1987;67(2):234-237. CrossRef PubMed
19. Solon JA, Kilpatrick NS. Aging-related education in entry-level physical therapy curricula: national benchmark data. *J Phys Ther Educ*. 1989;3(1):3.
20. Wong R, Stayeas C, Eury J, et al. Geriatric content in physical therapist education programs in the United States. *J Phys Ther Educ*. 2001;15(2):4.
21. Perez CDA. Integration of geriatric content in entry-level physical therapy education in the Philippines: a pilot study. *The Philippine Journal of Physical Therapy*. 2022;1(2). Accessed October 28, 2024. Online
22. Sharma A, Minh Duc NT, Luu Lam Thang T, et al. A Consensus-Based Checklist for Reporting of Survey Studies (CROSS). *J Gen Intern Med*. 2021;36(10):3179-3187. CrossRef PubMed
23. O'Sullivan T, Foley T, Timmons S, et al. Dementia content and delivery in physiotherapy curricula: an international study of entry level physiotherapy programmes in Ireland the United Kingdom and New Zealand. *Physiotherapy*. 2024;125:101410. CrossRef PubMed
24. Quick SM, Snowdon DA, Lawler K, et al. Dementia education for physiotherapy students: a questionnaire of australian and canadian entry-to-professional practice programs. *Physiother Can*. 2024;e20230100. CrossRef

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Arch Physioter 2025; 15: 77-89
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3233
REVIEW

Effectiveness of physical therapy interventions for coccydynia: a systematic review with a narrative synthesis

Mohammad Sidiq[1], Hariharasudhan Ravichandran[2,3], Balamurugan Janakiraman[4], Aksh Chahal[1], Richa Hirendra Rai[5], Ahmed Hamoud Alotaibi[6], Ali Hamoud Alotaibi[7], Abdullah Saad Alotaibi[8], Aminu Alhassan Ibrahim[9], Eid Abed Alharbi[7], Faizan Zaffar Kashoo[10], Hemlata Vats[5]

[1]Department of Physiotherapy, School of Allied Health Sciences, Galgotias University, Greater Noida, Uttar Pradesh, India
[2]Alvas College of Physiotherapy and Research Centre, Moodbidri, Dakshina Kannada, Karnataka, India
[3]Faculty of Physiotherapy, Meenakshi Academy of Higher Education and Research (MAHER), Tamil Nadu, India
[4]SRM College of Physiotherapy, Faculty of Medicine and Health Sciences, SRM Institute of Science and Technology (SRMIST), Kattankulathur, Tamil Nadu, India
[5]School of Physiotherapy, Delhi Pharmaceutical Sciences and Research University, New Delhi, India
[6]Department of Anesthesia, Northern Area Armed Forces Hospital, Hafar Al Batin, Saudi Arabia
[7]Department of Nursing, Northern Area Armed Forces Hospital, Hafar Al Batin, Saudi Arabia
[8]Department of Laboratory Sciences, Northern Area Armed Forces Hospital, Hafar Al Batin, Saudi Arabia
[9]Physiotherapy Department, Tishk International University, Erbil, Iraq
[10]Department of Physical Therapy and Health Rehabilitation, College of Applied Medical Sciences, Majmaah University, Al Majma'ah, Saudi Arabia

ABSTRACT

Introduction: Various physical therapy interventions for coccydynia have been evaluated, but their effectiveness has not yet been comprehensively synthesized. This systematic review aims to evaluate the effectiveness of physical therapy interventions in adults with coccydynia.

Methods: A systematic search of relevant randomized controlled trials (RCTs) was conducted in PubMed/MEDLINE, EMBASE, CINAHL, Scopus, Web of Science, Cochrane Central Register of Controlled Trials (CENTRAL), and Physiotherapy Evidence Database (PEDro). Outcomes of interest included pain, function, mobility, and patient satisfaction. Due to the heterogeneity of the included studies, a narrative synthesis was performed.

Results: A total of 515 adults with coccydynia across 10 studies were included in the review. Physical therapy interventions, including extracorporeal shock wave therapy, kinesiotaping plus exercise, levator anus stretching or massage, manipulation alone or manipulation plus electrotherapy or exercise, and muscle energy technique, showed significant improvements in pain and function in the short term. Additionally, kinesiotaping plus exercise showed significant short-term improvement in trunk mobility. In the intermediate term, manipulation alone and levator anus stretching or massage were effective at reducing pain, whereas manipulation alone was effective at improving function. In the long term, levator anus stretching or massage showed sustained improvement in pain.

Conclusions: Overall, physical therapy interventions led to short-term improvements in pain and function for adults with coccydynia. However, there is a need for high-quality studies with long-term follow-ups to compare the efficacy of various physical therapy interventions, both in isolation and in combination.

Keywords: Coccydynia, Function, Pain, Physical therapy interventions, Review, Randomized controlled trials

What's already known about this topic?

- *Coccydynia is a rare but painful and debilitating condition with uncertain pain sources and a lack of clinical guidelines.*
- *Although previous studies on physical therapy interventions show promising results, their effectiveness has not yet been comprehensively synthesized.*

What does this study add?

- *This is the first review to comprehensively synthesize the effectiveness of physical therapy interventions for adults with coccydynia.*
- *Of the ten studies evaluated, extracorporeal shock wave therapy, kinesiotaping, manipulation, massage, muscle energy technique, and stretching showed short-term improvements in pain, function, and trunk mobility.*

Received: July 30, 2024
Accepted: March 24, 2025
Published online: April 25, 2025

Corresponding author:
Aminu Alhassan Ibrahim
email: amenconafs@gmail.com

Introduction

Pain is a debilitating and distressing sensation that significantly impairs an individual's function and quality of life. Coccydynia, commonly known as tailbone pain, is a relatively rare medical condition characterized by discomfort or pain in

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

the coccyx region and is generally reported to be more prevalent in women (1,2). The pain often impacts the emotional well-being of patients, as it typically worsens during activities such as prolonged sitting, leaning back while seated, extended standing, and transitioning from sitting to standing (3). Primary coccydynia arises without any identifiable underlying cause, while secondary coccydynia may result from single-axis trauma, weight gain, change in postural biomechanics, childbirth, rapid weight-loss related to gastric by-pass surgery, degenerative spine disorders, infections, or previous surgeries leading to tissue inflammation and adhesions (2,4-10).

The coccyx, located at the lower end of the spine, forms a tripod structure with the ischial tuberosities. Muscle attachments on the coccyx help distribute weight evenly while sitting and provide structural support to surrounding areas. The shape of the coccyx may also be linked to the development of primary coccydynia (11). Moreover, individuals with a higher body mass index (BMI) are three times more likely to develop coccydynia compared to those with a normal BMI (10). Although most acute cases of coccydynia typically resolve within weeks to months with or without conservative treatment (3), some individuals may go on to experience persistent pain and develop chronic coccydynia, leading to debilitating symptoms that significantly interfere with daily life, with activities such as sitting, defecation, and sexual intercourse commonly being affected (10).

Effective management of coccydynia is necessary to avoid invasive procedures such as prolotherapy, radiofrequency ablations, and partial or complete coccygectomy. These procedures may have long-term consequences or result in further disability (12-17). Unfortunately, no clinical treatment guidelines for coccydynia are currently available to guide clinicians. As a result, conservative management options are considered first-line treatments and are generally preferred over surgical or minimally invasive procedures (3,18). Noninvasive methods include the use of specialized cushions, rest, non-steroidal anti-inflammatory drugs, and various physical therapy modalities (e.g., manipulation, soft-tissue techniques, electrotherapy, exercise, and patient education) to alleviate pain, prevent repetitive trauma, and restore near-normal biomechanics (19,20). Physical therapy modalities are tailored based on the patient's pain presentation, history, medical condition, anatomical involvement, and physiological findings, and in collaboration with functional assessments (20).

Studies evaluating physical therapy interventions for coccydynia have demonstrated promising results (21-33). A previous systematic review included 64 studies, of which only 5 were randomized controlled trials (RCTs), and concluded that non-surgical interventions are effective in providing pain relief for coccydynia (34). However, the review included minimally invasive procedures such as injection and ganglion block procedures, which are not considered physical therapy interventions and did not report functional outcomes. Additionally, the review only included studies published up to January 2020 (32). Given that newer trials (22, 28, 29, 35) have been published, and there has been a lack of a systematic review evaluating the effectiveness of these interventions, this is a key evidence gap for future reviews. Therefore, this systematic review aims to evaluate the effectiveness of physical therapy interventions for adults with coccydynia based on data from available RCTs.

Methods

Protocol registration

The protocol for this review was registered with the International Prospective Register of Systematic Reviews (PROSPERO), with registration number CRD42022344003. The Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (36) were followed for reporting this review. The PRISMA checklist is provided as Additional File 1.

Search strategy

A systematic literature search was conducted in the following databases: PubMed/MEDLINE, EMBASE, CINAHL, Scopus, Web of Science, Cochrane Central Register of Controlled Trials (CENTRAL), and Physiotherapy Evidence Database (PEDro) from their inception to March 31, 2024. Additionally, grey literature using Google was searched for potentially relevant studies. Citation chaining was employed for the selected studies to identify additional relevant studies using forward and backward citation tracking in PubMed/MEDINE, Scopus, and Web of Science. The search included key terms and synonyms, including both plural and singular forms identified through MeSH databases, combined using Boolean operators ("AND" and "OR") (Additional File 2).

Scientific and grey literature searches were further supplemented by manual searches of relevant review articles and reference lists from all included articles. Two reviewers (HR and BJ) independently performed the search, and any disagreements were resolved through discussion with the first author (MS). After removing duplicate articles, two reviewers (MS and AC) independently screened titles and abstracts to identify potentially relevant studies. Disagreements were resolved through discussion and, when necessary, with the involvement of a third reviewer (RHA and AHA Jr). Two additional reviewers (AHA Sr and ASA) independently evaluated the full texts for potential inclusion based on the review's inclusion and exclusion criteria.

Inclusion and exclusion criteria

The inclusion criteria were: (1) studies published in English with accessible full text, (2) RCTs or quasi-randomized trials involving male or female participants aged ≥18 years, (3) studies evaluating physical therapy interventions (e.g., manipulation, electrotherapy, exercise, massage, ergonomics/education, kinesiotaping, etc.) for patients diagnosed with coccydynia lasting at least 2 months, (4) studies evaluating pain, function/disability, mobility, or patient satisfaction as outcomes, and (5) studies comparing physical therapy interventions against each other or to comparator interventions such as sham, placebo, advice only, or usual care. The exclusion criteria were: (1) studies involving participants with coccydynia and other conditions such as pregnancy

or post-operative surgery, (2) scoping reviews, systematic reviews, case reports, opinion pieces, or case studies, and (3) studies on the effect of topical applications.

Data extraction

Data extraction was conducted independently by two reviewers (AHA Sr and ASA) in duplicate and compared after completion. A custom Excel Spreadsheet was designed to collect information about study details (author, publication year, setting/country), study design (e.g., RCT, quasi-randomized trial), participants (e.g., sample size, mean age), intervention (e.g., manipulation, exercise, electrotherapy), comparisons (e.g., sham, placebo, no-interventions, or other interventions), outcomes (e.g., pain intensity, disability/function, trunk mobility, patient satisfaction).

Risk of bias assessment

Two independent reviewers (AC and RHR) evaluated the risk of bias in the included studies. Any disagreements in grading were resolved through discussion, and if necessary, a third reviewer (HV) was consulted. The Cochrane Risk of Bias (RoB 2.0) tool (37) was used to assess the risk of bias in the included studies. The assessment was conducted for each outcome in the included studies (i.e., assessment at outcome level). Studies were evaluated across five domains of bias: 1) bias related to the randomization process, 2) bias arising from deviations in the intended intervention, 3) bias due to missing outcome data, 4) bias in the measurement of outcome, and 5) bias in the selection of the reported result. Each domain was categorized as low risk, some concerns, or high risk. The overall risk of bias was interpreted as follows:

(a) Low risk of bias: when all domains were judged to have a low risk of bias.
(b) Some concerns: when at least one domain raised some concerns but no domain was judged to be at high risk of bias.
(c) High risk of bias: when at least one domain was rated as high risk of bias, or multiple domains were judged to have some concerns that substantially lowered the confidence in the result.

Data synthesis

Initially, we planned to conduct a meta-analysis; however, this was deemed impractical due to heterogeneity in the interventions across studies. Although a few studies evaluated manipulation, variations in the approaches also precluded a meta-analysis. Therefore, a narrative synthesis of the studies was conducted based on the interventions and outcomes of interest. For available individual study data, treatment effects were extracted as mean differences (MD) and associated 95% confidence intervals (CI) using RevMan 5.4 software. As one study (25) reported median and range, the MD and 95% CI were calculated by estimating the mean and standard deviation (SD), as recommended by Wan et al. (38) as follows: Median = mean, SD = (b – a)/c, where b – a is the range, and c is 2 if the sample size (n) < 15, 4 if 15 ≤ n < 70, or n ≥ 70.

Results

Study selection

A total of 354 records were retrieved from the literature search, and after the removal of duplicates, the title and abstract of 342 studies were screened for potential eligibility (Fig. 1). Full texts of 231 studies were screened, of which 10 studies were deemed eligible (21-29,35) and finally included in the narrative synthesis based on the study stipulated criteria. The PRIMSA flow diagram, as depicted in Fig. 1, provides a summary of the study selection process.

Study characteristics

The included studies comprised a total of 563 participants, with a mean age ranging from 31 to 45 years. The intervention period spanned 1 to 4 weeks, and follow-up durations varied from 1 month to 2 years. The studies were conducted in several countries, including France (25,26), Turkey (28,35), Taiwan (24), Egypt (21), India (23,27), and Pakistan (22,29) (Table 1). All studies were hospital-based.

Regarding the outcome measures (Table 2), the pain was assessed using the following scales: seven studies used the Visual Analog Scale (VAS) (21,23-26,28,35), two studies used the Numerical Pain Rating Scale (NPRS) (22,29), one study used pain pressure threshold (PPT) (27), and another study used the (modified) McGill Pain Questionnaire (MPQ) (25). For functional outcome, two studies applied the Paris (functional coccydynia impact) Questionnaire (PQ) (25,28), four studies used the Oswestry Disability Index (ODI) (21,24,28,35), three studies used the Dallas Pain Questionnaire (DPQ) (22,25,29), and four studies evaluated pain-free sitting duration (22,23,27,29) as a functional measure. Mobility was assessed in one study using the Modified Schober Test (MMST) (21). Additionally, one study measured patient satisfaction using a 5-level self-satisfaction scale (24).

Interventions and comparators employed across the studies varied significantly and included the following: massage (26), mobilization (27), stretching (22,25,27-29), kinesiotaping (21), intrarectal manipulation (25,28), short wave diathermy (SWD) (25), coccygeal manipulation (23), conventional therapy (seat cushioning plus sitz bath plus phonophorosis) (25), transcutaneous electrical nerve stimulation (TENS) (23), reverse Kegel exercise (28), extracorporeal shock wave therapy (ESWT) (24,35), muscle energy technique (MET) (29), primal reflex release technique (PRRT) (22), and phonophoresis (21). For ease of synthesis, the interventions were grouped into six main categories (Table 2). All heating and electrical modalities used as comparators were collectively termed “electrotherapy.”

Manipulation [three studies (23,25,28)]: These studies employed a similar manipulation technique as the primary intervention but differed in their comparators. One (23) of the studies compared coccygeal manipulation plus electrotherapy (phonophoresis, TENS) against the same electrotherapy using VAS and PFSD in patients with clinical diagnosis of idiopathic coccydynia. Similarly, the second study (25) compared intrarectal manipulation plus electrotherapy (i.e., SWD) and electrotherapy alone using VAS, (modified) DPQ, (modified) MPQ, and PQ in patients with chronic coccydynia. The third

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

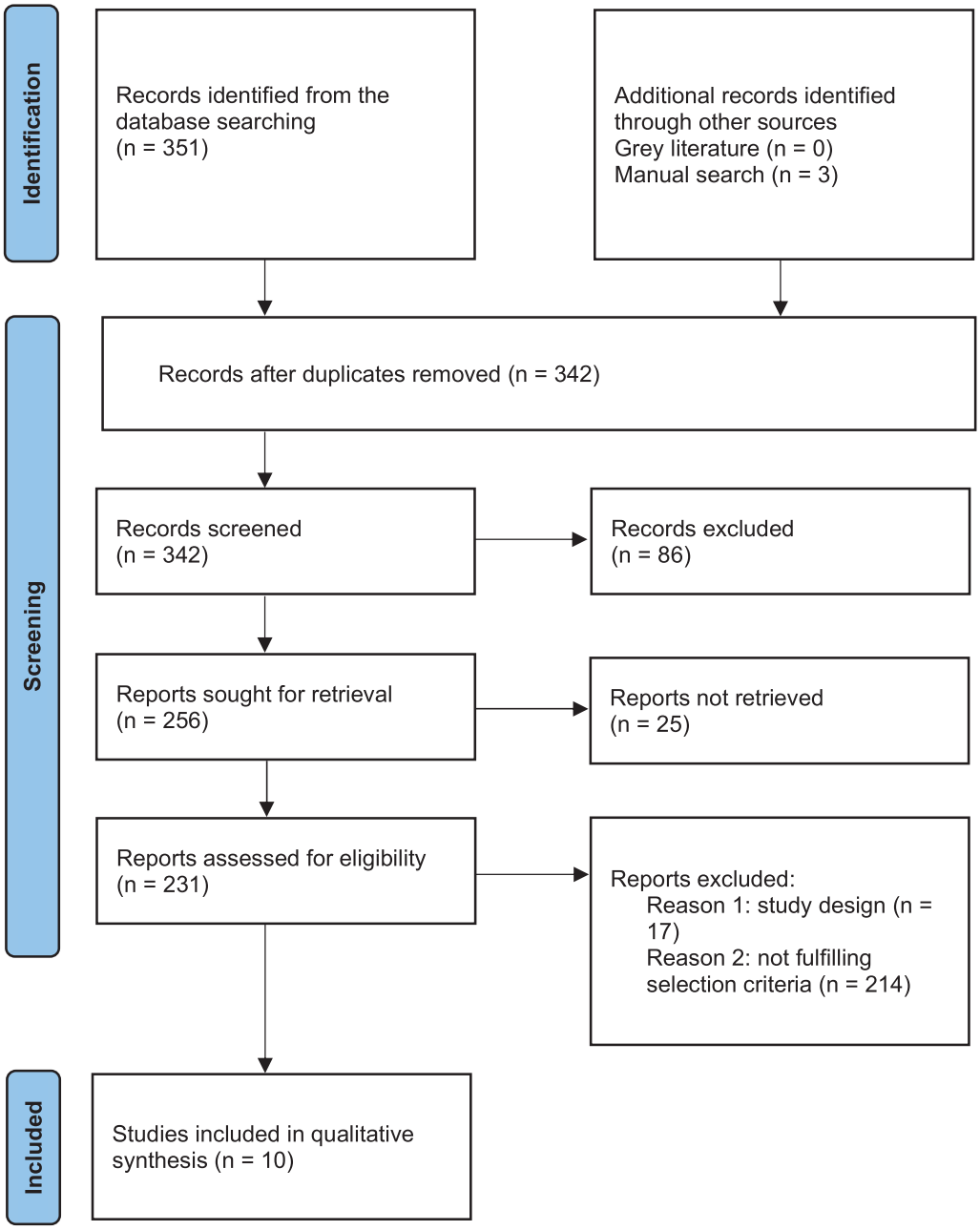


FIGURE 1 - PRIMSA flow diagram of the study selection process.

study (28) compared manipulation plus exercise (piriformis and iliopsoas stretching and reverse Kegal exercise) and exercise alone using the VAS, ODI, and PQ in patients with chronic coccydynia.

Stretching/massage [two studies (26,27)]: These studies targeted different muscle groups. One pilot study (26) compared levator anus stretch, levator anus and coccygeus massage, and sacrococcygeal mobilization using the VAS in patients with chronic coccydynia. The other study (27) compared piriformis and iliopsoas stretching, piriformis and iliopsoas stretching plus Maitland's rhythmic oscillatory thoracic mobilization, and conventional therapy (seat cushioning, sitz bath, and phonophoresis) using VAS and PFSD in patients with a clinical diagnosis of coccydynia.

Kinesiotaping [one study (21)]: This study compared combined kinesiotaping plus exercise and exercise alone using VAS, MMST, and ODI in patients with obesity-induced coccydynia.

Extracorporeal shock wave therapy [two studies (24,35)]: One study (35) compared the radial ESWT, focused ESWT, and sham ESWT using VAS and ODI in patients with coccydynia. The other study (24) compared ESWT and electrotherapy (SWD and IFT) using VAS, ODI, and self-reported satisfaction scores in patients with a first-time diagnosis of coccydynia.

Primal reflex release technique [one study (22)]: This study compared PRRT with piriformis and iliopsoas stretching using NPRS, DPQ, and PFSD in patients with a clinical diagnosis of coccydynia.

TABLE 1 - Characteristics of the included studies (*n* = 10)

Author & Year	Setting (country)	Design	Participants	Intervention	Comparator	Outcomes
Maigne and Chatellier. (2001) (26)	Hospital (France)	Pilot RCT	*N* = 74 Patients with chronic coccydynia (> 2 months) Mean age = 45.2 ± 14.8 years	Levator anus stretching (n = 25) Levator anus massage (n = 24) Duration = 3 min, 3 sessions for 10 days	Mobilization of coccyx (n = 25) Duration = 3 min, 3 sessions for 10 days	VAS assessed at baseline, 7 days, 30 days, 6 months, and at 2 years
Maigne et al. (2006) (25)	Hospital (France)	RCT	*N* = 102 Patients with chronic coccydynia (> 2 months) Mean age = 45.2 ± 11.5 years	Intrarectal manipulation (n = 51) Duration = 5 min, 3 sessions for 10 days	SWD (n = 51) Applied at the sacrum, 3 sessions for 10 days	VAS, (modified) MPQ, PQ, and (modified) DPQ at baseline, 1 month, and at 6 months follow-up
Khatri et al. (2011) (23)	Hospital (India)	RCT	*N* = 36 Patient with chronic coccydynia (> 2 months) Mean age = 31.0 ± 8.87 years	Coccygeal manipulation + phonophoresis (1MHz × 1 W/cm^2 × 8 min) + TENS (20–30 min) + coccygeal pillow advice, for 10 days	Phonophoresis (1MHz × 1 W/cm^2 × 8 min) + TENS (20–30 min) + coccygeal pillow advice), for 10 days	VAS, and PFSD at baseline and after 10 successive days
Lin et al. (2015) (24)	Hospital (Taiwan)	RCT	*N* = 41 Patients with a first-time diagnosis of coccydynia Mean age = 44.7 ± 14.8 years	ESWT, 2000 shots in the coccyx area per session for 4 sessions (5 Hz,3-4 bars) (n = 20) One session per week for 4 weeks	SWD combined with IFT for 20 min each thrice a week for 4 weeks	VAS, ODI, and Self-satisfaction score at baseline, 5th week, and 2 months
Mohanty et al. (2017) (27)	Hospital (India)	RCT	*N* = 48 Patients with a clinical diagnosis of coccydynia Mean age = not reported	Piriformis and iliopsoas stretching (n = 16) Piriformis and iliopsoas stretching plus Maitland thoracic mobilization (n = 16) Duration = 3 weeks	Conventional therapy (seat cushioning + Sitz bath + phonophorosis) (n = 16) Duration = 3 weeks	PPT using a modified syringe algometer and PFSD at baseline, after 3 weeks and 1 month follow-up
Abdel-Aal et al. (2020) (21)	Hospital (Egypt)	RCT	*N* = 60 Patients with obesity-induced coccydynia Age range = 45–60 years, with BMI > 32 kg/m^2	Kinesiotaping + exercise (n = 30) Duration = 3 weeks	Sham kinesiotaping + exercise (n = 30) Duration = 3 weeks	VAS, MMST, and ODI assessed at baseline, 3 weeks, and 1 month follow-up
Seemal et al. (2022) (22)	Hospital (Pakistan)	RCT	*N* = 46 Patients with a clinical diagnosis of coccydynia Mean age = 33.5 ± 11.3 years	PPRT + hot pack (n = 23) Duration = 12 sessions for 4 weeks	Piriformis stretching with hot pack (n = 23) Duration = 12 sessions for 4 weeks	NPRS, DPQ, and PFSD assessed at baseline and 1 month post-intervention
Zia et al. (2023) (29)	Hospital (Pakistan)	RCT	*N* = 50 Females with coccydynia Age range = 18–40 years Mean BMI Group 1 = 25.7 ± 2.34 kg/m^2, Group 2 = 26.2 ± 2.36 kg/m^2	MET to piriformis and iliopsoas (n = 25) Duration = 10 sessions (5 days/ week)	Static stretching of piriformis and iliopsoas (n = 25) Duration = 10 sessions (5 days/ week)	NPRS, (modified) DPQ, and PFSD at baseline and 2 weeks

(*Cont.*)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 1 - (*Cont.*)

Author & Year	Setting (country)	Design	Participants	Intervention	Comparator	Outcomes
Şah et al. (2023) (35)	Hospital (Turkey)	RCT	*N* = 60 Patients with coccydynia (1–3 months) Subacute cases Age range = 18–65 years Mean BMI = 26.2 ± 3.0 kg/m^2	Radial ESWT (8Hz, 1.6 bar pressure, 0.02-0.60 mJ/mm2 energy) (n = 20) Duration = 3 min 8 sec/session, 1 session/week for 4 weeks, along with oral painkiller	Focused ESWT (8Hz, 1.8 bar pressure, 0.02-0.60 mJ/mm2 energy) (n = 20) Sham ESWT (1Hz, 1 bar pressure, no energy) Duration = 3 min 8 sec/session, 1 session/week for 4 weeks, along with oral painkiller	VAS, and ODI at baseline, 1 month, 2 months and 4 months
Tufekci et al. (2024) (28)	Hospital (Turkey)	RCT	*N* = 46 Patients with chronic coccydynia (> 3 months) Mean age = 41.0 ± 7.47 years	Intrarectal manipulation of coccyx once a week for 4 weeks, + exercise (piriformis and iliopsoas stretching, + reverse Kegel exercise) (n = 23) 3 days/ week for 4 weeks	Exercise alone (as used for the intervention group) 3 days/week for 4 weeks (n = 23)	VAS, ODI, and PQ at 1 month and 6 months follow-up

ADL, Activities of daily living; BMI, Body mass index; DPQ, Dallas Pain Questionnaire; ESWT, Extracorporeal shock wave therapy; IFT, Interferential therapy; MET, Muscle energy technique; MMST, Modified-Modified Schober Test; NPRS, Numeric Pain Rating Scale; ODI, Oswestry Disability Index; PFSD, Pain-free sitting duration; PPT, pain pressure threshold; PRRT, Primal reflex release technique; PQ, Paris (functional coccydynia impact) Questionnaire; RCT, randomized controlled trial; SWD, Short wave diathermy; TENS, Transcutaneous electrical nerve stimulation; VAS, Visual Analog Scale; *N* = study sample size.

Muscle energy technique [one study (29)]: This study compared MET and static piriformis and iliopsoas stretching using NPRS, (modified) DPQ, and PFSD in females with coccydynia.

Two studies reported no adverse events associated with treatments (21,25).

Risk of bias within studies

The risk of bias assessment for each outcome in the included studies is shown in Table 3. Of the ten studies, four (50%) were rated as having a low risk of bias (21,24,26,28), four (40%) were rated as having a high risk of bias (22,23,25,27), and one (10%) was rated as having some concerns (21) for all the evaluated outcomes. A high risk of bias was predominantly observed in the domains related to bias in the measurement of outcome due to either unblinding of the outcome of assessors or inappropriate measurement of the outcome. Some concerns were observed in the domains related to bias arising from deviations in the intended intervention since the therapists delivering intervention knew the participants' treatment group and blinding participants was impossible due to the nature of the intervention. Additionally, the lack of detailed information about the randomization process resulted in some concerns in the domain related to bias in the randomization process. No domain was entirely free of a high risk of bias or some concerns. Notably, the outcome of PFSD was associated with a high risk of bias across all the included studies (Table 3).

Synthesis of results

Manipulation versus electrotherapy or exercise

Three studies (23,25,28) reported intrarectal manipulation to be effective at reducing pain and improving function in the short term (Table 4). One (23) of the studies reported significant improvements in pain [MD in VAS = 3.90, 95% CI 2.93-4.86) and function (MD in PFSD = 24.0, 95% CI 17.6-30.3] in favor of manipulation plus electrotherapy (phonophoresis and TENS) compared to the same electrotherapy at 10 days post-intervention. However, this study had a high risk of bias for all evaluated outcomes and lacked follow-up (23). Similarly, the second study (25) reported significant reduction in pain (MD in VAS = −14.5, 95% CI −24.2 to −4.70; MD in (modified) MPQ = −6.5, 95% CI −11.1 to −1.85) and improvement in function (MD in (modified) DPQ = −10.5, 95% CI −18.2 to −2.71; MD in PQ = −20.0, 95% CI −29.3 to −10.6) favoring manipulation over electrotherapy at 1-month post-intervention. Additionally, the manipulation group demonstrated better outcomes twice (22%) as often as the electrotherapy group (12%) at 6 months follow-up. This study (25), however, had a high risk of bias for all evaluated outcomes and incomplete data for the 6 months follow-up, limiting the ability to compute the

TABLE 2 - Summary of outcomes and interventions (*n* = 10)

Interventions	Study	Outcomes			
		Pain	Function	Mobility	Satisfaction
Intra-rectal manipulation	Tufekci et al. (2024) (28)	VAS	ODI, PQ		
	Maigne et al. (2006) (25)	VAS, (modified) MPQ	(modified) DPQ, PQ		
	Khatri et al. (2011) (23)	VAS	PFSD		
Piriformis and iliopsoas stretching, Piriformis and iliopsoas stretching plus thoracic mobilization	Mohanty et al. (2017) (27)	PPT	PFSD		
Levator anus massage, Levator anus stretching		VAS			
Extracorporeal shock wave therapy	Maigne and Chatellier. (2001) (26)	VAS	ODI		
	Şah et al. (2023) (35)	VAS	ODI		
Kinesiotaping plus exercise	Lin et al. (2015) (24)	VAS	ODI		Self-satisfaction score
Muscle energy technique	Abdel-Aal et al. (2020) (21)	NPRS	(modified) DPQ PFSD	MMST	
Primal reflex release technique	Zia et al. (2023) (29)	NPRS	(modified) DPQ PFSD		

DPQ, Dallas Pain Questionnaire; MMST, Modified-Modified Schober Test; MPQ, McGill Pain Questionnaire; NPRS, Numeric Pain Rating Scale; ODI, Oswestry Disability Index; PFSD, Pain-free sitting duration; PPT, Pain pressure threshold; PQ, Paris (functional coccydynia impact) Questionnaire; VAS, Visual Analog Scale

magnitude of the treatment effect. The third study (28) had a low risk of bias for all evaluated outcomes and reported a significant reduction in pain (MD in VAS = −13.4, 95% CI −20.3 to −6.58) and improvement in function (MD in ODI = −7.65, 95% CI −11.5 to −3.72 and MD in PQ = −12.6, 95% CI −19.6 to −5.57) favoring manipulation plus exercise compared to exercise alone at 1-month post-intervention. However, no therapeutic superiority was observed at the 6-month follow-up in any of the outcomes, although sustained benefits were observed more in the manipulation plus exercise group.

Levator anus stretching versus levator anus massage versus joint mobilization

One pilot RCT (26) with a low risk of bias reported a 25.7% success rate in pain reduction at 6 months, similar to the rate observed at 7 days, and a 24.3% success rate at 2 years follow-up for both levator anus stretching and levator anus and coccygeus massage. Although outcomes varied depending on the cause of coccydynia, both stretching and massage had similar and better results than mobilization after 7 days, at 6 months, and 2 years follow-up (Table 4). However, this study (26) had incomplete data to estimate the magnitude of treatment effects in terms of MD and 95% CI.

Piriformis and iliopsoas stretching with or without mobilization versus conventional therapy

One study (27) reported piriformis and iliopsoas stretching alone or combined with mobilization to be more effective than conventional therapy at reducing pain and improving function at 3 weeks post-intervention and 1-month follow-up ($p < 0.001$) (Table 4). Adding mobilization to stretching did not provide any additional benefit over stretching alone. However, the study had a high risk of bias for all evaluated outcomes and incomplete data, making it impossible to estimate the magnitude of the treatment effects.

Extracorporeal shock wave therapy (ESWT) versus sham versus electrotherapy

Two studies (24, 35) with a low risk of bias for all evaluated outcomes determined the effect of ESWT (Table 4). In one (35) of the studies, both focused and radial ESWT significantly reduced pain and improved function at 2 and 4 months compared to baseline. However, no significant difference in pain reduction was found between the two ESWT types, although focused ESWT showed a slight advantage over radial ESWT. For functional improvement, radial ESWT was superior to focused ESWT (MD = −11.2, 95% CI −21.3 to −2.33) and sham (MD = −18.0, 95% CI −28.0 to −7.92) at the 4-month follow-up only (35). The other study (24) reported a significant reduction in pain favoring ESWT over electrotherapy at 5th-week post-intervention (MD = −12.7, 95% CI −24.5 to −0.85) and 2 months follow-up (MD = −18.5, 95% CI −31.6 to −5.37). However, there was no significant difference in functional improvement between the groups. Additionally, 70% of patients in the ESWT group reported good to excellent satisfaction, which was higher than the satisfaction rate in the electrotherapy group (Table 4).

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 3 - Risk of bias assessments using RoB 2.0 for the included studies (*n* = 10)

Study (Author & Year)	Outcome	D1: Bias in randomization	D2: Bias from deviations	D3: Bias from missing data	D4: Bias in measurement	D5: Bias in reporting	Overall judgment
Maigne & Chatellier (2001) (26)	Pain (VAS)	Low	Low	Low	Low	Low	Low
Maigne et al. (2006) (25)	Pain (VAS)	Low	Some concerns	Low	High	Low	High
Maigne et al. (2006) (25)	Pain (modified MPQ)	Low	Some concerns	Low	High	Low	High
Maigne et al. (2006) (25)	Function (modified DPQ)	Low	Some concerns	Low	High	Low	High
Maigne et al. (2006) (25)	Function (PQ)	Low	Some concerns	Low	High	Low	High
Khatri et al. (2011) (23)	Pain (VAS)	Some concerns	Some concerns	High	High	Low	High
Khatri et al. (2011) (23)	Function (PFSD)	Some concerns	Some concerns	High	High	Low	High
Lin et al. (2015) (24)	Pain (VAS)	Low	Low	Low	Low	Low	Low
Lin et al. (2015) (24)	Function (ODI)	Low	Low	Low	Low	Low	Low
Lin et al. (2015) (24)	Self-satisfaction score	Low	Low	Low	Low	Low	Low
Mohanty et al. (2017) (27)	Pain (PPT)	Some concerns	Some concerns	Low	Low	Some concerns	High
Mohanty et al. (2017) (27)	Function (PFSD)	Some concerns	Some concerns	Low	High	Low	High
Abdel-Aal et al. (2020) (21)	Pain (VAS)	Low	Low	Low	Low	Low	Low
Abdel-Aal et al. (2020) (21)	Function (ODI)	Low	Low	Low	Low	Low	Low
Abdel-Aal et al. (2020) (21)	Mobility (MMST)	Low	Low	Low	Low	Low	Low
Seemal et al. (2022) (22)	Pain (NPRS)	Low	Low	Low	High	Low	High
Seemal et al. (2022) (22)	Function (modified DPQ)	Low	Low	Low	High	Low	High
Seemal et al. (2022) (22)	Function (PFSD)	Low	Low	Low	High	Low	High
Zia et al. (2023) (29)	Pain (NPRS)	Low	Low	Low	Low	Low	Low
Zia et al. (2023) (29)	Function (modified DPQ)	Low	Low	Low	Low	Low	Low
Zia et al. (2023) (29)	Function (PFSD)	Low	Low	Low	High	Low	High
Şah et al. (2023) (35)	Pain (VAS)	Some concerns	Low	Low	Low	Low	Some concerns
Şah et al. (2023) (35)	Function (ODI)	Some concerns	Low	Low	Low	Low	Some concerns
Tufekci et al. (2024) (28)	Pain (VAS)	Low	Low	Low	Low	Low	Low

Study (Author & Year)	Outcome	D1: Bias in randomization	D2: Bias from deviations	D3: Bias from missing data	D4: Bias in measurement	D5: Bias in reporting	Overall judgment
Tufekci et al. (2024) (28)	Function (ODI)	Low	Low	Low	Low	Low	Low
Tufekci et al. (2024) (28)	Function (PQ)	Low	Low	Low	Low	Low	Low

Abbreviations /Notes
D1 (Bias in randomization): Concerns about random sequence generation or allocation concealment.
D2 (Bias from deviations): Whether participants and personnel were blinded (when feasible) and whether interventions were delivered as intended.
D3 (Bias from missing data): How missing outcome data (dropouts, attrition) were handled.
D4 (Bias in measurement): Whether outcome assessment could be influenced by knowledge of the assigned intervention.
D5 (Bias in reporting): Whether all pre-specified outcomes and time points were reported or if selective reporting is suspected.
Dallas Pain Questionnaire; MMST, Modified-Modified Schober Test; MPQ, McGill Pain Questionnaire; NPRS, Numeric Pain Rating Scale; ODI, Oswestry Disability Index; PFSD, Pain-free sitting duration; PPT, Pain pressure threshold; PQ, Paris (functional coccydynia impact) Questionnaire; VAS, Visual Analog Scale

Kinesiotaping plus exercise versus exercise

One study (21) with some concerns related to the risk of bias in all evaluated outcomes reported kinesiotaping plus exercise to be more effective than exercise alone for improving pain (MD = −6.83, 95% CI −8.99 to −4.66 at 3 weeks; MD = −8.70, 95% CI −10.6 to −6.71 at 1 month), reducing disability (MD = −5.70, 95% CI −6.92 to −4.47 at 3 weeks; MD = −5.80, 95% CI −6.76 to −4.83 at 1 month), and increasing trunk flexion (MD = 0.80, 95% CI 0.23 to 1.36 at 3 weeks; MD = 0.53, 95% CI 0.18 to 0.87 at 1 month) (Table 4). However, the specific contribution of kinesiotaping alone could not be isolated, as it was only evaluated in combination with exercise.

Muscle energy technique (MET) versus stretching

One study (29), which had a low risk of bias overall but a high risk of bias for the PSFD outcome, reported MET to be superior to static stretching for improving pain (MD = −2.04, 95% CI −3.19 to −1.60) and function (MD in (modified) DPQ = −2.04, 95% CI −3.66 to −0.4; MD in PFSD = 37.4, 95% CI 19.4 to 55.3) at 2 weeks post-intervention (Table 4).

Primal reflex release technique (PRRT) versus stretching

One study (22) with a high risk of bias reported PRRT to be superior to stretching for improving pain (MD = −1.95, 95% CI −2.71 to −1.20) and function (MD in (modified) DPQ= −22.6, 95% CI −36.7 to −8.41; MD in PFSD = 274.3, 95% CI 196.6-352.0) at 1-month post-intervention (Table 4).

Discussion

Main findings and interpretation of the results

Pain medication and physical therapy are typically regarded as first-line treatments for musculoskeletal disorders. This review evaluated the effectiveness of physical therapy approaches for adults with coccydynia, which could contribute to the development of clinical guidelines for coccydynia. However, less than half of the studies in this review had a low risk of bias in all evaluated outcomes, suggesting that the overall findings should be interpreted cautiously. Reliable conclusions can only be drawn from studies free of bias.

Ten RCTs were included, evaluating the effects of various interventions (i.e., manipulation, stretching, massage, ESWT, kinesiotaping, MET, and PRRT) on pain and function (21–29,35), mobility (21), and patient satisfaction (24). Most interventions were passive, with a few studies combining them with other treatments such as kinesiotaping plus exercise (21) and manipulation plus electrotherapy (23) or exercise (28). Notably, no study explored active exercise as a standalone intervention, likely because manual treatments target the underlying misalignment or restricted movement in the sacrococcygeal joint, which is believed to be the primary perpetrator of pain.

Despite slight variations in the outcomes evaluated, current evidence suggests that, in the short term (up to 3 months post-intervention), physical therapy interventions including ESWT, kinesiotaping plus exercise, levator anus stretching or massage, manipulation alone or manipulation plus exercise or electrotherapy, and MET are effective at reducing pain and improving function (based on seven studies (21,24-26,28,29,35). Additionally, kinesiotaping plus exercise is effective at improving trunk mobility [based on one study (21)], whereas ESWT led to greater treatment satisfaction [based on one study (24)]. In the intermediate term (up to 6 months), manipulation alone and levator anus stretching or massage are effective at reducing pain [based on two studies (25,26)], whereas manipulation alone is effective at improving function [based on one trial (25)]. Only one study reported long-term (up to 2 years) pain reduction, which was associated with levator anus stretching or massage (26). No study evaluated the long-term effect of physical therapy intervention on function or trunk mobility.

Manipulation

Manipulation alone (25), or combined with electrotherapy (23) or exercise (28), was effective in alleviating pain in the short term (23,25,28) and the intermediate term (25). This approach may provide immediate relief by addressing anatomical anomalies such as misaligned coccyx reducing tension in the pelvic floor muscles. The addition of electrotherapy or exercise is believed to provide a positive reinforcement effect. Although one study (25) reported sustained benefits of manipulation for up to 6 months, significant methodological concerns in this study undermine the

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 4 - Main findings of the included studies (*n* =10)

Author & Year	Main findings (authors' synthesis)
Maigne and Chatellier. (2001) (26)	The success rate for manual treatments, defined as the reduction in pain from baseline, was 25.7% at 6 months, consistent with the 7-day results, and remained at 25.7% after 2 years. Outcomes varied based on the underlying cause of coccydynia, with massage and stretching demonstrating greater effects than mobilization. Additionally, individuals with normal coccyx mobility experienced better results, achieving a success rate of 43.8%, compared to those with luxation, hypermobility, or immobile coccyx.
Maigne et al. (2006) (25)	Compared to electrotherapy, manipulation led to significant improvements in VAS (MD = −14.5, 95% CI −24.2 to −4.70; $p = 0.004$), (modified) MPQ (MD = −6.5, 95% CI −11.1 to −1.85; $p = 0.007$), (modified) DPQ (MD = −10.5, 95% CI −18.2 to −2.71; $p = 0.009$), and PQ (MD = −20.0, 95% CI −29.3 to −10.6; $p = 0.001$) at 1-month post-intervention. At 6 months, the manipulation group demonstrated 11 good outcomes (22% of patients), with ≥ 60% improvement in the individual global score compared to ≥ 50% at 1 month. In contrast, the control group showed six good outcomes (12% of patients).
Khatri et al. (2011) (23)	Compared to electrotherapy (phonophoresis and TENS), manipulation significantly reduced VAS (MD = 3.90 95% CI 2.93 to 4.86; $p = 0.0001$) and improved PSFD (MD = 24.0, 95% CI 17.6 to 30.3; $p = 0.0002$) 10 days post-intervention.
Lin et al. (2015) (24)	ESWT led to a significant reduction in VAS compared to electrotherapy (SWD plus IFT) at 5th-week post-intervention (MD = −12.7, 95% CI −24.5 to −0.85; $p = 0.042$) and 2 months follow-up (MD = −18.5, 95% CI −31.6 to −5.37; $p = 0.009$). Approximately 70% of patients receiving ESWT reported good to excellent treatment satisfaction, which was significantly higher than those receiving electrotherapy ($p = 0.003$). No statistically significant difference in ODI between the groups at any time point ($p > 0.05$).
Mohanty et al. (2017) (27)	Piriformis and iliopsoas stretching, either alone or combined with mobilization, resulted in greater improvements in VAS and PFSD compared to conventional therapy at 3 weeks post-intervention and 1-month follow-up ($p < 0.001$).
Abdel-Aal et al. (2020) (21)	Compared to exercise alone, kinesiotaping plus exercise resulted in significantly greater improvements in VAS (MD = −6.83, 95% CI −8.99 to −4.66; $p < 0.001$ at 3 weeks post-intervention; MD = −8.70, 95% CI −10.6 to −6.71; $p < 0.001$ at 1-month follow-up), ODI (MD = −5.70, 95% CI −6.92 to −4.47; $p < 0.001$ at 3 weeks post-intervention; MD = −5.80, 95% CI −6.76 to −4.83; $p < 0.001$ at 1-month follow-up), and MMST (MD = 0.80, 95% CI 0.23 to 1.36; $p < 0.001$ at 3 weeks post-intervention; MD = 0.53, 95% CI 0.18 to 0.87; $p < 0.001$ at 1-month follow-up).
Seemal et al. (2022) (22)	Compared with stretching, PRRT resulted in significantly greater improvements in NPRS (MD = −1.95, 95% CI −2.71 to −1.20; $p < 0.001$), DPQ (MD = −22.6, 95% CI −36.7 to −8.41; $p = 0.003$), and PFSD (MD = 274.3, 95% CI 196.6 to 352.0; $p < 0001$) at 1-month post-intervention.
Zia et al. (2023) (29)	Compared to static stretching, MET resulted in significantly greater improvements in NPRS (MD = −2.40, 95% CI −3.19 to −1.60; $p < 0.001$), DPQ (MD = −2.04, 95% CI −3.66 to −0.41; $p = 0.017$), and PFSD (MD = 37.4, 95% CI 19.4 to 55.3; $p < 0.001$) at 2 weeks post-intervention.
Şah et al. (2023) (35)	Radial ESWT showed significant reductions in VAS scores 1-month post-intervention compared to baseline. Both focused and radial ESWT significantly reduced VAS and ODI scores at 2 and 4 months. For the between-group difference in VAS, focused ESWT was superior to sham at 1 month (MD = 1.50, 95% CI 0.40 to 2.59; $p = 0.017$) and 4 months (MD = −1.50, 95% CI −2.73 to −0.26; $p = 0.022$) but comparable to radial ESWT ($p > 0.05$). Radial ESWT was superior to sham (MD = −1.40, 95% CI −2.45 to −0.34; $p = 0.013$) at 2 months only. For the between-group difference in ODI, radial ESWT was superior to focused ESWT (MD = −11.2, 95% CI −21.3 to −2.33; $p = 0.020$) and sham (MD = −18.0, 95% CI −28.0 to −7.92; $p = 0012$) at 4 months follow-up.
Tufekci et al. (2024) (28)	Compared to exercise alone, manipulation plus exercise resulted in significantly greater improvements in VAS (MD = −13.4, 95% CI −20.3 to −6.58; $p < 0.001$), ODI (MD = −7.65, 95% CI −11.5 to −3.72; $p < 0.001$), and PQ (MD = −12.6, 95% CI −19.6 to −5.57–7.65; $p < 0.01$) at 1-month post-intervention. Though sustained effects were more pronounced in the manipulation plus exercise group at 6 months follow-up, there was no significant difference between the two interventions ($p > 0.05$).

CI, confidence interval; DPQ, Dallas Pain Questionnaire; ESWT, Extracorporeal shock wave therapy; IFT, Interferential therapy; MD, mean difference; MET, Muscle energy technique; MMST, Modified-Modified Schober Test; NPRS, Numeric Pain Rating Scale; ODI, Oswestry Disability Index; PFSD, Pain-free sitting duration; PRRT, Primal reflex release technique; PQ, Paris (functional coccydynia impact) Questionnaire; SWD, short wave diathermy; TENS, Transcutaneous electrical nerve stimulation; VAS, Visual Analog Scale

reliability of its conclusions. Moreover, the effects of manipulation as a standalone could not be substantiated in the studies using electrotherapy (23) or exercise (28) as an adjunct, though a combination approach is often recommended (39). Despite promising results, the sustainability of these benefits remains uncertain due to the lack of long-term follow-up (23,28). Therefore, while combining intrarectal manipulation and exercise (28) seems to be a valuable option compared to combining with electrotherapy (23), clinicians should remain cautious and consider incorporating follow-up assessments

to evaluate sustained benefits. Moreover, given that the application of transrectal manual techniques is feasible and acceptable in the Western population (40), further evaluation is warranted across diverse cultural settings. The technique involves digital rectal evaluation before and after the intervention to palpate the coccyx. As such, patient preferences, such as therapist gender, need to be considered.

Stretching and massage

Compared to joint mobilization, levator anus stretching or massage therapies were effective in alleviating pain both in the short and long term (26). Similarly, the superiority of piriformis and iliopsoas stretching with or without mobilization over conventional therapy was demonstrated for both pain and function in the short term (27). These therapies likely work by enhancing blood flow, reducing muscle tension, and increasing flexibility (41, 42). The demonstrated effects of levator stretching or massage across various follow-up points (7 days, 6 months, and 2 years) suggest that these interventions may offer both immediate and sustained benefits. However, the high risk of bias in the study demonstrating the efficacy of piriformis and iliopsoas stretching (27), coupled with the limited number of studies evaluating stretching as a treatment for coccydynia, highlights the need for more robust trials to elucidate these findings.

Kinesiotaping

Although kinesiotaping showed short-term improvements in pain, function, and trunk mobility (21) among obese with coccydynia, the effects could have also been attributed to exercise used as an adjunct. However, this dual approach likely offers synergistic benefits as kinesiotaping might provide support and reduce strain (43), while exercise helps strengthen muscles and improve function. This finding supports the integration of multimodal strategies in the management of coccydynia. Additionally, no side effects were reported in association with kinesiotaping (21), confirming its safety. As kinesiotaping application over the coccyx could facilitate bacterial invasion in the skin and lead to cellulitis, regular monitoring is mandatory (43). The lack of long-term follow-up and limited studies make it challenging to draw definitive conclusions about the sustainability effect of kinesiotaping as well as its efficacy as a standalone treatment for coccydynia.

Extracorporeal shock wave therapy (ESWT)

ESWT proved to be effective for alleviating pain and improving function, particularly in the short term (24, 35). Its therapeutic mechanism may include neovascularization (44), reduced substance P-immunoreactive neurons (45), and decreased levels of inflammatory mediators (46). High patient satisfaction levels further support its use (24). However, the type of ESWT may influence outcomes, with radial ESWT showing superior improvement in function compared to focused ESWT (35). Thus, tailoring the type of ESWT to individual patient needs could lead to better clinical outcomes. However, the lack of long-term follow-up data in existing studies limits understanding of its sustained effects, which is an obvious research gap for future studies.

Muscle energy technique (MET)

The findings that MET was better for pain and function over static stretching in the short term (29) suggest MET as one of the promising hands-on approaches for coccydynia. This technique involves active patient participation through isometric contraction and relaxation, which may stimulate joint proprioceptors, alter motor programming, and promote hypoalgesia (47). Moreover, increased mobility and function may be related to mechanisms promoting hypoalgesia, an increase in stretch tolerance, and possibly viscoelastic change in the muscle (47, 48). However, the limited number of studies calls for further research to confirm these benefits and establish standardized protocols for MET application in coccydynia.

Primal reflex release technique (PRRT)

Compared to piriformis stretching, PRRT shows superior effects in reducing pain and improving function in the short term (22). PRRT induces reciprocal inhibition between agonist and antagonist muscles, leading to a complementary suppression of the autonomic nervous system. This reduces reflexive muscle tone through the release of acetylcholine and serotonin, which contributes to pain reduction. This process may also positively impact the patient's psychological state. Nonetheless, the methodological weakness in the study examining PRRT (22) and the lack of adequate relevant trials require a cautious interpretation of the results. Further research with adequate follow-up is essential to support the efficacy of PRRT for coccydynia.

Limitations

Regarding the limitations of the included studies, there is a lack of adequate trials evaluating physical therapy interventions, particularly active treatments, for coccydynia. Many studies lacked true control groups, long-term outcome evaluation, and sufficient sample sizes. Although some studies were methodologically sound, many had critical issues in areas such as randomization, blinding of outcome assessors, and inappropriate outcome measurements, raising concerns about the validity of their results. Additionally, three studies (25–27) had incomplete data, which could not be retrieved.

With respect to the limitations of this review, first, a meta-analysis was not conducted due to substantial heterogeneity in the interventions employed by the included studies. Second, the small number of studies on various physical therapy interventions limits the robustness and generalizability of the conclusions. Lastly, the inclusion of only English-language published studies may have excluded other relevant research that could have contributed to the findings.

Recommendation for future research

Future high-quality RCTs with adequate sample size and long-term follow-up are needed to evaluate the effectiveness

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

of different physical therapy interventions either in isolation or combination to enable the pooling of data using a formal analysis like meta-analysis. Placebo-controlled trials are also crucial to establish treatment efficacy while minimizing bias and isolating the true impact of interventions. Key methodological areas, including randomization, blinding, and measurement of outcomes, should be addressed in future studies to ensure accurate quantification of effects and reliable evidence. Furthermore, studies should standardize intervention protocols, include cost-effectiveness analyses, and improve overall methodological reporting.

Conclusions

Based on the findings from the included studies, physical therapy interventions, including ESWT, kinesiotaping, manipulation, massage, MET, and stretching, show promise in improving pain and function in adults with coccydynia in the short term. However, the long-term effects of these interventions remain unclear. Further, high-quality RCTs are needed to compare the efficacy of various physical therapy interventions, both in isolation and in combination.

List of abbreviations

ADL: Activities of daily living; BMI: Body mass index; DPQ: Dallas Pain Questionnaire; ESWT, Extracorporeal shock wave therapy; IFT: Interferential therapy; MET: Muscle energy technique; MMST: Modified-Modified Schober Test; NPRS: Numeric Pain Rating Scale; ODI: Oswestry Disability Index; PFSD: Pain-free sitting duration; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis; PRRT: Primal reflex release technique; PQ: Paris (functional coccydynia impact) Questionnaire; RCT: Randomized controlled trial; SWD: Short wave diathermy; TENS: Transcutaneous electrical nerve stimulation; VAS, Visual Analog Scale

Disclosures

Conflict of interest: The authors declare that they have no competing interests.

Financial support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contributions: MS, BJ, AC, and HR developed the research questions and designed the overall study. HR and BJ developed and ran the literature search strategies. MS and AC conducted title and abstract screening with the contribution of RHA or AHA Jr. AHA Sr and ASA conducted full-text screening and data charting. MS and HR wrote the first draft of the manuscript. AHA Sr, ASA, FZK, and HV assisted in drafting and reviewing the manuscript and preparing tables and figures with AAH and HR contributions. AHA Jr, ASA, HV, EAA, and AAI contributed to the organization, analysis, and interpretation of the results. AAI and HR revised the final draft. All authors read and approved the final manuscript.

Ethical approval and consent to participate: Not applicable, as this study is a review and does not involve human participants.

Consent for publication: Not applicable. No identifying images or other personal or clinical details of participants are presented here.

Data availability statement: All data generated and/or analyzed during this study are available from the corresponding author upon reasonable request.

Systematic review registration: PROSPERO, CRD42022344003.

References

1. Fogel GR, Cunningham PY III, Esses SI. Coccygodynia: evaluation and management. J Am Acad Orthop Surg. 2004;12(1):49-54. PubMed
2. Nathan ST, Fisher BE, Roberts CS. Coccydynia: a review of pathoanatomy, aetiology, treatment and outcome. J Bone Joint Surg Br. 2010;92(12):1622-1627. PubMed
3. Lirette LS, Chaiban G, Tolba R, et al. Coccydynia: an overview of the anatomy, etiology, and treatment of coccyx pain. Ochsner J. 2014;14(1):84-87. PubMed
4. Gavriilidis P, Kyriakou D. Sacrococcygeal chordoma, a rare cause of coccygodynia. Am J Case Rep. 2013;14:548-550. PubMed
5. Kim HS, Yang SH, Park HJ, et al. Glomus tumor as a cause of coccydynia. Skeletal Radiol. 2013;42(10):1471-1473. PubMed
6. Maigne JY, Guedj S, Straus C. Idiopathic coccygodynia. Lateral roentgenograms in the sitting position and coccygeal discography. Spine. 1994;19(8):930-934. PubMed
7. Maigne JY, Pigeau I, Aguer N, et al. Chronic coccydynia in adolescents. A series of 53 patients. Eur J Phys Rehabil Med. 2011;47(2):245-251. PubMed
8. Pennekamp PH, Kraft CN, Stütz A, et al. Coccygectomy for coccygodynia: does pathogenesis matter? J Trauma. 2005;59(6):1414-1419. PubMed
9. Richette P, Maigne JY, Bardin T. Coccydynia related to calcium crystal deposition. Spine. 2008;33(17):E620-E623. PubMed
10. Garg B, Ahuja K. Coccydynia – A comprehensive review on etiology, radiological features and management options. J Clin Orthop Trauma. 2021;12(1):123-129. PubMed
11. Postacchini F, Massobrio M. Idiopathic coccygodynia. Analysis of fifty-one operative cases and a radiographic study of the normal coccyx. J Bone Joint Surg Am. 1983;65(8):1116-1124. PubMed
12. Antoniadis A, Ulrich NHB, Senyurt H. Coccygectomy as a surgical option in the treatment of chronic traumatic coccygodynia: a single-center experience and literature review. Asian Spine J. 2014;8(6):705-710. PubMed
13. El Mohsen Arafa A, El Rady Mahmoud A, Zayan M, et al. Coccygectomy for posttraumatic coccygodynia: A long-term prospective study. Curr Orthop Pract. 2017;28(5):484-488. CrossRef
14. Haddad B, Prasad V, Khan W, et al. Favourable outcomes of coccygectomy for refractory coccygodynia. Ann R Coll Surg Engl. 2014;96(2):136-139. PubMed
15. Hanley EN, Ode G, Jackson Iii BJ, et al. Coccygectomy for patients with chronic coccydynia: a prospective, observational study of 98 patients. Bone Joint J. 2016;98-B(4):526-533. PubMed
16. Karadimas EJ, Trypsiannis G, Giannoudis PV. Surgical treatment of coccygodynia: an analytic review of the literature. Eur Spine J. 2011;20(5):698-705. PubMed
17. Wray CC, Easom S, Hoskinson J. Coccydynia. Aetiology and treatment. J Bone Joint Surg Br. 1991;73(2):335-338. PubMed
18. Benditz A. Therapieoptionen bei der Coccygodynie. (Treatment options for coccygodynia). Orthopadie (Heidelb). 2024;53(2):100-106. PubMed
19. Shi Y, Wu W. Multimodal noninvasive non-pharmacological therapies for chronic pain: mechanisms and progress. BMC Med. 2023;21(1):372. PubMed
20. Delitto A, George SZ, Van Dillen L, et al. Orthopaedic Section of the American Physical Therapy Association. Low back pain. J Orthop Sports Phys Ther. 2012;42(4):A1-A57. PubMed
21. Abdel-Aal NM, Elgohary HM, Soliman ES, Waked IS. Effects of kinesiotaping and exercise program on patients with obesity-induced coccydynia: a randomized, double-blinded, sham-controlled clinical trial. Clin Rehabil. 2020;34(4):471-479. PubMed

22. Seemal P, Ayub A, Diilshad M, et al. Comparing primal reflex release technique and stretching exercises on pain and function in coccydynia. Iran Rehabil J. 2022;20(4):623-632. CrossRef
23. Khatri SM, Nitsure P, Jatti RS. Effectiveness of coccygeal manipulation in coccydynia: a randomized control trial. Indian J Physiother Occup Ther. 2011;5(3):110-112.
24. Lin SF, Chen YJ, Tu HP, et al. The effects of extracorporeal shock wave therapy in patients with coccydynia: a randomized controlled trial. PLoS One. 2015;10(11):e0142475. PubMed
25. Maigne JY, Chatellier G, Faou ML, et al. The treatment of chronic coccydynia with intrarectal manipulation: a randomized controlled study. Spine. 2006;31(18):E621-E627. PubMed
26. Maigne JY, Chatellier G. Comparison of three manual coccydynia treatments: a pilot study. Spine. 2001;26(20):E479-E483. PubMed
27. Mohanty PP, Pattnaik M. Effect of stretching of piriformis and iliopsoas in coccydynia. J Bodyw Mov Ther. 2017;21(3):743-746. PubMed
28. Tufekci O, Yilmaz K, Gercek H, et al. The effectiveness of manipulation in combination with exercise for patients with coccydynia: six months follow-up of a randomized controlled trial. Int J Osteopath Med. 2024;51:100711. CrossRef
29. Zia W, Bashir MS, Arshad MU, et al. Effects of muscle energy technique and static stretching of piriformis and iliopsoas in females with coccydynia. Journal of Xi'an Shiyou University. 2023;19(1):642-647. Natural Science Edition.
30. Marinko LN, Pecci M. Clinical decision making for the evaluation and management of coccydynia: 2 case reports. J Orthop Sports Phys Ther. 2014;44(8):615-621. PubMed
31. White WD, Avery M, Jonely H, et al. The interdisciplinary management of coccydynia: a narrative review. PM R. 2022;14(9):1143-1154. PubMed
32. Origo D, Tarantino AG, Nonis A, et al. Osteopathic manipulative treatment in chronic coccydynia: a case series. J Bodyw Mov Ther. 2018;22(2):261-265. PubMed
33. Haghighat S, Mashayekhi Asl M. Effects of extracorporeal shock wave therapy on pain in patients with chronic refractory coccydynia: a quasi-experimental study. Anesth Pain Med. 2016;6(4):e37428. PubMed
34. Andersen GØ, Milosevic S, Jensen MM, et al. Coccydynia-the efficacy of available treatment options: a systematic review. Global Spine J. 2022;12(7):1611-1623. PubMed
35. Şah V, Elasan S, Kaplan Ş. Comparative effects of radial and focused extracorporeal shock wave therapies in coccydynia. Turk J Phys Med Rehabil. 2023;69(1):97-104. PubMed
36. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372(71):n71. PubMed
37. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomized trials. BMJ. 2019;366:l4898. PubMed
38. Wan X, Wang W, Liu J, et al. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol. 2014;14:135. PubMed
39. Skelly AC, Chou R, Dettori JR, et al. Noninvasive nonpharmacological treatment for chronic pain: a systematic review update. Comparative effectiveness review no. 227. Agency for Healthcare Research and Quality; 2020.
40. Nourani B, Norton D, Kuchera W, et al. Transrectal osteopathic manipulation treatment for chronic coccydynia: feasibility, acceptability and patient-oriented outcomes in a quality improvement project. J Osteopath Med. 2023;124(2):77-83. PubMed
41. Weerapong P, Hume PA, Kolt GS. The mechanisms of massage and effects on performance, muscle recovery and injury prevention. Sports Med. 2005;35(3):235-256. PubMed
42. MacAuley D, Best T. Evidence-based sports medicine. 2nd ed. Blackwell Publishing; 2007.
43. Kase K, Wallis J, Kase T. Clinical therapeutic applications of the Kinesio taping ® method. 2nd ed. Kinesio Taping Association; 2003.
44. Wang CJ, Wang FS, Yang KD, et al. Shock wave therapy induces neovascularization at the tendon-bone junction. A study in rabbits. J Orthop Res. 2003;21(6):984-989. PubMed
45. Hausdorf J, Lemmens MA, Kaplan S, et al. Extracorporeal shockwave application to the distal femur of rabbits diminishes the number of neurons immunoreactive for substance P in dorsal root ganglia L5. Brain Res. 2008;1207:96-101. PubMed
46. Notarnicola A, Moretti B. The biological effects of extracorporeal shock wave therapy (eswt) on tendon tissue. Muscles Ligaments Tendons J. 2012;2(1):33-37. PubMed
47. Thomas E, Cavallaro AR, Mani D, et al. The efficacy of muscle energy techniques in symptomatic and asymptomatic subjects: a systematic review. Chiropr Man Therap. 2019;27(1):35. PubMed
48. Ballantyne F, Fryer G, McLaughlin P. The effect of muscle energy technique on hamstring extensibility: the mechanism of altered flexibility. J Osteopath Med. 2003;6(2):59-63. CrossRef

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 90-96
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3442
VIEWPOINT

Using AI to enhance scientific discourse by transforming journals into learning communities

Michael Rowe[1]

[1]School of Health and Care Sciences, University of Lincoln, Lincoln - UK

ABSTRACT

The introduction of generative AI into scientific publishing presents both opportunities and risks for the research ecosystem. While AI could enhance knowledge creation and streamline research processes, it may also amplify existing problems within the system. In this viewpoint article, I suggest that generative AI is likely to reinforce harmful processes unless scientific journals and editors use these technologies to transform themselves into vibrant knowledge communities that facilitate meaningful discourse and collaborative learning. I describe how AI could support this transformation by surfacing connections between researchers' work, making peer review more dialogic, enhancing post-publication discourse, and enabling multimodal knowledge translation. However, implementing this vision faces significant challenges, deeply rooted in the entrenched incentives of the current academic publishing system. Universities evaluate faculty based largely on publication counts, funding bodies rely on traditional metrics for grant decisions, and publishers benefit from maintaining existing models. Making meaningful change, therefore, requires coordinated action across multiple stakeholders who must be willing to accept short-term costs for long-term systemic benefits. The key to success lies in consistently returning to journals' core purpose: advancing scientific knowledge through thoughtful research and professional dialogue. By reimagining journals as AI-supported communities rather than metrics-driven repositories, we can better serve both the scientific community and the broader society it aims to benefit.

Keywords: Editor, Generative AI, Journal, Research Industrial complex, Scientific publication

Introduction

The introduction of generative AI into the research publication process presents unprecedented opportunities for journals and authors, as well as significant risks. While this technology could enhance the process of knowledge creation (1), generate novel hypotheses (2), and streamline the research process (3), it may also reinforce and amplify existing problems in scientific discourse and academic publishing (4).

One of the most fundamental of these problems is the conflict of interest between publishers and their journals, universities, and funding organizations. This ecosystem prioritizes metrics like publication counts and impact factors over meaningful contributions to knowledge creation, creating powerful incentives for researchers to focus on quantity over quality (4).

Received: December 19, 2024
Accepted: April 2, 2025
Published online: May 5, 2025

Study protocol number: N/A

Corresponding author:
Michael Rowe
email: mrowe@lincoln.ac.uk

This viewpoint article explores these risks, as well as the ways in which AI could be harnessed to improve scientific discourse. First, I examine how AI might exacerbate existing systemic problems in scientific publishing, intensifying the pressure for increased publication volume at the expense of meaningful contributions. Next, I discuss how AI could instead be used to transform journals into AI-supported learning communities, fostering more engaged and collaborative knowledge creation. Finally, I outline the challenges that must be addressed to implement these changes, including systemic barriers within academic publishing.

How a research industrial complex shapes scientific publishing

The concept of an "industrial complex"—first articulated by President Eisenhower in describing the military-industrial complex—helps illustrate the problematic dynamics in the academic publishing ecosystem. Just as the military-industrial complex creates self-perpetuating cycles that prioritize institutional interests over the public good, today's research publishing system has developed its own self-reinforcing patterns that might work against the interests of patients and practitioners (5). While the academic publishing ecosystem may not exactly mirror other industrial complexes, the analogy usefully highlights how interconnected institutional incentives can create self-perpetuating cycles that prioritize metrics over their purported missions. Just as defense contractors benefit from

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

continued military spending regardless of strategic outcomes, publishers, universities, and funding bodies have developed mutually reinforcing systems that reward publication volume and profit over meaningful scientific progress (6).

In the scientific community, this manifests in several important ways. The pressure to publish may lead researchers to make questionable choices about their use of AI to support scientific research and writing, where they adopt institutional metrics as being representative of their own values (6). In addition, the emphasis on novel findings can discourage crucial replication studies or research on routine clinical challenges that matter deeply to practitioners but may not generate exciting headlines. Again, the use of generative AI to articulate faux novelty in persuasive narratives (8) will influence what kinds of papers are published, as well as raise questions about which agendas are being privileged. And finally, the focus on quantitative publication metrics has blurred the boundary between clinically meaningful research, and research that serves, among other things, a journal's impact factor (7).

These dynamics are sustained by mutually reinforcing incentives across the research ecosystem. Universities evaluate faculty performance based largely on publication counts and journal impact factors. Funding bodies look to publication metrics when making grant decisions. Individual researchers need publications for career advancement, and journal publishers benefit from a steady stream of submissions. Each actor in this system is responding rationally to their incentives, but the collective result is a publishing culture that fails to serve the core mission of advancing scientific knowledge (8).

While there are limitations to the 'industrial complex' analogy (for example, publishers are not intentionally maintaining social and healthcare problems for continued profits), it nonetheless serves as a useful framework to explore the introduction of generative AI into the scientific process. There is already evidence that generative AI will enable researchers to produce papers at unprecedented speed—conducting analyses, drafting manuscripts, and responding to reviewer comments far faster than human researchers can do on their own (9). While this might seem like progress, it risks further divorcing research from careful reflection and clinical relevance. Without thoughtful guardrails, AI could amplify the volume-over-value tendency that the research industrial complex has triggered.

Mitigating the risks: reimagining journals as AI-supported communities

Scientific journals originally emerged as public records of conversations between scholars working on the most difficult problems of their time, serving as forums for ongoing dialogue (10). This original vision of journals as spaces for collaborative discourse stands in stark contrast to their current role as credentialing mechanisms within the research industrial complex. We might instead aim to recapture the original spirit of journals as platforms for scientific discourse, leveraging AI's capabilities to support conversation and dialog rather than using it to produce more papers with less value. The future of scientific publishing lies not in becoming more efficient article-processing platforms but in returning to the original purpose of journals: facilitating learning communities and discourse at the cutting edge of practice.

First, generative AI could help surface meaningful connections between different researchers' work that might otherwise go unnoticed. Instead of accelerating article production, AI might be used by journals and editors to analyze a submitted manuscript's methodology, findings, and theoretical framework to identify relevant ongoing discussions in the journal's community (11). For example, when a researcher submits a paper on a novel healthcare intervention, an AI system could not only flag related published papers but also highlight active discussions in other communities where clinicians are debating similar approaches or connect them with practitioners who have documented case studies in related areas. This shifts the focus from simply adding more to the literature toward joining an existing scholarly conversation, even when that conversation is happening elsewhere. It might also encourage authors and researchers to proactively connect to those existing discussions, embedding their work more effectively within those contexts. However, there is a risk that AI may prioritize surface-level keyword similarities over deeper methodological or theoretical coherence, leading to misleading associations. Editors and subject-matter experts should, therefore, verify AI-suggested connections, ensuring that flagged discussions align with the nuances of the research rather than being based on algorithmic correlations alone.

Second, AI could make peer review more dialogic rather than merely evaluative (or, in some cases, combative). Current large language models excel at asking probing follow-up questions and identifying unstated assumptions, which might help support all parties in a review process. This capability could be used to facilitate structured discussions between authors and reviewers, where the AI helps articulate points of confusion, surfaces potential counterarguments, and suggests areas where additional practitioner perspectives might be valuable as part of a multi-turn, long-context dialogue (12). Instead of reviewers (and, more likely, AI-based review systems) simply judging if a paper meets the publication criteria (13), the review process could facilitate a genuine dialogue about strengthening the work's contribution to practice while being supported by a critical yet constructive language model fine-tuned for peer review. While AI-supported review shows promise, there is nonetheless a concern that its suggestions might lack contextual sensitivity, leading to overly generic or misaligned critiques. There is also a risk that AI might introduce biases by disproportionately reflecting dominant perspectives embedded in its training data. Reviewers should, therefore, critically assess AI-generated feedback, using it as a starting point rather than a definitive judgment.

Third, AI could help make post-publication discourse more focused and productive. Rather than having fragmented discussions across unrelated journals, comment sections, social media, and separate response papers, AI could help organize and synthesize ongoing scholarly conversations about published work in a special section of the journal (14). It could identify key themes in practitioner responses,

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

highlight emerging consensus or disagreements, and help authors engage systematically with how their work is being interpreted and applied. This maintains the paper as a living document within an active community rather than a static artifact, where the AI-generated content is regenerated and updated in response to new publications and discussions from across the internet. However, we must also recognize that AI-driven content synthesis may inadvertently prioritize high-engagement topics over true scholarly debate, reinforcing existing narratives and failing to accurately capture nuanced disagreements. Editors and moderators should curate AI-generated summaries, ensuring a balanced representation of diverse perspectives. In addition, authors could verify and respond to AI-generated interpretations of their work.

And finally, generative AI's multimodal capabilities offer unique opportunities to make research more accessible and engaging across different contexts and audiences. Rather than limiting research dissemination to traditional PDF formats, journals could use AI to automatically generate alternative presentations of published work—from simplified explanations for practitioners and public audiences (15) to translations in multiple languages (16), to audio versions for listening while otherwise occupied (17), and eventually, video summaries that visualize key concepts and findings (18). This multimodal approach would help break down barriers to accessing scholarly work while enabling researchers to engage with the literature in ways that best suit their needs and preferences. However, the focus should remain on meaningful knowledge translation rather than simply multiplying formats—each alternative presentation should be thoughtfully designed to serve the journal's community and advance scholarly discourse. This reimagining of how research is shared could help journals fulfill their core purpose of facilitating learning and dialogue. While AI-generated summaries, translations, and visualizations can enhance accessibility, there is a risk of over-simplification, misinterpretation, or loss of critical nuance, particularly in technical or controversial subjects. Experts should review AI-generated content before dissemination, ensuring that translations retain accuracy and that generated content does not distort or misrepresent the original intent.

The key to these examples is to think of AI as a support for human-to-human scientific discourse and engagement rather than a means of producing more articles more quickly. This reimagining requires us to change our perspective of what a "journal" is. Rather than maintaining their gatekeeping and repository functions, scientific journals could become more dynamic spaces that foster ongoing dialogue between researchers, clinicians, educators, and society. These spaces would support multiple formats for sharing knowledge, with the aim of enhancing thoughtful human connection and understanding. But most importantly, they would redefine success through evidence of real-world impact rather than superficial citation counts. The value of AI is that it can help process and connect information at scale, enabling researchers to better embed their findings within broader contexts. The role of generative AI in scientific publishing should, therefore, be to support meaningful scholarly dialogue rather than obscure it with an avalanche of AI-generated content that serves only to feed the research industrial complex.

How to implement AI-supported learning communities in journals

The transformation of scientific journals from gatekeepers and content repositories into vibrant knowledge communities requires a fundamental reimagining of editorial roles and processes. This transformation represents not merely an adaptation to technological change but rather a return to journals' original purpose of facilitating learning and discourse in practice. In this section, several interconnected changes that journal editors might consider implementing as part of the journal transformation process are proposed.

Editors might reconceptualize peer review as an ongoing scholarly conversation, which takes as its starting point the practice of post-publication peer review (19). The traditional model of sequential rounds of peer review could be replaced with dynamic community dialogue, where standing review communities include generative AI agents, researchers, and practitioners who engage with manuscripts from submission through publication and beyond (13). Artificial intelligence can also support this transformation by identifying relevant reviewers based on expertise and interests, surfacing meaningful connections between submitted work and existing discussions in the community (12), and facilitating structured conversations by articulating points of confusion and potential counterarguments. This approach transforms peer review from a primarily evaluative process into a collaborative endeavor where AI acts as a mediator and guide, working alongside human reviewers to strengthen both the immediate work and its relevance to the broader scientific discourse.

The scope of publication formats must likewise expand beyond traditional research articles to accommodate a wider range of scholarly communication. Editors can create dedicated spaces for pre-publication discussion of protocols and methodologies, implementation experiences and case studies, practitioner perspectives, and ongoing discussions of published work. Again, AI can support this diversification by generating different versions of content tailored to various audiences while maintaining consistency, creating regularly updated topic-specific synthesis pages, and facilitating knowledge translation across languages and formats. These expanded formats acknowledge that scientific knowledge creation occurs through multiple channels and that different audiences may require different presentations of the same underlying research. This would also move us away from the idea that researchers should be rewarded per unit of publication and that engagement with transformed research findings may be more meaningful than simplistic publication metrics.

The role of editorial boards should also evolve away from primarily managing manuscript flow—which could be managed by dedicated AI agents—to actively cultivating the community. Board members could organize virtual journal clubs and themed discussions to connect researchers with relevant practitioners, educators, students, and other stakeholders, thereby creating opportunities for collaboration and sustained engagement. AI systems can assist by identifying emerging themes within the community that require deeper exploration, generating discussion summaries, connecting

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

participants with shared interests, and tracking ongoing conversations. This shift positions editorial board members as facilitators of scientific discourse rather than merely arbiters of publication decisions and "finders of peer-reviewers."

Fundamental to this transformation is the development of new impact metrics that better reflect journals' expanded role in the scientific community. Rather than relying primarily on citation counts and other quantitative metrics unrelated to real-world impact, journals should work with new stakeholders to create systems that track more authentic forms of research implementation, for example, changes in practice and community engagement across multiple platforms. Instead of focusing primarily on citation among academics as the principal measure of impact, editors could use AI to analyze a wide range of other signals of real-world engagement, tracking how research findings are influencing practice and identifying patterns in community engagement and implementation. These new metrics would provide a more nuanced and meaningful assessment of journal—and researcher—contribution to scientific progress. This approach would need to address potential challenges, such as the gaming of new metrics and difficulties in standardization across disciplines, and would require transparency in measurement methodologies.

Success in this transformation requires sustained commitment from the existing stakeholders in the scientific publishing ecosystem, as well as new stakeholders, including patients, service users, and members of society. Journal editors can lead the way in implementing these changes, but authors, reviewers, and practitioners must also adapt their approach to scientific communication. Authors and researchers should view publication not as the endpoint of research but as the beginning of a dialogue with their professional communities. Reviewers will need to embrace their role as contributors to ongoing scientific conversations rather than simply being called on to judge the quality of submitted manuscripts. Practitioners should actively engage in discussions about research priorities, methodology, and implementation, bringing their practical experience to bear on both the questions being asked and the ways that findings are translated into practice.

It should also be clear that the transition to AI-supported learning communities in scientific publishing does not necessitate an immediate or complete transformation of all journals. While some publications may choose to fully adopt this model, a more incremental approach is likely to be the most viable path forward. This could involve dedicating specific sections within journals to AI-supported discussions, such as post-publication forums, structured peer engagement spaces, or curated thematic dialogues facilitated by AI. These sections would operate alongside traditional journal functions, allowing for gradual integration and evaluation before broader adoption. By adopting a phased approach, journals can assess the effectiveness of AI-supported interactions while maintaining the integrity of their existing review and publication processes.

The integration of AI tools throughout this transformation must be guided by a clear understanding of their role: to enhance human networks and understanding rather than reinforcing the acceleration of article publication. By thoughtfully implementing these changes, journals and editors can better serve their communities while advancing scientific knowledge and improving clinical outcomes. This evolution represents not just an adaptation to this oncoming technological change but a return to the fundamental purpose of scientific journals as platforms for scientific discourse and professional development.

Navigating the transition to AI-supported learning communities

The transformation of scientific journals into knowledge communities faces significant systemic challenges deeply rooted in what I have been calling the research industrial complex. This interconnected system of incentives, metrics, and institutional practices and norms has created a self-perpetuating cycle that prioritizes publication volume over other, potentially more impactful, practices and activities. The introduction of generative AI into this ecosystem will only reinforce an accelerationist agenda that will quickly overwhelm scientific journals and their existing workflows and processes. Understanding and addressing these systemic constraints is important for any meaningful transformation of scientific publishing.

The most fundamental challenge lies in the deeply entrenched nature of current publication metrics within the broader academic ecosystem and the use of those metrics to inform promotion decisions, grant application success, and institutional reputation. Universities, funding bodies, and promotion committees rely heavily on quantitative metrics like publication count and journal impact factors, creating powerful institutional resistance to change (20). While individual journal editors might embrace the community-centered approaches suggested here, researchers—especially early-career academics—face strong systemic pressures to prioritize traditional publication metrics over meaningful engagement with their scientific communities. This path, reinforced by the introduction of generative AI, will further entrench misaligned incentives that might work against the substantive reform of scientific communication and publication.

Within the current metrics-driven system, authors will be pressured into using AI to accelerate article production and processing, further feeding the "publish or perish" mentality that already distorts academic incentives. Journal editors might respond to these AI advances by simply adapting their workflows to process more papers faster, seeing increased throughput as a competitive advantage relative to other journals. Given other challenges, like finding peer reviewers and managing the increased frequency of submissions, this risks creating an AI-driven paper mill (21) that serves career advancement metrics and publisher revenues while further degrading the quality of scientific discourse. The challenge for journal editors is to resist this pressure and instead harness AI's capabilities to deepen, rather than hasten, professional dialogue.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Overview of the AI-supported publishing model described in this paper

Current publishing system	AI-supported community model
Peer review process	
Primarily evaluative and gatekeeping, with reviewers judging manuscripts against publication criteria in isolation. Often slow, inconsistent, and lacking engagement between authors and reviewers.	AI facilitates dialogic review by identifying points of confusion, surfacing potential counterarguments, and connecting reviewers with appropriate expertise. Review becomes an ongoing scholarly conversation rather than a binary accept/reject decision.
Risks: Reviewing is viewed as a burden; bias and inconsistency in evaluation; limited opportunity for constructive dialogue.	*Risks*: AI suggestions lacking contextual sensitivity; potential to amplify biases from training data; over-reliance on automated feedback.
	Mitigation: Human reviewers critically assess AI-generated feedback, using it as a starting point rather than a definitive judgment. Editorial oversight ensures balanced representation.
Knowledge connections	
Research exists in silos with limited integration across studies, creating fragmentation where related work often goes unconnected.	AI analyses manuscripts to identify meaningful connections to existing work and ongoing discussions, embedding new research within broader contexts and conversations.
Risks: Information overload; researchers unable to keep up with expanding literature; redundant research.	*Risks*: Surface-level keyword matching rather than deeper methodological or theoretical connections; misleading associations.
	Mitigation: Editors and subject-matter experts verify AI-suggested connections, ensuring that flagged discussions align with research nuances.
Post-publication discourse	
Fragmented discussions across unrelated journals, comment sections, and social media. Limited engagement with published work after publication.	AI organizes and synthesizes ongoing scholarly conversations, identifies key themes in responses, and helps maintain papers as living documents within active communities.
Risks: Research treated as static artifacts; difficult to track impact beyond citations; limited practitioner engagement.	*Risks*: AI prioritizing high-engagement topics over scholarly debate, reinforcing existing narratives, and failing to capture nuanced disagreements.
	Mitigation: Editors curate AI-generated summaries to ensure balanced representation; authors verify and respond to AI-generated interpretations of their work.
Knowledge translation	
Research dissemination is limited to traditional PDF formats inaccessible to practitioners and public audiences.	AI generates alternative presentations tailored to different audiences - from simplified explanations to translations, audio versions, and visual summaries.
Risks: Research impact is limited by accessibility barriers and the disconnect between research and practice.	*Risks*: Oversimplification, misinterpretation, loss of critical nuance in technical or controversial subjects.
	Mitigation: Expert review of AI-generated content before dissemination, ensuring translations retain accuracy and don't misrepresent original intent.
Impact measurement	
Reliance on citation counts and journal impact factors, incentivizing quantity over quality and meaningful impact.	AI analyses multiple signals of real-world engagement and implementation, tracking how findings influence practice beyond academic citations.
Risks: Publication for career advancement rather than scientific progress, "salami slicing" of research, misaligned incentives.	*Risks*: Gaming of new metrics, difficulty standardizing impact measures across disciplines, resistance from established institutions.
	Mitigation: Multiple complementary metrics developed with stakeholder input; transparency in measurement methodologies; gradual transition to new evaluation standards.

The current publishing system generates significant revenue through traditional article processing charges and subscription models, creating financial disincentives for fundamental change. Maintaining active knowledge communities requires more sustained investment than simply processing manuscripts, yet traditional funding models do not currently support this shift. While AI could help manage some aspects of community engagement more efficiently, a

transition period would require significant investment in new infrastructure and processes, and individual researchers or journals attempting to break free from traditional metrics-driven publishing face significant risks.

However, there is hope. Forward-thinking editors and publishers could better harness the intellectual work of scientific publication that is currently freely given by engaged academics. Instead of volunteering their time for peer review in a model that serves mainly to increase the profit margins of already wealthy publishers, academics may be more likely to seek out opportunities to grow and develop as part of these new communities of learning centered around scientific journals.

This transformation is further complicated by the collective action problem - while most stakeholders might agree that the current system is sub-optimal (or actively harmful), no individual researcher, journal, or institution can safely step away from traditional metrics and publication practices while everyone else continues to use them. A researcher who focuses on 'building community' rather than increasing their publication rate risks being denied promotion. A journal that slows down to emphasize thoughtful discourse might see submissions decline. And universities that ignore traditional metrics could find themselves dropping in institutional rankings. Making meaningful change, therefore, requires coordinated action across multiple stakeholders, who must be willing to accept short-term costs for long-term systemic benefits. This coordination challenge helps explain why many previous attempts at reform have struggled to gain traction despite widespread recognition of the system's flaws. And while the system remains largely unchanged, I see the potential for generative AI to act as a forcing function.

Throughout all these challenges, the central question remains: how does the scientific community ensure that any reforms influenced by the integration of generative AI serve to enhance rather than merely accelerate scientific discourse? The answer lies in consistently returning to the core purpose of journals, which is to advance scientific knowledge and improve practice through thoughtful research and professional dialogue. This requires actively resisting the pressures of the research industrial complex while building new, AI-supported systems that better serve the needs of scientific communities and the broader areas of society they serve.

Conclusion

The rise of generative AI marks a critical juncture for scientific publishing. While AI could accelerate existing problems within the research industrial complex by enabling ever-faster article production, it also presents an opportunity to fundamentally reimagine scientific journals and the communities they serve. Rather than using AI to simply process more papers more quickly, we can use these technologies to transform journals from metrics-driven repositories of content into vibrant communities that facilitate meaningful discourse and collaborative learning. This transformation will require coordinated action across the scientific ecosystem, with authors, journals, institutions, and funding bodies working together to prioritize genuine scholarly dialogue over publication counts. The future of scientific publishing lies not in using AI to accelerate publication but in supporting the kind of thoughtful, community-driven discourse that advances both knowledge and practice.

This transformation may seem idealistic in the face of entrenched institutional incentives, but the alternative is to passively accept a system that prioritizes metrics over meaning. Rather than viewing these challenges as insurmountable, we should see them as an opportunity to deliberately shape the future of scientific discourse. By taking intentional steps to create an AI-supported publishing ecosystem that genuinely advances knowledge creation, we position ourselves as active architects of scientific progress rather than passive participants in an increasingly accelerated research industrial complex. This transformation will require coordinated action across the scientific ecosystem, with authors, journals, institutions, and funding bodies working together to prioritize genuine scholarly dialogue over publication counts. The future of scientific publishing lies not in using AI to accelerate publication but in supporting the kind of thoughtful, community-driven discourse that advances both knowledge and practice.

Acknowledgments

AI assistance disclosure: This article was written with the assistance of Claude, a large language model developed by Anthropic. The writing process involved multiple rounds of dialogue with the model, where initial ideas and arguments were iteratively refined through discussion. The model was used to help structure arguments, expand on key points, and improve clarity of expression. Some portions of the text were initially generated by Claude and then critically reviewed, edited, and integrated into the final manuscript. This collaborative process aimed to maintain scholarly rigor while leveraging AI capabilities to enhance the articulation of complex ideas. The author maintained full editorial control and responsibility for the final content, arguments, and conclusions presented in this article and is ultimately accountable for what is written. This disclosure aligns with what I believe are emerging best practices for transparency in AI-assisted academic writing.

Disclosures

Conflict of Interest: The author declares no conflicts of interest.

Financial support: The author received no specific funding or financial support for this work.

References

1. Khalifa M, Albadawy M. Using artificial intelligence in academic writing and research: an essential productivity tool. Comput Methods Programs Biomed Update. 2024;5:100145. CrossRef
2. Si C, Yang D, Hashimoto T. Can LLMs generate novel research ideas? A large-scale human study with 100+ NLP researchers. 2024. CrossRef
3. Bogin B, Yang K, Gupta S, et al. SUPER: evaluating agents on setting up and executing tasks from research repositories. 2024. CrossRef
4. Chu JSG, Evans JA. Slowed canonical progress in large fields of science. Proc Natl Acad Sci. 2021;118(41). CrossRef
5. Ezell JM. The health disparities research industrial complex. Soc Sci Med. 2024;351:116251. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

6. Nguyen C. Value capture. J Ethics Soc Philos. 2024;27(3). CrossRef
7. Callaway E. Beat it, impact factor! Publishing elite turns against controversial metric. *Nature*. 2016;535(7611):210-211. CrossRef
8. Edwards MA, Roy S. Academic research in the 21st century: maintaining scientific integrity in a climate of perverse incentives and hypercompetition. *Environ Eng Sci*. 2017;34(1):51-61. CrossRef PubMed
9. Noy S, Zhang W. Experimental evidence on the productivity effects of generative artificial intelligence. *Science*. 2023;381(6654):187-192. CrossRef PubMed
10. Baldwin M. Peer review. Encycl Hist Sci. 2020;4(1). CrossRef
11. Liu F, Xue S, Wu J, et al. Deep learning for community detection: progress, challenges and opportunities. In: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*. International Joint Conferences on Artificial Intelligence Organization; 2020:4981-4987. CrossRef
12. Tan C, Lyu D, Li S, et al. Peer review as a multi-turn and long-context dialogue with role-based interactions. 2024. CrossRef
13. Checco A, Bracciale L, Loreti P, et al. AI-assisted peer review. *Humanit Soc Sci Commun*. 2021;8(1):1-11. CrossRef PubMed
14. Jiang Y, Shao Y, Ma D, et al. Into the unknown unknowns: engaged human learning through participation in language model agent conversations. 2024. CrossRef
15. Ayre J, Mac O, McCaffery K, et al. New frontiers in health literacy: using ChatGPT to simplify health information for people in the community. *J Gen Intern Med*. 2024;39(4):573-577. CrossRef PubMed
16. Lion KC, Lin YH, Kim T. Artificial intelligence for language translation: the equity is in the details. *JAMA*. 2024;332(17):1427-1428. CrossRef PubMed
17. Dihan, Q.A., Nihalani, B.R., Tooley, A.A. *et al.* Eyes on Google's NotebookLM: using generative AI to create ophthalmology podcasts with a single click. *Eye* 39, 215–216 (2025). CrossRef
18. Liu Y, Zhang K, Li Y, et al. Sora: A Review on Background, Technology, Limitations, and Opportunities of Large Vision Models.2024. CrossRef
19. Hunter J. Post-publication peer review: opening up scientific conversation. *Front Comput Neurosci*. 2012;6:63. CrossRef PubMed
20. Lindner MD, Torralba KD, Khan NA. Scientific productivity: an exploratory study of metrics and incentives. *PLoS One*. 2018;13(4):e0195321. CrossRef PubMed
21. Eaton SE, Soulière M. Artificial intelligence (AI) and fake papers. COPE: Committee on Publication Ethics. 2023. Online (Accessed December 2024)

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Arch Physioter 2025; 15: 97-109
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3256
ORIGINAL RESEARCH ARTICLE

Differences in cervical mobility and pain sensitivity between patients with rotator cuff-related shoulder pain and asymptomatic subjects: a cross-sectional study

Néstor Requejo-Salinas[1,2], Rubén Fernández-Matías[3], Javier Gordo Herrera[1], Pablo Conesa López[1], Roy La Touche[1,4], Enrique Lluch-Girbés[5,6]

[1]Department of Physical Therapy, Superior Center for University Studies La Salle, Autonomous University of Madrid, Madrid - Spain
[2]International Doctoral School, Rey Juan Carlos University, Alcorcón - Spain
[3]Research Unit, Hospital Universitario Fundación Alcorcón, Alcorcón, Madrid - Spain
[4]Neuroscience and Craniofacial Pain Institute (INDCRAN), Madrid - Spain
[5]Department of Physical Therapy, University of Valencia, Valencian - Spain
[6]Pain in Motion Research Group, University of Valencia - Spain

ABSTRACT

Introduction: The cervical spine's role in shoulder pain remains unclear, particularly regarding its influence on shoulder pain severity and functional limitations. This study aimed to compare neck mobility, pain sensitivity, and strength between patients with rotator cuff-related shoulder pain (RCRSP) and asymptomatic controls and to explore associations between neck active range of motion (AROM) and shoulder outcomes.

Methods: A cross-sectional study was conducted with 50 patients with RCRSP and 50 asymptomatic controls. Neck AROM was measured with a CROM device, pressure pain thresholds (PPTs) with a digital algometer, and isometric neck strength with a handheld dynamometer. Shoulder pain and disability were assessed using the Visual Analog Scale (VAS) and the Shoulder Pain and Disability Index (SPADI).

Results: Compared to controls, patients with RCRSP exhibited reduced neck rotation toward the affected shoulder (mean difference: −5.19°; 95% CI: −8.84 to −1.38) and lower PPTs bilaterally (affected side: −1.49 kg/cm²; 95% CI: −1.99 to −1.00; unaffected side: −1.42 kg/cm²; 95% CI: −1.98 to −0.91). No differences were found in neck strength. Regression analysis showed that greater neck flexion, lateral flexion toward the affected side, and reduced protraction were associated with higher SPADI scores. Additionally, neck lateral flexion and rotation toward the affected side were negatively associated with shoulder pain intensity over the last week.

Conclusion: These findings suggest a potential interaction between the cervical spine and shoulder in RCRSP, underscoring the importance of a comprehensive assessment of both cervical and shoulder impairments in this condition.

Keywords: rotator cuff pain, neck, range of motion, pain threshold, muscle strength.

What is already known about this topic:

- *Previous research has suggested a relationship between the cervical spine and shoulder pain.*
- *Studies have reported alterations in the cervical region among individuals with shoulder pain compared to healthy controls. However, the specific nature of these impairments and their impact on shoulder function remain unclear due to limited research.*

What does the study add:

- *This cross-sectional study identifies a specific reduction in cervical rotation towards the affected shoulder and bilateral reductions in neck pain PPTs in patients with RCRSP, highlighting potential sensitization processes.*
- *Associations between specific cervical movements and shoulder outcomes, such as disability and pain intensity, provide new evidence of the interaction between cervical impairments and shoulder dysfunction.*
- *These findings underscore the importance of a comprehensive shoulder assessment that includes evaluating cervical spine impairments, which may contribute to shoulder dysfunction.*

Received: August 22, 2024
Accepted: March 24, 2025
Published online: May 12, 2025

This article includes supplementary material

Corresponding author:
Rubén Fernández-Matías
email: ruben.fernanmat@gmail.com

Introduction

Shoulder pain is a common complaint observed in the general population (1); the point prevalence of shoulder pain ranges from 6.9% to 26%, the 12-month prevalence from 4.7% to 55.2%, and the lifetime prevalence spans between 6.7% and 66.7% (1,2). Furthermore, shoulder pain results in high socioeconomic costs due to a significant decline in

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

the patient's ability to work and perform activities of daily living (3,4).

Among shoulder musculoskeletal pain conditions, rotator cuff disorders are the most common, affecting a total of 6.8-22.4% of individuals over 40 years old (5-7). Various terms have been suggested to describe pain originating in the rotator cuff, including subacromial pain syndrome (8). Rotator cuff disease (9) or rotator cuff-related shoulder pain (RCRSP) (10). The latter refers to a clinical presentation of shoulder pain and impaired shoulder movement and function, typically experienced during abduction and external rotation movements (10). It encompasses a range of conditions, including subacromial pain (impingement) syndrome, rotator cuff tendinopathy, and symptomatic partial and full-thickness rotator cuff tears (10).

The cervical spine has been considered a contributing factor or even the main source of pain in people with shoulder pain, including people with RCRSP (11). From a biomechanical perspective, adequate cervicothoracic mobility is an important requirement for proper shoulder complex function (12,13), and restricted mobility has been proposed as a risk factor for shoulder pain (14-17). In line with this, the Regional Interdependence model proposes that impairments in one region of the body can influence the musculoskeletal and neuromuscular function and symptoms in other remote regions of the body (18,19). In addition, this model states that neurophysiological, biopsychosocial, and somatovisceral systems can also influence musculoskeletal function both locally and at remote sites (19).

Regarding the link between the cervical spine and shoulder, individuals with shoulder pain may present with restricted cervical range of motion (ROM) (15,20-22). For example, Rebelatto et al. found that individuals with RCRSP had decreased active neck extension compared to healthy controls (22). Limited cervical ROM has even been proposed as a potential risk factor for developing sports-related shoulder injuries(16). On the other hand, proper activation of cervical muscles has been suggested to be necessary for upper limb tasks (23). In line with this, some studies have found altered activity patterns in the cervical flexor muscles in patients with shoulder pain (24,25). However, no significant differences have been found in overall maximal isometric neck strength between people with RCRSP compared to healthy individuals (22), although evidence on this topic is currently scarce. Moreover, research has identified variations in neck pain sensitivity, measured by pressure pain thresholds (PPTs), in people with shoulder pain compared to asymptomatic subjects. Some studies have reported heightened pain sensitivity on the symptomatic side of the neck (26), while others observed increased sensitivity bilaterally (on both the symptomatic and contralateral side), suggesting a broader sensitization process (22,27). However, other studies reported opposite findings (28,29), thus making it difficult to draw firm conclusions in this regard.

Given the inconsistencies and variability in previous findings, this study aimed to address key limitations in previous research, such as small sample sizes, lack of adjustment for confounding factors like age and sex, and limited exploration of cervical-shoulder relationships. By using a larger sample, employing robust statistical methods, and analyzing associations between cervical impairments and shoulder outcomes in greater depth, this study sought to provide more comprehensive insights into these interactions.

Specifically, this research aimed to analyze potential differences in various cervical sensorimotor variables, including neck active range of motion (AROM), neck muscle strength, and neck PPTs, between patients with RCRSP and asymptomatic controls. Additionally, possible associations between neck AROM and shoulder pain and disability in the RCRSP group were investigated. It was hypothesized that individuals with RCRSP would exhibit reduced neck mobility, increased neck pain sensitivity (i.e., reduced neck PPTs), and lower levels of neck muscle strength as compared to asymptomatic controls. Furthermore, reduced neck AROM would be expected to be associated with increased shoulder pain and disability in people with RCRSP.

Materials and methods

This study adhered to the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) guidelines (30) (Supplementary Material-1). Ethical approval was obtained from the Alcorcón Foundation Hospital Ethics Committee (HUFA) (Madrid, Spain). The study adhered to the Declaration of Helsinki principles (31).

Participants

Participants with RCRSP and asymptomatic individuals were recruited from three private physiotherapy clinics in the Community of Madrid (Spain). The recruitment period spanned from January 2022 to October 2023. Details of the study process can be found in Figure 1.

For the RCRSP group, inclusion criteria comprised: (i) subjects aged 18-65 years; (ii) unilateral shoulder pain lasting at least three months; (iii) pain reproduced during resisted muscle testing in shoulder abduction or external rotation that the participants identified as consistent with their usual shoulder pain (32); (iv) pain intensity ≥3 points on a Visual Analog Scale (VAS) and minimal or no pain at rest; and (v) familiar pain reproduced on at least three of the following five tests: Neer, Hawkins-Kennedy, Jobe, painful arc and external rotation resistance (33).

Exclusion criteria for the RCRSP group included: (i) neck pain within the last three months or a history of recurrent neck pain; (ii) history of prior shoulder surgery; (iii) signs or symptoms suggestive of radiculopathy, such as muscle weakness, hyporeflexia, or sensory changes; (iv) a positive Spurling or Arm Squeeze Test (34); (v) any systemic disease such as diabetes, fibromyalgia, rheumatoid arthritis, lupus, and/or neoplasms; (vi) shoulder pain of traumatic origin; (vii) passive shoulder external rotation ROM <45° or <50% compared to the contralateral shoulder, measured at 0° of shoulder abduction (35); (viii) positive sulcus sign for inferior instability and/or a positive drawer test and/or apprehension test for anterior and/or posterior instability (36); and (ix) current use of pain medication.

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

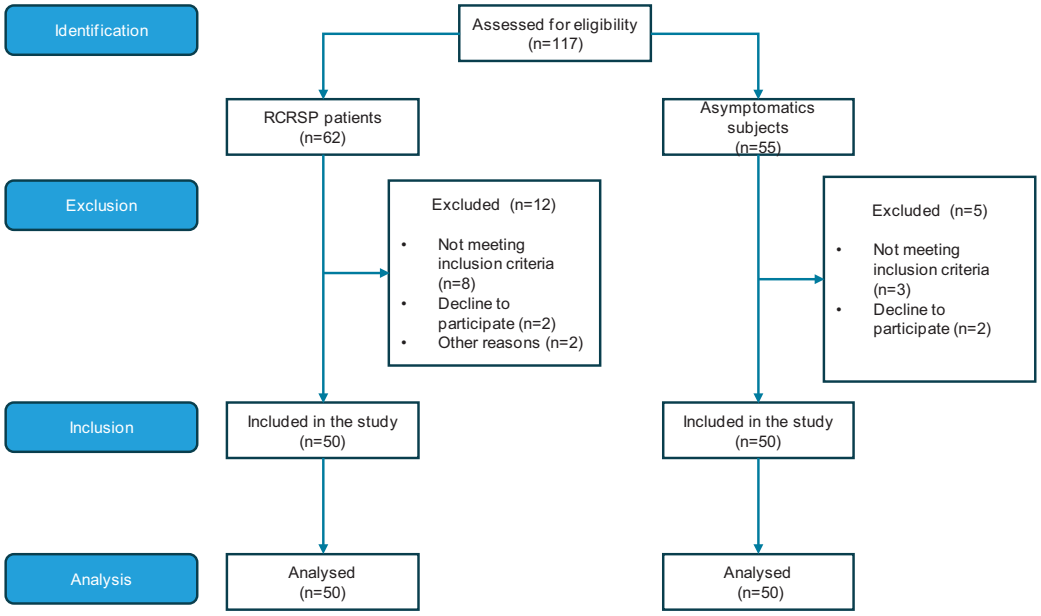


FIGURE 1 - Flow chart depicting the study process.

Asymptomatic controls were required to be aged 18-65 years. Exclusion criteria for this group included: (i) shoulder and/or neck pain in the last three months; (ii) neurological dysfunction of the upper limb; (iii) current use of pain medication; and (iv) previous history of shoulder surgery.

Sociodemographic data and psychological variables

Sociodemographic data, including sex, age, height, and weight, were recorded for each participant. In addition, psychological variables were also assessed to further characterize the study samples**.** Kinesiophobia was measured using the Tampa Scale of Kinesiophobia (TSK-11), which scores from 0 (less kinesiophobia) to 100 (more kinesiophobia) and has shown good reliability values (37). Catastrophizing levels were also measured using the Pain Catastrophizing Scale (PCS), which exhibits acceptable reliability values (38).

Shoulder outcomes

Pain intensity during the last week and current pain intensity were registered using a Visual Analog Scale (VAS). The VAS is presented as a 10-cm line anchored by verbal descriptors, usually “no pain” and “worst imaginable pain,” which has been demonstrated to be a reliable and valid pain measurement method (39). The minimal clinically important difference (MCID) for the VAS in people with rotator cuff disease is 1.37 cm (p = 0.0255) (40).

Shoulder disability was assessed using the Spanish-validated version of the Shoulder Pain and Disability Index (SPADI) (41). The SPADI is a self-administered questionnaire scoring from 0 (no disability) to 100 (maximum disability) (42). It has a reported test-retest reliability of 0.89-0.93 and a minimal detectable change (MDC95) of 12.2% (43).

Neck outcomes

Neck AROM was assessed using a CROM© device (Performance Attainment Associates, Roseville, MN; Fig. 2).

This device has demonstrated high intra- and inter-examiner reliability for measuring angular movements (ICC = 0.68-0.95; ICC = 0.79-0.99, respectively) and good validity (r = 0.93-0.98) (44). Furthermore, the CROM device has small measurement errors (SEMs ranging from 1.6° to 2.8° and MDCs ranging from 3.6° to 6.5°) (44). Additionally, the CROM© device allows us to measure the position of the head and its displacement in the sagittal plane from maximum protraction to maximum retraction, with excellent intra-rater reliability (ICC = 0.94-0.98), inter-rater reliability (ICC = 0.73-0.98) and an adequate validity [(r) = 0.47-0.78] (45). The following neck movements were assessed: flexion, extension, lateral flexion, rotation, protraction, and retraction. For each movement, three measurements were taken and averaged.

Maximal isometric neck strength was measured using a handheld dynamometer (MicroFET 2©, Hoggan Health Industries, West Jordan, UT, USA; Fig. 3), which has shown excellent intra- and inter-examiner reliability (ICC=0.73-0.95 and ICC = 0.78-0.86) (46-48), with a SEM of 0.59-0.87 Kgf (46) and a MDC of 1.62-4.85 in healthy patients (46, 47).

Maximal voluntary isometric contractions (MVICs) were assessed for neck flexion, extension, and lateral flexion. Participants performed three MVICs for each movement, each held for 5 seconds, with 30 seconds of rest between repetitions. Flexion strength was tested in a supine position, extension in a prone position, and lateral flexion in a side-lying position. The dynamometer was placed against the participant's forehead (flexion), occiput (extension), and temporal region (lateral flexion) to ensure standardized placement. Participants were instructed to generate maximal force against the device without compensatory movements, which were minimized using a

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

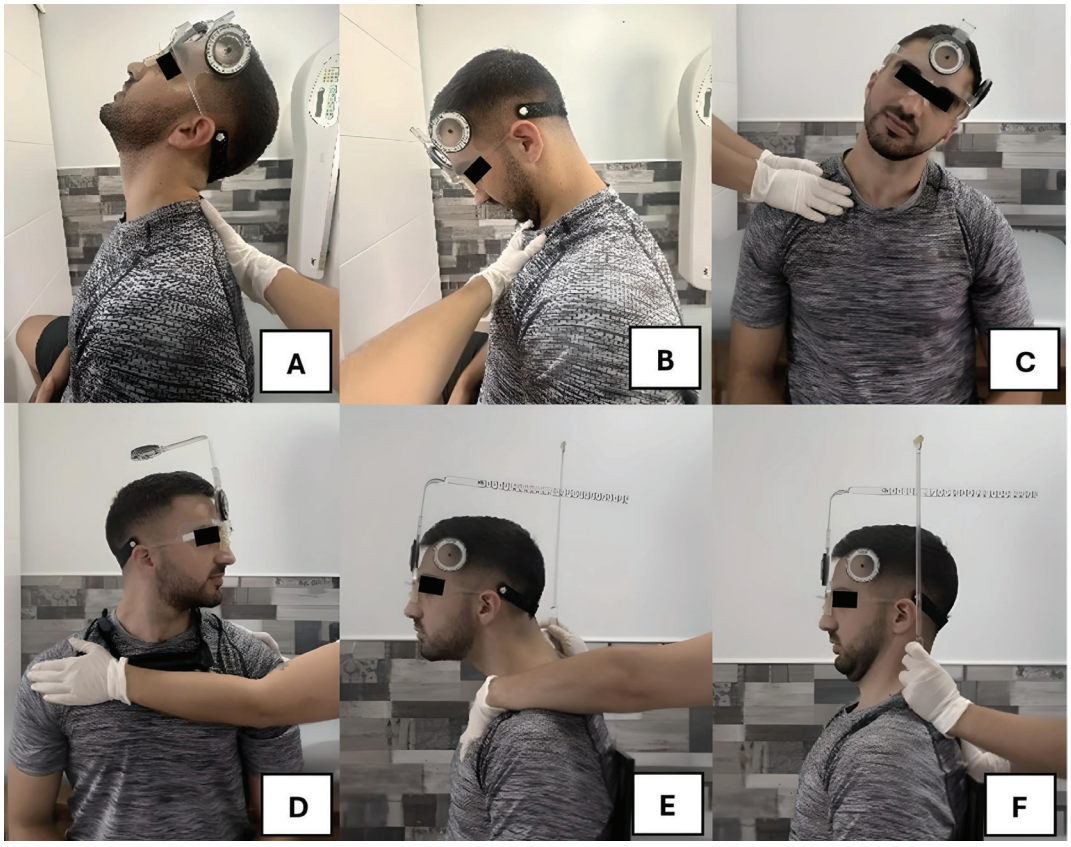


FIGURE 2 - Cervical AROM assessment. A: extension, B: flexion, C: lateral flexion, D: rotation, E: protraction, F: retraction.

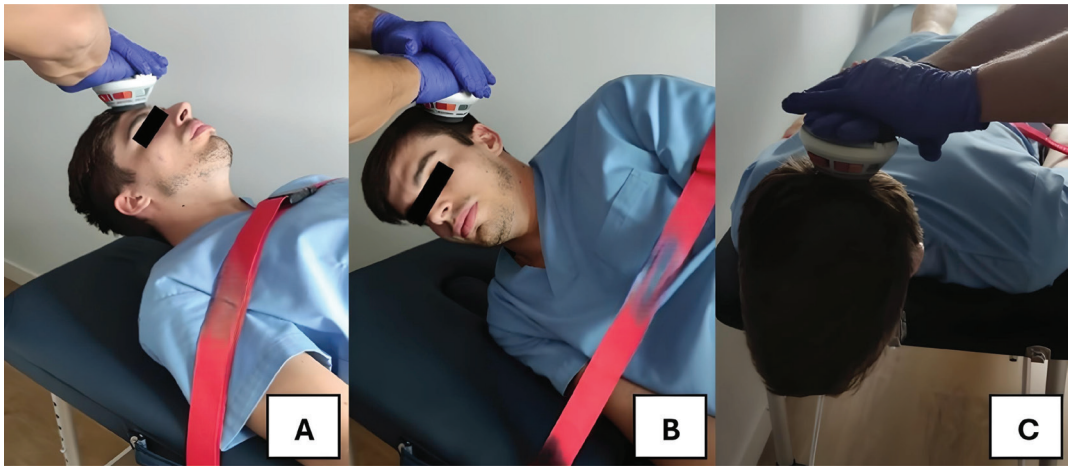


FIGURE 3 - Maximum isometric cervical strength assessment. A: flexion, B: lateral flexion, C: extension.

rigid strap secured at the T3 vertebra level. Verbal encouragement was provided to ensure maximum effort. The average of the three repetitions was used for analysis. An adjustable pillow was placed under the head during side-lying tests when necessary to maintain a neutral starting position (46).

Neck PPTs were assessed with a digital algometer Wagner FDX-25 (Wagner Instruments, Greenwich, CT) with a 1 cm^2 probe. Algometry is a reliable method for assessing neck PPTs, with high intra- and inter-examiner reliability (ICC = 0.97 and 0.73-0.91), an SEM of 0.18-0.64 kg/cm^2 (49,50) and an MDC of 0.43 kg/cm^2 (50) in asymptomatic individuals. The algometer was positioned perpendicularly to the participant's skin. Pressure was then gradually increased at a rate of approximately 0.5 kg/s until pressure was reported as painful. Three consecutive measurements were obtained with a 30-second rest period between repetitions. Neck PPTs were measured bilaterally on the C5-C6 zygapophyseal joints due to their segmental relationship with the shoulder (28).

The order of all measurements, including AROM, neck muscle strength, and PPTs, was randomized to minimize potential bias and ensure consistency across participants. All measurements were conducted by two trained assessors using standardized protocols.

In the asymptomatic group, all bilateral neck measurements (e.g., PPTs, AROM in rotation) were performed on one randomly selected side. As asymptomatic participants do not have a "symptomatic" or "non-symptomatic" side, this approach allowed meaningful comparisons with the corresponding sides in the symptomatic group while maintaining consistency in the analysis.

Sample size calculation

The sample size calculation was based on the expected width of the 95% confidence interval for the adjusted mean difference from a multivariable linear regression model in neck rotation for the covariates age, height, weight, and sex, with an assumed 80% power, following Kelley and Maxwell's proposal (51). Calculations were performed using the "ss.aipe.reg.coef" function in the R package 'MBESS' (Ken Kelley (2022). MBESS: The MBESS R Package. R package version 4.9.2).

A correlation of 0.30 was assumed between the covariates, and a biserial-point-biserial correlation coefficient value of 0.24 was assumed between the pain factor and the covariates. A biserial-point-biserial coefficient was also assumed between the pain factor and the dependent variable (neck rotation ROM) 0.24, as well as a correlation of 0.30 between the covariates and the dependent variable. Finally, a standard deviation of 7º for neck rotation ROM was estimated based on previous published literature (52). A total width of the 95% confidence interval of 6º (±3º accuracy) was considered acceptable. With these data, the necessary sample size was determined to be 99 subjects, which was rounded up to 100 (50 subjects with RCRSP and 50 asymptomatic controls) to maintain a 1:1 ratio in the pain factor.

Statistical analysis

All statistical analyses were conducted using the software R v.4.1.0 (R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria). The width of the confidence interval for all analyses was set a priori to 95%.

The intra-rater reliability of the measurement procedures was evaluated using an intraclass correlation coefficient with an absolute agreement ($ICC_{2,1}$) with the R package "irrICC."

For the descriptive analysis of quantitative variables, the mean, standard deviation (SD), median, first and third quartiles, and minimum and maximum values were reported. For categorical variables, the absolute frequencies and percentages were reported. Furthermore, the correlation matrix between all quantitative variables, histograms, and Q-Q plots was also reported, as well as measures of kurtosis and skewness, aiming to evaluate the distribution of the data.

For the analysis of between-group differences on outcome measures, an ordinary least squares regression was used to estimate adjusted between-group differences for age, sex, height, and weight, using the package "rms" (Harrell, 2022). A percentile bootstrap procedure with 5,000 samples was used to calculate 95% confidence intervals for each mean difference using the package "boot" (Canty and Ripley, 2022). This type of analysis is a better choice than the traditional t-test because it allows for the estimation of between-group differences, adjusting for the confounding effects of multiple covariates and providing a better understanding of the relationship between shoulder pain and the outcome measures. These adjusted between-group differences are interpreted as mean differences between shoulder pain and healthy subjects, assuming both groups have the same values in the covariates (i.e., eliminating their influence in the observed differences).

For the analysis of the relationship between cervical ROM measures (adjusted for age, sex, height, and weight) and shoulder disability measured with the SPADI as well as shoulder pain intensity during the last week, two beta regression models were conducted using the package "betareg" (53). These factors were adjusted due to their potential influence on cervical ROM (54-56). Beta regression was selected because this model can cope with bounded variables, which present with heteroskedasticity and usually have non-linear relationships with other variables. In addition, beta regression prevents the predicted values and the upper and lower bounds of the confidence intervals from falling outside the bounding limits of the outcome variable.

The effect size of the predictor variables was presented as the odds ratio for the outcome measure as per unit change in the predictor variable, with a value of 1 meaning no association, a value >1 meaning a positive association, and a value <1 meaning a negative association. In addition, diagnostic plots were also constructed for each model to evaluate their adequacy. Finally, plots of significant predictor values versus fitted values (with 95% percentile predictor intervals), holding all other predictors constant (and sex set to female), were also constructed to facilitate the interpretation of the significant relationships observed.

Results

The final sample was composed of 50 subjects with RCRSP (26 male) with a mean age of 41.1 (SD, 13.80) years, and a mean body mass index of 24.99 (SD, 3.28), and 50 healthy controls (26 male) with a mean age of 36.76 (SD, 13.36) years, and a mean body mass index of 25.23 (SD, 3.74) (Fig. 1). The characteristics of participants are presented in Table 1, and the full descriptive statistics, including median, quartiles, and measures of kurtosis and skewness, are presented in Supplementary Material 2. Furthermore, the normal Q-Q plots are presented in Supplementary Material 3, and the histograms in Supplementary Material 4.

The intra-rater reliability of the range of motion measurements was good for healthy subjects, with an $ICC_{2,1}$ value ranging from 0.838 to 0.968, and for pain subjects, with values ranging from 0.853 to 0.972. Regarding strength measurements, it was also good, with an $ICC_{2,1}$ value ranging from 0.973 to 0.981 for healthy subjects and from 0.97 to 0.982 for pain subjects. Finally, reliability was also good for PPT measurements, ranging from 0.91 to 0.93 for healthy subjects and from 0.916 to 0.944 for pain subjects.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 1 - Characteristics of participants

Variable	N	Mean	SD	Median	N	Mean	SD	Median
Group		Healthy group				Shoulder pain group		
Sex	50				50			
Male	26	52%			26	52%		
Female	24	48%			24	48%		
Age, years	50	36.76	13.36	34	50	41.1	13.80	41.5
Height, cm	50	170.6	9.465	170	50	169.98	9.55	169
Weight, kg	50	73.92	14.58	74.5	50	71.32	14.56	69.6
BMI, kg/m^2	50	25.23	3.74	25.2	36	24.99	3.28	25.55
Time with pain, days					50	469	647.76	182.5
Dominant side	50				50			
... Left	0	0%			7	14%		
... Right	50	100%			43	86%		
Affected side	0				50			
... Left					25	50%		
... Right					25	50%		
Previous shoulder pain (beyond the last three months)	50				50			
... No	45	90%			39	78%		
... Yes	5	10%			11	22%		
Previous neck pain (outside the last three months)	50				50			
... No	42	84%			35	70%		
... Yes	8	16%			15	30%		
Pain (actual)					50	3.95	2.55	3
Pain (last week)					50	4.85	1.86	4.2
SPADI (pain)					50	46.16	22.63	43
SPADI (disability)					50	26.8	21.9	24.38
SPADI (total)					50	32.738	19.36	32.31

Abbreviations: SD, standard deviation; BMI, body mass index; SPADI, Shoulder Pain and Disability Index.

Between-group differences in neck AROM

The descriptive statistics of all neck AROM outcome measures and the adjusted mean differences are presented in Table 2.

The adjusted analyses revealed a significant between-group difference in neck rotation, with the RCRSP group showing less neck rotation towards the affected shoulder (mean difference, −5.19°; 95% CI, −8.84° to −1.38°) compared to asymptomatic controls. There were no significant differences in the other neck AROM measures (Table 2).

Between-group differences in neck muscle strength

Regarding neck muscle strength, there were no significant adjusted mean between-group differences (Table 2).

Between-group differences in neck PPTs

The RCRSP group showed greater pain sensitivity in the cervical spine (i.e., lower neck PPTs) both in the side of the affected shoulder (mean difference, −1.49 kg/cm2; 95% CI, −1.99 to −1.00) and unaffected shoulder (mean difference, −1.42 kg/cm2; 95% CI, −1.98 to −0.91) compared to asymptomatic controls.

Relationship between neck AROM and shoulder pain and disability in the RCRSP group

The full model statistics and coefficients for each predictor are presented in Tables 3 and 4, and the diagnostic plots are presented in Supplementary Material 5.

TABLE 2 - Between-group differences

Variable	Healthy (n = 50), mean (SD)	Shoulder pain (n = 50), mean (SD)	Adjusted mean difference[#] (95% CI)[£]
Neck range of motion, degrees			
Neck protraction	5.50 (1.36)	4.94 (1.79)	-0.45 (-1.04 to 0.11)
Neck retraction	3.44 (0.99)	3.57 (1.13)	0.24 (-0.18 to 0.62)
Neck rotation (affected side)	71.74 (10.03)	64.95 (10.57)	-5.19 (-8.84 to -1.38) *
Neck rotation (unaffected side)		67.13 (11.52)	-2.90 (-6.50 to 1.01)
Neck lateral flexion (affected side)	41.03 (9.22)	40.60 (9.38)	1.13 (-1.83 to 4.34)
Neck lateral flexion (unaffected side)		39.89 (9.20)	-0.03 (-3.19 to 3.22)
Neck flexion	58.15 (10.80)	56.59 (11.53)	-0.97 (-5.16 to 3.09)
Neck extension	71.04 (12.51)	66.20 (15.53)	-2.81 (-7.83 to 1.91)
Neck strength, kgf			
Neck lateral flexion (affected side)	12.28 (5.90)	12.34 (6.41)	0.95 (-0.81 to 2.66)
Neck lateral flexion (unaffected side)		12.42 (6.75)	1.01 (-0.70 to 2.75)
Neck flexion	10.55 (5.66)	9.52 (5.56)	-0.08 (-1.66 to 1.41)
Neck extension	15.16 (6.23)	14.47 (6.83)	0.41 (-1.38 to 2.21)
Mechanosensitivity and psychological outcomes			
PPT (affected side), kg/cm^2	3.35 (1.63)	1.90 (1.01)	-1.49 (-1.99 to -1.00) *
PPT (unaffected side), kg/cm^2		1.97 (1.16)	-1.42 (-1.98 to -0.91) *
PCS	8.24 (9.66)	9.68 (9.06)	1.17 (-2.72 to 5.03)
TSK-11	19.14 (4.42)	23.42 (6.78)	3.80 (1.55 to 6.22) *

#Adjusted for age, height, weight, and sex. £Confidence intervals based on a percentile bootstrap procedure with 5,000 samples. Abbreviations: SD, standard deviation; CI, confidence interval; PPT, pressure pain threshold; PCS, Pain Catastrophizing Scale; TSK, Tampa Scale for Kinesiophobia. * $p < 0.05$

The evaluation of diagnostic statistics and plots of the first model, with the SPADI as the dependent variable, revealed two influential subjects (namely 2 and 28) with Cook's distance above 0.60. The estimated pseudo-R2 of this model was 0.36. The model was developed again without these two influential values, which improved the diagnostic plots (Supplementary Material 5) and revealed an increase in the pseudo-R2 value up to 0.47 (Table 3). This model showed a significant regression coefficient for age (OR per unit change = 1.03), meaning that for each increment in one unit of age, there is an estimated odds ratio of 1.03 for the expected increment in SPADI, neck flexion (OR per unit change = 1.03), neck protraction (OR per unit change = 0.85), and neck lateral flexion towards the side of the affected shoulder (OR per unit change = 1.04)

Since no relevant issues were identified in the examination of diagnostic plots (Supplementary Material 5), the full sample was used to develop a model for shoulder pain intensity during the last week. This model explained 33% of the variance in shoulder pain intensity during the last week, as measured by a pseudo-R2 value of 0.33. There were significant regression coefficients for age (OR per unit change = 1.02), weight (OR per unit change = 1.02), neck lateral flexion towards the affected shoulder (OR per unit change = 1.04), and neck rotation towards the affected shoulder (OR per unit change = 0.98) (Table 4).

The predicted value of SPADI and pain intensity as a function of the significant predictor parameters holding all other predictors constant at the mean are presented in Supplementary Material 6, aiming to improve the interpretation of the significant relationships observed.

Discussion

This observational study investigated differences in cervical sensorimotor variables between individuals with RCRSP and asymptomatic controls, including neck AROM, neck muscle strength, and neck PPTs. Potential associations between cervical AROM and shoulder pain and disability in individuals with RCRSP were also explored. Our results showed that people with RCRSP had a bilateral increase in pain sensitivity in the cervical region, as reflected by lower bilateral neck PPTs and reduced neck active rotation towards the affected shoulder compared to asymptomatic controls. Furthermore, significant associations were found between neck AROM in different directions (neck flexion, protraction, and lateral flexion toward the affected shoulder) and SPADI scores in the RCRSP group. Additionally, neck lateral flexion and rotation

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 3 - Beta regression for Shoulder Pain and Disability Index

Parameter	Regression (B) coefficient (95% CI)	OR
Intercept	−5.04 (−11.72 to 1.63)	0.01
Age	0.03 (0.01 to 0.05) *	1.03
Weight	0.02 (−0.01 to 0.04)	1.02
Height	0.01 (−0.04 to 0.05)	1.01
Sex (male = 1)	0.66 (−0.05 to 1.38)	1.94
Neck flexion	0.03 (0.01 to 0.05) *	1.03
Neck extension	−0.001 (−0.02 to 0.02)	1.00
Neck protraction	−0.16 (−0.29 to -0.03) *	0.85
Neck retraction	−0.11 (−0.34 to 0.12)	0.90
Neck lateral flexion (affected side)	0.04 (0.01 to 0.07) *	1.04
Neck lateral flexion (unaffected side)	−0.013 (−0.05 to 0.03)	0.97
Neck rotation (affected side)	−0.03 (−0.06 to 0.003)	0.99
Neck rotation (unaffected side)	0.01 (−0.02 to 0.04)	1.01

Standardized weighted residuals 2:

Min	1Q	Median	3Q	Max
−2.04	−0.89	−0.04	0.91	2.54

Phi coefficient (precision model with identity link):

	Estimate	Std. Error	z value	95% CI
(phi)	11.94	2.38	5.03	7.29 to 16.60

Type of estimator: maximum likelihood
Log-likelihood: 34.4 on 14 degrees of freedom
Pseudo R-squared: 0.474

Abbreviations: CI, confidence interval; 1Q, first quartile; 3Q, third quartile. OR refers to the odds ratio for the outcome measure per unit change in the predictor variable. * $p < 0.05$

towards the affected shoulder were associated with self-reported pain intensity during the last week in this group.

Neck AROM

Our study revealed a significant decrease in neck rotation towards the affected shoulder among patients with RCRSP compared to asymptomatic controls. This finding seems to align with previous studies where limitations in cervical ROM were reported in individuals with shoulder pain (15,20,22). However, Rebelatto et al. found a substantial reduction in cervical extension AROM, but not rotation, in patients with RCRSP, compared to controls (22). These differences may be due to several various factors, including the methodology used to select the neck side for neck AROM assessment in healthy subjects (dominance matched to their symptomatic counterparts in Rebelatto et al. study versus random selection in our case), differing inclusion criteria such as the duration of symptoms (4 weeks vs. 3 months) or pain intensity levels for the RCRSP group.

Cervical radiculopathy/radicular pain might be a contributing factor to shoulder pain in some patients (57). Cervical rotation and extension can narrow the intervertebral foramina (58,59), so limitations in these movements in people with shoulder pain, including RCRSP, as reported in Rebelatto and colleagues' study and ours, might indicate a protective mechanism to prevent further nerve compression in the presence of a subclinical cervical nerve root problem. Indeed, limited cervical ipsilateral rotation has been included as a diagnostic criterion for cervical radiculopathy (60).

Alternatively, these restrictions in AROM might represent a motor adaptation or avoidance strategy resulting from pain, which is a common phenomenon in musculoskeletal conditions (61). Furthermore, psychological factors such as fear of movement (kinesiophobia) may also influence these findings. In our study, patients with RCRSP exhibited significantly higher levels of kinesiophobia compared to asymptomatic participants, as reflected by their TSK-11 scores (mean difference: 3.80; 95% CI: 1.55-6.22). This aligns with the fear-avoidance model, which proposes that pain-related fear can result in reduced mobility and activity (62). Moreover, previous research has shown that psychological factors, including kinesiophobia, can significantly affect joint range of motion in musculoskeletal pain conditions (63).

Neck muscle strength

Few studies have examined the role of impaired function of cervical muscles in people with shoulder pain (22,24,25). We did not find significant differences in neck muscle strength between individuals with RCRSP and asymptomatic subjects, which is in accordance with Rebelatto et al.'s findings (22). However, neither of the two studies assessed maximal

TABLE 4 - Beta regression for pain intensity within the last week

Parameter	Regression (B) coefficient (95% CI)	OR
Intercept	–0.18 (–6.43 to 6.07)	0.84
Age	0.02 (0.003 to 0.04) *	1.02
Weight	0.02 (0.001 to 0.05) *	1.02
Height	–0.02 (–0.06 to 0.02)	0.98
Sex (male = 1)	0.40 (–0.24 to 1.05)	1.50
Neck flexion	0.02 (–0.004 to 0.04)	1.02
Neck extension	0.01 (–0.01 to 0.02)	1.01
Neck protraction	0.001 (–0.13 to 0.13)	1.00
Neck retraction	0.10 (–0.10 to 0.29)	1.10
Neck lateral flexion (affected side)	0.04 (0.01 to 0.07) *	1.04
Neck lateral flexion (unaffected side)	–0.02 (–0.05 to 0.01)	0.97
Neck rotation (affected side)	–0.03 (–0.05 to -0.001) *	0.98
Neck rotation (unaffected side)	0.01 (–0.02 to 0.03)	1.01

Standardized weighted residuals 2:

Min	1Q	Median	3Q	Max
-2.33	-0.82	-0.12	0.77	2.13

Phi coefficient (precision model with identity link):

	Estimate	Std. Error	z value	95% CI
(phi)	9.95	1.90	5.23	6.22 to 13.68

Type of estimator: maximum likelihood
Log-likelihood: 25.39 on 14 degrees of freedom
Pseudo R-squared: 0.33

Abbreviations: CI, confidence interval; 1Q, first quartile; 3Q, third quartile. OR refers to the odds ratio for the outcome measure per unit change in the predictor variable. * $p < 0.05$

isometric rotational strength, which could be important considering the differences in neck rotation found in our study.

Beyond maximal isometric neck strength, evaluating the specific function of the deep cervical musculature in people with RCRSP may warrant further investigation. Some studies have suggested a potential link between the function of the deep cervical muscles and upper limb (23-25). Therefore, analyzing the function of the deep cervical muscles by means of tests such as the craniocervical flexion test or the deep neck flexors endurance test in people with RCRSP may offer a more comprehensive understanding of cervical muscle (dys) function in this population (64).

Neck PPTs

Studies examining neck PPTs in patients with shoulder pain have yielded inconsistent results (22,25-28). Similar to our study, Rebelatto et al. found decreased PPT values measured in the upper trapezius in patients with RCRSP compared to healthy controls on both affected and unaffected sides. However, we observed larger differences in PPT values between patients and asymptomatic controls (−1.41 kg/cm² vs. −0.4 kg/cm² on the affected side, −1.52 kg/cm² vs. −0.66 and 0.4 kg/cm² on the unaffected side). Furthermore, reduced PPTs have been identified bilaterally at the C5-C6 zygapophyseal joint in male wheelchair basketball players with unilateral shoulder pain compared to asymptomatic basketball players (with or without wheelchairs) (27). Another study by Hidalgo-Lozano et al. reported lower neck PPTs in the levator scapulae of subjects with Subacromial Impingement Syndrome (SIS) compared to healthy subjects (26). However, this study only included right-handed subjects with dominant side involvement, and measurements were taken solely on that side. Additionally, the study had a small sample size (10 healthy subjects and 12 patients with SIS).

Hidalgo-Lozano et al. also conducted a similar study comparing PPTs on the dominant side of the levator scapulae, upper trapezius, sternocleidomastoid, and anterior scalene muscles in elite swimmers with and without shoulder pain, as well as asymptomatic elite athletes (25). The results indicated reduced PPT levels in swimmers with shoulder pain compared to elite athletes across all muscles. However, no significant differences were found between elite swimmers with and without shoulder pain. Other studies have reported no significant differences in neck PPTs measured over the upper trapezius, levator scapulae, and articular pillar of the C5-C6 zygapophyseal joint when comparing the affected and unaffected sides of patients with SIS to the dominant side of healthy controls (28).

Several factors could contribute to the observed inconsistencies in neck PPT results in people with shoulder pain, including RCRSP, such as variations in inclusion criteria, sample size, measurement site (e.g., upper trapezius vs. C5-C6 zygapophyseal joints), or failure to adjust for potential confounders.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

The bilateral increase in pain sensitivity observed within the cervical region of RCRSP patients in this study may reflect both peripheral and central sensitization processes (65). While the role of central nervous system sensitization in shoulder pain remains inconsistent (66), our findings of widespread mechanical hyperalgesia in the neck region suggest potential alterations in central pain processing among RCRSP patients. The inclusion of a remote PPT measurement (e.g., tibial anterior) may have provided a better opportunity to investigate the role of centrally-mediated mechanisms.

Hyperactivity of the upper trapezius muscle, which has been reported in patients with RCRSP (67), may also contribute to reduced PPTs in this population. Excessive activation of this muscle could lead to increased nociceptive input or myofascial pain syndromes, which are known to reduce pain pressure thresholds (68). Additionally, psychological factors, such as fear-avoidance beliefs and kinesiophobia, may further exacerbate pain sensitivity in this population (69).

Associations between neck AROM and shoulder pain and disability

The results suggest that neck mobility may be closely linked to shoulder function, reinforcing the concept of an interplay between these regions. Arm elevation requires adequate cervicothoracic mobility to ensure efficient scapular and shoulder mechanics (12,13). A lack of mobility in this region could impair the functionality of the rotator cuff and scapular muscles, potentially increasing the risk of shoulder dysfunction (15,70). However, given the modest strength of the associations observed, these results should be interpreted with caution.

Implications for clinical practice and future research

This study highlights the importance of evaluating the cervical spine in individuals diagnosed with RCRSP. The assessment of neck AROM may complement clinical evaluation when used alongside validated questionnaires for assessing shoulder pain and disability levels. Additionally, reduced cervical rotation towards the side of the shoulder with RCRSP may indicate a contribution of the cervical spine in the patient's symptoms (e.g., subclinical cervical nerve root problem). As no standardized protocols for assessing the cervical spine in patients with shoulder pain, including people with RCRSP are not available (71), future research should work on this aim.

Further research is needed to clarify the relationship between cervicothoracic mobility and shoulder pain and function. Longitudinal studies in initially asymptomatic populations, assessing cervical mobility prior to the onset of shoulder pain, could help determine whether cervical impairments are a cause or consequence of shoulder pain. Predictive models could also identify key factors, including cervicothoracic mobility, associated with the onset or persistence of shoulder pain. Interventional studies are warranted to evaluate the effects of improving cervicothoracic mobility on pain and functional outcomes in patients with RCRSP.

Moreover, investigating the role of cervical motor control and deep cervical musculature in RCRSP could provide valuable insights. As discussed earlier, targeted tests, such as the cranio-cervical flexion test and the deep neck flexor endurance test, may enhance our understanding of cervical muscle dysfunction and its impact on shoulder pain and function in this population.

Although the present study did not assess the interrelationships between cervical ROM, muscle strength, and pain sensitivity, this remains an important area for future research. Given the complexity of sensorimotor interactions in the cervical region, studies should explore how these variables influence one another and whether specific impairments contribute to overall dysfunction in individuals with RCRSP. Such investigations would require larger sample sizes and advanced statistical models, such as mediation or moderation analyses, to provide a more comprehensive understanding of these interactions.

Study limitations

Several limitations should be considered when interpreting the results of this study. First, assessors were not blinded to participants' pain status, potentially introducing bias and reducing the study's internal validity. Second, the levels of participants' physical activity were not assessed, which could have influenced the observed differences in PPTs (72). Third, the absence of a neck disability questionnaire (e.g., Neck Disability Index) may have led to the inclusion of asymptomatic individuals with functional limitations in the neck despite not reporting neck pain, potentially affecting the results (70). Additionally, although participants with shoulder or neck pain within the last three months were excluded, we cannot rule out the possibility of residual impairments in those with a remote history of pain. Future studies should consider incorporating objective assessments, such as imaging or functional tests, to confirm the absence of such impairments. Moreover, the generalizability of the study's results may be limited to populations with RCRSP and may not necessarily extend to other shoulder conditions, thus affecting external validity. Finally, the cross-sectional design precludes establishing causality, leaving open the question of whether our observations in the cervical spine represent a predisposing factor to shoulder pain, a consequence of the painful shoulder process, or an underlying cervical cause of shoulder pain.

Conclusion

This study investigated cervical variables in individuals with RCRSP, revealing differences in cervical AROM and PPTs compared to asymptomatic controls. The observed associations between cervical variables and shoulder pain and disability further support the concept of regional interdependence in people with shoulder pain. Our findings emphasize the need for a comprehensive approach to RCRSP assessment and management including potentially contributing factors from the cervical spine.

Acknowledgments

We gratefully acknowledge all participating patients involved in the study.

Disclosures

Conflict of interest: The authors declare no conflict of interest.

Financial support: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contributions: Conceptualization (NRS, RLT, RFM, ELG), data curation(RFM), formal analysis (RFM, NRS), investigation (NRS, JGH, PCL), methodology (NRS, RLT, RFM, JGH, PCL, ELG), project administration (ELG, NRS), resources (NRS, JGH, PCL), supervision (ELG, NRS), writing – original draft (NRS, JGH, PCL), writing – review & editing(NRS, RLT, RFM, JGH, PCL, ELG).

Data availability statement: All data generated or analyzed during this study are included in this published article and its supplementary information section.

References

1. Luime JJ, Koes BW, Hendriksen IJM, et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. Scand J Rheumatol. 2004;33(2):73-81. CrossRef PubMed
2. Lucas J, van Doorn P, Hegedus E, et al. A systematic review of the global prevalence and incidence of shoulder pain. BMC Musculoskelet Disord. 2022;23(1):1073. CrossRef PubMed
3. Greenberg DL. Evaluation and treatment of shoulder pain. Med Clin North Am. 2014;98(3):487-504. CrossRef PubMed
4. Linaker CH, Walker-Bone K. Shoulder disorders and occupation. Best Pract Res Clin Rheumatol. 2015;29(3):405-423. CrossRef PubMed
5. Kuhn JE. Prevalence, natural history, and nonoperative treatment of rotator cuff disease. Oper Tech Sports Med. 2023;31(1): 150978. CrossRef
6. Littlewood C, May S, Walters S. Epidemiology of rotator cuff tendinopathy: a systematic review. Shoulder Elbow. 2013;5(4): 256-265. CrossRef
7. Yamamoto A, Takagishi K, Osawa T, et al. Prevalence and risk factors of a rotator cuff tear in the general population. J Shoulder Elbow Surg. 2010;19(1):116-120. CrossRef PubMed
8. Diercks R, Bron C, Dorrestijn O, et al. Dutch Orthopaedic Association. Guideline for diagnosis and treatment of subacromial pain syndrome: a multidisciplinary review by the Dutch Orthopaedic Association. Acta Orthop. 2014;85(3):314-322. CrossRef PubMed
9. Zadro JR, O'Keeffe M, Ferreira GE, et al. Diagnostic labels for rotator cuff disease can increase people's perceived need for shoulder surgery: an online randomized controlled trial. J Orthop Sports Phys Ther. 2021;51(8):401-411. CrossRef PubMed
10. Lewis J, McCreesh K, Roy JS, et al. Rotator cuff tendinopathy: navigating the diagnosis-management conundrum. J Orthop Sports Phys Ther. 2015;45(11):923-937. CrossRef PubMed
11. Katsuura Y, Bruce J, Taylor S, et al. Overlapping, masquerading, and causative cervical spine and shoulder pathology: a systematic review. Global Spine J. 2020;10(2):195-208. CrossRef PubMed
12. Crosbie J, Kilbreath SL, Hollmann L, et al. Scapulohumeral rhythm and associated spinal motion. Clin Biomech (Bristol, Avon). 2008;23(2):184-192. CrossRef PubMed
13. Roren A, Nguyen C, Zauderer J, et al. Arm elevation involves cervical spine 3-D rotations. Ann Phys Rehabil Med. 2020;63(4):372-375. CrossRef PubMed
14. Norlander S, Aste-Norlander U, Nordgren B, et al. Mobility in the cervico-thoracic motion segment: an indicative factor of musculo-skeletal neck-shoulder pain. Scand J Rehabil Med. 1996;28(4):183-192. Accessed April 7, 2016. CrossRef PubMed
15. Norlander S, Gustavsson BA, Lindell J, et al. Reduced mobility in the cervico-thoracic motion segment--a risk factor for musculoskeletal neck-shoulder pain: a two-year prospective follow-up study. Scand J Rehabil Med. 1997;29(3):167-174. PubMed
16. Devaney LL, Denegar CR, Thigpen CA, et al. Preseason neck mobility is associated with throwing-related shoulder and elbow injuries, pain, and disability in college baseball pitchers. Orthop J Sports Med. 2020;8(5):2325967120920556. CrossRef PubMed
17. Bullock GS, Thigpen CA, Zhao H, et al. Neck range of motion prognostic factors in association with shoulder and elbow injuries in professional baseball pitchers. J Shoulder Elbow Surg. 2025;34(2):421-429. CrossRef PubMed
18. Wainner RS, Whitman JM, Cleland JA, et al. Regional interdependence: a musculoskeletal examination model whose time has come. J Orthop Sports Phys Ther. 2007;37(11):658-660. CrossRef PubMed
19. Sueki DG, Cleland JA, Wainner RS. A regional interdependence model of musculoskeletal dysfunction: research, mechanisms, and clinical implications. J Man Manip Ther. 2013;21(2): 90-102. CrossRef PubMed
20. Medeni OC, Camci E, Duzgun I. The effect of active cervical lateral flexion range on shoulder joint range of motion in adhesive capsulitis patients: a pilot study. Ann Rheum Dis. 2013;72(suppl 3):A1097-A1097. CrossRef
21. Weh L, Ehlers K. Rotator syndrome of the shoulder and mobility of the cervical spine. Z Rheumatol. 1989;48(5):223-228. PubMed
22. Rebelatto MN, Alburquerque-Sendín F, Gava V, et al. Cervical sensitivity, range of motion and strength in individuals with shoulder pain: a cross-sectional case control study. Musculoskelet Sci Pract. 2023;67:102834.CrossRef PubMed
23. Gurfinkel VS, Lipshits MI, Lestienne FG. Anticipatory neck muscle activity associated with rapid arm movements. Neurosci Lett. 1988;94(1-2):104-108. CrossRef PubMed
24. Asker M, Ravnanger J, Bjørnstad T, Skillgate E. Correlation between neck motor control impairment and shoulder pain in elite male handball players. J Sci Med Sport. 2014;18:e76. CrossRef
25. Hidalgo-Lozano A, Calderón-Soto C, Domingo-Camara A, et al. Elite swimmers with unilateral shoulder pain demonstrate altered pattern of cervical muscle activation during a functional upper-limb task. J Orthop Sports Phys Ther. 2012;42(6):552-558. CrossRef PubMed
26. Hidalgo-Lozano A, Fernández-de-las-Peñas C, Alonso-Blanco C, et al. Muscle trigger points and pressure pain hyperalgesia in the shoulder muscles in patients with unilateral shoulder impingement: a blinded, controlled study. Exp Brain Res. 2010;202(4):915-925. CrossRef PubMed
27. Ortega-Santiago R, González-Aguado ÁJ, Fernández-de-Las-Peñas C, et al. Pressure pain hypersensitivity and referred pain from muscle trigger points in elite male wheelchair basketball players. Braz J Phys Ther. 2020;24(4):333-341. CrossRef PubMed
28. Alburquerque-Sendín F, Camargo PR, Vieira A, et al. Bilateral myofascial trigger points and pressure pain thresholds in the shoulder muscles in patients with unilateral shoulder impingement syndrome: a blinded, controlled study. Clin J Pain. 2013;29(6):478-486. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

29. Coronado RA, Kindler LL, Valencia C, et al. Thermal and pressure pain sensitivity in patients with unilateral shoulder pain: comparison of involved and uninvolved sides. J Orthop Sports Phys Ther. 2011;41(3):165-173.CrossRef PubMed
30. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol. 2008;61(4):344-349.CrossRef PubMed
31. General Assembly of the World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. J Am Coll Dent. 2014;81(3):14-18. PubMed
32. Requejo-Salinas N, Lewis J, Michener LA, et al. International physical therapists consensus on clinical descriptors for diagnosing rotator cuff related shoulder pain: A Delphi study. Braz J Phys Ther. 2022;26(2):100395. CrossRef PubMed
33. Michener LA, Walsworth MK, Doukas WC, et al. Reliability and diagnostic accuracy of 5 physical examination tests and combination of tests for subacromial impingement. Arch Phys Med Rehabil. 2009;90(11):1898-1903. CrossRef PubMed
34. Thoomes EJ, van Geest S, van der Windt DA, et al. Value of physical tests in diagnosing cervical radiculopathy: a systematic review. Spine J. 2018;18(1):179-189. CrossRef PubMed
35. Hanchard NCA, Goodchild L, Thompson J, et al. Evidence-based clinical guidelines for the diagnosis, assessment and physiotherapy management of contracted (frozen) shoulder: quick reference summary. Physiotherapy. 2012;98(2):117-120. CrossRef PubMed
36. Warby SA, Ford JJ, Hahne AJ, et al. Effect of exercise-based management on multidirectional instability of the glenohumeral joint: a pilot randomized controlled trial protocol. BMJ Open. 2016;6(9):e013083. CrossRef PubMed
37. Gómez-Pérez L, López-Martínez AE, Ruiz-Párraga GT. Psychometric properties of the Spanish version of the Tampa Scale for Kinesiophobia (TSK). J Pain. 2011;12(4):425-435. CrossRef PubMed
38. García Campayo J, Rodero B, Alda M, et al. Validation of the Spanish version of the Pain Catastrophizing Scale in fibromyalgia. Med Clin (Barc). 2008;131(13):487-492. CrossRef PubMed
39. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. J Clin Nurs. 2005;14(7):798-804. CrossRef PubMed
40. 40. Tashjian RZ, Deloach J, Porucznik CA, et al. Minimal clinically important differences (MCID) and patient acceptable symptomatic state (PASS) for visual analog scales (VAS) measuring pain in patients treated for rotator cuff disease. J Shoulder Elbow Surg. 2009;18(6):927-932. CrossRef PubMed
41. Membrilla-Mesa MD, Cuesta-Vargas AI, Pozuelo-Calvo R, et al. Shoulder pain and disability index: cross cultural validation and evaluation of psychometric properties of the Spanish version. Health Qual Life Outcomes. 2015;13(1):200. CrossRef PubMed
42. Roach KE, Budiman-Mak E, Songsiridej N, et al. Development of a shoulder pain and disability index. Arthritis Care Res. 1991;4(4):143-149.CrossRef PubMed
43. Membrilla-Mesa MD, Cuesta-Vargas AI, Pozuelo-Calvo R, et al. Shoulder pain and disability index: cross cultural validation and evaluation of psychometric properties of the Spanish version. Health Qual Life Outcomes. 2015;13:200. CrossRef
44. Audette I, Dumas JP, Côté JN, et al. Validity and between-day reliability of the cervical range of motion (CROM) device. J Orthop Sports Phys Ther. 2010;40(5):318-323. CrossRef PubMed
45. Elizagaray-García I, Obispo-Villamayor Á, Prats-Martínez C, et al. Measurement proprieties of the CROM instrument for assessing head posture, neck retraction and protraction. Musculoskelet Sci Pract. 2024;71:102950. CrossRef PubMed
46. Carnevalli APO, Bevilaqua-Grossi D, Oliveira AIS, et al. Intrarater and inter-rater reliability of maximal voluntary neck muscle strength assessment using a handheld dynamometer in women with headache and healthy women. J Manipulative Physiol Ther. 2018;41(7):621-627. CrossRef PubMed
47. Krause DA, Hansen KA, Hastreiter MJ, et al. A comparison of various cervical muscle strength testing methods using a handheld dynamometer. Sports Health. 2019;11(1):59-63. CrossRef PubMed
48. Vannebo KT, Iversen VM, Fimland MS, et al. Test-retest reliability of a handheld dynamometer for measurement of isometric cervical muscle strength. J Back Musculoskelet Rehabil. 2018;31(3):557-565. CrossRef PubMed
49. Chesterton LS, Sim J, Wright CC, et al. Interrater reliability of algometry in measuring pressure pain thresholds in healthy humans, using multiple raters. Clin J Pain. 2007;23(9):760-766. CrossRef PubMed
50. Walton DM, Macdermid JC, Nielson W, et al. Reliability, standard error, and minimum detectable change of clinical pressure pain threshold testing in people with and without acute neck pain. J Orthop Sports Phys Ther. 2011;41(9):644-650. CrossRef PubMed
51. Kelley K, Maxwell SE. Sample size for multiple regression: obtaining regression coefficients that are accurate, not simply significant. Psychol Methods. 2003;8(3):305-321. CrossRef PubMed
52. Ferrario VF, Sforza C, Serrao G, et al. Active range of motion of the head and cervical spine: a three-dimensional investigation in healthy young adults. J Orthop Res. 2002;20(1):122-129. CrossRef PubMed
53. Cribari-Neto F, Zeileis A. Beta regression in R. J Stat Softw. 2010;34(2):1-24. CrossRef
54. Kocur P, Tomczak M, Wiernicka M, et al. Relationship between age, BMI, head posture and superficial neck muscle stiffness and elasticity in adult women. Sci Rep. 2019;9(1):8515. CrossRef PubMed
55. Allen WE, Lin JJ, Barfield WB, et al. Shoulder motion decreases as body mass increases in patients with asymptomatic shoulders. JSES Int. 2020;4(3):438-442. CrossRef PubMed
56. Pan F, Arshad R, Zander T, et al. The effect of age and sex on the cervical range of motion – a systematic review and meta-analysis. J Biomech. 2018;75:13-27. CrossRef PubMed
57. Mimori K, Muneta T, Komori H, et al. Relation between the painful shoulder and the cervical spine with narrow canal in patients without obvious radiculopathy. J Shoulder Elbow Surg. 1999;8(4):303-306. CrossRef PubMed
58. Kitagawa T, Fujiwara A, Kobayashi N, et al. Morphologic changes in the cervical neural foramen due to flexion and extension: in vivo imaging study. Spine. 2004;29(24):2821-2825. CrossRef PubMed
59. Chang V, Basheer A, Baumer T, et al. Dynamic measurements of cervical neural foramina during neck movements in asymptomatic young volunteers. Surg Radiol Anat. 2017;39(10):1069-1078. CrossRef PubMed
60. Wainner RS, Fritz JM, Irrgang JJ, et al. Reliability and diagnostic accuracy of the clinical examination and patient self-report measures for cervical radiculopathy. Spine. 2003;28(1):52-62. CrossRef PubMed
61. Merkle SL, Sluka KA, Frey-Law LA. The interaction between pain and movement. J Hand Ther. 2020;33(1):60-66. CrossRef PubMed

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

62. Vlaeyen JWS, Kole-Snijders AMJ, Boeren RGB, et al. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. Pain. 1995;62(3):363-372. CrossRef PubMed
63. Özden F, Tuğay N, Karaman ÖN, et al. The relationship of fear of movement with pain, range of motion and function in patients with shoulder pathologies. Bull Fac Phys Ther 2021 261. 2021;26(1):1-5. CrossRef
64. Martins F, Bento A, Silva AG. Within-session and between-session reliability, construct validity, and comparison between individuals with and without neck pain of four neck muscle tests. PM R. 2018;10(2):183-193. CrossRef PubMed
65. den Boer C, Terluin B, van der Wouden JC, et al. Tests for central sensitization in general practice: a Delphi study. BMC Fam Pract. 2021;22(1):206. CrossRef PubMed
66. Noten S, Struyf F, Lluch E, et al. Central pain processing in patients with shoulder pain: a review of the literature. Pain Pract. 2017;17(2):267-280. CrossRef PubMed
67. Houry M, Bonnard M, Tourny C, et al. Kinematic, electromyographic and isokinetic measurements for the management of shoulder subacromial pain syndrome: A systematic review. Clin Biomech. 2023;107:106029. CrossRef
68. Wytrążek M, Huber J, Lipiec J, et al. Evaluation of palpation, pressure algometry, and electromyography for monitoring trigger points in young participants. J Manipulative Physiol Ther. 2015;38(3):232-243. CrossRef PubMed
69. Valenza-Peña G, Martín-Núñez J, Heredia-Ciuró A, et al. Relationship between pain intensity, physical factors, pronociceptive pain modulation profile and psychological vulnerability on upper limb disability in older patients with chronic shoulder pain. Int J Environ Res Public Health. 2022;19(22):15006. CrossRef PubMed
70. Libardoni TC, Armijo-Olivo S, Bevilaqua-Grossi D, et al. Relationship between intensity of neck pain and disability and shoulder pain and disability in individuals with subacromial impingement symptoms: a cross-sectional study. J Manipulative Physiol Ther. 2020;43(7):691-699. CrossRef PubMed
71. Walker T, Salt E, Lynch G, et al. Screening of the cervical spine in subacromial shoulder pain: a systematic review. Shoulder Elbow. 2019;11(4):305-315. CrossRef PubMed
72. Tesarz J, Schuster AK, Hartmann M, et al. Pain perception in athletes compared to normally active controls: a systematic review with meta-analysis. Pain. 2012;153(6):1253-1262. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 110-117
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3457
ORIGINAL RESEARCH ARTICLE

Association of physiotherapy dose with motor recovery in early subacute phase after stroke: results from a multicenter Italian study

Silvia Salvalaggio[1*], Nicola Valè[2,3*], Thomas Bowman[4], Davide Cattaneo[4,5], Cristina Allera Longo[6], Serena Bocini[7], Viviana Bonci[8], Michele Gennuso[8], Francesca Marazzini[9], Francesco Giuseppe Materazzi[10,11], Elisa Pelosin[12,13], Martina Putzolu[14], Andrea Turolla[11,15], Susanna Mezzarobba[12,13,16], Fabiola Giovanna Mestanza Mattos[4], Simone Gambazza[17**]**

[1]Laboratory of Computational Neuroimaging, IRCCS San Camillo Hospital, Venice - Italy; [2]Department of Neuroscience, Biomedicine and Movement Sciences, University of Verona, Italy; [3]Department of Engineering for Innovation Medicine, University of Verona, Italy; [4]IRCCS Fondazione Don Carlo Gnocchi, Milan - Italy; [5]Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milan - Italy; [6]Department of Rehabilitation, San Carlo Borromeo Hospital, Milano - Italy; [7]Division of Physical and Rehabilitation Medicine, Fondazione Opera San Camillo, Presidio di Torino, Torino - Italy; [8]Department of Neurological Sciences, Neurorehabilitation clinic, AOU delle Marche, Ancona - Italy; [9]AIAS di Milano Onlus, Milan - Italy; [10]Montecatone Rehabilitation Institute, Imola (BO), Italy; [11]Department of Biomedical and Neuromotor Sciences (DIBINEM), Alma Mater Studiorum - Università di Bologna, Bologna - Italy; [12]Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genoa, Genoa - Italy; [13]IRCCS Ospedale Policlinico San Martino, Genoa - Italy; [14] Department of Experimental Medicine (DIMES), Section of Human Physiology, University of Genoa, Genoa - Italy; [15]IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna - Italy; [16]Azienda Sanitaria universitaria Giuliano Isontina (ASUGI), Trieste - Italy; [17]Department of Clinical Sciences and Community Health, Laboratory of Medical Statistics, Biometry and Epidemiology "G.A. Maccacaro," Dipartimento di Eccellenza 2023-2027, University of Milan, Milan - Italy

*These Authors equally contributed to the work. **These Authors equally contributed to the work.

ABSTRACT

Introduction: The literature investigating the factors associated with functional recovery after stroke suggests that, in the early subacute phase, time is the factor that most significantly drives the recovery. However, it is unclear whether the dose of physiotherapy (PT) delivered is equally associated with recovery of motor function and independence.
Methods: A multivariable modeling of data from a multicenter longitudinal prospective cohort study investigating the contents of neurological PT interventions in Italy was developed, with the aim to estimate the association between the dose of PT received in the early subacute phase after stroke and recovery of walking, motor function, and independence.
Results: A total of 96 patients were included in the analyses. PT dose seems associated with recovery of manual dexterity, measured by the Box and Block test (BBT), but not with the level of independence nor walking function. The probability of achieving a score higher than 31 at the BBT is 46% (95%CI: 18-76) after 10 hours and 76% (95%CI: 42-93) after 20 hours of PT, respectively.
Conclusions: In the early subacute phase after stroke, the higher the dose of PT provided, the better the probability of significant recovery of upper limb (UL) dexterity should be.

Keywords: Dose, Expected recovery, Physiotherapy, Stroke

What is already known in this topic:

- *In the first 3 months after stroke, time is the most influential factor associated with recovery of motor functions.*

What does the study add:

- *In the first 3 months after stroke, increasing the physiotherapy (PT) dose seems associated with improvement of outcomes for manual dexterity but may require more nuanced approaches or additional interventions to show measurable impacts on broader, more complex domains of recovery.*

Received: January 8, 2025
Accepted: April 6, 2025
Published online: May 13, 2025

This article includes supplementary materials

Clinical Trial Protocol number: clinical trial.gov (ID: NCT04386863)

Corresponding author:
Thomas Bowman
email: tbowman@dongnocchi.it

Introduction

Stroke is the second cause of death and one of the leading causes of disability worldwide (1). The consequent neurological impairment dramatically impacts the daily independence and quality of life of stroke survivors (2). Functional recovery is thought to be driven by a combination of spontaneous biological processes and mechanisms involving motor re-learning and the acquisition of new skills (3). While behavioral

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

restitution occurs mostly in the acute (1-7 days) and early subacute phase of recovery (7 days to 3 months), adaptation and compensation can continue to sustain improvement at the activity level in the chronic phase after stroke (4). Recently, an Italian multicenter study highlighted that in neurological physiotherapy (PT), walking recovery consistently emerges as the primary goal, followed by upper limb (UL) functional recovery (5). Moreover, the average hours spent on gait training were approximately twice that of those allocated to manipulation exercises in people with stroke (PwS). Interestingly, independence in activities of daily living (ADLs) was ranked only fifth in priority.

Walking recovery after a stroke typically occurs within the first 3-6 months, with most of the improvement occurring in the first 10 weeks, but with different rates based on stroke severity and individual factors (6). Analogously, the independence in ADLs, as measured by the Barthel Index (BI), achieves the highest recovery rate typically within the first month after discharge. This improvement continues for 6 months, with many PwS reaching a functional plateau.

A previous study found that initial BI scores, age, and cognitive status play a key role in modifying recovery trajectories. Those subjects with higher BI at discharge, especially more than 60 points, show a tendency to maintain or improve their ADLs independence over time (7). Other evidence showed that PwS experienced recovery in different domains, including upper limb (UL) function and gait, following a logarithmic-like trajectory, reaching a plateau around 10 weeks after the event (8).

Despite the crucial role of rehabilitation in fostering sensorimotor recovery after stroke, there is still inconclusive evidence on the interaction between rehabilitation and neurological mechanisms involved in behavioral restitution after stroke, especially in the early subacute phase (4). Moreover, studies investigating changes in UL quality of movement after rehabilitation suggested that rehabilitation is likely to induce adaptive and compensatory mechanisms rather than behavioral restitution (9). Possibly reflecting these considerations, in the subacute phase, time has been found to be the major driver of recovery (10), whereas, in later phases, rehabilitative treatments promote clinically significant effects. Indeed, there is strong evidence suggesting that a high dose of practice can induce clinically significant changes in the recovery of UL motor function, (11,12) lower limb strength (13), and walking activity (14). The dose of PT in neurological rehabilitation is defined by specific measures such as the intensity (defined as the amount of physical or mental work put forth by the patient during a particular movement or series of movements, exercise, or activity during a therapy session, and as the session's length in minutes), frequency (e.g., number of sessions per week) and duration (e.g., how many sessions in total or for how many weeks) (15,16). Recently, research on this topic has begun to explore the influence of the dose of rehabilitation, offering initial insights suggesting that the dose may indeed be associated with better functional recovery after strokes (17-20). However, these studies have predominantly focused on specific domains, like UL function and activity, and a clear statistical association or interaction between PT dose and motor recovery has not been demonstrated. Moreover, this approach has not yet been applied to more general domains, such as quality of life and levels of independence.

Aims

The primary aim of this study was to investigate the association between PT dose and motor recovery in PwS. The secondary aim was to identify clinical and demographic features that might be associated with rehabilitation-induced motor recovery.

In summary, our hypothesis is that the PT dose may be an additional component of motor recovery of specific UL functions during the early subacute phase after stroke, whereas it could be more challenging to discern its impact in broader domains like gait recovery and independence in ADLs.

Methods

Study design

Data for this longitudinal study are retrieved from a multicenter longitudinal prospective cohort study investigating the contents of neurological PT interventions in Italy (5). The multicenter network comprised seven facilities including UO di Riabilitazione Specialistica, Presidio ospedaliero San Carlo Borromeo, ASST Santi Paolo e Carlo, Milan (Center 1); Azienda Ospedaliera Universitaria Ospedali Riuniti di Ancona Torrette (Center 2); Dipartimento di Neuroscienze, Riabilitazione, Oftalmologia, Genetica e Scienze Materno Infantili, University of Genova (Center 3); Presidio Sanitario San Camillo, Torino (Center 4); Università degli Studi di Trieste (Center 5); IRCCS San Camillo, Venezia (Center 6); IRCCS Santa Maria Nascente, Fondazione Don Carlo Gnocchi, Milan (Center 7).

Details on the ethics committees and recruitment methods are described elsewhere (5).

Participants

For the purpose of our work, we analyzed only PwS included in the acute and early sub-acute phase (0-12 weeks) after stroke and with rehabilitation goals of improving manipulation, gait, or independence in ADL. Individuals under 18 years of age or those unable to comprehend the study protocol were excluded.

Physiotherapy interventions

PwS completed PT sessions according to the Italian National Healthcare System (NHS). Each physiotherapist in charge of the PwS was free to choose the most appropriate intervention according to the patient's need as in daily clinical practice, then completed the PT Interventions Form (i.e., a checklist gathering information on PT goals, e.g., walking or hand dexterity) and intervention details (e.g., PT hours, type of exercises performed, etc.). This allowed us to characterize the PT dose (in hours) dedicated to each specific goal and to investigate PT effects using the appropriate outcome measure. To minimize the assessment burden, physiotherapists were requested to fill out the "PT Intervention Form" only at the conclusion of the rehabilitation program.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Clinical assessment and outcome measures

All the participants were clinically assessed at two time points: baseline and after the intervention (follow-up). An experienced clinical researcher who was not involved in the PT sessions conducted all the clinical assessments. Motor function was assessed before and after treatment using the following outcome measures: Modified Barthel Index (BI) score (21), a measure of physical disability used widely to assess behavior relating to ADL with higher scores indicating a lower level of disability; Two-Minute Walk Test (2MWT) distance (22) to evaluate walking performance, that is the distance walked in two minutes along a 30 m walkway; Box and Blocks test (BBT) score (23), which measures unilateral gross manual dexterity and consists in the number of blocks carried out from a box to a close one overtaking a barrier in one-minute. For quantification of stroke severity, the National Institutes of Health Stroke Scale (NIHSS) (24) was used (0-42 points, where 0 represents intact neurological functions and 42 very severe stroke).

The dose of PT was reported both as total dose (i.e., total hours of therapy for each goal – manual dexterity, walking, and independence) and through specific metrics, including frequency (number of sessions per week), intensity (minutes for the session) and total duration (total number of sessions), for each rehabilitation center involved in the project.

Statistical analysis

Data are presented as mean and ±1 standard deviation (SD) or median and interquartile range (IQR), as appropriate. Metrics of interest are also reported as the difference between follow-up and baseline, expressed as the mean difference with 95% confidence intervals (CI) and as the count and percentage of individuals above the minimal clinically important difference (MCID). For the subacute PwS, the MCID known from the literature is 6 points for the BBT (25) and 31 meters for the 2MWT (26), while it is not known yet for the BI. The Wilcoxon signed-rank test was used to assess whether paired samples showed statistically significant differences, defined as a p-value of <0.05. We investigated the association between PT dose (hours) and the outcomes of interest (e.g., BI, 2MWT and BBT at follow-up) by fitting three semiparametric ordinal regression models: (i) the BI-model included age (years), sex (male/female), weeks post-stroke, and baseline Barthel score as adjusting covariates; (ii) the 2MWT-model included age, sex, weeks post-stroke, baseline 2MWT distance, and baseline BI score as covariates. Age and 2MWT distance were modeled using restricted cubic splines with three knots at 0.10, 0.50, and 0.90 quantiles; and (iii) the BBT model was adjusted for weeks post-stroke and baseline BBT score on top of the dose. The Huber-White method was applied to account for correlated responses across different rehabilitation centers. Models were selected based on minimizing the Akaike Information Criteria function. The choice of variables for each model was determined by the authors' expertise, which was within the limitations of the available data. In the BI model, dose refers to the total hours of PT received. In the 2MWT model, the dose specifically reflected the hours dedicated to improving gait capacity and endurance, while in the BBT model, it referred to the hours focused on enhancing hand dexterity.

Each model was validated using 200 bootstraps. Performance was reported as an optimism-corrected R^2 index and as Van Houwelingen-Le Cessie's heuristic shrinkage factor. Model estimates are accompanied by 95%CI and are expressed as predicted means and exceedance probabilities. All the analyses were performed using R Core Team 2023, version 4.4.0, with rms and Hmisc packages added (27).

Results

In total, 119 PwS were screened for inclusion and exclusion criteria, and 96 were included in the present study. Among them, 78 performed PT aiming at improving walking capacity and 31 aiming at improving manual dexterity and upper limb function. Noteworthy, some PwS underwent both UL and walking rehabilitation. In those cases, for each analysis, only the PT dose associated with the specific aim was considered. Clinical and demographic data on the included sample are reported in Table 1. A description of specific dose parameters is reported for each center involved in supplementary materials (Table S1).

TABLE 1 - Sample description at baseline

Variables at baseline	BI-model (N = 96)	2MWT-model (N = 78)	BBT-model (N = 31)
Sex (male/female)	56 (58%)/40 (42%)	44 (56 %)/34 (44%)	19 (61%)/12 (39%)
Age (years)	69.4 (12.3) 71.0 [63.5; 78.5]	70.7 (11.1) 73 [64; 79]	67.5 (12.2) 69 [61; 75]
Time from lesion (weeks)	3.4 (2.8) 2.4 [1.3; 4.4]	3.4 (2.9) 2.4 [1.3; 4.4]	3.1 (3.0) 1.9 [1.0; 3.7]
Unknown (n)	1	–	–
NIHSS score	5.4 (3.32) 5 [3; 7]	5.2 (3.27) 5 [3; 6]	5.46 (4.39) 3 [2; 9]
Unknown (n)	62	53	20
PT dose (hours)	31 (20) 25.0 [1.4, 44.0]	12.5 (9.5) 9.5 [5.6; 17.5]	9.9 (8.1) 7.5 [3.9; 14.0]
Unknown (n)	7	–	–

Values are reported as counts and percentages (%), mean(sd), and median [Q1; Q3]. NIHSS: National Institute of Health Stroke Scale; BI: Barthel Index; 2MWT: 2 minutes walking test; BBT: box & blocks test; sd: standard deviation; Q1: first quartile; Q3: third quartile; PT: physiotherapy.

Baseline and follow-up analysis for the three groups showed statistically significant improvements in the specific outcome measures (Table 2). The 59.3% (16/27) and 49.2% (37/75) of PwS showed an improvement in the BBT score and 2MWT distance equal to or above the MCID (e.g., ≥6 blocks and ≥31 m, respectively).

Results of the BI model showed that time since stroke (χ^2 = 6.6; p = 0.01) and BI at baseline (χ^2 = 41.9; p < 0.001) are the only factors associated with the BI score at follow-up (Fig. 1). With an average dose of 31 hours, the mean expected BI is 87 (95%CI: 83.5-90.4). Sex (χ^2 = 2.3; p = 0.133), dose

TABLE 2 - Performance improvement after physiotherapy

Outcome measure	Baseline	Follow-up	Mean difference [95% CI]	p-value
BI [points] (N = 96)	54.1 (23.8)	80.58 (20.1)	26.3 [21.5–31.1]	<0.001
2MWT [m] (N = 78)	39.1 (42.9)	77.4 (43.8)	38.8 [30.5–47.2]	<0.001
BBT [blocks] (N = 31)	17.4 (17.0)	27.1 (16.7)	9.7 [5.9–13.5]	<0.001

Values are reported as mean (standard deviation) for baseline and follow-up scores; BI: Barthel Index; 2MWT: 2 minutes walking test (in meters); BBT: box and blocks test; CI: confidence interval.

(χ^2 = 0.15; p = 0.702), and age (χ^2 = 1.2; p = 0.281) do not show evidence of association with BI at follow-up.

Figure 2 shows exceedance probabilities of follow-up BI scoring ≥50, ≥75, and ≥91 (e.g., moderate, mild, and minimal functional dependency, respectively) as a function of time. Probabilities of Barthel ≥91 at one week are 59.1% (95% CI: 37.5-77.7), 54.7% (95%CI: 34.1-73.9) at two weeks and 45.8% (95%CI: 25.4-67.7) at 4 weeks.

Results of 2MWT-model (Fig. 3) show that only 2MWT at baseline (χ^2 = 43.6; p < 0.001) is statistically associated with the 2MWT follow-up, with no evidence of association for age (χ^2 = 5.5; p = 0.064), time since stroke (χ^2 = 0.9; p = 0.334), dose (χ^2 = 0.4; p = 0.531), sex (χ^2 = 0.2; p = 0.638) and BI at baseline (χ^2 = 0.02; p = 0.878).

Results of the BBT model (Fig. 4) show that dose (χ^2 = 6.4; p = 0.011) and BBT at baseline (χ^2 = 36.9; p < 0.001) are statistically associated with BBT at follow-up. Ten and twenty hours of PT are associated with a final mean BBT of 30 (95% CI: 26-34) blocks and 36 (95%CI: 30-42) blocks. Weeks since stroke are not statistically associated with BBT score (χ^2 = 0.7; p = 0.399), with a small difference when starting one or three weeks after stroke, 29 (95%CI: 24-34) blocks, and 28 (95%CI: 24-32) blocks, respectively.

Exceedance probabilities of scoring above 50 and 90 percentiles of the predicted mean BBT are represented in Figure 5. Probabilities of BBT ≥31 blocks are 46% (95%CI: 18-76) and 76% (95%CI: 42-93) when doing 10 and 20 hours of PT, respectively.

Overall, R^2 overfitting corrected is 35%, 47%, and 74% in the BI, 2MWT, and BBT models, respectively. Overfitting (γ) is above 0.86 except in the 2MWT model, presenting with γ = 0.76, indicating some overfitting. Detailed coefficients of each model are presented in Supplementary Materials (Table S2).

Discussion

The main message of this study is that in the early subacute phase after stroke, PT dose supports manual dexterity recovery; timely intervention is key for regaining independence, but the factors influencing walking recovery remain unclear.

We did not find an association between the PT dose and the BI. However, we found that the probability of regaining a level of independence decreases over time following a stroke, suggesting that earlier initiation of rehabilitation is linked to better recovery outcomes, whereas rehabilitation dose seems less relevant.

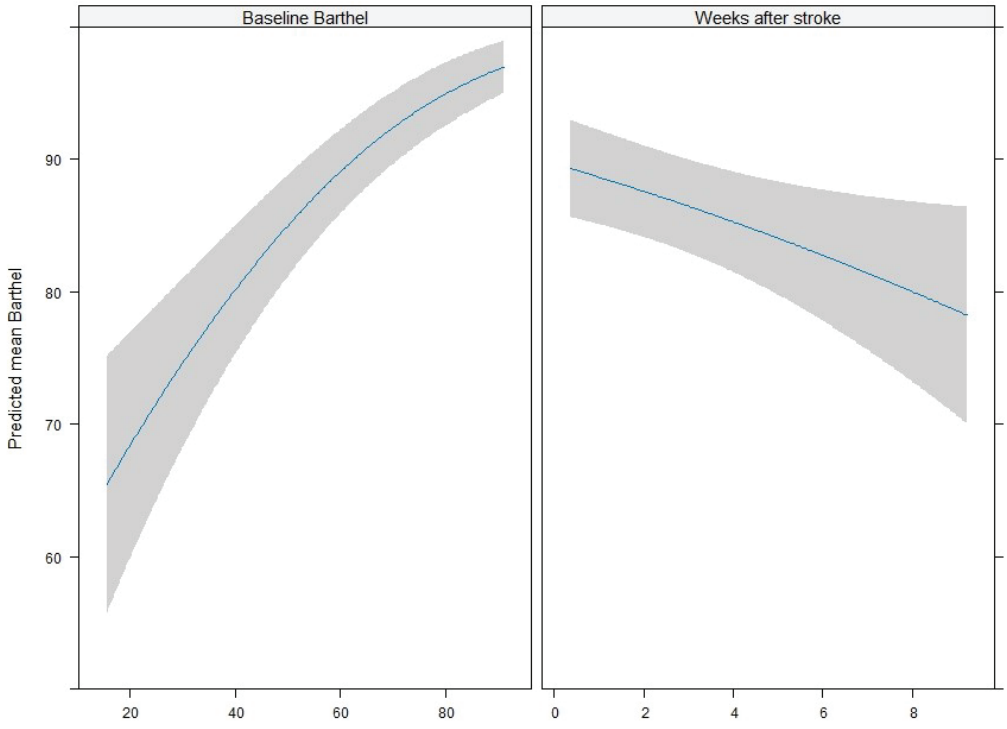


FIGURE 1 - Partial effect plot of BI-model. Partial effect plots show how baseline Barthel and weeks after stroke are associated with the predicted mean BI score. Grey bands denote 95%CI.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

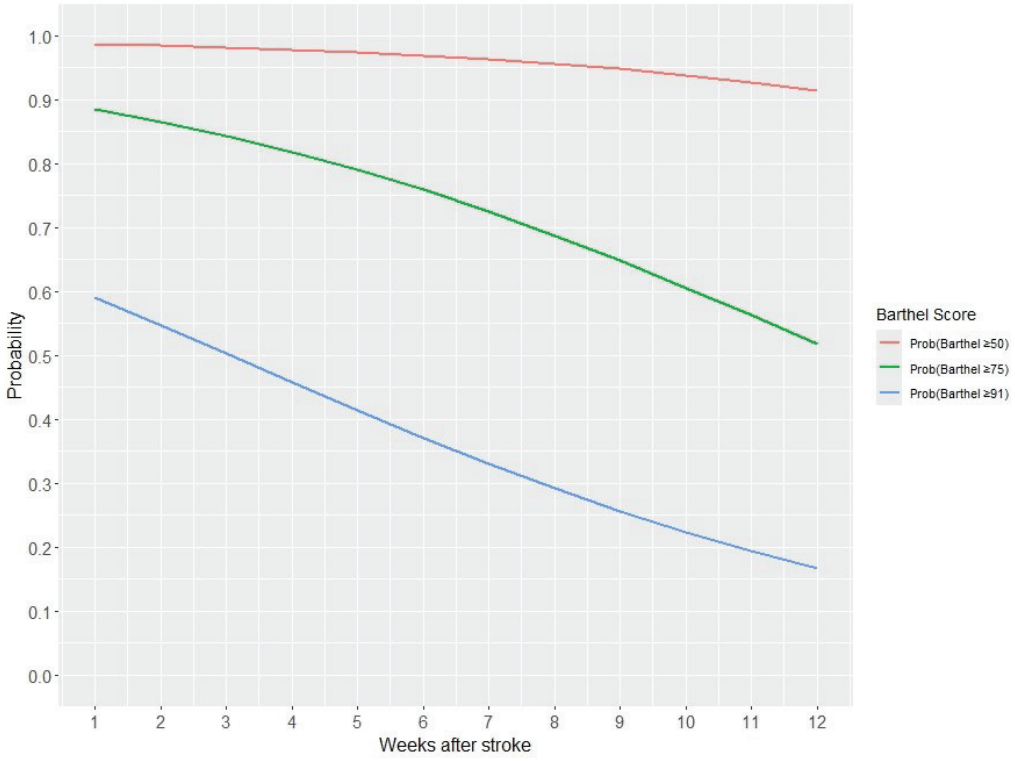


FIGURE 2 - Exceedance probabilities of Barthel Index as a function of time since stroke. Exceedance probabilities (i.e., Prob) of scoring equal or above clinically relevant BI thresholds, according to the fitted model.

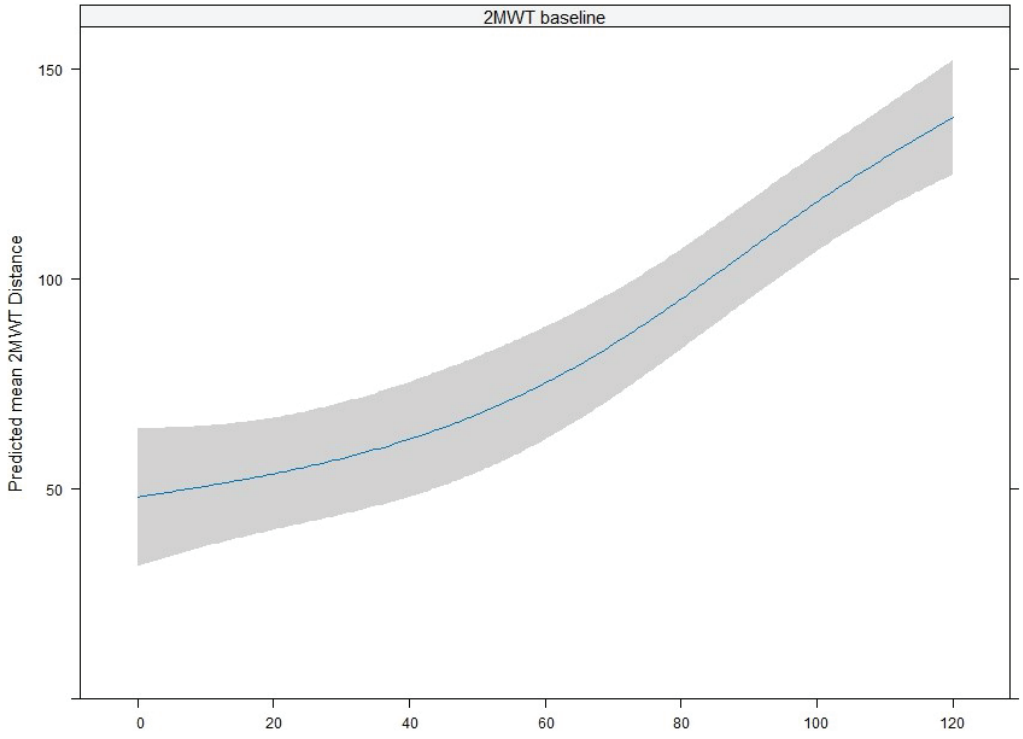


FIGURE 3 - Partial effect plot of 2MWT-model. Partial effect plot of baseline 2MWT [m] distance against 2MWT model prediction, with evidence of non-linearity (χ^2 = 6.9; p = 0.009). Grey bands denote 95%CI.

Given the results of the 2MWT model, it is worth considering that the 2MWT primarily assesses endurance rather than speed. Therefore, it is possible that the training provided was not endurance-specific, which would clarify the absence of association with PT dose. Moreover, the domain of walking is quite broad and likely influenced by factors

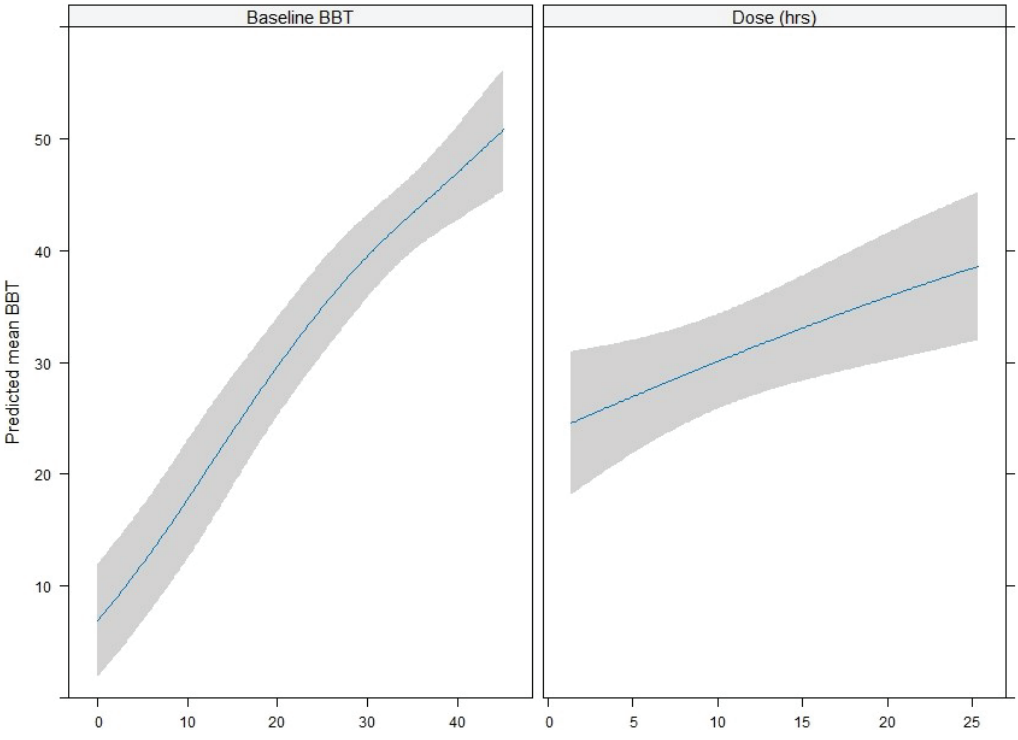


FIGURE 4 - Partial effect plot of BBT-model. Partial effect plot of mean Baseline BBT and Dose. Grey bands denote 95%CI.

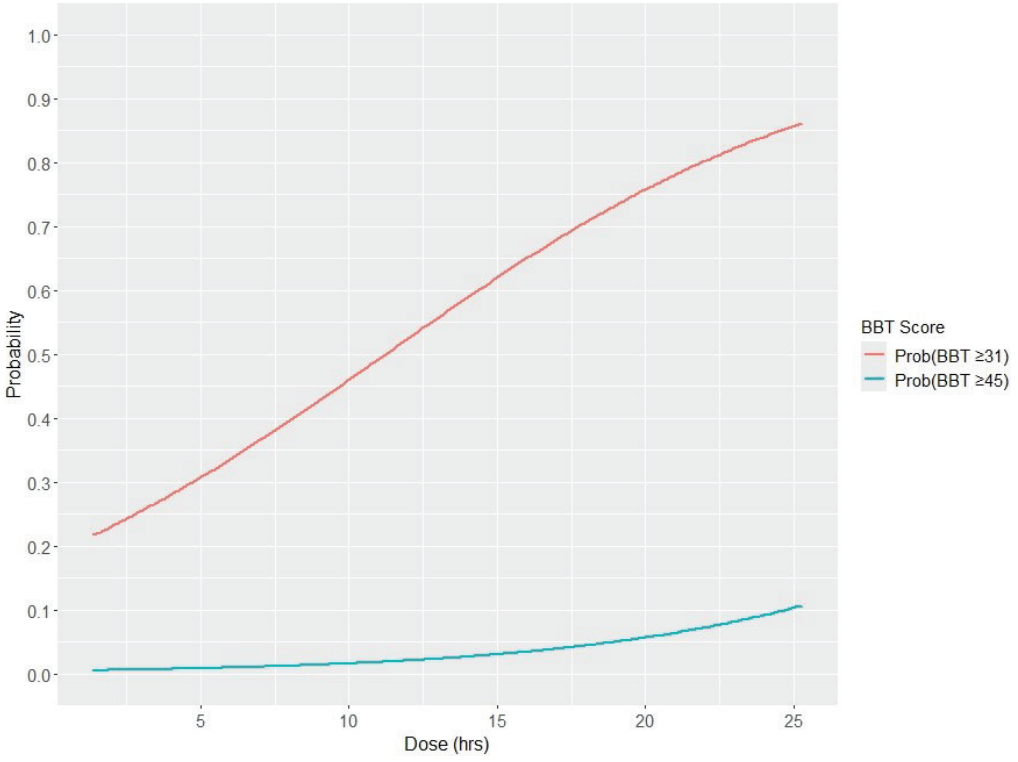


FIGURE 5 - Probabilities of BBT improvement according to increasing PT dose. Exceedance probabilities (i.e., Prob) of scoring equal or above 50 (BBT > 31) and 90 (BBT > 45) percentiles of mean BBT, according to the fitted model.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

beyond specific training interventions, as well as other elements not directly related to training, suggesting a more complex interplay of variables contributing to overall performance. Indeed, walking is a motor activity that requires both sensorimotor and cognitive functions, such as muscle strength, balance, proprioception, cardiovascular fitness, and visual abilities (28-32).

Considering the relevance of time since stroke, available literature suggested that starting PT early during the acute to subacute stroke phase with a higher dose may enhance walking capacity, but the evidence remains limited (14,33). For example, a previous study found that PwS can improve their walking recovery with higher doses (2 hours, 5 days/week) of aerobic and stepping activity within a critical time period for neurological recovery (14).

Another relevant aspect related to walking recovery is the specificity of the intervention. For instance, gait-oriented training (30 minutes, three to five times per week) showed promising results, but the factors associated with gait recovery are still unclear (32). Although the results are not statistically significant, it is interesting to note that PwS who were able to walk >60 meters in 2 minutes before treatment seem to recover more rapidly than those walking shorter distances. Clinically, this finding suggests that PwS should be encouraged to walk more to support recovery, while those unable to do so require careful monitoring and guidance toward alternative movement strategies. Training should be tailored to baseline performance, recognizing that progress may vary. In particular, PwS walking less than 60 meters in the first weeks after stroke should receive closer supervision and targeted support for alternative mobility approaches.

The PT dose was found to be associated with the BBT score in the early subacute phase after stroke in PwS undergoing PT.

For expected recovery of manual dexterity, its probability of recovery increases with increasing PT dose. The relationship between manual dexterity and dose suggests that part of the observed recovery may not solely be attributed to spontaneous recovery but rather to the effects of rehabilitation. This contrasts with results found for walking function and level of independence, where such association was not established. Moreover, for manual dexterity, the timing of initiation of rehabilitation treatment, whether starting a week earlier or later, has little effect on the overall outcome. Our results suggest that maintaining consistent and targeted work overtime should be encouraged throughout the subacute phase of recovery after the stroke. Worth mentioning is the fact that our findings fit into a broader scientific context where it is still unclear what works and why with PwS. We may not be able to conclude that rehabilitation dose does not impact some recovery domains. Perhaps simply not enough was provided. Additionally, it remains uncertain whether the interventions used were based on principles of motor control, motor learning, or neuroplasticity, as recently suggested in the field (34,35). Our results reflect the current reality of rehabilitation in Italian settings but should serve as a warning to adopt a broader perspective, both on recovery mechanisms and on the decision-making processes behind the interventions proposed.

Limitations of our study should be mentioned. First, there could be a recruitment bias since participants were selected according to the availability of centers. Furthermore, we examined the PT dose in terms of the total amount of hours, not frequency, number of repetitions, or specific time on task, which would be more accurate parameters. Finally, since the recruited PwS were following the rehabilitation program provided by the Italian NHS, it is possible that, in addition to PT, they were also undergoing concurrent therapies for which it was not possible to track the dosage (e.g., occupational therapy). This is due to a standardized data collection method across all centers that prioritized the acquisition of simple but consistent data over complex information storage, which would have led to a high risk of data loss. Moreover, the dose used in our models refers specifically to the goal-specific dose, not the total dose or the dose provided by individual centers. The latter was included only descriptively, as making inferences about the effects of different doses would have required a different study design with a priori hypotheses (e.g., randomized controlled trial). Instead, this work reflects conventional practices within the Italian NHS.

Conclusion

In the early subacute phase after stroke, PT dose seems associated with recovery of manual dexterity, not with level of independence or walking function recovery. Our findings suggest the idea that in the early subacute phase of recovery of the UL dexterity, *the more rehabilitation, the better the outcome*. Future research on this topic should use a prospective study design to validate our findings, including treatment frequency and intensity as crucial components of the overall treatment dose.

Acknowledgments

The authors acknowledge all the physiotherapists for their contribution to data collection and patients' treatment at the NHS centers involved in the project.

Disclosures

Conflict of interest: The authors declare they have no conflict of interest.

Financial support: The study was partly funded by the Italian Ministry of Health – Ricerca Corrente 2024 and partly supported by the Italian Physiotherapy Association (AIFI).

Data availability statement: The data presented in this study are available on request from the corresponding author.

Authors contribution: Conceptualization: T.B. and D.C.; methodology: T.B., F.G.M.M., F.M., A.T., S.M. and D.C.; data collection: T.B., F.G.M.M., S.S., F.M., C.A., S.B., V.B., M.G., F.G.M., E.P., M.P. and A.T.; data analysis: N.V., S.G., S.S.; writing – original draft preparation: S.S., N.V., S.G.; writing – review and editing: S.S., F.M., C.A., S.B., V.B., M.G., F.G.M., E.P., M.P., A.T.; supervision: S.M., F.G.M.M. and D.C.; project administration: T.B, F.G.M.M. and D.C.

References

1. Feigin VL, Brainin M, Norrving B, et al. World Stroke Organization (WSO): Global Stroke Fact Sheet 2022. Int J Stroke. 2022;17(1):18-29. CrossRef PubMed
2. Hayward KS, Kramer SF, Thijs V, et al. A systematic review protocol of timing, efficacy and cost effectiveness of upper limb

therapy for motor recovery post-stroke. Syst Rev 2019; 8: 187. CrossRef
3. Levin MF, Kleim JA, Wolf SL. What do motor "recovery" and "compensation" mean in patients following stroke? Neurorehabil Neural Repair. 2009;23(4):313-319. CrossRef PubMed
4. Kwakkel G, Stinear C, Essers B, et al. Motor rehabilitation after stroke: European Stroke Organisation (ESO) consensus-based definition and guiding framework. Eur Stroke J. 2023;8(4): 880-894. CrossRef PubMed
5. Bowman T, Mestanza Mattos FG, Salvalaggio S, et al. Classification and quantification of physical therapy interventions across multiple neurological disorders: an Italian multicenter network. J Clin Med 2023; 12(20):6483 CrossRef
6. Kennedy C, Bernhardt J, Churilov L, et al. Factors associated with time to independent walking recovery post-stroke. J Neurol Neurosurg Psychiatry. 2021;92(7):702-708. CrossRef PubMed
7. Li QX, Zhao XJ, Wang Y, et al. Value of the Barthel scale in prognostic prediction for patients with cerebral infarction. BMC Cardiovasc Disord 2020; 20: 14. CrossRef
8. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. Lancet. 2011;377(9778):1693-1702. CrossRef PubMed
9. Bernhardt J, Hayward KS, Kwakkel G, et al. Agreed definitions and a shared vision for new standards in stroke recovery research: the Stroke Recovery and Rehabilitation Roundtable Taskforce. Neurorehabil Neural Repair. 2017;31(9):793-799. CrossRef PubMed
10. Kwakkel G, Kollen B, Twisk J. Impact of time on improvement of outcome after stroke. Stroke. 2006;37(9):2348-2353. CrossRef PubMed
11. Ward NS, Brander F, Kelly K. Intensive upper limb neurorehabilitation in chronic stroke: outcomes from the Queen Square programme. J Neurol Neurosurg Psychiatry. 2019;90(5): 498-506. CrossRef PubMed
12. Daly JJ, McCabe JP, Holcomb J, et al. Long-dose intensive therapy is necessary for strong, clinically significant, upper limb functional gains and retained gains in severe/moderate chronic stroke. Neurorehabil Neural Repair. 2019;33(7): 523-537. CrossRef PubMed
13. Veerbeek JM, van Wegen E, van Peppen R, et al. What is the evidence for physical therapy poststroke? A systematic review and meta-analysis. PLoS One 2014; 9: e87987. CrossRef
14. Klassen TD, Dukelow SP, Bayley MT, et al. Higher Doses Improve Walking Recovery During Stroke Inpatient Rehabilitation. Stroke. 2020;51(9):2639-2648. CrossRef PubMed
15. Goikoetxea-Sotelo G, van Hedel HJA. Defining, quantifying, and reporting intensity, dose, and dosage of neurorehabilitative interventions focusing on motor outcomes. Front Rehabil Sci 2023; 4: 1139251. CrossRef
16. Conn VS, Chan KC. How Much, how often, how long? Addressing dosage in intervention studies. West J Nurs Res. 2016;38(1): 3-4. CrossRef PubMed
17. Salvalaggio S, Turolla A, Ando M, et al. Prediction of rehabilitation induced motor recovery after stroke using a multidimensional and multi-modal approach. Front Aging Neurosci 2023; 15: 1205063. CrossRef
18. Salvalaggio S, Cacciante L, Maistrello L, et al. Clinical predictors for upper limb recovery after stroke rehabilitation: retrospective cohort study. Healthcare (Basel) 2023; 11 CrossRef
19. Salvalaggio S, Gambazza S, Ando' M, et al. Modeling upper limb rehabilitation-induced recovery after stroke: the role of attention as a clinical confounder. Phys Ther. 2025;105(2):pzae148 CrossRef PubMed
20. Salvalaggio S, Gianola S, Andò M, et al. Predictive factors and dose-response effect of rehabilitation for upper limb induced recovery after stroke: systematic review with proportional meta-analyses. Physiotherapy. 2024;125:101417. CrossRef PubMed
21. Kamdi MKA, Shafei MN, Musa KI, et al. Comparison of the Modified Barthel Index (MBI) score trends among workers with stroke receiving robotic and conventional rehabilitation therapy. Cureus 2023; 15: e34207. CrossRef
22. Bohannon RW, Wang YC, Gershon RC. Two-minute walk test performance by adults 18 to 85 years: normative values, reliability, and responsiveness. Arch Phys Med Rehabil. 2015;96(3):472-477. CrossRef PubMed
23. Bertoni R, Cattaneo D, Grosso C, et al. Distribution and relation of two arm function tests, Box and Blocks test and Nine Hole Peg test, across disease severity levels and types of multiple sclerosis. Mult Scler Relat Disord 2022; 59: 103683. CrossRef
24. Goldstein LB, Bertels C, Davis JN. Interrater reliability of the NIH stroke scale. Arch Neurol. 1989;46(6):660-662. CrossRef PubMed
25. Mathiowetz V, Volland G, Kashman N, Weber K. Adult norms for the Box and Block Test of manual dexterity. Am J Occup Ther. 1985;39(6):386-391. CrossRef PubMed
26. Bowman T, Mestanza Mattos FG, Allera Longo C, et al. The minimally clinically important difference in the 2-minute walk test for people in the subacute phase after a stroke. Topics in Stroke Rehabilitation:2024, 1-9. CrossRef
27. R Foundation for Statistical Computing. R: A Language and Environment for Statistical Computing, 2019. Online (Accessed January 2025)
28. Fujita T, Sato A, Ohashi Y, et al. Amount of balance necessary for the independence of transfer and stair-climbing in stroke inpatients. Disabil Rehabil. 2018;40(10):1142-1145. CrossRef PubMed
29. Blennerhassett JM, Levy CE, Mackintosh A, et al. One-quarter of people leave inpatient stroke rehabilitation with physical capacity for community ambulation. J Stroke Cerebrovasc Dis. 2018;27(12):3404-3410. CrossRef PubMed
30. Rudomin G, Keren O, Katz Leurer M. Development of a prediction model for ascent and descent staircase independence during the sub-acute rehabilitation phase in individuals post-stroke. NeuroRehabilitation. 2021;48(4):523-532. CrossRef PubMed
31. Tiedemann AC, Sherrington C, Lord SR. Physical and psychological factors associated with stair negotiation performance in older people. J Gerontol A Biol Sci Med Sci. 2007;62(11): 1259-1265. CrossRef PubMed
32. An M, Shaughnessy M. The effects of exercise-based rehabilitation on balance and gait for stroke patients: a systematic review. J Neurosci Nurs. 2011;43(6):298-307. CrossRef PubMed
33. Wu WX, Zhou CY, Wang ZW, et al. Effect of early and intensive rehabilitation after ischemic stroke on functional recovery of the lower limbs: a pilot, randomized trial. J Stroke Cerebrovasc Dis 2020; 29: 104649. CrossRef
34. Branscheidt M, Hadjiosif AM, Anaya MA, et al. Reinforcement learning is impaired in the sub-acute post-stroke period. Neurorehabil Neural Repair. 2025 CrossRef PubMed
35. Hadjiosif AM, Krakauer JW, Haith AM. Did we get sensorimotor adaptation wrong? Implicit adaptation as direct policy updating rather than forward-model-based learning. J Neurosci. 2021;41(12):2747-2761. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 118-130
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3209
REVIEW

Clinical relevance of the effects of robotic rehabilitation for upper limb recovery after stroke in randomized studies: a systematic review with meta-analysis

Sofia Verola[1], Alessandro Ugolini[2], Leonardo Pellicciari[3], Mauro Di Bari[4,5], Matteo Paci[1,4]

[1]Department of Allied Health Professions, Azienda USL Toscana Centro, Florence - Italy
[2]Independent Researcher, Empoli (FI) - Italy
[3]IRCCS Istituto delle Scienze Neurologiche di Bologna, Bologna - Italy
[4]Department of Experimental and Clinical Medicine, University of Florence, Florence - Italy
[5]Department of Medicine and Geriatrics, Unit of Geriatrics, Azienda Ospedaliero-Universitaria Careggi, Florence - Italy

ABSTRACT

Introduction: Recent randomized clinical trials (RCTs) provide evidence on the effects of robot-assisted training (RAT) for upper limb impairments in stroke subjects; however, evidence on the clinical relevance of these differences is lacking. This study aimed to perform a systematic review with meta-analyses of RCTs on clinical relevance, expressed as minimal clinically important difference (MCID), of RAT to improve independence in activities of daily living, arm function, and impairments in patients with stroke.

Methods: Four databases were searched. RCTs investigating RAT aimed at recovering motor and functional skills of the upper limb in adult post-stroke patients were included. MCID values were retrieved from specific databases. Two independent reviewers performed screening, data extraction, and assessment of methodological quality. Meta-analyses for both statistical significance and clinical relevance were performed. Clinical relevance was expressed as a standardized MCID overall score (SMOS) for each outcome measure, calculated as the difference between mean outcome measures in experimental and control groups divided by corresponding MCID, when available.

Results: Eighty-five studies were included. Conventional meta-analyses showed that RAT, compared to control, had significant effects in the domains of activities of daily living, dexterity, arm function, and strength, but not on pain. Meta-analyses for clinical relevance reported non-clinically relevant differences between groups for all domains.

Conclusion: RAT produces some significant improvements for the upper limb, but these differences are not clinically relevant when compared to other therapies. Improvements in using the RAT in clinical practice may not be more clinically relevant than other therapies for stroke patients.

Keywords: Clinical relevance, Data interpretation, Epidemiologic methods, Rehabilitation, Stroke

What is already known about this topic?

- *Randomized clinical trials usually focus their results on statistical significance rather than clinical relevance.*

What this study adds?

- *Robotic rehabilitation for the upper extremity after stroke produces statistically significant, but not clinically relevant, changes when compared to other therapies.*

Received: July 16, 2025
Accepted: April 24, 2025
Published online: May 23, 2025

This article includes supplementary material

Corresponding author:
Leonardo Pellicciari
email: leonardo.pellicciari@gmail.com

Introduction

Randomized clinical trials (RCTs) are intervention studies that apply an experimental design to establish causal relationships between interventions and outcomes. In an RCT, two or more groups are compared, and differences among them are reported in terms of statistical significance or estimating the effect size. However, small outcome changes that are statistically significant may lack clinical relevance.

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

Statistical significance means that two events, represented by two variables, occur jointly more often than what can be anticipated only by the effect of the chance (1), possibly – but not necessarily – because they are causally related. The effect size is used as an indicator of the magnitude of the effect of an independent variable on the dependent one.

However, to judge whether one treatment is more effective than another, simply knowing whether a difference exists is not enough; we need to know how big the difference is (2).

From a clinical perspective, it is more useful to report the results of an RCT in terms of their clinical relevance, that is, changes that patients acknowledge as beneficial or harmful (3).

The perception of relevant changes can be estimated within or across individuals or groups (4); while the Minimal Important Change (MIC) refers to the smallest change in health status perceivable by patients compared to a baseline, the minimal clinically important difference (MCID) represents the smallest difference between groups that is considered by patients to be clinically relevant (5).

Upper limb impairments are frequent after a stroke, causing activity limitations, participation restrictions, and poor quality of life (6). Previous studies showed that robot-assisted training (RAT), compared to other therapies, is associated with statistically significant improvement in upper limb motor function recovery after stroke (7,8). On the other hand, more recent findings reported no statistically significant differences between RAT and conventional therapy (9-11).

Despite a growing body of evidence on RAT effects, the clinical relevance of this treatment approach has never been synthesized within systematic reviews. Therefore, this systematic review aimed to assess the clinical relevance of RAT, expressed as MCID, in improving activities of daily living, arm function, and upper limb impairments in patients recovering after a stroke, compared to control treatments.

Methods

Study design

The study protocol was *a priori* registered on the PROSPERO database (Number: CRD42023387467). This systematic review was reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (12) (Supplemental Material 1).

Data sources and searches

To identify eligible studies, systematic literature searches were conducted on MEDLINE (via PubMed), EMBASE, CINAHL, and PEDro databases from their inception through February 2023. The following search strings were used for MEDLINE, EMBASE, and CINHAL: (stroke OR CVA OR cerebrovascular accident) AND (upper limb OR arm OR upper extremity OR hand) AND (robot* OR robotics OR electro-mechanic OR exoskeleton device) AND (RCT OR randomized controlled trial OR clinical trial OR controlled trial). For PEDro searches, the following keywords were used: stroke AND robotics, with "clinical trial" as a filter. Moreover, we checked the reference list of the included and related articles to identify other relevant publications (cross-referencing).

Search strategy for clinical relevance (MCID)

For each assessment tool used in the RCTs retrieved, the value of MCID was searched in the Rehabilitation Measures Database, provided by the National Institute on Disability, Independent Living, and the Rehabilitation Research Shirley Ryan AbilityLab (https://www.sralab.org/rehabilitation-measures, last access, March 2023) and Stroke Engine (https://strokengine.ca/en/, last access: March 2023) using the term "assessment" and applying a search restriction to the "stroke severity" field.

When the MCID value of a measurement instrument was not available, we carried out further research on PubMed, Google Scholar, and EMBASE, from their inception through February 2023, using the following keywords: ((MCID OR "Minimal clinically important difference") AND stroke AND "*name of the outcome measure*").

MCID values calculated in post-stroke populations were preferred; when they were unavailable, the MCID values estimated on different patient populations were also considered. When different MCID values were obtained with the same method on the same population, the lowest MCID value criterion was chosen.

Study selection and eligibility

Articles published in English or Italian were included in this systematic review if they reported findings from RCTs assessing the effects of RAT on motor and functional skills of the upper limb in adult post-stroke patients. Conference proceedings and studies that administered RAT to participants in the control group were excluded. No time restriction was applied.

Two authors (AU and SV) independently screened and assessed the studies retrieved for inclusion with the support of dedicated software (Rayyan®, Qatar Computing Research Institute. Qatar; www.rayyan.ai). First, duplicated records were removed; then, AU and SV excluded non-pertinent articles based on title and abstract; finally, the full texts were retrieved and assessed for definitive inclusion. Reasons for exclusion during the full-text reading were given for each study excluded. Any disagreements regarding inclusion were resolved through the involvement of a third researcher (MP).

Assessment of the risk of bias

The methodological quality of the trials included was assessed with the PEDro score (13), which was extrapolated from the PEDro database. When the PEDro score was unavailable, it was calculated independently by two authors (LP, MP), with the involvement of a third researcher (SV) in case of disagreement. This score includes eleven items; each scored as 0 (criterion not satisfied or unreported) or 1 point (criterion satisfied); because the eligibility criterion is not included in the final score, the total score ranges from 0 (worse methodological quality) to 10 points (highest methodological quality).

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

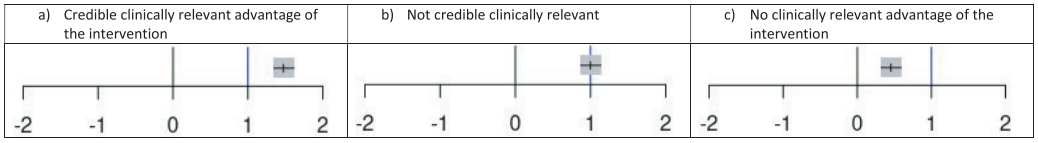


FIGURE 1 - Interpretation of the meta-analysis for clinical relevance.

Data extraction

Data were extrapolated independently by two authors (SV and AU), involving a third author (MP) in case of disagreement. The following data were extracted: author, year of publication, outcome measures with related statistics (i.e., sample size, mean, standard deviation (SD), median, inter-quartile range, minimum and maximum values, both at baseline and post-treatment), type of intervention in the experimental and control groups, duration of the study and the timing of the first follow-up. In the case of cross-over trials, the data from the first period only (i.e., before the crossing-over) was used.

Statistical analysis

Quantitative data synthesis was performed; in the meta-analyses, articles for which the MCIDs of the outcome measures used were not available were excluded.

Preliminarily, manipulation of the original outcome measures was sometimes required, as higher scores indicated better performance in some tools and worse performance in others. Thus, the scores were reversed in the latter to express all the results in the same direction of improving performance with increasing scores.

To obtain a traditional meta-analysis, the findings of each study were expressed in terms of standardized mean difference (SMD), with the corresponding 95% confidence interval (CI), and combined using both fixed effects and random effects models (Restricted Estimator Maximum Likelihood method).

The procedure to assess the clinical relevance of each study and to obtain an estimate of the pooled effect can be described as follows. In any given article, when an MCID value was available for the main outcome measure, the clinical relevance of the original findings was expressed as a standardized MCID overall score (SMOS), calculated as the difference between the mean of the outcome measures in the experimental and control groups divided by the corresponding MCID. The 95% CI associated with the SMOS point estimate was calculated using a pooled SD normalized by the MCID of the corresponding outcome measure. From the individual studies, the pooled estimate of the SMOS and the corresponding 95% CI were calculated, again using fixed and random effects models. For both the individual study and the pooled estimates, the interpretation of the SMOS is as follows (Figure 1): a) when the SMOS and the lower limit of its 95% CI are one or greater, the advantage offered by the experimental intervention over the control would reach or exceed the MCID and, therefore, the result would credibly be of clinical relevance; b) when the SMOS is one or greater but the lower limit of its 95% CI is not, the advantage offered by the experimental intervention would not reach an acceptable level of certainty in terms of clinical relevance; (c) when the SMOS point estimate is below 1, the experimental intervention would not offer any clinically relevant advantage. The interpretation of SMOS point estimate negative values and 95% CI would be similar but in favor of the control intervention. A detailed description of the calculations used to study clinical relevance with examples is provided in Supplementary Material 2.

All the results obtained were summarized in tables with forest plots.[12] Data analysis was performed with R, a free software environment for statistics, using the "meta" package [13]; in particular, the "metacont" function was used for the analysis of continuous data.

Statistical heterogeneity between eligible studies was assessed using the I^2 statistic and the Cochran Q test (χ^2) (14). The I^2 statistic was interpreted according to the following guide(14): 0-40% = not important statistical heterogeneity; 30-60% = moderate heterogeneity; 50-75% = substantial heterogeneity; 75-100% = considerable heterogeneity.

In case of heterogeneity, subgroup analyses were performed, grouping studies based on involved upper extremity segments, classified as proximal, distal, or both, and based on the robotic types, coded as unilateral, bilateral, or both. Additional sub-analyses were conducted by grouping studies based on the therapy content of control groups. Finally, a meta-regression using the duration of interventions as an independent factor was performed.

Results

Study selection

After removing duplicates, 5,302 articles were obtained. After reviewing the abstract and title, 5,190 studies were excluded for failing to meet the inclusion criteria, leaving 112 articles for full-text screening. After excluding 28 other articles, eighty-five RCTs were included in the systematic review. A list of the included articles is reported in Supplemental Material 3. Supplemental Material 4 reports a list of the studies excluded after full-text reading (during the study selection and cross-reference phases) with reasons for exclusion.

Since not all MCID values were found for all outcome measures or their subscales, seven RCTs were excluded, and 78 studies were included in the meta-analyses. The study selection process is shown in Figure 2.

MCID selection

MCID values of all measures were found except nine. All retrieved MCID values were referred to stroke populations, except for six outcome measures: QuickDASH (Disability of the Arm, Shoulder, and Hand) and Numeric Rating Pain Scale (NPRS) were referred to musculoskeletal disorders, Medical Research Council, Ashworth Scale, and Modified Ashworth

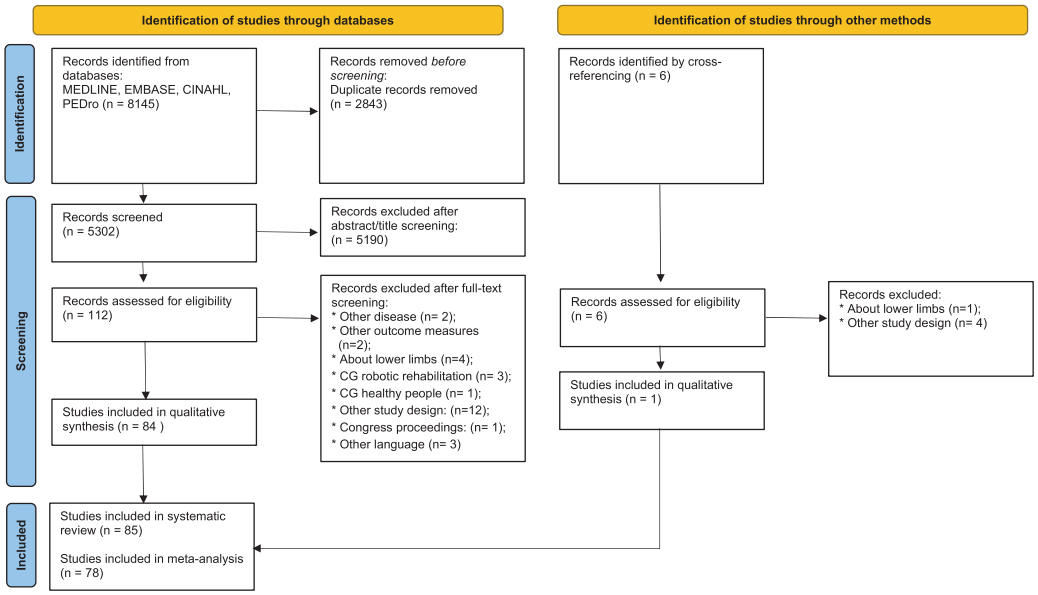


FIGURE 2 - Flow chart of the study procedure.

Scale to neurological population other than stroke and handgrip strength to the heterogeneous population.

For the analysis, outcome measures were pooled according to six domains (activities of daily living, arm function, muscle strength, dexterity, muscle tone, and pain). MCID values of scales assessing quality of life were not found. Domains, MCID values, and related references are reported in Supplemental Material 5. The classification of outcomes is based on the review by Mehrholz et al. (7). The three original outcomes were expanded within the domain of impairments, including dexterity, muscle tone, and pain, since additional outcomes were found in the included trials.

Characteristics of the included studies

Detailed characteristics of the included studies are reported in Table 1. In the systematic review, 3452 patients with stroke were included, with a mean±SD number per study of 40.6±40.8 (range: 11, 224). The experimental treatments included a high variability of devices, while the most commonly used control intervention was defined as "conventional physiotherapy" or equivalent terms (n = 60), followed by occupational therapy (n = 11). The included studies used 77 different outcome measures, the most commonly used of which was the Fugl-Meyer Assessment. The treatment duration was 45.2±25.8 days (range: 12, 180).

TABLE 1 - Characteristics of the included studies

Study	N	Experimental group	Control group	Outcome	Treatment duration*
Abdullah et al. 2011	20	Name not available	Conventional physiotherapy	Likert scale	70
Aisen et al. 1997	20	MIT - MANUS	Conventional physiotherapy	MSS, FIM, FMA (UE)	14
Ang et al. 2014	14	Haptic Knob	Conventional physiotherapy	FMA (UE)	12
Aprile et al. 2020	224	Motore Humanware, Amadeo Tyromotion, Pablo Tyromotion	Conventional physiotherapy	NRPS, FMA (UE), BI, MI	42
Bayindir et al. 2022	32	HandTutor	Conventional physiotherapy	FMA (UE), BBT, NHPT	35
Brokaw et al. 2014	11	ARMin III HandSOME	Conventional physiotherapy	ARAT	30
Budhota et al. 2021	44	H-Man	Conventional physiotherapy	Grip, FMA (UE), ARAT	42
Burgar et al. 2011	35	MIME	Conventional physiotherapy	Ashworth, WMFT (FAS), FIM, FMA, MP	60
Bustamante Valles et al. 2016	20	TheraDrive, NESS H200, NESS L300	Conventional physiotherapy and occupational therapy	BBT, FMA (UE)	49
Calabrò et al. 2019	50	Amadeo	Conventional physiotherapy	FMA (UE), NHPT	56
Cameirao et al. 2011	16	Rehabilitation Gaming System	Occupational therapy or non-specific interactive games	FMA (H/W), FMA (Arm), BI, FMA (UE), CAHAI, MI	84

(*Continued*)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 1 - (*Continued*)

Study	N	Experimental group	Control group	Outcome	Treatment duration*
Carpinella et al. 2020	40	Braccio di Ferro	Conventional physiotherapy	FIM, FMA (UE)	28
Chen et al. 2021	20	Name not available	Conventional physiotherapy	FMA (H/W), FMA (Arm), FMA (UE), ARAT, BI	28
Chen et al. 2022	31	The Hand of hope	Conventional physiotherapy	MAL (QOM), MAL (AOU), WMFT (TIME), WMFT (FAS), FMA (UE), ARAT	28
Chinembiri et al. 2020	45	The Fourier M2 machine	Occupational therapy	BI, FMA (UE), FMA (Arm)	42
Connelly et al. 2010	16	PneuGlove	Conventional physiotherapy	FMA (H/W), FMA (UE), BBT, Pinch, Grip	42
Conroy et al. 2011	41	InMotion 2.0	Intensive conventional arm exercise program	SIS (Mobility), SIS (ADL), SIS (Hand), FMA (UE), WMFT	42
Coskunsu et al. 2022	24	Hand of hope	Conventional physiotherapy	MAL (QOM), MAL (AOU), ARAT	21
Daly et al. 2005	12	InMotion2	Functional neuromuscular stimulation and motor learning	FMA	84
Daunoraviciene et al. 2018	34	Armeo Spring	Occupational therapy	HAD, FIM, ACE-R, FMA (UE)	84
De Araújo et al. 2011	12	Name not available	Conventional physiotherapy	MAS (Elbow), MAS (W/H), FMA (H/W), FMA (Arm), FMA (UE)	56
Dehem et al. 2019	45	REAplan1	Conventional physiotherapy	BBT, ABILHAND, Activlim, WMFT (FAS), SIS (ADL), FMA (UE)	63
Fasoli et al. 2004	56	MIT-MANUS	Conventional physiotherapy	MRC, FIM (Self-care), FIM, FMA (UE)	49
Fazekas et al. 2007	30	REHAROB	Bobath	Rivermead arm score, MAS (Elbow), FMA (Arm), ROM, FIM (Self Care)	20
Franceschini et al. 2019	48	InMotion2	Conventional physiotherapy	MAS (Shoulder), MAS (Elbow), FMA (UE), pROM	90
Frisoli et al. 2022	26	L-EXOS	Conventional physiotherapy	FMA (W/H), FMA (Arm), FMA (UE), MAS	42
Gandolfi et al. 2019	32	Armotion	Conventional physiotherapy + botulinum toxin	FMA (UE)	35
Grigoras et al. 2016	25	Hybrid glove	Conventional physiotherapy	SIS (ADL), FMA (H/W), BBT, FMA (UE), FMA (Arm)	21
Gueye et al. 2021	50	Armeo Spring	Standard rehabilitation program + conventional physiotherapy	MoCA, FIM, FMA (UE)	21
Hesse et al. 2005	44	Bi-Manu-Track	Conventional physiotherapy + electromyography electrical stimulation	MAS, MRC, FMA (UE)	42
Hesse et al. 2014	50	Bi-Manu-Treck, RehaDigit, Reha Slide, Reha Slide duo	Double sessions of individual arm therapy	BBT, MAS, MRC, ARAT, FMA (UE), BI	28
Housman et al. 2009	34	T-WREX	Conventional exercises	MAL (AOU), MAL (QOM), Grip, FMA (UE)	42
Hsieh et al. 2011	12	Bi-Manu-Track	Conventional physiotherapy	ABILHAND, MAL (QOM), MAL (AOU), MRC, FMA (UE)	30

Study	N	Experimental group	Control group	Outcome	Treatment duration*
Hsieh et al. 2012	36	Bi-Manu-Track	Intensive therapist-administered control therapy	FMA (H/W), FMA (Arm), FMA (UE)	30
Hsieh et al. 2014	32	Bi-Manu-Track	Conventional physiotherapy	MAL (AOU), MAL (QOM), WMFT (FAS), FMA (H/W), FMA (Arm), FMA (UE)	30
Hsieh et al. 2017	31	Bi- Manu- Track	Task-oriented approach	MRS, Grip, BBT, SIS (strength), SIS (ADL), SIS (Mobility), SIS (Hand), FMA (UE), FIM	28
Hsiu et al. 2019	43	Bi-Manu-Track	Usual care	FMA (Arm), MAL (QOM), MAL (AOU), FMA (UE)	28
Hsiu et al. 2021	34	TIGER	Occupational therapy + usual care	MAL (QOM), MAL (AOU), FMA (Arm), FMA (UE), BBT	63
Hwang et al. 2016	17	Amadeo	Conventional physiotherapy	MAS (Elbow), Pinch, FMA (Arm), FMA (H/W), Grip, SIS (ADL), NHPT	30
Iwamoto et al. 2019	12	Hybrid Assistive Limb - SJ	Occupational physiotherapy	MAL (AOU), MAL (QOM), Grip, MI, BI, FIM	28
Kim et al. 2017	38	InMotion ARM	Conventional physiotherapy	VAS, K-SDQ, pROM	30
Klamroth-Marganska 2014	73	ARMin	Conventional physiotherapy	FMA (UE), WMFT (FAS), WMFT (TIME), MAL (QOM), MAS, Grip	60
Kuo et al. 2022	18	TIGER	Task-specific motor training	MAL (QOM), MAL (AOU), FMA (UE), BBT	28
Kutner et al. 2010	18	Hand Mentor robotic system	Therapist-supervised repetitive task practice	SIS (Hand), SIS (Mobility), SIS (ADL)	21
Lee et al. 2016	58	Robot Neuro-X	Conventional physiotherapy	MAS, BI	14
Lee et al. 2018	30	REJOYCE	Occupational therapy	BI, FMA (UE)	60
Lee et al. 2021	29	Gloreha	Conventional physiotherapy	Grip, FMA (W/H), FMA (Arm), FMA (UE), BBT	42
Lencioni et al. 2021	32	Braccio di ferro	Usual care	FMA	nr
Liao et al. 2011	40	Bi-Manu-Track	Dose-matched active control therapy	ABILHAND, MAL (QOM), MAL (AOU), FIM, FMA (UE)	28
Lin et al. 2022	222	Arm1	Conventional physiotherapy	FMA (Arm), FMA (UE)	21
Lo et al. 2010	25	Name not available	Intensive conventional physiotherapy	Ashworth, FMA, SIS (ADL), WMFT	84
Lum et al. 2002	27	MIME	Neurodevelopmental therapy	FIM, BI	60
Lum et al. 2006	16	MIME	Conventional physiotherapy	FIM, MP	180
Masiero et al. 2006	20	NeReBot	Conventional physiotherapy	MRC (Elbow), FMA (H/W), FMA (Arm), FIM- mot, upMI	30
Masiero et al. 2007	35	NeReBot	Conventional physiotherapy	MRC (Deltoid), MRC (Biceps), FMA (H/W), FMA (Arm)	35
Masiero et al. 2011	21	NeReBot	Conventional physiotherapy	MAS, MRC (Biceps), MRC (Deltoid), FAT, FMA (H/W), FMA (Arm), BBT, FMA (UE), FIM-mot	90
Masiero et al. 2014	30	NeReBot	Conventional physiotherapy	MAS, MRC (Biceps), MRC (Deltoid), FAT, FMA (H/W), FMA (Arm), FMA (UE), BBT, FIM-mot	35
MC Cabe et al. 2015	25	InMotion2	Motor learning	FMA (Arm), FMA (H/W), FMA (Coordination)	35

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 1 - (*Continued*)

Study	N	Experimental group	Control group	Outcome	Treatment duration*
Orihuela Espina et al. 2015	17	Amadeus Tyromotion	Occupational therapy	FMA (H/W), MI	75
Page et al. 2013	16	Myomo	Repetitive task-specific practice	SIS (Strength), COPM-P, COPM-S, SIS (ADL), FMA (UE), SIS (Mobility)	56
Rabadi et al. 2008	20	MIT-Manus	Occupational therapy + conventional physiotherapy	VAS, MAS, FMA (Arm), FMA (H/W), MSS (Arm), ARAT, MSS (Hand), MP	56
Ramos-Murguialday et al. 2019	28	Name not available	Random orthoses movements + conventional physiotherapy	GAS, FMA (H/W), FMA (Arm), MAS, MAL	30
Ranzani et al. 2020	33	ReHaptic Knob	Conventional physiotherapy	FMA (W/H), FMA (Arm), FMA (UE), BBT	28
Reinkensmeyer et al. 2012	26	Pneu-WREX I	Conventional tabletop therapy	FMA (UE), MAL (AOU), MAL (QOM), Grip, BBT	56
Rémy Néris et al. 2021	215	Armeo Spring	Conventional physiotherapy	FMA (UE), ARAT, FIM, SIS (Hand)	28
Rodgers et al. 2020	511	MIT-Manus, InMotion	Conventional physiotherapy	FMA (UE), ARAT, BI, SIS (ADL), SIS (Hand), SIS (Mobility), NRPS	90
Sale et al. 2013	20	Amadeo Robotic System	Occupational therapy	MRC, MAS, BBT, FMA (UE), MI	35
Sale et al. 2014	53	MIT-MANUS/ InMotion2	Conventional physiotherapy	MAS (Elbow), MAS (Shoulder), FMA (UE), MI, pROM	42
Singh et al. 2022	27	Name not available	Conventional physiotherapy	MAS, BI, FMA (UE), FMA (W/H), FMA (Arm)	28
Straudi et al. 2019	40	Reo Therapy System	Conventional physiotherapy	FMA (UE)	21
Susanto et al. 2015	19	Name not available	Non-assisted finger training groups	FMA (H/W), FMA (Arm), FMA (UE), WMFT, ARAT	35
Takahashi et al. 2016	56	ReoGo	Conventional physiotherapy + self-training	MAL (QOM), MAL (AOU), FMA, WMFT	42
Takebayashi et al. 2022	78	ReoGo-J	Conventional physiotherapy	FMA (UE)	70
Taravati et al. 2022	45	ReoGo	Conventional physiotherapy	FMA (W/H), FMA (UE), FIM, Grip	28
Taveggia et al. 2016	27	Armeo Spring	Physical and rehabilitation medicine+ conventional physiotherapy	MAS, NRPS, MI, FIM	84
Terranova et al. 2021	51	InMotion	Conventional physiotherapy	FMA (UE), WMFT	84
Timmermans et al. 2014	22	Haptic Maste	Task-oriented non-robotic arm-hand training	FMA (UE), ARAT, MAL (AOU), MAL (QOM), MAL	56
Tomic et al. 2017	26	ArmAssist	Conventional arm training	WMFT FAS (Arm), WMFT-FAS, FMA (Arm), FMA (UE), BI	21
Vanoglio et al. 2016	30	Gloreha	Conventional hand rehabilitation	NHPT, Pinch, Grip, QuickDASH, MI	42
Villafane et al. 2017	32	Gloreha	Conventional physiotherapy + occupational therapy	MAS, NIHSS, VAS, QuickDASH, BI, MI	21
Volpe et al. 2000	56	MIT-Manus	Conventional physiotherapy	MSS (Hand), FMA (H/W), FMA (Arm), MSS (Arm), MP	30
Volpe et al. 2008	21	Mit-Manus	Intensive movement protocol	VAS, FMA (H/W), FMA (Arm), MAS, ARAT, SIS (Hand), MP	90

Study	N	Experimental group	Control group	Outcome	Treatment duration*
Wolf et al. 2015	99	Hand Mentor Pro	Home exercise program	FMA (H/W), FMA (Arm), FMA (UE), ARAT	56
Wu et al. 2012	28	Bi-Manu-Track	Therapist-based bilateral arm training	MAL (AOU), MAL (QOM), FMA (Arm), FMA (H/W), FMA (UE), SIS (Strength), SIS (ADL), SIS (Hand)	28
Yoo et al. 2013	22	ReoGo	Conventional physiotherapy	Grip, BI, BBT, WMFT	42

Abbreviations: ACE-R, Addenbrooke Cognitive Examination-Revised; Active ROM, Active Range of Motion; ADL, activities of daily living; AMAT, Arm Motor Ability Test; ARAT, Action Research Arm Test; BBT, Box and Block Test; BI, Barthel Index; CAHAI, Chedoke Arm and Hand Activity Inventory; CMSA, Chedoke McMaster Stroke Assessment of the Arm and Hand; COPM-P, Canadian Occupational Performance Measure Performance; COPM-S, Canadian Occupational Performance Measure Satisfaction; FAS, Functional ability scale; FAT, Frenchay Arm Test; FIM, Functional Independence Measure; FMA, Fugl-Meyer Motor Assessment; FMA (H/W), Fugl-Meyer Motor Assessment Hand/Wrist; FMA (UE) , Fugl-Meyer Motor Assessment Upper Extremity; GAS, Goal Attainment Scale; FIM mot, Functional Independence scale- motor; HAD, Hamilton Rating Scale for Depression; K-SDQ: Korean version of the Shoulder Disability Questionnaire; MAL, Motor Activity Log; MAL AOU, Motor Activity Log Amount Of Movement; MAL QOM, Motor Activity Log Quality Of Movement; MAS, Modified Ashworth Scale; MAS (W/H) Modified Ashworth Scale Wrist and Hand; mBI, modified Barthel index; MI, Motricity Index; MIME, mirror image motion enabler; MIT-Manus, robotic device developed at the Massachusetts Institute of Technology; MoCA, Montreal Cognitive Assessment; MP, Motor Power Scale; MRC, Medical Research Council; MRS, Modified Rankin Scale; MSS, Motor Status Score; NHPT, Nine hole peg test; NIHSS, National Institutes of Health Stroke Scale; NRPS, Numeric Rating Pain Scale; pROM, passive Range Of Motion; QuickDASH: the short version of the Disabilities of the Arm Shoulder and Hand; ROM, Range Of Movement; SIS, Stroke Impact Scale; UE, upper extremity; VAS, Visual Analogue Scale; MI, Motricity Index upper extremity; WMFT, Wolf Motor Function Test
* Treatment duration is reported in days

Risk of bias in included studies

The methodological quality of the included studies is reported in Table 2. The PEDro score was 6.1±1.3 (range: 2, 8). Items 2, 10, and 11 were the most frequently satisfied (98.8%, 97.6%, and 96.4%, respectively), while items 3, 5, and 6 were the least often satisfied (35.7%, 2.4%, and 0.0%, respectively).

Meta-analyses

The summary results of the meta-analyses are reported in Table 3. The forest plots considering the statistical significance for each domain are detailed in Supplemental Material 6, and those reporting the clinical relevance for each domain in Supplemental Material 7.

TABLE 2 - Methodological quality of the included studies

Study	Item 1	Item 2	Item 3	Item 4	Item 5	Item6	Item 7	Item 8	Item 9	Item 10	Item 11	Total score
Abdullah et al. 2011	Y	Y	Y	Y	N	N	Y	Y	N	Y	N	6
Aisen et al. 1997	Y	N	N	Y	N	N	Y	Y	N	N	Y	4
Ang et al. 2014	Y	Y	N	Y	N	N	Y	Y	N	Y	Y	6
Aprile et al.2020	Y	Y	N	Y	N	N	Y	N	Y	Y	Y	6
Bayindir et al. 2022	N	Y	Y	Y	N	N	N	Y	N	Y	Y	6
Brokaw et al. 2013	Y	Y	N	N	N	N	Y	N	N	Y	Y	4
Budhota et al. 2021	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	8
Burgar et al. 2011	Y	Y	N	Y	N	N	Y	Y	N	Y	Y	6
Bustamante Valles et al. 2016	Y	Y	N	N	N	N	N	N	N	Y	Y	3
Calabrò et al. 2019	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	8
Cameirao et al. 2011	Y	Y	N	Y	N	N	Y	N	N	Y	Y	5
Carpinella et al. 2020	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	8
Chen et al.2021	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	8
Chen et al. 2022	Y	Y	N	Y	N	N	Y	N	N	Y	Y	5

(*Continued*)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - (*Continued*)

Study	Item 1	Item 2	Item 3	Item 4	Item 5	Item6	Item 7	Item 8	Item 9	Item 10	Item 11	Total score
Chinembiri et al. 2020	Y	Y	N	N	N	N	Y	Y	N	Y	Y	5
Connelly et al. 2010	Y	Y	N	Y	N	N	N	Y	N	Y	N	4
Conroy et al. 2011	Y	Y	N	Y	N	N	Y	Y	N	Y	Y	6
Coskunsu et al. 2022	Y	Y	N	Y	N	N	Y	N	N	Y	Y	5
Daly et al., 2005	Y	Y	N	Y	N	N	Y	Y	N	N	Y	5
Daunoraviciene et al. 2018	N	Y	N	Y	N	N	N	Y	N	Y	Y	5
De Araújo et al. 2011	N	Y	N	Y	N	N	N	Y	N	Y	Y	5
Dehem et al. 2019	Y	Y	Y	Y	N	N	Y	N	Y	Y	Y	7
Fasoli et al. 2004	Y	Y	N	Y	N	N	Y	Y	N	Y	Y	6
Fazekas et al. 2007	N	Y	N	N	N	N	N	Y	N	N	N	2
Franceschini et al. 2019	Y	Y	Y	Y	N	N	Y	N	N	Y	Y	6
Frisoli et al. 2022	Y	Y	N	Y	N	N	N	N	N	Y	Y	4
Gandolfi et al. 2019	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	8
Grigoras et al. 2016	Y	Y	N	Y	N	N	N	Y	Y	Y	Y	6
Gueye et al. 2021	Y	Y	N	Y	N	N	N	Y	N	Y	Y	5
Hesse et al. 2005	Y	Y	Y	N	N	N	Y	Y	Y	Y	Y	7
Hesse et al. 2014	Y	Y	N	Y	N	N	Y	Y	Y	Y	Y	7
Housman et al. 2009	Y	Y	N	Y	N	N	Y	N	N	Y	Y	5
Hsieh et al. 2011	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	8
Hsieh et al. 2012	N	Y	N	Y	N	N	Y	Y	Y	Y	Y	7
Hsieh et al. 2014	N	Y	Y	Y	N	N	Y	Y	Y	Y	Y	8
Hsieh et al. 2017	N	Y	Y	Y	N	N	Y	Y	N	Y	Y	7
Hsiu et al. 2019	N	Y	Y	Y	N	N	Y	Y	Y	Y	Y	8
Hsiu et al. 2021	Y	Y	Y	Y	N	N	Y	Y	N	Y	Y	7
Hwang et al. 2016	Y	Y	N	Y	N	N	Y	Y	N	Y	Y	6
Iwamoto et al. 2019	Y	Y	N	Y	N	N	N	N	N	Y	Y	4
Kim et al. 2017	Y	Y	Y	Y	N	N	N	Y	N	Y	Y	6
Klamroth Marganska et al. 2014	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	8
Kuo et al. 2022	Y	Y	Y	Y	N	N	Y	Y	N	Y	Y	7
Kutner et al. 2010	N	Y	Y	Y	N	N	Y	N	N	Y	Y	6
Lee et al. 2016	Y	Y	N	Y	N	N	N	N	N	Y	Y	4
Lee et al. 2018	Y	Y	N	Y	N	N	Y	Y	N	Y	Y	6
Lee et al. 2021	Y	Y	N	Y	N	N	Y	Y	Y	Y	Y	7
Lencioni et al. 2021	Y	Y	N	Y	N	N	Y	N	N	Y	Y	5
Liao et al. 2011	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	8
Lin et al. 2022	Y	Y	N	Y	N	N	Y	Y	Y	Y	Y	7
Lo et al. 2010	Y	Y	N	Y	N	N	Y	Y	Y	Y	Y	7

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Study	Item 1	Item 2	Item 3	Item 4	Item 5	Item6	Item 7	Item 8	Item 9	Item 10	Item 11	Total score
Lum et al. 2002	N	Y	N	Y	N	N	Y	Y	N	Y	Y	6
Lum et al. 2006	N	Y	N	N	N	N	Y	N	N	Y	Y	4
Masiero et al. 2006	Y	Y	N	Y	N	N	Y	N	N	Y	Y	5
Masiero et al. 2007	Y	Y	Y	Y	N	N	Y	N	Y	Y	Y	7
Masiero et al. 2011	N	Y	N	Y	N	N	Y	Y	N	Y	Y	6
Masiero et al. 2014	Y	Y	N	Y	N	N	Y	N	Y	Y	Y	6
MC Cabe et al. 2015	N	Y	N	Y	N	N	Y	Y	Y	Y	Y	7
Orihuela Espina et al. 2015	Y	Y	N	Y	N	N	N	Y	Y	Y	Y	6
Page et al. 2013	Y	Y	N	Y	N	N	Y	N	N	Y	Y	5
Rabadi et al. 2008	Y	Y	Y	N	N	N	Y	Y	N	Y	Y	5
Ramos-Murguialday et al. 2019	N	Y	N	N	N	N	Y	Y	N	Y	Y	5
Ranzani et al. 2020	Y	Y	N	Y	N	N	Y	N	Y	Y	Y	6
Reinkensmeyer et al. 2012	Y	Y	N	Y	N	N	Y	Y	N	Y	Y	6
Rémy Néris et al. 2021	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	8
Rodgers et al. 2020	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	8
Sale et al. 2013	N	Y	N	N	N	N	Y	Y	Y	Y	Y	6
Sale et al. 2014	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	8
Singh et al. 2022	Y	Y	Y	Y	N	N	Y	N	N	Y	Y	6
Straudi et al. 2019	N	Y	Y	Y	N	N	Y	N	Y	Y	Y	7
Susanto et al. 2015	N	Y	N	Y	N	N	Y	Y	Y	Y	Y	7
Takahashi et al. 2016	Y	Y	N	N	N	N	Y	Y	N	Y	Y	5
Takebayashi et al. 2022	Y	Y	N	Y	N	N	Y	Y	Y	Y	Y	7
Taravati et al. 2022	Y	Y	N	Y	N	N	Y	N	N	Y	Y	5
Taveggia et al. 2016	N	Y	N	Y	N	N	Y	Y	Y	Y	Y	7
Terranova et al. 2021	Y	Y	N	Y	N	N	N	N	Y	Y	Y	5
Timmermans et al. 2014	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	8
Tomic et al. 2017	Y	Y	N	Y	N	N	Y	Y	Y	Y	Y	7
Vanoglio et al. 2016	Y	Y	Y	Y	N	N	Y	Y	N	Y	Y	7
Villafane et al. 2017	N	Y	N	Y	N	N	Y	Y	Y	Y	Y	7
Volpe et al. 2000	Y	Y	N	Y	N	N	Y	Y	N	Y	Y	6
Volpe et al. 2008	Y	Y	N	Y	N	N	Y	Y	N	Y	Y	6
Wolf et al. 2015	Y	Y	N	Y	N	N	Y	Y	Y	Y	Y	7
Wu et al. 2012	N	Y	Y	Y	Y	N	Y	Y	N	Y	Y	8
Yoo et al. 2013	N	Y	N	Y	N	N	Y	N	N	Y	Y	5

Notes: Item 1: Eligibility criteria, Item 2: Random allocation, Item 3: Concealed allocation, Item 4: Baseline comparability, Item 5: Blind subjects, Item 6: Blind therapists, Item 7: Blind assessors, Item 8: Adequate follow-up, Item 9: Intention-to-treat analysis, Item 10: Between-group comparisons, Item 11: Point estimates and variability

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

In terms of statistical significance, the domains of activity of daily living (SMD = 0.29; 95% CI: 0.15; 0.43), dexterity (SMD = 0.19; 95% CI: 0.01; 0.37), arm function (SMD = 0.27; 95% CI: 0.17; 0.38), and strength (SMD = 0.44; 95% CI: 0.17; 0.71) showed a significant effect in favor of the experimental group. Conversely, a non-significant effect in favor of the experimental treatment was observed for the domain of pain (SMD = 0.53; 95% CI: −0.15; 1.20), while a non-significant effect in favor of the control group was achieved in the domain of muscle tone (SMD = −0.02; 95% CI: −0.26; 0.23).

In terms of clinical relevance, the forest plots reported non-clinically relevant findings for all the domains considered (SMOS < 1.00 for all meta-analyses).

Heterogeneity ranged from negligible to substantial for all analyses, ranging from 0% to 85% for the meta-analyses considering statistical significance and from 0% to 91% for those exploring clinical relevance.

TABLE 3 - Summary of the meta-analyses

Domain	Statistical significance								Clinical relevance							
	N	*Fixed*	*95% CI*	*Random*	*95% CI*	*Q*	*Tau²*	*I²*	*N*	*Fixed*	*95% CI*	*Random*	*95% CI*	*Q*	*Tau²*	*I²*
Activity of daily living	34	0.20	0.11, 0.30	0.29	0.15, 0.43	58.52	0.06	44	34	0.13	0.08, 0.18	0.17	0.08, 0.26	66.43	0.02	50
Dexterity	18	0.19	0.01, 0.37	0.19	0.01, 0.37	14.84	0.00	0	18	0.30	0.06, 0.53	0.30	0.06, 0.53	16.26	0.00	0
Arm function	73	0.26	0.19, 0.33	0.27	0.17, 0.38	140.18	0.09	49	73	0.30	0.23, 0.37	0.45	0.27, 0.64	204.74	0.27	65
Muscle Tone	19	0.07	−0.09, 0.22	−0.02	−0.26, 0.23	40.37	0.16	55	19	−0.08	−0.17, 0.02	0.01	−0.22, 0.23	54.37	0.11	67
Pain	7	0.12	−0.02, 0.26	0.53	−0.15, 1.20	39.71	0.72	85	7	0.31	0.11, 0.51	0.65	−0.13, 1.42	70.55	0.95	91
Strength	23	0.40	0.24, 0.56	0.44	0.17, 0.71	56.95	0.26	61	23	0.39	0.24, 0.54	0.48	0.18, 0.78	90.79	0.24	76

Abbreviations: CI, confidence interval; I^2, I^2 statistics; N, number of studies included in the meta-analysis; Q, Q statistics

Subgroup analyses

When the specific interventions applied in the control groups were considered, the following subgroup analyses were performed: conventional physiotherapy, occupational therapy, conventional physiotherapy plus occupational therapy, intensive therapy, task-oriented approach, virtual reality, and others.

Heterogeneity showed a general reduction in subgroup analyses, although it remains substantial or considerable for some domains. However, the results were similar to those found in the general meta-analyses, both in terms of statistical significance and clinical relevance (Supplemental Materials 8 and 9). The meta-regression analyses using the duration of interventions as an independent factor are shown in Supplemental Material 10.

Discussion

This systematic review with meta-analysis showed that the RAT, compared to other physiotherapy, produces statistically significant improvements in activities of daily living, dexterity, arm function, and strength but not in muscle tone and pain in patients with stroke. However, none of these improvements were clinically relevant.

Our results showed that the RAT produced significant effects in improving daily living activity, dexterity, arm function, and strength but not in muscle tone (15). RAT uses active or actively assisted movements. This technique is used in clinical practice when the goal is to improve function, dexterity and strength. Instead, usually, passive techniques and stretching are used to reduce muscle tone (fundamentals on which RAT is not based) (16). Therefore, it could be reasonable that RAT is not able to reduce muscle tone more than other therapies. Several recent systematic reviews on RAT effects in patients with stroke have been published, reporting different conclusions. The majority of reviews report that RAT was effective in improving upper limb motor function recovery (or, at least, some of the function-related outcomes) when compared to other treatments (7,8,17). In contrast, other reviews reported no significant difference between RAT and conventional physiotherapy (9-11). Conflicting results among these studies are concerned with the study methodology and the different strategy analyses. However, the conclusions of these studies were only based on the statistical significance of the differences and not their clinical relevance. Hence, the findings of an RCT or meta-analysis might be significant but irrelevant in practice.

Our results showed that clinical relevance was never reached for all the outcome domains. Discussing clinical relevance is useful for those who must apply the results of clinical research in daily clinical practice. From the perspective of evidence-based practice, it is certainly crucial that the result of the clinical study is statistically significant (i.e., that the difference evaluated is not attributable to chance but to the real

effectiveness of the treatment). Yet, it is even more crucial that a result is clinically relevant (i.e., that the effect size is of such a magnitude that it is useful for clinically improving the patient's condition), implying that, even if true within the study, these results may not be so useful in determining a clinically significant change in the patient's condition in clinical practice.

In the research field, authors of scientific articles are more likely to report and discuss the statistical significance of a result rather than its clinical relevance. Verhagen et al. (18) analyzed frequencies and proportions of reporting clinical relevance of RCTs published between 2000 and 2018 in six major physiotherapy peer-reviewed journals; over 40% of RCTs failed to report the clinical relevance of the results. However, the attention to clinical relevance has been increasing with time; many publications still fail to document and interpret study findings accordingly. Two previous studies systematically investigated this issue by analyzing RCT trials for low back pain (19,20), recommending reporting clinically important and statistically significant differences in meta-analyses to identify the real, meaningful effect. Van Tulder et al. (19) transformed means and standard deviations to a 0 to 100 scale and considered results clinically important when their magnitude was 20% or more for pain and 10% or more for functioning. Differently, Gianola et al. (20) compared between-group differences with the planned MCID reported in the sample size calculation, determining if the effect size reached clinical relevance. The authors classified clinical relevance as 'achieved' if the point estimate of the mean difference was equal or greater than the a priori planned MCID and 'not achieved' in the other case. In both studies, only a minority of RCT findings were statistically significant and clinically relevant. Unlike previous investigations, we performed analyses using an approach similar to the mean difference calculation (i.e., when the studies all assess the same domain but measure it using different scales). This approach allows us to (I) estimate the effect using a scale-free measure, (II) clarify the size of the effect, and (III) compare and summarize results when studies use different outcome scales (21).

Eighty-five studies were included in the systematic review, while 78 studies were included in the meta-analysis because none of the MCIDs for each outcome measure for the seven studies were found; indeed, we retrieved 27 out of 36 MCID values. The lack of known clinically important values (i.e., MCID values) may hinder researchers from reporting and interpreting their findings concerning clinical relevance. Future research that aims to determine MCIDs for core outcome measures is necessary. In addition, not all the MCIDs included in this review were related to stroke populations, which should be considered when results are interpreted. The MCID construct is not clearly defined, and different index categories could cause uncertainty in the results (22). For example, within-group methods might be used to calculate the MCID, providing more probably the value of the within-group MIC (23). Finally, grouping outcomes in domains could have sometimes been arbitrary since the construct of some measures included more than one domain (e.g., arm function and activities of daily living or dexterity).

An additional concern may be related to the content of the so-called "conventional physiotherapy." Previous studies on physiotherapy interventions in neurological conditions, including stroke (24-26), reported inadequate descriptions of the interventions, especially when they are defined as "conventional physiotherapy." Without strong evidence for the superiority of one treatment over others, conventional physiotherapy (and equivalent terms) may be related to a variety of practice approaches since they are considered complex interventions.

High heterogeneity was found for all analyses, despite sub-group analyses, according to upper extremity segments, robotic types, and content of control groups, not clarifying the heterogeneity. However, this finding was not unexpected because the studies retrieved differed in treatment dosages and applications and control groups' treatment contents (27). In fact, potential factors influencing outcomes in stroke survivors are the dose and intensity of therapy (28). These variables have not been included in the analyses because they are underreported in the included studies. The quality of reporting of interventions used in RCTs could be improved following the Template for Intervention Description and Replication (TIDieR) (29) or the Consensus on Exercise Reporting Template (CERT) (30). However, the results of subgroup analyses in terms of both statistical significance and clinical relevance confirmed those obtained by overall analyses, also when a meta-regression was conducted using the duration of interventions as a covariate.

Finally, our results should be used with caution: considering that the MCID is influenced by the clinical conditions and the phases post-stroke of the subjects studied (31), including also MCID from populations other than stroke patients could have influenced the results of this study.

Conclusion

In conclusion, RAT produces significant clinical improvements for the upper limb in some domains (i.e., activity of daily living, dexterity, arm function, and strength), but these improvements are not clinically relevant when compared with other therapies. Using the RAT in clinical practice can produce improvement that may not be clinically relevant for patients with stroke. Future research in this field should consider the clinical relevance when interpreting results. Because a substantial heterogeneity remained unexplained for some analyses, pooled data should be interpreted with caution.

Disclosures

Conflict of interest: The authors reported no conflict of interest

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author's contributor role: Conceptualization: AU, MP; Data curation: SV, AU; Formal Analysis: AU; Investigation: SV; Methodology: LP, MP; Project administration: SV; Resources: AU; Software: AU; Supervision: MP; Validation: MDB; Visualization: MDB; Writing – original draft: LP, MP; Writing – review & editing: all authors

Data availability statement: Data sharing not applicable: No new data were created or analyzed in this study

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

References

1. Tal J. Strategy and statistics in clinical trials: a non-statisticians guide to thinking, designing and executing. Academic Press; 2011.
2. Kamper SJ. Interpreting outcomes 1 — change and difference: linking evidence to practice. J Orthop Sports Phys Ther. 2019;49(5):357-358. CrossRef PubMed
3. Kamper SJ. Interpreting outcomes 2 — statistical significance and clinical meaningfulness: linking evidence to practice. J Orthop Sports Phys Ther. 2019;49(7):559-560. CrossRef PubMed
4. de Vet HC, Terluin B, Knol DL, et al. Three ways to quantify uncertainty in individually applied "minimally important change" values. J Clin Epidemiol. 2010;63(1):37-45. CrossRef PubMed
5. Qin Z, Zhu Y, Shi DD, et al. The gap between statistical and clinical significance: time to pay attention to clinical relevance in patient-reported outcome measures of insomnia. BMC Med Res Methodol. 2024;24(1):177. CrossRef PubMed
6. Morris JH, van Wijck F, Joice S, et al. Predicting health related quality of life 6 months after stroke: the role of anxiety and upper limb dysfunction. Disabil Rehabil. 2013;35(4):291-299. CrossRef PubMed
7. Mehrholz J, Pohl M, Platz T, et al. Electromechanical and robot-assisted arm training for improving activities of daily living, arm function, and arm muscle strength after stroke. Cochrane Database Syst Rev. 2018;9(9):CD006876. CrossRef PubMed
8. Ferreira FMRM, Chaves MEA, Oliveira VC, et al. Effectiveness of robot therapy on body function and structure in people with limited upper limb function: a systematic review and meta-analysis. PLoS One. 2018;13(7):e0200330. CrossRef PubMed
9. Zhang L, Jia G, Ma J, et al. Short and long-term effects of robot-assisted therapy on upper limb motor function and activity of daily living in patients post-stroke: a meta-analysis of randomized controlled trials. J Neuroeng Rehabil. 2022;19(1):76. CrossRef PubMed
10. Chien, WT, Chong YY. Robot-assisted therapy for upper-limb rehabilitation in subacute stroke patients: a systematic review and meta-analysis. Brain Behav. 2020;10(8):e01742. CrossRef
11. Mehrholz J, Pollock A, Pohl M, et al. Systematic review with network meta-analysis of randomized controlled trials of robotic-assisted arm training for improving activities of daily living and upper limb function after stroke. J Neuroeng Rehabil. 2020;17(1):83. CrossRef PubMed
12. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. CrossRef PubMed
13. Maher CG, Sherrington C, Herbert RD, et al. Reliability of the PEDro scale for rating quality of randomized controlled trials. Phys Ther. 2003;83(8):713-721. CrossRef PubMed
14. Higgins JPT. T.J., Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors), Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023). Cochrane, 2023. Online. (Accessed July 2024)
15. Straudi S, Baluardo L, Arienti C, et al. Effectiveness of robot-assisted arm therapy in stroke rehabilitation: an overview of systematic reviews. NeuroRehabilitation. 2022;51(4):559-576. CrossRef PubMed
16. Brusola G, Garcia E, Albosta M, et al. Effectiveness of physical therapy interventions on post-stroke spasticity: an umbrella review. NeuroRehabilitation. 2023;52(3):349-363. CrossRef PubMed
17. Yang X, Shi X, Xue X, et al. Efficacy of robot-assisted training on rehabilitation of upper limb function in patients with stroke: a systematic review and meta-analysis. Arch Phys Med Rehabil. 2023;104(9):1498-1513. CrossRef PubMed
18. Verhagen A, Stubbs PW, Mehta P, et al. Comparison between 2000 and 2018 on the reporting of statistical significance and clinical relevance in physiotherapy clinical trials in six major physiotherapy journals: a meta-research design. BMJ Open. 2022;12(1):e054875. CrossRef PubMed
19. van Tulder M, Malmivaara A, Hayden J, et al. Statistical significance versus clinical importance: trials on exercise therapy for chronic low back pain as example. Spine. 2007;32(16):1785-1790. CrossRef PubMed
20. Gianola S, Castellini G, Corbetta D, et al. Rehabilitation interventions in randomized controlled trials for low back pain: proof of statistical significance often is not relevant. Health Qual Life Outcomes. 2019;17(1):127. CrossRef PubMed
21. Cummings P. Arguments for and against standardized mean differences (effect sizes). Arch Pediatr Adolesc Med. 2011;165(7):592-596. CrossRef PubMed
22. Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. Curr Opin Rheumatol. 2002;14(2):109-114. CrossRef PubMed
23. Dekker J, de Boer M, Ostelo R. Minimal important change and difference in health outcome: an overview of approaches, concepts, and methods. Osteoarthritis Cartilage. 2024;32(1):8-17. CrossRef PubMed
24. Paci M, Risaliti F, Pellicciari L, Reporting of "usual care" as the control group in randomized clinical trials of physiotherapy interventions for multiple sclerosis is poor: a systematic review. 2022;43(9):5207-5216. CrossRef PubMed
25. Arienti C, Buraschi R, Pollet J, et al. A systematic review opens the black box of "usual care" in stroke rehabilitation control groups and finds a black hole. Eur J Phys Rehabil Med. 2022;58(4):520-529. CrossRef PubMed
26. Lohse KR, Pathania A, Wegman R, et al. On the reporting of experimental and control therapies in stroke rehabilitation trials: a systematic review. Arch Phys Med Rehabil. 2018;99(7):1424-1432. CrossRef PubMed
27. Morone G, Palomba A, Martino Cinnera A, et al. Systematic review of guidelines to identify recommendations for upper limb robotic rehabilitation after stroke. Eur J Phys Rehabil Med. 2021;57(2):238-245. CrossRef PubMed
28. Lang CE, Lohse KR, Birkenmeier RL. Dose and timing in neurorehabilitation: prescribing motor therapy after stroke. Curr Opin Neurol. 2015;28(6):549-555. CrossRef PubMed
29. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014;348(mar07 3):g1687. CrossRef PubMed
30. Slade SC, Dionne CE, Underwood M, et al. Consensus on Exercise Reporting Template (CERT): explanation and elaboration statement. Br J Sports Med. 2016;50(23):1428-1437. CrossRef PubMed
31. Podurgiel J, Piscitelli D, Denegar C. Challenges in applying minimal clinically important difference: a critical review. Int J Rehabil Res. 2024;47(1):10-19. CrossRef PubMed

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Arch Physioter 2025; 15: 131-137
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3291
ORIGINAL RESEARCH ARTICLE

Application and adaptation of the Virtual Eggs Test for assessing hand dexterity in subjects with stroke: a pilot study

Michele Piazzini[1], Tommaso Ciapetti[2], Marco Baccini[2], Stefano Doronzio[1,2], Lucia Angelini[3,4], Paola Mucci[3,4], Robinson Guachi[3-5], Marco Controzzi[3,4], Francesca Cecchi[1,2]

[1]Dipartimento di Medicina Sperimentale e Clinica, UniFi, Firenze - Italia
[2]Unità di Ricerca Neuromotoria, IRCCS Fondazione don Carlo Gnocchi ONLUS, Firenze - Italia
[3]The Biorobotics Institute, Scuola Superiore Sant'Anna, Pisa - Italy
[4]Department of Excellence in Robotics & AI, Scuola Superiore Sant'Anna, Pisa - Italy
[5]Department of Mechatronics, International University of Ecuador, Quito - Ecuador

ABSTRACT

Objectives: To verify the feasibility of the Virtual Eggs Test (VET) and establish the ranges of fragilities of the Virtual Eggs (VEs) for assessing dexterity of stroke subjects and to collect feedback to improve its usability.
Methods: An observational non-profit study, with a pre-market medical device. It was conducted at a hospital neurorehabilitation unit. Nine subjects with chronic stroke (5 males; mean age 55.8 ± 18.9) performed the pilot with their paretic arm. Time to complete the test (TT), the number of failures for each VE, the threshold (T), and participants' self-reported comfort in performing the test were measured.
Results: The T varied among participants from 1.70 to 4.88 N/N. The average TT was 20.1 ± 6.5 minutes (range 11-33). Only one subject found the test uncomfortable.
Conclusions: The study found that the VET, with minor modifications, is feasible in stroke subjects. It might be useful for assessing sensorimotor impairment in both the affected and the less affected arm. Its metric properties and normative values in the healthy population will be determined in a study currently underway.

Keywords: Movement, Stroke, Technology, Upper extremity

What is already known about this topic:

- *Currently, there are many assessment instruments for hand dexterity in stroke subjects; these tests assess the motor capacity of the hand, but do not assess sensory and motor capacity simultaneously.*

What does the study add:

- *The Virtual Egg Test application could significantly enhance the evaluation of sensorimotor impairments, potentially leading to better-targeted rehabilitation strategies and improved outcomes for individuals affected by stroke.*

Introduction

Stroke is a major challenge, causing disability across all age groups and significantly impacting quality of life and the healthcare system (1). It affects over 80 million people worldwide, with about 13.5 million new strokes each year (2). The high prevalence of stroke represents an important economic and social as well as ethical concern (3). The degree of autonomy of basic (BADLs) and instrumental (IADLs) activities of daily life (ADLs) is a key outcome measure to verify the evolution of recovery over time and the effectiveness of different treatments (4). The extent of disability resulting from the acute event can vary greatly based on the severity and site of injury. In severe cases, sensory and motor functions may be entirely or nearly lost, resulting in a complete or nearly complete loss of motor skills and sensory perception. Motor impairments in persons with stroke include poor voluntary control of movements, poor coordination, and poor muscle strength (5). This is often accompanied by impairment of tactile and proprioceptive perception, further impacting the ability to manipulate objects and perform precise movements (6).

Sensory information is crucial for movement control. The cerebellum utilizes internal models to generate corrective motor commands based on discrepancies between actual and

Received: September 5, 2024
Accepted: May 12, 2025
Published online: June 5, 2025

This article includes supplementary material

Corresponding author:
Marco Controzzi, Francesca Cecchi
email: marco.controzzi@santannapisa.it; fcecchi@dongnocchi.it

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

predicted sensory inputs from a copy of the motor command sent to the muscles to trigger these corrective actions (7,8). Specifically, according to the Discrete Event-driven Sensory feedback Control (DESC) theory, sensory inputs related to the salient discrete mechanical events of manipulation (i.e., contact, lifting, leaning, and release) are used in the central nervous system to define the beginning and end of the phases of manipulation (9). Each of these phases corresponds to a subgoal of the overall action of grasping objects, for which grasping forces and their timing must be optimized (9-12).

Studies on grip force control carried out on stroke survivors (13,14) show that, following a stroke event, grip force control is impaired. In fact, the force exerted is excessive compared to healthy controls; moreover, the grip force exerted by the hand most affected by the consequences of the stroke (contralesional hand) shows a higher variance than the hand less affected (ipsilesional). This may also arise from sensory impairments because intact sensory functions are essential for the modulation of grip strength based on the manipulated object (15).

During many activities in daily life, force must be adjusted according to the weight and fragility of the object being manipulated, and this adjustment requires a complex interaction between sensory information and motor output. The integration of force and position sensory signals allows the central nervous system to deal with the uncertainty of these signals: when grasping fragile objects, force and finger position are correlated, but interactions with rigid objects result in smaller changes in position and larger changes in force than interactions with fragile objects (16). Furthermore, when moving an object, the gripping force must be modulated in parallel with the movement-induced fluctuations of the inertial load, and intact sensory feedback seems necessary for accurate force scaling. In fact, it has been demonstrated that subjects with chronic sensory loss (17) or acute, drug-induced anesthesia (18,19), while being able to anticipate changes in load forces, employ inefficiently high grip forces compared to controls, thus showing an alteration in grip force regulation according to inertial loads. This sensorimotor integration was also found to be impaired in stroke survivors, who frequently show excessive grip forces (6,20).

To evaluate upper arm and hand dexterity, several tests have been developed and validated over the years, such as the Box and Block Test (BBT) (21), the Nine Hole Peg Test (9HPT) (22), the Minnesota Dexterity Test (MDT) (23), the Functional Dexterity Test (24), and the Action Research Arm Test (25). These tests require prehension, manipulation, and/or moving from one point to another of objects, in some cases of varying shapes and sizes, and provide a score linked primarily to the success of the performance and the time required to complete the test. None of them examines the kinematics (trajectory and smoothness of the movement performed) and kinetics (force of prehension) of the action, which may be used to assess the sensorimotor integration. An exception is the use of complex motion analysis systems, inertial sensors (26,27), or robotic devices (28) are required, which are not generally available in clinical practice. However, such an assessment could provide valuable insights into the sensory-motor function of the hand and might also be valuable in detecting subtle sensory-motor impairments on the less affected side. In fact, though the cerebral lesion affects mostly the contralateral side, some ipsilateral deficits can also be demonstrated in both the acute and the chronic phase post-stroke (29-31).

Researchers of the Scuola Superiore Sant'Anna, Pisa, developed a test that assesses both the fine regulation of grip strength that is needed when manipulating fragile objects and gross dexterity. The test, named Virtual Eggs Test (VET), was originally designed for the assessment of grip strength control in myoelectric prosthesis wearers and consists of picking from a platform, transporting, and replacing objects of different fragility, named Virtual Eggs (VEs) (32,33). VEs are small, roughly cubic-shaped objects, equipped inside with two magnets, so that their facing walls collapse, simulating breaking, when the force of prehension exerted on them exceeds the magnetic force of attraction. Failure occurs with a relative displacement of the gripping walls of the device. The test uses 20 VEs, characterized by a progressively lower force of attraction and therefore progressively more "fragile." The task required in the VET is on the whole similar to that required in the Box and Block Test (even the object manipulated is quite similar in shape and size), with the important difference that the subject must not only be able to grasp and transport from one point to another a small object, but also to regulate the grasping force, which must progressively decrease by manipulating increasingly fragile VEs. In this way, it is possible to assess the ability to fine-tune the force without having to use complex and expensive instrumentation required to measure the grip force exerted.

The VET has been widely used to evaluate the effectiveness of sensory feedback in myoelectric hands (32,34-36), and data on healthy populations have also been collected (37). The aim of this pilot study was to collect insights for the adaptation of the VET in chronic stroke survivors, test its feasibility, and provide information about the VEs' range of fragility for this specific population.

Methods

Study design and setting

This study was an observational, non-profit study, with a pre-market medical device. It was conducted at a hospital neurorehabilitation unit. The study received approval from the local Ethics Committee on 16/03/2023 under the number 23327_spe.

The protocol was developed as part of the Fit for Medical Robotics project (Fit4MedRob, National Plan for Complementary Investments to the National Recovery and Resilience Plan—Legislative Decree no. 59, May 6, 2021, converted with amendments by Law no.101, July 1, 2021—Research Initiatives for Innovative Technologies and Pathways in Health and Care) which aspires to supplement current rehabilitation and care models for individuals of all ages with motor, sensory or cognitive impairments, from hospital to home care, by means of new robotic and digital technologies.

The protocol of the study comprises the following steps:

- Perform a modified procedure of the VET, dubbed the pilot, with the aim of identifying the values of the VEs to be used for assessing post-stroke subjects.

- Assess the comfort during the execution of the test. The participant was asked to answer the question "How comfortable was it for you to perform the task?" on a Likert scale between 0 (i.e., very uncomfortable) and 7 (very comfortable).

The complete description of the VET, the assessment protocol, and the adaptations implemented to assess hand dexterity in stroke survivors is provided in the Supplementary material.

Participants

Participants were selected among patients referred to the hospital neurorehabilitation unit, based on the following inclusion criteria:

- Age>18.
- First ischemic or hemorrhagic stroke occurred for 6 months; sufficient motor skills to perform the test (score ≥19 on the Motricity Index item "Pinch Grip" and total Motricity Index score at the paretic upper extremity ≥57 (38).
- Willingness to participate in the study and sign the consent form.

Participants were excluded in case of:

- Severe wrist and hand hypertension (modified Ashworth Scale >3) (39).
- severe visual and oculomotor deficits.
- Cognitive deficits that impede comprehension of the task, as evidenced by A Mini-Mental State Examination (MMSE) score <25.
- Other neurological or musculoskeletal pathologies that impact upper limb function and clinical instability.
- Psychiatric comorbidities.
- A lack of understanding of the Italian language.

Procedure

All participants performed the VET only once with their paretic arm, following the protocol described above. The test was conducted in an isolated room without distractions. The examiner recorded the total time required to complete the test, including the time required to instruct the participant and the trials required for the participant to become familiar with the task.

The test ended after the completion of fifteen trials. At the end of the test, participants were asked to answer this question: "How comfortable was it for you to perform the task?," choosing from a 7 points Likert-type scale from −3 to 3, where: 0 = neither comfortable nor uncomfortable; −3/3 = very uncomfortable/comfortable; −2/2 = uncomfortable/comfortable; -1/1 = somewhat uncomfortable/comfortable.

Moreover, the raters were asked to rate the difficulty in understanding the instructions by participants and their comfort in performing the test, using a Likert scale as above. They were also asked to indicate any critical issues they encountered.

Data analysis

The main aim of the pilot is to derive the set of values of the VE for the study with post-stroke survivors. For this reason, we computed the average of failures across the nine subjects for each VE transported. In addition, we derived the threshold at which stroke survivors fail to carry the object correctly. This was retrieved by applying the up/down method. The threshold (T) is calculated as the average of the values of the stimuli of the least four trials, which are typically reversals from a trial with a more fragile VE (failed) to a trial with a less fragile VE (successful), or vice versa. The threshold limit is converged after a certain number of trials (15 trials in the present study).

Results

Characteristics of the participants

Nine subjects (five males and four females, average age 55.8 ± 18.9, range 29-81; distance from stroke 6-52 months) were consecutively screened for eligibility and were enrolled after signing the informed consent. Their demographic and clinical characteristics are shown in Table 1.

TABLE 1 - Characteristics of participants

ID	Age	TSS (months)	Paretic side	Dominant side	Stroke type	MI Upper Limb	MI pinch	MMSE
ID1	59	52	L	R	ischemic	60	26	28
ID2	42	22	R	R	ischemic	66	26	29
ID3	32	6	R	R	ischemic	66	22	28
ID4	76	10	R	R	ischemic	85	26	26
ID5	54	24	L	R	ischemic	61	22	27
ID6	29	23	L	L	ischemic	92	33	27
ID7	81	17	R	R	ischemic	92	33	22
ID8	55	21	R	R	ischemic	85	26	30
ID9	74	51	R	R	ischemic	92	33	29

ID = Participant's Identification number; TSS = Time since stroke (months); MI = Motricity Index; MMSE = Mini-Mental State Examination.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Test results

The VEs within the range 10.98-28.05 were broken by 1 time out of 5 transports by 50% of the participants. This means that the analysis of the VE with thresholds ≥ 10.98 might not provide additional information about the participant in terms of gross dexterity.

The threshold varied among participants from 1.70 to 4.88 N/N (mean T = 3.22 ± 1.26 N/N) (Figure 1-Figure 2), and the time needed to complete the test ranged from 11 to 33 minutes (mean 20.1 ± 6.5). Data from each participant are reported in Table 2.

TABLE 2 - Threshold measured in each participant

	T (N/N)	Time
ID01	1.95	33
ID02	4.75	18
ID03	2.19	11
ID04	2.19	19
ID05	4.27	28
ID06	4.88	18
ID07	3.96	19
ID08	1.70	18
ID09	3.14	17

ID = identification number; T = threshold; VE = Virtual Egg; Time = Time of administration in minutes

Times of VET phases

The analysis of the duration of different VET phases showed very inconsistent results with several artifacts, which were most often due to uncertainties by the participant in grasping and releasing the VE during the test. Therefore, no results regarding VET phase times are reported.

Results of the comfort of use questionnaire by participants and the feasibility questionnaire by evaluators

The results of the comfort questionnaire by participants show that the task was rated favorably by most participants (Table 3).

According to the feedback of the evaluators, the participants were able to easily understand the execution instructions, except for one subject who reported some difficulty (Table 3). The evaluators found the test quite feasible, but identified one critical issue: the difficulty in detecting whether the most fragile VEs were broken during the test. The breakage was easily perceived when the trial was conducted with VEs with a higher breakage threshold, because in this case, an easily distinguishable noise was produced. With the more fragile VEs, on the other hand, the noise was almost imperceptible, and the examiner had to carefully look at the VE throughout the test.

Discussion

This study suggests the feasibility of the VET protocol in a population of stroke survivors. The assessment setting was rated as very comfortable or comfortable by all participants except one subject. In addition, the time required to complete the test was relatively short, although significantly longer than that required by classical dexterity tests such as the BBT and 9HPT. For both tests, no conclusive data on the administration time is available, but it has been estimated that both can be completed in about 5 minutes (21,40). We think that the additional time required to complete the VET would be quite acceptable, provided that the information about the ability of grip strength tuning is proved to be relevant in the assessment of upper limb functioning in stroke survivors. This deals with the validity of the VET, which will be addressed, along with reliability, in a future study. However, we decided to make minor modifications in the test protocol to reduce the administration time and the fatigue of participants, as explained below.

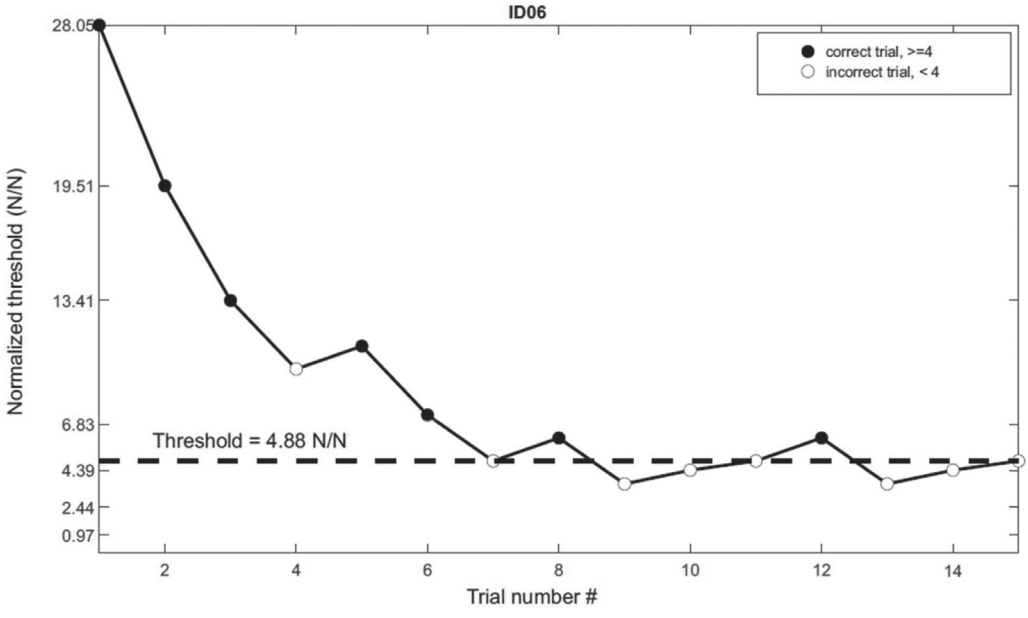


FIGURE 1 - Pilot result of the subject (ID06) with the highest (4.88 N/N) threshold.

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

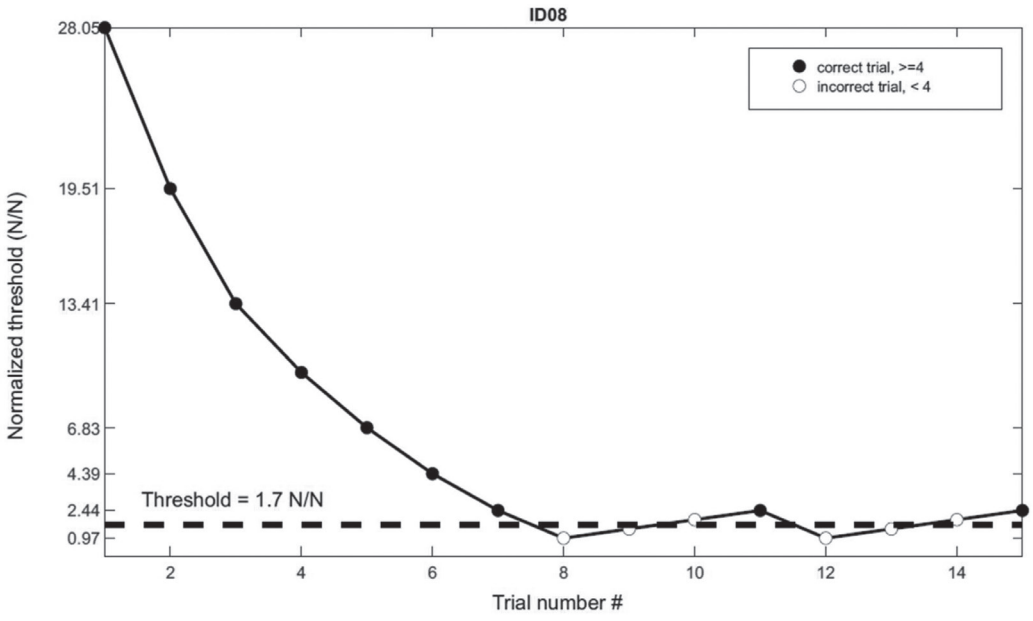


FIGURE 2 - Pilot result of the subject (ID08) with the lower (1.70 N/N) threshold.

TABLE 3 - Questionnaire results

	Occurrence						
Likert Scale / Question	−3	−2	−1	0	1	2	3
Comfort questionnaire by participants:							
How comfortable was it for you to perform the task?		1		1	3	1	3
Feasibility questionnaire by evaluators:							
How difficult was it for the participants to understand the instructions?		1					8
How comfortable were the participants in performing the test?					2	3	4

Currently, apart from assessments that require complex robotic or computerized devices, the subject's ability to control grip strength may be evaluated by means of the Strength Dexterity Test (SDT) (41). This test, which also takes longer than classical dexterity tests, measures the ability to fully compress by a lateral pinch a set of springs with plastic end caps without buckling; the mechanical properties of slender springs are used to quantify the subject's ability to dynamically control fingertip forces. The SDT, however, assesses the ability to finely control the direction of the force exerted between the two fingers as the required force increases, rather than the amount of force that must be exerted to manipulate objects of varying fragility. Furthermore, this ability is assessed by the SDT in a relatively static condition, without involvement of other joints of the upper limb. In contrast, most manipulation tasks require close coordination between the activation of proximal and distal muscles, so that the grip force must be adjusted during a simultaneous variable activation of other muscle groups. From this point of view, the VET most closely resembles the common manipulative tasks of everyday life.

Preliminary data on the application of the VET in a healthy population showed that the average threshold for these subjects is less than 2 (37), whereas the average threshold for stroke survivors presented in our pilot study is 3.7, with high variability among participants. Such variability in the performance of stroke survivors was expected, since the severity of motor impairment largely varies among these patients. This result suggests that the VET can discriminate between healthy subjects and stroke survivors, but this finding needs to be confirmed in a future study with adequate sample sizes. It also appears necessary to verify the reliability of the test and to collect normative data in the healthy population.

The VET might have a promising role in the clinical assessment of subtle motor deficits in the limb ipsilateral to brain injury, which should be referred to the less affected, rather than the healthy, side (42). A relatively recent study (43) compared the performance in a forward reaching task of individuals with stroke outcomes, age, and gender-matched healthy controls, collecting kinematic data by means of an optoelectronic motion analysis system. Subjects with stroke were assessed both in the acute and the subacute phase. The

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

authors found a consistently worse performance of stroke participants who showed reduced smoothness of movements, range of motion, and power. These deficits were more pronounced in the acute phase and in subjects with greater motor deficits on the affected side. Some data have also been published on the ability of classical dexterity tests, such as the MDT and the Purdue Pegboard Test (30) and the 9HPT (44), to detect motor deficit in the ipsilateral side of stroke subjects, but the differences compared to healthy controls were minimal, although significant. For example, Son et al. (44) found that stroke participants took on average 2 seconds more than healthy individuals to complete the 9HPT. Moreover, data were collected in very limited samples (27-30 participants with stroke outcomes). The VET promises to be a simple, more sensitive clinical tool that can detect motor deficits in the less affected side after stroke since it includes both quantitative (time to complete the task) and qualitative (ability to finely control grip strength) measures.

Based on the results of the pilot, the set VEs should have a range of thresholds between 1 and 10.5 (N/N) to increase the accuracy of the outcome. The execution of the pilot was based on an up/down procedure, and this can be exploited as the revision of the original protocol of the VET (37) to minimize the time required to execute the test.

The answers to the questionnaires did not reveal any issues in performing the VET by the enrolled subjects. However, the addition of visual feedback for the VEs with lower break thresholds might increase the detection of the failure. This was also reported by the raters in a study involving amputees (37).

This study has some limitations. First, although calculating the threshold as the mean of a number of reversals is a widely used method in psychophysics (45), in the present study, the limited number of reversals may have led to an inaccurate threshold estimate for some participants. In future studies with the VET, we aim to replace the current criterion for the termination of the test (15 trials completed) with the criterion of having reached five reversals, and to calculate the threshold as the mean of all such reversals. Moreover, the software couldn't accurately measure VET phase durations for all participants, especially when tremor caused uncertainty in grasping or releasing the VE. To solve this problem, the software will be modified to calculate lifting times (initial grasp to complete VE lifting) and releasing times (initial VE contact with the platform to full release), and thus have more accurate and detailed information on the timing of the different phases. Conceivably, these additional metrics would improve phase discrimination and provide valuable insights into the hand dexterity of stroke survivors. An intrinsic limitation of the VET, in fact, is the impossibility of measuring the movement kinematics, for which different and more expensive instrumentation would be required. However, the time needed to grasp and transfer the VE, if accurately measured, could be an important parameter for quantifying both gross and fine hand dexterity. When handling the first, 'stronger VEs, the difficulty in fine-tuning the grip forces recedes into the background, and the subject can concentrate on the speed of the movement. Conversely, as the breaking threshold decreases, acceleration must be minimized, since the greater the acceleration, the greater the force that must be exerted to compensate for inertial load, thus increasing the risk of "breaking" the VE.

In conclusion, we believe the VET is a promising tool for evaluating hand dexterity in stroke survivors, even if some modifications can be implemented to increase its potential. A new study will be carried out with the aim of verifying the reliability and validity of VET in stroke survivors on both the paretic and the less affected limb and identifying normative data for the healthy population.

Disclosures

Conflict of interest: The authors declare no conflict of interest

Financial support: Italian Ministry of Research, under the complementary actions to the NRRP "Fit4MedRob - Fit for Medical Robotics" Grant (# PNC0000007).

Grant numbers:
- Italian National Institute for Insurance against Accidents at Work (INAIL) within the BioARM project, grant n° PR19-RR-P3.
- The European Commission under the DeTOP project (LEIT-ICT-24-2015, Grant no. 687905), B-CRATOS project (FETOPEN-01-2018-2019-2020, Grant no. 965044)
- Italian Ministry of Research, under the project CALLIOPE (PRIN2022, Grant no. 2022PRHZM3)

References

1. Katan M, Luft A. Global burden of stroke. Semin Neurol. 2018;38(2):208-211. CrossRef PubMed
2. Lindsay MP, Norrving B, Sacco RL, et al. World Stroke Organization (WSO): Global Stroke Fact Sheet 2019. Int J Stroke. 2019;14(8):806-817. CrossRef PubMed
3. Feigin VL, Norrving B, Mensah GA. Global burden of stroke. Circ Res. 2017;120(3):439-448. CrossRef PubMed
4. Cecchi F, Cassio A, Lavezzi S, et al. Redefining a minimal assessment protocol for stroke rehabilitation: the new "Protocollo di Minima per l'ICtus" (PMIC2020). Eur J Phys Rehabil Med. 2021;57(5):669-676. CrossRef PubMed
5. Byl N, Roderick J, Mohamed O, et al. Effectiveness of sensory and motor rehabilitation of the upper limb following the principles of neuroplasticity: patients stable poststroke. Neurorehabil Neural Repair. 2003;17(3):176-191. CrossRef PubMed
6. Blennerhassett JM, Carey LM, Matyas TA. Grip force regulation during pinch grip lifts under somatosensory guidance: comparison between people with stroke and healthy controls. Arch Phys Med Rehabil. 2006;87(3):418-429. CrossRef PubMed
7. Franklin DW, Wolpert DM. Computational mechanisms of sensorimotor control. Neuron. 2011;72(3):425-442. CrossRef PubMed
8. Wolpert DM, Diedrichsen J, Flanagan JR. Principles of sensorimotor learning. Nat Rev Neurosci. 2011;12(12):739-751. CrossRef PubMed
9. Johansson R, Flanagan J. Tactile sensory control of object manipulation in humans. Senses Compr Ref. 2008;6:67-86. CrossRef
10. Johansson R, Edin B. Predictive feed-forward sensory control during grasping and manipulation in man. Biomed Res. 1993;14. Online (Accessed September 2024)
11. Johansson RS, Cole KJ. Sensory-motor coordination during grasping and manipulative actions. Curr Opin Neurobiol. 1992;2(6):815-823. CrossRef PubMed

12. Johansson RS, Flanagan JR. Coding and use of tactile signals from the fingertips in object manipulation tasks. Nat Rev Neurosci. 2009;10(5):345-359. CrossRef PubMed
13. Hermsdörfer J, Hagl E, Nowak DA, Marquardt C. Grip force control during object manipulation in cerebral stroke. Clin Neurophysiol. 2003;114(5):915-929. CrossRef PubMed
14. Quaney BM, Perera S, Maletsky R, et al. Impaired grip force modulation in the ipsilesional hand after unilateral middle cerebral artery stroke. Neurorehabil Neural Repair. 2005;19(4):338-349. CrossRef PubMed
15. Dresp-Langley B. Grip force as a functional window to somatosensory cognition. Front Psychol. 2022;13:1026439. CrossRef PubMed
16. Geelen JE, van der Helm FCT, Schouten AC, et al. Sensory weighting of position and force feedback during pinching. Exp Brain Res. 2023;241(8):2009-2018. CrossRef PubMed
17. Nowak DA, Glasauer S, Hermsdörfer J. How predictive is grip force control in the complete absence of somatosensory feedback? Brain. 2004;127(Pt 1):182-192. CrossRef PubMed
18. Nowak DA, Hermsdörfer J, Glasauer S, et al. The effects of digital anaesthesia on predictive grip force adjustments during vertical movements of a grasped object. Eur J Neurosci. 2001;14(4):756-762. CrossRef PubMed
19. Monzée J, Lamarre Y, Smith AM. The effects of digital anesthesia on force control using a precision grip. J Neurophysiol. 2003;89(2):672-683. CrossRef PubMed
20. Kanzler CM, Averta G, Schwarz A, et al. A low-dimensional representation of arm movements and hand grip forces in post-stroke individuals. Sci Rep. 2022;12(1):7601. CrossRef PubMed
21. Mathiowetz V, Volland G, Kashman N, et al. Adult norms for the Box and Block Test of manual dexterity. Am J Occup Ther. 1985;39(6):386-391. CrossRef PubMed
22. Mathiowetz V, Weber K, Kashman N, et al. Adult norms for the Nine Hole Peg Test of finger dexterity. Occup Ther J Res. 1985;5(1):24-38. CrossRef
23. Tesio L, Simone A, Zebellin G, et al. Bimanual dexterity assessment: validation of a revised form of the turning subtest from the Minnesota Dexterity Test. Int J Rehabil Res. 2016;39(1):57-62. CrossRef PubMed
24. Aaron DH, Jansen CWS. Development of the Functional Dexterity Test (FDT): construction, validity, reliability, and normative data. J Hand Ther. 2003;16(1):12-21. CrossRef PubMed
25. Yozbatiran N, Der-Yeghiaian L, Cramer SC. A standardized approach to performing the action research arm test. Neurorehabil Neural Repair. 2008;22(1):78-90. CrossRef PubMed
26. Wagner JM, Rhodes JA, Patten C. Reproducibility and minimal detectable change of three-dimensional kinematic analysis of reaching tasks in people with hemiparesis after stroke. Phys Ther. 2008;88(5):652-663. CrossRef PubMed
27. Hussain A, Balasubramanian S, Roach N, et al. SITAR: a system for independent task-oriented assessment and rehabilitation. J Rehabil Assist Technol Eng. 2017;4:2055668317729637. CrossRef PubMed
28. Morone G, Palomba A, Martino Cinnera A, et al. "CICERONE" Italian Consensus Conference on Robotic in Neurorehabilitation. Systematic review of guidelines to identify recommendations for upper limb robotic rehabilitation after stroke. Eur J Phys Rehabil Med. 2021;57(2):238-245. CrossRef PubMed
29. Noskin O, Krakauer JW, Lazar RM, et al. Ipsilateral motor dysfunction from unilateral stroke: implications for the functional neuroanatomy of hemiparesis. J Neurol Neurosurg Psychiatry. 2008;79(4):401-406. CrossRef PubMed
30. Pandian S, Arya KN. Motor impairment of the ipsilesional body side in poststroke subjects. J Bodyw Mov Ther. 2013;17(4):495-503. CrossRef PubMed
31. Metrot J, Froger J, Hauret I, et al. Motor recovery of the ipsilesional upper limb in subacute stroke. Arch Phys Med Rehabil. 2013;94(11):2283-2290. CrossRef PubMed
32. Clemente F, D'Alonzo M, Controzzi M, et al. Non-invasive, temporally discrete feedback of object contact and release improves grasp control of closed-loop myoelectric transradial prostheses. IEEE Trans Neural Syst Rehabil Eng. 2016;24(12):1314-1322. CrossRef PubMed
33. Controzzi M, Clemente F, Pierotti N, et al. Evaluation of hand function transporting fragile objects: the Virtual Eggs Test, proceedings of the Myoelectric Controls Symposium 2017. Online (Accessed September 2024)
34. Mastinu E, Clemente F, Sassu P, et al. Grip control and motor coordination with implanted and surface electrodes while grasping with an osseointegrated prosthetic hand. J Neuroeng Rehabil. 2019;16(1):49. CrossRef PubMed
35. Valle G, D'Anna E, Strauss I, et al. Hand control with invasive feedback is not impaired by increased cognitive load. Front Bioeng Biotechnol. 2020;8:287. Accessed February 6, 2024. Online CrossRef PubMed
36. Chai G, Wang H, Li G, et al. Electrotactile feedback improves grip force control and enables object stiffness recognition while using a myoelectric hand. IEEE Trans Neural Syst Rehabil Eng. 2022;30:1310-1320. CrossRef PubMed
37. Controzzi M, Angelini L, Randi P, et al. Assessing hand function in trans-radial amputees wearing myoelectric hands: the Virtual Eggs Test (VET). 2024. CrossRef
38. Sunderland A, Tinson D, Bradley L, et al. Arm function after stroke. An evaluation of grip strength as a measure of recovery and a prognostic indicator. J Neurol Neurosurg Psychiatry. 1989;52(11):1267-1272. CrossRef PubMed
39. Harb A, Kishner S. Modified Ashworth Scale. In: StatPearls. StatPearls Publishing; 2024, Online (Accessed September 2024).
40. Strokengine. Nine Hole Peg Test (NHPT). Online (Accessed September 2024)
41. Valero-Cuevas FJ, Smaby N, Venkadesan M, et al. The strength-dexterity test as a measure of dynamic pinch performance. J Biomech. 2003;36(2):265-270. CrossRef PubMed
42. Kitsos GH, Hubbard IJ, Kitsos AR, et al. The ipsilesional upper limb can be affected following stroke. ScientificWorldJournal. 2013;2013(1):684860. CrossRef PubMed
43. Scano A, Guanziroli E, Mira RM, et al. Biomechanical assessment of the ipsilesional upper limb in post-stroke patients during multi-joint reaching tasks: a quantitative study. Front Rehabil Sci. 2022;3:943397. CrossRef PubMed
44. Son S, Nam S, Kang K, et al. Relationship between ipsilateral motor deficits on the less-affected side and motor function stage on the affected side. J Korean Phys Ther. 2018;30(6):234-238. CrossRef
45. Kingdom FAA, Prins N. Psychophysics: A Practical Introduction. Academic Press; 2016. Online (Accessed September 2024)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 138-147
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3448
METHODOLOGY REPORT

A new test protocol and device for measuring the cranio-cervical flexion test in participants with bruxism

Mareile Jung, Annika Zind, Markus J Ernst

Zürich University of Applied Sciences, Institute of Physiotherapy, Winterthur - Switzerland

ABSTRACT

Introduction: The craniocervical flexion test (CCFT) has been proposed to assess muscular stabilization using deep neck flexors. Reliability of the CCFT using a pressure biofeedback unit has been regarded as doubtful, and the level of evidence is reportedly low.

The aim of the current study is to pilot test an alternative measurement protocol by using a new digital device during the staged CCFT when compared to surface electromyography (sEMG) of neck flexor muscles in participants with bruxism.

Methods: Cross-sectional pilot study including five participants with bruxism and five controls. Measuring five incremental stages (20-100%), from a maximum force of 17 N for the CCFT and parallel to sEMG measurements of bilateral sternocleidomastoideus (SCM) and masseter muscles. SEMG data were normalized to their activity during maximal voluntary contraction. Performance during the staged CCFT was fed back via a smartphone screen.

Results: A two-way repeated measures ANOVA found a significant stage ($F = 32.56$, $df = 4$, $p = 0.001$) but not group (0.30, $df = 1$, $p = 0.65$) main effect for SCM activity, with both groups demonstrating higher normalized sEMG activity for incremental stages of the CCFT. Neither group nor stage effects were found for normalized masseter activity during the staged CCFT.

Conclusion: A new measurement protocol and test device for the CCFT were examined in participants with and without bruxism. In parallel, sEMG used showed differences in SCM activity for incremental test stages. Group differences could not be found.

Keywords: Bruxism, Craniocervical flexion test, Measurement protocol, NOD device, Pilot study, Superficial EMG

Introduction

Muscular stabilization of the neck via the deep neck flexors (DNF) has been found important in everyday life (1). Impaired function is associated with problems such as neck pain (2), headaches (3-5), or craniomandibular dysfunction (CMD) (6). The craniocervical flexion test (CCFT) has been proposed for assessing the muscular stabilization by the DNF, which is provided predominantly by the longus capitis and longus colli muscles (7). The CCFT assesses three components: first, the ability to increase contractions of the DNF progressively and correctly to five inner range positions of craniocervical flexion. Second, to determine that stage or threshold someone can contract correctly without substitution by superficial muscles (sternocleidomastoideus, i.e. SCM and scalene muscles), for 10 seconds (i.e. activation score), and third to evaluate the holding capacity, which is the repeatedly correct activation of the DNF, usually up to ten times ten second holds (i.e. performance index) (8,9). The CCFT has commonly been investigated by using a pressure biofeedback unit (i.e. PBU) in clinical settings (2,10). In laboratory settings, additional measurements of superficial muscles by surface electromyography (sEMG) have also been performed (11,12). Intra- and interrater-reliability of the clinical CCFT using the PBU has, however, a conflicting level of evidence across some recent systematic reviews, as often only healthy participants had been examined, and measurement errors account for approximately 15 % of the whole range of the scale from 20 to 30 mmHg (2,10,13). Uthaikhup and Jull reported shortfalls in achieving incremental 2 mmHg values on the scale and during the CCFT by using the PBU in asymptomatic elderly of 5-30% (14). Studies reporting shortfalls or the opposite, overshooting during the CCFT in symptomatic populations and by using the PBU, are lacking. Further doubts regarding the PBU emerge from its scale with 2 mmHg interval units, making reported measurement errors

Received: December 24, 2024
Accepted: May 12, 2025
Published online: June 6, 2025

This article includes supplementary material.

Clinical trial number:
NCT 06037798.

Corresponding author:
Markus J Ernst
email: markus.ernst@zhaw.ch

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

below that one unit a theoretical value only (2). As reliability is necessary for validity, the latter is also restricted.

The NOD is a hand-held dynamometer (HDD) with three incorporated functions. The biofeedback function has been introduced to test the CCFT, but only one study by Lonati et al. has recently studied its concurrent validity in comparison to the PBU for examining the CCFT, and reported a strong correlation of r = 0.92 (15).

Lonati et al. have used a maximum value of 17 Newton (N) for the CCFT by using the NOD, which corresponds to 30 mmHg by using the PBU (15) (see also supplemental Material). In clinical settings, superficial muscle activity can be controlled for by palpation of the corresponding muscles (9). However, more accurate measurements are needed in laboratory setups, and especially when using a new test protocol and device. Alalawi et al. have measured maximal and submaximal cranio-cervical flexion strength by using the NOD parallel to sEMG of the SCM in healthy and two groups of neck pain patients (16). The maximum achievable value of 17 N, used by Lonati et al., corresponded to approximately 32% of the maximum strength for the healthy and 36-38% for the two symptomatic groups (16). Directly measuring DNF muscle activity with EMG techniques is difficult to perform, accordingly surface EMG measurements of the SCM were most frequently performed in parallel (12,17-19), as an inverse relationship between the activity of deep and superficial flexor muscles, for different populations like neck pain, whiplash and headache subjects has been demonstrated (5,12,18,20-22). Further significant differences for normalized root mean square values of superficial muscle activity between healthy controls and various patient groups, like neck pain, whiplash, cervicogenic headache, and migraine, have been reported (5,12,20,23-25).

Bruxism has been found to be associated with Temporomandibular disorders (TMD) (26). Co-contraction of neck muscles during teeth grinding has been shown in various studies (27-30), but not in association with the CCFT. Parafunctional bruxism with TMD has been found significantly associated with neck disability (26). However, authors reported that cervical impairments were unlikely to be associated with bruxism, but with a pain mechanism leading to increased sensitivity in the cranio-cervical region. Unfortunately, that study did not report results on the CCFT (26).

In order to examine the common notion of an impaired function of the DNF in participants with bruxism, a new measurement protocol using the NOD will be piloted in participants with and without bruxism.

The first aim of the current study is to examine the CCFT by introducing an alternative measurement protocol and device, the "NOD" (http://www.to-nod.com/) for measuring the CCFT in research settings and when compared to EMG of the superficial neck flexors, i.e., the SCM.

A further aim of this study was to examine whether participants with bruxism show a different performance of the CCFT together with higher activity of the superficial neck flexor or masseter muscles, as has already been found in other symptomatic groups. It was hypothesized that participants with bruxism showed an increased activation of the SCM and the masseter during the CCFT.

Methods

Study design

Pilot study using a cross-sectional design. The study was approved by the local ethics committee (ethics committee of the canton Zürich, Switzerland), identification number: 2023-01641, and registered at clinicaltrials.gov: NCT 06037798.

Recruitment and Inclusion

Participants were recruited via physiotherapy practices or at the university with flyers or direct references to the study. Those interested in participating were screened by an online questionnaire using REDCap (31) for eligibility. Persons with neurological conditions, primary headaches, mental or inflammatory diseases, known pregnancies, and current or ≥ three months TMD complaints during the last 12 months were excluded. Furthermore, subjects were excluded when reporting neck pain (32) with an intensity of >3 out of 10 on the numeric rating scale (NRS) at the start of the measurements.

Subjects with bruxism were further screened for complaints such as pain, aching, or stiffness within muscles of mastication, sensitized teeth, either self-reports or by partners of grinding or chattering with the teeth. On the day of measurement, further objective signs of bruxism like hypertrophy of the masseter muscle, cheek and/or tongue impressions, abraded tooth tips or exposed tooth necks were examined.

After inclusion, activation of DNF and sEMG of superficial neck flexors were examined in all participants, with and without bruxism, in a single session at the movement laboratory of the participating university.

Measurement protocol

After cleaning the skin of both cheeks and the region above both collar bones, with alcohol to decrease impedance and electrode skin interface noise, one pair of bipolar, oval, self-adhesive and wet sEMG electrodes (44.8 × 22 mm) with a centre-to centre distance of 2cms, each for the masseter and the sternal part of the SCM were attached on both sides (see Fig. 1) (33). EMG signals were recorded using the Cometa system with PICO EMG sensors. The amplification was set at 1000×. A band-pass filter integrated into the hardware was applied with cut-off frequencies set at 10 Hz (high-pass) and 500 Hz (low-pass) (34). The EMG sampling rate was 2000 Hz. EMG signals were transmitted to a receiver via Bluetooth.

For all tests, participants lay on their backs with their legs bent (crooked lying) (see Fig. 1).

In total, five tests were conducted:

1. Measuring the *resting tone of both the masseter and SCM muscles* for 30 seconds, once. Participants were asked to breathe normally but avoid swallowing.
2. Measuring the activation of the deep neck flexor muscles with the *staged CCFT*. The NOD device with a silicon pad attached was placed under the neck of the participant, enabling a nodding movement without lifting the head.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

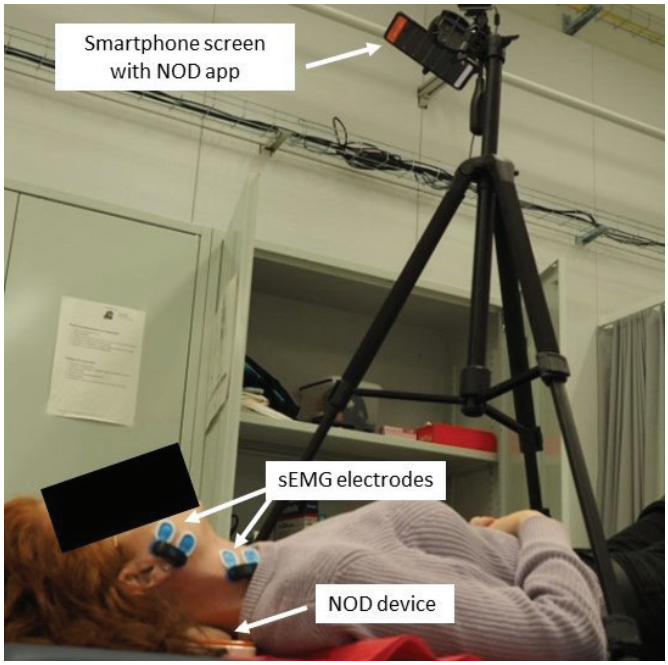


FIGURE 1 - Measurement setup.

A plastic foil was placed under the head to facilitate friction-free upper cervical flexion by sliding the head on the treatment table. If needed, persons with kyphosis were additionally supported with towels below the head or neck.

Measurements were recorded by using the "biofeedback mode" of the NOD. Minimum and maximum values were set manually at 0 and 17 N, respectively (15). The device was calibrated before every single test. An Android operating smartphone with the NOD app installed was fixed to a tripod, which was placed above the participant (Fig. 1). Participants were instructed and familiarized with the CCFT to nod without lifting the head, keeping their jaw muscles relaxed, and not to push their neck backwards. They were allowed to practice once or twice before measurements were recorded. Participants were given five seconds to adjust their contraction to each corresponding and incremental stage and to hold it for a further 10 seconds each, at 20, 40, 60, 80, and 100% of the maximum of a 17 N value. Accurate performance was fed back via the smartphone screen.

If the participants could hold steadily the requested %-line of force, and only occasionally had a "wobble" that did not reach or exceed the adjacent % lines, or repeatedly (three times within five seconds) crossed the requested centre line toward the adjacent %-lines, that stage was considered passed (see Figs 2 and 3).

For each stage, the activity in the SCM and masseter muscles was measured by sEMG.

3. *A "head lift-off" test* was used to measure the maximal voluntary contractions (MVC) of the neck flexor muscles.

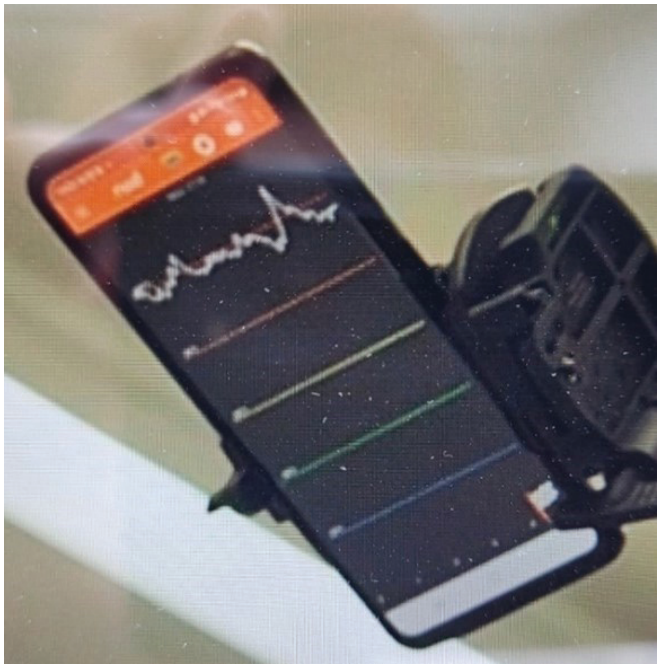


FIGURE 2 - Example of a participant who passed the final stage at 100%.

FIGURE 3 - Example of a participant who failed the 2nd (40%) stage.

Participants were fixed to the treatment table via a belt over their upper chest. The NOD was set to dynamometer mode. Participants had to lift their head for approximately 3 cm while maintaining upper cervical flexion (nodding). While doing so, they pressed their forehead against the NOD with maximum effort for five seconds in three runs with 30 seconds' rest in between (35-37). Surface EMG of the SCM muscles was recorded in parallel. The peak force was recorded by the App, and the mean of three trials was used for subsequent analysis. One familiarization and warm-up trial, asking participants to push with approximately half of their strength, was conducted prior to the recordings.

4. *Clenching the teeth*: The maximum force for bilateral masseter muscles by clenching the teeth on a wooden spatula and for five seconds in two runs was measured and used for further analysis (38-40). In addition, the activity of both SCMs was recorded. One familiarization and warm-up trial, asking participants to clench with approximately half of their strength, was conducted prior to the recordings.
5. *Nodding and clenching*: Participants were asked to perform a nod to the second stage (40%) of the CCFT while subsequently and maximally clenching their teeth on a wooden spatula. The capacity of holding the nod while clenching the teeth was evaluated. The same evaluation as for test 2 was applied here. If the 40%-line has been crossed three times within 5 seconds, the test is considered failed. One familiarization or warm-up trial was conducted prior to the recordings.

Data processing and analysis

Post-processing of all sEMG data was done with proEMG (version 2.1.3.12). The raw EMG signals were rectified and subsequently low-pass (450 Hz) and high-pass (20 Hz) filtered (Butterworth filter) to enable amplitude analysis. Maximum root mean square (RMS) values for the entire period of each test were calculated using a continuous smoothing RMS window of 1000 ms with 0.5 s overlap.

Data was examined for normal distribution by using the Shapiro-Wilk test. If the data were not normally distributed, non-parametric tests were applied subsequently.

All sEMG data of the SCM and masseter muscles during the *staged CCFT* and the *nodding and clenching test* were normalized to their activity during maximal voluntary contraction (MVC) and presented as a percentage.

Normalized values were examined for side differences first and averaged if no side differences were found (Wilcoxon signed-rank test).

Differences in sEMG activity for the SCM and masseter muscles between groups and for all tests but the *staged CCFT* were examined by using the Wilcoxon signed-rank test.

Frequencies of failed test performance during the *staged CCFT* and between groups were examined using Fisher's exact test.

For the *staged CCFT* group, differences in sEMG activity for the SCM and masseter muscles were examined using a two-way repeated measures ANOVA. Bootstrapping using 1000 iterations of subsamples was performed to provide confidence intervals and test statistics. For all statistical group comparisons, alpha was set at 0.05. For all statistical analyses, R Statistical Software (v4.4.1, R Core Team, 2024) (41) and the packages afex, epiR, epitools, ggplot2, gridExtra, plyr, and reshape were used (42-48).

Boxplots and interaction plots are provided for graphical illustrations.

Results

During enrollment, fifteen persons were assessed for eligibility using an online questionnaire on REDCap. Five people had to be excluded. Four did not meet the inclusion criteria, and one person did not complete the eligibility questionnaire. Finally, ten participants, five for each group, were included.

Both groups consisted of two men and three women, with a mean age ± SD of 35.8 ±11.0 years for the bruxism group and 39.4 ±16.5 for the control group, respectively.

In the bruxism group, two people each stated that they grind or clench during the day or at night, and one person during both time points.

Resting tone

There were no significant group differences in resting muscle activity for the SCM ($p = 0.69$) or the masseter muscle ($p = 0.55$) (see supplementary material)

Staged CCFT

During the staged CCFT, both groups showed an increased normalized sEMG activity of the SCM and masseter muscles. Whilst normalized sEMG values for the masseter muscle showed non-normal distribution across all stages of the CCFT, the opposite was true for normalized SCM values.

The two-way repeated measures ANOVA found a significant stage ($F = 32.56$, $df = 4$, $p = 0.001$) but no group (0.30, $df = 1$, $p = 0.65$) main effect (for mean values of SCM activity) with both groups demonstrating higher normalized sEMG activity with each stage of the CCFT (see Fig. 4).

The two-way repeated measures ANOVA for normalized masseter muscle activity during the staged CCFT found no stage (F 5.71, $df = 4$, $p = 0.28$) and no group effect (F: 0.54, $df = 1$, $p = 0.55$) (see Fig. 5).

From a total of 25 test stages performed among five participants per group, nine levels could not be controlled in the bruxism group, while three levels could not be controlled in the control group. The odds ratio of failing a test stage among participants with bruxism compared to controls was 4.01 (95% confidence interval: 0.883-26.71, $p = 0.10$)

Four participants of the bruxism group failed to control the highest stage (100%), whereas two participants of the control group could not pass this level.

Head lift-off

The control group showed larger strength values during the lift-off test, with a median value of 82.1 Newton and a range from 35.5 to 112.5 N, compared to the bruxism group with a median of 44.1 Newton and a range from 28.2 to 110.2 Newton. Differences were statistically not significant ($p = 0.42$). No statistically significant differences were found between groups for the sEMG activity for both SCM ($p = 0.42$) and masseter muscles ($p = 0.22$) during the lift-off test.

Clenching teeth

There were no significant group differences in the activity of SCM ($p = 0.42$) or masseter muscle ($p = 1$) during maximal teeth clenching.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

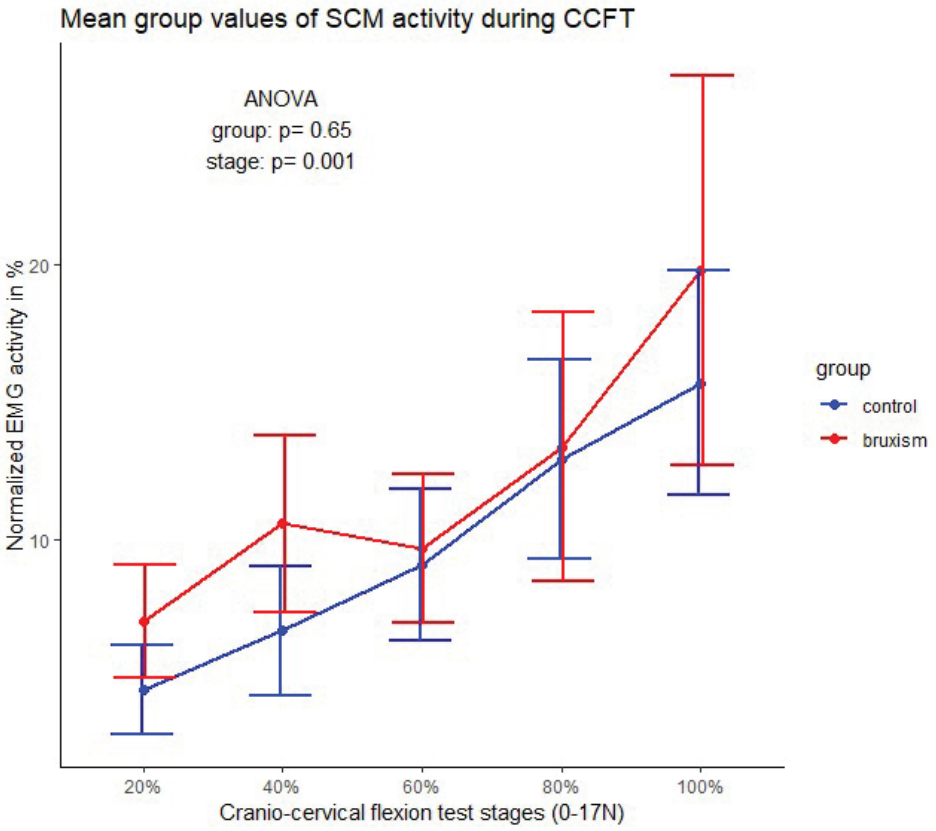


FIGURE 4 - Line plot with group means and standard errors of normalized sEMG activity of the SCM during the CCFT.

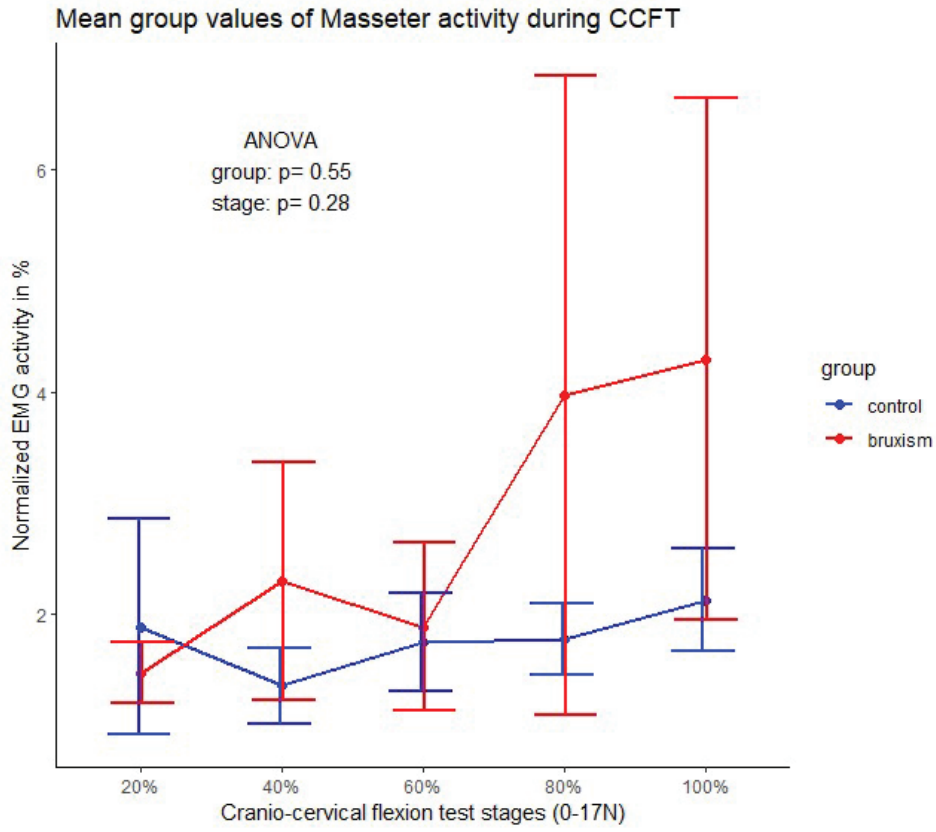


FIGURE 5 - Line plot with group means and standard errors of normalized sEMG activity of the masseter muscles during the CCFT.

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Nodding and clenching

After cranio-cervical flexion to test stage 2 (40%), and subsequently maximal clenching of the teeth, the control group showed on average a lower activity of the normalized masseter muscle, which remained not significant with a median (interquartile range): bruxism: 48.86 (27.88) vs. control: 40.19 (9.34, p = 0.84) (Fig. 6).

The odds ratio of failing to perform the CCFT correctly while maximally clenching the teeth among participants with bruxism compared to controls was 3.61 (95% confidence interval: 0.230-224.19, p = 0.58).

Discussion

A new test protocol for the CCFT, by using the NOD, has been examined in the current pilot study. The NOD offers detailed feedback via an App on a smartphone or tablet screen for patients and therapists, which can help identify someone's impairment in correctly activating the deep neck flexor muscles during the CCFT. Five %-test-stages were manually set between 0 and 17 N and needed to be achieved by a cranio-cervical flexion action with the NOD device placed under the neck of the participant. Further criteria were defined, such as how correctly holding ten seconds of cranio-cervical flexion at these stages should be rated on the display, as described in the methods section.

By additionally using sEMG, we were able to provide preliminary evidence that both participants with bruxism and healthy controls may have an increased activity of superficial neck flexors during incremental stages of the CCFT, which has not been found for the masseter muscles.

The CCFT is designed to test the correct activation of the DNF during five incremental stages. Ideally, the activity of superficial muscles should not be pronounced, and associated movements, such as retraction or oral parafunctions, should not occur (9). Several studies have already shown increased activity of the superficial muscles while performing the CCFT for different patient groups, such as primary headaches (5,49,50), cervicogenic headache (25,51), whiplash-associated disorders (20) or neck pain (1,11,12).

Even if the current study has only found significant differences between incremental test stages of normalized activity for the SCMs (p = 0.001) during the staged CCFT, a tendency towards increased activity could also be recognized for the masseter muscles, but remained, for both groups, below that of the SCM muscles.

The amount of EMG activity in the SCM muscles and for the bruxism group is comparable to values found in other pathologies or disorders like migraine (5,23,25) or mild neck pain (24), though it tends to be lower than in whiplash-associated disorders (20), tension-type headache (25), and moderate (24) or chronic neck pain (1,12).

So far, most clinical and laboratory studies have used the PBU for measuring or treating patients with DNF exercises as a feedback tool (2,4,17,52,53), despite evidence that its reliability has been reported to be controversial, with large measurement errors (2,10). Reasons for the limitations are partially methodological ones, as many studies have examined only healthy participants (2,53). Another issue that affects reliability might be the PBU itself and its scale, with

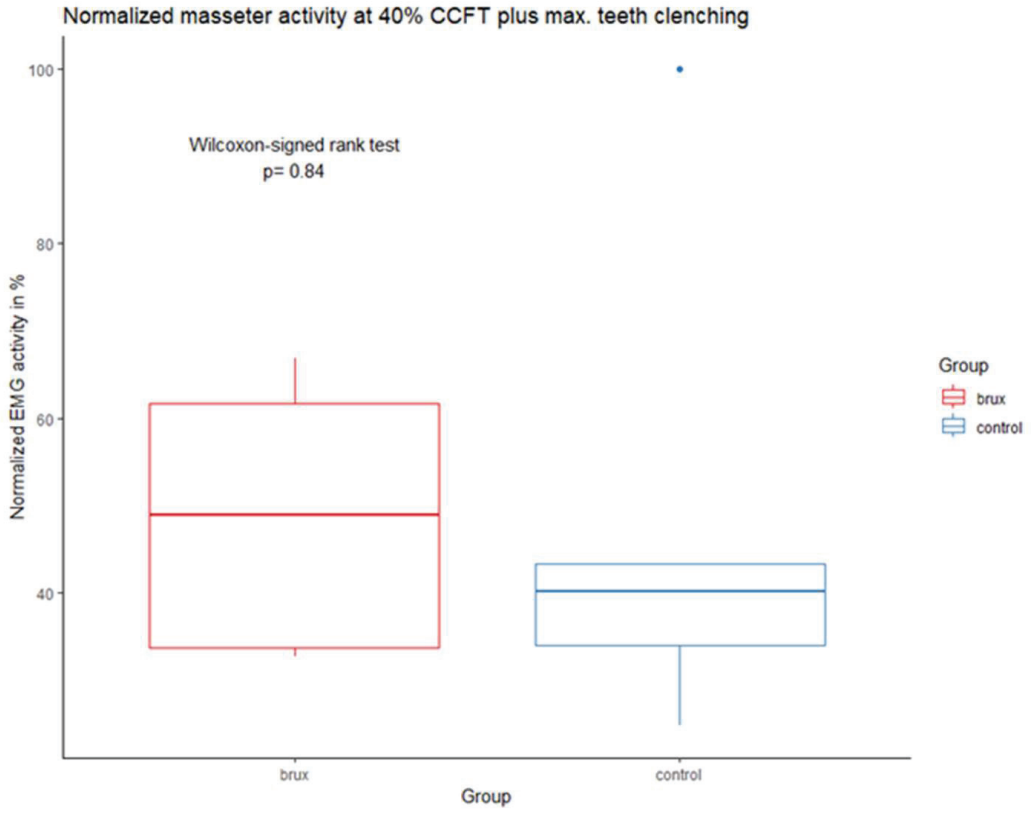


FIGURE 6 - Line plot with group mean values and standard errors of normalized sEMG activity of the masseter muscles during clenching teeth and at CCFT stage 2.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

2 mmHg intervals between 20 and 30 mmHg, leaving only six values that can be achieved when including the starting value of 20 mmHg. To overcome these limitations, a "performance index" and "cumulative performance index" have additionally been established, but their use is time-consuming in clinical practice (8). Further, whether the manometer gauge of the PBU can exactly display the value someone can repeatedly and correctly hold in clinical settings is still doubtful. Uthaikhup and Jull have shown mean shortfalls of approximately 0.6 mmHg for the upper stage of 30 mmHg for elderly asymptomatic individuals (14). No study has examined potential shortfalls or the opposite overshooting, using the PBU in symptomatic populations. The NOD can overcome these shortcomings by offering a display with considerably more options for % values. Besides those percentages used in the current study, further interim values can be tested, too. By using additional settings via the NOD app, one can even add % thresholds in which a correct test needs to be performed, like a ±5% target error. At the end of the test, the App provides a "force score index," which might give an indication of the performance for that stage (see Fig. 7)

FIGURE 7 - further NOD settings and "force score index."

Unfortunately, these settings are currently not described in the NOD manual and were, at the time of the current study, also not known by the authors and accordingly not implemented in the current study. Furthermore, the scientific value of any "force score index", using a 5% or larger target error, needs to be first determined in future studies with symptomatic populations.

Despite these technical innovations, for examining the CCFT with the NOD, especially in clinical settings, without the use of EMG, therapists still need good observation and manual palpation skills to detect compensatory movements and/or superficial muscle activation, as it has been described before (1,54).

At the end of the testing procedure, a final report can be extracted from the App, making pre- to post-treatment comparisons possible, but this also needs to be scientifically evaluated first.

Clinically, bruxism participants showed a tendency for lower masseter muscle activation during maximal teeth clenching, while during other tests involving neck flexor muscle activity, the bruxism group had a tendency for more activity in the masseter muscles, which, however, also remained not significant. These results are in line with those described by Palinkas et al., who reported lower EMG activity for both masseter muscles during a maximal clenching task in cases with bruxism when compared to controls, which, similar to our results, remained not significant (39). Testa et al. used high-density sEMG in neck pain patients without TMD compared to controls and found significantly more masseter activity in the neck pain group during a maximal tooth clenching task and a changed activation pattern during submaximal clenching or chewing tasks (55). However, they did not specifically report on bruxism (55). Whether masseter activation in participants with dysfunctions in the neck and head region really differs when non-masticatory functions are tested needs to be examined in future cross-sectional or longitudinal studies.

Between-group results of the current study need to be regarded with caution, as this was not the primary objective. No specific a-priori hypothesis has been stated, and the rather small sample size may have introduced a beta error of not finding significant group differences. However, the findings of the current study can serve as preliminary evidence in planning future studies, including an appropriate sample size calculation.

Carrasco-Uribarren et al. developed another new device to measure the CCFT (56). In their study, the CCFT was measured by assessing gliding between two surfaces within their device and initiated by the occiput on the outer surface. Whether gliding might have also been initiated by neck extensors via retraction remains unknown, as no EMG measurements had been performed in parallel, and a fulcrum

for cranio-cervical flexion, like it is present for the NOD or the PBU, is apparently not present in their device (56).

Nevertheless, friction-free gliding of the occiput on the treatment table has also been established in the current study, though rather improvised using a plastic foil. This practical limitation has not been mentioned in clinical or laboratory studies so far, though obviously, cranio-cervical flexion might be impeded by friction between hair and the texture of various treatment tables. However, more elaborate material needs to be used in future studies.

Clinical implications

The current measurement protocol and using the NOD might be suitable for testing the CCFT in laboratory settings. In clinical settings, at least the same precautions to detect substitution strategies as have been described for using the PBU (1,20,54), need to be considered.

Limitations

Testing and data analysis were not blinded and could have introduced an observation bias. The sample size was low, as the study was primarily intended as a pilot study, exploring the feasibility of the NOD device to examine the CCFT in research studies.

Subjective examination by questionnaires was rather short and superficial. However, participants were further and objectively examined by the first author for typical signs of bruxism like hypertrophy of the masseter muscles, cheek and/or tongue impressions, abraded tooth tips or exposed tooth necks (57). Additional questioning regarding stress or depression leading to typical habits like grinding or clenching should be conducted in subsequent studies. Nevertheless, bruxism is, at least controversially discussed, for being a pain syndrome or a behavioural dysfunction only (58). Subsequent studies should screen all participants with a questionnaire about neck complaints, e.g., by using the neck disability index (NDI) to narrow down the probability of an impaired CCFT. An independent evaluation by dentists may also be performed in larger subsequent studies, but was disregarded for this pilot study.

Further evaluation of TMJ and upper cervical spine mobility, including pain responses to passive examination, can help to better identify bruxism patients who may suffer from comorbidities.

Even though none of the participants complained that the screen was too far away, a tripod is not ideal for use in future clinical and laboratory practice. A mobile stand that can be attached to a treatment plinth should be tested in future studies.

Conclusion

The staged CCFT, by using the current measurement protocol and the NOD device, might be a useful alternative to the currently used PBU in laboratory and clinical settings. The "force index score", provided by the NOD, might be a useful adaptation to the current protocol, but needs to be examined for its scientific value in future studies. Patients with subjective and objective signs of bruxism showed similar and increasing activity of the sternocleidomastoideus and masseter muscles, like bruxism-free control subjects.

Acknowledgments

The authors thank:

– All participants
– Marco Barbero and Corrado Cescon from the University of Applied Sciences and Arts of Southern Switzerland (SUPSI) for helping to set up the measurement protocol with the NOD and providing the supplementary material.
– Gabriela Carvalho from the Furtwangen University, Germany, for specific hints to further NOD settings.

Author's contributor role

MJ conceived the study and designed the study protocol; was responsible for recruiting and examining participants in the movement laboratory, and writing the manuscript. AZ helped measure, process, and analyze the EMG data. MJE contributed to the design, helped during the measurements, analyzed the data, and wrote the manuscript. All authors read and approved the final manuscript. All authors agreed to be accountable for all aspects of the work, ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Data availability statement

Data is available by reasonable request.

Disclosures

Conflict of interest: The authors declare no conflict of interest.

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Falla DL, Jull GA, Hodges PW. Patients with neck pain demonstrate reduced electromyographic activity of the deep cervical flexor muscles during performance of the craniocervical flexion test. Spine. 2004;29(19):2108-2114. CrossRef PubMed
2. Romeo A, Baccini M, Carreras G, et al. Reliability, validity, and responsiveness of the craniocervical flexion test in people who are asymptomatic and patients with nonspecific neck pain: a systematic review and meta-analysis. Phys Ther. 2022;102(7):pzac054. CrossRef PubMed
3. García-Pérez-de-Sevilla G, Gónzalez-de-la-Flor Á, Martín-Vera D, et al. Deep cervical muscles and functionality in patients with chronic tension-type headache: an observational study. Medicina (Kaunas). 2022;58(7):917. CrossRef PubMed
4. Anarte-Lazo E, Carvalho GF, Schwarz A, et al. Differentiating migraine, cervicogenic headache and asymptomatic individuals based on physical examination findings: a systematic review and meta-analysis. BMC Musculoskelet Disord. 2021;22(1):755. CrossRef PubMed
5. Florencio LL, Ferracni GN, Chaves TC, et al. Analysis of head posture and activation of the cervical neck extensors during a low-load task in women with chronic migraine and healthy

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

participants. J Manipulative Physiol Ther. 2018;41(9):762-770. CrossRef PubMed
6. Greenbaum T, Dvir Z, Emodi-Perelmam A, et al. Relationship between specific temporomandibular disorders and impaired upper neck performance. Eur J Oral Sci. 2020;128(4):292-298. CrossRef PubMed
7. Jull GA, O'Leary SP, Falla DL. Clinical assessment of the deep cervical flexor muscles: the craniocervical flexion test. J Manipulative Physiol Ther. 2008;31(7):525-533. CrossRef PubMed
8. James G, Doe T. The craniocervical flexion test: intra-tester reliability in asymptomatic subjects. Physiother Res Int. 2010;15(3):144-149. CrossRef PubMed
9. Jull G, Barrett C, Magee R, Ho P. Further clinical clarification of the muscle dysfunction in cervical headache. Cephalalgia. 1999;19(3):179-185. CrossRef PubMed
10. Araujo FXD, Ferreira GE, Scholl Schell M, et al. Measurement properties of the Craniocervical Flexion Test: a systematic review. Phys Ther. 2020;100(7):1094-1117. CrossRef PubMed
11. Falla D, O'Leary S, Farina D, et al. Association between intensity of pain and impairment in onset and activation of the deep cervical flexors in patients with persistent neck pain. Clin J Pain. 2011;27(4):309-314. CrossRef PubMed
12. Jull G, Falla D. Does increased superficial neck flexor activity in the craniocervical flexion test reflect reduced deep flexor activity in people with neck pain? Man Ther. 2016;25:43-47. CrossRef PubMed
13. Selistre LFA, Melo CS, Noronha MA. Reliability and validity of clinical tests for measuring strength or endurance of cervical muscles: a systematic review and meta-analysis. Arch Phys Med Rehabil. 2021;102(6):1210-1227. CrossRef PubMed
14. Uthaikhup S, Jull G. Performance in the cranio-cervical flexion test is altered in elderly subjects. Man Ther. 2009;14(5):475-479. CrossRef PubMed
15. Lonati D. Corrado Cescon, Bonafine A, et al. P1.70- Validation of a digital dynamometer for assessing the cranio-cervical flexion test in volunteers with neck pain. In: XXV ISEK Conference Proceedings. 2024. Online (Accessed December 2024)
16. Alalawi A, Devecchi V, Gallina A, et al. Assessment of neuromuscular and psychological function in people with recurrent neck pain during a period of remission: cross-sectional and longitudinal analyses. J Clin Med. 2022;11(7):2042. CrossRef PubMed
17. Blomgren J, Strandell E, Jull G, et al. Effects of deep cervical flexor training on impaired physiological functions associated with chronic neck pain: a systematic review. BMC Musculoskelet Disord. 2018;19(1):415. CrossRef PubMed
18. Florencio LL, de Oliveira AS, Will-Lemos T, et al. Muscle endurance and cervical electromyographic activity during submaximal efforts in women with and without migraine. Clin Biomech (Bristol). 2021;82:105276. CrossRef PubMed
19. Scholl Schell M, de Araujo FX, da Rosa Almeida R, et al. Does focus of attention alter craniocervical flexion test motor learning? A randomized controlled trial. Hum Mov Sci. 2020;74:102709. CrossRef PubMed
20. Jull GA. Deep cervical flexor muscle dysfunction in whiplash. J Musculoskeletal Pain. 2000;8(1-2):143-154. CrossRef
21. Amiri M, Jull G, Bullock-Saxton J, et al. Cervical musculoskeletal impairment in frequent intermittent headache. Part 2: subjects with concurrent headache types. Cephalalgia. 2007;27(8):891-898. CrossRef PubMed
22. Rodrigues A, Florencio LL, Martins J, et al. Craniocervical flexion test in patients with migraine: discriminative validity and accuracy. Int J Clin Pract. 2021;75(7):e14248. CrossRef PubMed
23. Zito G, Jull G, Story I. Clinical tests of musculoskeletal dysfunction in the diagnosis of cervicogenic headache. Man Ther. 2006;11(2):118-129. CrossRef PubMed
24. Johnston V, Jull G, Souvlis T, et al. Neck movement and muscle activity characteristics in female office workers with neck pain. Spine. 2008;33(5):555-563. CrossRef PubMed
25. Jull G, Amiri M, Bullock-Saxton J, et al. Cervical musculoskeletal impairment in frequent intermittent headache. Part 1: subjects with single headaches. Cephalalgia. 2007;27(7):793-802. CrossRef PubMed
26. Piekartz HV, Rösner C, Batz A, et al. Bruxism, temporomandibular dysfunction and cervical impairments in females – results from an observational study. Musculoskelet Sci Pract. 2020;45:102073. CrossRef PubMed
27. Eriksson PO, Zafar H, Nordh E. Concomitant mandibular and head-neck movements during jaw opening-closing in man. J Oral Rehabil. 1998;25(11):859-870. CrossRef PubMed
28. Häggman-Henrikson B, Nordh E, Eriksson PO. Increased sternocleidomastoid, but not trapezius, muscle activity in response to increased chewing load. Eur J Oral Sci. 2013;121(5):443-449. CrossRef PubMed
29. Giannakopoulos NN, Schindler HJ, Hellmann D. Co-contraction behaviour of masticatory and neck muscles during tooth grinding. J Oral Rehabil. 2018;45(7):504-511. CrossRef PubMed
30. Rodríguez K, Miralles R, Gutiérrez MF, et al. Influence of jaw clenching and tooth grinding on bilateral sternocleidomastoid EMG activity. Cranio. 2011;29(1):14-22. CrossRef PubMed
31. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009;42(2):377-381. CrossRef PubMed
32. Guzman J, Hurwitz EL, Carroll LJ, et al. A new conceptual model of neck pain: linking onset, course, and care: the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. J Manipulative Physiol Ther. 2009;32(2)(suppl):S17-S28. CrossRef PubMed
33. Barbero M, Merletti R, Rainoldi A. Atlas of Muscle Innervation Zones. Springer Milan; 2012, CrossRef.
34. Besomi M, Devecchi V, Falla D, et al. Consensus for experimental design in electromyography (CEDE) project: checklist for reporting and critically appraising studies using EMG (CEDE-Check). J Electromyogr Kinesiol. 2024;76:102874. CrossRef PubMed
35. Florencio LL, de Oliveira AS, Carvalho GF, et al. Cervical muscle strength and muscle coactivation during isometric contractions in patients with migraine: a cross-sectional study. Headache. 2015;55(10):1312-1322. CrossRef PubMed
36. Dumas JP, Arsenault AB, Boudreau G, et al. Physical impairments in cervicogenic headache: traumatic vs. nontraumatic onset. Cephalalgia. 2001;21(9):884-893. CrossRef PubMed
37. Tolentino GA, Bevilaqua-Grossi D, Carvalho GF, et al. Relationship Between Headaches and Neck Pain Characteristics With Neck Muscle Strength. J Manipulative Physiol Ther. 2018;41(8):650-657. CrossRef PubMed
38. Giannakopoulos NN, Schindler HJ, Rammelsberg P, et al. Co-activation of jaw and neck muscles during submaximum clenching in the supine position. Arch Oral Biol. 2013;58(12):1751-1760. CrossRef PubMed
39. Palinkas M, Bataglion C, de Luca Canto G, et al. Impact of sleep bruxism on masseter and temporalis muscles and bite force. Cranio. 2016;34(5):309-315. CrossRef PubMed
40. Kilinc HE, Onan D, Ulger O. Investigation of masticatory muscle thickness and mechanosensitivity of cervical and masticatory muscles in myofascial temporomandibular disorder patients

with bruxism: a cross-sectional study. Musculoskelet Sci Pract. 2024;70:102919. CrossRef PubMed
41. R Core Team. R: A Language and Environment for Statistical Computing. Online (Accessed December 2024)
42. Singmann H, Bolker B, Westfall J, et al. afex: analysis of factorial experiments. Published online September 1, 2024. Online (Accessed December 2024)
43. Stevenson M, Sergeant E, Heuer C, et al. epiR: tools for the analysis of epidemiological data. Published online November 19, 2024. Online (Accessed December 2024)
44. Aragon TJ, Fay MP, Wollschlaeger D, et al. epitools: epidemiology tools. Published online March 22, 2020. Online (Accessed December 2024)
45. Wickham H, Chang W, Henry L, et al. ggplot2: create elegant data visualizations using the grammar of graphics. Published online April 23, 2024. Online (Accessed December 2024)
46. Auguie B, Antonov A. gridExtra: miscellaneous functions for "Grid" graphics. Published online September 9, 2017. Online (Accessed December 2024)
47. Wickham H. plyr: tools for splitting, applying and combining data. Published online October 2, 2023. Online (Accessed December 2024)
48. Wickham H. reshape: flexibly reshape data. Published online April 12, 2022. Online (Accessed December 2024)
49. Liang Z, Thomas L, Jull G, et al. Cervical musculoskeletal impairments in migraine. Arch Physiother. 2021;11(1):27. CrossRef PubMed
50. Fernandez-de-las-Peñas C, Pérez-de-Heredia M, Molero-Sánchez A, et al. Performance of the craniocervical flexion test, forward head posture, and headache clinical parameters in patients with chronic tension-type headache: a pilot study. J Orthop Sports Phys Ther. 2007;37(2):33-39. CrossRef PubMed
51. Getsoian SL, Gulati SM, Okpareke I, et al. Validation of a clinical examination to differentiate a cervicogenic source of headache: a diagnostic prediction model using controlled diagnostic blocks. BMJ Open. 2020;10(5):e035245. CrossRef PubMed
52. Villanueva-Ruiz I, Falla D, Lascurain-Aguirrebeña I. Effectiveness of specific neck exercise for nonspecific neck pain; usefulness of strategies for patient selection and tailored exercise—a systematic review with meta-analysis. Phys Ther. 2022;102(2):pzab259. CrossRef PubMed
53. Pérez-Fernández T, Armijo-Olivo S, Liébana S, et al. A novel use of inertial sensors to measure the craniocervical flexion range of motion associated to the craniocervical flexion test: an observational study. J Neuroeng Rehabil. 2020;17(1):152. CrossRef PubMed
54. Falla DL, Campbell CD, Fagan AE, et al. Relationship between cranio-cervical flexion range of motion and pressure change during the cranio-cervical flexion test. Man Ther. 2003;8(2):92-96. CrossRef PubMed
55. Testa M, Geri T, Gizzi L, et al. High-density EMG reveals novel evidence of altered masseter muscle activity during symmetrical and asymmetrical bilateral jaw clenching tasks in people with chronic nonspecific neck pain. Clin J Pain. 2017;33(2):148-159. CrossRef PubMed
56. Carrasco-Uribarren A, Marimon X, Portela A, et al. A novel device for assessment and treatment of upper cervical spine: test-retest reliability study. J Clin Med. 2023;12(5):1954. CrossRef PubMed
57. Manfredini D, Colonna A, Bracci A, et al. Bruxism: a summary of current knowledge on aetiology, assessment and management. Oral Surg. 2020;13(4):358-370. CrossRef
58. Voß LC, Basedau H, Svensson P, et al. Bruxism, temporomandibular disorders, and headache: a narrative review of correlations and causalities. Pain. 2024;165(11):2409-2418. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 148-157
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3476
ORIGINAL RESEARCH ARTICLE

Validation of a Novel device for Assessing Neck Muscle Strength

Michail Arvanitidis[1], Hon Hin Ken Mak[1], Eduardo Martinez-Valdes[1], Marco Barbero[2], Deborah Falla[1]

[1]Centre of Precision Rehabilitation for Spinal Pain (CPR Spine), School of Sport, Exercise and Rehabilitation Sciences, College of Life and Environmental Sciences, University of Birmingham, Birmingham - UK
[2]Rehabilitation Research Laboratory 2rLab, Department of Business Economics, Health and Social Care (DEASS), University of Applied Sciences and Arts of Southern Switzerland (SUPSI), Manno - Switzerland

ABSTRACT

Background: The Neuromuscular Cranio-Cervical Device (NOD) was originally designed to evaluate Cranio-Cervical Flexion Test performance but can also be used as a handheld dynamometer for testing other muscle groups, including neck muscle strength. It offers a potential alternative to the Multi-Cervical Unit (MCU), a fixed dynamometer, more closely aligned with isokinetic dynamometry, the gold standard. However, its validity and reliability need to be established. This study aimed to evaluate concurrent validity compared to the MCU and inter- and intra-rater reliability of the NOD for measuring neck flexion and extension muscle strength.
Methods: Twenty participants were assessed for neck flexion/extension strength whilst in a seated position, with the measurements repeated over three sessions. Concurrent validity was assessed by comparing NOD measurements to the MCU using Pearson correlation coefficients, and reliability was determined using Intraclass Correlation Coefficients (ICCs).
Results: Concurrent validity was strong for extension (r = 0.954) but lower for flexion (r = 0.705), indicating some variability in flexion measurements. Inter-rater reliability was good to excellent for both flexion (ICC = 0.931) and extension (ICC = 0.896). Intra-rater reliability for extension was good to excellent (ICC = 0.893), while flexion ranged from moderate to excellent (ICC = 0.844).
Conclusions: The NOD is a valid tool, particularly for extension measurements, although further refinement of testing is needed to improve the accuracy for flexion strength measurements. It is also reliable for both extension and flexion, showing promise as a practical, affordable, portable tool with real-time feedback for the assessment of neck muscle strength in clinical settings.

Keywords: Handheld Dynamometer, Multi-Cervical Unit, Neck Muscle Strength, Concurrent Validity, Inter-rater Reliability, Intra-rater Reliability

What is already known about this topic:

- *The NOD, initially designed for use during the Cranio-Cervical Flexion Test, can also be used as a handheld dynamometer for neck strength testing, though its validity and reliability remain unclear.*

What does the study add:

- *This study establishes the NOD as a valid and reliable tool for assessing neck strength, particularly neck extension strength, providing a portable, practical alternative to fixed devices for clinical and research applications.*

Introduction

Muscle strength, typically defined as the force generated by muscles during a maximal isometric contraction (1), is a crucial measure in clinical practice for the assessment of physical function and for guiding rehabilitation progress (2). People with neck pain, like other musculoskeletal disorders, commonly present with reduced strength (3), and resistance training programmes are often prescribed to address these deficits (4,5). Additionally, individuals in remission from neck pain, particularly following a whiplash injury, frequently display reduced neck strength, with lower flexion strength being a predictor of higher neck disability at a 6-month follow-up (6). Assessing neck muscle strength is therefore essential to establish baseline physical function, monitor rehabilitation progress, and guide prognostic decisions. However, traditional manual muscle testing (MMT) methods, where clinicians provide resistance and estimate the force generated by the individual, have been criticized for their subjectivity, variability in application based on clinician skill, and lack of standardization (7). These limitations are especially problematic when testing neck flexion and

Received: January 27, 2025
Accepted: May 12, 2025
Published online: June 10, 2025

This article includes supplementary material

Corresponding author:
Michail Arvanitidis
email: m.arvanitidis@bham.ac.uk

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

extension, where no opposite side or limb is available for comparison, and factors like individual positioning and clinician experience further affect measurement accuracy (7). A recent systematic review and meta-analysis (8) confirmed that manual neck strength measurements lack clinical reliability, raising concerns about their ability to track changes in muscle function over time.

With advancements in technology, more objective methods for measuring muscle strength have emerged, such as force transducers, fixed frame dynamometry, and isokinetic dynamometry. Isokinetic dynamometers are devices that control movement speed and provide accommodating resistance throughout a predetermined range of motion (9). Isokinetic dynamometry, provides valuable insights into muscle strength across varying speeds, allowing for detailed analysis of muscle performance under controlled conditions. It is widely used for peripheral joints (10-14) but is also applied to the spine, particularly for assessing trunk muscles (15-21). For neck strength measurements, fixed frame dynamometry is more commonly used, with the Multi-Cervical Rehabilitation Unit (MCU) being a well-known example, due to its high reliability and validity (22). The MCU is a more sophisticated device, offering controlled testing conditions and greater stability, making it more comparable to isokinetic dynamometers, the gold standard (23) for objective muscle strength assessment. However, handheld dynamometers (HHDs) have become popular among clinicians due to their convenience, affordability, and ease of use. Various studies have assessed the reliability (2,24-33) and validity (24,29,30,32) of different HHDs. These studies used different protocols, including variations in participant positioning, which can significantly affect results. For example, Krause et al., (2018) demonstrated that neck strength measurements differ significantly between sitting and lying positions (28). Studies evaluating HHDs in seated positions, the position used when evaluating neck muscle strength with the MCU, are particularly relevant. However, only five studies (24,27-29,33) have assessed neck strength in seated participants, with only two (24,29) evaluating the concurrent validity of the HHD. Although these studies demonstrated high validity and reliability, neither directly compared an HHD to the MCU (i.e., they compared between HHDs and handheld mode versus a wall-mounted configuration), leaving a gap in the understanding of the accuracy of HHDs for the measurement of neck muscle strength.

Recently, a new handheld dynamometer called "*Neuromuscular Cranio-Cervical Device*" (NOD; OT Bioelettronica, Turin, Italy) was developed, offering distinct advantages over existing HHDs. The NOD provides greater stability through its two-handed operation and allows real-time visualization of the force produced via Bluetooth, enhancing interaction, as individuals can see their force output. The NOD's portability and enhanced features make it a promising candidate, but its validity, specifically concurrent validity, and reliability need to be rigorously tested. Therefore, this study aims to evaluate the validity of the NOD handheld dynamometer for measuring neck flexor and extensor muscle strength, in comparison to the MCU, and to also assess its reliability. This comparison will provide essential insights into the NOD's potential use in both clinical and research settings.

Methods

Design and setting

This validity and reliability study was conducted in accordance with the Declaration of Helsinki, approved by the Ethics Committee of the University of Birmingham (approval number: MCR2122_05). The methods and reporting of findings were carried out in accordance with the Guidelines for Reporting Reliability and Agreement Studies (GRRAS) (34).

Data was collected between 01 April 2022 and 01 July 2022 at a laboratory located within the Centre of Precision Rehabilitation for Spinal Pain, School of Sport, Exercise and Rehabilitation Sciences at the University of Birmingham, UK. All participants provided written, informed consent before participating in the study and attended the laboratory on three separate days.

The study involved two raters and three sessions. Session 1 was conducted by Rater A (HHKM), who tested the participants using the handheld dynamometer (NOD) and the MCU (randomized order; rest of 15 mins between). This session was used to assess the validity of the handheld dynamometer compared to the MCU device. After 48 hours (session 2), Rater B (MA) performed the same neck muscle strength measurements using the handheld dynamometer to evaluate inter-rater reliability. The final session (session 3), conducted seven days after session one, was performed by Rater A using the handheld dynamometer to assess intra-rater reliability, capturing potential variability over different days while ensuring no significant changes in strength.

Session 2 was carried out within 48 hours after session 1 to achieve a balance between fatigue recovery and homogeneity, as conducting both measurements within the same session might over-exert participants and introduce performance bias (35). A time interval longer than one day allows for sufficient recovery while maintaining consistency in the maximal voluntary contraction (MVC). Session 3 took place seven days after session 1, which was five days after session 2, to further ensure adequate recovery and minimize the effects of fatigue, ensuring stability in MVC measurements over time.

Participants

A total of 20 participants (8 males, 12 females) were recruited from the student and staff population of the University of Birmingham using various recruitment methods, including email, posters, and peer contacts. To ensure a homogeneous sample and minimize confounding factors, this initial validation study was conducted exclusively with individuals without neck pain.

The sample size was calculated based on a moderate expected effect size of Pearson's $r = 0.6$ (36) for the assessment of concurrent validity using G*Power software version 3.1 (37,38). The parameters included an expected power (β) of 0.8, a significance level (α) of 0.05, and a two-tailed test. An initial sample size of 17 participants was determined. To account for a potential drop rate of 15%, the final required sample size was increased to 20 participants.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Inclusion and exclusion criteria

The study included individuals between the ages of 18 and 60 years old who reported pain-free neck movements. Individuals were excluded from the study if they had a history of neck pain that required treatment from a healthcare professional, cervical radiculopathy or myelopathy, spinal fractures or dislocation, spinal or thoracic surgery, osteoporosis, spinal infection, previous whiplash injury or spinal trauma, or any condition that required hospital medical care in the past 12 months. These criteria aimed to recruit an asymptomatic population with no significant prior history of neck conditions or injuries, ensuring participant safety and minimizing potential confounding factors.

Instrumentation and raters

The NOD dynamometer (OT Bioelettronica, Turin, Italy) is a 2-in-1 handheld dynamometer and biofeedback device consisting of a force cell with magnetic attachments. To record neck muscle strength, the NOD was connected to an Android tablet via Bluetooth, and its associated app was used. A silicon pad was magnetically attached to the force cell to enhance grip and participant comfort. The MCU consists of an armchair with an adjustable seat, back support and armrest, and a head fixation/assembly system where a force cell is positioned to measure the force of participants (22).

The neck muscle strength measurements from the NOD were recorded directly using the Android app in Newtons (NOD app, OT Bioelettronica, Turin, Italy). For the MCU device, the force signal was extracted via an Ethernet cable connected to a Forza device (OT Bioelettronica, Turin, Italy), which amplified the force signal (sampling rate: 100 Hz). The Forza was then connected to a data acquisition board (National Instruments/Emerson, Austin, Texas, USA), which was connected to a laptop, allowing real-time visualization, recording, and calculation of the peak force through a custom MATLAB script (version 9.9.0.1718557, R2020b, The MathWorks Inc., Natick, Massachusetts, USA). The force signal from the MCU was initially recorded in volts and subsequently converted into Newtons. Both dynamometers were calibrated once at the start of the study.

Both raters were male and experienced in using each device. Rater A had one year of experience using dynamometers, while Rater B had five years of clinical and research experience using dynamometers.

NOD measurements

The NOD handheld dynamometer (OT Bioelettronica, Turin, Italy) was used to assess MVCs of neck flexion and neck extension (randomized a priori using a randomization app). For these measurements, participants were secured with a strap around their waist while seated in a specific chair to minimize upper body compensation during force exertion (Fig. 1). For the measurement of neck flexion, the NOD was placed just superior to the participants' eyebrows (Fig. 1A), while for neck extension, the NOD was positioned against the occipital region of the skull (Fig. 1B). For each predetermined direction, participants performed one practice trial (one submaximal and one MVC contraction) to familiarize themselves with the procedure and to ensure the raters exerted a static force equal to the participants' force. After this familiarization, participants performed three MVCs in the specified direction. Verbal prompts were also provided by the raters to encourage a true maximal contraction. Participants were instructed to rest for two minutes between each MVC. The same procedure was then repeated for the opposite direction. These measurements were performed by Rater A (Session 1 and Session 3) and by Rater B (Session 2, 48 hours after Session 1).

Multi-Cervical Unit measurements

Participants were seated and stabilized using straps across their chest to prevent excessive upper body compensation during force exertion. Cervical flexion was measured by positioning the force cell of the MCU machine just above the participants' eyebrows, with a stabilization band placed behind the occipital protuberance and a stabilization clamp securing the back of the head (Fig. 1C, D). For neck extension, the force cell was positioned perpendicular to the occipital protuberance, and a stabilization band was secured just superior to the eyebrows (Fig. 1D). To familiarize participants with the testing procedure, they performed one submaximal and one MVC in the pre-determined direction (e.g., flexion). Participants were then instructed to perform three MVCs in that direction, with verbal encouragement provided by the rater, and a two-minute rest between each MVC. This protocol was repeated in the opposite direction (e.g., extension). These measurements were performed only during session 1, by Rater A.

Outcome measures and statistical analyses

Appropriate outcome measures and statistical analyses were selected in line with the GRRAS guidelines (34). Descriptive analysis was conducted to summarize the participants' demographics, with data presented as mean ± standard deviation. For statistical analysis, the mean peak isometric force from three measurements in both flexion and extension MVCs was used. Data normality was assessed using the Shapiro-Wilk test. Once normality was confirmed, parametric tests were used.

Concurrent validity was assessed by calculating Pearson's correlation coefficients (r) and the corresponding 95% CIs between the NOD and MCU measurements taken by rater 1 in session 1. Validity thresholds were classified as poor ($r < 0.50$), moderate ($0.50 \leq r < 0.70$), good ($0.70 \leq r < 0.90$), or excellent ($r \geq 0.90$), similarly to others (24). Pearson's correlation (*r*) coefficients were computed using SPSS Statistics, version 29 (IBM, USA).

The reliability of the NOD handheld dynamometer was evaluated using ICCs (3, k) and their 95% Confidence Intervals (CIs), calculated using a single-rating, absolute-agreement, two-way mixed-effects model, with two raters across 20 subjects. ICC values and their 95% CIs were calculated for inter-rater reliability (rater 1 vs rater 2 in session 2) and intra-rater reliability (rater 1 between sessions 1 and 3). Based on the criteria from (39), reliability was classified as poor ($ICC < 0.50$), moderate ($0.50 \leq ICC < 0.75$), good ($0.75 \leq ICC < 0.90$), or excellent ($ICC \geq 0.90$).

To assess measurement precision and interpretability, the Standard Error of Measurement (SEM) and Minimal

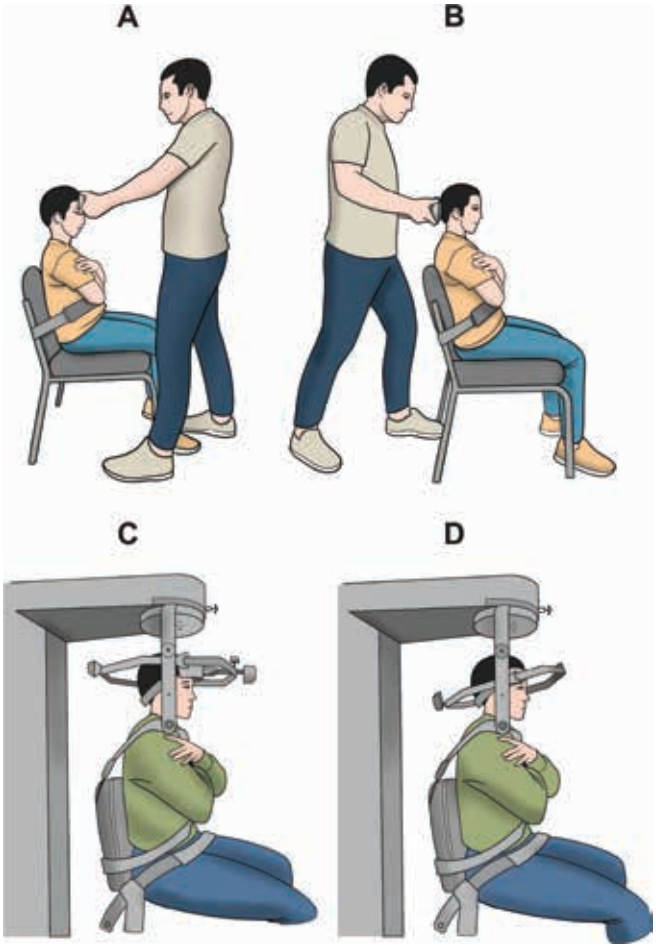


FIGURE 1 - Neck muscle strength testing for flexion and extension using two different methods: a handheld dynamometer (NOD) (A, B) and the Multi-Cervical Unit (MCU) (C, D). In panels A and B, the individual was seated with their arms crossed and a belt around the waist to stabilize the trunk, minimizing compensatory movements. The experimenter held the handheld dynamometer (NOD) on the forehead (A) for flexion and on the occiput (B) for extension, applying force while ensuring consistent contact with the head. In panels C and D, the individual is seated with their arms crossed and secured at the waist and chest with belts to stabilize the trunk and prevent extraneous movement. The head is positioned within the adjustable frame of the MCU, with the device set to the appropriate attachment for each movement. For extension (D), the attachment is tilted 15 degrees so that the point of force application is correctly aligned on the occiput.

Detectable Change (MDC) were calculated. SEM provides an estimate of the expected random variation in scores when no real change has occurred, reflecting the standard deviation of measurement error associated with each specific movement (flexion or extension) due to variability either between raters (inter-rater reliability) or across different sessions with the same rater (intra-rater reliability) (40,41).

For each movement, the pooled standard deviation (SD_{pooled}) was calculated by combining the standard deviations of the measurements across raters (for inter-rater reliability) or across sessions with the same rater (for intra-rater reliability). This represents the overall variability due to differences between or within raters and was calculated as:

$$SD_{pooled} = \sqrt{\frac{\sum(n_i - 1) \times SD_i^2}{\sum(n_i - 1)}}$$

where: SD_i is the standard deviation of measurements for each set, and n_i is the sample size for that set. The SEM was calculated as:

$$SEM = SD_{pooled} \times \sqrt{1 - ICC}$$

where the *ICC* reflects the test–retest reliability of the measurement. The MDC, representing the smallest real change beyond measurement error with 95% confidence (40,41), was then calculated as:

$$MDC = SEM \times \sqrt{2} \times 1.96$$

These metrics are crucial for identifying meaningful changes in clinical assessments and research, ensuring that observed differences are not due to random variability.

Values for SEM and MDC are reported in Newtons (N) for both flexion and extension, with separate calculations for inter-rater and intra-rater reliability. These measures provide insights into the consistency and accuracy of the NOD for assessing neck muscle strength across different testing conditions and serve as benchmarks for interpreting significant changes in muscle performance.

Scatter plots with their respective regression lines were created to visually assess the concurrent validity between MCU and NOD measurements, illustrating the strength and direction of their relationship. Additionally, Bland-Altman plots were created using GraphPad Prism version 10.3.1 (GraphPad Software, San Diego, California, USA) to visually assess the agreements and biases for the validity between the NOD and MCU measurements, as well as between the inter- and intra-rater NOD measurements for reliability. For all analyses, statistical significance was set at $\alpha = 0.05$.

Results

Participants and neck muscle strength

All 20 recruited participants (8M, 12F) attended all three sessions of the study. The characteristics of the participants are documented in Table 1. Additionally, the mean isometric neck muscle force of participants measured in each session are presented in Table 2.

TABLE 1 - Characteristics of participants

Participants' characteristics (N = 20)	Mean ± SD
Age (years)	20.7 ± 2.23
Height (cm)	165.2 ± 8.88
Weight (kg)	56.6 ± 8.07
BMI (kg/m²)	20.61 ± 1.77

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - Mean isometric neck muscle forces of all participants across sessions measures by both instruments. Data is reported as mean ± SD

Session	Equipment	Movement	Force
Session 1	*NOD*	*Flexion*	22.0 ± 5.97 N
		Extension	30.5 ± 6.63 N
	MCU	*Flexion*	19.2 ± 5.31 N
		Extension	30.7 ± 7.16 N
Session 2	*NOD*	*Flexion*	22.8 ± 5.82 N
		Extension	31.0 ± 5.99 N
Session 3	*NOD*	*Flexion*	23.4 ± 5.16 N
		Extension	30.3 ± 6.32 N

Concurrent validity

Poor to good agreement was found between the measurements of NOD and MCU for neck flexion strength. There was a significant positive correlation between the two variables, *r(18) = 0.705 [0.382, 0.875]*, *p<0.001.* (Fig. 2A). Good to excellent agreement was found for neck extension strength. There was a significant positive correlation between the two variables, *r(18) = 0.954 [0.885, 0.982], p<0.001. (*Fig. 2B)

Bland-Altman plots are presented in Figure 2C for neck flexion strength and 2D for neck extension strength. For flexion, the mean bias was −2.87 N, with limits of agreement ranging from −11.44 N to 5.70 N. For extension, the mean bias was 0.22 N, with limits of agreement ranging from −4.44 N to 4.88 N. For neck flexion, the bias suggests that NOD tends to slightly underestimate the values compared to MCU. However, for neck extension, the bias is very small, indicating a close agreement between the two methods. Most of the data points for both flexion and extension fall within the limits of agreement, showing good agreement overall, but the wider range for flexion indicates more variability in the measurements.

Reliability

Inter-rater reliability

Good to excellent reliability was found for neck flexion; the mean ICC was 0.931, with a 95% CI ranging from 0.833 to 0.972 (*F(19,19) = 30.0, p<0.001*). Similarly, good to excellent reliability was observed for neck extension; the mean ICC was 0.896, with a 95% CI ranging from 0.760 to 0.957 (*F(19,19) = 17.9, p<0.001*).

Bland-Altman plots are presented in Figure 3A and 3B. For both flexion and extension, most of the data points fall within the limits of agreement, suggesting no significant mean biases in the inter-rater measurements of NOD for either flexion (−0.75 N) or extension (−0.48 N).

Intra-rater reliability

Moderate to excellent intra-rater reliability was observed for neck flexion measurements; the mean ICC was 0.844, with a 95% CI ranging from 0.626 to 0.937 (*F(19,19) = 13.8, p<0.001*). Good to excellent intra-rater reliability was found for neck extension; the mean ICC was 0.893, with a 95% CI ranging from 0.751 to 0.956 (*F(19,19) = 17.0, p<0.001*).

Bland-Altman plots are shown in Figure 3C and 3D. For flexion, most of the data points fell within the limits of agreement, but a higher mean bias of −1.38 N was found when comparing the NOD measurements of neck flexion in session 1 against session 3. For extension, most of the data points also fell within the limits of agreement, with no significant bias observed (mean bias = 0.22 N).

The ICC (95% CI), SEM, and MDC values for both inter-rater and intra-rater reliability for neck flexion and extension strength measurements are summarized in Table 3.

TABLE 3 - Reliability measures

Characteristic	Movement	ICC (95% CI)	SEM (N)	MDC (N)
Inter-rater reliability	Flexion	0.931 [0.833, 0.972]	1.549	4.292
	Extension	0.896 [0.760, 0.957]	2.037	5.646
Intra-rater reliability	Flexion	0.844 [0.626, 0.937]	2.204	6.110
	Extension	0.893 [0.751, 0.956]	2.118	5.869

Discussion

The NOD showed good to excellent agreement with the MCU for neck extension, but the agreement for flexion was lower, ranging from poor to good. Overall, the NOD proved to be a reliable instrument, especially for neck extension, though some methodological refinements may be needed to improve its accuracy for neck flexion strength measurements.

Concurrent validity

The agreement between the NOD and MCU for neck flexion strength was lower than expected, ranging from only poor to good. One explanation for this lies in the different measurement setups for each device. The MCU employs a more rigid setup with a metal head brace and multiple chest straps, providing greater stabilization of the head and trunk. In contrast, the NOD relies on the rater to provide static resistance, which is inherently more variable. Moreover, the lack of precise head positioning markers in the NOD setup may introduce small discrepancies in initial positioning of the dynamometer against the person's forehead. It is also possible that the MCU setup, by restricting movement, encourages participants to perform a combined movement of neck flexion and protraction, whereas the NOD, with its greater freedom of movement, allows participants to perform a more natural neck flexion. This difference in movement direction could potentially affect the direction of force applied and

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

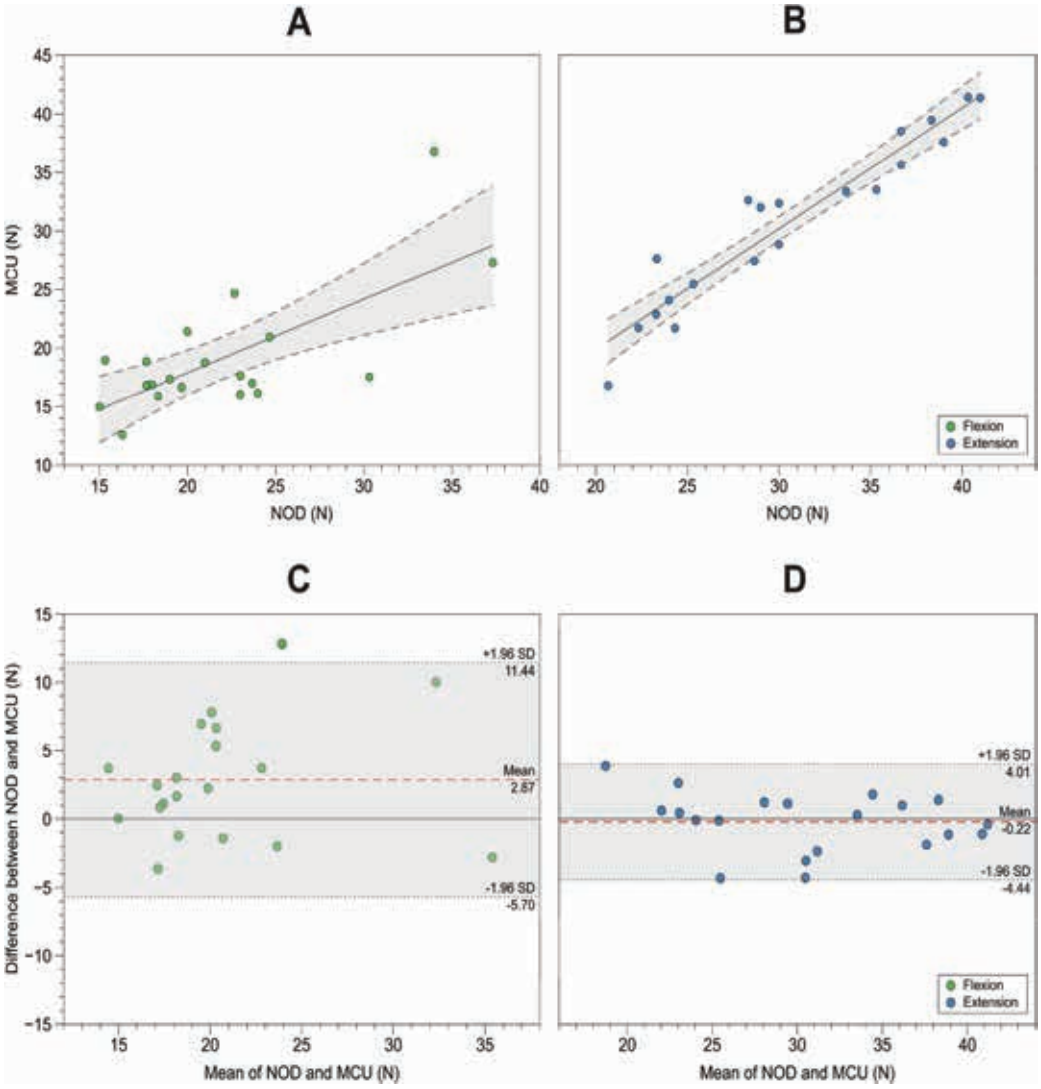


FIGURE 2 - Scatter plots with regression lines (A, B) and Bland-Altman plots (C, D) comparing NOD measurements with MCU for both neck flexion (A, C) and neck extension (B, D). In panels A and B, the x-axis represents the NOD measurements, and the y-axis represents the MCU measurements. The solid black lines indicate the regression lines, while the shaded areas represent the 95% CIs. These plots illustrate the relationship between the two methods, highlighting the strength and direction of their correlation for both flexion (A) and extension (B). In panels C and D, the x-axis represents the average of the measurements from the NOD and MCU methods, while the y-axis shows the difference between the NOD and MCU measurements. The red dashed line indicates the mean bias, reflecting the average difference between the two methods, and the black dotted lines represent the 95% limits of agreement, showing the range within which most differences are expected to fall. These plots assess the validity of the NOD method by comparing its measurements to the MCU for both flexion (C) and extension (D), highlighting any systematic bias and the level of agreement between the two methods.

contribute to the variability in force measurements. These findings are consistent with Ashall's et al., (2021) (24), who observed similar challenges when comparing measurements obtained with a HHD versus a wall-mounted dynamometer for neck flexion strength. In their study, the Pearson correlation for flexion strength was comparable to our results, with reduced agreement in flexion compared to extension. This suggests that the nature of handheld dynamometry introduces greater variability in flexion measurements due to the increased challenge in maintaining a stable, neutral head position. In addition, flexion movements tend to involve more compensatory muscle activity, especially if the head position deviates from neutral (42). These compensatory movements may lead to inconsistent force generation, resulting in lower agreement between the NOD and MCU.

In contrast, the NOD demonstrated good to excellent agreement with the MCU for neck extension strength. The rigidity provided by the seated posture and back support likely contributed to the higher consistency in extension measurements. Additionally, extension movements are generally less prone to compensatory trunk and head movements compared to flexion, resulting in more stable muscle recruitment patterns (43,44). Interestingly, these findings align with those of González-Rosalén et al., (2021) (45), who demonstrated that handheld dynamometry performs well for extension across different joints when forces remain below 200 N. Given that the neck extensors are less prone to fatigue due to their higher proportion of slow-twitch fibres (46), this could also explain the better agreement observed in extension compared to flexion, where fatigue may play a more significant role.

Reliability

Inter-rater reliability

The inter-rater reliability of the NOD was found to be good to excellent for both flexion and extension, consistent with the available literature on using HHDs to assess neck

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

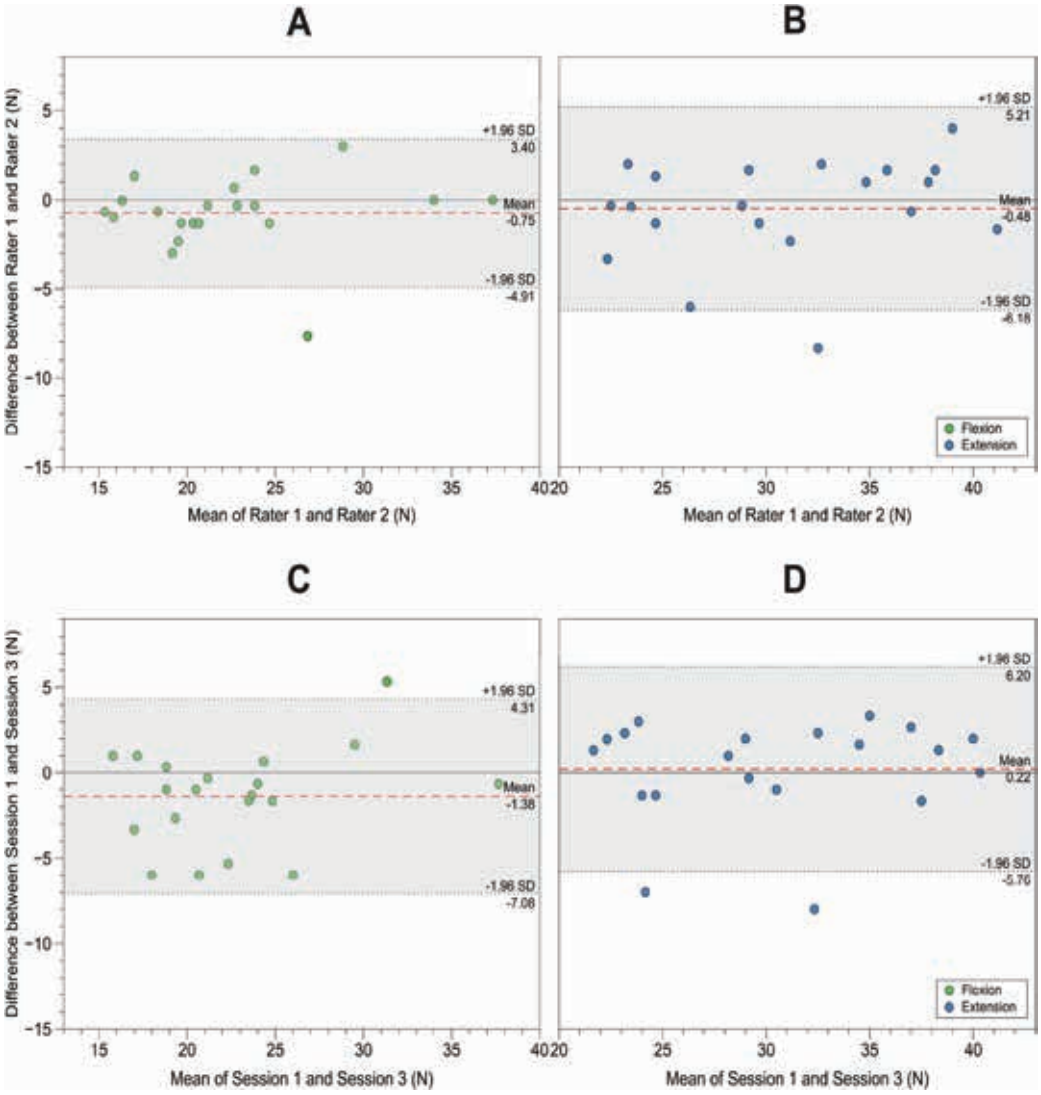


FIGURE 3 - Bland-Altman plots for inter-rater (A, B) and intra-rater (C, D) measurements of neck flexion and extension. In panels A and B, the x-axis represents the average of the two raters' measurements, while the y-axis indicates the difference between their measurements. The red dashed line represents the mean bias, showing the average difference between raters, and the black dotted lines indicate the upper and lower 95% limits of agreement, outlining the range within which most of the differences between the raters are expected to fall. In panels C and D, the x-axis represents the average of the measurements taken by the same rater across sessions 1 and 3, while the y-axis represents the difference between those repeated measurements. The red dashed line represents the mean bias, indicating the average difference between the repeated measurements by the same rater, and the black dotted lines mark the upper and lower 95% limits of agreement, showing the range within which most differences are expected to fall. Panels A and C correspond to neck flexion measurements, while panels B and D correspond to neck extension measurements.

muscle strength in a seated position. For example, Kubas et al., (2017) (29) reported similar reliability for the measurements of neck flexion and extension using an HHD. However, our study had a smaller 95% CI (0.760-0.957 vs 0.53-0.97), likely due to the larger sample size, which reduces the variability of the estimations. In contrast, the study by Kubas et al., (2021) (29) had a smaller sample (n = 10), which may have contributed to greater variability. Additionally, their study tested multiple directions (flexion, extension, side flexion, and rotation), possibly increasing muscle fatigue, which is known to affect force output (47). Our protocol limited the number of MVCs, likely reducing fatigue and contributing to more consistent reliability.

Intra-Rater Reliability

Moderate to excellent intra-rater reliability was observed for flexion using the NOD. This aligns with five previous studies examining the intra-rater reliability of HHD measurements for neck flexion in a seated position (24,27-29,33), all of which reported good reliability (ICC > 0.75). However, the 95%CIs in these studies varied substantially. For example, the study by Kubas et al., (2021) (29) reported a wide CI (0.08 to 0.96), while Vannebo et al., (2018) (33) reported a much narrower CI (0.92 to 0.97), indicating more consistent measurements. The differences in study design, including participant characteristics, time intervals between sessions, and the types of HHDs used, likely contributed to the variability in CIs. The study by Vannebo et al., (2018) (33) stands out for its higher intra-rater reliability, potentially due to participants being seated against a wall, which provided better proprioceptive input and trunk control, reducing compensatory movements and leading to more consistent measurements.

The intra-rater reliability of the NOD for neck extension strength was demonstrated to be good to excellent, aligning with four previous studies on neck extension strength assessed using HHDs with the participant in a seated position (24,27-29). Most of these studies reported similar

results, except for the study by Krause et al., (2019) (28), who reported a much wider CI range (0.21 to 0.82). This discrepancy could be due to over-exertion and muscular fatigue, as their study involved a higher number of submaximal holds and MVCs across multiple directions. In contrast, the current study required fewer MVCs, likely reducing fatigue and contributing to more consistent intra-rater measurements, as also mentioned for inter-rater reliability.

Systematic bias and learning Effect

A key finding in this study was the presence of systematic bias, particularly for the neck flexion measurements. Bland-Altman plots revealed a high bias in intra-rater flexion measurements, where force output increased by 6.27% from session 1 to session 3. This likely reflects a learning effect, where participants become more familiar with the testing procedure (33,48). The learning effect has been well documented in isometric dynamometry, where participants exhibit greater force production in later sessions due to improved technique and familiarity with the equipment (49). The absence of a similar bias in extension measurements may be due to the greater stability and consistency of extension movements as mentioned above. The back support provided during testing limits compensatory movements, ensuring that the neck extensors are more consistently recruited in the same manner across sessions.

Clinical implications

The NOD's portability and real-time feedback make it a valuable tool for clinical and research settings. However, the neck flexion results should be interpreted with some caution, as the lower agreement with the MCU suggests that variability in the testing setup may impact the accuracy of neck flexion strength measurements. Further refinement of the NOD's design or the development of standardized protocols for flexion testing could help mitigate this issue. Despite the challenges with neck flexion strength measurement, the NOD's excellent reliability and validity in extension suggest it is a reliable tool for tracking changes in neck strength over time, particularly in rehabilitation settings or in environments where portability is critical.

Methodological considerations

While this study demonstrated the validity and reliability of the NOD for the measurement of neck strength in a seated position, the sample was limited to young, asymptomatic individuals, which may restrict the generalisability of the findings to clinical populations with neck pain or older adults. The modest overall sample size, given that larger cohorts are typically advised to strengthen reliability estimates, and the relatively narrow range of demographics, particularly in terms of age and BMI, further limit the external validity of the results, as these factors are known to influence both muscle strength and measurement reliability (50). Inter- and intra-rater reliability were assessed with only one pair of raters, so variability introduced by different examiner combinations remains to be explored. Although the use of the NOD in the testing procedure adopted in this study could be considered to be reliant on the experience and physical strength of the examiner, we mitigated this by stabilizing the trunk with a strap, visually monitoring participants to ensure that they were primarily using their neck muscles with minimal trunk movement, and having both an experienced and a less experienced examiner perform the measurements, which increased our confidence in the reliability of the results. Lastly, the observed learning effect in flexion measurements suggests that further refinement of testing protocols may be needed to minimize systematic bias over multiple sessions.

Conclusions

This study provides preliminary evidence that the NOD is a reliable and valid tool for measuring neck muscle strength, particularly for neck extension. The good to excellent reliability across both inter-rater and intra-rater measurements confirms its consistency and potential suitability for clinical and research applications. Although its accuracy for neck flexion strength is influenced by some systematic bias and variability, the NOD remains a promising and reliable instrument. It may serve as a practical tool for frequent, reliable assessments in both clinical and research settings.

Acknowledgments

We would like to thank Dr. David Jiménez-Grande for his assistance with the experimental setup to extract force signals from the MCU and for developing the custom MATLAB script. We also extend our gratitude to all the participants of this study for their time and effort. This research was supported by the UK Space Agency [grant number ST/W003058/1].

Disclosures

Conflict of Interest: The authors declare no conflict of interest

Financial Support: This research was supported by the UK Space Agency [grant number ST/W003058/1].

Author Contributions: M.A., H.H.K.M., E.M.-V., M.B. and D.F. conceived and designed research; M.A. and H.H.K.M. performed experiments; M.A. and H.H.K.M. analysed data; M.A., H.H.K.M., E.M.-V., M.B. and D.F. interpreted results of experiments; M.A. and M.B. prepared figures: M.A. and H.H.K.M. drafted manuscript; M.A., H.H.K.M., E.M.-V., M.B. and D.F. edited and revised manuscript; M.A., H.H.K.M., E.M.-V., M.B. and D.F. approved final version of manuscript.

Data Availability Statement: The data supporting the findings of this study are provided in the supplementary material as an Excel file.

References

1. Enoka RM. Neuromechanics of human movement. Fifth edition ed. Champaign, IL: Human Kinetics; 2015. CrossRef
2. Versteegh T, Beaudet D, Greenbaum M, Hellyer L, Tritton A, Walton D. Evaluating the reliability of a novel neck-strength assessment protocol for healthy adults using self-generated resistance with a hand-held dynamometer. Physiother Can. 2015;67(1):58-64. CrossRef PubMed
3. Miranda IF, Wagner Neto ES, Dhein W, Brodt GA, Loss JF. Individuals With Chronic Neck Pain Have Lower Neck Strength Than Healthy Controls: A Systematic Review With

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Meta-Analysis. J Manipulative Physiol Ther. 2019;42(8):608-622. CrossRef PubMed
4. Falla D, Lindstrøm R, Rechter L, Boudreau S, Petzke F. Effectiveness of an 8-week exercise programme on pain and specificity of neck muscle activity in patients with chronic neck pain: a randomized controlled study. Eur J Pain. 2013;17(10):1517-1528. CrossRef PubMed
5. Ylinen JJ, Häkkinen AH, Takala EP, et al. Effects of neck muscle training in women with chronic neck pain: one-year follow-up study. J Strength Cond Res. 2006;20(1):6-13. CrossRef PubMed
6. Alalawi A, Devecchi V, Gallina A, Luque-Suarez A, Falla D. Assessment of Neuromuscular and Psychological Function in People with Recurrent Neck Pain during a Period of Remission: Cross-Sectional and Longitudinal Analyses. J Clin Med. 2022;11(7):2042. CrossRef PubMed
7. Cuthbert SC, Goodheart GJ Jr. On the reliability and validity of manual muscle testing: a literature review. Chiropr Osteopat. 2007;15(1):4. CrossRef PubMed
8. Selistre LFA, Melo CS, Noronha MA. Reliability and Validity of Clinical Tests for Measuring Strength or Endurance of Cervical Muscles: A Systematic Review and Meta-analysis. Arch Phys Med Rehabil. 2021;102(6):1210-1227. CrossRef PubMed
9. Osternig LR. Isokinetic dynamometry: implications for muscle testing and rehabilitation. Exerc Sport Sci Rev. 1986;14:45-80. CrossRef PubMed
10. Arvanitidis M, Falla D, Sanderson A, Martinez-Valdes E. Does pain influence control of muscle force? A systematic review and meta-analysis. Eur J Pain. 29(2):e4716 CrossRef PubMed
11. Chamorro C, Armijo-Olivo S, De la Fuente C, Fuentes J, Javier Chirosa L. Absolute Reliability and Concurrent Validity of Hand Held Dynamometry and Isokinetic Dynamometry in the Hip, Knee and Ankle Joint: Systematic Review and Meta-analysis. Open Med (Wars). 2017;12(1):359-375. CrossRef PubMed
12. Contreras-Hernandez I, Arvanitidis M, Falla D, Negro F, Martinez-Valdes E. Achilles tendon morpho-mechanical parameters are related to triceps surae motor unit firing properties. J Neurophysiol. 2024;132(4):1198-1210. CrossRef PubMed
13. Lesnak JB, Anderson DT, Farmer BE, Katsavelis D, Grindstaff TL. Ability of Isokinetic Dynamometer to Predict Isotonic Knee Extension 1-Repetition Maximum. J Sport Rehabil. 2019;29(5):616-620. CrossRef PubMed
14. Sørensen L, Oestergaard LG, van Tulder M, Petersen AK. Measurement Properties of Isokinetic Dynamometry for Assessment of Shoulder Muscle Strength: A Systematic Review. Arch Phys Med Rehabil. 2021;102(3):510-520. CrossRef PubMed
15. Arvanitidis M, Bikinis N, Petrakis S, et al. Spatial distribution of lumbar erector spinae muscle activity in individuals with and without chronic low back pain during a dynamic isokinetic fatiguing task. Clin Biomech (Bristol). 2021;81:105214. CrossRef PubMed
16. Arvanitidis M, Jiménez-Grande D, Haouidji-Javaux N, Falla D, Martinez-Valdes E. People with chronic low back pain display spatial alterations in high-density surface EMG-torque oscillations. Sci Rep. 2022;12(1):15178. CrossRef PubMed
17. Arvanitidis M, Jiménez-Grande D, Haouidji-Javaux N, Falla D, Martinez-Valdes E. Eccentric exercise-induced delayed onset trunk muscle soreness alters high-density surface EMG-torque relationships and lumbar kinematics. Sci Rep. 2024;14(1):18589. CrossRef PubMed
18. Arvanitidis M, Jiménez-Grande D, Haouidji-Javaux N, Falla D, Martinez-Valdes E. Low Back Pain-Induced Dynamic Trunk Muscle Control Impairments Are Associated with Altered Spatial EMG-Torque Relationships. Med Sci Sports Exerc. 2024;56(2):193-208. CrossRef PubMed
19. Estrázulas JA, Estrázulas JA, de Jesus K, de Jesus K, da Silva RA, Libardoni Dos Santos JO. Evaluation isometric and isokinetic of trunk flexor and extensor muscles with isokinetic dynamometer: A systematic review. Phys Ther Sport. 2020;45:93-102. CrossRef PubMed
20. García-Vaquero MP, Barbado D, Juan-Recio C, López-Valenciano A, Vera-Garcia FJ. Isokinetic trunk flexion-extension protocol to assess trunk muscle strength and endurance: Reliability, learning effect, and sex differences. J Sport Health Sci. 2020;9(6):692-701. CrossRef PubMed
21. Reyes-Ferrada W, Chirosa-Rios L, Martinez-Garcia D, Rodríguez-Perea Á, Jerez-Mayorga D. Reliability of trunk strength measurements with an isokinetic dynamometer in non-specific low back pain patients: A systematic review. J Back Musculoskelet Rehabil. 2022;35(5):937-948. CrossRef PubMed
22. Chiu TT, Sing KL. Evaluation of cervical range of motion and isometric neck muscle strength: reliability and validity. Clin Rehabil. 2002;16(8):851-858. CrossRef PubMed
23. Althobaiti S, Falla D. Reliability and criterion validity of hand-held dynamometry for measuring trunk muscle strength in people with and without chronic non-specific low back pain. Musculoskelet Sci Pract. 2023;66:102799. CrossRef PubMed
24. Ashall A, Dobbin N, Thorpe C. The concurrent validity and intrarater reliability of a hand-held dynamometer for the assessment of neck strength in semi-professional rugby union players. Phys Ther Sport. 2021;49:229-235. CrossRef PubMed
25. Carnevalli APO, Bevilaqua-Grossi D, Oliveira AIS, Carvalho GF, Fernández-De-Las-Peñas C, Florencio LL. Intrarater and Interrater Reliability of Maximal Voluntary Neck Muscle Strength Assessment Using a Handheld Dynamometer in Women With Headache and Healthy Women. J Manipulative Physiol Ther. 2018;41(7):621-627. CrossRef PubMed
26. Cibulka MT, Herren J, Kilian A, Smith S, Mahmutovic F, Dolles C. The reliability of assessing sternocleidomastoid muscle length and strength in adults with and without mild neck pain. Physiother Theory Pract. 2017;33(4):323-330. CrossRef PubMed
27. Geary K, Green BS, Delahunt E. Intrarater reliability of neck strength measurement of rugby union players using a handheld dynamometer. J Manipulative Physiol Ther. 2013;36(7):444-449. CrossRef PubMed
28. Krause DA, Hansen KA, Hastreiter MJ, Kuhn TN, Peichel ML, Hollman JH. A Comparison of Various Cervical Muscle Strength Testing Methods Using a Handheld Dynamometer. Sports Health. 2019;11(1):59-63. CrossRef PubMed
29. Kubas C, Chen YW, Echeverri S, et al. Reliability and Validity of Cervical Range of Motion and Muscle Strength Testing. J Strength Cond Res. 2017;31(4):1087-1096. CrossRef PubMed
30. Martins F, Bento A, Silva AG. Within-Session and Between-Session Reliability, Construct Validity, and Comparison Between Individuals With and Without Neck Pain of Four Neck Muscle Tests. PM R. 2018;10(2):183-193. CrossRef PubMed
31. Shahidi B, Johnson CL, Curran-Everett D, Maluf KS. Reliability and group differences in quantitative cervicothoracic measures among individuals with and without chronic neck pain. BMC Musculoskelet Disord. 2012;13(1):215. CrossRef PubMed
32. Tudini F, Myers B, Bohannon R. Reliability and validity of measurements of cervical retraction strength obtained with a hand-held dynamometer. J Man Manip Ther. 2019;27(4):222-228. CrossRef PubMed
33. Vannebo KT, Iversen VM, Fimland MS, Mork PJ. Test-retest reliability of a handheld dynamometer for measurement of isometric cervical muscle strength. J Back Musculoskelet Rehabil. 2018;31(3):557-565. CrossRef PubMed

34. Kottner J, Audigé L, Brorson S, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol. 2011;64(1):96-106. CrossRef PubMed
35. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998;26(4):217-238. CrossRef PubMed
36. Chinn S. The assessment of methods of measurement. Stat Med. 1990;9(4):351-362. CrossRef PubMed
37. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods. 2009;41(4):1149-1160. CrossRef PubMed
38. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007;39(2):175-191. CrossRef PubMed
39. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-163. CrossRef PubMed
40. Furlan L, Sterr A. The Applicability of Standard Error of Measurement and Minimal Detectable Change to Motor Learning Research-A Behavioral Study. Front Hum Neurosci. 2018;12:95. CrossRef PubMed
41. Portney LG, Watkins MP. Foundations of Clinical Research: Applications To Practice. Pearson/Prentice Hall; 2015.
42. Lee DK, Moon DC, Hong KH. Effect of neck flexion restriction on sternocleidomastoid and abdominal muscle activity during curl-up exercises. J Phys Ther Sci. 2016;28(1):90-92. CrossRef PubMed
43. Bogduk N, Mercer S. Biomechanics of the cervical spine. I: normal kinematics. Clin Biomech (Bristol). 2000;15(9):633-648. CrossRef PubMed
44. Feipel V, Rondelet B, Le Pallec J, Rooze M. Normal global motion of the cervical spine: an electrogoniometric study. Clin Biomech (Bristol). 1999;14(7):462-470. CrossRef PubMed
45. González-Rosalén J, Benítez-Martínez JC, Medina-Mirapeix F, Cuerda-Del Pino A, Cervelló A, Martín-San Agustín R. Intra- and Inter-Rater Reliability of Strength Measurements Using a Pull Hand-Held Dynamometer Fixed to the Examiner's Body and Comparison with Push Dynamometry. Diagnostics (Basel). 2021;11(7):1230. CrossRef PubMed
46. Mannion AF, Dumas GA, Cooper RG, Espinosa FJ, Faris MW, Stevenson JM. Muscle fibre size and type distribution in thoracic and lumbar regions of erector spinae in healthy subjects without low back pain: normal values and sex differences. J Anat. 1997;190 (Pt 4)(Pt 4):505-13. CrossRef
47. Cvetko E, Karen P, Eržen I. Myosin heavy chain composition of the human sternocleidomastoid muscle. Ann Anat. 2012;194(5):467-472. CrossRef PubMed
48. Ylinen J, Ruuska J. Clinical use of neck isometric strength measurement in rehabilitation. Arch Phys Med Rehabil. 1994;75(4):465-469. CrossRef PubMed
49. Blazevich, A. J., Gill, N. D., Deans, N., & Zhou, S. (2007). Lack of human muscle architectural adaptation after short-term strength training. Muscle & nerve, 35(1), 78–86. CrossRef
50. Liu J, Liu P, Ma Z, et al. The effects of aging on the profile of the cervical spine. Medicine (Baltimore). 2019;98(7):e14425. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 158-164
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3340
ORIGINAL RESEARCH ARTICLE

Head Pain Reproduction and Resolution Behavior in Response to Sustained Mobilization of the Upper Cervical Spine: A Case Series

Damien C. Cummins[1], Lucy C. Thomas[2], Peter G. Osmotherly[1]

[1]School of Health Sciences, The University of Newcastle, Callaghan - Australia
[2]School of Health and Rehabilitation Sciences, The University of Queensland, St Lucia - Australia

ABSTRACT

Background: Reproduction and resolution of head pain with sustained upper cervical mobilization has been proposed as a diagnostic indicator of Cervicogenic Headache (CGH). However, there has been little focus on describing head pain response during the application of these maneuvers. The purpose of this study was to describe the head pain response during the performance of sustained upper cervical spine mobilization in people with probable CGH.

Methods: In this case series of 20 individuals with probable CGH as defined under the International Headache Society criteria, were assessed by one experienced physiotherapist who administered five sustained upper cervical spine mobilization techniques to each participant. Reproduction of head pain during mobilization was noted. At the start, during, and end of each mobilization technique, change in head pain intensity and time taken to achieve head pain resolution was recorded.

Results: Sixteen of the 20 participants experienced reproduction and resolution of head pain within 90 seconds of sustained upper cervical mobilization. This phenomenon occurred no more frequently with the mobilization of C2 than with C1. Eight patients reported this on the dominant head pain side only, 6 patients experienced this bilaterally. Neck pain was present in 13 of the 20 participants.

Conclusion: From this case series, it seems that neither the duration (measured in seconds) nor the magnitude of reduction in head pain intensity was markedly different across dominant compared to non-dominant head pain sides or across cervical levels, indicating reproduction and resolution behavior is irregular. The presence of CGH without neck pain is possible.

Keywords: Headache, Cervical Spine, Manual Examination, Sustained Joint Mobilization

What is already known about this topic?

- *Reproduction and resolution of head pain with sustained cervical spine mobilization have been suggested as an effective strategy for the diagnosis and management of CGHs.*

What does this study add?

- *This case series provides a complete documented description of the response to sustained cervical spine mobilization in a CGH population in terms of occurrence, intensity, spread, and resolution of head pain.*

Introduction

Cervicogenic Headache (CGH) is characterized by referred head pain from musculoskeletal dysfunction in the upper three cervical segments (1). Convergence of the trigeminal nerve complex with the upper cervical spinal nerve complex in the Trigeminocervical Nucleus (TCN) is thought to be the cause of referred pain to the head region (2). Approximately 15 to 20% of chronic and recurrent headache sufferers experience head pain from a cervicogenic origin (3).

Local cervical pain or reproduction of headache symptoms with upper cervical spine palpation is considered a key component necessary to determine if a headache is of cervical origin (4,5). Whilst physical findings indicating patterns of cervical musculoskeletal impairment including postural variances,(6) active range of movement limitations (7), and muscular impairments (8) have been associated with CGH, the presence of painful joints in the upper three cervical spinal levels as detected with manual examination is the physical examination finding that most clearly distinguishes CGH from other headache disorders (9).

Received: October 14, 2024
Accepted: June 1, 2025
Published online: June 20, 2025

Corresponding author:
Peter G. Osmotherly
email: Peter.Osmotherly@newcastle.edu.au

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

Reproduction and resolution of head pain with sustained upper cervical joint mobilization achieved by maintaining thumb pressure to passively stress a particular intervertebral segment has been proposed to be an important diagnostic indicator of CGH (10). Recent work has demonstrated that reproduction and resolution of head pain with sustained upper cervical joint mobilization can distinguish people with CGH from asymptomatic individuals with a sensitivity of 78% and specificity of 90% (11).

Head pain referral from the upper cervical spine occurs due to the convergence of cervical and trigeminal afferents in the TCN where nociceptive afferents from the C1 to C3 spinal nerves converge with second-order neurons which also receive afferents from the first division of the trigeminal nerve (2). Hyperexcitability of nociceptive second order neurons in the TCN may result from dysfunction in deep somatic tissues of the upper cervical spine (12) resulting in increased sensitivity to afferent information from the trigeminal field which is perceived as pain (13). Sustaining a passive mobilization maneuver, whereby head pain is reproduced and then lessens in intensity or resolves within 90 seconds, is suggested to cause a modulation in hyperexcitability of the second-order neurons of the TCN and diminution of pre-existing central sensitization (10,13-15).

The reduction of head pain following mobilization of the upper cervical spine has been previously described. Malo-Urries et al investigated the effect of translational mobilization in individuals with CGH reporting immediate improvements in pain intensity, range of movement, and pain pressure threshold (16). The mechanism proposed for this change is the mechanical stimulus of the mobilization technique could activate the descending pain inhibitory systems via corticospinal projections from the periaqueductal grey matter causing immediate hypoalgesic effects (16,17). However, these techniques used oscillatory techniques. Further effects have been hypothesized relating to the application of sustained end-range mobilization techniques including the reduction in the effects of reflex muscle contraction on periarticular and capsular structures with subsequent reduction in peripheral afferent discharge due to mechanical adaptation of encapsulated endings of joint nerves, and reduction in intraarticular pressure resulting in reduced peripheral afferent discharge and thereby decreased pain perception (18).

There is emerging evidence of the existence and usefulness of the reproduction and resolution phenomena in the identification and management of CGH (11,13). However, there has been no attempt to date to describe the head pain in terms of occurrence and duration to head pain resolution, and the location and intensity of head pain referral in response to testing of sustained upper cervical mobilization. For this procedure to be considered clinically useful, particularly as a diagnostic indicator of CGH, it is important to clarify that the reproduction of head pain reflects the patient's head pain experience, and that the subsequent resolution includes a significant reduction or abolition of this head pain. This series of patient cases describes these head pain responses during the performance of sustained upper cervical spine mobilization in people with probable CGH.

Cervicogenic Headache (CGH) patients

Patients described in this case series were consecutively recruited from a private physiotherapy clinic in Newcastle, Australia. Determination of an individual as presenting with probable CGH was consistent with the criteria stipulated in the International Classification of Headache Disorders version 3 (ICDH-3)(19) for CGH with the exception of the diagnostic blockade. Patients were defined as having probable CGH since, under International Headache Society criteria, a definitive diagnosis of CGH is not possible without the use of diagnostic blocks (2). Explicitly, the criteria used for probable CGH and inclusion in this case series were:

- Unilateral or side-dominant headache
- Associated neck pain or stiffness
- Headache frequency at least once per week
- History of episodic headaches for at least the previous three months
- Aged over 18 years and,
- No manual cervical treatment received in the past three months.

Patients were not included in this case series description if the headache was not of cervical origin according to the subjective screening criteria developed by the International Headache Society (19), they experienced a headache strongly associated with autonomic symptoms such as dizziness or visual disturbances, or they had any condition which might contraindicate manipulative therapy (20). Individuals with features suggestive of migraine with and without aura or bilateral symmetrical headache typifying tension-type headache were not included (19).

Examination procedures and data collection

Age, gender, and dominant head pain side were recorded for each patient. Two headache questionnaires, the La Trobe University Headache Questionnaire 1.0 (21) and the Headache Impact Test (HIT-6) (22), were administered to define the characteristics of the headaches experienced by patients.

The La Trobe University Headache Questionnaire 1.0 was used to assess headache frequency, intensity, and duration. Correlation coefficients ranging from 0.51 to 0.80 for the elements of this questionnaire have been reported matched with headache diary data (23) The Headache Impact Test (HIT-6) questionnaire measured the frequency of functional disability and evaluated emotional effects of the headache. This questionnaire has demonstrated internal consistency and test-retest reliability. When discriminating across diagnostic and headache groups, relative validity coefficients of 0.82 and 1.00 and internal consistency were observed when comparing HIT-6 total scores (22).

All patients were then assessed by an experienced physiotherapist who had 28 years of clinical experience and was currently treating at least 40 patients with headaches per week. Five sustained mobilization tests were administered by the physiotherapist to each participant. The tests assessed head pain reproduction and resolution with upper

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

cervical mobilization tests. All tests were performed during one assessment session for each participant.

First, with the patient in prone lying, a centrally directed posterior to anterior (P-A) passive accessory sustained mobilization technique was applied to the spinous process of C3 as described by Maitland (24) (Fig. 1a). Following this, specific unilateral techniques were applied bilaterally, following the guidelines described by Watson and Drummond (10). The C2 unilateral technique involved applying sustained P-A pressure to the superior articular surface of the axis (C2) with the participant's head in approximately 30 degrees of contralateral rotation with the patient in prone lying (Fig. 1b). The C1 unilateral technique, applied bilaterally, involved the application of a sustained P-A pressure on the middle third region of the posterior arch of the atlas (C1). With the patient supine lying, the starting position of the participant's head was in 20 degrees of contralateral rotation. With the mobilizing thumb maintained on C1, an ipsilateral rotation force of no more than 5 degrees was applied and sustained by the other hand placed on the patient's forehead (Fig. 1c).

FIGURE 1 - Mobilization techniques applied to the upper cervical spine. (a) Central C3 spinous process posterior to anterior (P-A) mobilization, (b) Left unilateral C2 sustained mobilization, (c) Right unilateral C1 sustained mobilization.

The pressure of sustained mobilization applied to each segment was determined by the patient's response. All applied passive pressures moved into joint resistance to a magnitude whereby usual head pain was reproduced or where either local neck discomfort or tissue resistance limited motion. When head pain was provoked during any sustained upper cervical mobilization technique, at the end of each procedure the participant reported the location and intensity of head pain produced both initially and following 90 seconds of sustained mobilization. Familiarity of head pain, i.e. head pain similar to the pain usually experienced, was also noted. These reports were recorded using a body chart and an eleven-point Numerical Rating Scale (NRS). The NRS has proven convergent validity with correlation coefficients ranging from 0.79 to 0.95 (25).

Any technique failing to produce or increase head pain within 15 seconds was recorded as 'clear' on the assessment. Resolution was determined by the participant reporting a reduction in the produced head pain within 90 seconds as measured by the NRS. Furthermore, the head pain was required to be similar to the participant's previous experiences of head pain.

Data Analysis

Reproduction of head pain was presented descriptively. At the start, during and end of each mobilization technique, changes in head pain intensity and time taken to achieve head pain resolution were recorded. In individuals where reproduction and resolution of usual head pain on the dominant head pain side were evident, NRS scores on initial provocation and following 90 seconds of sustained mobilization were compared using the Wilcoxon Signed-rank test. The significance level was set as $p<0.05$.

Results

Patient Demographics

Demographic and headache characteristics of the participants are presented in Table 1. The age range of headache participants was 23 to 75 years (mean 45.8 years, sd 14.8). All headache participants reported a dominant pain side. Neck pain was present in 13 of the 20 participants.

TABLE 1 - Characteristics of the headache patients assessed

Age in years (mean (SD))	45.8 (14.8)
Gender (N (%))	
Male	6 (30)
Female	14 (70)
HIT-6 score* (mean (SD))	60.6 (5.4)
Headache frequency (N (%))	
<1/month	0 (0)
1 per month	0 (0)
2-3 per month	2 (10)
1 per week	3 (15)
2-3 per week	5 (25)
4-5 per week	5 (25)
Daily	4 (20)
>1 per day	1 (5)
Headache intensity (median (IQR))	10 (5,12)
Headache duration (N (%))	
<1 hour	4 (20)
1-2 hours	7 (35)
3-5 hours	4 (20)
6-8 hours	2 (10)
9-12 hours	0 (0)
13-24 hours	2 (10)
1-2 days	1 (5)
>2 days	0 (0)
Dominant head pain side (N (%))	
Left	6 (30)
Right	14 (70)
Neck pain reported (N (%))	
No	7 (35)
Yes	13 (65)

Legend: HIT-6 = Headache Impact Test-6, IQR = interquartile range
*Score range of HIT-6 questionnaire 36 to 78.

Patient response to mobilization testing

Sixteen out of 20 (80%) participants experienced head pain reproduction and resolution within 90 seconds of the sustained upper cervical mobilization (Fig. 2). Four participants failed to experience reproduction and resolution of their usual head pain with the sustained upper cervical mobilization. Table 2 shows the response of each participant to each mobilization technique administered.

Effects of specific mobilization techniques

Each participant was assessed using five separate sustained cervical mobilization techniques. In the 16 participants who reported reproduction and resolution of head pain with any sustained cervical mobilization technique, 42.5% of applied mobilization techniques reproduced familiar head pain with resolution. Fifty percent of reproduced and resolving head pain occurred during the mobilization technique applied to C2, 47% during application of sustained mobilization at C1 and 3% upon mobilization of C3 centrally.

Relationship between mobilization testing and dominant head pain side.

Figure 2 illustrates the response to mobilization testing and its relationship to the dominant head pain side in the 16 participants who reported reproduction and resolution of head pain with sustained mobilization. Of these 16 people, the median head pain NRS score of 5 (iqr 3,6) on initial provocation reduced to 1 (iqr 0,1) following the sustained mobilization technique ($p < 0.01$). Reproduction and resolution of head pain only on the non-dominant side only occurred in two individuals where mobilization of C1 elicited this response.

Discussion

This case series aimed to describe head pain response during the performance of sustained upper cervical spine mobilization techniques. Reproduction of any head pain was always perceived on the same side as the side of the upper cervical mobilization technique, with the exception of the centrally applied C3 spinous process mobilization. This technique was perceived equally on both sides of the head. Among the 16 headache participants who demonstrated reproduction and resolution of head pain with sustained upper cervical mobilization, neither the duration (measured in seconds) nor the magnitude of reduction in head pain intensity (measured on the NRS scale) appeared markedly different across dominant head pain sides nor across segmental levels, indicating that reproduction and resolution of head pain are not dependent on head pain side dominance nor specific segmental levels.

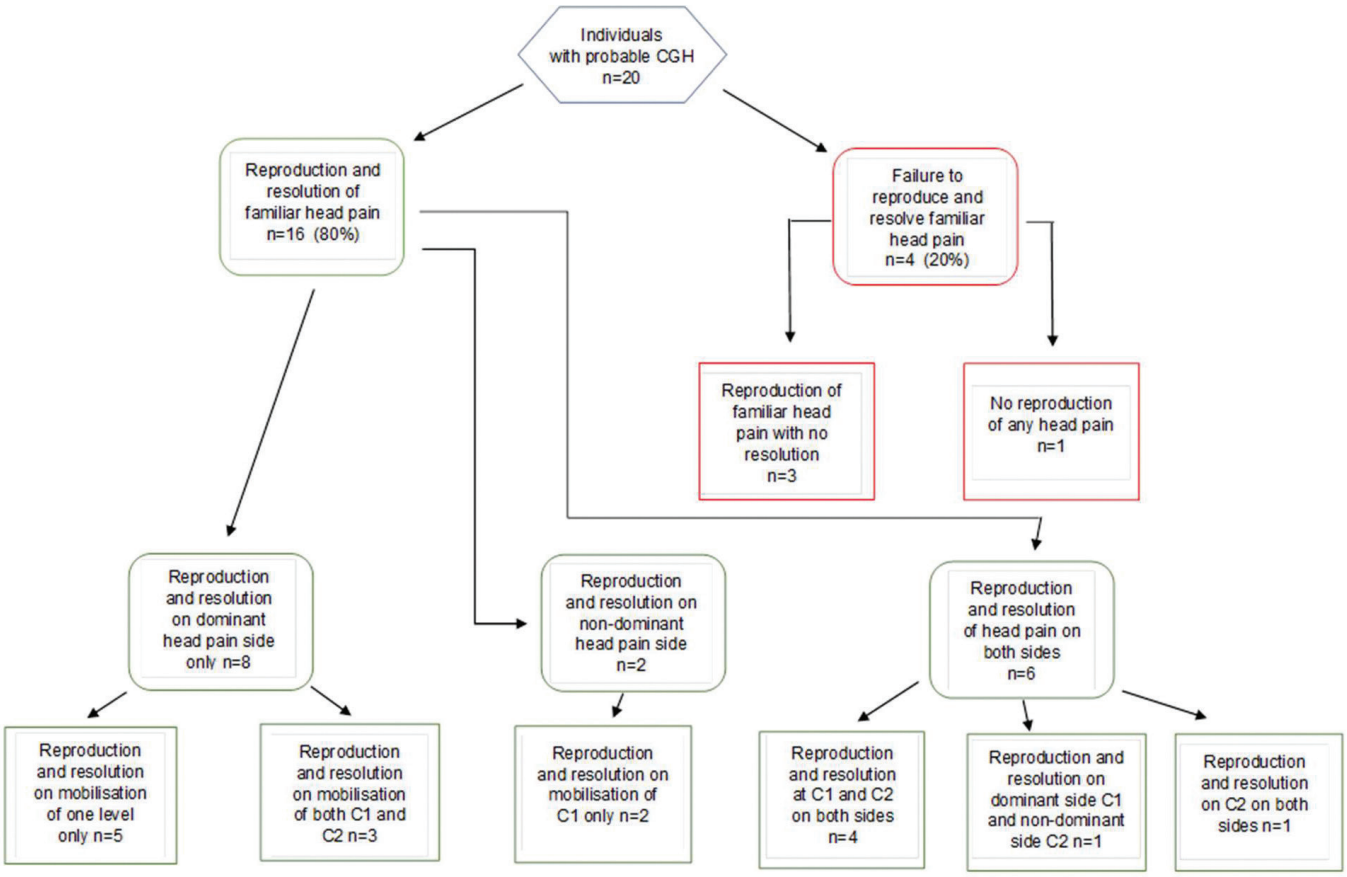


FIGURE 2 - Reproduction and resolution responses within the patient sample to sustained mobilization of the upper cervical segments.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - Palpation results with respect to dominant head pain side

Participant ID	C3 central PA					C2 left unilateral PA					C2 right unilateral PA					C1 left unilateral PA					C1 right unilateral PA					Head pain reproduced and resolved YES = 1 No = 0	Dominant head pain side
	familiar/not familiar	resolved/non resolved	palpation time (secs)	NRS start	NRS finish	familiar/not familiar	resolved/non resolved	palpation time (secs)	NRS start	NRS finish	familiar/not familiar	resolved/non resolved	palpation time (secs)	NRS start	NRS finish	familiar/not familiar	resolved/non resolved	palpation time (secs)	NRS start	NRS finish	familiar/not familiar	resolved/non resolved	palpation time (secs)	NRS start	NRS finish		
Headache 1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Left
Headache 2	0	0	0	0	0	0	0	0	0	0	1	0	90	2	2	1	0	90	2	2	1	1	90	3	1	1	Right
Headache 3	1	0	90	3	3	0	0	0	0	0	1	0	90	5	5	1	0	90	5	5	1	0	90	5	5	0	Right
Headache 4	0	0	0	0	0	1	0	90	2	4	0	0	0	0	0	1	1	90	2	1	1	0	90	4	4	1	Right
Headache 5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	90	6	0	0	0	0	0	0	1	Right
Headache 6	0	0	0	0	0	1	1	70	2	0	1	1	90	4	2	1	1	90	3	0	1	1	90	3	1	1	Right
Headache 7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	90	3	1	0	1	90	2	1	1	Left
Headache 8	0	0	0	0	0	0	0	0	0	0	1	1	90	6	0	0	1	30	2	0	1	1	90	5	4	1	Right
Headache 9	0	0	0	0	0	0	0	0	0	0	1	0	90	3	4	1	0	90	1	1	1	0	90	1	1	0	Right
Headache 10	1	1	60	8	0	1	1	70	7	0	1	1	90	8	3	1	1	90	8	2	1	1	90	8	3	1	Right
Headache 11	0	0	0	0	0	1	1	90	7	4	1	0	90	5	5	1	1	90	7	5	0	0	0	0	0	1	Left
Headache 12	0	0	0	0	0	1	1	90	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	Left
Headache 13	0	0	0	0	0	1	1	70	6	0	1	1	70	5	0	1	1	45	3	0	1	1	90	3	1	1	Right
Headache 14	1	0	90	5	5	1	1	90	5	3	1	0	90	4	7	1	0	90	7	9	1	1	90	4	2	1	Left
Headache 15	0	0	0	0	0	1	1	90	2	0	1	1	50	3	0	1	1	90	3	2	1	1	90	2	1	1	Left
Headache 16	1	0	90	1	3	0	0	0	0	0	1	1	90	3	1	0	0	0	0	0	1	1	90	3	2	1	Right
Headache 17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	2	0	0	0	0	0	0	Right
Headache 18	0	0	0	0	0	1	1	90	9	4	1	1	90	6	4	1	0	90	6	6	1	0	90	9	9	1	Left
Headache 19	0	0	0	0	0	0	0	0	0	0	1	1	90	6	1	1	0	0	0	0	1	0	90	2	2	1	Right
Headache 20	0	0	0	0	0	0	1	40	3	0	1	1	75	5	0	0	0	0	0	0	0	1	60	2	0	1	Right

Legend:

- Experienced reproduction and resolution of familiar head pain
- Reproduction and resolution of familiar head pain match with a dominant side
- Reproduction of familiar head pain with no resolution
- Reproduction and resolution of non-familiar head pain

NRS = Numerical Rating Scale pain intensity rating. Familiar = reproduction of familiar head pain (yes = 1, no = 0). Resolve = reduction in the produced head pain within 90 seconds of upper cervical joint mobilization by at least one point on the NRS scale (yes = 1, no = 0).

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Production and resolution of head pain never previously experienced occurred during the application of some techniques suggesting that stressing upper cervical levels with sustained mobilization may produce and resolve head pain. However, it should be considered that only reproduction and resolution of *usual* head pain with sustained upper cervical mobilization may be considered a distinguishing feature of CGH.

The description of head pain responses in this series of cases also highlights the potential for multisegmental dysfunction in the upper cervical spine as a contributor to CGH. Whilst findings largely matched the application of the unilateral technique to the dominant head pain side, two participants displayed reproduction and resolution of head pain on the non-dominant head pain side. A further six headache participants displayed reproduction and resolution of head pain with mobilization testing of both the left and right sides. This potentially indicates the involvement of bilateral cervical dysfunction. Although CGH has typically been described as a unilateral headache without alteration or side shift (5,26,27), the potential for bilaterality has long been recognized (5,26,28). In cases where this presentation occurs, it is usually considered to be a consequence of the duplication of the joint disorder on the contralateral side, sometimes referred to as "unilaterality on two sides" (28). The existence of more than one locus of pain in CGH has been previously noted in respect of diagnosis by anesthetic blockade where presentations of suspected bilateral CGH are addressed through the application of multiple blockades (4). The existence of this situation is entirely plausible when we consider that CGH may be a common pathway for a number of pain-generating disorders in the cervical spine area potentially involving numerous innervated tissues (5).

Seven individuals in this sample did not experience neck pain, instead presenting with other evidence of upper cervical dysfunction such as limited passive motion. This finding is consistent with the most recent International Headache Society criteria (ICHD-3) description for CGH which states that CGH is "usually, but not invariably accompanied by neck pain". The acknowledgment of the existence of CGH without associated neck pain represents the development in the accepted description of CGH over time as a presentation that usually begins as a pain in the neck or occiput (26,27) and possibly extends into the ipsilateral arm (28) to a condition precipitated by cervical spine dysfunction but not necessarily accompanied by pain experienced in the cervical region (19).

A nociplastic pain mechanism has been suggested to be present in some individuals with chronic CGH with hyperalgesia evident in upper cervical segments and the TCN (29-31). This mechanism, whereby pain is maintained by amplified processing or decreased inhibition at multiple levels within the central nervous system (31,32), may be present alongside nociceptive pain derived from upper cervical spine joint dysfunction (33,34). The failure of head pain to resolve within 90 seconds following provocation as observed in three individuals in our case series may indicate the presence of nociplastic pain in these individuals. As nociplastic pain may be associated with preceding prolonged nociceptive inputs from peripheral structures (35), it is likely that central sensitivity may be modulated by reducing the peripheral nociceptive input (36). Sustaining the mobilization manoeuvre may result in the diminution of pre-existing sensitization by lessening referred head pain leading to reduced head pain referral and resolution of the pain (13).

In interpreting the descriptions of head pain and its response to sustained mobilization described in this case series, it should be considered that this work represents a small sample of individuals seeking care for their headache disorder in a single private clinic. Hence, the generalizability of these descriptions to a larger population of CGH patients may be limited. Additionally, patients in this series were classified as having probable CGH due to the absence of confirmation of the diagnosis by diagnostic blockade. This raises an increased possibility of misdiagnosis compared to samples derived from interventional procedures.

Conclusion

Reproduction and resolution of headache with sustained upper cervical mobilization findings in the patients reported in this case series largely matched the dominant head pain side. The bilaterality of findings was evident in this clinical series. This case series also demonstrated that CGH can present without neck pain despite underlying cervical dysfunction.

Disclosures

Conflict of interest: The authors declare no conflict of interest.

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author contributions: PO: conceptualization, data curation, formal analysis, methodology, project administration, supervision, writing original draft, writing review and editing. DC: conceptualization, data curation, investigation, methodology, writing original draft, writing review and editing. LT: conceptualization, methodology supervision, writing original draft, writing review and editing.

Data availability statement: Data available on request

References

1. Bogduk N. The neck and headaches. Neurol Clin. 2004;22(1):151-171, vii. CrossRef PubMed
2. Bogduk N, Govind J. Cervicogenic headache: an assessment of the evidence on clinical diagnosis, invasive tests, and treatment. Lancet Neurol. 2009;8(10):959-968. CrossRef PubMed
3. Nilsson N. The prevalence of cervicogenic headache in a random population sample of 20-59 year olds. Spine. 1995;20(17):1884-1888. CrossRef PubMed
4. Antonaci F, Fredriksen TA, Sjaastad O. Cervicogenic headache: clinical presentation, diagnostic criteria, and differential diagnosis. Curr Pain Headache Rep. 2001;5(4):387-392. CrossRef PubMed
5. Sjaastad O, Fredriksen TA, Pfaffenrath V; The Cervicogenic Headache International Study Group. Cervicogenic headache: diagnostic criteria. Headache. 1998;38(6):442-445. CrossRef PubMed
6. Farmer PK, Snodgrass SJ, Buxton AJ, Rivett DA. An investigation of cervical spinal posture in cervicogenic headache. Phys Ther. 2015;95(2):212-222. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

7. Hall T, Robinson K. The flexion-rotation test and active cervical mobility--a comparative measurement study in cervicogenic headache. Man Ther. 2004;9(4):197-202. CrossRef PubMed
8. Jull G, Amiri M, Bullock-Saxton J, Darnell R, Lander C. Cervical musculoskeletal impairment in frequent intermittent headache. Part 1: subjects with single headaches. Cephalalgia. 2007;27(7):793-802. CrossRef PubMed
9. Zito G, Jull G, Story I. Clinical tests of musculoskeletal dysfunction in the diagnosis of cervicogenic headache. Man Ther. 2006;11(2):118-129. CrossRef PubMed
10. Watson DH, Drummond PD. Head pain referral during examination of the neck in migraine and tension-type headache. Headache. 2012;52(8):1226-1235. CrossRef PubMed
11. Cummins DC, Rivett DA, Thomas LC, Osmotherly PG. Reproduction and resolution of familiar head pain with upper cervical spine sustained joint mobilization may help identify cervicogenic headaches: a case-control study. J Man Manip Ther. 2023 Jun;31(3):198-205. PubMed
12. Schaible H-G, Grubb BD. Afferent and spinal mechanisms of joint pain. Pain. 1993;55(1):5-54. CrossRef PubMed
13. Watson DH, Drummond PD. Cervical referral of head pain in migraineurs: effects on the nociceptive blink reflex. Headache. 2014;54(6):1035-1045. CrossRef PubMed
14. Korr IM. Proprioceptors and somatic dysfunction. J Am Osteopath Assoc. 1975;74(7):638-650. PubMed
15. Pickar JG. Neurophysiological effects of spinal manipulation. Spine J. 2002;2(5):357-371. CrossRef PubMed
16. Malo-Urriés M, Tricás-Moreno JM, Estébanez-de-Miguel E, Hidalgo-García C, Carrasco-Uribarren A, Cabanillas-Barea S. Immediate effects of upper cervical translatoric mobilization on cervical mobility and pressure pain threshold in patients with cervicogenic headache: a randomized controlled trial. J Manipulative Physiol Ther. 2017;40(9):649-658. CrossRef PubMed
17. Núñez-Cabaleiro P, Leirós-Rodríguez R. Effectiveness of manual therapy in the treatment of cervicogenic headache: A systematic review. Headache. 2022;62(3):271-283. CrossRef PubMed
18. Zusman M. Spinal manipulative therapy: review of some proposed mechanisms, and a new hypothesis. Aust J Physiother. 1986;32(2):89-99. CrossRef PubMed
19. Headache Classification Committee of the International Headache Society (IHS). The International Classification of Headache Disorders, 3rd edition. Cephalalgia. 2018;38(1):1-211. CrossRef PubMed
20. Refshauge KM, Parry S, Shirley D, Larsen D, Rivett DA, Boland R. Professional responsibility in relation to cervical spine manipulation. Aust J Physiother. 2002;48(3):171-179. CrossRef PubMed
21. Niere K, Robinson P. Determination of manipulative physiotherapy treatment outcome in headache patients. Man Ther. 1997;2(4):199-205. PubMed PMID:11440533 CrossRef PubMed
22. Kosinski M, Bayliss MS, Bjorner JB, et al. A six-item short-form survey for measuring headache impact: the HIT-6. Qual Life Res. 2003;12(8):963-974. CrossRef PubMed
23. Niere K, Jerak A. Measurement of headache frequency, intensity and duration: comparison of patient report by questionnaire and headache diary. Physiother Res Int. 2004;9(4):149-156. CrossRef PubMed
24. Maitland G, Hengeveld E, Banks K, English K. Maitland's Vertebral Manipulation. Elsevier Butterworth Heinemann; 2005.
25. Finch E, Brooks D, Stratford PW. Physical rehabilitation outcome measures: a guide to enhance clinical decision making. 2nd ed. BC Decker Inc; 2002.
26. International Headache Society. 11. Headache or Facial Pain Associated with Disorder of Cranium, Neck, Eyes, Ears, Nose, Sinuses, Teeth, Mouth or Other Facial or Cranial Structures. Cephalalgia. 1988;8(7)(suppl):61-64.
27. Merskey H, Bogduk N. Classification of chronic pain. 2nd ed. ed: IASP Press; 1994. 57-8 p.
28. Sjaastad O, Fredriksen TA. Cervicogenic headache: criteria, classification and epidemiology. Clin Exp Rheumatol. 2000;18(2) (suppl 19):S3-S6. PubMed
29. Chua NH, van Suijlekom HA, Vissers KC, Arendt-Nielsen L, Wilder-Smith OH. Differences in sensory processing between chronic cervical zygapophysial joint pain patients with and without cervicogenic headache. Cephalalgia. 2011;31(8):953-963. CrossRef PubMed
30. Mingels S, Dankaerts W, van Etten L, Bruckers L, Granitzer M. Exploring multidimensional characteristics in cervicogenic headache: relations between pain processing, lifestyle, and psychosocial factors. Brain Behav. 2021;11(10):e2339. CrossRef PubMed
31. Fitzcharles M-A, Cohen SP, Clauw DJ, Littlejohn G, Usui C, Häuser W. Nociplastic pain: towards an understanding of prevalent pain conditions. Lancet. 2021;397(10289):2098-2110. CrossRef PubMed
32. Shraim MA, Massé-Alarie H, Hall LM, Hodges PW. Systematic review and synthesis of mechanism-based classification systems for pain experienced in the musculoskeletal system. Clin J Pain. 2020;36(10):793-812. CrossRef PubMed
33. Kosek E, Clauw D, Nijs J, et al. Chronic nociplastic pain affecting the musculoskeletal system: clinical criteria and grading system. Pain. 2021;162(11):2629-2634. CrossRef PubMed
34. Shraim MA, Sluka KA, Sterling M, et al. Features and methods to discriminate between mechanism-based categories of pain experienced in the musculoskeletal system: a Delphi expert consensus study. Pain. 2022;163(9):1812-1828. CrossRef PubMed
35. Abadiyan F, Hadadnezhad M, Khosrokiani Z, Letafatkar A, Akhshik H. Adding a smartphone app to global postural re-education to improve neck pain, posture, quality of life, and endurance in people with nonspecific neck pain: a randomized controlled trial. Trials. 2021;22(1):274. CrossRef PubMed
36. Fernández-de-Las-Peñas C, Courtney CA. Clinical reasoning for manual therapy management of tension type and cervicogenic headache. J Man Manip Ther. 2014;22(1):44-50. CrossRef PubMed

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Arch Physioter 2025; 15: 165-173
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3381
REVIEW

Evaluation and synthesis of physiotherapy protocols for Femoroacetabular Impingement Syndrome (FAIS): a scoping review

Céline Berguerand, Noemie Rossi, Rahel Caliesch

School of Health and Sciences, HES-SO Valais-Wallis - Switzerland

ABSTRACT

Introduction: The physiotherapeutic treatment of femoroacetabular impingement syndrome (FAIS) is a topic of growing interest in the literature. The aim of this study is to present all of the treatment modalities used in scientific research in order to analyze the extent to which the protocols are explicit.
Method: This is a scoping review. The literature search was performed using the Cochrane, Embase and PubMed databases. The data was collected in various tables and the protocols were assessed using the Template for Intervention Description and Replication (TIDieR) and the Consensus on Exercise Reporting Template (CERT) tools.
Results: Twenty-four studies were selected and 30 protocols were analyzed. The most frequently reported treatment modalities were strengthening (n = 25), manual therapy (n = 22) and stretching (n = 21). The average total score was 47% for studies evaluated by the TIDieR checklist and 40% for studies evaluated by the CERT checklist.
Discussion: The treatment modalities identified are similar to those used for other musculoskeletal conditions. Specific treatments were found and were generally consistent with the clinical characteristics of FAIS. The lack of precision in the reporting of interventions compromises their clinical use. The same lack of detail is noted for other physiotherapeutic interventions in the musculoskeletal field.
Conclusion: Numerous treatment methods are presented in the literature. However, the protocols lack in explicitness and the use of the TIDieR and CERT evaluation grids is not widespread.

Keywords: Femoroacetabular Impingement, Exercises, Physical Therapy, Intervention Protocol, TIDieR, CERT

What is already known about this topic?

- *A lot of treatment modalities are reported in the literature for the treatment of FAIS. The study protocols are described with varying degrees of precision, for example, the number of exercises prescribed, which compromises the evidence-based practice.*

What does this study add?

- *This study gives an overview of the different modalities for the treatment of FAIS, from manual therapy to physical exercises. It also evaluates the explicitness of the protocols outlined in selected studies by the use of appropriate tools.*

Received: November 12, 2024
Accepted: June 1, 2025
Published online: June 24, 2025

This article includes supplementary material

Clinical Trial Protocol number:
DOI 10.17605/OSF.IO/GCQJX

Corresponding author:
Céline Berguerand
email: berguerandceline@gmail.com

Introduction

FAIS is defined as a premature and symptomatic contact between the acetabulum and the proximal part of the femur (1). It can provoke severe pain, often impacting the participation in sports and activities of daily life (2). When poorly managed, FAIS can cause labral and cartilage lesions, progressing to osteoarthritis (3). FAIS is a concept recently described in the literature and is present in 40% of hip pathologies (2). Unfortunately, the current literature focuses mainly on more advanced degenerative hip diseases such as osteoarthritis

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

and there are few meta-analyses on the physiotherapeutic rehabilitation of FAIS.

Physiotherapists and doctors are often unfamiliar with this condition. It is difficult to diagnose and it is not always clear how this type of pathology should be managed. Whether for conservative treatment or post-operative rehabilitation, the treatment modalities used in the literature can vary, from manual therapy to strengthening exercises. Understanding the best course of action can be difficult.

Study protocols are sometimes difficult to find and can lack precise information about the content of their interventions. For example, some important treatment parameters such as the type of exercise prescribed, the total number of sessions, or the number of sets per exercise are often poorly described. It is important that the intervention protocols are well described in order to understand which treatment was used and which obtained the best results. The detailed description of interventions also enables the transfer of results to clinical practice.

Therefore, the goal of this study is to investigate the physiotherapeutic protocols and exercises used for FAIS and to assess their level of explicitness.

Method

A scoping review was performed according to the guidelines for scoping reviews in Chapter 11 of the JBI Manual for Evidence Synthesis (4). The Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) checklist published by Tricco et al. (5) was used to structure the entire document. The scoping review protocol was registered in osf.io. (DOI 10.17605/OSF.IO/GCQJX)

Search strategy

A web-based literature search was conducted on the 22nd of November 2023 with no date restriction, on the following databases: Cochrane, Embase, and PubMed. Additional records were not identified through the grey literature, but the reference lists of relevant studies and meta-analyses on the topic were checked.

Study selection criteria

Studies were included if their participants were aged 16 years old or older, with a diagnosis of FAIS, with or without labral injuries. A rehabilitation protocol must have been used for preoperative, postoperative, or conservative physiotherapeutic treatment. Studies were excluded if the rehabilitation protocol was not tested on the participants, and if the participants had other pathologies, such as arthritis, cysts, or dysplasia. Studies with mixed pathologies, meta-analyses, and systematic reviews were excluded.

Study selection process

The search strategy was created with different terms linked with FAIS. Details can be seen in Appendix 1 (Search strategy).

Two independent reviewers screened all titles and abstracts based on the selection criteria. A second screening was made by the examination of full texts. If two articles used the same rehabilitation protocol, the most recent article or the one with the better explanation was selected.

Data extraction

Data were collected using three separate Microsoft Excel sheets: study characteristics (author, publication year, publication type,…), protocol characteristics (number of treatment weeks, series, repetitions,…), and protocol content (treatment modalities).

To ensure consistency in data extraction, two authors independently extracted data from three protocols. A step-by-step data extraction guide was then developed based on this process. The remaining protocols were equally divided, with half assessed by author CB and half by author NR.

The explicitness of the treatment protocols was assessed using the TIDieR checklist (6) and the CERT (7). The TIDieR checklist is based on 12 items with a total possible score of 12 and the CERT checklist is based on 16 items with a total possible score of 19. These two checklists can be used in 3 different cases: for authors to report their interventions, in systematic reviews to evaluate the reporting of exercises, and for readers to better understand how the therapy has been or can be used. They represent the most recent and specific framework for assessing rehabilitation protocols. The TIDieR checklist is an extension of the CONSORT statement SOURCE and the CERT checklist is based on the EQUATOR Network (8). The TIDieR checklist is designed for the reporting of any type of intervention in scientific research. The CERT checklist is more specific for the reporting of exercise programs across all evaluative study designs for exercise research (7).

Statistical analysis

The scores of the TIDieR and CERT checklists (total score and score per item) were calculated with an Excel formula. The “NA” was used for items not reported or not described by the authors.

Results

Study selection

Twenty-five studies were chosen after the whole selection process. Details are presented in Figure 1: Flow chart.

Study characteristics

Following the *Ascension* Classification (9) of the level of evidence of study designs, eight case studies (10,10–17) (level of evidence = 5), six case series (11,18–22) (level of evidence = 4), two cohort studies (23,24) (level of evidence = 2), four pilot RCTs (17,25–27) (level of evidence = 1) and five RCTs (28–32) (level of evidence = 1) were selected.

Six protocols (24,28,30,32,33) referred to protocols from another article. Only the best-presented and/or the most complete versions were selected for this review.

In total, 31 protocols were analyzed since six articles (22,25,27,30,31,34) used two different protocols. Forty-eight percent were postoperative (n = 15) (12,16–19,

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

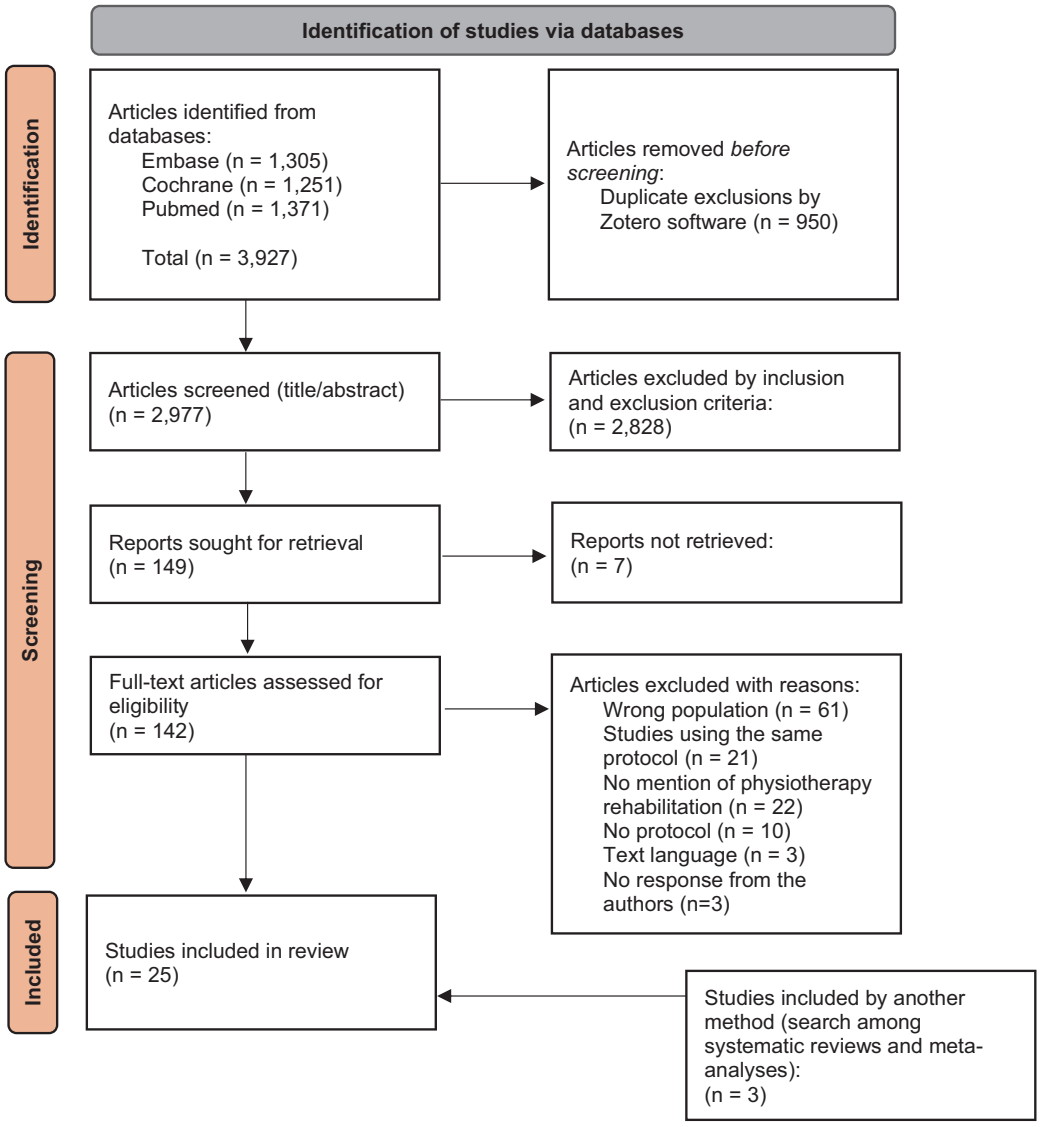


FIGURE 1 - Flow chart of the study selection process.

21,23–25,30–35), 45% were conservative (n = 14) (10,11,13–15,17,20,22,27,28,30,34) and 6% were preoperative (n = 2) (25,26) protocols. Details can be found in appendix 2 (Study characteristics).

Protocol characteristics

Twenty-six protocols (84 %) (10,12–21,24–28,30,31,33–35) mentioned the duration of treatment, which was on average (± sd) 17,4 weeks (± 21.85). Fewer protocols (42 %) (13,14,16–19,26,27,30,32–34) detailed the number of treatment sessions which was on average 21.64 (± 16.6). Twenty-four protocols stated the number of supervised sessions, with an average of 11.77 supervised sessions (± 6.78). Only 13 protocols (17,18,25–28,30,31,33) out of the 31 mentioned an adhesion assessment method. The number of series per exercise was presented in 58 % (10,12–18,26,27,30,31,33,34) of the protocols and the number of repetitions in 65 % (10,12–18,25–27,30,31,33,34). A home exercise program was provided in 65 % (10,12,13,15,17,18,20,24,26–28,30,31,33–35) of the protocols and the material used was reported in 48 % (14,15,23–28,30,31,33,34) of the case. Eleven protocols (18,25–28,30–33) tailored their program to participants' specific symptoms. This does not include the study cases, for which a tailored program was the norm. Progression criteria were indicated in 20 protocols (65 %) (10,12,14,16,18–20,22,24,26–28,30–34) and the goal of the treatment was indicated in 21 protocols (68 %)(10,12–15,17–20,24–26,28,30–35). Details can be found in appendix 3 (Protocol characteristics).

Content of the rehabilitation protocols

Active and passive treatment modalities were used as shown in Figure 2. Details about the content of the protocol can be found in appendix 4 (Content of the rehabilitation protocols).

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

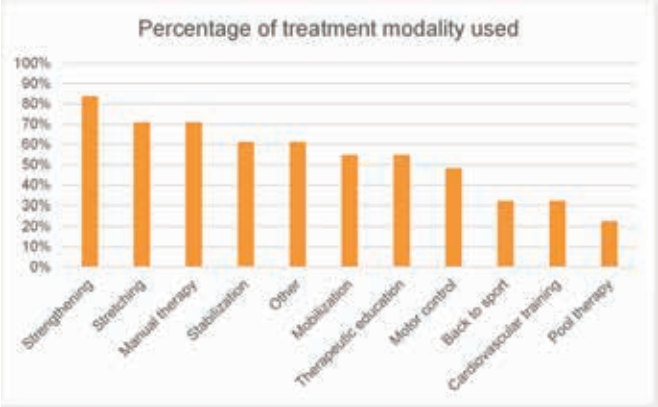


FIGURE 2 - Percentage of use of each treatment modality.

Strengthening was the most frequently used treatment modality (n = 26) (10–19,21,22,24–28,30–35), with emphasis on different muscular groups as presented in Figure 3.

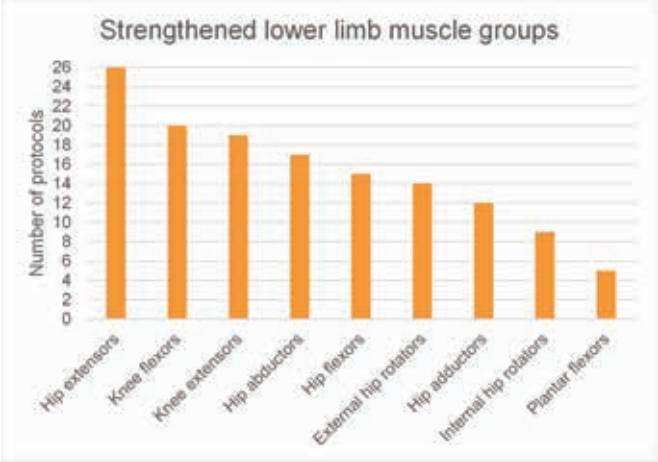


FIGURE 3 - Lower limb muscle groups most cited in the 26 strengthening protocols.

Other active modalities used were mobilization (n = 17) (12,17,19,20,22–25,27,28,30–33,35) in all three directions of movement, stabilization (n = 18) (10–12,14,15,18, 19,21,22,24,25,27,28,30–35) of different body parts such as the trunk or the hips or motor control, which was global or joint-specific. Some studies also proposed exercises in a swimming pool. Passive modalities were also proposed, such as stretching (n = 22) (12–15,18–20,22,24,25,27,28,30–35) in different directions of movement (Fig. 4) and manual therapy (n = 22) (10,12,14–19,22,24,25,27,28,30–35) on soft tissue or on joints. Patient education was often mentioned (n = 17) (11,12,16–18,20,22,25,27,28,31–34) including activity modification or postural advice.

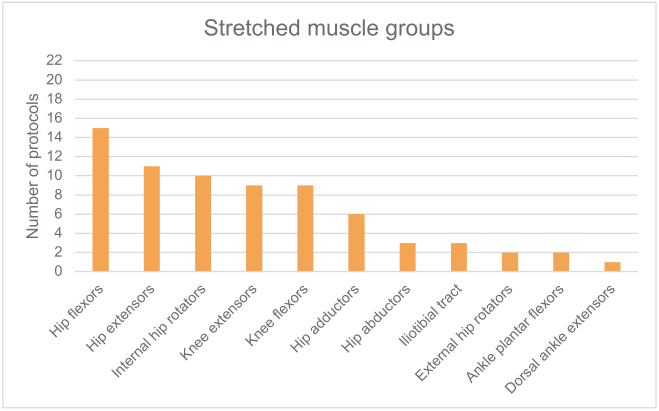


FIGURE 4 - Lower limb muscle groups cited in 22 stretching protocols.

Protocol explicitness

Only 2 protocols used the TIDieR checklist to rate their interventions. Figure 5 shows the score for each protocol and Figure 6 presents the average per item. The average score was 5.58 (± 3.24) and the median was 6.

None of the protocols selected used the CERT checklist to evaluate their intervention. Figure 7 shows the score of each protocol and Figure 8 presents the average per item. The average score was 7.35 (± 4.31) and the median was 8.

Discussion

Synthesis of results

Content of the protocols

The first goal of this study was to investigate which treatment modalities were used for FAIS rehabilitation. It was discovered that the most frequently trained direction of movement was hip extension even though it has been shown that patients who suffer from FAIS have a lack of strength in all movement directions (27,36–38). Strengthening is important but motor control also has a place in the rehabilitation of FAIS. Only three protocols (12,16,28) mentioned motor control exercises.

Trunk motor control was infrequently trained despite evidence showing that a lack of flexion in the lumbar spine can lead to compensation in the hips with increased flexion during sit-to-stand, which can lead to the development of FAIS (39).

Stabilization should also be trained by patients with FAIS. Freke et al. (37) showed that before a hip arthroscopy for FAIS, patients had a significant lack of unipedal balance compared to the control group. An improvement was observed after a three- and six-month rehabilitation, however, the results were still lower than in the control group. Global stabilization such as unipedal balance exercises should therefore also be performed during the final phase of rehabilitation and as part of conservative treatment.

Only a few articles mentioned manual therapy of the associated joints such as the knees, the sacroiliac joint, or

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

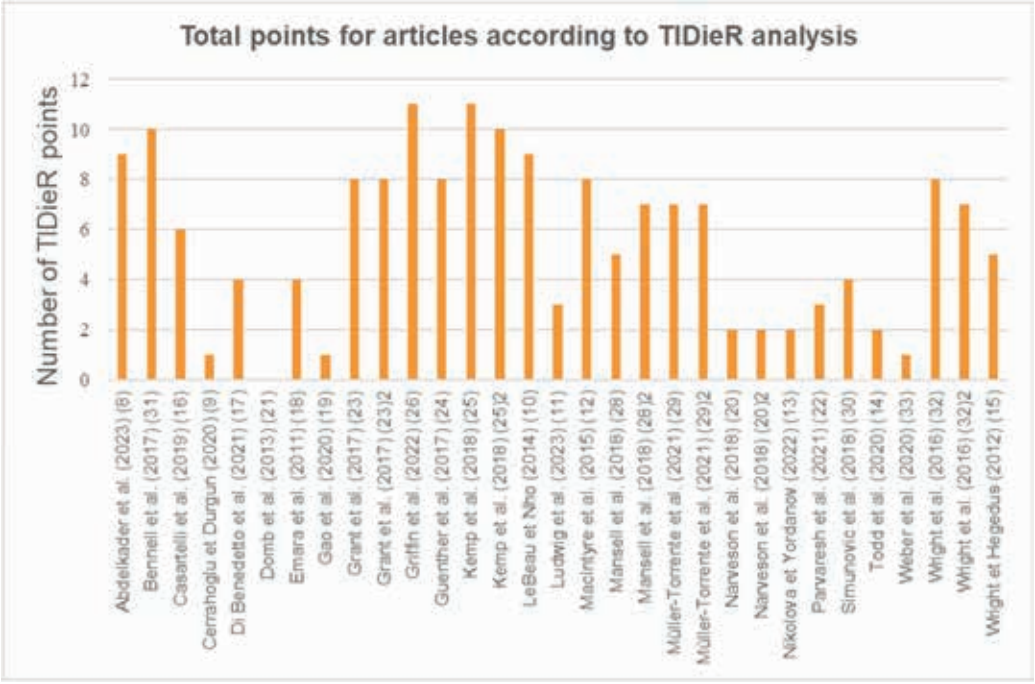


FIGURE 5 - Number of total points for articles following TIDieR evaluation.

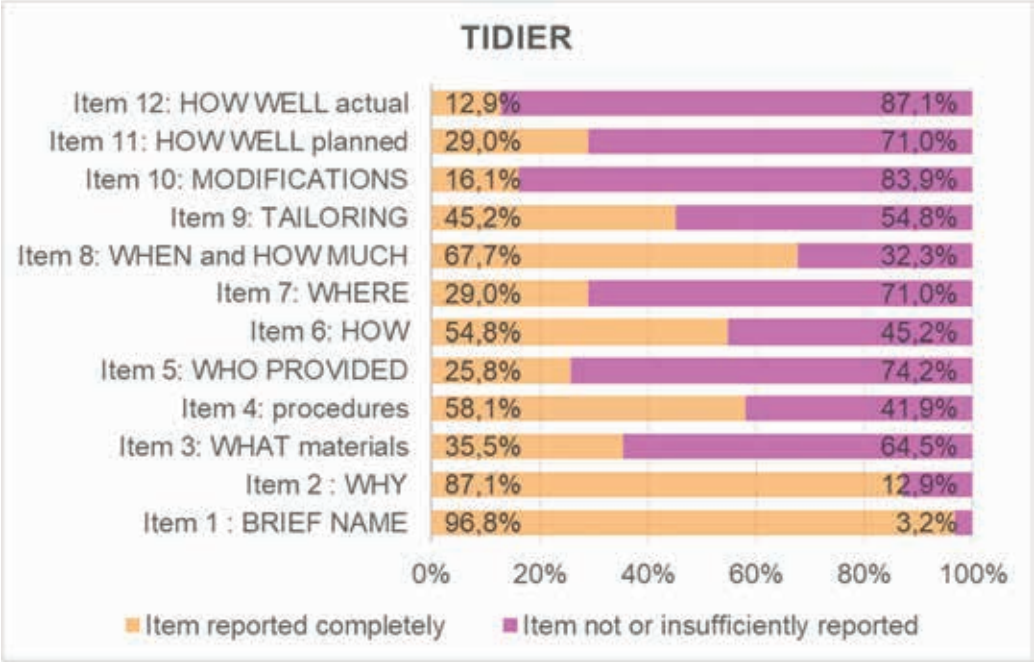


FIGURE 6 - Percentage of TIDieR items completely, not, or insufficiently reported according to all selected protocols.

the spine. Manual therapy of these joints can also influence the hip joint since their biomedical function is interdependent (40).

Manual therapy of soft tissues of the hip was often described in the protocols since it is recommended by Kuhns et al. (41) in the early stage of treatment post-surgery or for chronic hip pain. Twenty minutes per session is recommended to reduce stiffness. Manual therapy can improve the range of motion, and prevent global irritation, tendinopathy, or scar adherences (41,42).

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

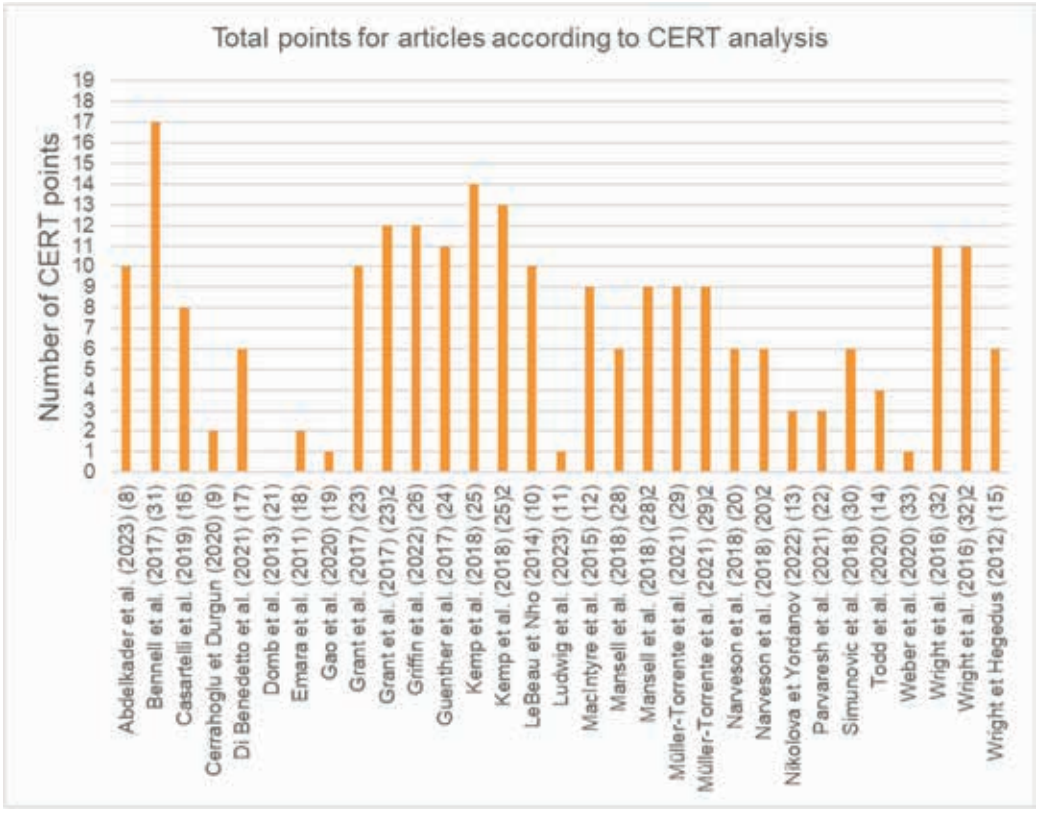


FIGURE 7 - Number of total points for articles following CERT evaluation.

Stretching of muscles and the capsule was often mentioned but there was little mention of pain tolerance during the stretching, even though it is suspected to exacerbate pain in patients with FAIS (43). However, psoas stretching associated with abdominal drawing-in can significantly increase the active range of motion with fewer compensations (44).

Protocol explicitness

The second goal of this study was to evaluate the intervention reporting. The results show that protocols are not explicit enough. Items concerning the name and the goal of the intervention were better described in the TIDieR checklist. Adherence is one important factor in the effectiveness of a treatment. Despite this, item 12 (*How well, actual*) and item 10 (*Modification*) produced very low scores. An explanation for this could be that these items are usually evaluated during or after an intervention and cannot be pre-planned. Similar conclusions can be drawn for the CERT checklist. Items related to protocol details were better described than those concerning how the intervention was delivered.

The variability in protocol duration, number of sessions, and adherence assessment affects the validity and the interpretation of study results. For instance, if a treatment modality is deemed statistically effective or superior to another, it is crucial to understand the specifics: how it was applied, how often, and to what extent. This clarity is essential for accurate interpretation and in order to replicate the results in clinical practice.

Future perspectives

Content of the protocols

Guidelines for sub-acromial conflict include similar treatment modalities as those presented here, except for the inclusion of electrotherapy, laser therapy, and shockwave therapy (45). It is interesting to compare the results of this scoping review with treatment recommendations for postoperative rehabilitation for FAIS (41). Similar treatment modalities were recommended. However, motor control exercises and pelvic stabilization were not included in the guidelines. Future research should investigate whether there is a lack of motor control or pelvic stabilization in the FAIS population.

Protocol explicitness

Protocols are not described in enough detail and the TIDieR and CERT checklists are too infrequently applied in studies on FAIS. These results are analogous to those of other studies of musculoskeletal conditions such as ACL injuries (46), back pain (47), cervical pain (48), rotator cuff pathologies (49), or Achilles tendon injuries (50). These studies reported between 5 and 8 out of 19 points for the CERT checklist, and half of the possible points on the TIDieR checklist.

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

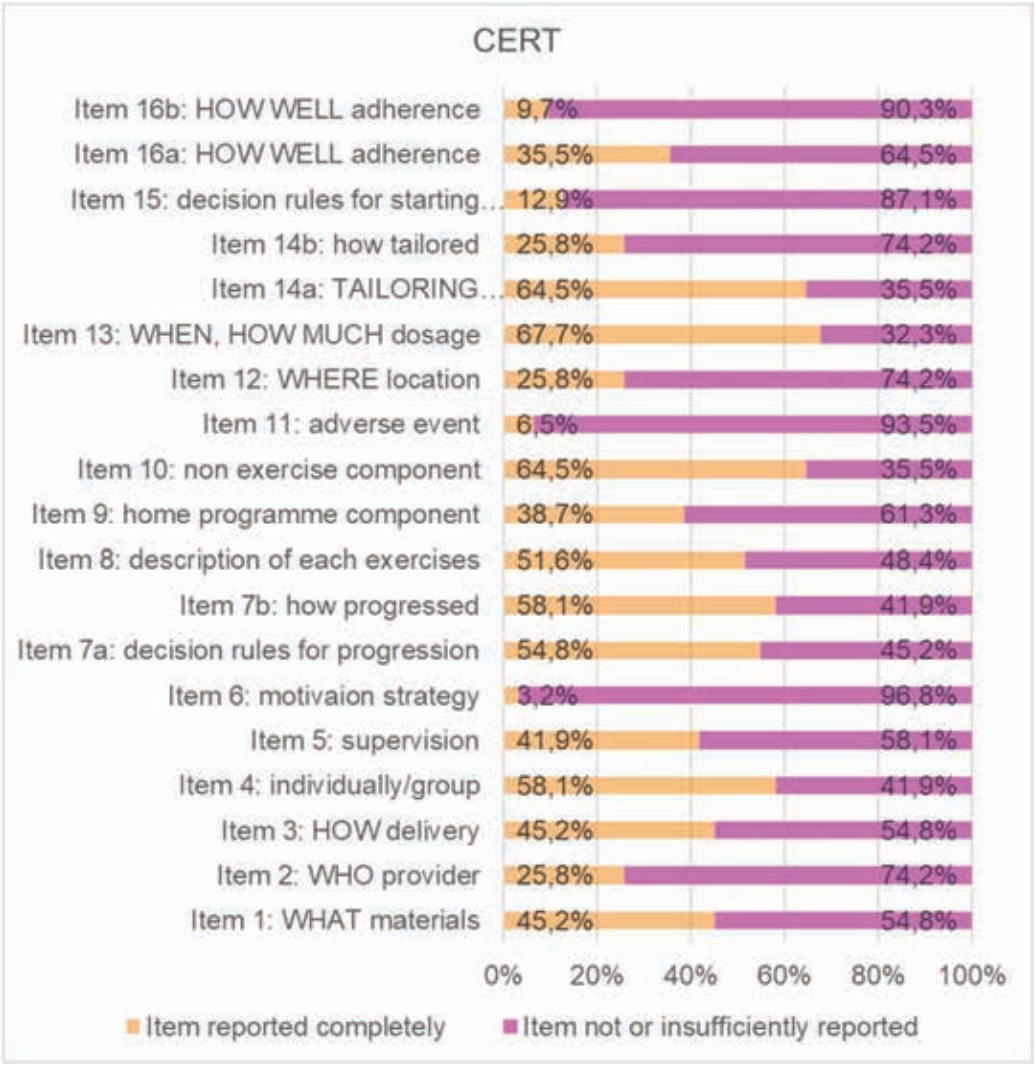


FIGURE 8 - Percentage of CERT items completely, not, or insufficiently reported according to all selected protocols.

A paper from Esterberger et al. (51) about hip pain in general showed similar results for the CERT checklist: 13 out of 19 items had similar scores (less than 25% difference) Only exercise therapy treatments were included and therefore only the CERT checklist was used. Our study can be considered as an update of the study from Esterberger et al. This study is more complete since other treatment modalities were included and assessed with the TIDieR checklist. The TIDieR and the CERT checklists are relatively recent frameworks, having been introduced in 2016, and thus are not yet widely recognized or used by authors. To improve intervention reporting, journals should require their implementation.

Strengths and limits

This article summarises the evidence in the field of FAIS and provides a large overview of the available treatment modalities and presents them in a useful table. Two assessment grids were selected to evaluate all of the treatment modalities, not only exercise-based interventions. This enhances the validity of the results. While these tools are valuable for assessing intervention reporting, their ratings are not entirely objective, despite existing guidelines for their use. To increase precision, a range of grades for each item could be useful (for example from 0 to 4, 0 = not described at all, 1 = barely described, 2 = described with some missing information, 3 = almost perfectly described, and 4 = perfectly described). Average and median values were used but no interpretation of cut-offs of CERT and TIDieR results exists. The creation of such an evaluation could be useful for future research.

Conclusion

This paper aims to be a reference for clinicians for the treatment of FAIS. It highlights the lack of evidence in reporting intervention in the scientific literature. The TIDieR and the CERT checklists could be revised to improve precision but are already useful tools to increase clarity in the reporting of interventions in future research.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Acknowledgment

Simone Gafner and Sophie Carrard for the supervision.

Laura Marguerettaz and Chris Brown for the substantive translation.

This is the final version of this article as stated in CrossRef

Disclosures

Conflict of interest: The authors declare no conflict of interest.

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Byrd JWT. Femoroacetabular impingement in athletes: current concepts. Am J Sports Med. 2014;42(3):737-751. CrossRef PubMed
2. Trigg SD, Schroeder JD, Hulsopple C. Femoroacetabular Impingement Syndrome. Curr Sports Med Rep. 2020;19(9): 360-366. CrossRef PubMed
3. Charles T, Bloemers N, Deneubourg M, Schuind F, Jayankura M. Femoro-acetabular conflict : A cause of osteoarthritis of the young subject. Rev Med Brux. 2019;40(4):318-323.
4. Peters MD, Godfrey CM, Mclnerney P, Munn Z, Tricco AC, Khalil H. Scoping reviews (2020 version). In: JBI Manual for Evidence Synthesis. JBI; 2020:chap 11, CrossRef.
5. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and Explanation. Ann Intern Med. 2018;169(7):467-473. CrossRef PubMed
6. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014;348(mar07 3):g1687. CrossRef PubMed
7. Slade SC, Dionne CE, Underwood M, et al. Consensus on Exercise Reporting Template (CERT): modified Delphi Study. Phys Ther. 2016;96(10):1514-1524. CrossRef PubMed
8. Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340(mar23 1):c332-c332. CrossRef
9. Matucheski M. LibGuides: Evidence-Based Practice: Levels of Evidence and Study Designs. 2023. Online (Accessed November 2024)
10. Abdelkader N, Leonardelli C, Howitt S. Olympic fencer with femoroacetabular impingement syndrome uses active rehabilitation to avoid surgery: A case report. J Bodywork Mov Ther. 2023;33: 1-7. CrossRef
11. Cerrahoğlu L, Durgun ÖE. Bilateral femoroacetabular impingement syndrome: A case report. Meandros Med Dent J. 2020;21(3):266-269. CrossRef PubMed
12. LeBeau RT, Nho SJ. The use of manual therapy post-hip arthroscopy when an exercise-based therapy approach has failed: a case report. J Orthop Sports Phys Ther. 2014;44(9):712-721. CrossRef PubMed
13. Ludwig O, Schneider G, Kelm J. Improvement of Groin Pain in a Football Player with Femoroacetabular Impingement via a Correction of the Pelvic Position-A Case Report. J Clin Med. 2023;12(23):7443. CrossRef PubMed
14. MacIntyre K, Gomes B, MacKenzie S, D'Angelo K. Conservative management of an elite ice hockey goaltender with femoroacetabular impingement (FAI): a case report. J Can Chiropr Assoc. 2015;59(4):398-409. PubMed
15. Nikolova A, Yordanov YP. Conservative Physiotherapy Treatment of Femoroacetabular Impingement. Acta Med Bulg. 2022;49(3):50-53. CrossRef
16. Todd C, Karlsson J, Baranto A. Resolving anterior hip pain in a young male footballer following arthroscopic surgery for Femoroacetabular Impingement Syndrome: A case report. J Bodyw Mov Ther. 2020;24(1):63-68. CrossRef PubMed
17. Wright AA, Hegedus EJ. Augmented home exercise program for a 37-year-old female with a clinical presentation of femoroacetabular impingement. Man Ther. 2012;17(4):358-363. CrossRef PubMed
18. Casartelli NC, Bizzini M, Maffiuletti NA, et al. Exercise Therapy for the Management of Femoroacetabular Impingement Syndrome: Preliminary Results of Clinical Responsiveness. Arthritis Care Res (Hoboken). 2019;71(8):1074-1083. CrossRef PubMed
19. Di Benedetto P, Vidi D, Buttironi MM, Mancuso F, Ricciarelli M, Causero A. Outcomes in arthroscopic surgery and proposal rehabilitative treatment in femoral acetabular impingement syndrome. Acta Biomed. 2022;92(S3):e2021575. CrossRef PubMed
20. Emara K, Samir W, Motasem H, Ghafar KAE. Conservative treatment for mild femoroacetabular impingement. J Orthop Surg (Hong Kong). 2011;19(1):41-45. CrossRef PubMed
21. Gao F, Zhang B, Hu B, et al. Outcomes of Hip Arthroscopy for Femoroacetabular Impingement in Chinese Patients Aged 50 Years or Older. Orthop Surg. 2020;12(3):843-851. CrossRef PubMed
22. Narveson JR, Haberl MD, Nathan Vannatta C, Rhon DI. Conservative treatment continuum for managing Femoroacetabular Impingement Syndrome and Acetabular Labral Tears in surgical candidates: a case series. Int J Sports Phys Ther. 2018;13(6):1032-1048. CrossRef PubMed
23. Domb BG, Stake CE, Botser IB, Jackson TJ. Surgical dislocation of the hip versus arthroscopic treatment of femoroacetabular impingement: a prospective matched-pair study with average 2-year follow-up. Arthroscopy. 2013;29(9):1506-1513.CrossRef PubMed
24. Parvaresh KC, Wichman DM, Alter TD, Clapp IM, Nho SJ. High rate of return to tennis after hip arthroscopy for patients with femoroacetabular impingement syndrome. Phys Ther Sport. 2021;51:45-49. CrossRef PubMed
25. Grant LF, Cooper DJ, Conroy JL. The HAPI 'Hip Arthroscopy Prehabilitation Intervention' study: does pre-habilitation affect outcomes in patients undergoing hip arthroscopy for femoroacetabular impingement? J Hip Preserv Surg. 2017;4:85. CrossRef
26. Guenther JR, Cochrane CK, Crossley KM, Gilbart MK, Hunt MA. A Pre-Operative Exercise Intervention Can Be Safely Delivered to People with Femoroacetabular Impingement and Improve Clinical and Biomechanical Outcomes. Physiother Can. 2017; 69(3):204-211. CrossRef PubMed
27. Kemp JL, Coburn SL, Jones DM, Crossley KM. The Physiotherapy for Femoroacetabular Impingement Rehabilitation STudy (physioFIRST): A Pilot Randomized Controlled Trial. J Orthop Sports Phys Ther. 2018;48(4):307-315. CrossRef PubMed
28. Griffin DR, Dickenson EJ, Achana F, et al. Arthroscopic hip surgery compared with personalised hip therapy in people over 16 years old with femoroacetabular impingement syndrome: UK FASHIoN RCT. Health Technol Assess. 2022;26(16):1-236. CrossRef PubMed
29. Griffin DR, Dickenson EJ, Wall PDH, et al; FASHIoN Study Group. Protocol for a multicentre, parallel-arm, 12-month, randomised, controlled trial of arthroscopic surgery versus conservative care for femoroacetabular impingement syndrome (FASHIoN). BMJ Open. 2016;6(8):e012453. CrossRef PubMed

30. Mansell NS, Rhon DI, Meyer J, Slevin JM, Marchant BG. Arthroscopic Surgery or Physical Therapy for Patients With Femoroacetabular Impingement Syndrome: A Randomized Controlled Trial With 2-Year Follow-up. Am J Sports Med. 2018;46(6):1306-1314. CrossRef PubMed
31. Müller-Torrente A, Puig-Torregrosa J, Montero-Navarro S, et al. Benefits of a specific and supervised rehabilitation program in femoroacetabular impingement patients undergoing hip arthroscopy: A randomized control trial. J Clin Med. 2021;10(14):3125. CrossRef PubMed
32. Simunovic N, Heels-Ansdell D, Thabane L, Ayeni OR; FIRST Investigators. Femoroacetabular Impingement Randomised controlled Trial (FIRST) - a multi-centre randomized controlled trial comparing arthroscopic lavage and arthroscopic osteochondroplasty on patient important outcomes and quality of life in the treatment of young adult (18-50 years) femoroacetabular impingement: a statistical analysis plan. Trials. 2018;19(1):588. CrossRef PubMed
33. Bennell KL, Spiers L, Takla A, et al. Efficacy of adding a physiotherapy rehabilitation programme to arthroscopic management of femoroacetabular impingement syndrome: a randomised controlled trial (FAIR). BMJ Open. 2017;7(6):e014658. CrossRef PubMed
34. Wright AA, Hegedus EJ, Taylor JB, Dischiavi SL, Stubbs AJ. Non-operative management of femoroacetabular impingement: A prospective, randomized controlled clinical trial pilot study. J Sci Med Sport. 2016;19(9):716-721. CrossRef PubMed
35. Weber AE, Nakata H, Mayer EN, et al. Return to Sport After Hip Arthroscopy for Femoroacetabular Impingement Syndrome in NCAA Division I Athletes: Experience at a Single Institution. Orthop J Sports Med. 2020;8(5):2325967120918383. CrossRef PubMed
36. Diamond WT. Hinman R, et al. Isometric and isokinetic hip strength in symptomatic femoroacetabular impingement. J Sci Med Sport. 2015;19:e69-e70. CrossRef
37. Freke M, King M, Crossley K, Sims K, Semciw A. Acute and Subacute Changes in Dynamic Postural Control After Hip Arthroscopy and Postoperative Rehabilitation. J Athl Train. 2022;57(5):494-501. CrossRef PubMed
38. Mayne E, Memarzadeh A, Raut P, Arora A, Khanduja V. Measuring hip muscle strength in patients with femoroacetabular impingement and other hip pathologies: A systematic review. Bone Joint Res. 2017;6(1):66-72. CrossRef PubMed
39. Fader RR, Tao MA, Gaudiani MA, et al. The role of lumbar lordosis and pelvic sagittal balance in femoroacetabular impingement. Bone Joint J. 2018;100-B(10):1275-1279. CrossRef PubMed
40. Sueki DG, Cleland JA, Wainner RS. A regional interdependence model of musculoskeletal dysfunction: research, mechanisms, and clinical implications. J Man Manip Ther. 2013;21(2):90-102. CrossRef PubMed
41. Kuhns BD, Weber AE, Batko B, Nho SJ, Stegemann C. A four-phase physical therapy regimen for returning athletes to sport following hip arthroscopy for femoroacetabular impingement with routine capsular closure. Int J Sports Phys Ther. 2017;12(4):683-696. PubMed
42. Bialosky JE, Bishop MD, Price DD, Robinson ME, George SZ. The mechanisms of manual therapy in the treatment of musculoskeletal pain: a comprehensive model. Man Ther. 2009;14(5):531-538. CrossRef PubMed
43. Emary P. Femoroacetabular impingement syndrome: a narrative review for the chiropractor. J Can Chiropr Assoc. 2010;54(3):164-176. PubMed
44. Park KY, Jeon IC, Hwang UJ, Jung SH, Ha SM, Yim SY. Effects of abdominal drawing-in maneuver with or without prior iliopsoas stretching on gluteus maximus activity during prone hip extension. J Back Musculoskelet Rehabil. 2022;35(2):331-339. CrossRef PubMed
45. Pieters L, Lewis J, Kuppens K, et al. An Update of Systematic Reviews Examining the Effectiveness of Conservative Physical Therapy Interventions for Subacromial Shoulder Pain. J Orthop Sports Phys Ther. 2020;50(3):131-141. CrossRef PubMed
46. Vlok A, van Dyk N, Coetzee D, Grindem H. Exercise Descriptors That Determine Muscle Strength Gains Are Missing From Reported Anterior Cruciate Ligament Reconstruction Rehabilitation Programs: A Scoping Review of 117 Exercises in 41 Studies. J Orthop Sports Phys Ther. 2022;52(2):100-112. CrossRef PubMed
47. Davidson SRE, Kamper SJ, Haskins R, et al. Exercise interventions for low back pain are poorly reported: a systematic review. J Clin Epidemiol. 2021;139:279-286. CrossRef PubMed
48. Grandeo J, Favaro L, Rhon DI, Young JL. Exercise therapy reporting in clinical trials for chronic neck pain: A systematic review. Musculoskeletal Care. 2022;20(4):796-811. CrossRef PubMed
49. Major DH, Røe Y, Grotle M, et al. Content reporting of exercise interventions in rotator cuff disease trials: results from application of the Consensus on Exercise Reporting Template (CERT). BMJ Open Sport Exerc Med. 2019;5(1):e000656. CrossRef PubMed
50. Christensen M, Zellers JA, Kjær IL, Silbernagel KG, Rathleff MS. Resistance Exercises in Early Functional Rehabilitation for Achilles Tendon Ruptures Are Poorly Described: A Scoping Review. J Orthop Sports Phys Ther. 2020;50(12):681-690. CrossRef PubMed
51. Estberger A, Kemp JL, Thorborg K, Pålsson A, Ageberg E. Are Exercise Therapy Protocols For The Treatment of Hip-Related Pain Adequately Described? A Systematic Review of Intervention Descriptions. Int J Sports Phys Ther. 2023;18(1):38-54. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 174-183
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3254
ORIGINAL RESEARCH ARTICLE

Motor and cognitive changes across 2 years: a longitudinal study in early Multiple Sclerosis

Elisa Gervasoni[1], Claudio Solaro[2], Denise Anastasi[3], Rachele Di Giovanni[4], Marco Rovaris[1], Giampaolo Brichetto[5,6], Paolo Confalonieri[7], Andrea Tacchino[5], Ilaria Carpinella[1], Davide Cattaneo[1,8]

[1]IRCSS Fondazione Don Carlo Gnocchi, Milano - Italy
[2]Neurology Unit Galliera Hospital, Genova - Italy
[3]Department of Biomedical Sciences, University of Sassari, Sassari - Italy
[4]Rehabilitation Department, CRRF Mons L Novarese, Moncrivello (VC) - Italy
[5]Scientific Research Area, Italian Multiple Sclerosis Foundation (FISM), Genoa - Italy
[6]AISM Rehabilitation Servicie of Genoa, Genoa - Italy
[7]Multiple Sclerosis Center, IRCCS Foundation "Carlo Besta" Neurological Institute, Milan - Italy
[8]Department of Physiopathology and Transplants, University of Milan, Milan - Italy

ABSTRACT

Background: Functional disorders appear early in the course of the disease and develop over time, impacting participation and quality of life. However, little is known about functional deterioration assessed by clinical and instrumented tools in longitudinal studies.

Methods: We included 63 PwMS with Expanded Disability Status Scale (EDSS) ≤2.5 points, disease duration ≤5 years, and age (Mean ± Standard Deviation) 38.7 ± 10.7 years. Participants were assessed at baseline and after 2 years with clinical and instrumented evaluations. Data on disability and functional disorders were collected using EDSS, Six-Minute Walk Test (6MWT), Multiple Sclerosis Walking Scale-12 (MSWS-12), Fatigue Severity Scale (FSS), and Symbol-digit modality test (SDMT), while instrumented data measuring complexity and intensity of balance disorders and gait symmetry-regularity-instability were extracted from wearable devices.

Results: Clinical scales (EDSS, 6MWT, FSS, MSWS-12, and SDMT) did not show a statistically significant deterioration when baseline and 2-year follow-up were compared: EDSS (Median and min-max) from 2 (0-2.5) to 1.5 (0-4.5)points, 6MWT from 566.2 ± 80.4 to 573.9 ± 94.7m, FSS from 2.89 (0.89-7) to 2.67(1-7)points, MSWS-12 from 25 (20-65) to 25 (20-78.3) points, and SDMT from 55.34 ± 14.7 to 61.4 ± 15.5 points. We observed similar results in instrumented variables: complexity from −0.15 ± 1.06 to −0.38 ± 1.08 [au], intensity from 0.00 ± 0.69 to −0.17 ± 0.78 [au], gait regularity from 0.87 ± 0.07 to 0.88 ± 0.08 [au], gait symmetry from 80.24 ± 20.24 to 83.0 ± 8.85 [au], and gait instability from 0.67 ± 0.11 to 0.70 ± 0.13 [au].

Conclusion: We hypothesized that subtle functional deteriorations would be detectable over two years. However, our data showed a functional stability of the disease at follow-up. This held even when an instrumented assessment was added to assess subtle functional disorders.

Keywords: Assessment, Functional disorders, Multiple sclerosis, Rehabilitation, Wearable sensors

What is already known about the topic:

- *Evidence has shown that even participants with EDSS ≤ 2.5 present functional deficits when assessed with specific functional tests.*

What does the study add:

- *Our study is the first longitudinal multicenter study that aims to provide a comprehensive picture of all the functional domains using objective, subjective, and instrumented tools to describe changes over time.*

Received: August 9, 2024
Accepted: June 7, 2025
Published online: July 1, 2025

Corresponding author:
Ilaria Carpinella
email: icarpinella@dongnocchi.it

Background

Functional impairments in early Multiple Sclerosis

Multiple Sclerosis (MS) is a disabling autoimmune disease of the central nervous system, showing a large heterogeneity in its clinical presentation and course. Among functional disorders, balance disorders, fatigue, and cognitive impairments

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

are the most common deficits in people with MS (PwMS), significantly impacting activities of daily living (1).

Notably, both static and dynamic balance disorders can emerge early, even in PwMS, who maintain normal walking ability and worsen over time, increasing the risk of falls and loss of independence (2). Fatigue, reported by up to 95% of PwMS, is one of the most debilitating symptoms, often persisting for years, exacerbating during relapses, and sometimes continuing even after recovery. Cognitive impairments frequently affect information processing speed, episodic memory, attention, and executive function, further contributing to disability. Although cognitive deficits have been observed in the early stages of the disease, recent studies suggest that motor impairments may be more prevalent than cognitive impairments in the early stages of MS (3,4).

So far, among clinician-reported outcome measures, the Expanded Disability Status Scale (EDSS) is the most commonly used tool to quantify disability status and disease progression in MS. Although its clinical usefulness has been well-reported over the years, several drawbacks have also been acknowledged (5). In particular, EDSS is not able to detect which specific functional deficits evolve during the course of the disease relative to the others. Compared to EDSS, both performance outcomes and patient-reported outcomes provide a more comprehensive assessment across different functional domains such as balance (e.g., Fullerton Balance Activity Scale), gait (e.g., 6-Minute walk test), upper limb function (e.g., 9-hole peg test), and cognition (Brief International Cognitive Assessment for MS) that are currently used to monitor disease progression from both clinical and patient perspective (6).

The use of technology in Functional Assessment

To overcome the lack of an objective perspective in terms of clinical dependency and measurement accuracy, the use of technology can be a valid solution. Instrumented assessments, particularly wearable inertial measurement units (IMUs) incorporating accelerometers, gyroscopes, and magnetometers, provide a sensitive and cost-effective way to detect subtle, clinically unobservable changes in function. These tools can assess balance and gait without requiring large spaces or extended evaluation times, offering objective measurements that capture small fluctuations in motor function over time (7).

These tools can assess balance and gait without requiring large spaces or extended evaluation times, offering objective measurements that capture subtle impairment not detectable by clinical scales (8,9), changes due to rehabilitation interventions (10), and small fluctuations over time in several aspects of motor function (11,12), particularly gait and balance (7).

For example, indexes such as sway complexity and sway intensity can describe irregular and large trunk sway during static balance tests (13). Specifically, low sway complexity indicates increased sway smoothness and regularity and reflects the loss of complexity and adaptability of the balance control system due to the reduction and/or impairments of its structural components and their interaction (14). Besides, a larger sway intensity means an increased amplitude and velocity of trunk sway, indicating the difficulty for PwMS in maintaining a stable position and a center of mass far from the limits of the base of support (15). As for walking, a recently published framework identified a set of validated instrumental metrics relevant to describing the different domains of walking in MS, such as gait rhythm, stability, symmetry, and variability (16).

Current gaps and study rationale

To date, few longitudinal studies have independently examined the progression of cognitive decline, gait impairment, and balance deterioration over time (17-20). Galea et al. reported significant declines in gait and balance, even though PwMS did not experience clinical relapse over a 12-month period (20). Moreover, Damasceno et al. found that stable cognitive function over a two-year observation period was associated with better long-term outcomes (17).

However, no longitudinal cohort studies have been performed yet to assess both motor and cognitive functions using clinical and instrumented tools to detect changes over time. Given the importance of characterizing specific deficits, identifying subtle deteriorations caused by the progression of the disease, and optimizing treatment strategies, this study aimed at investigating short-term motor and cognitive trajectories in early diagnosed, non-disabled PwMS using a combination of clinical, subjective, and instrumented tools. We hypothesize that even in the early stages of MS, measurable functional deteriorations can be detected over time, which could help guide more personalized and effective treatment strategies.

Methods

Participants

We recruited a convenience consecutive sample of PwMS diagnosed following the McDonald criteria (21) from three Italian centers.

The inclusion criteria were: confirmed MS diagnosis, EDSS ≤2.5 (i.e., no locomotor disability), disease duration ≤5 years, stable disease course clinically defined as <0.5-point increase in the EDSS disability score over the last 3 months, age ≥18 years and ability to give written informed consent. We excluded participants with diagnoses of major depression, severe joint and/or bone disorders interfering with balance and gait (based on clinical records), and cardiovascular or other concomitant neurological diseases.

The study was conducted in accordance with the Declaration of Helsinki and approved by the local Ethics Committee. Each participant gave the informed consent to participate.

Procedures

Participants were examined at two time points, baseline and after 2 years. At each of the time points, participants underwent a neurological assessment for disability, performance tests, and patient-reported outcomes scales on different domains (i.e., endurance, walking, balance, upper limb, cognition). We used written and standardized instructions, and the same test order was used at both time points, asking

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

participants to complete all questionnaires before the clinical and instrumented assessments.

The whole assessment was performed in a single session, allowing PwMS to rest approximately 2-5 minutes after each mobility test based on PwMS' needs. When participants had an exacerbation on the day of the assessment, measurements were postponed until the exacerbation had subsided or complaints had stabilized.

Researchers with experience in MS administered the assessments and were trained in dedicated practice sessions to ensure standardization among centers.

Outcome Measures

Disability

Expanded Disability Status Scale (EDSS) that ranges from 0 (normal neurological signs) to 10 (death due to MS). In specific, when a subject is independent in walking, the score ranges between 0 and 4.5, depending on the distance walked, while a score from 5.0 to 9.5 describes an increased difficulty in walking (22).

Walking

The Six Minute Walking Test (6MWT) measures walking endurance. Participants were instructed to safely walk as fast as possible to cover a maximum distance in 6 minutes on a 30 m walkway with turns at both ends. We recorded the total distance covered during the 6MWT (23). Moreover, PwMS performed the 6MWT while wearing three wireless inertial sensors (MTw, XSens, NL) secured to the lower back, at L5 level, right and left shanks (see Angelini et al. for further details) (7). Each sensor consisted of a 3D accelerometer (±160 m/s 2 range), a 3D gyroscope (±1200 deg/s range), and a 3D magnetometer (±1.5 Gauss). Sensor signals were recorded at 75 Hz. We calculated: Stride regularity (increasing values from 0 to 1 indicate higher stride regularity); Gait instability (larger score means increased instability, as detailed in Caronni et al. (24); Gait symmetry (increasing values from 0 to 100% indicate more symmetrical gait); and Cadence (computed as 60/Tstride, where Tstride is the stride duration) (11).

Walking speed was measured with the Timed 25 Foot Walk (T25FW). Participants were asked to walk twice at a fast but safe walking speed, and the mean of the two trials was calculated. The Twelve-Multiple Sclerosis Walking Scale (MSWS-12) was administered to rate perceived walking ability. Each item was scored on a 1-5 scale; the higher the score, the higher the perceived walking difficulties (25).

Balance

The Fullerton Advanced Balance-Short version scale (FAB-S) is a 6-item balance scale with a 5-point ordinal scale (0-4) for each item. The maximum score on the scale is 24 points, with increasing scores meaning better balance performance. The FAB-S scale is suitable for comparing participants with high balance skills with other scales (26). The subjects executed the test wearing an inertial sensor (MTw, Xsens, NL) secured at the sternum level with an elastic band. Signals from the inertial sensor were acquired with a sampling frequency of 75 Hz. Sway Complexity and Sway Intensity metrics, defined by McCallum et al. (27) and detailed in Carpinella et al. (13), were calculated. Increasing values of Sway Complexity indicate higher adaptability and automaticity of the balance control system, while decreasing values of Sway Intensity represent smaller amplitude and velocity of trunk sway and, consequently, better standing balance.

Fatigue

The Fatigue Severity Scale (FSS) is a tool designed to assess the impact of fatigue on daily function. It consists of nine items related to fatigue. The maximum score is 7, which indicates a high level of fatigue (28).

Upper limb

To assess upper limb function, we used the Nine Hole Peg Test (NHPT) and Manual Ability Measure-36 (MAM-36). The NHPT measures manual dexterity, and the mean of two trials was considered (29). MAM-36 was used to rate perceived manual ability on a scale ranging from 0 (impossible to complete any activities) to 144 (every activity is accomplished without any difficulty) (30).

Cognition

To assess learning, memory, and processing speed in PwMS, we used the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). It is a cognitive battery composed of three tests: the Symbol Digit Modalities Test (SDMT) to assess information speed, the California Verbal Learning Test-II (CVLT) for verbal learning and memory, and the revised version of the Brief Visuospatial Memory Test (BVMT) for visuospatial memory (31).

Statistical analysis

Mean (SD) values or median (min-max) were used to describe the magnitude of functional disorders, and linear models were used to identify univariate and multivariate statistically significant differences between baseline and follow-up scores, taking into account confounders such as age, disease duration, education, and medication. We checked for residual distribution and the presence of influential points. In the case of non-normal distribution of residuals, we used robust linear models using MM methods. Moreover, percentages of abnormal scores according to normative values were used to report functional disorders at baseline and at 2-year follow-up. The McNemar test was computed to check the differences between these percentages. R v3.6.2 was used for data analyses. $p < 0.05$ was considered statistically significant.

Results

In this paper, we reported data on 63 out of 82 (3) participants who completed all the assessments within 2 years. Participants were mostly female (n = 37, 60%), and all had a relapsing-remitting MS phenotype. At baseline, 13% of PwMS were taking no MS- Disease Modifying Therapy

(DMT), 62% first line, and 25% second line. These percentages were slightly different after 2 years: 7% no MS-DMT, 50% first-line MS-DMT, and 43% second-line MS-DMT. In detail, 44 PwMS maintained the same medication at follow-up; 8 PwMS (taking no MS-DMT at baseline) were taking respectively first-line DMT (n = 3) and second-line DMT (n = 2). Moreover, 8 PwMS changed medication from first-line to second-line DMT, and one subject changed from first-line DMT to no MS-DMT. Clinical longitudinal data are detailed in Table 1.

TABLE 1 - Clinical longitudinal data at baseline and at follow-up after 2 years.

	Baseline (N = 63)	2-year Follow-up (N = 63)	Univariate Estimate (Basel vs FU); p-value; and 95% CI	Multivariate Estimate (Basel vs FU); p-value; and 95% CI
EDSS (points)*	2 (0-2.5)	1.5 (0-4.5)	0.07; 0.50; −0.15 0.30	−0.03; 0.18; −0.30 0.23
Age (years)	38.7 (10.7)	40.6 (10.8)	-	
Disease duration (years)	2.3 (1.8)	4.3 (1.9)	-	
Education (years)	14.8 (3.6)	14.8 (3.6)	-	
6MWT (m)	566.2 (80.4)	573.9 (94.7)	12.80; 0.06; −0.04 25.77	12.7; 0.51; −3.63 29.00
FAB_S (points)*	23 (14-24)	23.5 (13-24)	−0.29; 0.22; −0.77 0.18	−0.40; 0.98; −1.01 −0.21
FSS (points)*	2.89 (0.89-7)	2.67 (1-7)	−0.01; 0.89; −0.3 0.26	−0.04; 0.94; −0.38 0.29
T25FW (sec)	4.30 (0.82)	4.11 (1.16)	**−0.24;0.04; −0.47 -0.006**	−0.25; 0.9; −0.55 0.05
9HPT (sec)				
Dominant hand	20.3 (3.4)	19.0 (3.0)	−1.01; 0.001; −1.75 −0.41	-5.00; 0.04; -9.58 -0.21
Non-dominant hand	21.6 (4.5)	20.4 (3.6)	−1.0, 0.001; −1.66 −0.40	−1.10; 0.39; −6.32 2.53
MSWS-12 (points)*	25 (20-65)	25 (20-78.3)	−0.13; 0.88; −1.90 1.63	0.37; 0.17; −1.49 2.24
MAM-36 (points)*	144.0 (94-144)	138.48 (104-144)	−0.0001; 0.44; −0.0002 0.0001	−2.37; 0.37; −0.61 0.15
SDMT (points)	55.34 (14.7)	61.4 (15.5)	**6.11; <0.001; 3.61 8.61**	**6.6; 0.02; 3.53 9.67**
CVLT (points)	10.4 (3.6)	13.0 (3.5)	**2.65; <0.01; 1.95 3.36**	2.21; 0.41; 1.41 3.02
BWMT (points)	10.50 (3.53)	11.42 (3.21)	**2.62; 0.02; 0.12 2.15**	0.88; 0.55; −0.13 1.89

*Median (min-max); EDSS: Expanded Disability Status Scale; MS: Multiple Sclerosis; 6MWT: Six Minute Walk Test; FAB_S: Fullerton Advance Balance scale_Short; FSS: Fatigue Severity Scale; T25FW: Timed 25 Foot Walk Test; 9HPT: Nine Hole Peg Test; MSWS-12: Multiple Sclerosis Walking Scale Twelve; MAM-36: Manual Ability Measurement_36; SDMT: Symbol Digit Modalities Test; CVLT: California Verbal Learning Test; BWMT: Brief Visuo-spatial Memory Test; * using robust linear estimation methods

Clinical changes between baseline and follow-up

Median EDSS scores were reported in Table 1, showing no or mild improvements between baseline and 2-year follow-up. This was true also considering age, sex, education, and DMD_line as covariates that were statistically significantly ($p < 0.05$) associated with SDMT for Age: Beta = 0.04, P = 0.03, and Education: Beta = 0.23, P= 0.09.

However, we can observe a different EDSS distribution over the years (Fig. 1). To note, at baseline, all PwMS were lower than 2.5 EDSS points, while after 2-year 24% of them worsened on EDSS (EDSS change ≥ 0.5 points), 12% improved, and only 63% were stable. Considering a 1-point change in the EDSS, we observed 9 (14%) PwMS deteriorated and 7 (11%) PwMS improved.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

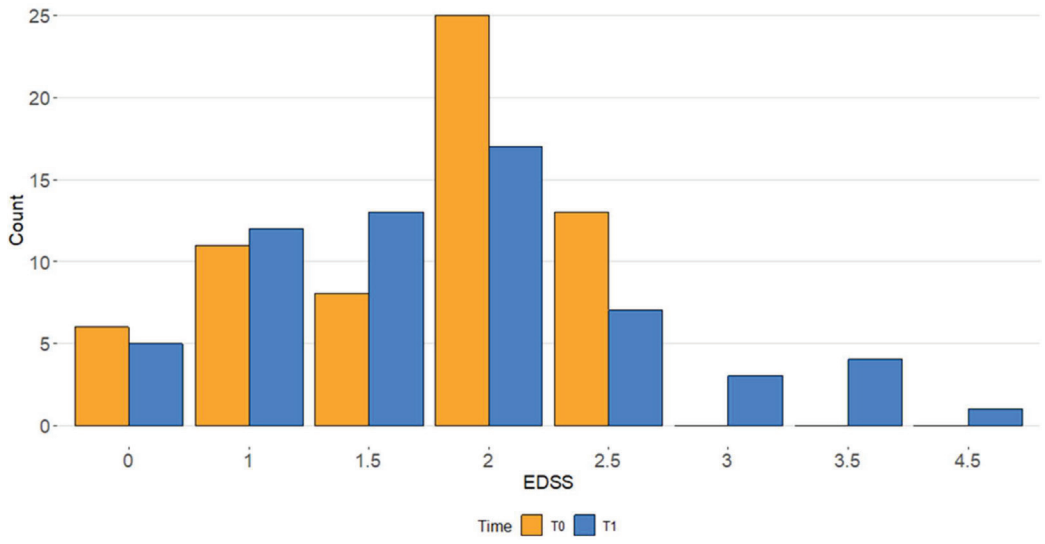


FIGURE 1 - Number of participants per different EDSS scores
EDSS: Expanded Disability Status Scale (EDSS); Count: number of PwMS; T0: Baseline; T1: Follow-up after 2 years

At baseline, MSWS-12 was the clinical measure showing the largest percentage of abnormal scores, with 60% of participants reporting perceived walking limitations (see also Supplementary Fig. 1). This percentage only slightly increased after 2 years (66%, p = 0.42).

When considering walking endurance, 27% of PwMS had an abnormal score in the 6MWT at baseline that significantly decreased after 2 years (13%, p = 0.03). Improvements were less evident for gait speed, with 18% of PwMS showing an abnormal score in the T25FWT at baseline and 14% at the 2-year follow-up (p = 0.68). This trend was reversed for balance with a percentage of abnormal score in the FAB_S of 46% at baseline and 49% after 2 years (p = 0.78, Fig. 2).

We did not observe changes in fatigue (FSS), with PwMS showing a percentage of 21% abnormal scores both at baseline and after 2 years (p = 0.99, Fig. 2).

One-third (33%) of PwMS showed abnormal NHPT scores at baseline that significantly decreased after 2 years (17%, p = 0.02). This reduced percentage of abnormal scores in manual dexterity was not accompanied by a higher percentage of abnormal scores in the perceived manual ability measured by the MAM-36 that significantly worsened from 35% at baseline to 48% at 2-year follow-up (p = 0.02, Fig. 2). In supplementary Figure 2. We also reported MAM-36 score at baseline and after 2 years, indicating worse performances on items more related to strength activities (ITEM 11: Wring a

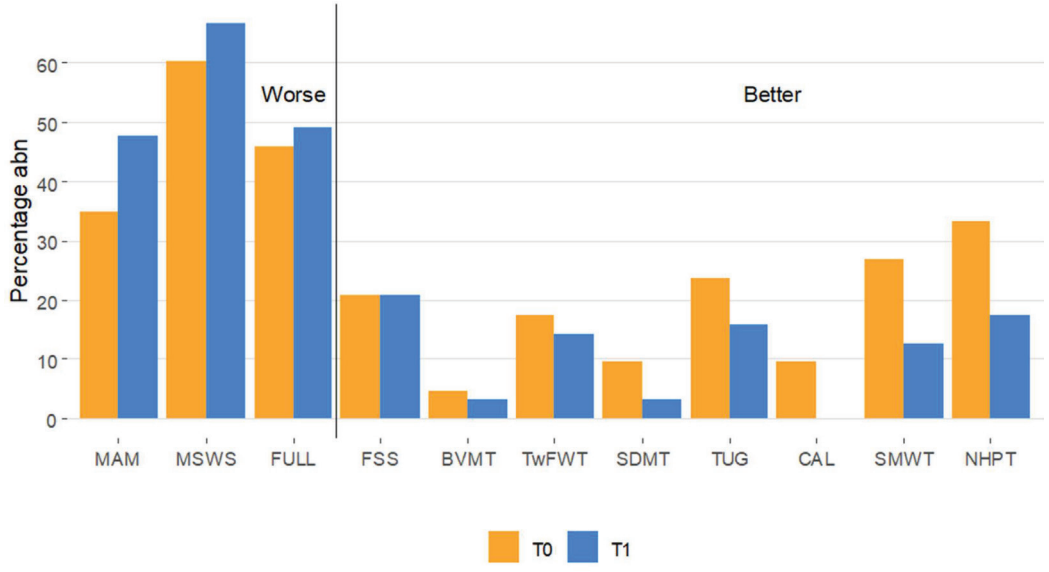


FIGURE 2 - Percentage of abnormal clinical scores at baseline and at follow-up
Percentage abn: Percentage abnormal scores; MAM: Manual Ability Measure-36; MSWS: Twelve-Multiple Sclerosis Walking Scale; FULL: Fullerton Advanced Balance_Short version scale; FSS: Fatigue Severity Scale; BVMT: Brief Visuospatial Memory Test TWFWT: Timed 25 Foot Walk; SMDT: Symbol Digit Modalities Test; TUG: Timed Up and GO; CAL: California Verbal Learning Test-II; SMWT: Six Minute Walking Test NHPT: Nine Hole Peg Test

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

towel; ITEM 16: Cut nails with a nail clipper; ITEM 26: Open a medication bottle with a child-proof top; ITEM 34: Use a hammer or screwdriver)

Unexpectedly, we did not find any relevant cognitive disorders at baseline. Indeed, they were the same or even better after 2 years: 10% had abnormal scores on the SDMT at baseline, 4% after 2-year (p = 0.22); 10% on the CVLT at baseline, 0% after 2-year (p = 0.13); and 5% on the BVMT at baseline, 3% after 2-year (p = 0.99, Fig. 2).

Instrumented changes between baseline and follow-up

As shown in Table 2, we observed statistically significant improvement in gait symmetry, stable results in cadence and gait regularity while stride length deteriorated when adjusting for covariates (Education: Beta = 0.01, P = 0.01)

Additionally, we obtained controversial results when considering static and dynamic balance indexes. Indeed, a statistically significant worsening of Sway Complexity during standing (i.e., reduced automaticity and adaptability of

TABLE 2 - Changes in instrumented balance and walking outcomes at baseline and 2-year follow-up

	Baseline	2-year Follow-up	Univariate Estimate (Basel vs FU); p-value; and 95% CI	Multivariate Estimate (Basel vs FU); p-value; and 95% CI
Static and dynamic Balance				
Complexity* (a.u.)	−0.15 ± 1.06	−0.38 ± 1.08	**−0.25, 0.001, −0.40 −0.10**	−0.22, 0.42, −0.41 −0.02
Intensity^ (a.u.)	0.00 ± 0.69	−0.17 ± 0.78	−0.22,0.06, −041 −0.02	**−0.10, 0.03, −0.25 0.06**
Gait Instability^ (a.u.)	0.67 ± 0.11	0.70 ± 0.13	**0.03, 0.04, 0.001 – 0.06**	0.05,0.34,0.021 0.09
Gait Quality				
Gait Symmetry* (a.u.)	80.24 ± 20.24	83.00 ± 8.85	**2.64, 0.001, 1.06 4.23**	2.58, 0.44, 0.40 4.76
Gait Regularity* (a.u.)	0.87 ± 0.07	0.88 ± 0.08	0.007, 0.28, -0.006 0.02	0.0002, 0.40, −0.018 0.018
General gait parameters				
Cadence* (stride/min)	63.79 ± 5.14	64.34 ± 5.77	0.56, 0.11, −0.14 1.27	0.25. 0.08, −0.65 1.15
Stride length* (m)	1.56 ± 0.19	1.54 ± 0.18	−0.02, 0.13, −0.05 0.006	**−0.02, 0.008 −0.06 0.01**

* Larger or ^ smaller values mean better performance

balance control) was accompanied by a statistically significant improvement in Sway Intensity (i.e., reduced amplitude and velocity of trunk sway). Regarding dynamic balance, Gait Instability showed a small worsening at a 2-year follow-up when adjusting for covariates (Disease duration: beta = 0.01, P = 0.04, Table 2). We found no statistically significant differences between baseline and 2-year follow-up in the percentage of PwMS showing abnormal values in static and dynamic balance: 27% had abnormal scores in sway complexity at baseline, 35% after 2-year (p = 0.18); 27% in sway intensity at baseline, 13% after 2-year (p = 0.06); and 51% on the gait instability at baseline, 57% after 2-year (p = 0.60, Fig. 3).

No differences were found in gait quality (Gait symmetry: 54% at baseline, 42% after 2-year; Gait regularity: 25% at baseline; 14% after 2-year, Fig. 3) and in general gait parameters (Cadence: 21% at baseline; 14% after 2-year; Stride length: 32% at baseline; 23% after 2-year, Fig. 3).

Discussion

In the present study, we evaluated functional changes over 2 years in a group of PwMS in the early phase of the disease, with no disability and fully able to walk. The first finding of this work highlights that the disability level remained stable in the majority of the sample over the years. Moreover, contrary to our initial hypothesis, also functional outcomes such as fatigue, upper limb, and cognition remained stable over the years. Finally, instrumented data showed mixed results, suggesting that it is still unclear whether instrumented variables could depict balance and gait changes in short-term follow-up.

One of the main results of the study is that the mean EDSS score remained stable over 2-years. Eighty-six percent of PwMS maintained a stable EDSS score, while 24% deteriorated, although they did not exceed the threshold for moderate disability, thus indicating a preserved walking ability. However, disability accrual observed in several subjects may indicate unfavorable disease progression, urging the development of predictive models to tailor interventions. Our data are in line with a previous study by Galea et al. that reported a worsening in 24% of subjects in a similar population of PwMS. (20) and other studies on DMT showed a stable course of the disease in most of the recruited participants in the years right after diagnosis (32). These results are corroborated by no changes in fatigue, upper limb function, and cognition at follow-up. This also holds for the percentage

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

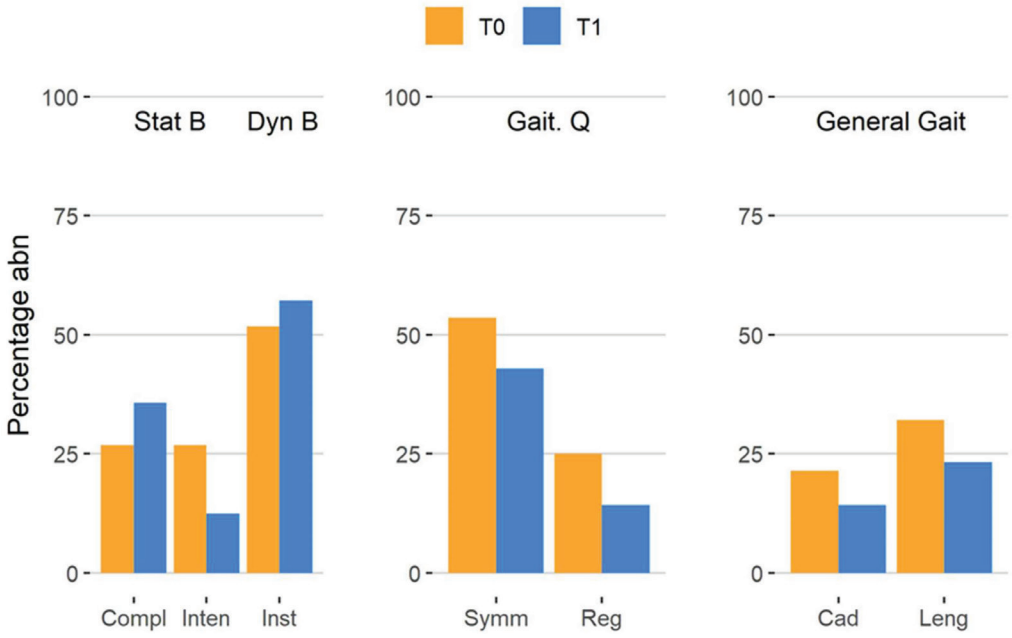


FIGURE 3 - Percentage of abnormal scores of instrumented variables at baseline and at follow-up
Stat B: Static Balance; Dyn B: Dynamic Balance; Compl: Complexity; Inten: Intensity; Inst: Instability; Gait Q: Gait Quality; Symm: Symmetry; Reg: Regularity; Cad: Cadence; Leng: Stride Length

of abnormal scores, which was unchanged across domains except for balance, perceived walking ability, and upper limb perception.

Our results obtained using a general measure of disability are parallel with changes we observed considering specific clinical scales. In our preceding cross-sectional study (3), we reported that fatigue was abnormal in 25% of PwMS; thus, we expected an increased percentage of PwMS affected by fatigue with a higher percentage of abnormal scores after two years. Surprisingly, fatigue remained stable in the majority of the sample, and only a few subjects worsened over the 2-year period, highlighting the stability of this symptom during this observational period. Our results are in agreement with other longitudinal studies. One by Damasceno et al. reported changes in fatigue over 2-year similar to our data (17). The other by Chang et al. that assessed 322 PwMS in the early phase reported no difference in fatigue after a short follow-up of 1-year. (33)

A lack of deterioration was also observed in the upper limb. Manual dexterity performances are similar between our sample and healthy subjects, indicating we assessed subjects that still did not show any upper limb dysfunction. Contrary to our expectation, we observed a statistically significant but no clinical improvement in hand performance at 2 years, possibly indicating that lack of manual dexterity impairments and test practice effects may have influenced our results on manual dexterity due to the repeated execution of the 9HPT. The lack of real clinical improvements in upper limb dexterity is confirmed by PRO data. Indeed, although we showed improvements in manual dexterity over the years, PwMS perception of the use of the upper limb during daily activity seemed to be unchanged. However, the relationship between perceived and objective deterioration should be considered in light of cross-sectional studies reporting there is a weak association between objective upper limb function and subjective perception in samples with an overall lower disability, while the association is strong when considering people with a higher level of disability (34). Interestingly, we observed that PwMS's perception was lower in activities more associated with strength (e.g., opening a bottle or wringing out a towel) than dexterity, indicating deficits more at the impairment level than functional level (see Supplementary Fig. 2).

Changes observed and perceived in motor function were accompanied by no sign of deterioration in cognition, especially in executive function, with a 6-point improvement over 2 years (35). Our results are in line with previous prospective studies reporting that (17,18,36) cognitive functions remained unchanged at short follow-ups, indicating that longer observations are crucial to understanding whether cognitive impairments are stable or are likely to progress.

Concerning mobility, around half of PwMS still had abnormal scores in balance tests that did not change compared to baseline assessment, confirming that balance disorder remains one of the most prevalent impairments in this population (3). Also, follow-up assessments still confirmed worse scores in tasks with body/head turns, highlighting the presence of somatosensory impairments (9). This is worth considering when planning rehabilitation interventions. In the specific gait, PwMS were faster (25FWT) at 2 years. However, this result was not confirmed by the endurance test, and subjective data showed a slightly higher percentage of abnormal scores on the perception of walking ability, indicating that the patient's perspective is different from the clinical tests in describing changes in gait.

Instrumented data on balance yielded controversial results since the worsening in Sway Complexity was

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

accompanied by an improvement in Sway Intensity during upright standing. We can speculate that worsening in sway complexity over the years indicates less automatic and more conscious postural control and the need for increased attentional investment to maintain balance (37). It could also be speculated that the lack of changes observed after 2 years is the result of compensatory mechanisms taking place to maintain balance at an optimal level. The assessment of cortical activity at the two time points may unveil these compensatory mechanisms.

Moreover, our longitudinal analysis of instrumented gait variables confirmed this trend of improvement in symmetry with a slight deterioration in stride length and gait instability. Our data are in line with a previous cross-sectional study by Kalron et al. (38) that observed no significant differences in gait variability in terms of step length, step time, step width, and single support in PwMS with low disability, while a statistically significant increase when PwMS reached higher level of disability. Similarly, our previous cross-sectional study on a subsample of the same population showed that PwMS who worsened in EDSS showed higher gait abnormalities (increased gait instability, decreased symmetry, and lower stride regularity) compared to stable PwMS confirming this association between high disability and gait abnormality (11). While our findings are consistent with these earlier studies, the longitudinal analysis of instrumented gait variables in this population represents a novelty, making direct comparisons with existing research challenging.

Taken together, the main point of this study is that a 2-year observation period is too short to reveal a deterioration of the disease at the functional level, suggesting that a longer timeframe and a deeper analysis at the impairment level should be considered to monitor MS functional deterioration and improvement. Moreover, It could also be speculated that the lack of changes observed after 2 years is the result of brain adaptation taking place to maintain functions at an optimal level. Specifically, we can postulate that the lack of changes at the behavioral level is due to compensatory brain mechanisms, which translate into bilateral recruitment or the involvement of areas not typically involved in motor control. The combined assessment of behavior and cortical activity at the two time points may unveil these compensatory mechanisms.

Strengths and limitations

The strength of this work is that we reported for the first time a comprehensive longitudinal picture of motor and cognitive disorders assessed both with clinical and instrumented tests in non-disabled PwMS. However, we should state some limitations. The first main limitation is that the follow-up period was too short to reveal deterioration in this population, although we did not address the potential for non-linear progression in early MS. Second, even if the follow-up was performed after 2 years, we tested cognition without providing different versions of the test. Third, we did not collect EDSS functional scores that could detail the impairment level of each subject. Fourth, our results are based on the analysis of a single cohort with a relatively small sample size that may limit the generalizability of findings and the precision of the estimations. Another limitation is the potential insufficient sensitivity of the tasks performed and the tools employed. More complex assessments, such as prolonged treadmill walking or the use of more advanced technologies (e.g., optoelectronic systems, electromyography, and force plates), could have enhanced the sensitivity of the evaluation. However, this would have compromised the ecological validity of the assessment, which was intentionally designed to closely approximate real-life conditions. Lastly, we did not specifically perform clinical scales aimed at assessing deficits at the impairment level.

Conclusion

Although our expectation was that PwMS performances would have worsened across domains after 2 years or, at least, we expected that the addition of instrumented assessment could depict subtle changes in balance and gait performances. However, our data showed functional stability of the disease after 2 years. To guide clinicians towards more personalized medicine, future research with larger cohorts, longer follow-ups, and biomarker assessments are needed. This will be crucial for the development of predictive models that can guide the selection of optimal interventions to mitigate the progression of the disease.

Disclosures

Conflict of interest: The authors declare that there is no conflict of interest.

Financial support: This work was supported by the Italian Multiple Sclerosis Foundation—FISM (FISM grant 2016, N16/17/F14) and by the Italian Ministry of Health (Fondi di Ricerca Corrente 2024).

Ethical approval: The study was approved by the ethical committees of all three participating centers. Written informed consent was obtained from all participants. The project was conducted in accordance with the Declaration of Helsinki and with local ethical guidelines.

Author contributions: EG: conceptualization (lead); data curation (supporting); formal analysis (equal); funding acquisition (lead); investigation (lead); methodology (equal); project administration (lead); resources (lead); supervision (lead); writing review and editing (lead). CS: Conceptualization (equal); data curation (supporting); funding acquisition (equal); investigation (equal); methodology (equal); project administration (equal); resources (supporting); supervision (equal); writing review and editing (lead). DA: Data curation (equal); funding acquisition (supporting); investigation (supporting); methodology (supporting); supervision (supporting); writing review and editing (equal). RDG: Data curation (equal); investigation (supporting); writing review and editing (supporting). AT: Conceptualization (equal); data curation (supporting); funding acquisition (equal); investigation (equal); methodology (equal); project administration (supporting); resources (supporting); supervision (equal); writing review and editing (equal). GB: Conceptualization (equal); data curation (supporting); funding acquisition (equal); investigation (equal); methodology (equal); project administration (supporting); resources (supporting); supervision (equal); writing review and editing (equal). PC: Conceptualization (supporting); resources (equal); writing review and editing (equal). MR: Conceptualization (supporting); resources (equal); writing review and editing (equal). IC: conceptualization (lead); formal analysis (lead); investigation (equal); methodology (lead); software (lead); writing original draft (lead). DC:

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

conceptualization (lead); data curation (supporting); formal analysis (equal); funding acquisition (lead); investigation (lead); methodology (equal); project administration (lead); resources (lead); supervision (lead); writing review and editing (lead).

References

1. Moreno-Navarro P, Gomez-Illán R, Carpena-Juan C, et al. Understanding the deterioration of gait, postural control, lower limb strength and perceived fatigue across the disability spectrum of people with multiple sclerosis. J Clin Med. 2020;9(5):1385. CrossRef PubMed
2. Beckerman H, Kempen JCE, Knol DL, et al. The first 10 years with multiple sclerosis: the longitudinal course of daily functioning. J Rehabil Med. 2013;45(1):68-75. CrossRef PubMed
3. Cattaneo D, Gervasoni E, Anastasi D, et al. Prevalence and patterns of subclinical motor and cognitive impairments in non-disabled individuals with early multiple sclerosis: a multicenter cross-sectional study. Ann Phys Rehabil Med. 2022;65(1):101491. CrossRef PubMed
4. Thrue C, Riemenschneider M, Hvid LG, et al. Time matters: early-phase multiple sclerosis is accompanied by considerable impairments across multiple domains. Multiple Sclerosis Journal. 2021;27:1477-85. CrossRef
5. Hobart JC, Riazi A, Lamping D L, et al. Measuring the impact of MS on walking ability The 12-Item MS Walking Scale (MSWS-12). 2003. CrossRef
6. Zaratin P, Vermersch P, Amato MP, et al. The agenda of the global patient reported outcomes for multiple sclerosis (PROMS) initiative: progresses and open questions. Vol 61. Multiple Sclerosis and Related Disorders. Elsevier B.V; 2022.
7. Angelini L, Carpinella I, Cattaneo D, et al. Is a wearable sensor-based characterisation of gait robust enough to overcome differences between measurement protocols? A multi-centric pragmatic study in patients with multiple sclerosis. Sensors (Switzerland). 2020 Jan 1;20(1). CrossRef
8. Carpinella I, Gervasoni E, Anastasi D, et al. Instrumentally assessed gait quality is more relevant than gait endurance and velocity to explain patient-reported walking ability in early-stage multiple sclerosis. Eur J Neurol. 2021;28(7):2259-2268. CrossRef PubMed
9. Carpinella I, Gervasoni E, Anastasi D, et al. Walking with horizontal head turns is impaired in persons with early-stage multiple sclerosis showing normal locomotion. Front Neurol. 2022;12:821640. CrossRef PubMed
10. Hilfiker R, Vaney C, Gattlen B, et al. Local dynamic stability as a responsive index for the evaluation of rehabilitation effect on fall risk in patients with multiple sclerosis: a longitudinal study. BMC Res Notes. 2013;6(1):260. CrossRef PubMed
11. Gervasoni E, Anastasi D, Di Giovanni R, et al. Uncovering subtle gait deterioration in people with early-stage multiple sclerosis using inertial sensors: a 2-year multicenter longitudinal study. Sensors (Basel). 2023;23(22):9249. CrossRef PubMed
12. Vienne-Jumeau A, Quijoux F, Vidal PP, et al. Wearable inertial sensors provide reliable biomarkers of disease severity in multiple sclerosis: a systematic review and meta-analysis. Annals of Physical and Rehabilitation Medicine. 2020;63:138-47. CrossRef
13. Carpinella I, Anastasi D, Gervasoni E, et al. Balance impairments in people with early-stage multiple sclerosis: boosting the integration of instrumented assessment in clinical practice. Sensors (Basel). 2022;22(23):9558. CrossRef PubMed
14. Roeing KL, Wajda DA, Sosnoff JJ. Time dependent structure of postural sway in individuals with multiple sclerosis. Gait Posture. 2016;48:19-23. CrossRef
15. Prosperini L, Fortuna D, Giannì C, et al. The diagnostic accuracy of static posturography in predicting accidental falls in people with multiple sclerosis. Neurorehabil Neural Repair. 2013;27(1):45-52. CrossRef PubMed
16. Karatsidis A, Angelini L, Scaramozza M, et al. Characterizing gait in people with multiple sclerosis using digital data from smartphone sensors: a proposed framework. Multiple Sclerosis Journal. 2025;31(5):512-528 CrossRef
17. Damasceno A, Pimentel-Silva LR, Damasceno BP, et al. Exploring the performance of outcome measures in MS for predicting cognitive and clinical progression in the following years. Mult Scler Relat Disord. 2020;46:102513. CrossRef PubMed
18. Amato MP, Ponziani G, Siracusa G, et al. Cognitive dysfunction in early-onset multiple sclerosis: a reappraisal after 10 years. Arch Neurol. 2001;58. CrossRef
19. Dreyer-Alster S, Menascu S, Dolev M, et al. Longitudinal relationships between disability and gait characteristics in people with MS. Sci Rep. 2022;12(1):3653. CrossRef PubMed
20. Galea MP, Cofré Lizama LE, Butzkueven H, et al. Gait and balance deterioration over a 12-month period in multiple sclerosis patients with EDSS scores ≤ 3.0. NeuroRehabilitation. 2017;40(2):277-284. CrossRef PubMed
21. Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. The Lancet Neurology. 2018;17:162-173.
22. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983;33(11):1444-1452. CrossRef PubMed
23. Goldman MD, Marrie RA, Cohen JA. Evaluation of the six-minute walk in multiple sclerosis subjects and healthy controls. Mult Scler. 2008;14(3):383-390. CrossRef PubMed
24. Caronni A, Gervasoni E, Ferrarin M, et al. Local dynamic stability of gait in people with early multiple sclerosis and no-to-mild neurological impairment. IEEE Trans Neural Syst Rehabil Eng. 2020;28(6):1389-1396. CrossRef PubMed
25. Solaro C, Trabucco E, Signori A, et al. Italian validation of the 12-item multiple sclerosis walking scale. Mult Scler Int. 2015;2015:540828. CrossRef PubMed
26. Mestanza Mattos FG, Gervasoni E, Anastasi D, et al. Assessing balance in non-disabled subjects with multiple sclerosis: validation of the Fullerton Advanced Balance Scale. Mult Scler Relat Disord. 2020;42:102085. CrossRef PubMed
27. MacCallum RC, Widaman KF, Preacher KJ, et al. Sample size in factor analysis: the role of model error. Multivariate Behav Res. 2001;36(4):611-637. CrossRef PubMed
28. Krupp LB, Larocca NG, Muir-Nash J, et al. The Fatigue Severity Scale application to patients with multiple sclerosis and systemic lupus erythematosus. CrossRef
29. Solaro C, Grange E, Di Giovanni R, et al. Nine Hole Peg Test asymmetry in refining upper limb assessment in multiple sclerosis. Mult Scler Relat Disord. 2020;45:102422. CrossRef PubMed
30. Solaro C, Di Giovanni R, Grange E, et al. Italian translation and psychometric validation of the Manual Ability Measure-36 (MAM-36) and its correlation with an objective measure of upper limb function in patients with multiple sclerosis. Neurol Sci. 2020;41(6):1539-1546. CrossRef PubMed
31. Goretti B, Niccolai C, Hakiki B, et al. The Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS): normative values with gender, age and education corrections in the Italian population. BMC Neurol. 2014;14(1):171. CrossRef PubMed
32. Hauser SL, Bar-Or A, Comi G, et al. Ocrelizumab versus interferon beta-1a in relapsing multiple sclerosis. N Engl J Med. 2017;376(3):221-234. CrossRef PubMed

33. Chang YT, Kearns PKA, Carson A, et al. Network analysis characterizes key associations between subjective fatigue and specific depressive symptoms in early relapsing-remitting multiple sclerosis. Mult Scler Relat Disord. 2023;69:104429. CrossRef PubMed
34. Grange E, Solaro C, Di Giovanni R, et al. The correlation between 9-HPT and patient-reported measures of upper limb function in multiple sclerosis: a systematic review and meta-analysis. Journal of Neurology. 2023;270:4179-4191. CrossRef
35. Benedict RH, DeLuca J, Phillips G, et al. Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. Multiple Sclerosis Journal. 2017;23(5):721-733. CrossRef
36. Longinetti E, Englund S, Burman J, et al. Trajectories of cognitive processing speed and physical disability over 11 years following initiation of a first multiple sclerosis disease-modulating therapy. J Neurol Neurosurg Psychiatry. 2023. PubMed
37. Donker SF, Roerdink M, Greven AJ, et al. Regularity of center-of-pressure trajectories depends on the amount of attention invested in postural control. Exp Brain Res. 2007;181(1):1-11. CrossRef PubMed
38. Kalron A. Gait variability across the disability spectrum in people with multiple sclerosis. J Neurol Sci. 2016;361:1-6. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 184-191
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3405
ORIGINAL RESEARCH ARTICLE

Primary care rehabilitation after Knee Replacement – a cross sectional study

Elin Östlind[1,2], Marcus Ljung[3,4], Caroline Ståhl[5], Anna Cronström[1], Thérèse Jönsson[1,6]

[1]Department of Health Sciences, Lund University, Lund - Sweden
[2]Dalby Healthcare Centre, Dalby - Sweden
[3]Department of Orthopedics, Vrinnevi Hospital, Norrköping - Sweden
[4]Department of Health, Medicine, and Caring Sciences, Linköping University, Linköping - Sweden
[5]Department of Orthopedics, Ängelholms Hospital, Ängelholm - Sweden
[6]Department of Orthopedics, Skåne University Hospital, Lund - Sweden

ABSTRACT

Background: Rehabilitation after Knee Replacement (KR) surgery often entails an extensive rehabilitation in primary care but evidence-based high-quality guidelines are lacking. There is also a knowledge gap regarding current rehabilitation modalities applied in primary care in Sweden. This study aimed to (I) describe rehabilitation in primary care after KR and (II) explore physiotherapists' perceptions of patients' challenges during the rehabilitation.
Methods: A cross-sectional, web-based survey was conducted among Swedish physiotherapists working in primary care. Questions were categorical or open-ended and related to current rehabilitation practices, treatment modalities, and physiotherapists' perceptions of patients' challenges in rehabilitation after KR. Data were described descriptively and open-ended answers were analyzed with quantitative and qualitative content analysis.
Results: In total, 202 physiotherapists answered the survey. Rehabilitation focused on home exercises with recurrent physiotherapy visits. Common treatment modalities were knee range of motion exercises, strength training, and stationary cycling. Key rehabilitation challenges included the following categories: *Patients are unprepared*, *Challenging to find the optimal load*, and *Restoring function and trust in the knee*.
Conclusion: Rehabilitation after KR in Swedish primary care seems to be in line with previously recommended international treatment modalities. According to the physiotherapists in this study, some of the key challenges that patients faced were not being prepared for the severe pain regaining function and trust in the knee, balancing load/recovery, and resuming physically demanding activities.

Keywords: Knee Replacement, Rehabilitation, Primary Care, Physiotherapy

What's already known about this topic?

- *Although rehabilitation in primary care after KR is common, lengthy, and challenging for patients, there is no knowledge of current rehabilitation modalities used by physiotherapists in Sweden.*

What does the study add?

- *Swedish primary care physiotherapists seem to apply previously, internationally recommended active treatment modalities in KR rehabilitation. Key challenges for the patients were severe pain, knee swelling, balancing load/recovery, and resuming physically demanding activities.*

Received: November 29, 2024
Accepted: June 1, 2025
Published online: July 3, 2025

This article includes supplementary material

Corresponding author:
Elin Östlind
email: elin.ostlind@med.lu.se

Introduction

The first-line core treatment for Knee Osteoarthritis (OA) includes education, structured exercise, and, when necessary, weight loss (1,2). If first-line treatment is insufficient, KR can be a suitable option for managing end-stage OA (2). KR is one of the most common orthopedic procedures, and its prevalence is expected to rise as the number of individuals

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

with knee OA continues to grow (3). Standard treatment for end-stage OA is a total KR while unicompartmental KR is used less frequently (4).

In Sweden, a fast-track perioperative protocol is applied for elective KR which has resulted in shorter hospital stays (5). In contrast to the short hospital stays, the rehabilitation after KR is often experienced as long and challenging according to qualitative studies (6,7). Rehabilitation after KR seems to require more physiotherapy resources compared to Hip Replacement (HR) (8,9). Although most knee range of motion (ROM) is regained within the first three months after surgery, quadriceps strength imbalances between the operated and non-operated leg can persist for a year or longer after the surgery (10,11). Rehabilitation following KR typically includes strength training and ROM exercises, although these recommendations are based on limited and unclear evidence (12). In fact, a recent systematic review highlighted the lack of evidence regarding effective rehabilitation after KR (13). In Sweden, 20,622 primary elective KR were reported during 2023, and the prevalence of individuals with at least one KR is 1.5% per 100,000 Swedish residents, according to the Swedish Arthroplasty Register (14). Patients undergoing KR in 2023 had a mean age of 69.4 years, 55.4% were women and 37% had a Body Mass Index ≤30.

Given the high prevalence of KR and possible existing regional differences in care management, a national interprofessional working group was appointed within Sweden's national system for knowledge-based healthcare management in 2023 to develop a person-centered and cohesive clinical pathway/guideline for patients undergoing KR (15). The guideline should incorporate the clinical pathway including surgery, peri-operative, post-operative care, and rehabilitation, and be based on reliable international guidelines, systematic reviews, or individual studies (16). The working group included orthopedic surgeons, an anesthesiologist, nurses, an occupational therapist, physiotherapists, a general practitioner, and a patient representative.

In developing the clinical pathway/guideline for KR, it was noted that the existing international clinical guidelines and recommendations are mostly focused on the surgical procedure and peri-operative care and not the post-operative rehabilitation (17). There are some clinical practice guidelines that provide general recommendations for KR, however, these are either not evidence-based or supported by limited or unclear evidence (12). A scoping review by Krysa et al. (12) provided some recommendations regarding rehabilitation after KR, but it is unclear if these treatment modalities are applied by physiotherapists in primary care in Sweden. This knowledge gap, including patients' referral from orthopedic clinics to primary care, is important to address when developing a clinical pathway/guideline. Furthermore, understanding physiotherapists' perceptions of the challenges patients face during rehabilitation after KR is also essential to explore, especially in light of previous research highlighting the patients' difficulties and the mismatch between patients' rehabilitation results and expectations (6,7).

Therefore, the aims of this study were (I) to describe rehabilitation in Swedish primary care after KR and (II) physiotherapists' experiences of patients' challenges during the rehabilitation.

Methods

Design and setting

A cross-sectional web-based survey study was conducted within physiotherapy practice in primary care. Swedish primary care is characterized by a decentralized healthcare system with variations across different regions. These regional differences impact accessibility and quality of care, as each region is responsible for organizing and financing its healthcare services (18). Primary care in Sweden is primarily financed through taxes and thus, healthcare is available to all residents at a relatively low out-of-pocket cost (19). Physiotherapy is an integral part of primary care in Sweden, where physiotherapists work at health care centers or operate their own private practices. Patients can access physiotherapy either through referral or by directly contacting a physiotherapist of their choice (20). According to healthcare data from Region Skåne in southern Sweden, most KR patients undergo rehabilitation in primary care and attend a median of 15 visits (interquartile range: 9–23).

Data collection and participants

An online, public survey was created to identify current rehabilitation practices, treatment modalities, and recommendations for patients after all types of KR. The survey aimed to reach physiotherapists working in primary healthcare who treated patients after KR. They were requested to provide responses in line with local guidelines (if existing) or established best practices in their geographic area. The online survey was published in November 2023 and was open until January 2024. A link to the survey was distributed through the Swedish Physiotherapy Association, social media, and by connections with physiotherapists working in different regions.

Survey development

A survey tool, Sunet Survey was used to construct the digital questionnaire and manage the responses. The survey was constructed in connection with the development of the national care pathway for patients undergoing KR. Three of the authors (EÖ, ML, CS) were part of the working group for the care pathway and were responsible for constructing and distributing the survey. They are all physiotherapists with clinical experience in primary care (EÖ) and orthopedic clinics conducting KRs (ML, CS). A patient representative in the working group also contributed feedback in this process. A pilot version of the survey was tested by five physiotherapists in primary care and minor language adjustments were made after their feedback. No personal information about the respondents was collected, only the type of workplace was requested (i.e., private primary care, own establishment, regional primary care, municipal primary care).

The questions in the survey were related to rehabilitation in primary health care including referral from orthopedic clinics

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

after surgery, rehabilitation modalities, and recommendations regarding walking. One question focused on patients' challenges during the rehabilitation. The questions were categorical (binary or multiple-choice) or open-ended. The open-ended questions allowed the respondents to provide elaborative responses regarding their perceptions of important considerations in different phases of the rehabilitation. The survey questions related to this study are included in Appendix I.

Data analysis

EÖ and TS had the major responsibility for the analysis with contributions from the other co-authors. The distribution of the categorical questions regarding referral to primary care after KR, rehabilitation modalities and recommendations concerning walking were synthesized in Microsoft Excel, calculated and descriptively presented. Open-ended responses regarding *treatment and walking* were summarized in the text.

The free-text answers responding to the perceived challenges for the patients were pragmatically analyzed with manifest quantitative and qualitative content analysis (21,22). All free text that responded to the specific aim were seen as a unit of analysis. The responses were read several times by EÖ and TS. Each response could entail several challenges, for example, knee swelling *and* pain. The identified challenges were seen as meaning units. During the analysis, the meaning units were repeatedly compared and thereafter grouped in similar subcategories and summarized in actual numbers of responses per subcategory using Microsoft Excel and the software program N'Vivo (23). The subcategories were grouped in categories based on their content. Data was reorganized several times before reaching the results. The results were presented in text with representative quotes from respondents' free-text answers. The number represents the ID of the respondent.

Ethics

Ethical approval was not required for this study as no personal data was collected.

Results

The survey was responded to by 202 physiotherapists in 14 (of 21) Swedish regions. Fifty-six percent worked in the three largest regions in Sweden (Stockholm, Västra Götaland, and Skåne). The physiotherapists worked in regional primary care (58%), private primary care (38%), private practice (3%), or other (1%).

Referral to primary care

The first contact between the patient and the primary care physiotherapist after the surgery can be established in different ways according to the respondents. Patients can be referred from the orthopedic clinic, make an appointment themselves, and in some cases, they already have an established contact with the physiotherapist prior to surgery. In cases where patients contact the physiotherapist themselves, most respondents (72%) preferred that the first appointment after KR would be scheduled before the surgery. Regarding timing for the first post-operative visit, *2–3 weeks* was the most common response (57%), followed by *within a week (36%).*

Rehabilitation modalities

Treatment was most frequently (73%) delivered as home exercises combined with recurring physiotherapist visits. Group sessions and digital solutions were only reported to be used by a few physiotherapists (7%). Almost all respondents reported using a combination of active knee range ROM, strength training, and stationary cycling in KR rehabilitation (Fig. 1).

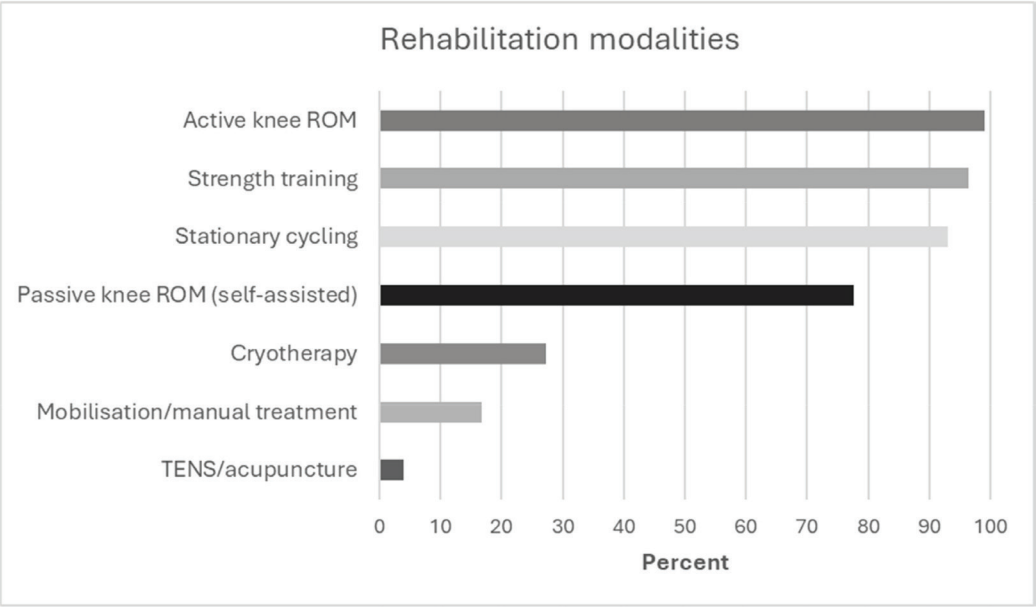


FIGURE 1 - Reported use of treatment modalities in primary care rehabilitation after KR surgery.

Regaining knee ROM was emphasized in the early post-operative phase of the rehabilitation. The respondents often recommended the patients to use a stationary bike or a peanut/bobath ball as equipment in partially unloaded exercises to increase knee flexion.

> Range of motion is prioritized during the initial period. Load to the point of pain and inform the patient that pain and swelling are common during the first few months and that this is normal. Resp. 21

Mobility exercises were combined with progressive strength training with an initial focus on the unloaded contraction of m. quadriceps proceeding to functional exercises involving other muscle groups such as the hip abductors and extensors. The respondents frequently recommended functional exercises such as rising from a chair and stair climbing.

Recommendations concerning walking

Sixty-one percent reported having a general recommendation for walking aids. The respondents were mostly in agreement regarding the type of walking aids and the duration of the period with walking aids. Two crutches were recommended unless the patient required a walker or other walking aid. Using crutches was recommended for at least 4–6 weeks or for as long as the patient walked with a limp.

Recommendations on walking distance were individual to a high extent, and most respondents emphasized that the walking time/distance should be limited during the first weeks after KR. Knee pain and joint swelling were frequently mentioned factors that should guide the walking time/distance. Some respondents were less restrictive and recommended as much loading as possible and 20–30 minutes of daily walking in the first weeks after the surgery.

> Listen to the body. Stay inside during week one and walk as little distance as possible the first 2 weeks after surgery. Then individually tailored recommendations. Resp. 97

> I usually recommend a daily walk of 20 minutes the first time (after surgery). Then increase when 5–6 weeks have passed depending on the patient's conditions. Resp. 153

Closely monitoring the patient, guiding them in balancing activity/load with recovery, and providing support during the challenging period following the surgery was seen as important.

Challenges during the rehabilitation

184 respondents (91%) answered an open-ended question related to patients' difficulties during the rehabilitation. Three main categories were identified, each with four to five subcategories. The main categories, subcategories, and the number of comments are presented in Table I.

TABLE 1 - Physical therapists' perception of patients' challenges during the rehabilitation

Main categories	Subcategories	Counted comments (n)
Patients are unprepared	Lengthy rehabilitation	16
	To endure the first tough weeks	9
	Misconceived expectations	9
	Handling life-long limitations	5
Challenging to find the optimal load	Manage pain and swelling	42
	Adherence to exercise	27
	Activity balance	13
	Setbacks with too much loading	9
Restore function and trust in the knee	Regain knee range of motion	61
	Walk without limping	34
	Resume physically demanding work and activities	27
	Kinesiophobia	8
	Trusting the knee	6

Patients are unprepared

The respondents wrote that patients could be surprised by the often severe pain during the first post-operative period. Furthermore, the patients were not always aware that the rehabilitation might be very lengthy. Sometimes, patients had expected to be able to return to work or activities more quickly. One respondent answered that many patients might be surprised that they don't recover as quickly as they had expected. Pre-operative information about the rehabilitation period was mentioned as extremely important.

> Patients often do not adequately absorb the information from the orthopedic surgeon regarding time, pain, and the rehabilitation process. Patients are often surprised that they experience such significant pain after the surgery. Resp. 57

The rehabilitation after the surgery often continues for a long period. The duration of the period with recurring follow-up visits differs according to the respondents, but a common period seems to be between three to six months with a transition to self-care only. It can take up to a year or more for full recovery after KR, even though considerable improvements often occur within the first few months.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

> Some people are surprised that it takes so long before it feels ok, that it takes time before it feels like "my knee" and not my "operated knee", which can take one year and sometimes longer than that. Resp. 68

Challenging to find the optimal load

Finding the appropriate balance between load and recovery was mentioned as difficult for the patients by approximately half of the respondents. This seemed to be especially important during the first post-operative period. Too much load led to increased knee swelling and pain. Knee ROM and pain were often driven by swelling, which could worsen if the patient overloaded the knee. The respondents emphasized the importance of guiding patients to find the optimal load to manage knee swelling effectively. Identifying the optimal load and not exceeding it was something that the respondents often discussed with their patients. Several respondents noted that younger patients were often eager to return to work and daily activities, leading them to discontinue the use of walking aids or place excessive strain on the knee prematurely.

> For younger patients: Eager to start with heavier loads, they may need to be slowed down as they often experience increased swelling and pain with too rapid progression. Resp. 121

The physiotherapists experienced that an increase in pain after loading could cause a setback for many patients and that the setback might have a negative impact on their motivation and mental health. An important aspect that the respondents wrote was that they offered support to the patients and helped them find the appropriate balance between load and recovery. Another challenge that the respondents acknowledged was patients' limited motivation and low adherence to exercise. Patients who had a low adherence to the rehabilitation protocol may risk a worse outcome after the surgery. Continuation and consistency in the rehabilitation were recommended to improve adherence.

> The motivation to exercise as much as required after knee arthroplasty is not an easy rehabilitation process that happens on its own; much depends on the patient. Resp. 121

Restore function and trust in the knee

Regaining knee ROM was seen as a priority during the early phase of the rehabilitation, and it was also mentioned as a common struggle that the patients had, especially in knee extension. The respondents considered it important that the patients push the knee to full ROM. Many patients had difficulties with regaining quadriceps strength and a symmetrical gait pattern. The respondents stated that this was probably because the patients had been limping for quite a while before the surgery and this was difficult to change.

> The most difficult part for some is finding an 'optimal' walking pattern after the surgery, i.e., avoiding the limping gait they had before the surgery. Resp. 68

Some patients have difficulties trusting their knee and understanding that the pain they are experiencing does not mean that there is something wrong with the KR. Kinesiophobia may be an obstacle for some patients, especially in the early post-operative period when the pain is worse. The respondents also reported that the patients might have a fear of doing something that might "destroy" the KR and force them to have a revision surgery.

Discussion

This study explores rehabilitation in primary care following KR surgery. The most reported treatment modalities included a combination of active ROM exercises and strength training. Passive treatments like therapist-guided mobilization and acupuncture were less frequently utilized. The physiotherapists identified several key challenges for the patients during the rehabilitation process, with the difficulties of regaining ROM and managing pain and swelling in the knee being the most reported.

Since patients are typically discharged from the hospital shortly after surgery, early support from a physiotherapist may be crucial. Most physiotherapists in our study expressed a preference for scheduling the first post-operative session prior to surgery, to avoid delays in rehabilitation. If patients booked their appointment only after the surgery, they risked facing a long wait for an appointment and, in some cases, they did not prioritize physiotherapy due to post-surgery pain. This opinion echoes findings from a Japanese study, where patients who booked appointments after surgery had to wait several weeks for their first physiotherapy session, a delay that was deemed suboptimal (24).

The treatment modalities suggested in this study mainly align with the results in a scoping review of clinical practice recommendations for post-operative joint replacement rehabilitation (12). That review recommended early joint mobilization and physiotherapy, while treatments like continuous passive motion and transcutaneous electrical nerve stimulation therapy (TENS) lacked sufficient evidence to support or oppose their use. In the review by Krysa et al., *physiotherapy* encompassed several modalities such as strength training, exercise, ROM exercise, manual therapy, aquatic exercise, balance and mobility training. In our study, few respondents reported that they used manual therapy and, given the responses in free-text, it was mainly used as an add-on if the initial active ROM exercise did not have the desired effect on knee mobility.

According to the respondents, there are many challenges facing KR patients. A lack of preparedness for the struggles during the rehabilitation was identified as a factor that negatively impacted the recovery process. The respondents expressed that many patients were unaware of the initial pain and swelling and that it may take a long time to regain

knee function. This lack of awareness can lead to frustration and may negatively impact their adherence to rehabilitation protocols according to the respondents. This finding is consistent with earlier studies showing that patient expectations significantly influence rehabilitation outcomes (13,14). Preoperative education about the rehabilitation process, emphasizing that the first weeks after surgery may be difficult and that the recovery might take a long time, could play a crucial role in managing expectations and improving overall satisfaction and outcomes for the patients. Future studies may consider exploring the effects of pre-operative information on post-surgery outcomes such as pain, health-related quality of life, and psychological distress. Finding the balance between loading the knee and recovery appears to be a common struggle for patients, particularly in the early stages of rehabilitation. The respondents in our study emphasized that patients tend to either overdo physical activity or remain too cautious, both of which can hinder recovery. The respondents were fairly consistent in their recommendations that the patients should limit their walking distances and use walking aids early in the rehabilitation to prevent overload and knee swelling. The ability to walk without limitation was identified in a systematic review as one of the most important aspects of recovery for patients after a KR (25). Regaining symmetrical walking without limping was also expressed as important by the respondents in our study but it was also acknowledged as a major challenge for the patients. Other important recovery aspects identified in the previously mentioned systematic review were pain reduction, returning to activities of daily living, recreational activity, and knee ROM (25). Several of these factors that were considered important in the recovery process were also identified as challenges in our study.

Low adherence to recommendations and exercises among the patients may lead to a worse outcome of rehabilitation according to the respondents. This is not unique to this group of patients and is especially common in home-based exercise (26,27). The respondents emphasized that the exercises needed to be repeated several times per day, especially in the early phase of rehabilitation to regain knee ROM and reduce swelling. Previous research has identified self-motivation as one of the predicting factors of adherence, which was also reported in our study (27). A strong therapeutic alliance between the patient and physiotherapist as well as continuous support from physiotherapist were highlighted as important for adherence in our study and have been previously described as factors that may enhance adherence to exercise (24). Patients who are well-informed and understand the importance of doing the exercises might have more successful outcomes (28). Of the patient difficulties reported, regaining knee ROM was the most mentioned factor. Regaining knee ROM early was repeatedly emphasized as important by the respondents. There is reason to place emphasis on this, as previous studies have shown an association between early (six-weeks) post-operative knee ROM and self-reported satisfaction with knee ROM (29). The maximum gain of knee ROM is achieved within twelve weeks post-operatively according to an observational study (10). Thus, to make the most of the possibilities for improvement within this time, an early post-operative physiotherapy contact is recommended.

Regaining muscle strength, walking without limping, and returning to physically demanding work and activities were, according to the respondents, challenges that the patients were faced with later in the rehabilitation process. The respondents expressed that the patients often had difficulties in regaining a normal gait pattern because they had been limping for a long time before the KR. These results are in line with the results from a recent systematic review and meta-analysis reporting that lower speed, stride length, cadence, and longer stance phase on the non-operated leg were detected in patients at least six months after KR (30). Difficulties in regaining quadriceps strength are common and may be an important reason for functional impairment and the long-term alterations in gait (11,31) which was also expressed by the respondents in our study. Although the recommended period of sick leave after KR for patients with physically demanding work in Sweden is up to six months on a full-time basis, followed by partial sick leave (32), the respondents expressed that many patients found it difficult to return to physically demanding work or activities that required sufficient muscle strength. This result is in line with the results from a Swedish registry study reporting that long sick leave is common after KR and that 12–17% were still sick-listed one year after the surgery (33). Individuals at risk of long-term sick leave should be identified at an early phase to facilitate appropriate intervention at an early stage. A recent study has suggested that the *Work, Osteoarthritis, or Joint Replacement questionnaire* score at three-month follow-up could be used to predict the ability to return to work-related activities (34).

Strengths and limitations

One strength of this study is the combination of closed- and open-ended questions which allows the respondents to fully reflect their views. There were a geographical spread of the respondents and most Swedish regions were represented. The use of an anonymous digital survey might have reduced the influence of social desirability bias since the respondents could feel secure that they wouldn't be identified (35). There are also weaknf responses were low. Siesses in this study that should be mentioned. Considering the (probably) several thousand physiotherapists working in primary care in Sweden, the number of responses was low. Since this was a public survey, we were not able to determine the response rate. However, the consistency in the responses suggests that the outcome would likely have remained unchanged, even with additional responses. Self-selection bias might have influenced the results since the respondents were not targeted but had a choice to use the link to the survey.

There is a risk that the questions did not cover all relevant and important aspects of KR rehabilitation such as patient education and self-management. Furthermore, despite that the alternatives for rehabilitation modalities were based on the authors' and patient representative's experiences, we might have omitted important modalities. We also did not specify our definition of different treatment modalities in the survey. Strength training especially may encompass

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

different definitions among physiotherapists which may lead to misclassification of treatment modalities. Lastly, this study focuses on physiotherapists' perceptions of patients' challenges after KR surgery, and future studies in the same project will reveal if this agrees with the patients' experiences.

Conclusion

Physiotherapists in primary care in Sweden prioritized knee ROM, strength training, and stationary cycling in rehabilitation after KR. These treatment modalities were mainly in line with those recommended in previous research. Handling pain and knee swelling, regaining function and trust in the knee, balancing load/recovery, and resuming physically demanding activities and work were among the challenges that patients faced during rehabilitation according to the physiotherapists. Sufficient pre-operative information, effective pain management, adherence to exercise, and accepting the situation were seen as important factors to overcome the challenges.

Disclosures

Conflict of interest: The authors declare no conflict of interest.

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Moseng T, Vliet Vlieland TPM, Battista S, et al. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis: 2023 update. Ann Rheum Dis. 2024;83(6):730-740. CrossRef PubMed
2. Conley B, Bunzli S, Bullen J, et al. Core Recommendations for Osteoarthritis Care: A Systematic Review of Clinical Practice Guidelines. Arthritis Care Res (Hoboken). 2023;75(9):1897-1907. CrossRef PubMed
3. Shichman I, Roof M, Askew N, Nherera L, Rozell JC, Seyler TM, et al. Projections and Epidemiology of Primary Hip and Knee Arthroplasty in Medicare Patients to 2040-2060. JBJS Open Access. 2023 Feb 28;8(1):e22.00112. CrossRef
4. Hu H, Li P, Liu Z, Lv H, Yang X, Liu P. Comparative long-term outcomes of unicompartmental and total knee arthroplasty in knee osteoarthritis patients: a systematic review and meta-analysis. Front Surg. 2024;11:1405025. CrossRef PubMed
5. Pilot P, Bogie R, Draijer WF, Verburg AD, van Os JJ, Kuipers H. Experience in the first four years of rapid recovery; is it safe? Injury. 2006;37(suppl 5):S37-S40. CrossRef PubMed
6. Bunzli S, O'Brien P, Klem N, et al. Misconceived expectations: patient reflections on the total knee replacement journey. Musculoskeletal Care. 2020;18(4):415-424. CrossRef PubMed
7. Skogö Nyvang J, Hedström M, Iversen MD, Andreassen Gleissman S. Striving for a silent knee: a qualitative study of patients' experiences with knee replacement surgery and their perceptions of fulfilled expectations. Int J Qual Stud Health Well-being. 2019;14(1):1620551. CrossRef PubMed
8. Boekesteijn R, Smolders J, Busch V, Keijsers N, Geurts A, Smulders K. Objective monitoring of functional recovery after total knee and hip arthroplasty using sensor-derived gait measures. PeerJ. 2022;10:e14054. CrossRef PubMed
9. Smith TO, Dainty JR, Clark EM, Whitehouse MR, Price AJ, MacGregor AJ. Demographic and geographical variability in physiotherapy provision following hip and knee replacement. An analysis from the National Joint Registry for England, Wales, Northern Ireland and the Isle of Man. Physiotherapy. 2020;106:1-11. CrossRef PubMed
10. Mehta S, Rigney A, Webb K, et al. Characterizing the recovery trajectories of knee range of motion for one year after total knee replacement. Physiother Theory Pract. 2020;36(1):176-185. CrossRef PubMed
11. Singla R, Niederer D, Franz A, et al. The course of knee extensor strength after total knee arthroplasty: a systematic review with meta-analysis and -regression. Arch Orthop Trauma Surg. 2023;143(8):5303-5322. CrossRef PubMed
12. Krysa JA, Ho C, O'Connell P, Pohar Manhas K. Clinical practice recommendations for prehabilitation and post-operative rehabilitation for arthroplasty: A scoping review. Musculoskeletal Care. 2022;20(3):503-515. CrossRef PubMed
13. Konnyu KJ, Thoma LM, Cao W, et al. Rehabilitation for Total Knee Arthroplasty: A Systematic Review. Am J Phys Med Rehabil. 2023;102(1):19-33. PubMed
14. W-Dahl A. Kärrholm J, Rogmark C, Johansson O, Ighani Arani P, Mohaddes M, et al. Annual report 2024. The Swedish Arthroplasty Register. Online (Accessed November 2024)
15. National system for knowledge-driven management within Swedish healthcare. 2024. Online (Accessed November 2024)
16. Nationelltsystem för kunskapsstyrning Hälso- och sjukvård. Generiskt-ramverk-kunskapsstod-v1.6-2024-06-26.pdf. Online (Accessed November 2024)
17. Wainwright TW, Gill M, McDonald DA, et al. Consensus statement for perioperative care in total hip replacement and total knee replacement surgery: Enhanced Recovery After Surgery (ERAS®) Society recommendations. Acta Orthop. 2020;91(1):3-19. CrossRef PubMed
18. Socialstyrelsen.About the Swedish healthcare system. 2020. Online (Accessed November 2024)
19. Anell A, Glenngård AH, Merkur S. Sweden health system review. Health Syst Transit. 2012;14(5):1-159. PubMed
20. Swedish Association of Physiotherapists. Physiotherapy in primary care. Online (Accessed November 2024)
21. Krippendorff K. Content Analysis: An Introduction to Its Methodology. SAGE Publications, Inc.; 2019. Online (Accessed November 2024)
22. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today. 2004;24(2):105-112. CrossRef PubMed
23. Feng, X., & Behar-Horenstein, L. Maximizing NVivo Utilities to Analyze Open-Ended Responses. The Qualitative Report. 2019;24(3): 563-571 CrossRef
24. Makimoto K, Fujita K, Konno R. Review and synthesis of the experience of patients following total hip or knee arthroplasty in the era of rapidly decreasing hospital length of stay. Jpn J Nurs Sci. 2020;17(4):e12361. CrossRef PubMed
25. Khatri C, Ahmed I, Dhaif F, et al. What's important for recovery after a total knee replacement? A systematic review of mixed methods studies. Arch Orthop Trauma Surg. 2024;144(5):2213-2221. CrossRef PubMed
26. Moore AJ, Holden MA, Foster NE, Jinks C. Therapeutic alliance facilitates adherence to physiotherapy-led exercise and physical activity for older adults with knee pain: a longitudinal qualitative study. J Physiother. 2020;66(1):45-53. CrossRef PubMed
27. Essery R, Geraghty AWA, Kirby S, Yardley L. Predictors of adherence to home-based physical therapies: a systematic review. Disabil Rehabil. 2017;39(6):519-534. CrossRef PubMed
28. Collado-Mateo D, Lavín-Pérez AM, Peñacoba C, et al. Key Factors Associated with Adherence to Physical Exercise in

Patients with Chronic Diseases and Older Adults: An Umbrella Review. Int J Environ Res Public Health. 2021;18(4):2023. CrossRef PubMed
29. Stephens JD, Hurst JM, Morris MJ, Berend KR, Lombardi AV Jr, Crawford DA. Correlation Between Patient-Reported "Happiness" With Knee Range of Motion and Objective Measurements in Primary Knee Arthroplasty. J Arthroplasty. 2022;37(6S):S105-S109. CrossRef PubMed
30. Marino G, De Capitani F, Adamo P, Bolzoni L, Gatti R, Temporiti F. Long-term gait analysis in patients after total knee arthroplasty: A systematic review and meta-analysis. Gait Posture. 2024;113:75-98. CrossRef PubMed
31. Pua YH, Seah FJT, Clark RA, Lian-Li Poon C, Tan JWM, Chong HC. Factors associated with gait speed recovery after total knee arthroplasty: A longitudinal study. Semin Arthritis Rheum. 2017;46(5):544-551. CrossRef PubMed
32. Socialstyrelsen. Knäartros. 2021. Online (Accessed November 2024)
33. Stigmar K, Dahlberg LE, Zhou C, Jacobson Lidgren H, Petersson IF, Englund M. Sick leave in Sweden before and after total joint replacement in hip and knee osteoarthritis patients. Acta Orthop. 2017;88(2):152-157. CrossRef PubMed
34. van Zaanen Y, Hoozemans MJM, Kievit AJ, et al; Expect TO work Consortium. Predictive Validity of the Work, Osteoarthritis, or Joint Replacement Questionnaire for Return to Work After Total Knee Arthroplasty: A 12-Month Multicenter Prospective Cohort Study. J Arthroplasty. 2025;40(3):625-631. CrossRef PubMed
35. Larson RB. Controlling social desirability bias. Int J Mark Res. 2019;61(5):534-547. CrossRef

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 192-193
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3538
CORRESPONDENCE

Comments on: Impact of direct access on the quality of primary care musculoskeletal physiotherapy: a scoping review from a patient, provider, and societal perspective

Jean-François Kaux, Marc Schiltz, Patrick Linden, Geoffrey Brands, Bruno Schoonejans, Thierry Lejeune

The « Coupole » includes the Royal Belgian Society of Physical and Rehabilitation Medicine (RBSPRM), the Belgian Association of Medical Unions (Absym), and the Belgian Group of Specialists (GBS)

Dear Editor,

We read with interest the scoping review by Cattrysse et al. (1) on the impact of direct access (DA) to physiotherapy for musculoskeletal disorders. We acknowledge the authors' efforts but would like to share our concerns about several critical issues.

This review includes five systematic reviews with a significant risk of bias (cf ROBIS tool), without meta-analysis, and four primary studies, none of which are controlled, randomized prospective clinical trials, published in low-impact and/or suspected predatory journals. Most report short-term outcomes, making the level of evidence for DA conclusions quite low.

The first step in treating musculoskeletal disorders is establishing a diagnosis, requiring extensive medical training to distinguish serious disease from common disorders. Moreover, DA assumes that physiotherapy is the best therapeutic option for musculoskeletal disorders, which is not always the case. Some conditions require alternative or complementary treatments that only a physician can properly assess and prescribe. A holistic patient assessment is essential to avoid misdiagnosis or delayed diagnosis, which can have serious consequences. In many countries, only doctors are trained and authorized to diagnose. However, their training in musculoskeletal disorders—often seen by general practitioners—should probably be improved. Physiotherapists' training varies widely by country and even within countries, with some acquiring musculoskeletal skills through additional training, leading to significant variability in their diagnostic and management abilities. The authors highlight red and yellow flags in low back pain but do not sufficiently address misdiagnosis risks. (2) Red flags were developed to help doctors identify rare symptomatic low back pain among common cases, but they are not diagnostic criteria, have poor sensitivity and specificity, and do not replace clinical reasoning. (3,4)

This study does not consider key contextual factors such as health system structure, economy, availability of health professionals, training disparities, and medical/paramedical practice regulations. Care organization is also crucial: do disciplines work in isolation or within structured networks? Thus, clinical trial results from one country should not be extrapolated to another or used for policy decisions without careful consideration.

Moreover, many healthcare systems lack an independent regulatory body for physiotherapy, unlike the Medical Council governing physicians. This absence raises concerns about clinical practice monitoring, professional standards enforcement, and handling of medical errors or complaints. Any DA consideration should include creating or strengthening such regulatory bodies to ensure patient safety and professional accountability.

The review suggests DA physiotherapy may reduce consultation, imaging, and medication costs but lacks rigorous economic analysis. No health economist contributed, and hidden costs (e.g., delayed diagnoses or inappropriate treatments) were overlooked. Additionally, claims that no significant adverse events were reported are weakened by limited patient safety data on diagnostic errors or missed serious conditions. Without robust safety data, stating that DA physiotherapy is as safe as physician-led care is premature.

In conclusion, while this review presents potential DA benefits, its limitations cannot be ignored. High-quality clinical and economic studies are needed to assess DA risks and benefits properly. A cautious, well-regulated approach prioritizing interdisciplinary collaboration is essential to ensure safe, effective patient care. (5)

Received: March 16, 2025
Accepted: June 25, 2025
Published online: July 14, 2025

Corresponding author:
Jean-François Kaux
email: jfkaux@uliege.be

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

Disclosures

Conflict of interest: The authors declare no conflict of interest.

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Cattrysse E, den Broeck JV, Petroons R, et al. Impact of direct access on the quality of primary care musculoskeletal physiotherapy: a scoping review from a patient, provider, and societal perspective. Arch Physiother. 2024;14:20-28. CrossRef PubMed
2. Bourassa M, Kolb WH, Barrett D, et al. Guideline adherent screening and referral: do third year Doctor of Physical Therapy students identify red and yellow flags within descriptive patient cases? a United States based survey study. J Man Manip Ther. 2023;31(4):253-260. CrossRef PubMed
3. Tsiang JT, Kinzy TG, Thompson N, et al. Sensitivity and specificity of patient-entered red flags for lower back pain. Spine J. 2019;19(2):293-300. CrossRef PubMed
4. Hartvigsen J, Hancock MJ, Kongsted A, et al.; Lancet Low Back Pain Series Working Group. What low back pain is and why we need to pay attention. Lancet. 2018;391(10137):2356-2367. CrossRef PubMed
5. Nijs J, Kosek E, Chiarotto A, et al. Nociceptive, neuropathic, or nociplastic low back pain? The low back pain phenotyping (BACPAP) consortium's international and multidisciplinary consensus recommendations. Lancet Rheumatol. 2024;6(3):e178-e188. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 194
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3666
CORRESPONDENCE

Authors' reply to: "Comments on: Impact of direct access on the quality of primary care musculoskeletal physiotherapy: a scoping review from a patient, provider, and societal perspective"

Erik Cattrysse[1], Jona Van den Broeck[1], Robin Petroons[1], Amber Teugels[1], Aldo Scafoglieri[1], Emiel van Trijffel[1,2]

[1]Experimental Anatomy Research Group (EXAN), Department of Physiotherapy, Human Physiology and Anatomy (KIMA), Faculty of Physical Education and Physiotherapy, Vrije Universiteit Brussel, Brussels - Belgium
[2]Ziekenhuisgroep Twente, ZGT Academy, Almelo - Netherlands

Thank you for your detailed reading of our scoping review titled "Impact of direct access on the quality of primary care musculoskeletal physiotherapy: a scoping review from a patient, provider, and societal perspective." We appreciate the critical engagement and the opportunity to respond to the points raised in both your letter and the accompanying commentary.

It is acknowledged that careful assessment of study selection and quality is fundamental in any review process. While scoping reviews differ from systematic reviews in their objectives and methodology, we recognize the importance of ensuring methodological rigor and transparency. The concern regarding the perceived low impact or potentially predatory nature of the included studies warrants clarification. Of the five systematic reviews incorporated in our study, four were published in Q1-ranked journals and one in a Q2 journal. The single studies were primarily drawn from Q2 and Q3 journals, with one published in a Q1 journal. While there was variation in study quality—as assessed using the ROBIS tool—this was transparently reported and critically appraised in the review. As a scoping review, our objective was to map the available evidence on direct access (DA) in musculoskeletal physiotherapy rather than to assess intervention effectiveness through meta-analysis. The application of PRISMA-ScR guidelines facilitated comprehensive and transparent reporting, including limitations inherent in the included literature.

Regarding the diagnostic role of physiotherapists and the issue of patient safety, we share your view that diagnosing musculoskeletal disorders (MSKDs) is a complex and critical task requiring appropriate training. Our intention was not to suggest that physiotherapists replace physicians in this role. Instead, our review sought to describe how DA is currently implemented in different contexts, where physiotherapists typically operate within well-defined scopes of practice and often in collaboration with medical professionals. While red and yellow flags are indeed not diagnostic tools in themselves, we referenced them as part of structured screening approaches used in clinical reasoning. This aspect of practice was explored in several studies we included, and we appreciate the opportunity to clarify this point.

The issue of heterogeneity in training and healthcare systems was also addressed in our discussion. We acknowledge that such variability limits the generalizability of findings and emphasized this point by cautioning against broad extrapolation. The review aimed to synthesize and present the current body of evidence on DA while highlighting the importance of contextual adaptation in clinical and policy applications.

On the subject of regulatory frameworks, we agree that effective governance is crucial to ensuring patient safety and professional accountability. While an in-depth evaluation of regulatory mechanisms fell outside the scope of our review, the importance of professional oversight and interprofessional collaboration was noted, and World Physiotherapy was cited as a global body exemplifying these principles. This remains a key area for further investigation.

Concerns regarding the paucity of comprehensive economic evaluations and limited safety data are well-founded. Our findings similarly indicated a lack of robust data in these domains. Consequently, the review called for more rigorous studies, particularly those involving health economic analyses and systematic monitoring of adverse outcomes, in order to inform policy and practice more effectively.

In summary, we acknowledge the value of continued scholarly dialogue in this evolving field. While our scoping review aimed to map the current state of knowledge regarding DA in MSK physiotherapy, it also identified critical gaps and areas for further inquiry. Your feedback contributes meaningfully to that process, and we thank you for your engagement.

On behalf of the authors.

Disclosures

Conflict of interest: The authors declare no conflict of interest.

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Received: June 26, 2025
Accepted: June 26, 2025
Published online: July 14, 2025

Corresponding author:
Erik Cattrysse
email: Erik.cattrysse@vub.be

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

Arch Physioter 2025; 15: 195-205
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3360
ORIGINAL RESEARCH ARTICLE

What do people search online about sciatica? What answers do they get from Dr. Google? A mixed analysis of Italian web-based data

Massimo Esposto[1], **Matteo Cioeta**[1], **Giuseppe Giovannico**[1], **Valerio Barbari**[2], **Alessandra Arca**[2]

[1]Department of Medicine and Health Science "Vincenzo Tiberio", University of Molise, c/o Cardarelli Hospital, C/da Tappino, Campobasso - Italy
[2]Department of Human Neurosciences, Sapienza University of Rome, Rome - Italy

ABSTRACT

Introduction: Spine-related leg pain (SRLP), also commonly referred to as sciatica, is characterized by many unanswered questions. Nowadays, patients search for information online to find answers not provided by clinicians. Therefore, this work aimed to (i) collect the main questions and keywords searched by the Italian general population related to SRLP, (ii) analyze the content of the most clicked web pages, (iii) investigate its alignment with evidence-based recommendations, and (iv) assess their credibility and (v) readability.

Methods: SEMrush was used to collect questions, keywords, search volumes, and web pages. Credibility was assessed through the QUality Evaluation Scoring Tool (QUEST). Readability was evaluated by the Gulpease index. Thematic and content analyses were performed.

Results: Monthly search volumes ranged from 4,400 to 33,100. QUEST scores ranged from 1 to 14 out of 28. Readability scores ranged from 20 to 47 out of 100. Common themes highlighted the roles of medications and physiotherapy, the value of doing effortless activities, and the importance of sleep. Content analysis showed that 11 codes (41%) were aligned with, 7 (26%) were contradicted by, and 26 (96%) were entirely omitted from at least one web page.

Conclusion: The findings suggest that inaccurate, poorly credible, and difficult-to-read information on SRLP are common online, leading to plausible targets for the educational intervention of clinicians dealing with SRLP patients. Clinicians are now more aware of the questions that their patients could be asking to Google and the answers it provides about SRLP.

Keywords: Disinformation, Public Health, Health Education, Health Literacy, Consumer Health, Sciatica

What is already known about this topic?

- *Patients with sciatica reportedly need clear and simple information. Consequently, they might look for information online. Furthermore, patients generally feel the need to be guided by their clinicians through the information retrieved on the Internet.*

What does the study add?

- *Inaccurate, poorly credible, and difficult to read information on sciatica are common online. Clinicians are now more aware of the questions and answers retrieved on Google. This knowledge will help clinicians tailor their educational intervention.*

Introduction

Sciatica is an old and umbrella term that can refer to any back or leg pain (1,2). However, in most cases, it usually refers to pain that radiates from the glute and lower back area to the lower leg and foot (3). Nevertheless, due to its heterogeneity, research studies found it difficult to provide clinicians and patients with a clear picture of its incidence, prevalence, prognosis, and most effective treatment options (4). For this and other reasons, it was recently suggested to abandon the term sciatica and replace it with spine-related leg pain (SRLP) (4). While older investigations suggested its incidence to be 5 cases per 1000 people (5), a more recent one reported a mean of about 20 cases per 1000 people per year from 2013 to 2021 in Dutch general practice (6). The reported prevalence varied from 1.2% to 43% when systematically considering all the studies published until 2006 (7), whereas recent investigations resulted in an 11% and 10% prevalence among patients presenting to their Denmark and Italian general practices, respectively (8,9). Although the prognosis is generally favorable, high-quality evidence

Received: October 28, 2025
Accepted: July 7, 2025
Published online: July 22, 2025

This article includes supplementary material

Corresponding author:
Massimo Esposto
email: massimo.esposto.physio@gmail.com

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

for this is lacking as well, as it is challenging to explore these aspects due to both heterogeneity and ethical issues (10,11).

A globally recognized and less dubious feature of SRLP, instead, is the suffering and psychological distress that characterize the experiences and reduced quality of life of these patients (12–14). Indeed, patients with SRLP show significant associations between their symptoms and psychological factors (e.g., fear, anxiety, catastrophizing) as well as coping strategies (e.g., isolation, acting out defensive style) resulting in a decreased quality of life (15,16). Qualitative research so far has consistently shown that one of the main needs of these patients is the provision of clear and simple information about their condition (17,18).

Nowadays, one of the most common and accessible strategies to gather relevant information is to search specific keywords and questions online to find the answers that science and healthcare professionals are not always able to provide the patient with. The easiest and most widely used web browser for finding such answers is Google. (19). Previous works showed that approximately 50% of patients presenting with various conditions have navigational needs (i.e., the need to seek out additional types of online health-related information and from a greater variety of sources) (20,21). Furthermore, a common theme among patients is their expectation to be guided towards the highest quality information by their healthcare professionals (22). Indeed, research suggested that not only did patients with SRLP align with these behaviors, but some were not even aware of the need to critically appraise online content (23). Unfortunately, Google's websites do not necessarily report credible and easily comprehensible information (24,25). In fact, recent evidence related to tendons and low back pain highlighted the need for higher-quality online health-related content (26,27).

For these reasons, it is essential for clinicians to be aware of the main questions asked by the general population about SRLP and, even more importantly, understand what information patients – among other people – may have already encountered on websites when they come to the clinician's clinic. To the authors' knowledge, such an analysis of web-based data, specifically concerning SRLP, has never been performed in Italian or in any other language. Additionally, the existing literature so far has not thoroughly investigated these themes for several musculoskeletal disorders.

Considering these premises, the aims of this work were to (i) collect the main questions and keywords searched by the Italian general population concerning SRLP, (ii) analyze the content of the most clicked web pages for each of the question found previously, and (iii) assess credibility and (iv) readability of these web pages. Following the reviewers' suggestions, a post-hoc analysis was conducted to investigate whether the content of the most clicked web pages aligned with evidence-based recommendations.

Methods

This study consists of cross-sectional analyses of web-based data concerning SRLP. It followed the methodology of the analyses performed by Rio et al. and Hauber et al. on tendons and low back pain, respectively (26,27). However, details and significant differences are outlined in the following paragraphs.

Where applicable, the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) guideline was followed to report data (28). See the Supplementary File 1 for further details.

Questions and search terms

SEMrush (Inc., Boston, MA, 2008) was used to collect the most frequently searched questions and keywords about SRLP by the Italian general population (29). SEMrush is an online platform that offers a range of tools for marketing research and analysis. It uses a machine learning algorithm that allows it to collect accurate data regardless of the searcher's location and history of searches. The search was run on August 07th, 2024 through the "Keyword Magic Tool" for the term "sciatica". This latter term was preferred to others due to its popularity among the Italian general population and to a higher total volume of searches, proved by a preliminary analysis, compared to other similar keywords.

Keywords and questions were collected and ordered according to their volume of search. Keywords significantly redundant with each other were excluded, and the first ten were listed and reported along with their search volume. Questions asked to Google followed the same procedure: duplicates or significantly similar questions were removed, and the first five were listed and reported. For each of these questions, the five most consulted web pages were collected. Figure 1 outlines the research process.

Credibility

Credibility of every web page was assessed through the QUality Evaluation Scoring Tool (QUEST) by two independent authors (ME, AA) and disagreements were solved through consensus (30). In addition, the authors reached out to the developers of QUEST and obtained further information on how to accurately interpret its domains. This tool was selected due to its reliability values, ease of use, and ability to assess important domains not considered by other tools (e.g., JAMA benchmark).

The QUEST tool involves 7 items investigating authorship, attribution, conflict of interest, currency, complementarity, and tones (30). Each item has its own distinct numerical evaluation scale in order to assign the weighted score that best describes the item, which must then be summed with the others for a total score ranging from 0 to 28 – higher scores are indicative of higher credibility of the online information. It was validated for the screening of health-related online information quality, showing high levels of intra-rater reliability (Cohen's $k = 0.7 - 1.0$; $p < .05$) (30).

Readability

Readability of each web page was assessed online through the Gulpease index (31) autonomously by a single author (ME). This index is automatically calculated using a fixed formula that takes into account multiple variables: the number of sentences, letters, and words. It consists of a score

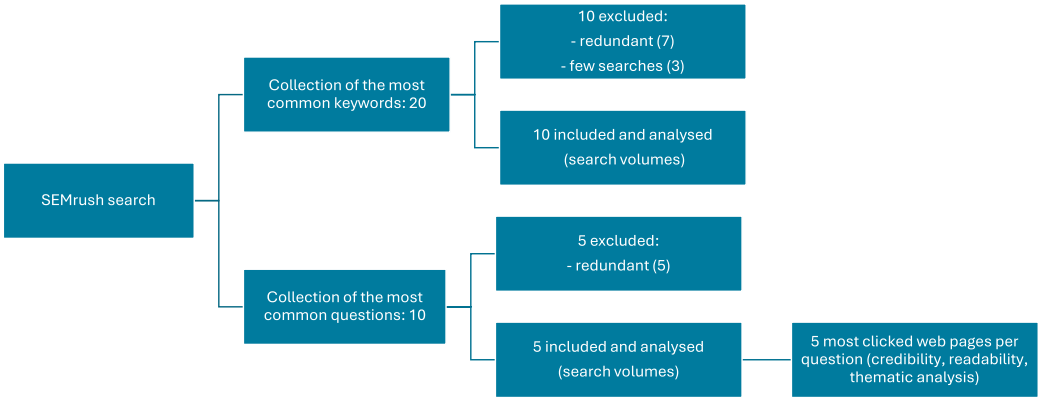


FIGURE 1 - Flowchart of the research procedure.

ranging from 0 to 100 where higher scores equal higher readability. Another output of the Gulpease index is the level of comprehensibility according to three different levels of the reader's educational background (i.e., elementary, medium and higher education).

This index was developed by Italian linguists at "La Sapienza" University of Rome specifically for the Italian language, and its validation was achieved through the analysis of several texts' comprehensibility performed by people with different levels of education (32).

Qualitative analyses

Thematic analysis

The most consulted web pages were then qualitatively analyzed following the six steps and recommendations of the thematic analysis (33). Specifically, (i) the authors familiarized themselves with the texts of all the web pages related to a single question and identified potential items of interest; (ii) initial codes were generated to categorize data relevant to answering the question, and codes across different web pages began to emerge; (iii) major and broad themes were created to identify patterns among web pages; (iv) initial themes were reviewed and modified based on their relevance to the initial question; (v) final themes were developed and named; (vi) a report was produced detailing these themes and how they helped answer the initial question. Themes, along with subthemes were found inductively and iteratively, as recommended. The relevant text of all web pages answering the same question was analyzed by one author without software support (ME) and by a second one (MC) supported by ATLAS.ti (34). Results were then compared and integrated (35). The same process occurred for each question.

Content analysis

A content analysis was conducted to determine whether the content of web pages aligned with current guidelines' recommendations. This analysis followed the established methodology guidelines (36,37). The document developed by the National Institute for Health and Care Excellence (NICE) was chosen due to its scientific rigour and plausible publication date (38) – the initial codebook was developed by one author (MC) to summarize diagnosis and treatment codes. Two other authors (ME, AA) supported this process by providing feedback and leading to the final document (Supplementary File 2). To enhance clarity and comprehensibility, the authors ensured that each code of the developed codebook related to one or more than one similar recommendations provided by the NICE guideline. Web pages were analyzed by one author (ME) who categorized each NICE code into "aligned", "misaligned", or "not reported". Online text that did not fit into any a-priori determined codes was excluded and not taken into consideration. The reporting of this analysis highlighted the total number and percentages of web pages aligning, misaligning, and not mentioning each code extracted by the NICE guidelines.

Results

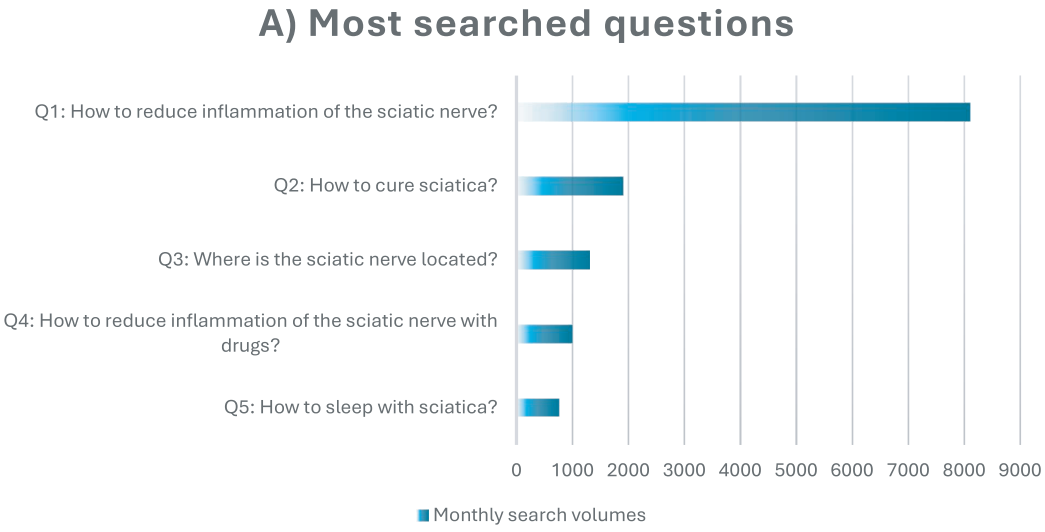
The five most searched questions on sciatica by the Italian general population are listed in Figure 2, *panel A*. Search volumes of questions ranged from 750 to 8100 monthly. The Uniform Resource Locators (URLs) of the five most clicked web pages for each of the five questions are outlined in the Supplementary File 3. The total number of web pages was 19, due to three of these being among the five most clicked web pages for more than one question.

The most searched keywords included remedies for sciatica (e.g., quick ones, homemade, exercise), ways to manage pain and inflammation both with and without drugs, sciatic symptoms, and the sciatic nerve itself (Fig. 2, *panel B*). Search volumes ranged from 4400 to 33100 per month.

Credibility

Credibility assessment of the five most consulted web pages per question resulted in total score values ranging from 1 to 14 out of 28 possible points. Notably, none of the web pages scored more than one on the *Attribution* item, meaning that none of the web pages referred to at least one identifiable scientific article. Conflicts of interest were high for 9 out of 19 different web pages, especially concerning *Question 5*. Ten different web pages supported the patient-physician relationship. None of the web pages showed a score higher than

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu



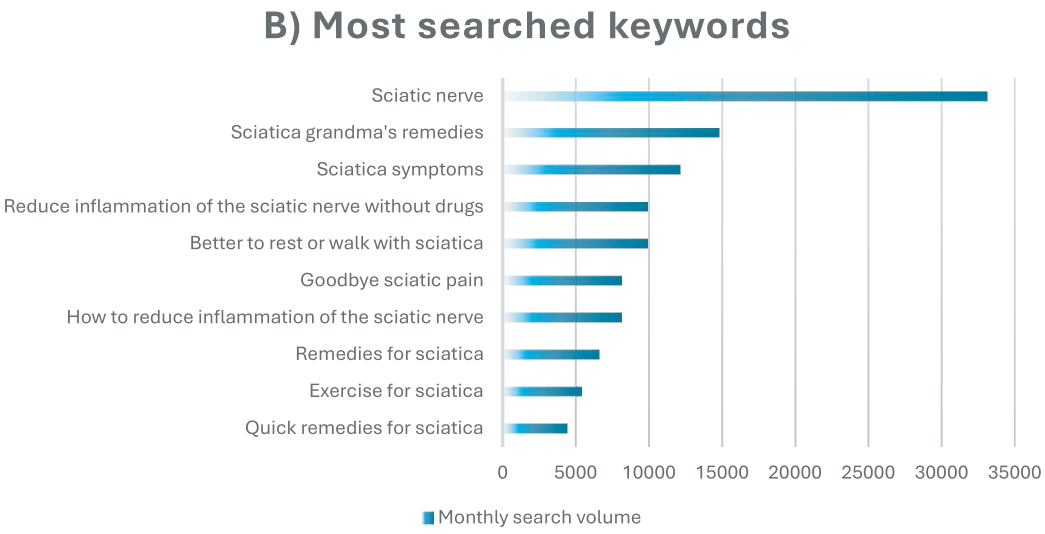


FIGURE 2 - Graphical representation of the most searched questions (panel A) and keywords (panel B) related to SRLP and their respective search volumes.

one on the *Tones* item and 15 of them resulted in authors supporting their claims without discussing the limitations but using cautious vocabulary (i.e., score of 1). See Table 1 for further details.

Readability

Readability assessment of the five most clicked web pages per question showed low Gulpease indexes (Table 2). Indexes ranged from 47 to 20, and it was estimated that people with an elementary educational background would find 100% (i.e., 25) of the web pages "*almost incomprehensible*". People with a medium education would find 21 out of 25 web pages "*very difficult*" to understand, with the remaining four rated as "*almost incomprehensible*". Subjects with a higher educational background would rate 16 web pages as "*easy*" to understand, 8 as "*difficult*", and one as "*very difficult*".

Qualitative analyses

Thematic analysis

Qualitative analysis of the most clicked web pages highlighted the following themes per question (Fig. 3): Question 1 included "Non-Steroidal Anti-Inflammatory Drugs (NSAIDs) first, stronger meds then", "Rest initially, move eventually but… avoid efforts with your back!", "Physiotherapy to stretch & mind your posture!", and "Explore homemade remedies"; Question 2 highlighted "Acute pain: rest (not too much) & meds", "Subacute pain & prevention: physiotherapy", "Last resort: surgery", and "Favorable prognosis: 6 weeks"; Question 3 resulted in "The longest nerve in the body", and "Many places, many functions"; Question 4 included "NSAIDs as first line therapy", and "Do not misuse meds"; Question 5 resulted in "As difficult as it is important", "There are ways to ease falling asleep", and "Different postures to sleep".

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

TABLE 1 - Assessment of credibility of online information through the QUality Evaluation Scoring Tool (QUEST) and its domains showed for each of the most clicked web pages per question

Web page	QUality Evaluation Scoring Tool (QUEST)							
	Authorship[a] (0-2)	Attribution[b] (0-3) x 3	Type of study[c,d] (0-2)	Conflict of interest[e] (0-2) x 3	Currency[f] (0-2)	Complementarity[g] (0-1)	Tones[h] (0-2) x 3	Total score (0-28)
Question 1: *How to reduce inflammation of the sciatic nerve?*								
1.a[i]	0	1	0	0	0	1	1	7
1.b	2	0	0	0	2	1	1	8
1.c[j]	0	0	0	2	0	0	1	9
1.d	2	0	0	1	2	0	0	7
1.e[k]	2	0	0	2	2	1	1	14
Question 2: *How to cure sciatica?*								
2.a	2	0	0	2	1	0	1	12
2.b	2	0	0	0	2	0	1	7
2.c[k]	2	0	0	2	2	1	1	14
2.d[i]	0	1	0	0	0	1	1	7
2.e	0	0	0	2	2	0	1	11
Question 3: *Where is the sciatic nerve located?*								
3.a[i]	0	1	0	0	0	1	1	7
3.b	2	0	0	1	2	0	1	10
3.c	2	1	0	1	2	1	0	13
3.d	0	0	0	0	0	1	1	4
3.e	0	0	0	0	0	1	1	4
Question 4: *How to reduce inflammation of the sciatic nerve with drugs?*								
4.a[i]	0	1	0	0	0	1	1	7
4.b	0	0	0	1	2	1	1	9
4.c[j]	0	0	0	2	0	0	1	9
4.d	0	0	0	1	2	1	1	9
4.e	0	0	0	1	2	1	0	6
Question 5: *How to sleep with sciatica?*								
5.a[i]	0	1	0	0	0	1	1	7
5.b	1	0	0	0	2	0	1	6
5.c	1	0	0	0	2	0	1	6
5.d	1	0	0	0	0	0	0	1
5.e	0	0	0	0	0	1	1	4

[a]0 – No indication of authorship or username; 1 – All other indications of authorship; 2 – Author's name and qualification clearly stated.
[b]0 – No sources; 1 – Mention of expert source, research findings (though with insufficient information to identify the specific studies), links to various sites, advocacy body, or other; 2 – Reference to at least one identifiable scientific study, regardless of format (e.g., information in text, reference list); 3 – Reference to mainly identifiable scientific studies, regardless of format (in >50% of claims).
[c]for articles scoring 2 or 3 on *Attribution*.
[d]0 – In vitro, animal models, or editorials; 1 – All observational work; 2 – Meta-analyses, randomized controlled trials, clinical studies.
[e]0 – Endorsement or promotion of intervention designed to prevent or treat condition (e.g., supplements, brain training games, foods) within the article; 1 – Endorsement or promotion of educational products & services (e.g., books, care home services); 2 – Unbiased information.
[f]0 – No date present; 1 – Article is dated but 5 years or older; 2 – Article is dated within the last 5 years.
[g]0 – No support of the patient-physician relationship; 1 – Support of the patient-physician relationship.
[h]0 – Fully supported (authors fully and unequivocally support the claims, strong vocabulary such as "cure", "guarantee", and "easy", mostly use of non-conditional verb tenses ("can", "will"), no discussion of limitations); 1 – Mainly supported (authors mainly support their claims but with more cautious vocabulary such as "can reduce your risk" or "may help prevent", no discussion of limitations); 2 – Balanced/cautious support (authors' claims are balanced by caution, includes statements of limitations and/or contrasting findings)
[i]same web page
[j]same web page
[k]same web page

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - Indexes of readability and level of comprehensibility for each of the most clicked web pages per question

Web page	Gulpease index	Elementary education	Medium education	Higher education
Question 1: *How to reduce inflammation of the sciatic nerve?*				
1.a[a]	43	Almost incomprehensible	Very difficult	Easy
1.b	38	Almost incomprehensible	Very difficult	Difficult
1.c[b]	42	Almost incomprehensible	Very difficult	Easy
1.d	37	Almost incomprehensible	Very difficult	Difficult
1.e[c]	41	Almost incomprehensible	Very difficult	Easy
Question 2: *How to cure sciatica?*				
2.a	45	Almost incomprehensible	Very difficult	Easy
2.b	32	Almost incomprehensible	Almost incomprehensible	Difficult
2.c[c]	41	Almost incomprehensible	Very difficult	Easy
2.d[a]	43	Almost incomprehensible	Very difficult	Easy
2.e	37	Almost incomprehensible	Very difficult	Difficult
Question 3: *Where is the sciatic nerve located?*				
3.a[a]	43	Almost incomprehensible	Very difficult	Easy
3.b	35	Almost incomprehensible	Almost incomprehensible	Difficult
3.c	42	Almost incomprehensible	Very difficult	Easy
3.d	47	Almost incomprehensible	Very difficult	Easy
3.e	36	Almost incomprehensible	Very difficult	Difficult
Question 4: *How to reduce inflammation of the sciatic nerve with drugs?*				
4.a[a]	43	Almost incomprehensible	Very difficult	Easy
4.b	46	Almost incomprehensible	Very difficult	Easy
4.c[b]	42	Almost incomprehensible	Very difficult	Easy
4.d	31	Almost incomprehensible	Almost incomprehensible	Difficult
4.e	20	Almost incomprehensible	Almost incomprehensible	Very difficult
Question 5: *How to sleep with sciatica?*				
5.a[a]	43	Almost incomprehensible	Very difficult	Easy
5.b	47	Almost incomprehensible	Very difficult	Easy
5.c	47	Almost incomprehensible	Very difficult	Easy
5.d	47	Almost incomprehensible	Very difficult	Easy
5.e	40	Almost incomprehensible	Very difficult	Difficult

[a]same web page; [b]same web page; [c]same web page

Answers to how one could cure (i.e., *Question 1*) or reduce the inflammation causing their sciatica, both with (i.e., *Question 4*) and without (i.e., *Question 2*) medications, focused on the use of NSAIDs, frequently alerting about their adverse effects, and on the role of stronger medications in cases of severe and persistent symptoms: "*NSAIDs like ibuprofen and naproxen can help decrease the inflammation. When pain persists, stronger medications like corticosteroids and opioids could be prescribed. However, […] their long-term use is not recommended*".

Another key point discussed by many web pages was the role of rest and physical activity. Most of these advocated in favor of gentle activities as soon as possible, to decrease the negative effects of inactivity: "*Unlike what is commonly thought, if you have sciatica, it is not recommended to spent days in the bed: while it can help decreasing your pain, it quickly backfires and hampers the healing and recovery processes*".

Furthermore, the role of physiotherapy practices was commonly discussed. While some suggested starting with physiotherapy in a subacute phase, others recommended to wait until the symptoms were solved and as a means of prevention: "*When the pain is too severe, it is better to avoid manipulations, massages, or exercises that could move some vertebras and worsen the symptoms. […] Once pain-free, instead, exercising helps to prevent new attacks*".

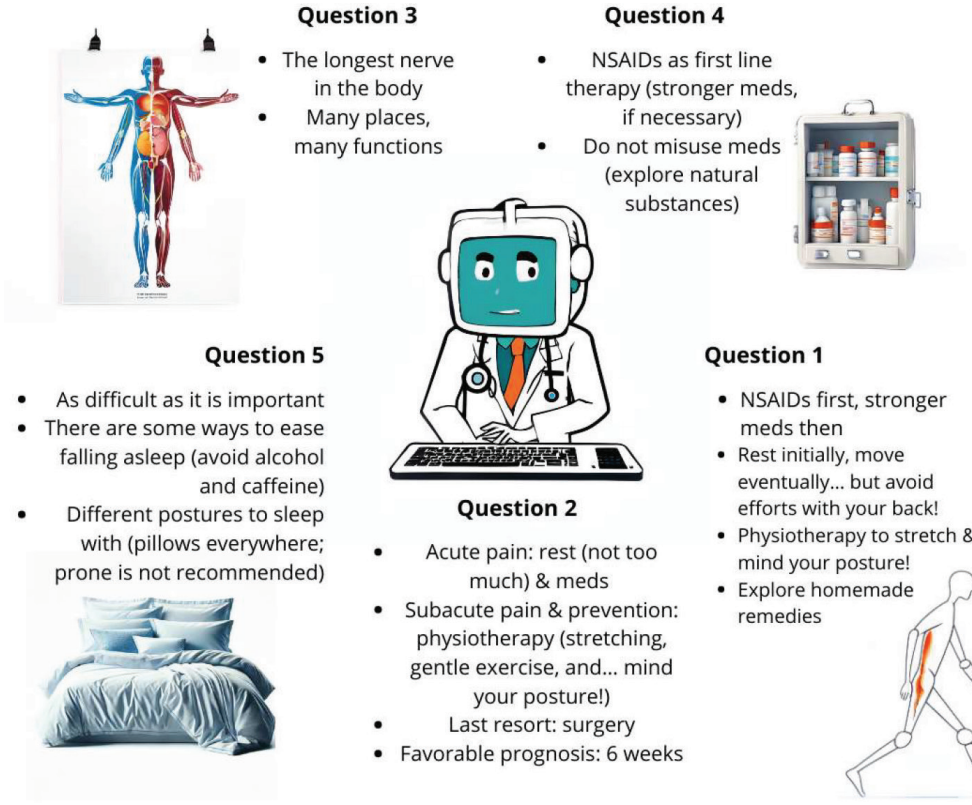


FIGURE 3 - Representation and list of all the main themes per question: Question 1, "How to reduce inflammation of the sciatic nerve?"; Question 2, "How to cure sciatica?"; Question 3, "Where is the sciatic nerve located?"; Question 4, "How to reduce inflammation of the sciatic nerve with drugs?"; Question 5, "How to sleep with sciatica?".

However, all the web pages were consistent with the suggestion of avoiding certain postures, exercises, and efforts as these would be bad for the patients' backs and likely lead to further damage: "*[…] stretching and soft activities are very helpful, and use your legs instead of your back when lifting weights to dampen the weight and keep the right posture*".

Given the favorable prognosis, about which all web pages agreed on, the patient was recommended to wait before considering more invasive options of treatment (e.g., surgery) and to look out for specific signs and symptoms suggesting a malignant pathology (e.g., tumour, cauda equina) and requiring urgent investigations (e.g., imaging tests).

Answers to "where the sciatic nerve is located" (i.e., *Question 3*) were similar among web pages and provided in-depth explanations of its course and functions throughout the leg.

Online information on how to sleep with sciatica (i.e., *Question 5*) highlighted the invaluable role of sleep for one's general health and how frequently it is impaired by this condition: "*Sleep is crucial for both mental and physical health. However, falling asleep could be very hard if you have night pain and burning sensations from your back to your foot. […] A poor-quality sleep seems to negatively impact the perception of pain*".

Several postures and pillows' positionings were thoroughly described together with advice about what to do and not to do to fall asleep more easily: "*To improve your sleep quality, you can follow these suggestions: choose a high-quality mattress […] to help you keep a good posture at night […]; use extra pillows […]. Memory foam pillows […] allow you to relax your spine; relax and soften your muscles before going to sleep. […] with a hot shower […]. And stretching exercises […]; decrease the amount of alcohol and caffeine […]*".

Content analysis

Out of the twenty-seven extracted codes, eleven (41%) resulted in at least one web page reporting online content in line with guidelines' recommendations. The remaining sixteen (63%) did not show online content aligned with their recommendations, either because it was misaligned (four codes, 15%) or because it was not mentioning any recommendations (twelve codes, 44%). Three codes (i.e., #8, #11, and #26) showed at least one web page aligning and one misaligning to them, accounting for a total of seven codes (26%) contradicted by web pages. Twenty-six (96%) codes resulted in at least 78% of web pages not reporting any information about the content of their recommendation. Further details on the codes and web pages alignment can be found in the Supplementary File 2 and Table 3, respectively.

Discussion

To the best of the authors' knowledge, this is the first study investigating the questions and keywords searched online by the Italian general population related to SRLP. This web-based data analysis has not been performed in any other language or country so far. Additionally, this is the first work to thoroughly analyze credibility, readability, and content of the most clicked web pages.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 3 - Summary of the extracted codes and content analysis of the included web pages

"To do" recommendation	No (%) of aligned web pages	No (%) of misaligned web pages	No (%) of non-reporting web pages
Diagnosis			
1. Exclude specific causes of symptoms (e.g., cancer, infection, trauma, inflammatory diseases)	3 (15.8%)	0 (0%)	16 (84.2%)
2. Consider using risk stratification tools (e.g., STarT Back) to inform shared decision-making about stratified management	0 (0%)	0 (0%)	19 (100%)
3. Based on risk stratification, offer simpler and less intensive support for people likely to improve quickly, while offering more complex and intensive support for those at risk of poorer outcomes	0 (0%)	0 (0%)	19 (100%)
4. Explain that imaging may not be needed	0 (0%)	0 (0%)	19 (100%)
5. Consider imaging in specialist settings of care (e.g., musculoskeletal interface clinic or hospital) if the results are likely to change management	0 (0%)	0 (0%)	19 (100%)
Treatment			
6. Provide advice and information to help patients self-manage their symptoms	18 (94.7%)	0 (0%)	1 (5.3%)
7. Consider a group exercise program that accounts for people's specific needs, preferences and capabilities	2 (10.5%)	0 (0%)	17 (89.5%)
8. Consider manual therapy, only as part of a treatment package including exercise	1 (5.3%)	2 (10.5%)	16 (84.2%)
9. Consider psychological therapies, only as part of a treatment package including exercise	0 (0%)	0 (0%)	19 (100%)
10. Consider a combined physical and psychological program for people with persistent symptoms	1 (5.3%)	0 (0%)	18 (94.7%)
11. Promote and facilitate return to work and normal activities	3 (15.8%)	1 (5.3%)	15 (78.9%)
12. If the patient is already taking medicines, explain the risks of continuing them	0 (0%)	0 (0%)	19 (100%)
13. Be aware of the risk of harms and limited evidence of benefit from the use of non-steroidal anti-inflammatory drugs (NSAIDs)	1 (5.3%)	0 (0%)	18 (94.7%)
14. If prescribing NSAIDs, take into account the person's risk factors, the lowest effective dose, and the shortest possible period of time	0 (0%)	0 (0%)	19 (100%)
15. Consider epidural injections of local anaesthetic and steroid in people with acute and severe sciatica	1 (5.3%)	0 (0%)	18 (94.7%)
16. Consider spinal decompression when non-surgical treatment has not improved pain or function and their radiological findings are consistent with sciatic symptoms	3 (15.8%)	0 (0%)	16 (84.2%)
"Not to do" recommendation	**No (%) of aligned web pages**	**No (%) of misaligned web pages**	**No (%) of non-reporting web pages**
Diagnosis			
17. Do not routinely offer imaging in non-specialist settings	6 (31.6%)	0 (0%)	13 (68.4%)
Treatment			
18. Do not offer traction	0 (0%)	0 (0%)	19 (100%)
19. Do not offer rocker sole shoes	0 (0%)	0 (0%)	19 (100%)
20. Do not offer foot orthotics	0 (0%)	0 (0%)	19 (100%)
21. Do not offer ultrasound	0 (0%)	1 (5.3%)	18 (94.7%)
22. Do not offer percutaneous electrical nerve simulation (PENS)	0 (0%)	0 (0%)	19 (100%)
23. Do not offer transcutaneous electrical nerve simulation (TENS)	0 (0%)	1 (5.3%)	18 (94.7%)
24. Do not offer interferential therapy	0 (0%)	0 (0%)	19 (100%)
25. Do not offer belts or corsets	0 (0%)	2 (10.5%)	17 (89.5%)
26. Do not offer gabapentinoids, other antiepileptics, oral corticosteroids or benzodiazepines as there is no overall evidence of benefit and there is evidence of harm	1 (5.3%)	3 (15.8%)	15 (78.9%)
27. Do not offer opioids for managing persistent symptoms	0 (0%)	2 (10.5%)	17 (89.5%)

Searches performed by the Italian general population on Google yielded keywords and questions mainly concerning various ways to manage their SRLP (e.g., how to sleep, how to reduce inflammation with or without drugs). This was consistent with qualitative studies showing that patients perceived web resources as necessary for understanding and accessing treatment options (23). Another set of searches was related to the location of the sciatic nerve, possibly highlighting the concerns and fears arising from the experience of pain throughout the whole lower limb (18).

In line with previous research, credibility values were primarily low with none of the web pages scoring higher than 14 out of 28 possible points, as measured by the QUEST (Table 1) (24,25). Particularly concerning was that only two web pages referred to scientific research findings, albeit without providing sufficient information to identify the studies. The remaining web pages did not reference any external source. This finding, combined with substantial conflicts of interest and biased information, underscores the necessity of critically appraising online health-related content (25). Moreover, the challenges that patients face in this process emphasize the crucial role that healthcare professionals play, not only delivering updated information but also in guiding patients effectively through navigating online resources (22,25).

Another consistent finding across all web pages was their poor readability, with scores as low as 20 out of 100 possible points on the Gulpease scale (Table 2) (32). Notably, none of the web pages scored higher than 47 points, reinforcing the need for health-related content to be written more clearly (39,40). Indeed, the educational background analysis showed that only a limited number of web pages were classified as "easy" to read and understand, and even then, this was true only for individuals with higher educational levels. This finding aligns with other research studies in health-related fields (39,40), suggesting that the poor readability of online information may partially explain the relationship between lower educational attainment and poorer clinical outcomes among patients with SRLP (41–43).

The results coming from the content analysis suggested a persistent discrepancy between NICE guideline recommendations and the information most commonly featured on web pages (Table 3). Several core points (e.g., discouraging routine imaging, minimizing reliance on certain medications, risk stratification) were often undervalued, oversimplified, or even contradicted. Moreover, many web pages appeared to emphasize short-term strategies while offering minimal insight into interventions that guidelines strongly recommend, including consistent self-management support and exercise-based treatments. This mismatch reflects a broader – but not yet well explored – trend observed for other musculoskeletal conditions such as tendinopathies (unpublished data) and low back pain – online resources frequently lag behind current best practices and rarely provide nuanced, evidence-based perspectives. (26)

Notably, this work showed that people searching about sciatica online often do so in highly practical terms, focusing on immediate relief (e.g., "How to reduce inflammation," "How to cure sciatica,"). Web pages thus gravitated toward "user-friendly" suggestions like rest, NSAIDs, avoiding strenuous exercise, and trying homemade remedies, in a perpetuating vicious cycle. While this aligns with what musculoskeletal patients typically seek and expect (44), it also raises the more complex and widely debated question: "*Should healthcare professionals give patients what they want?*" (45). Indeed, across questions 1, 2, and 4 (on reducing inflammation with or without drugs and how to cure sciatica), the main themes included prioritizing medications (e.g., NSAIDs first, stronger meds if pain persists), some acknowledgment of physiotherapy (i.e., typically only after an initial rest phase), and repeated warnings to avoid "effort" with the back. While these short-term approaches may resonate with individuals seeking immediate symptom relief, they do not necessarily align with guideline-adherent practices that advocate for stratified care and exercise-based programs (38,46), thereby possibly influencing the answer to the aforementioned question (45).

Despite acknowledging a generally favorable prognosis, the online guidance of avoiding exertions or "bad" postures, and prioritizing pain relief at all costs, could inadvertently reinforce fear of "harmful" movement. This may potentially delay the return to daily activities or undermine the perceived importance of physiotherapy and routine activities during both the acute and subacute phases (47,48). In fact, established guidelines recommend swift but measured engagement in exercise and advocate watchful waiting before considering imaging, underscoring that routine imaging tests can be counterproductive unless their results are expected to alter clinical management (38,46). To address the widespread demand for and performance of unnecessary imaging, the Choosing Wisely Campaign was developed in the United States and subsequently adopted by several other countries (49,50). Recently, a Choosing Wisely Leaflet was developed in Denmark to assist patients in understanding why imaging tests might not be needed (51). Hopefully, this tool will be adopted by other European countries, including Italy, and lead to fewer unnecessary tests.

From an educational standpoint, our findings reinforce that health literacy interventions must address more than just readability and credibility – they also need to correct for content gaps or inaccuracies regarding guideline-based advice. Clinical implications, while secondary to the educational focus here, are nonetheless significant. Patients may arrive with entrenched beliefs caused by web pages promoting outdated practices (e.g., routinary imaging tests, ultrasounds, belts) or downplaying guideline-endorsed interventions. Addressing these misconceptions early could help align patient expectations with more effective, guideline-adherent management strategies, therefore hoping for better clinical outcomes (18,23).

Limitations and future directions

While this study has strengths such as strong methods (e.g., QUEST, Gulpease index), it also presents some limitations: (i) searches were performed in Italian, thereby limiting the interpretation of results to Italy and its general population; (ii) these findings relate to the Italian general population and they are not a specific collection of what patients with SRLP search online, even though these are likely to be the people who contribute the most to the overall searches; (iii) findings pertain to the month before the search was run on SEMrush, possibly leading

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

to an unknown degree of variability as time goes by; (iv) the study focused solely on Google, overlooking emerging technologies such as chatbots and artificial intelligence, which are gaining increasing traction – it also neglected social media, which is now widely used across all age groups and where pages related to health and rehabilitation topics can easily be found; and (v) this study did not assess the strength of the recommendations in its content analysis, thereby potentially limiting the interpretation of results related to aligning and misaligning web pages.

Considering the findings and their limitations, future research should replicate this study in other languages and analyze other online technologies (e.g., artificial intelligence, social media) to provide a broader picture of the online information that patients with SRLP may come across. Furthermore, categorizing web pages based on their source type (e.g., institutional websites, commercial platforms, blogs, or forums) could be beneficial, as it would allow healthcare professionals to offer patients targeted recommendations regarding the reliability and readability of specific online resources. Finally, it would be insightful to explore the thoughts and feelings of actual patients confronting online information and how these impact the relationship with their clinicians' advice and management program.

In light of our findings, a further future direction involves enhancing the role of scientific associations, public health institutions, and universities through proactive efforts aimed at developing and disseminating high-quality, up-to-date information on SRLP. For instance, collaborating with trusted patient-centric platforms might help ensure that credible, evidence-based material reaches broader audiences. Moreover, partnerships between healthcare professionals and universities could foster the creation of tailored educational resources (e.g., leaflets, videos, and interactive modules) that meet patients' literacy needs and reduce the prevalence of misleading or incomplete content.

Conclusions

Healthcare professionals are now more aware of the questions that their patients could be asking Google and the answers it provides about SRLP. Moreover, low levels of credibility, readability, and alignment with evidence-based recommendations of the most clicked web pages highlighted the need for clinicians to support patients with their understanding of the condition and its treatment options. A further and crucial step forward would include the commitment of healthcare professionals to share evidence-based information on the Internet, thus contributing to a more helpful health literacy of patients with SRLP and improving the outcomes of a shared decision-making process.

Disclosures

Conflicts of interest: The authors declare no conflicts of interest.

Financial support: This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Authors' contributions: ME: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing- Original draft preparation, Writing- Reviewing and Editing. MC: Data curation, Formal analysis, Methodology, Software, Visualization, Writing- Reviewing and Editing. GG: Methodology, Supervision, Visualization, Writing- Reviewing and Editing. VB: Methodology, Supervision, Visualization, Writing- Reviewing and Editing. AA: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing- Original draft preparation, Writing- Reviewing and Editing.

Data availability statement: The data presented in this study are available on request from the corresponding author.

References

1. Truumees E. A history of lumbar disc herniation from Hippocrates to the 1990s. Clin Orthop Relat Res. 2015;473(6):1885-1895. CrossRef PubMed
2. Lin CW, Verwoerd AJH, Maher CG, et al. How is radiating leg pain defined in randomized controlled trials of conservative treatments in primary care? A systematic review. Eur J Pain. 2014;18(4):455-464. CrossRef PubMed
3. Frymoyer JW, Cats-Baril WL. An overview of the incidences and costs of low back pain. Orthop Clin North Am. 1991;22(2):263-271. CrossRef PubMed
4. Schmid AB, Tampin B, Baron R, et al. Recommendations for terminology and the identification of neuropathic pain in people with spine-related leg pain. Outcomes from the NeuPSIG working group. Pain. 2022;164(8):1693-1704. PubMed
5. Cherkin DC, Deyo RA, Loeser JD, Bush T, Waddell G. An international comparison of back surgery rates. Spine. 1994;19(11):1201-1206. CrossRef PubMed
6. Khorami AK, Chiarotto A, de Schepper EIT, Bindels PJE, Koes BW, van den Driest JJ. The prevalence, incidence and management of low back pain with radiating leg pain in Dutch general practice: A population-based cohort study in the Rijnmond Primary Care Database. Eur J Pain. 2024;28(8):1378-1386. CrossRef PubMed
7. Konstantinou K, Dunn KM. Sciatica: review of epidemiological studies and prevalence estimates. Spine. 2008;33(22):2464-2472. CrossRef PubMed
8. Hartvigsen L, Hestbaek L, Lebouef-Yde C, Vach W, Kongsted A. Leg pain location and neurological signs relate to outcomes in primary care patients with low back pain. BMC Musculoskelet Disord. 2017;18(1):133. CrossRef PubMed
9. Polistina C. La diagnosi di lombalgia e lombosciatalgia in Medicina Generale: epidemiologia ed appropriatezza delle indagini diagnostiche. Health Search Istituto di Ricerca della SIMG. 2018;25(2):2-5. Online (Accessed October 2024)
10. Ostelo RWJG. Physiotherapy management of sciatica. J Physiother. 2020;66(2):83-88. CrossRef PubMed
11. Vroomen PCAJ, de Krom MC, Wilmink JT, Kester ADM, Knottnerus JA. Lack of effectiveness of bed rest for sciatica. N Engl J Med. 1999;340(6):418-423. CrossRef PubMed
12. Inoue S, Taguchi T, Yamashita T, Nakamura M, Ushida T. The prevalence and impact of chronic neuropathic pain on daily and social life: A nationwide study in a Japanese population. Eur J Pain. 2017;21(4):727-737. CrossRef PubMed
13. Ryan C, Roberts L. 'Life on hold': the lived experience of radicular symptoms. A qualitative, interpretative inquiry. Musculoskelet Sci Pract. 2019;39:51-57. CrossRef PubMed
14. Konstantinou K, Hider SL, Jordan JL, Lewis M, Dunn KM, Hay EM. The impact of low back-related leg pain on outcomes as compared with low back pain alone: a systematic review of the literature. Clin J Pain. 2013;29(7):644-654. CrossRef PubMed
15. Tutoglu A, Boyaci A, Karababa IF, et al. Psychological defensive profile of sciatica patients with neuropathic pain and its relationship to quality of life. Z Rheumatol. 2015;74(7):646-651. CrossRef PubMed

16. Ionescu D, Iacob CI, Brehar FM, Avram E. The role of catastrophizing and basic psychological needs satisfaction on health-related quality of life and pain in patients with lumbar disc herniation. Front Psychol. 2023;14:1147254. CrossRef PubMed
17. Ong BN, Konstantinou K, Corbett M, Hay E. Patients' own accounts of sciatica: a qualitative study. Spine. 2011; 36(15):1251-1256. CrossRef PubMed
18. Goldsmith R, Williams NH, Wood F. Understanding sciatica: illness and treatment beliefs in a lumbar radicular pain population. A qualitative interview study. BJGP Open. 2019;3(3): bjgpopen19X101654. CrossRef PubMed
19. Mohamed A, Ismail I. A performance comparative on most popular internet web browsers. Procedia Comput Sci. 2022;215: 589-597. CrossRef
20. Cocco AM, Zordan R, Taylor DM, et al. Dr Google in the ED: searching for online health information by adult emergency department patients. Med J Aust. 2018;209(8):342-347. CrossRef PubMed
21. Lee K, Hoti K, Hughes JD, Emmerton LM. Consumer use of "Dr Google": a survey on health information-seeking behaviors and navigational needs. J Med Internet Res. 2015;17(12):e288. CrossRef PubMed
22. Lee K, Hoti K, Hughes JD, Emmerton L. Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions. J Med Internet Res. 2014;16(12):e262. CrossRef PubMed
23. Ryan C, Pope CJ, Roberts L. Why managing sciatica is difficult: patients' experiences of an NHS sciatica pathway. A qualitative, interpretative study. BMJ Open. 2020;10(6):e037157. CrossRef PubMed
24. Ferreira G, Traeger AC, Machado G, O'Keeffe M, Maher CG. Credibility, accuracy, and comprehensiveness of internet-based information about low back pain: a systematic review. J Med Internet Res. 2019;21(5):e13357. CrossRef PubMed
25. Peterson S, Rainey N, Weible K. Who writes this stuff? Musculoskeletal information quality and authorship of popular health websites: A systematic review. Musculoskelet Sci Pract. 2022;60:102563. CrossRef PubMed
26. Hauber SD, Robinson K, Fechner R, Pate JW, O'Sullivan K. Credibility, readability and content analysis of treatment recommendations for adolescents with nonspecific back pain published on consumer websites. Eur J Pain. 2024;28(9):1571-1584. CrossRef PubMed
27. Rio E, Couch J, Perrot M, Ganderton C. Hey Google? When patients consult Dr Google about tendinopathy; an analysis of search terms and questions from four English speaking regions across six tendons. JOSPT Open. 2024;2(3):233-239. CrossRef
28. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370(9596):1453-1457. CrossRef PubMed
29. Semrush. 2008. Online (Accessed October 2024)
30. Robillard JM, Jun JH, Lai JA, Feng TL. The QUEST for quality online health information: validation of a short quantitative tool. BMC Med Inform Decis Mak. 2018;18(1):87. CrossRef PubMed
31. Web&Multimedia. Carlo Alberto Villa. Calcola il grado di leggibilità dei tuoi testi con la formula gulpease. 2019. Online (Accessed October 2024)
32. Lucisano P, Piemontese ME. Gulpease: una formula per la predizione della leggibilita di testi in lingua italiana. Scuola e città. 1988;110–24. Online (Accessed October 2024)
33. Terry G, Hayfield N, Clarke V, Braun V. Thematic analysis. The SAGE handbook of qualitative research in psychology. 2017;2(17–37):25. CrossRef
34. ATLAS.ti Mac. ATLAS.ti Scientific Software Development GmbH. 2023. Online (Accessed October 2024)
35. Ronzani CM, da Costa PR, da Silva LF, Pigola A, de Paiva EM. Qualitative methods of analysis: an example of Atlas. TI™ Software usage. Revista Gestão & Tecnologia. 2020;20(4):284-311. CrossRef
36. Krippendorff K. Content analysis: An introduction to its methodology. Sage publications; 2018. Online (Accessed October 2024)
37. Neuendorf KA. The content analysis guidebook. Sage publications; 2017. Online (Accessed October 2024)
38. National Institute for Health and Care Excellence (NICE). Low back pain and sciatica in over 16s: assessment and management. 2020 Dec 11; Online (Accessed October 2024)
39. Barrett DR, Boone JD, Butch JO, Cavender JA, Sole G, Wassinger CA. A critical appraisal of web-based information on shoulder pain comparing biomedical vs. psychosocial information. J Shoulder Elbow Surg. 2023;32(1):e23-e32. CrossRef PubMed
40. Austin RR, Jantraporn R, Schulz C, Zhang R. Navigating Online Health Information: Assessing the Quality and Readability of Dietary and Herbal Supplements for Chronic Musculoskeletal Pain. Comput Inform Nurs. 2024 Aug 1;42(8):547-554. CrossRef PubMed
41. Cobo Soriano J, Sendino Revuelta M, Fabregate Fuente M, et al. Predictors of outcome after decompressive lumbar surgery and instrumented posterolateral fusion. Eur Spine J. 2010;19(11):1841-1848. CrossRef PubMed
42. Dionne CE, Von Korff M, et al. Formal education and back pain: a review. J Epidemiol Community Health. 2001 Jul;55(7):455-68. CrossRef PubMed
43. Maslak JP, Jenkins TJ, Weiner JA, et al. Burden of sciatica on us Medicare recipients. J Am Acad Orthop Surg. 2020;28(10): e433-e439. CrossRef PubMed
44. Subialka JA, Smith K, Signorino JA, Young JL, Rhon DI, Rentmeester C. What do patients referred to physical therapy for a musculoskeletal condition expect? A qualitative assessment. Musculoskelet Sci Pract. 2022;59:102543. CrossRef PubMed
45. Lord Ferguson S. Should we give patients what they want? Patient expectations and financial pressures need to be addressed to increase uptake of evidence-based practice. Musculoskelet Sci Pract. 2023;66:102831. CrossRef PubMed
46. Stochkendahl MJ, Kjaer P, Hartvigsen J, et al. National Clinical Guidelines for non-surgical treatment of patients with recent onset low back pain or lumbar radiculopathy. Eur Spine J. 2018;27(1):60-75. CrossRef PubMed
47. Thomas JS, France CR. Pain-related fear is associated with avoidance of spinal motion during recovery from low back pain. Spine. 2007;32(16):E460-E466. CrossRef PubMed
48. Oosterhuis T, Smaardijk VR, Kuijer PPF, Langendam MW, Frings-Dresen MHW, Hoving JL. Systematic review of prognostic factors for work participation in patients with sciatica. Occup Environ Med. 2019;76(10):772-779. CrossRef PubMed
49. Domenighetti G, Vernero S. Looking for waste and inappropriateness: if not now, when. Intern Emerg Med. 2014;9:S1-S7.
50. Levinson W, Kallewaard M, Bhatia RS, Wolfson D, Shortt S, Kerr EA; Choosing Wisely International Working Group. 'Choosing Wisely': a growing international campaign. BMJ Qual Saf. 2015;24(2):167-174. CrossRef PubMed
51. Djurtoft C, O'Hagan E, Laursen MD, et al. Co-creating a Choosing Wisely leaflet supporting the reduction of imaging usage in low back pain management – A multi-method study. Patient Educ Couns. 2025;135:108730. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 206-213
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3482
ORIGINAL RESEARCH ARTICLE

Biomechanical of bilateral heel rise, and its association with balance, functional mobility, and walking speed in older adults

Cristian Caparrós-Manosalva[1,5,6], Rodrigo Guzmán-Venegas[2], Francisca Gajardo-Garrido[3], Marion González-Fuentes[3], Víctor Pino-Domínguez[3], Jessica Espinoza-Araneda[1,5,6], Ivan Palomo[4,5,6], Nacim Molina[4,5,6]

[1]Department of Human Movement Sciences, Faculty of Health Sciences, Universidad de Talca, Talca - Chile
[2]School of Kinesiology, Faculty of Medicine, Universidad de los Andes, Santiago - Chile
[3]School of Kinesiology, Faculty of Health Sciences, Universidad de Talca, Talca - Chile
[4]Thrombosis Research Center, Medical Technology School, Department of Clinical Biochemistry and Immunohematology, Faculty of Health Sciences, Universidad de Talca, Talca - Chile
[5]Interuniversity Center for Healthy Aging - Chile
[6]VITALIS Longevity Center, Faculty of Health Sciences, Universidad de Talca, Talca - Chile

ABSTRACT

Introduction: Aging advancing decreases ankle-foot strength and mobility, affecting gait and balance control. The heel-rise (HR) task requires the ankle-foot to control different biomechanical demands. It is still unclear whether these demands during HR are associated with functional performance in older adults. The aim was to describe the association between HR biomechanical parameters and single-leg stability, functional mobility, and walking speed in community-dwelling older adults.

Methods: Sixty-nine older adults (73.0, SD 6.8 years) were tested on a force platform performing bilateral rapid HR in the rise and drop phases. The biomechanical parameters measured were peak force and time, impulse, root mean square and displacement of the center of pressure (CoP), as well as displacement and velocity of the center of mass (CoM), and vertical stiffness. Functional performance was assessed through balance using the single-leg stance test (SLS), functional mobility with the Timed Up & Go test (TUG), and walking speed (WS). Associations between functional tests and biomechanical parameters were determined using correlation tests.

Results: HR peak strength and time showed a medium to large association with TUG and WS but not SLS. CoP anteroposterior displacement showed a large association in the drop phase with all functional tests but not in the rise phase. CoM velocity and vertical stiffness were associated with all tests in both phases.

Conclusion: Older adults HR biomechanical parameters are more closely associated with functional mobility and walking speed tests (TUG and WS) than with static balance tests such as SLS.

Keywords: Older Adults, Physical Functional Performance, Heel Rise, Ankle, Biomechanics

What is already known about this topic?

- *Older adults often experience decreased ankle and foot strength and mobility, which can significantly impact their gait and balance.*
- *Achieving tasks such as heel rise are key to maintaining functional performance in older adults.*

What does the study add?

- *Ankle and foot biomechanical parameters during heel rise are associated with performance on standard functional tests for older adults.*
- *The biomechanical demands of a bilateral rapid heel rise offer interesting insight into functional capabilities important for daily mobility in older age.*

Received: January 31, 2025
Accepted: July 15, 2025
Published online: July 30, 2025

This article includes supplementary materials

Corresponding author:
Nacim Molina
email: nacim.molina@gmail.com

Introduction

The strength, power, and mobility of the ankle-foot complex play a significant role in controlling postural balance, functional capacity, and the risk of falls in older adults (1,2). The strength and the muscular power of the ankle plantarflexors may serve as significant predictors of physical functional performance in older adults, particularly in

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

assessments such as the Timed Up & Go Test (TUG) and maximum walking speed (3-5). A common task in these functional tests is the heel rise (HR), a key milestone in the capacity of weight transfer during tasks such as walking. It is also used to identify ankle-foot dysfunctions, determine the strength of ankle plantarflexors, or prescribe exercises to recover ankle-foot functions (6-8). For proper function of the ankle plantarflexors during HR, the ankle and foot are required to be able to control different biomechanical demands, such as force, position, or speed, both during HR (concentric work) and when controlling body weight (Bw) when lowering the heel on the ground (eccentric work) (9). The behavior of these biomechanical parameters during impulse tasks, such as rising and dropping the heel of the foot, may influence functional performance; however, little has been discussed about these aspects in older adults.

The advancing age leads to a decrease in the strength and mobility of the ankle-foot, as shown, for example, by an annual rate of loss of ankle plantarflexor strength of approximately 2.3% in older adults (10), which impacts the phases of walking and balance control (11). During the push-off in walking, the foot must increase its rigidity allowing greater activity of the foot's intrinsic muscles so that the ankle plantarflexors can adequately transmit their force and thus achieve adequate elevation (7,12). In this sense, older adults have shown a 24%-37% decrease in strength of the extrinsic and intrinsic muscles of the ankle and foot compared to young subjects (13), which would limit adequate mobility and stability of the foot to achieve sufficient HR (14). On the other hand, the HR implies a decrease in the base of support of the foot and shifts the CoM anteriorly, increasing the demand for postural control to avoid losing balance (15). Older adults employ a more proximal posture control strategy than younger subjects (16), implying a greater biomechanical demand on the ankle-foot complex through the CoP to rise and control the heel drop without losing balance. HR has been reported to present numerous health benefits, as it enables the development or maintenance of lower extremity muscle strength and stability (17), thereby improving balance and muscle strength and consequently reducing the incidence of falls (2).

During HR, like walking, sufficient ankle and foot mobility, as well as increased midfoot stability, are required for adequate action of the ankle plantarflexors (14). Biomechanical parameters during HR, such as ground reaction force (GRF), CoP, and CoM, and their association with performance on functional tests frequently assessed in older adults, would allow for a better understanding of the prescription of plantar flexor strengthening exercises as a measure to mitigate fall risk. However, the association between these biomechanical parameters and performance on functional and stability tasks in older adults has been little discussed. The study aimed to describe the association between the biomechanical parameters of HR and single-leg stability, functional mobility, and walking speed in community-dwelling older adults. We hypothesize that the biomechanical parameters of heel rise would be associated with locomotion and, therefore, deteriorate, just as functional tests and balance are affected in older adults.

Methods

A cross-sectional study was conducted, inviting all community-dwelling, self-sufficient adults aged 60 and above who participated in a multimodal program offered at various primary care centers across the city. A total of 86 individuals were assessed between December and January 2024, of whom sixty-nine completed all evaluations (Fig. 1). Participants with evident balance disorders or who used walk-assistive devices were excluded (excluded = 3). During the initial interview, information was also collected regarding the presence of stroke sequelae, neurodegenerative diseases, diabetic foot, tendon injuries or fractures of the ankle or foot, or relevant medical history (excluded = 2). Additionally, individuals who failed to complete one or more functional tests (excluded = 10) or who did not understand the instructions and performed the tests poorly were excluded (excluded = 2). All participants received information about the study's purpose before providing their written informed consent, which was approved by the Scientific Ethics Committee of the University Institution in accordance with the guidelines of the Declaration of Helsinki.

The sample size was calculated based on an estimated medium correlation size of 0.3, which is the ratio reported for associations between ankle and foot strength and functional tests in older adults (18). Considering a significance level of 0.05 and a power of 80%, the calculation obtained indicated that 64 participants were needed.

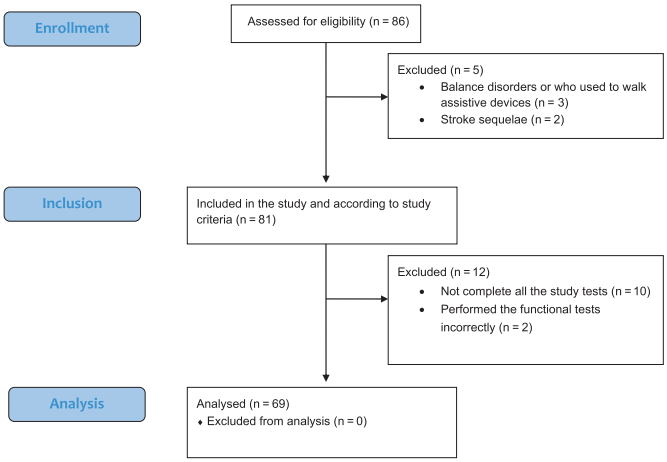


FIGURE 1 - Study participant flow chart.

Procedure

The anthropometric assessment was carried out according to the standardized protocol of the International Society for the Advancement of Kinanthropometry (ISAK) (19). Height was measured using a portable stadiometer (Seca 213, Seca Corporation, Germany), and Bw was assessed with an eight-electrode bioelectrical impedance scale (Huawei Scale 3 Pro, China). Waist and calf circumferences were measured using a flexible tape measure, following the methodological recommendations described by Lera et al. (2014) (20).

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Then, the functional tests of the TUG, single leg stance (SLS), and walking speed (WS) were assessed. An obstacle-free corridor was used for the TUG test, with a cone placed 3 meters from a chair with a backrest and a height of 40 cm. Each participant was instructed to start from a seated position, stand up without assistance, walk three meters, turn 180°, return to the starting point, and sit down again. The test was performed twice, asking the participant to complete it in the shortest possible time. The time was recorded with a stopwatch, and the shortest time obtained from the two attempts was used for analysis (21). The SLS was used to assess static balance. Before lifting the leg, the subject was instructed to cross their arms over the chest. A stopwatch was used to measure the time the subject could maintain balance on one leg. Timing began when the foot was lifted off the ground and ended if: (a) the arms were uncrossed, (b) the raised foot moved or touched the ground, (c) the supporting foot shifted to maintain balance, or (d) the maximum of 30 seconds was reached. The procedure was repeated three times, and the longest time obtained on the dominant leg was recorded for analysis (22). The WS test was conducted along a 4-meter corridor, measuring the time it took participants to comfortably walk a central 3-meter distance, with 1 meter allocated before and after for the acceleration and deceleration phases (23). The procedure was applied once, and the recording was used for analysis.

For the bilateral HR test, a clinician-guided warm-up was previously performed for 5 minutes, which included general mobility and flexibility exercises for the ankle and foot, followed by 30 seated HR tests and familiarization with the standing HR test through three repetitions to ensure adequate test performance. Then, each participant stood barefoot, with feet parallel, on a force platform (Bertec, USA) in front of a wall. They were instructed to rise both heels "as high and as fast as possible". The lowering of the heels was self-determined without receiving instructions for its control. During the execution of the test, each participant was instructed according to the procedure previously used for the bilateral heel rise (3,6). The participant could lightly support themselves against the wall with their fingers without assisting or pushing their body against it during the test. This was only used to avoid losing balance and ensure vertical displacement. Three valid repetitions were performed, eliminating those that presented visually noticeable asymmetry when lifting both heels, low elevation of both heels and loss of balance.

The force platform was configured to obtain the vertical ground reaction force (vGRF) and CoP signal at 200 Hz. The signals were smoothed post hoc with a 2nd-order filter and a 10 Hz low-pass filter, and the vGRF was adjusted to each participant's Bw. The signal onset was determined when the curve exceeded 3 SD of the pre-task resting signal average (24).

Figure 2 presents the variables obtained according to the phases defined for the HR. The rise phase was determined from the onset to the lowest point of the curve, and the drop phase was determined from the lowest point of the curve to the second peak of force.

The biomechanical parameter of peak force was determined from the top of the vGRF signal, and peak time was

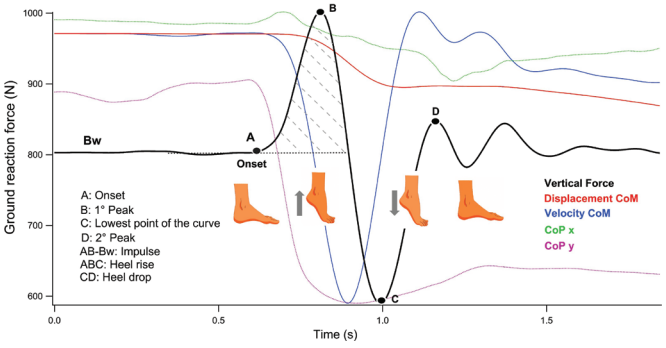


FIGURE 2 - Graphical summary of an example of the variables obtained during the Heel rise (HR) test. In the graph, the GRF is plotted on the vertical axis, and time in seconds is plotted on the horizontal axis. At the beginning of the recording, we observe the GRF line (in black) when the participant stands on the platform and registers their Bw. They then initiate the HR (point A: onset) by raising the heel (points ABC: rise phase) to the toe position (point C). They then lower the heel (points CD: drop phase) until the entire foot is in contact with the ground. The red line represents the displacement of the CoM, the blue line represents the CoM velocity, and the green and purple lines represent the displacements of the center of pressure in x and y, respectively. In addition, the impulse (shaded area under the curve at points AB and the projection of Bw) is shown in the GRF record.

the difference between the onset and peak force time (24). Impulse was obtained from the area under the curve of the rise phase between the onset and when the signal equaled the Bw (25). Vertical stiffness (Kv) was obtained using a one-dimensional mass-spring model (26), which comprises a single mass (e.g., body mass) supported by a single spring-like system, represented, for example, by the legs. Therefore, this model enabled the calculation of vertical stiffness (Kv) by analyzing the relationship between the HR and the length change during the linear movement, as determined by the vertical displacement of the CoM and the vGRF. The CoM displacement was obtained by applying a double integration of the vertical acceleration with respect to time obtained from the vGRF (a = F/m), as reported (27). Kv (Bw/m), velocity (m/s), and CoM displacement (m) were obtained in both phases of the HR. The CoP in the anteroposterior (CoP y) and lateral (CoP x) axes were obtained during the HR. The Root Mean Square (RMS) value, which determines the average absolute displacement of the CoP around its mean position and CoP displacement in each axis, was analyzed for both the rise and drop phases.

Statistical analysis

Data were tabulated and presented as means and standard deviations for each HR phase and functional test. Normality assumptions were checked using the Kolmogorov-Smirnov test ($p > 0.05$). Associations between functional tests and biomechanical parameters were examined using bivariate correlations with Spearman rank correlation coefficients (*rho*). For the interpretation of the associations, the magnitude of rho in the range 0.1-0.29 was considered a small association, 0.3-0.49 a medium association, and >0.5 a large association (28). All analyses and graphs were performed

using GraphPad Prism (GraphPad Software Inc., USA), with a significance level of 0.05.

Results

Table 1 describes the characteristics of the participants in the study sample. The older adults had an average age of 74 (62-88) years, a height of 1.54 (1.4-1.7) m, a weight of 70.3 (43.6-106.8) kg, a BMI of 29.4 (20.2-44.7) kg/m2, and 61 (88%) women and 8 (12%) men. In addition, they had a waist circumference of 94.9 (71-126) cm and a calf circumference of 36.1 (27-45) cm.

TABLE 1 - Characteristics of the study sample participants. Mean, standard deviation (SD), and minimum-maximum values are presented

Characteristics	n total = 69		
Sex (%women / %men)	88% / 12%		
	Mean	SD	Min-Max
Age, y	74.1	6.8	62.2-88.7
HR. Bpm	74.9	9.7	54.0-108.0
SBP, mmhg	134.1	19.6	104.0-190.0
DBP, mmhg	76.2	10.0	52.0-104.0
Weight, kg	70.3	15.2	43.6-106.8
Height, m	1.54	0.7	1.39-1.72
BMI, kg/m²	29.4	4.8	20.2-44.7
WC, cm	94.9	10.8	71.0-126.0
HC, cm	104.6	9.7	89.0-132.0
MCC, cm	36.1	3.6	27.0-45.0

Abbreviations: Y = years, HR = Heart rate, SBP = Sistolic blood pressure, DBP = Diastolic blood pressure, bpm = beat per minute, mmHg = millimeter of mercury, kg = Kilograms, cm = centimeters, kg/m^2 = kilograms per square meter, WC = Waist circumference, HC = Hip circumference, MCC = Max Calf circumference.

Heel rise during the rise and drop phases, and functional tests

Table 2 summarizes the results of the functional test performance and biomechanical parameters during the rise and drop phases. It was observed that the older subjects showed a CoPy (anteroposterior axis) displacement of 0.093, SD 0.02 m, a CoM velocity of 0.40, SD 0.09 m/s, a CoM vertical displacement of 0.05, SD 0.03 m and a vertical stiffness of 10.0, SD 6.2 Bw/m during the rise phase. On the other hand, during the drop phase, a CoPy displacement of 0.006, SD 0.00 m, CoM velocity of 0.21, SD 0.05 m/s, CoM displacement of 0.08, SD 0.05 m, and a vertical stiffness of 13.7, SD 13.7 Bw/m were obtained.

Correlations between the biomechanical parameters and the functional tests

The HR correlations with the functional tests (Table 3) showed a large correlation between the functional tests with the peak force (TUG: *rho* = -0.566 (-0.711, -0.374), *p*<0.001; WS: *rho* = 0.567 (0.376, 0.712), *p*<0.001), a medium correlation of the TUG and WS with the peak time (TUG: *rho* = 0.385 (0.157, 0.575), *p* = 0.001; WS: *rho* = -0.384 (-0.574, -0.155), *p* = 0.001), and no correlation of the impulse with the functional tests (*p*>0.05). In the rise phase, in general, the functional tests did not present a correlation with the displacement and RMS of the CoP (*p*>0.05), but this was not the case in the drop phase, where a significant correlation was observed with the anteroposterior displacement with the functional tests (TUG: *rho* = -0.562 (-0.708, -0.369), *p*<0.001; WS: *rho* = 0.558 (0.364, 0.705), *p*<0.001).

On the other hand, vertical stiffness showed a small to medium correlation in the rise phase (TUG: *rho* = -0.434 (-0.613, -0.214), *p*<0.001; WS: *rho* = 0.474 (0.261, 0.643), *p*<0.001), and also in the drop phase (TUG: *rho* = -0.387 (-0.576, -0.159), *p* = 0.001; WS: *rho* = 0.445 (0.226, 0.621), *p*<0.001). Similar behavior was shown by the CoM velocity in the rise phase (TUG: *rho*= -0.576 (-0.719, -0.387), *p*<0.001; WS: *rho* = 0.546 (0.349, 0.697), *p*<0.001) and in the drop phase (TUG: *rho* = -0.592 (-0.730, -0.408), *p*<0.001; WS: *rho* = 0.571 (0.380, 0.715), *p*<0.001). The CoM displacement did not present correlations with the functional tests in either of the two phases of the HR (*p*>0.05). All correlations were also illustrated in figures, see supplementary material.

Correlations between biomechanical parameters and the balance test

Correlations of HR biomechanical parameters with balance tests (see also Table 3) showed a medium correlation between SLS performance and peak force (*rho* = 0.387 (0.159, 0.576), *p* = 0.001), but no correlation with peak time and impulse (*p*>0.05). CoP displacement and RMS showed a significant large correlation between anteroposterior displacement and SLS (*p* = 0.511 (0.306, 0.671), *p*<0.001) only in the drop phase.

Vertical stiffness showed a small correlation in the rise phase (SLS: *rho* = 0.262 (0.020, 0.475), *p* = 0.029) and in the drop phase (SLS: *rho* = 0.278 (0.037, 0.488), *p* = 0.020). CoM velocity showed a medium and large correlation with SLS in the rise and drop phases (*rho* = 0.383 (0.154, 0.573), *p* = 0.001; *rho* = 0.543 (0.346, 0.695), *p*<0.001, respectively). As with the functional tests, the correlations with the balance test were illustrated in the figures (see also supplementary material).

Discussion

The present study aimed to describe the association between the biomechanical parameters of HR and single-leg stability, functional mobility, and walking speed in community-dwelling older adults. The main results show that greater peak force and shorter peak time in the rise phase during HR are more associated with better performance in walking and speed tests, such as the TUG and WS, than with balance (SLS). Additionally, a greater anteroposterior displacement of the CoP was associated with better performance in the drop phase of the three functional tests but not in the rise phase. Finally, greater vertical stiffness and velocity of the CoM were associated with better functional performance in the tests in both phases.

Postural balance impairment in older adults has been frequently associated with greater oscillation of the CoP due

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - Summary of biomechanical parameters of the heel rise and functional tests in older adults. Values obtained during the rise and drop phases and functional tests in older adults are presented. Values are presented as mean (standard deviation)

Heel rise	Mean (SD)	Heel drop	Mean (SD)	Functional test	Mean (SD)
Fpeak (Bw)	1.34 (0.1)	n.a.	n.a.		
Impulse (Ns)	179.0 (40.9)	n.a.	n.a.	TUG (s)	8.7 (1.9)
Time Peak (ms)	149.0 (31.3)	n.a.	n.a.		
RMS CoP x (m)	0.015 (0.01)	RMS CoP x (m)	0.016 (0.01)		
RMS CoP y (m)	0.060 (0.02)	RMS CoP y (m)	0.067 (0.03)	SLS (s)	10.2 (10.4)
Displacement Cop x (m)	0.022 (0.01)	Displacement CoP x (m)	0.013 (0.01)		
Displacement CoP y (m)	0.093 (0.02)	Displacement Cop y (m)	0.006 (0.00)		
Vertical Stiffness (Bw/m)	10.0 (6.2)	Vertical Stiffness (Bw/m)	13.7 (13.7)		
Displacement CoM (m)	0.05 (0.03)	Displacement CoM (m)	0.08 (0.05)	WS (m/s)	1.19 (0.2)
Velocity CoM (m/s)	0.40 (0.09)	Velocity CoM (m/s)	0.21 (0.05)		

Abbreviations: Fpeak: Force peak (maximum vertical ground reaction force during hell rise); Bw: body weight; ms: milliseconds; RMS: Root Mean Square; CoPx: center of pressure mediolateral direction; CoPx: center of pressure anteroposterior direction; m: meter; CoM: center of mass; m/s: meter/second; n.a.: not applicable; TUG: timed Up & Go test; SLS: single leg stance; WS: walking speed.

TABLE 3 - Correlations between biomechanical parameters of the heel rise and the performance of functional tests. Spearman rank correlation coefficient values and confidence Interval 95% are presented, and significant values are shown in bold

Outcomes	TUG			SLS			WS		
	rho	CI 95%	p	rho	CI 95%	p	rho	CI 95%	p
Heel rise									
Fpeak (Bw)	-0.566	-0.711, -0.374	**<0.001**	0.387	0.159, 0.576	**0.001**	0.567	0.376, 0.712	**<0.001**
Impulse (Ns)	0.013	-0.230, 0.256	0.910	-0.102	-0.337, 0.144	0.401	0.045	-0.199, 0.286	0.707
Time Peak (ms)	0.385	0.157, 0.575	**0.001**	-0.133	-0.365, 0.113	0.273	-0.384	-0.574, -0.155	**0.001**
RMS CoP x (m)	0.246	0.00339, 0.462	**0.041**	-0.073	-0.311, 0.173	0.549	-0.151	-0.380, 0.0958	0.215
RMS CoP y (m)	-0.224	-0.443, 0.0197	0.063	0.202	-0.042, 0.425	0.094	0.206	-0.038, 0.428	0.088
Displacement CoP x (m)	-0.087	-0.323, 0.159	0.476	-0.012	-0.255, 0.231	0.919	0.149	-0.097, 0.378	0.221
Displacement CoP y (m)	-0.117	-0.350, 0.129	0.337	0.095	-0.151, 0.331	0.436	0.110	-0.136, 0.344	0.366
Vertical Stiffness (Bw/m)	-0.434	-0.613, -0.214	**<0.001**	0.262	0.020, 0.475	**0.029**	0.474	0.261, 0.643	**<0.001**
Displacement CoM (m)	0.196	-0.049, 0.419	0.105	-0.095	-0.331, 0.151	0.435	-0.247	-0.463, -0.004	**0.040**
Velocity CoM (m/s)	-0.576	-0.719, -0.387	**<0.001**	0.383	0.154, 0.573	**0.001**	0.546	0.349, 0.697	**<0.001**
Heel drop									
RMS CoP x (m)	0.073	-0.172, 0.311	0.546	-0.012	-0.255, 0.231	0.915	0.012	-0.231, 0.255	0.919
RMS CoP y (m)	-0.161	-0.389, 0.085	0.184	0.146	-0.100, 0.376	0.229	0.120	-0.126, 0.353	0.323
Displacement CoP x (m)	-0.270	-0.482, -0.029	**0.024**	0.102	-0.144, 0.337	0.400	0.226	-0.017, 0.445	0.061
Displacement CoP y (m)	-0.562	-0.708, -0.369	**<0.001**	0.511	0.306, 0.671	**<0.001**	0.558	0.364, 0.705	**<0.001**
Vertical Stiffness (Bw/m)	-0.387	-0.576, -0.159	**0.001**	0.278	0.037, 0.488	**0.020**	0.445	0.226, 0.621	**<0.001**
Displacement CoM (m)	0.018	-0.225, 0.260	0.879	0.049	-0.196, 0.289	0.686	-0.109	-0.344, 0.137	0.368
Velocity CoM (m/s)	-0.592	-0.730, -0.408	**<0.001**	0.543	0.346, 0.695	**<0.001**	0.571	0.380, 0.715	**<0.001**

Abbreviations: Fpeak: Force peak (maximum vertical ground reaction force during hell rise); Bw: body weight; ms: milliseconds; RMS: Root Mean Square; CoPx: center of pressure mediolateral direction; CoPx: center of pressure anteroposterior direction; m: meter; CoM: center of mass; m/s: meter/second; TUG: timed Up & Go test; SLS: single leg stance; WS: walking speed.

to impaired postural control, which increases the risk of falls (29). In dynamic tasks such as HR, a greater anterior displacement of the CoP and CoM, as well as increased work of the ankle plantarflexor muscles, has been observed (30). Our results showed that a greater anteroposterior displacement of the CoP during the drop phase was associated with better performance in the three functional tests studied, but not during the rise phase. Few studies have analyzed the drop phase during HR in older adults; however, physiotherapists are known to prescribe eccentric ankle plantarflexor training when controlling heel drop (31,32). Therefore, the ability to better control balance in different tasks would depend not only on the concentric capacity of the ankle plantar flexors but also on the eccentric control capacity required to control movement (29). Furthermore, foot and ankle biomechanics are affected by advancing age, with increased stiffness, decreased range of motion, and postural changes in the foot, resulting in decreased propulsion capacity, for example, during walking (33). Along these lines, plantarflexor training has shown benefits in heel rise strength, dynamic balance (34), and in reducing the risk of falls, through a reduced fear of falling and increased fall self-efficacy in older adults (8).

In functional transfer tasks, such as gait, the ankle-foot complex needs to be able to transition between flexible and stiff conditions in order to adequately act during the push-off and absorption of foot impact in the stance phase (35). Our results showed that in both phases of HR, rise and drop, greater vertical stiffness and vertical CoP velocity were associated with better performance in functional tests in older adults. Although it has been described that both muscle strength and joint flexibility decrease in advanced ages, which would cause a lower functional capacity in balance and walking (36), older adults have shown that ankle mechanical stiffness is maintained in activities such as walking due to neural and musculoskeletal changes to maintain ankle stability (37). This could explain the association found between the biomechanical parameters and the functional tests analyzed. However, neither the plantarflexor muscle strength nor ankle range of motion was part of this study, making it necessary to incorporate these control parameters when evaluating HR.

Regarding the clinical implications of this study, the fast bilateral heel rise provides a compact window into several capacities—strength, power and balance control—that are important for everyday mobility in late life. Decades of work with the unilateral version already show that the number of pain-free rises an older person can complete mirrors their static and dynamic balance scores (38), and the test itself is both easy to teach and highly reliable across sessions and examiners (39). When heel rise metrics are normalized to body mass, they track plantar-flexor torque almost as well as an isokinetic dynamometer, yet need only a stopwatch or force plate.

More recently, researchers have begun to connect those same metrics with downstream outcomes that physiotherapists care about. Slower gait speed, shorter strides, and longer double-support phases—kinematic fingerprints of fall-prone elders—co-vary with weaker or slower heel-rise performance, suggesting a shared physiological bottleneck in ankle power generation (40). Sadeh et al. (2023) study on sit-to-stand mechanics shows a similar result. Compared to younger adults, seniors take longer to stand, flex their trunks more, and sway more afterward, all hallmarks of reduced ankle stiffness and delayed plantar-flexor activation (41). Interestingly, even brief interventions can make a significant impact. An eight-week balance block that blended progressive calf-raise drills with sensory-challenge exercises cut TUG times and trimmed fall-risk scores in community-dwelling participants over 70 (8).

Technological tools have also taken on importance in these areas. Pressure mats in footwear, inertial units, and small force sensors attached to the heel can record elevation height, velocity, and even fatigue-induced asymmetries, turning a two-minute field test into a clinical-quality data stream without the need to strap the patient into bulky laboratory equipment (42). Together, these developments position the heel rise not merely as a snapshot of calf strength but as a versatile, intervention-responsive marker that physiotherapists can use to screen, stratify, and track older adults who are slipping toward mobility loss or recurrent falls.

This study is not without limitations. An important aspect was that ankle plantarflexor muscle strength and ankle joint range of motion were not measured, which would limit the interpretation of some functional performance results in older adults, which have been shown to be related to strength and mobility capacity. Studies analyzing biomechanical factors in this type of dynamic testing should consider strength and mobility capacities in older adults, as well as attempt to compare these parameters with younger or middle-aged individuals. Another limitation was that the sample of participants presented a significant difference in the number of women compared to older men. Although it has been reported that sex does not influence both the concentric and eccentric phases of the HR (9), maximum strength values were normalized concerning Bw to control for differences obtained when using absolute strength values. Another factor that could limit the analysis of the results was the methodological configuration used, which allowed finger support during the HR execution. Although other studies have proposed this configuration (3,6), no form of control was applied to this strategy, nor was the number of participants who used or required this support considered, which should be taken into account in future applications of this test in older adults.

Conclusion

Older adults present biomechanical parameters of HR that are more closely associated with functional tests, such as the TUG and WS, than with balance tests (SLS). The fast HR showed associations with vGRF, anteroposterior displacement control, vertical stiffness, and vertical speed with tasks of greater dynamic demand, such as TUG and WS. Bilateral rapid heel rise and its biomechanical demands offer interesting insight into functional capabilities important for daily mobility in old age.

Acknowledgments

The authors thank all the community-dwelling older people who gave their time to participate in this study.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Disclosures

Conflict of interest: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Financial support: This research received no specific grant from public, commercial, or not-for-profit funding agencies

Authors' contributions: Conceptualization: CCM, NM; Formal analysis: CCM, FGG, MGF, VPD; Investigation: CCM, NM; Methodology: CCM, RGV; Project administration: NM, IP; Writing – original draft: CCM, NM, RGV; Writing – review & editing: CCM, RGV, JEA, IP, NM

Data Availability Statement: Data available on request: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy policy of the institution.

References

1. Menz HB, Morris ME, Lord SR. Foot and ankle characteristics associated with impaired balance and functional ability in older people. J Gerontol A Biol Sci Med Sci. 2005;60(12):1546-1552. CrossRef PubMed
2. Menz HB, Morris ME, Lord SR. Foot and ankle risk factors for falls in older people: a prospective study. J Gerontol A Biol Sci Med Sci. 2006;61(8):866-870. CrossRef PubMed
3. Hester GM, Ha PL, Dalton BE, et al. Rate of Force Development as a Predictor of Mobility in Community-dwelling Older Adults. J Geriatr Phys Ther. 2021;44(2):74-81. CrossRef
4. Kanayama A, Minami M, Yamamoto S, et al. Examination of the Impact of Strength and Velocity of the Knee and Ankle on Gait Speed in Community-Dwelling Older Adults. Healthcare (Basel). 2022;10(10):2093. CrossRef PubMed
5. Stenroth L, Sillanpää E, McPhee JS, et al. Plantarflexor Muscle-Tendon Properties are Associated With Mobility in Healthy Older Adults. J Gerontol A Biol Sci Med Sci. 2015;70(8):996-1002. CrossRef PubMed
6. André HI, Carnide F, Borja E, Ramalho F, Santos-Rocha R, Veloso AP. Calf-raise senior: a new test for assessment of plantar flexor muscle strength in older adults: protocol, validity, and reliability. Clin Interv Aging. 2016;11:1661-1674. CrossRef PubMed
7. Jeong HJ, Mueller MJ, Zellers JA, Hastings MK. Midfoot and ankle motion during heel rise and gait are related in people with diabetes and peripheral neuropathy. Gait Posture. 2021;84:38-44. CrossRef PubMed
8. Maritz CA, Silbernagel KG. A Prospective Cohort Study on the Effect of a Balance Training Program, Including Calf Muscle Strengthening, in Community-Dwelling Older Adults. J Geriatr Phys Ther. 2016;39(3):125-131. CrossRef PubMed
9. Ugbolue UC, Yates EL, Ferguson K, et al. Electromyographic Assessment of the Lower Leg Muscles during Concentric and Eccentric Phases of Standing Heel Raise. Healthcare (Basel). 2021;9(4):465. CrossRef PubMed
10. Winegard KJ, Hicks AL, Sale DG, Vandervoort AA. A 12-year follow-up study of ankle muscle function in older adults. J Gerontol A Biol Sci Med Sci. 1996;51(3):B202-B207. CrossRef PubMed
11. Ishikawa M, Komi PV, Grey MJ, Lepola V, Bruggemann GP. Muscle-tendon interaction and elastic energy usage in human walking. J Appl Physiol (1985). 2005;99(2):603-8. CrossRef
12. Hastings MK, Woodburn J, Mueller MJ, Strube MJ, Johnson JE, Sinacore DR. Kinematics and kinetics of single-limb heel rise in diabetes related medial column foot deformity. Clin Biomech (Bristol). 2014;29(9):1016-1022. CrossRef PubMed
13. Spink MJ, Fotoohabadi MR, Menz HB. Foot and ankle strength assessment using hand-held dynamometry: reliability and age-related differences. Gerontology. 2010;56(6):525-532. CrossRef PubMed
14. Houck JR, Neville C, Tome J, Flemister AS. Foot kinematics during a bilateral heel rise test in participants with stage II posterior tibial tendon dysfunction. J Orthop Sports Phys Ther. 2009;39(8):593-603. CrossRef PubMed
15. Wang J, Wang Y, Zhou B, Wang L, Lai Z. Age-Related Reduction of Foot Intrinsic Muscle Function and the Relationship with Postural Stability in Old Adults. Clin Interv Aging. 2024;19:1005-1015. CrossRef PubMed
16. Boyas S, Medd ER, Beaulieu S, Boileau A, Lajoie Y, Bilodeau M. Older and young adults adopt different postural strategies during quiet bipedal stance after ankle plantarflexor fatigue. Neurosci Lett. 2019;701:208-212. CrossRef PubMed
17. Flanagan SP, Song JE, Wang MY, Greendale GA, Azen SP, Salem GJ. Biomechanics of the heel-raise exercise. J Aging Phys Act. 2005;13(2):160-171. CrossRef PubMed
18. Spink MJ, Fotoohabadi MR, Wee E, Hill KD, Lord SR, Menz HB. Foot and ankle strength, range of motion, posture, and deformity are associated with balance and functional ability in older adults. Arch Phys Med Rehabil. 2011;92(1):68-75. CrossRef PubMed
19. Esparza-Ros F, Vaquero-Cristóbal R, Marfell-Jones M. International standards for anthropometric assessment. International Society for the Advancement of Kinanthropometry (ISAK): Glasgow, UK, 2019; 154.
20. Lera L, Ángel B, Sánchez H, et al. [Validation of cut points of skeletal muscle mass index for identifying sarcopenia in Chilean older people]. Nutr Hosp. 2014;31(3):1187-97. CrossRef
21. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc. 1991;39(2):142-148. CrossRef PubMed
22. Springer BA, Marin R, Cyhan T, Roberts H, Gill NW. Normative values for the unipedal stance test with eyes open and closed. J Geriatr Phys Ther. 2007;30(1):8-15. CrossRef PubMed
23. Bohannon RW, Wang YC. Four-Meter Gait Speed: Normative Values and Reliability Determined for Adults Participating in the NIH Toolbox Study. Arch Phys Med Rehabil. 2019;100(3):509-513. CrossRef PubMed
24. Maffiuletti NA, Aagaard P, Blazevich AJ, Folland J, Tillin N, Duchateau J. Rate of force development: physiological and methodological considerations. Eur J Appl Physiol. 2016;116(6):1091-1116. CrossRef PubMed
25. Zellers JA, Baxter JR, Grävare Silbernagel K. Functional Ankle Range of Motion but Not Peak Achilles Tendon Force Diminished With Heel-Rise and Jumping Tasks After Achilles Tendon Repair. Am J Sports Med. 2021;49(9):2439-2446. CrossRef PubMed
26. McMahon JJ, Comfort P, Pearson S. Lower Limb Stiffness: Effect on Performance and Training Considerations. Strength Condit J. 2012;34(6):94-101. CrossRef
27. Donelan JM, Kram R, Kuo AD. Simultaneous positive and negative external mechanical work in human walking. J Biomech. 2002;35(1):117-124. CrossRef PubMed
28. Ellis PD, ed. PD E. Effect sizes and the interpretation of results. In: The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results. Cambridge University Press; 2010:1-2. CrossRef
29. Brech GC, Bobbio TG, Cabral KN, et al. Changes in postural balance associated with a woman's aging process. Clinics (Sao Paulo). 2022;77:100041. CrossRef PubMed
30. Ridge ST, Rowley KM, Kurihara T, et al. Contributions of Intrinsic and Extrinsic Foot Muscles during Functional Standing Postures. BioMed Res Int. 2022;2022(1):7708077. CrossRef PubMed

31. Jeong S, Lee DY, Choi DS, Lee HD. Acute effect of heel-drop exercise with varying ranges of motion on the gastrocnemius aponeurosis-tendon's mechanical properties. J Electromyogr Kinesiol. 2014;24(3):375-379. CrossRef PubMed
32. Lee SM, Cynn HS, Yoon TL, Lee JH. Effects of different heel-raise-lower exercise interventions on the strength of plantarflexion, balance, and gait parameters in stroke survivors. Physiother Theory Pract. 2017;33(9):706-715. CrossRef PubMed
33. Menz HB. Biomechanics of the Ageing Foot and Ankle: A Mini-Review. Gerontology. 2015;61(4):381-388. CrossRef PubMed
34. Jeon MY, Jeong H, Petrofsky J, Lee H, Yim J. Effects of a randomized controlled recurrent fall prevention program on risk factors for falls in frail elderly living at home in rural communities. Med Sci Monit. 2014;20:2283-2291. CrossRef PubMed
35. Akuzawa H, Imai A, Iizuka S, Matsunaga N, Kaneoka K. The influence of foot position on lower leg muscle activity during a heel raise exercise measured with fine-wire and surface EMG. Phys Ther Sport. 2017;28:23-28. CrossRef PubMed
36. Boyer KA, Johnson RT, Banks JJ, Jewell C, Hafer JF. Systematic review and meta-analysis of gait mechanics in young and older adults. Exp Gerontol. 2017;95:63-70. CrossRef PubMed
37. Krupenevich RL, Clark WH, Sawicki GS, Franz JR. Older Adults Overcome Reduced Triceps Surae Structural Stiffness to Preserve Ankle Joint Quasi-Stiffness During Walking. J Appl Biomech. 2020;36(4):209-216. CrossRef PubMed
38. Hashish R, Samarawickrame SD, Wang MY, Yu SS, Salem GJ. The association between unilateral heel-rise performance with static and dynamic balance in community dwelling older adults. Geriatr Nurs. 2015;36(1):30-34. CrossRef PubMed
39. Hébert-Losier K, Wessman C, Alricsson M, Svantesson U. Updated reliability and normative values for the standing heel-rise test in healthy adults. Physiotherapy. 2017;103(4):446-452. CrossRef PubMed
40. Marques NR, Spinoso DH, Cardoso BC, Moreno VC, Kuroda MH, Navega MT. Is it possible to predict falls in older adults using gait kinematics? Clin Biomech (Bristol). 2018;59:15-18. CrossRef PubMed
41. Sadeh S, Gobert D, Shen KH, Foroughi F, Hsiao HY. Biomechanical and neuromuscular control characteristics of sit-to-stand transfer in young and older adults: A systematic review with implications for balance regulation mechanisms. Clin Biomech (Bristol). 2023;109:106068. CrossRef PubMed
42. Pires IM, Ponciano V, Garcia NM, Zdravevski E. Analysis of the Results of Heel-Rise Test with Sensors: A Systematic Review. Electronics (Basel). 2020;9(7):1154. CrossRef

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 214-238
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3336
REVIEW

What are the evidence-based medical management approaches for the concussed youth athlete? A scoping review

Tom McKeever[1], **Michael Leavitt**[1], **Stephanie Valentin**[1], **David F Hamilton**[2]

[1]School of Health and Social Care, Edinburgh Napier University, Edinburgh - UK
[2]Research Centre for Health, Glasgow Caledonian University, Glasgow - UK

ABSTRACT

Background: Concussion management in youth sport relies on the experience of adults pitch-side as to injury recognition, removal, and management decisions. Little consensus exists on the consistency of pitch-side and medical pathway management approaches.
Objectives: A scoping review was completed to identify and synthesise primary research as to the management of the concussed youth athlete.
Methods: A systematic search was completed in Medline, CINAHL, PubMed, SPORTDiscus, OVID emcare, Web of Science, ScienceDirect, and Cochrane Library databases and Google Scholar from inception to 1st March 2025 according to PRISMA-ScR guidelines. Primary research studies that provided outcome data on management approaches for concussed youth athletes at all stages post-injury were included. Articles were synthesised and reported in themes.
Results: 36 studies were included and four themes identified: Exercise, activity, and neuromuscular training interventions (n = 14), Pitch-side, sub-acute management and monitoring (n = 6), Novel treatment interventions (n = 5), and individual management in specific populations (n = 11).
Conclusion: Published primary research for concussed youth athlete management approaches with outcome data supports early activity recommendations and sub-threshold exercise programmes. Few data exist for alternative management approaches. Case reports/ series saturated this review, with limited generalisable data reported on. Further transparency on concussion injury reporting and management, with outcome data, is warranted.

Keywords: Adolescent, Brain concussion, Brain injuries, Patient care management, Rehabilitation, Sports

What is already known about this topic?

- *Concussion management approaches are based on consensus group agreements and recommendations. Previous reviews provide rationale for generalised management approaches for injured athletes. Recommended approaches often require resources and adequate training, limiting implementation.*

What this study adds?

- *This study reviews the primary data sources that currently stand as rationale for concussion injury medical management approaches, and highlights the lack of such sources, particularly from populations outside of the United States of America (USA)/ Canada.*

Introduction

Participation in sport is encouraged by schools, governments, and healthcare professionals for the physical, mental, and social benefits it can provide. Despite this, sports with an increased incidence of head contact injuries and subconcussive forces could negatively impact long-term brain health (1,2). Concussions or mild traumatic brain injuries (mTBI) are common within the youth athletic and sporting populations (3). The term concussion is well-used, however, poorly articulated. It broadly relates to the cluster of symptoms that results following a mild brain injury (4). McRory et al. defined a sports-related concussion as "a traumatic brain injury that is defined as a complex pathophysiological process affecting the brain, induced by biomechanical forces" (4). Many authors use the term 'concussion' to categorise neurological deficits following head trauma as a result of biomechanical forces (5). Whilst a traumatic brain injury (TBI) is often diagnosed using imaging such as magnetic resonance

Received: October 9, 2024
Accepted: July 15, 2025
Published online: August 4, 2025

This article includes supplementary material

Corresponding author:
Tom McKeever
email: t.mckeever2@napier.ac.uk

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

imaging (MRI) or computerised tomography (CT), the use of imaging for a concussed patient will often show no obvious structural change to the brain (6). The lack of a definitive clinical or diagnostic test for concussion proves challenging for medical professionals, such as pitch-side physiotherapists, when deciding when to remove athletes from play and in attempts to treat the concussion. Where initial mild head injuries are missed during play, in rare circumstances, a further brain insult can result in second impact syndrome, with fatal consequences for individuals (7).

The overall global annual prevalence of concussions in all age groups is estimated to be more than 6 per 1000 people (8). Data from the United States of America (USA) suggests that 1 in 5 adolescents have sustained a concussion (9). Around 1-in-3 concussions are induced by a sports-based mechanism, and in Canada, contact sports are the leading cause of concussion in young people (10,11). Concussion prevalence and injury susceptibility may be increased for the youth athlete due to differing pathophysiological responses in the developing brain (12-14). Youth female athletes are at greater risk of sustaining a concussion, potentially due to a reduced biomechanical threshold tolerance for impacts (15). However, female players are more likely to report a concussive injury to coaches, and there is preliminary evidence that the symptom burden is higher in youth female athletes (15,16).

In community youth sport settings, coaches and parents often lead the identification and management of sports injuries. If a grass-roots sports team is well-resourced, it may have additional first-aid staff or physiotherapy cover. In the United Kingdom (UK), the Government published its concussion guidance for grassroots sport in April 2023 (17), and other countries have followed suit (18,19). However, literature suggests there is a lack of knowledge of the Sports Scotland guidance publication from 2018, *"If in doubt, sit them out"*, amongst football coaches (20,21). This highlights the challenges of the translation of guidance documents to effective pitch-side management strategies in the community game. Despite increased resources, even in professional sports, there remain concerns over under-reporting, misdiagnosis, and poor pitch-side management of concussions (22). Events as recent as the 2022 Fédération Internationale de Football Association (FIFA) World Cup have highlighted that even the most prestigious sporting tournaments can fall short in terms of concussion protocol standards (23). Despite advances in public health messaging and formalised guidance and return-to-play protocols being published, both immediate and subsequent management of the concussed youth requires scientific data to support approaches taken. Empirical evidence has guided many of the recent publications on concussion management approaches, yet there remains no current available Scottish Intercollegiate Guidelines Network (SIGN) or National Institute for Health and Care Excellence (NICE) publications on youth concussion. Concussion guidelines have been developed from the consensus statement from the 5th and 6th international concussion conferences in Berlin, 2016, and Amsterdam, 2022 (4,24). Whilst the consensus statement acknowledges the barriers and challenges of sideline youth sport concussion management, often due to a lack of availability of medical staff, there is limited discussion of the data-driven approaches to support and rehabilitate a concussed youth. Piedade et al. reported on the poor standardisation of concussion management across age groups, and also the lack of high-quality evidence to guide practice (25). This lack of data is concerning as evidence-based practice is dependent on appropriate information as to comparative management to support treatment algorithms and best practice rehabilitation choices. Physiotherapists, working with concussed youth athletes pitch-side, or indeed in primary or secondary care settings, require data to inform their treatment and management approaches. Therefore, the aim of this study was to complete a systematic search of published primary research literature on the medical management, rehabilitation approaches, and outcome data. For the concussed youth athlete across all stages of injury management.

Methods

A scoping review of the literature was undertaken in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for scoping reviews (PRISMA-ScR) guidelines (26). Our protocol and search strategy were pre-published on the open science framework website – CrossRef This protocol included our study design.

Authorship group statement

Our author team consisted of three men and one woman, which was the makeup of the PhD supervisory team. All authors reside in the United Kingdom, and two authors were from Scotland, one from Chicago in the United States, and one from the Netherlands. The paper was led by the main author, who is a PhD student and supported by the supervisory team. All team members are chartered physiotherapists.

Information sources and search strategy

We applied the population, concept, and context (PCC) criteria to inform our search strategy that aimed to find primary research articles that reported outcome data on medical management approaches for the concussed youth athlete worldwide. The search strategy was devised in conjunction with a specialist librarian and an electronic search of the following databases was conducted from inception to 1st March 2025 in Cumulative Index to Nursing and Allied Health Literature (CINAHL), PubMed, SPORTDiscus, Cochrane Library, MEDical Literature Analysis and Retrieval System Online (MEDLINE), Ovid Emcare, Web of Science and the Cochrane Library. Boolean operators and MeSH terms were employed in the searches as detailed in the supplemental data. A hand search of subject-relevant journals, using the search term 'concussion', was also completed. Keywords were identified from initial database searches and screening. Such chosen keywords were used in a free text search of Google Scholar and further searches of grey literature. A simplified search strategy, using eight subject heading keywords only, was completed in ScienceDirect and Ovid emcare due to database limiters, using such chosen keywords. Reference

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

lists of included manuscripts were hand-searched for additional information sources. We applied an English language restriction, but no other filters to the search.

Eligibility Criteria and Study Selection

Due to the nature of our research objectives, all primary data capture study types were considered for inclusion, provided they were published as peer-reviewed research articles and reported on outcome data for pitch-side, medical, and rehabilitative management approaches for the concussed youth athlete. We limited our search to participants up to and including the age of 19 years, which is the World Health Organization's end range age for adolescents (27). This focused our review specifically on adolescent and child athletes involved in pre-college and pre-university level sport.

We employed a broad definition of concussion and included terminology synonymous with 'concussion' for the search. All stages of sport-related concussion management were considered, including: identification and initial assessment; primary and secondary care management; and home-care management. Studies reporting on more serious forms of traumatic brain injuries (moderate or severe TBIs) were excluded, due to differing management approaches and alternative diagnoses. All levels of sporting participation were considered for this study. We excluded expert opinion or narrative articles, and excluded material that was not available as a full-text publication, such as conference abstracts. Studies relating to concussion epidemiology and diagnostics were excluded. We also excluded articles that reported standardised hospital pathway data, as these articles did not provide discernible information on interventional management approaches and their outcomes. Manuscripts conducting retrospective analyses were excluded if they were reporting on secondary data that was not collected by the study team for their own sole purpose of analysis. We also excluded articles that included participants who had not sustained a sports-related concussion (SRC).

A three-part screening strategy was employed to identify relevant articles. Covidence systematic review management software was used to process database results and screen articles (Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia). Two investigators independently carried out the searches and screened by title using the online Covidence software platform (TM and SV]. Abstracts were reviewed independently by the same two investigators, and a consensus was reached for full text inclusion. In the event of disagreement or doubt, manuscripts were included for full text review. Full texts were reviewed by the same two reviewers independently, and the final selection was agreed upon by consensus with a third independent reviewer.

Data extraction and synthesis

Data extraction forms were created, and two researchers [TM and SV] independently extracted the data from included articles. The following characteristics of each study were extracted to a bespoke excel database and data was extracted as per our published protocol [Author/s and Year of publication, geographical location, study design, level of evidence (28), study aims, population and sample size (Including number of male and female participants, methods, intervention type and comparator, theme, outcomes and measurement and results)]. The results were synthesised and pooled by themes and sub-themes.

Results

The search generated 11640 articles (Medline: 4038; CINAHL: 1794; Web of Science: 1982; PubMed: 1772; OVID Emcare: 436; Science Direct: 455; Cochrane Library: 446; SPORTDiscus: 717). Following the removal of duplicates, 5410studies were evaluated against the eligibility criteria. Title and abstract screening was completed, resulting in the removal of 5095 studies which clearly did not meet the inclusion criteria, 315 studies were eligible for full-text review. Most articles were excluded as they discussed non-sports-related concussions (n = 89), were secondary data analysis, narrative reviews, and expert opinions (n = 59), or contained participants over the age of 19 (n = 32). The remaining excluded articles resulted from no full text available (n = 56), no medical outcome data provided (n = 12), no management approaches discussed (n = 4), an epidemiological study (n = 5), and more serious brain injury being discussed (n = 5). Thirty-six studies were retained from the initial search. A further 42 articles were identified through citation screening, 41 of which were excluded. The final included studies numbered 36. Full details are displayed in the PRISMA flowchart (Figure 1).

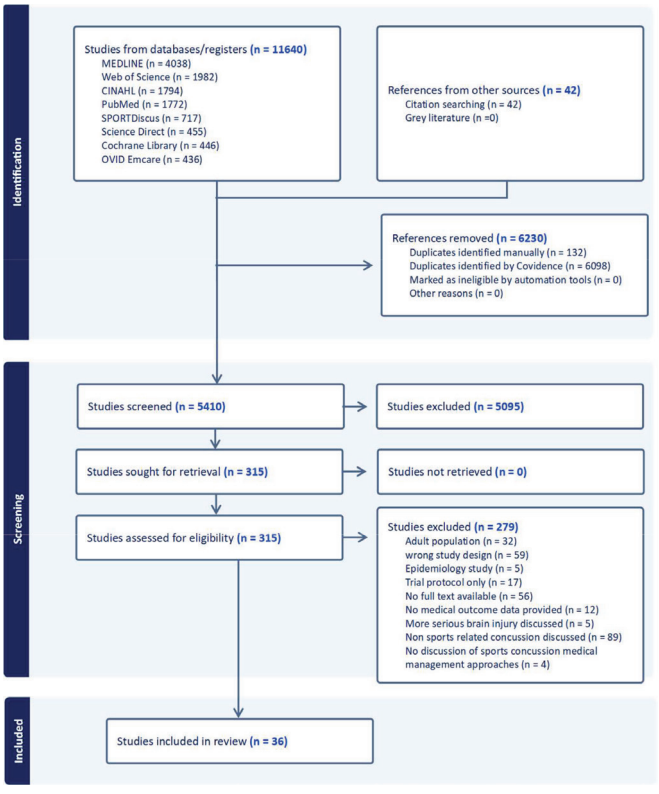


FIGURE 1 - PRISMA flow diagram.

Study characteristics

The majority of articles were from America (n = 25) (29-53) and Canada (n = 8) (54-61). The remaining studies were carried out in Switzerland (n = 1) (62), New Zealand (n = 1) (63), and Ireland (n = 1) (64). Various study design were reported: Prospective cohort studies (n = 9) (Level 4 evidence) (30-33, 54-57,64), case reports/series (n = 11) (Level 7 evidence) (41-48,60,52-53,63), randomised controlled trials (RTCs) (n = 12) (Level 2 evidence) (35-40,49,50,52,58-60), case-control studies (n = 2) (Level 4 evidence) (29, 51), quasi-experimental studies (n = 1) (Level 3 evidence) (61) and a feasibility trial design (n = 1) (Level 4 evidence) (34). Our categorisation of evidence level was informed by Melnyk and Fineout-Overholt (2023) (28). Across all studies, participants' ages ranged from ten to nineteen years old. Most studies reported on both male and female youth participants (64%, n = 23) (29,31-40,44,48-57,61), with 25% of studies reporting on male participants only (n = 9) (30,42-43,46-47,59-60,64) and the remaining 11% (n = 4) reporting on sole female participants in individual case reports (41,45,53,62). No trials reported on medical management outcome data for female-only participants. The results from this scoping review also highlight the saturation of data published from North America (69%, n = 25 from the USA and 22%, n = 8 from Canada). Only two studies were published in Europe (6%) and one from New Zealand (3%). Interestingly, there were no studies found from the United Kingdom (UK) that met the inclusion criteria for this scoping review.

Study themes

Four key themes were identified: (1) Exercise, activity, and neuromuscular training interventions (n = 14), (2) Pitch-side and sub-acute management, and monitoring (n = 6), (3) Novel treatment interventions (n = 5), and (4) Individual management in specific populations (n = 11). Individual study population details, study aims, methods, interventions, and results can be found in Table 1.

Theme 1 – Exercise, activity, and neuromuscular training interventions

1.1 Exercise/activity interventions efficacy

Eight studies reviewed the impact of exercise-based interventions or early activity levels for concussed youth athletes (29,33,37,40,49,54,56,59). Five of these specifically evaluated the utilisation of exercise interventions on recovery and post-concussion symptom scale (PCSS) Scores (37,40,49,56,59). Sub-threshold exercise programs (STEP) were found to support a reduction in PCSS scores and improved recovery times (40,49,56,59). Leddy et al. reported that symptom scores appeared to reduce in the exercise group; however, the reduction in scores was not statistically significant (37). Relative rest and placebo-like stretching interventions were less effective than STEP in relation to days to recovery and symptom reduction (37,40).

The remaining three studies in this theme evaluated physical activity levels post-concussion injury and reported outcome data (29,33,54). Early light-moderate exercise engagement after concussion was beneficial, and those not engaging in light exercise prolonged their return to play (33,54). Vigorous levels of physical activity post-injury delay recovery time (54). Completing exercise bouts lasting more than 15 minutes reduced the risk of developing PCS (29).

1.2 Early aerobic exercise testing safety and effects

Four studies examined the safety and efficacy of early exercise interventions post-concussion for the youth athlete (35,39,58,61). Two studies evaluated the Buffalo Concussion Treadmill test (BCTT), a provocative exercise testing protocol. Completing such a test within 1-10 days post-injury did not delay participant recovery (39). Participants with symptomatic SRC (SSRC) had a lower peak rate of perceived exertion (RPE), Oxygen consumption, Carbon dioxide production, and minute ventilation compared with control participants when completing the test (61).

Active rehabilitation and sub-maximal exercise interventions were safe, well tolerated, and reportedly provided therapeutic benefit for recently concussed participants (35,60). Chan et al. reported six adverse events during their study, yet concluded that these were unrelated to the intervention. Symptom exacerbation was experienced during initial sessions (resolving within 24 hours); however, they were able to tolerate 15 minutes of aerobic exercise by the final session, symptom-free (60). Worts et al. reported that early exercise interventions working at 40-60% of heart rate max yielded clinical improvements for those with a recent SRC (35). These studies support the safety of the use of such testing protocols and early exercise interventions for concussion management.

1.3 Neuromuscular training intervention

The two remaining studies under this theme explored outcomes following specialist rehabilitative interventions (34,38). One study was a feasibility trial for neuromuscular training (NMT) following a concussion injury and reported that it was possible to complete with their target population, and no adverse events occurred (34). A follow-up study by the same lead author established that an 8-week NMT programme after concussion may provide benefit to participants and reduce sports-related injuries and time lost from play over the following year from concussion (38).

Theme 2 – Pitch-side management approaches, sub-acute assessment, and monitoring

2.1 Removal vs no removal from play following concussion

Three studies evaluated recovery time and outcomes following removal from play versus no removal from play, post-concussion injury (31-32,57). Olsen et al. reported that youths who presented with early symptom onset (Initially symptom onset less than or equal to 15 mins from time of injury) were found to be more likely to be removed from play, compared to those with delayed symptom onset (Initial symptom onset more than 15 mins from time of injury). Those with delayed symptom onset took longer to return to play (RTP) following injury (57). Elbin et al. reported that delayed removal from play following concussion contributed to worse neurocognitive performance (ImPACT scores) and more

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

severe symptom presentations. Recovery time was also found to be enhanced by immediate removal from play following injury (32). Zydna et al. found that athletes with increased initial symptom burden and later presentation to clinical assessment had prolonged recovery; however, there were no differences in overall recovery time between those who continued to play post-concussion versus those who didn't (31).

2.2 Assessment and monitoring

Three studies reported on the assessment and monitoring of the concussed youth athlete (30,55,64). The Cogstate brief battery (CBB) and King-Devick (K-D) test were unable to consistently provide accurate results for concussion assessment and monitoring (64). Emery et al. evaluated factors associated with clinical recovery in youth ice hockey players. And found that delayed presentation for clinical assessment (>7 days) was associated with delayed RTP clearance. Poorer SCAT3/5 tandem gait scores were associated with delayed clinical recovery in the players, alongside higher symptom evaluation scores and greater headache severity (55). Kelleher et al. reviewed outcomes of youth athletes presenting to a mTBI clinic, with RTP success achieved within 16.9 days 9 (mean) (Range 5-39). However, added limited data to guide concussion symptom management (30).

Theme 3 – Novel treatment interventions

Five studies reported on novel treatment interventions for concussed youth athletes (36,44,50-52). Of those, two studies aimed to establish the safety and efficacy of head and neck cooling therapy following injury (37,50), and a further study reviewed participant response to face cooling therapy (51). For head and neck cooling, the authors reported a greater and faster reduction in symptoms reported (from SCAT5 symptom evaluation) when compared to a control group using between-group mean differences at each visit and time point (36,50). Congeni et al. reported that the mean differences between groups (95% CI) at similar post-treatment time points yielded significantly greater and earlier reduction in symptoms for the treatment group during the baseline scores post-treatment, the following 72-hour visit pre- and post-treatment only. At the 10-day and 28-day visits, there was a reduced between-group mean difference of symptom scores up until the study endpoint, with a similar total recovery trajectory. Smith et al. reported no difference in total symptom scores pre-intervention between groups. The reduction in symptoms scores for the intervention group were reported as follows: a 14.4% reduction (95% CI, 6.4-22.4) at the initial visit, post-treatment; a 25.5% (95% CI, 15.3-35.1) greater reduction at 72 hour post initial treatment (following further treatment); and a 3.4% (95% CI, 28.1-14.9) greater reduction at the 10-day visit when compared with the control group (50). Similarly to Congeni et al., this study demonstrated head cooling efficacy for reducing symptom burden earlier; however, the control group's total recovery trajectory was also similar. Haider et al. reported no between-group differences in parasympathetic parameter responses for those who were concussed and those who weren't. Interestingly, they found that in athletes with a prior history of concussion, heart rate and heart rate variability responses were significantly different from the control group, suggesting reduced adaptability of the autonomic nervous system (51).

A further novel treatment option studied was the use of docosa-hexaenoic acid (DHA) supplementation and its effects on recovery from SRC. This study found that no adverse effects were associated with DHA supplementation, and there was no improvement in recovery times for those taking DHA (52).

The remaining study charted hyperbaric oxygen therapy for the treatment of mTBI. No adverse effects were charted, and with limited data presented, the authors reported that HB02 may be an effective treatment following concussion, but their results were inconclusive. The efficacy of the intervention was not evaluated, and the purpose was to describe the effect and report any adverse events, of which there were none (44).

Theme 4 – Individual accounts of concussion management in specific populations

Eleven of the included studies were case reports/series. These reports were heterogeneous in terms of the participant characteristics and interventions reported. These case series/reports provide individual accounts of concussion management in specific populations and remain poorly generalisable. These reports can be groups into sub-themes as outlined below.

4.1 Concussion and mental health

Three case reports/series charted interventions and outcomes of youth athletes requiring more specialist input and transitions of care to mental health services/ teams following concussion and with such complaints: persistent mood disorder and anxiety (n = 1 participant) (45); post-concussion symptoms in combination with pre-existing mood disorders (n = 2) (47); prolonged PCS and post-traumatic stress symptoms (n = 1) (62). All cases report on the need for enhanced multidisciplinary team communication and management (45,47,62). One study reports the need for improved mental health screening in all athletes to identify risk factors for PCS (47). Only one case reported positive outcomes (45), whilst the remaining two cases reported poor, lasting mental health complaints in the young athletes who required ongoing psychological support (47,62).

4.2 Specialist medical case reports/series

A further four studies report on specialist medical interventions or the case of concussion cases with medical complications or delayed recovery (42,46,48,53). Lunardini et al. (2009) reported on treating a concussed individual with diabetes and reported the potential for abnormal blood sugars to have contributed to delayed recovery. The individual in this case had a history of poor blood sugar control and had a protracted recovery of 6 months and required specialist medical input and academic protections (42).

Stiller-Ostrowski reported one case of a young athlete with a cranial nerve palsy following a concussive injury. The patient

in this case remained symptomatic at five months post-injury, with visual disturbance. He was provided with advice and risk education from a specialist neuro-ophthalmologist; however, he returned to full contact sport at month six, reporting a full improvement in symptoms, subjectively (46). Bramley et al. evaluated the use of steroids for severe post-traumatic headaches, following a concussion incident, and reported improvements for both participants with physician monitoring and medication management (48). McNally et al. reported on the requirement for more advanced multidisciplinary and neuropsychology treatment input for an injured female youth horse rider, who had failed to respond to usual concussion care. She recovered by day 75, post injury, with an ongoing mild functional gait disturbance and remained on medication for her underlying psychiatric conditions (53).

4.3 Chiropractic management

Two case reports/series evaluated outcomes following chiropractic interventions (43,63). Olson et al. charted the recovery of three youth football players with concussion and reported improvements in cervical spine range of movement, muscular spasms, and neck pain alongside associated 'concussive/vestibular symptoms' following manual therapy interventions (43). The authors positively frame their interventions to have enhanced return to play outcomes. Hunt et al. charted the recovery of a youth skateboarder presenting for manual therapy interventions post-concussion (63). The authors report the potential for their intervention to have improved symptoms but draw no conclusion due to their expression of a weak evidence base for improvement claims.

4.4 Normative return to play rehabilitation

A further two studies in this review chart rehabilitative approaches for athletes with concussions (41,60). Gunter et al. reported on the physical therapy management of a female equestrian athlete with a history of multiple concussions and charts a return to sport at 8 weeks with no concerning complications (41). Marsh et al. reported on a standardised return to play protocol for a young hockey player, who was able to return to sport after three weeks with no ongoing symptoms (60). Both cases provide a narrative account of practitioner intervention and report on outcomes.

TABLE 1 - Data extraction – Participant and study details

Authors	Participant details				Study Design
	N total **N. control (CG) & intervention groups (IG)**	**N, male (M) & Female (F)**	**Age range (years)**	**Mean Age (years) ± standard deviation (SD)**	
Theme 1 – Exercise, activity, and neuromuscular training interventions					
1.1 Exercise/activity interventions efficacy					
Chrisman et al., 2019	30 IG = 19 CG = 11	M = 12 F = 18	12–18	15.5 ± 1.6	Pilot RCT
Gagnon et al., 2016	10	M = 7 F = 3	14–18	N/A	Pilot prospective cohort study
Micay et al., 2018	15 IG CG	M = 15 8 7	14–18	 15.8 ± 1.2 15.6 ± 1.0	RCT
Leddy et al., 2019	103 AE = 52 Stretching group (SG) = 51	M = 55 F = 48 M = 28 F = 24 M = 27 F = 24	13–18	 15.3 ± 1.6 15.4 ± 1.7	Multicenter prospective RCT
Leddy et al., 2021	118 AE group = 61 SG = 57	M = 74 F = 44 M = 38 F = 23 M = 36 F = 21	13–18	 15.5 ± 1.4 15.9 ± 1.4	Multicentre prospective RCT
Seehusen et al., 2021	32 RTP ≥ 28 Days Post-Injury = 17 RTP < 28 Days Post-Injury = 15	M =18 F = 14 M = 9 F = 8 M =9 F = 6	12–18	 14.9 ± 1.9 15.4 ± 1.5	Observational prospective cohort study
Lishchynsky et al., 2019	30 Low activity group (LAG) = 15 High activity group (HAG) = 15	M = 25 F = 5 M = 11 F = 4 M = 14 F = 1	12–17	N/A	Prospective cohort study
Neely et al., 2023	49 PPCS group No PPCS group	M = 24 F = 25 M = 11 F = 12 M = 13 F = 13	12–18	 14.7 ± 1.9 14.9 ± 1.7	Prospective observational case-control study

(*Continued*)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 1 - (*Continued*)

Authors	Participant details				Study Design
	N total **N. control (CG) & intervention groups (IG)**	**N, male (M) & Female (F)**	**Age range (years)**	**Mean Age (years) ± standard deviation (SD)**	
1.2 Early aerobic exercise testing safety and effects					
Morissette et al., 2020	74		14–19		Quasi-experimental nonrandomized study
	Symptomatic SRC (SSRC) = 34	M = 19 F = 15		16.4 ± 1.2	
	CG = 40	M = 13 F = 27		15.9 ± 0.8	
Chan et al.,2018	19	M = 5 F = 14	12–18	15.5 ± 1.7	Single-site RTC comparing TAU to TAU plus AR with blinded assessors
	Treatment at usual (TAU) = 9	M = 1 F = 8		15.1 ± 1.42	
	Active rehabilitation (AR) group = 10	M = 4 F = 6		15.9 ± 1.66	
Leddy et al., 2018	54	M= 37 F = 17	14–19		Prospective RTC
	BCTT = 27	M = 18 F = 9		15.19 ± 1.45	
	CG = 27	M = 19 F = 8		15.63 ± 1.36	
Worts et al., 2021	30	M = 20 F = 10	13–18		RCT
	SRC group = 19	M = 13 F = 6		15.8 ± 1.4	
	Healthy group = 11	M = 7 F = 4		16.3 ± 0.9	
1.3 Neuromuscular training intervention					
Howell et al., 2021	27	M = 17 F = 10	12–18		Feasibility trial.
	NMT = 11	M = 7 F = 4		14.7 ± 1.7	
	CG = 14	M = 9 F = 6		14.8 ± 1.5	
Howell et al., 2022	27	M = 16 F = 11	12–18		block-stratified RCT
	NMT group = 11	M = 7 F = 4		14.7 ± 1.7	
	CG = 16	M = 9 F = 7		15.3 ± 1.8	
Theme 2 – Pitch-side management approaches, sub-acute assessment, and monitoring					
2.1 Removal VS no removal from play following concussion					
Olson et al., 2020	144	M = 88 F = 56	N/A	14.6 ± 1.8	Prospective cohort study - qualitative data collection
	Delayed symptom onset group = 24	M = 15 F = 9		14.2 ± 2.0	
	Early symptom onset group = 120	M = 73F = 47		14.68 ± 1.7	
Elbin et al., 2016	69	M = 51 F = 18	12–19		prospective, repeated measures design
	RFP = 35	M = 27 F = 8		15.61 ± 1.65	
	Continued to play (CTP) = 34	M = 24 F = 10		15.35 ± 1.73	
Zynda et al., 2022	441	M = 217 F = 224	13–18		A prospective cohort study
	PLAY = 231	M = 109 F = 122		14.9 ± 1.29	
	NO PLAY = 210	M = 108 F = 102		14.8 ± 1.33	
2.2 Assessment and monitoring					
Cosgrave et al., 2023	135	M = 135	15–19	16.7 ± 0.8	Prospective cohort study.
	Prior history of concussion = 61	M = 61		16.7 ± 0.8	
	No prior history of concussion = 74	M = 74		16.6 ± 0.9	
Emery et al., 2021	329	M = 279 F = 50	11–17	N/A	prospective sub-cohort study
Kelleher et al., 2014	28	M = 28	13–18	15.3*	Pilot study plus cross-sectional survey
	Survey respondents = 266	M/F = Not reported		15.5*	

Authors	Participant details				Study Design
	N total N. control (CG) & intervention groups (IG)	N, male (M) & Female (F)	Age range (years)	Mean Age (years) ± standard deviation (SD)	
Theme 3 – Novel treatment interventions					
Congeni et al., 2022	45 Treatment group = 28 CG = 27	M = 28 F = 27 M = 14 F = 14 M =14 F = 13	12–17	 14.7 ± 1.6 14.25 ± 1.4	A randomized, nonblinded pilot trial.
Smith et al., 2024	167 CG = 88 Treatment group = 79	M= 82 F = 85 M = 37 F =51 M = 45 F = 34	12–19	 15.1 ± 1.6 15.2 ± 1.6	Amulti-site,prospective, longitudinal RCT
Roby et al., 2021	8	M = 5 F = 3		16.0 ± 1.2	Case series
Miller et al., 2022	40 CG = 20 Treatment group = 20	M = 27 F = 13 M = 13 F = 7 M = 14 F = 6	14–18	 15.97 ± 1.03 16.02 ± 1.19	Double-blind, parallel-group, pilot RCT
Haider et al., 2025	47 CG = 24 Concussed group = 23	M = 26 F = 21 M = 14 F = 10 M = 12 F = 11	13–18	 15.83 ± 1.55 15.48 ± 1.16	Prospective case-control study
Theme 4 – Individual accounts of concussion management in specific populations					
4.1 Chiropractic management					
Olson et al., 2018	3	M = 3	11–16	N/A	Case series
Hunt et al., 2018	1	M = 1	13	N/A	Case report
4.2 Concussion and mental health					
Sherry et al., 2021	1	F = 1	18	N/A	Case report
Tucker et al., 2019	2	M = 2	16–17	N/A	Case series
Waldmeier-Wilhelm et al., 2019	1	F = 1	14	N/A	Case report
4.3 Specialist medical case reports/series					
Lunardini et al., 2009	1	M = 1	16	N/A	Case report
Stiller-Ostrowski 2010	1	M = 1	19	N/A	Case report
Bramley et al., 2012	2	M = 1 F = 1	10–15	N/A	Case series
McNally et al., 2024	1	F = 1	16	N/A	Case report
4.4 Normative return-to-play rehabilitation					
Gunter et al., 2018	1	F = 1	14	N/A	Case report
Marsh, Fraser, and Marsh. 2013	1	M = 1	14	N/A	Case report

*Reported as average age. Did not specify mean/ median

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - Data extraction - Aims, methods, outcome measures and results

Authors	Study Aims/ Purpose	Methods/ Interventions	Primary Outcome Measure/s	Results
Theme 1 – Exercise, activity, and neuromuscular training interventions				
Exercise/activity interventions efficacy				
Chrisman et al., 2019	Evaluate feasibility/ acceptability for STEP with minimal in-person visits for persistent SRC, and efficacy for improving symptoms, health-related quality of life, and fear-avoidance.	• Randomised to CG (home stretching program(HSP)) or IG (6-week STEP). • In-person assessments at study entry and at 6 weeks (post-intervention). • Online weekly assessments of concussion symptoms during intervention, and surveys (PedsQL, FOPQ-P/C) at baseline, 6 weeks, 3 and 6 months. • Accelerometry was completed for 5– 7 days at baseline and at 6 weeks	• Feasibility and acceptance. • Health behaviour inventory. • PedsQL© • Physical activity (Accelerometry) • Fear of Pain Questionnaire – Parental/Child (FOPQ-P/FOPQ-C).	• Symptoms during exercise reported by 2/11 participants in CG and 7/19 in IG, No dropouts to symptom exacerbation. • Intervention significantly reduced symptoms (controlling: age, sex, and prior concussion). • The rate of symptom improvement is slower among youth with chronic symptoms (9–22 weeks & >22 weeks) compared to acute symptoms (<9 weeks). • PedsQL© improved in IG compared to CG. • FOPQ-P reduced overall, but not improved for IG. No differences were found between FOPQ-C scores between groups.
Gagnon et al., 2016	Examine the effectiveness of AR intervention for adolescents slow-to-recover after SRC.	• Participants (SCR slow recovery with persistent symptoms ≥ 4 weeks post injury). • The AR program lasted 6 weeks with testing throughout (10-day intervals).	• Post-concussion symptoms (PCS) using the post-concussion scale (PCSS).	• Mean time of assessment post injury: 7.9 weeks (range: 3.6–26.2 weeks). Significant decrease of PCS at 6 weeks.
Micay et al., 2018	Feasibility of post-acute structured AE intervention in concussed and symptomatic adolescent athletes.	• Randomized (IG or CG). • CG -6-stage activity progression, assessed at 1,2,3,4 weeks. • RTP decisions by a blinded sports medicine physician. • IG completed standardised AE intervention on day 6 post-injury, plus a six-stage activity progression protocol. • AE - 8 sessions utilising a step-wise approach (duration/ intensity).	• Symptom status pre-post exercise sessions/ completion of intervention. • Clinical recovery: symptom status at weeks 1, 2, 3, and 4. • Medical RTP clearance date.	• AEs are not associated with symptom exacerbation. • PCSS were lower than pre-season scores post-exercise, but not statistically significant. • No statistically significant difference in mean time to medical RTP between groups (36.1 ± 18.5) in the EG and (29.6 ± 15.8) in the UCG. • Significant association between time to medical clearance and acute PCSS at day 5 postinjury (p=0.04). • IG-Significant decrease in symptoms between weeks 1 and 3. (Same result not achieved in UCG until week 4).
Leddy et al., 2019	Effectiveness of sub-symptom threshold AE vs stretching programme during acute recovery following SRC.	• Randomised to AE or SG < 10 days of sustaining SRC. • Treating physicians were blinded to the allocated groups. • AE-daily use of bike at target HR. • SG completed the progressive stretching program. (20 mins per day - HR monitor)	• Days to recovery. • Return to normal exercise tolerance on a treadmill. (Normal recovery <30 days, delayed >30 days).	• AE participants recovered in 13 days (median). • The SG took 17 days (Median) to recover. • There was a non-significant incidence of lower delayed recovery in the AE group, compared to the control group.

Authors	Study Aims/ Purpose	Methods/ Interventions	Primary Outcome Measure/s	Results
Leddy et al., 2021	Evaluate safety, efficacy, and adherence with prescribed early targeted heart rate sub-symptom threshold AE, compared to stretching exercise for SRC recovery.	• Completed PCSI and BCTT at initial visit and during up to 3 subsequent weekly visits, unless recovered sooner. • Randomly assigned to AE or placebo SG < 10 days of SRC. • AE - exercise at home at up to 90% of max HR-20 min a day. • SG - light stretches and breathing exercises that would not elevate HR.	• Clinical recovery within the 4-week intervention period. • Development of PCS beyond 28 days post-injury.	• The AE group was more likely to recover within 4 weeks after injury compared to the SG. • During the entire 4-month follow-up period, the median days to recovery for the AE group was 14 days (IQR 10–25) versus 19 days (13–31) for the SG. • There was also a reported 48% reduced risk of developing PCS in AE. • There were no adverse events reported for the study duration.
Seehusen et al., 2021	To examine the association between objective physical activity metrics and clinical recovery after concussion.	• Recruited following a concussion diagnosis and presenting within 14 days with a PCSI score ≥ 9. • Evaluated at two time points + reported PCSI scores: Initial visit and at RTP clearance visit. • Between visits, participants wore an activity tracking device (calculated average: steps per day, sessions per week; time spent during each specific exercise session; and maximum HR during each session.	• RTP clearance • PCSI • Activity data (steps per day/ exercise frequency/ exercise duration/ exercise intensity.	• 17 participants (53%) required ≥28 days to RTP. • Those ≥28 days RTP reported significantly greater symptom severity at initial assessment than those who RTP within 28 days of injury. • First 2 weeks, the RTP clearance ≥ 28 group took significantly fewer steps per day, exercised fewer days per week, significantly lower average total exercise volume per week compared with the <28 day RTP group. • The highest classification accuracy between groups was ≥10,250 average steps/day and completing ≥4 exercise sessions per week, and ≥135 total minutes per week.
Lishchynsky et al., 2019	To evaluate the association between the amount of MVPA during the first 3 days following concussion diagnosis and time to medical clearance (days) to RTP in youth ice hockey players.	• Completed SCAT3, vestibulo-ocular/ cervical spine function tests. • Physicians were blinded to activity and baseline symptom reports. • Participants are advised to rest until the acute symptoms resolution. • Exposure: amount of time (minutes) spent in MVPA in the initial 3 days (72 h) immediately following concussion diagnosis. MVPA was measured with raw accelerometer data categorised into (sedentary, light, moderate, vigorous) • 3 days exposure window to ensure accurate initial physical activity.	• SCAT3 (PCSS scores) • RTP clearance timescale. • Symptom duration. • Number of days to initiation of activity (injury date to initiation of activity) • Date of return to school (self-reported).	• Participants in both groups presented to the clinic a median of 4 days post-injury (low activity: IQR 3-5 days and high activity: IQR 3-7 days). • The median total number of symptoms reported at initial appointment was 13 out of 22 (IQR 9-20) for the LAG and 10 (IQR 6-14) for the HAG. • Median symptom severity score LAG = 31 (IQR 14-51) and LAG = 14 (IQR 8-29). • Median amount of time LAG spent in MVPA = 110.7 min (IQR: 76.2–131.0 min). HAG = median of 217.2 min in MVPA (IQR: 184.2-265.2 min). • Four participants (13.3%) did not miss any school days and RTP the day after injury. • HAG took more time for RTP clearance (p = 0.041). No significant differences in secondary measures.

(*Continued*)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - (*Continued*)

Authors	Study Aims/ Purpose	Methods/ Interventions	Primary Outcome Measure/s	Results
Neely et al., 2023	Determine whether sleep behaviour (duration/ timing) and/or physical activity (steps/ day, exercise frequency, duration, intensity) during the first month post-concussion are associated with PCS development.	• completed an initial medical/ injury history questionnaire. • Participants remotely tracked physical activity, sleep, and exercise over 2 weeks following their initial assessment.	• PCSI. • Main outcome variable - persisting post-concussion symptoms (PPCS): A symptom resolution of >28 days or <28 days (defined as no PPCS).	• 47% developed PCSS. A higher proportion reporting PCSS had an ADHD diagnosis. • The PPCS group took significantly fewer steps per day and exercised fewer days per week. They also spent more time in bed awake. Exercise duration/ intensity and sleep times were not significantly different between groups. • Odds of PPCS significantly increased with fewer exercise sessions per week. • Completing more exercise sessions lasting >15 minutes during the recovery period reduced the risk of PPCS.
1.2 Early aerobic exercise testing safety and effects				
Morissette et al., 2020	To compare cardiorespiratory (CR) response to a graded AE challenge between adolescents with symptomatic sport-related concussion (SSRC) and healthy control subjects.	• Graded AE challenge using BCTT. • Heart rate (HR), Rate of perceived exertion (RPE), and symptom status data were collected every minute during treadmill testing, whereas cardiorespiratory data were recorded every minute using 20-s averages. • The BCTT is either the participant reached volitional fatigue or the participant experienced a symptom-limited threshold (An increase in self-reported symptom status by two or more points using a 10-point standardized Likert scale). • Peak HR, blood pressure (BP), and cardiorespiratory data were recorded, and a post-exercise PCSS was completed.	• HR • consumption (VO_2) • Carbon dioxide production (VCO_2) • Minute ventilation (VE) • Taken at rest/ test termination. • Change from rest in variables (ΔHR, ΔVO_2, ΔVCO_2, and ΔVE) during the first five stages of the BCTT.	• Symptoms/symptom severity were higher in the SSRC group compared with CG ($p < 0.0005$). • SSRC group achieved a lower peak stage during BCTT compared to CG ($p < 0.0005$), lower peak RPE ($P = 0.023$), lower relative peak VO_2, VCO_2, and VE compared with CG ($p < 0.0005$). • SSRC - lower peak SBP($p < 0.0005$), higher peak diastolic BP (DBP) ($p < 0.0005$). • Peak HR + %HRmax is lower among male SSRC patients compared to male controls ($p < 0.0005$). Female SSRC - higher peak HR + %HRmax ($p = 0.016$) ($p = 0.018$). • The highest BCTT stage was reached lower level among females vs males ($p = 0.04$), peak relative VO_2, VCO_2, and VE all lower in female control vs male controls ($p < 0.0005$). • No group-based differences ΔHR, ΔVO_2, ΔVCO_2 ΔVE. • RPE is higher among SSRC patients compared with CG.
Chan et al., 2018	To examine the safety and tolerability of an AR program for adolescents who are slow to recover from an SRC. Secondary aim: To estimate the treatment effect for this intervention.	• Participants were randomized to either TAU or TAU plus an AR program. • TAU consisted of symptom management and RTP advice, return-to-school facilitation, and physiatry consultation. • The AR program involved in-clinic sub-symptom threshold aerobic training, coordination exercises, and visualisation and imagery techniques with a physiotherapist (mean, 3.4 sessions) as well as a HEP, over 6 weeks.	• Adverse events. • In-clinic symptom exacerbation during AE. • PCSS	• 6 adverse events in each group. The most common was PCSS change (>10 points) from one week to the next. Adverse events were reportedly unrelated to the intervention. • AR group, there were worsening symptoms (>1 on a 0-10 scale) in 37% (11/30) of clinic sessions involving AE. Most symptom exacerbations occurred in the 1st or 2nd AE session. All symptom exacerbations resolved within 24 hours, in most (10 of 11 instances) within 15 minutes of discontinuing exercise. • All participants were able to tolerate 15 minutes AE without symptom exacerbation by their last clinic session.

Authors	Study Aims/ Purpose	Methods/ Interventions	Primary Outcome Measure/s	Results
Leddy et al., 2018	Evaluate the systematic assessment of exercise tolerance in adolescents shortly after SRC and determine the prognostic utility of this assessment.	• Participants were randomized to BCTT within 10 days of sustaining a concussion or standardised care (no test). • Intervention: BCTT on the day of presentation to the clinic (visit 1), Secondary follow-up visit, approx. 14 days after, or approximately. 21 days post-injury.	• Days to recovery and typical (≤21 days) versus prolonged recovery (>21 days).	• Days to recovery + typical vs. prolonged recovery were not significantly different between groups. Symptom scores decreased over 14 days in both groups, and symptom scores did not increase the day after visit 1 (Intervention group), BCTT. • A lower HR threshold on visit 1 was strongly associated with prolonged recovery time and prognosis. BCTT used within one week post SRC did not delay recovery.
Worts et al., 2021	Examine acute effects (pre, during, post-intervention) of two different AE intensities/ rest on autonomic, oculomotor, vestibular function, and symptom burden in patients with SRC and compare their responses to sex-matched, age-stratified, non-concussed (HEALTHY) student-athletes.	• Two study groups were observed: (1) student athletes treated for an SRC within 3-7 days post-injury. (2) non-concussed (HEALTHY) athletes. • All participants were randomised to a single session of one of three 20-minute interventions using a double-block design (Sex: Male/Female; Age: 13–15 and 16–18) • (1) 40% of Age-Predicted HRMAX (40HR). • (2) 60% of Age-Predicted HRMAX (60HR). • (3) seated, non-exercise (NOEX). • Participants were fitted with test equipment (Blood pressure cuff and HR monitor)	• HR variability (HRV) using an electrocardiogram (ECG) strap. • Blood pressure. • Mean arterial pressure (MAP) was derived from SBP and DBP. • RPE using Borg's Category Ratio-15 scale. • King-Devick (K-D) test. • Vestibular Oculomotor Screening (VOMS). • PCSS.	• SRC group reported greater provocation during VOMS (more symptoms & more severe symptoms than HEALTHY participants). • Near point convergence (NPC) distance and K-D times did not exhibit an injury status group effect. 15/16 SRC group completed AE session, 13/15 who completed AE displayed improvements beyond reliable change on at least one of the clinical test measures. None displayed a group effect during the orthostatic stressor. • In the SRC and HEALTHY groups that exercised, participants generally exhibited improvements in clinical measures, even while exerting at 40–60% of HRMAX. • Pre-exercise HRV and MAP were significantly different (p's<0.001) during treatment but returned to pre-exercise values within 5 min of recovery in both groups. Both groups exhibited similar reductions pre- to post-intervention for symptom severity and count ($p < 0.05$), three VOMS items ($p < 0.05$) not K-D time.
1.3 Neuromuscular training intervention				
Howell et al., 2021	To determine the feasibility of an 8-week neuromuscular training (NMT) program initiated upon RTP clearance following concussion.	• 3 assessments: 14 days post-concussion, after RTP, and 8 weeks after Visit 2. • Visit 1: Independent physician assessment. No study interventions between Visits 1 and 2. • Following RTP clearance, participants attended visit 2: Randomized to NMT or CG. Attended a final assessment 8 weeks after Visit 2. • NMT: 2x per week with athletic trainer.	• PCSI • Patient Reported Outcomes Measurement Information System • Tampa Scale of Kinesiophobia. • Reaction time, single/dual task gait, and balance outcomes.	• No adverse events in either participant group during the intervention period. • The median time to symptom resolution was 17 days [interquartile range (IQR) = 12, 43] for the intervention group and 19 days [IQR = 11, 49] for the standard-of-care group. • The median time to medical clearance to RTP for the intervention group was 22 [IQR = 15, 49] days, and 35 [IQR = 15, 62] days for the standard-of-care group. • 55% of the NMT group were able to complete all 16 sessions.

(*Continued*)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - (*Continued*)

Authors	Study Aims/ Purpose	Methods/ Interventions	Primary Outcome Measure/s	Results
Howell et al., 2022	To examine the efficacy of an NMT intervention on acute sports-related time-loss injury (SRTLI) over the subsequent year relative to standard of care.	• Twenty-seven youth athletes assessed initially post-concussion and after RTP clearance, randomly assigned to NMT intervention or standard of care group. • The intervention included guided strength exercises with a landing stabilisation focus. Standard of care received no recommendations. • For the subsequent year, a completed log of sports-related injuries.	• Time from post-concussion RTP clearance until injury. • Incidence of musculoskeletal injuries, plus the total number of injuries sustained over the monitoring year, and injury severity.	• First year after RTP clearance, SRTLI is more common among CG relative to NMT intervention (75% [95% CI, 48%-93%] vs 36% [95% CI, 11-69%]). • Adjusting for age and sex, the hazard of subsequent injury in the CG group was 3.56 times (95% CI, 1.11-11.49; p = .0334) that of the NMT group. • Age and sex-adjusted injury incidence was 10.2 per 1000 competitive exposures (95% CI, 3.7-28.4) in the CG & 3.4 per 1000 (95% CI, 0.9-13.4) in the NMT group. • Adjusting for age & sex, Injury incidence was higher for the CG vs NMT intervention.
Theme 2 – Novel treatment interventions				
Congeni et al., 2022	To determine the safety and efficacy of head and neck cooling when applied up to 8 days after concussion among adolescent athletes.	• Within 8 days of SRC: Randomized CG/ treatment group. • The treatment group received both standard treatment and head and neck cooling therapy using the Pro-2cool device. • SCAT5 symptom scores charted post-injury assessment visit, before and after treatment/ standard treatment (0-8 days); 72 hours, before/ after treatment or standard treatment (±24 hours from initial visit); 10 days (±3 of initial visit), 4 weeks (28 days ± 7 days)	• SCAT5 symptom severity score assessment	• Significant main effects (group and time, $p < 0.01$ for each), as well as interaction between time and group ($p = 0.017$) for absolute differences in SCAT5 total symptom severity scores from injury to follow-up. • Between groups mean differences (95% CI) at each time points demonstrated significantly greater and earlier reduction in symptoms for the treatment group: initial visit posttreatment 14.1 (6.7-21.4), 72 hours visit pre-treatment 9.2 (-3.3- 15.1), 72 hours visit post-treatment 15.5 (8.2-22.8), 10 days visit 5.8 (1.9-13.5), and 4 weeks visit 0.7 (-9.5-10.9). • No significant adverse events were reported..
Smith et al., 2024	To determine whether an investigational head–neck cooling device, Pro2cool, can better reduce symptom severity compared with standard post-concussion care in early adolescent athletes after a SRC.	• Randomised to CG or treatment group (standard care plus head–neck cooling therapy using Pro2cool device applied at two different post-injury time points). • SCAT5 completed (IVPreT) within 8 days of injury and immediately following this, received control/ treatment. • Participants returned 72 hours after the IVPreT for the 2nd round of treatment. SCAT5 tests completed pre-/post-intervention/control. SCAT5 scores were gathered at day 10 + 28.	• SCAT5 symptom severity scores. • Adverse reactions.	• The treatment group experienced a 14.4% greater reduction in SCAT5 symptom severity scores at the initial visit post-treatment, a 25.5% greater reduction at the 72-hour visit post-treatment, and a 3.4% greater reduction at the 10-day visit compared with CG. • Significant group absolute differences from baseline SCAT5 scores favouring the Pro2cool treatment arm at the initial visit post-test, at the 72-hour visit post-test, and at the 10-day visit. • 36 adverse events reported, 13 associated with the device (3 - moderate severity). No significant association between experimental group/event severity.

Authors	Study Aims/ Purpose	Methods/ Interventions	Primary Outcome Measure/s	Results
Roby et al., 2021	Describe the effect of HBO_2 therapy on reducing initial symptom burden in acutely concussed high school student-athletes compared to two different placebo treatments.	• Eight participants randomised and blinded to: (1) Hyperbaric therapy with 100% oxygen (HBO_2), 100% O2 at pressure (n = 3) (2) Hyperbaric therapy with medical-grade air (HBA) (n = 3) (3) 100% O_2 normo-baric therapy (O_2 Therapy) (100% O_2 therapy with negligible pressure) (n = 2). • PCSS was completed before and after treatment.	• PCSS • SAC • Number of days from injury until the physician permitted the student-athlete to return to activity.	• No statistical analysis was reported between treatment arms. The below % reported represents a total symptom score reduction. • HBO_2 RTP average of 13.7 ± 5.1 days (median = 15 days). Participant 3 (90.2%), 6 (77.8%), and 7 (93.5%). • HBA Therapy RTP average of 13.0 ± 5.7 days (median = 13 days). Participants 2 (87.5%) and 5 (81.1%). • O_2 therapy RTP average of 19.0 ± 16.5 days (median = 11 days). Participants 1 (83.3%) and Participant 8 (95.7%). Participant 4 (93.1%) at visit 4.
Miller et al., 2022	Examine the use of Docosa-hexaenoic acid (DHA) for the treatment of SRC in an adolescent population. (To also determine if adverse reactions occurred with higher dosage of DHA and to examine athlete outcomes following DHA supplementation.)	• Participants were randomised to the control or intervention arm within 4 days of suspected SRC. • Participants were advised to take 4 DHA capsules daily (total 2000mg per day). • Participants were evaluated at weeks 1,2,4, and 12.	• SCAT3 • Adverse effects reported by a sports medicine physician. • Drug compliance. • Time to 'symptom-free' and time to initiation of RTP.	• The use of 2000 mg per day of DHA was well tolerated. • This supplementation did not significantly improve recovery times in adolescent athletes.
Haider et al., 2025	To evaluate the Face cooling (FC) response in adolescent athletes with SRC while they were symptomatic after clinical recovery.	• Participants attended the research lab within 10 days of injury (visit 1) and at 2 weeks following successful graded return to play. • FC treatment was applied following a standardised approach, and parasympathetic responses were recorded.	• HRV/Mean arterial BP • HR/MAP/Cardiac Output (CO)/ Stroke volume (SV) • Root mean square of the successive differences (RMSSD) of R-R interval (RRI) • Low frequency (LF) and High frequence (HF) ratio	• During FC, no significant differences were found between groups at the initial visit in rate of change in HR, mean arterial BP, RMSSD, or LF/HF ratio. • No differences amongst those with delayed recovery from concussion. • Previous history or concussion had a significant effect on HR and HRV responses to FC intervention.

(*Continued*)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - (*Continued*)

Authors	Study Aims/ Purpose	Methods/ Interventions	Primary Outcome Measure/s	Results
Theme 3 – Pitch-side management approaches, sub-acute assessment, and monitoring				
3.1 Removal VS no removal from play following concussion				
Olson et al., 2020	Examine: Delayed symptom onset in youth athletes sustaining SRC; Effect of symptom onset on initial symptom severity, length of recovery, and consequent development of delayed recovery; Impact of symptom onset on pitch-side management.	• Prospective recruitment of youth athletes sustaining SRC from the Pan AM concussion programme. • Sideline assessment data not recorded. • Reported standardised procedure for athletes at baseline/ and following injury (Pan AM concussion programme). All patients were assessed following an SRC by a single neurosurgeon. Follow-up included a weekly visit or dictated by the rate of recovery. Recovery was defined as the resolution of symptoms or a normal neuro exam. Patients completed the graded RTP (GRPT) protocol prior to full RTP.	• Delayed symptom onset (DSO) ≥ 15 mins from injury. • Early symptom onset (ESO) <15 mins since injury. • Length of clinical recovery (injury date - final follow-up). • Delayed recovery reported as >28 days to resolution.	• Median time to initial assessment was 5.0 days, and the median PCSS scores were 22. • DSO athletes were assessed at 6.5 days, whilst ESO athletes were assessed at 5.0 days. There were no significant differences between median symptom scores at initial medical assessment between ESO/DSO athletes. • DSO were cleared to play at 23 days (median), and those with ESO at 21 days (median) following injury. • SRC patients who experienced ESO were significantly more likely to be removed from play (RFP) compared to those with DSO.
Elbin et al., 2016	To compare recovery time and related outcomes between athletes who were immediately RFP and athletes who continued to play with an SRC.	• Compared neurocognitive performance, symptoms, and recovery time between 35 athletes immediately removed after an SRC (REMOVED group), compared to 34 athletes who continued to play (PLAYED group) with SRC. • Neurocognitive and symptom data were obtained at baseline and at 1 to 7 days and 8 to 30 days after an SRC.	• ImPACT scores • Recovery time (defined as the total number of days from the date of injury to the date of receiving RTP clearance).	• PLAYED took longer to recover than the REMOVED group (44.4 ± 36.0 vs 22.0 ± 18.7 days; p = .003) & were 8.80 times more likely to demonstrate protracted recovery (≥21 days) (p < .001). • Removal from play status was associated with the greatest risk of protracted recovery (adjusted odds ratio, 14.27; p = .001) compared with predictors (eg, sex). • PLAYED group exhibited significantly worse neurocognitive and greater symptoms vs REMOVED.
Zynda et al., 2022	Identify the frequency of continued play (CP) following SRC, defined as continuing athletic activity on the same day following a suspected SRC, characteristics associated with CP, and whether CP was associated with worse outcomes.	• Participants were divided into two groups: Those who continued to play following SRC (PLAY) and those who did not (NO PLAY). NO PLAY - Immediately removed from play. • Data were collected prospectively at the time of initial clinic visit and three months following enrolment. • Assessments at enrolment: concussion symptom log, mBESS, ImPACT. Psychological assessments included: Generalized Anxiety Disorder Scale (GAD-7), Patient Health Questionnaire (PHQ-8), Brief Resilience Scale (BRS).	• Concussion symptom log • ImPACT scores • mBESS scores • GAD-7 • PHQ-8 • BRS • Time to symptom resolution. • Time to RTP	• 52.4% were in the PLAY group and had less severe balance problems from the day of injury (p = 0.016), but reported trouble with falling asleep, concentrating, and memory at their initial visit (p < 0.05). • No significant group differences were found when comparing remaining symptoms from the day of injury, symptom log. Symptom scores – injury day not significantly different. • PLAY group: worse ImPACT scores (p = 0.042). • No significant difference between groups for time to symptom resolution, time to initiate RTP (p = 0.381), or RTP time (p = 0.776). • Greater symptom severity score at the initial visit and longer time to presentation were associated with prolonged recovery in both groups (p < 0.01).

Authors	Study Aims/ Purpose	Methods/ Interventions	Primary Outcome Measure/s	Results
3.2 Assessment and monitoring				
Cosgrave et al., 2023	To characterize the incidence, severity, and recovery of SRC in schoolboy rugby players and explore whether the SCAT, Cogstate Brief Battery (CBB), and the K-D test can be used to monitor concussion status through to full recovery.	• Participants attended for baseline testing during preseason and were followed up over the 2016-2017 playing season. • Participants diagnosed with SRC during the season attended for post-concussion testing on a weekly basis until fully recovered. At each post-concussion appointment, the full battery of investigations was repeated. • Participants were prescribed an individualised rehabilitation programme, which is considered the gold standard management of concussion and was not part of the study in itself. • Participants adhered to the Irish Rugby Football Union (IRFU) concussion guidelines (IRFU Guide to Concussion in Amateur Rugby Union) for their age group.	• Questionnaire (sleep, mood, stress, anxiety, depression, quality of Life • symptom checklist, concussion predictive and modifying factors) • SCAT3 • CBB • K-D test	• 64 (48%) reported symptoms on baseline SCAT. Sixteen concussed participants, 9 (56.3%), had a history of SRC. • CBB was consistent with clinically assessed recovery on 27 (51.9%) occasions. • K-D test was consistent with clinically assessed recovery on 32 (64%) occasions. Return to baseline for K-D test - 11 days(mean time) (±8.7, range 2-33 days). • 19 (38%) occasions, both the CBB + K-D test were consistent with the clinical assessment of recovery. On seven occasions, there was agreement on the presence of concussion, and on 12 occasions, recover agreement. • 10 (20%) occasions participants passed CBB and K-D when not deemed clinically recovered from concussion. 1 (2%) occasion, a participant failed CBB and K-D when they were deemed recovered from concussion. • CBB and K-D test results differed on 20 (40%) occasions. CBB agreed on 7 (14%) of occasions. K-D agreed with assessment-13 (26%) occasions.
Emery et al., 2021	To assess factors associated with clinical recovery after concussion in youth ice hockey players.	• Players with suspected SRC were referred to a sports medicine physician for diagnosis, and SCAT3/5 was completed. • Two accelerated failure time models are used to estimate days to RTP clearance. • Other covariates were time to physician first visit (≤7 and >7 days), age group (11-12, 13-14, and 15-17 years), sex, league type (body checking and no body checking), tandem stance (modified Balance Error Scoring System result ≥4 errors), number of previous concussions.	• Symptom scores • RTP clearance (days) • Time to first clinician visit • SCAT3/5 data • Time to clinical recovery was recorded as the time between concussion and physician clearance to RTP.	• The median time to clinical recovery was 18 days. Longer time to first physician visit (>7 days) (time ratio [TR], 1.637 [95% confidence interval (CI), 1.331-1.996]) and greater symptom severity (TR, 1.016 [95% CI, 1.012-1.020]) were predictors of longer recovery. • Longer time to first physician visit, headache (moderate/severe), and poorer tandem stance were predictors of longer recovery. • RTP longer for players with >7 days to assessment. • Poorer scores on tandem gait, higher symptom evaluation score, and greater headache severity were associated with longer time to clinical recovery.

(*Continued*)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - (*Continued*)

Authors	Study Aims/ Purpose	Methods/ Interventions	Primary Outcome Measure/s	Results
Kelleher et al., 2014	To implement standardised care via a mild traumatic brain injury (mTBI) clinic (pilot program to provide concussion evaluation and assist RTP) and survey players post-season.	• 405 students playing football were selected. • Athletes with SRC either attended a local healthcare provider or the mTBI clinic. Appointments included mTBI symptom screening + neurological exam. • Symptomatic athletes - rescreening occurred after the RTP algorithm specified interval, and, if appropriate, athletes were cleared for graded RTP. If symptoms persisted beyond 2 weeks, referral to a specialist was performed.	• IMPACT scores • ACE evaluation • Survey: concussion injury reporting and medical management journey.	• 17/28 students had a previous concussion history. The mean amount of time until RTP was 16.9 days. • 266 respondents to the survey. 40 (15%) reported mTBI symptoms during the season. Of the 15%, 9 (22.5%) did not report their symptoms to anyone, and 18 (45%). • Unreported mTBI episodes are lower (22.5%) than published self-reported mTBI rates.
Theme 4 – Individual accounts of concussion management in specific populations				
4.1 Chiropractic management				
Olson et al., 2018	To describe the chiropractic management and return-to-play of 3 student football players who had experienced a concussion.	• Concussion, neurological and musculoskeletal evaluation, computerised neurocognitive (ImPACT) and balance testing. • The course of treatment was initiated, consisting of: 1. Instrument-assisted soft tissue mobilization, 2. Myofascial release (MRT)Concussion. Spinal manipulative therapy (SMT) and therapeutic exercises were provided to help treat cervical spine symptoms and dysfunction.	• Symptomatology • ImPACT scores • SCAT3 • RTP	• **CASE 1: Final visit:** Memory composite (verbal, visual), visual motor composite, reaction time composite, and impulse control composite scores returned to baseline. RTP - 12 days. Monitored throughout the year - no injury/ symptom exacerbations. • **CASE 2 Final visit:** Completed RTP protocol + RTP (24 days) - no further setbacks. • **CASE 3 Follow-up:** Follow-up took place 12 days after the initial visit. The participant had been symptom-free since the day after his initial treatment and RTP. Computer neurocognitive test scores – baseline levels.
Hunt et al., 2018	Concussed skateboarder undergoing chiropractic treatment.	• Intervention: Chiropractic care. No control. Intervention included - Evaluated for subluxation and 'adjusted' where necessary, twice per week.	• Symptom relief was reported by the patient.	• 5th visit - patient reported overall 80% improvement and reported 'feeling freer'/ sight and breathing had improved. Ongoing 'sore head' when sneezing. • Reported resolution of pupillary light reflex testing and photophobia by the 8th visit.
4.2 Concussion and mental health				
Sherry et al., 2021	Presentation of a case of a high-achieving 18-year-old female athlete (Rower) and her transitioning concussion care to mental health care.	• Seen by a neuropsychologist in clinic 4 days post-injury. Ongoing referrals to a vestibular therapist. • Symptoms persisted, plus anxiety/mood clinical profile. • Referred to cognitive behavioral therapy (CBT) and psychiatry. • She declined the referral to counselling due to distress and time constraints. Commenced on sertraline medication. • Decided to return to full rowing participation as a return to activity was expected to improve her mental health.	• Symptoms: cognitive/ ocular-motor/ Headache/ vestibular/ Anxiety/ mood. • PCSS	• The participant had been on sertraline for approximately 2 months and reported feeling "back to normal." She reported resolution of all concussion symptoms. • Significant improvement in her mood, anxiety, and no further panic attacks. Full participation in rowing and passed formal exertion testing with the school athletic trainer. • Cleared for RTP and academic activities and discharged. Advised to remain on antidepressant under the guidance of a psychiatry and her family doctor. • 6 months after being discharged, she was continuing with sertraline and felt it made a "big difference in her anxiety."

Authors	Study Aims/ Purpose	Methods/ Interventions	Primary Outcome Measure/s	Results
Tucker, DeCastro, and Brock 2019	To discuss two cases of post-concussion syndrome in adolescent athletes with a past personal or family history of attention-deficit/ hyperactivity disorder, anxiety, and depression, treated by a combination of cognitive behavioural therapy and medication, with ongoing persistent symptoms.	• **Case 1** - A 17-year-old male athlete (medical history of attention deficit/hyperactivity disorder (ADHD)) sustained an SRC while participating in football practice. • Amitriptyline 50 mg daily treatment was initiated for posttraumatic headaches and other associated symptoms. • **CASE 2 -** 16-year-old male with history of a simple febrile seizure during childhood + family history of posttraumatic stress disorder with depression. Followed the RTP protocol following a concussion. Further concussion sustained 1 month later. Return to school (RTS) was unsuccessful due to ongoing headaches, dizziness, and difficulty concentrating.	• Symptomatology • Mental health status	• **CASE 1:** Return to activity failed due to headaches and difficulty concentrating. The participant reported bullying. • Ongoing irritability, anxiety, and insomnia. Transitioned to home education. • Ongoing nystagmus, saccadic eye movements, impaired motor development, and vestibular impairment. New panic attacks (started on short-term lorazepam 1 mg). Did not attend RTS and instead gained employment in his family's landscaping business. • **CASE 2:** Difficulty with balance and right eye nystagmus. Transitioned to homebound education, started on amitriptyline. • Unsuccessful at the second RTS attempt at 5 months. Able to RTS after the school year. Ongoing depressive symptoms (commenced counselling and fluoxetine). Required hospitalization - mental health. • Developed grand mal seizures (started anti-epileptics.
Waldmeier-Wilhelm et al., 2019	To describe a case involving a 14-year-old girl who developed a pronounced and prolonged post-concussive syndrome and subsequent posttraumatic stress symptoms after (mild) traumatic brain injury	• Clinic presentation - worsening of school performance. 4 months before, involved in a snowboard accident (2-3 minutes of unconsciousness). • Evaluation in the ED was unremarkable, and the patient was discharged. • Persisting daily headache, concentration, memory problems, fatigue, dizziness, and blurred vision. • Neuropsychological assessment + physical therapy initiated. • Sports activities and school visits were reduced again, and oral magnesium was initiated. Relaxation therapy + psychological support continued. Occupational therapy declined.	• Symptomatology • Neurological exam • MRI Report	• 14 months after the injury, when returning to the injury site, a relapse of symptoms occurred. The patient complained about episodes of blurred vision, stomach aches, concentration problems, and sleep problems. Episodes of cardiac presyncope were observable. • Repeated neurological and neuropsychological examinations did not show any abnormalities. There were no obvious symptoms of depression. As the general physical examination was normal as well, no further additional investigations were initiated. • The relapse of symptoms was interpreted as a sign of posttraumatic stress, triggered by the confrontation with the location where the accident occurred (mountain area). Patient refused psychiatric evaluation- psychological coaching was continued.

(*Continued*)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - (*Continued*)

Authors	Study Aims/ Purpose	Methods/ Interventions	Primary Outcome Measure/s	Results
4.3 Specialist medical case reports/series				
Lunardini et al., 2009	Report on treated patient: Treating post-concussion symptoms in a young diabetic person	• Young male athlete with an SRC whilst playing football. Suffered headache, dizziness, partial loss of vision left eye, and right arm numbness. Initial high blood sugars (240 mg/dL). • Athlete RFP and within 30 minutes (dizziness and headache symptoms). The patient was reviewed by the physician and sent home with a diagnosed grade 2 concussion + post-injury instructions. • Concentration difficulties/ headaches. • 6 weeks - referred to a neuropsychologist.	• Symptoms reported • Initial high blood sugars • Concentration and delayed recall testing (Using the SAC test) • Neuropsychologist review (at 8 weeks)	• Concentration and delayed recall improved when tested over the following week. • Following Neuropsychologist review (at 8 weeks), headaches had reduced to once every three days and reported functioning 90-95% of normal. • The young athlete was cleared after 4 months, following completion of a controlled aerobic exertion test and resolution of all symptoms. In an interview with the participant at 4 years after the initial injury, he reported that it took 6 months to fully return to his normal 'thinking speed'. No long-term effects reported. • Known to have a history of poor blood glucose control.
Stiller-Ostrowski 2010	Case of a National Collegiate Athletic Association Division I men's lacrosse athlete with a fourth cranial nerve injury as the result of a minor traumatic blow	• Minor traumatic blow during lacrosse. Referred to neurology/ ophthalmology. MRI of the orbits was normal. • Diagnosed with contusion/ stretch injury to the oculomotor nerve. • Prognosis of full resolution of symptoms (3-4 weeks) and advised to continue to use the eye normally. • Visual disturbances not resolved by 4 weeks. A battery of tests was performed in week 5, and the original diagnosis was rejected. • Diagnosed with palsy of the right superior oblique muscle secondary to right CN IV palsy. Prognosis of full recovery 3 to 4 months postinjury.	• Symptoms • Injury risk • RTP clearance	• 5 months postinjury - No diplopia about 90% of the time (Double vision occurring in extreme left gaze). Diplopia did not have any functional implications. • The neuro-ophthalmologist informed the patient and his family of the risk of returning to high-velocity sports with even intermittent diplopia, but the patient received medical clearance to return to competitive lacrosse. • By 6 months post-injury, he reported subjective normal vision. At that time, his risk of reinjury was equivalent to his preinjury risk. The patient was referred for a custom-moulded mouthguard. • Able to return to full competition during the fall 2008 season and competed fully in both the fall 2008 and spring 2009 competitive seasons with no recurrence of symptoms. 16 months post-injury, the patient is back to his pre-injury level of competition.

Authors	Study Aims/ Purpose	Methods/ Interventions	Primary Outcome Measure/s	Results
Bramley et al., 2012	To report on the effectiveness of steroids for severe posttraumatic headache, along with recommended treatment strategies for acute pain management and prevention.	• **Case 1:** 15-year-old with worsening symptoms. Standard 6-day methylprednisolone 4-mg dose pack initiated. Resulted in significant headache reduction/ improved mood. • Completed steroid taper/400 mg ibuprofen. Discontinue acetaminophen with codeine. • **CASE 2 -** A 10-year-old boy referred with severe headache, dizziness, and vomiting 21 days post SRC. • Ongoing pain, and a standard 6-day tapering methylprednisolone 4-mg dose pack was added. • 7 days later, symptoms returned. Amitriptyline increased to 50 mg, started 14-day methylprednisolone.	• Symptom reporting • medication use • Return to school/ physical activity.	• **CASE 1:** 13 days post injury, ongoing headache caused by activity resumption. Methylprednisolone completed. • Sleep remained disturbed, melatonin was added, plus increased trazodone. Continued gabapentin/ibuprofen. Instructed to start school on half days. • RTS full time at 8 weeks after injury - close to baseline with only occasional headache. Trazodone weaned to 50 mg, to discontinue in 1 week if sleep is resolved. • Melatonin/gabapentin was discontinued a few weeks later. • 12 weeks after the injury patient had returned to baseline. • **CASE 2:** After a 14-day tapering course of medication, subjective decrease in frequency/intensity of headaches and complete relief of emesis. The headache gradually resolved and did not return after the completion of the second course of steroids. • 11 weeks post injury, returned to school full time, increasing physical activity, and symptom-free.
McNally et al., 2024	To report on a case of worsening concussion symptoms and an MDT approach to care	• 16-year-old female who sustained a concussion falling off a horse, and who failed to respond to standard concussion care. • After 2 months of worsening symptoms, she was referred to an MDT clinic. • Interventions included: Stopping amitriptyline, continuing to take magnesium and riboflavin supplements; Physiotherapy – vestibular and neck treatment; Supervised exercise programme with an athletic trainer; PT; Advice on improving lifestyle (hydration and not skipping meals).	• Behaviour rating scales • SCAT5 symptom scale • PedsQL • NeuroQOL Sleep • Activities-Specific Balance Confidence Scale. • Neuropsychological testing	• Full recovery of symptoms at final treatment session (Visit 3 – 75 days post injury). • Continued to present with mild functional gait disturbance. • Cleared by neurologist to complete a gradual return to horse riding at 82 days post injury. • Patient continued on medications for her underlying anxiety and post-traumatic stress disorder (PTSD).
4.4 Normative return-to-play rehabilitation				
Gunter et al., 2018	Physical therapy management and adapted return-to-sport protocol for an equestrian athlete with history of multiple concussions.	• 14-year-old female with 3rd concussion in 3 years. • Return-to-riding protocol tailored. • 8 physical therapy sessions over 4 weeks for vestibular training, aerobic conditioning, and cervical/ core exercises, plus equestrian exercises.	• Symptom reporting • Dynamic visual acuity test • Joint position error • BESS	• Final evaluation: Patient reported no symptoms at rest, with exercise, or when testing the vestibular-ocular reflex. Improvements were noted in the dynamic visual acuity test, joint position error, and BESS, with changes in the BESS exceeding minimal detectable change. • The patient completed the full return-to-riding protocol in 8 weeks and returned to competition.

(*Continued*)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - (*Continued*)

Authors	Study Aims/ Purpose	Methods/ Interventions	Primary Outcome Measure/s	Results
Marsh et al., 2013	Report on treated patient: Case report of young 14 year 14-year-old athlete with hockey hockey-related concussion	• Diagnosed with a mild concussion and discontinued hockey/ activity that exacerbated symptoms. • Patient advised to return for 1 1-week review.	• Symptoms reported • Follow-up attendance • RTP and symptom reporting.	• Symptoms: Feeling 'fuzzy' and frontal headache. Did not attend 1 week follow-up, but instead was reviewed 3 weeks later. • Completed the stages of RTP, with no significant symptoms reported. No timescale reported.

Discussion

The aim of this scoping review was to complete a systematic search of published primary research literature on the medical management, rehabilitation approaches, and outcome data for the concussed youth athlete across all stages of injury management. Following a rigorous search, we identified studies reporting a variety of intervention approaches and grouped these into four key themes. Thirty-six studies were included in the review, 14 studies reported on exercise, early activity and neuromuscular rehabilitation interventions, six studies reported on pitch-side management of concussion or ongoing monitoring/management of injured youth athletes, five studies reported on the use and outcomes of novel treatment interventions and 11 studies reported on individual case report accounts of concussion management in specific populations.

Exercise, activity, and neuromuscular training interventions

In this review, we found level 2 and 4 evidence (RCTs and cohort studies, respectively) that remaining active following a concussion was beneficial, provided that participants worked at a sub-symptom threshold (40,49,56,59). Relative rest and exercise interventions with a low aerobic exercise requirement, such as stretching programmes, were generally less effective at reducing symptoms and facilitating return to play (37,40). Generally, those who took fewer steps, or exercised vigorously had poorer outcomes than their counterparts working within a graded STEP (54). There was evidence to suggest that exercise sessions lasting more than 15 minutes reduced the risk of developing further post-concussion symptoms (29). This is particularly useful in planning and supporting youth athletes with their activity and exercise plans post-concussion, and the evidence found in this section of the review continues to support the safety and efficacy of early aerobic exercise interventions, provided the athlete is working at a sub-symptom threshold. In order to establish these thresholds, provocative exercise testing such as the Buffalo Concussion Treadmill test (BCTT) can be used. The evidence found in this review reported that completing such a test within 10 days post-concussive injury did not delay participant recovery and cardiovascular responses differed between control and SRC participants, but that the test was safe to administer (39,61). STEP programmes, alongside provocative aerobic exercise testing, were found to be safe and provided therapeutic benefits (35,61). By grading up exercise in a controlled and progressive manner, those with recent SRC were able to build up to tolerating 15 minutes of aerobic exercise by the final session in one study (60). Chan et al. reported that symptom exacerbation was limited to 24 hours, with resolution within this period, and the majority of exacerbations were self-limited to 15 minutes before resolution (60). This is particularly useful for those clinicians administering exercise interventions for concussed youth athletes, and provides evidence that symptom exacerbation is not an indicator to cease STEP interventions. Limited data were presented for neuromuscular training programmes, with two studies suggesting that NMT is safe and may provide longer-term benefits, such as reduced future sports-related injuries over future seasons (34,38). The evidence here may help clinicians develop post-concussion, future injury prevention programmes to support their athletes throughout future sporting seasons.

Pitch-side management approaches, sub-acute assessment, and monitoring

There were only six studies reporting on pitch-side management, ongoing assessment, and monitoring. Youth athletes whose symptoms arose after 15 minutes were more likely to have remained on the field of play, as opposed to those with a more rapid onset of symptoms. Those with delayed symptom presentation were slower to RTP following their concussion diagnosis (31-32,57). Immediate symptom recognition may not only be difficult for responsible pitch-side adults, but also for the youth athletes themselves. Symptoms of concussion, such as disorientation and confusion, may contribute to reduced injury reporting abilities. Concussion injuries remain Challenging to diagnose due to the heterogeneity of clinical presentations and the lack of definite diagnostic and objective tests. Concussion reporting from high school athletes to their respective coaches or responsible adults is poor, and those transitioning into college/university level sport, who have sustained previous concussions, are more likely to remain on the pitch following a future potential concussive injury (65). Athlete reporting behaviour is known to be a concern, with often a myriad of reasons for poor injury reporting. Coaches may also be limited in their concussion injury recognition and early pathway management, thus limiting effective identification, removal from play, and necessary onward medical management or referral (23,66). Furthermore, medical primary and secondary care clinicians' knowledge of concussion management serves as a barrier to best practice concussion care, often with a lack of specific concussion education as a precipitating factor (67,68). More

basic medical management messaging, such as ”if in doubt, sit them out” and “recognize and remove”, has been targeted to support less experienced and knowledgeable stakeholders (24). However, what comes after this in terms of management is somewhat more ambiguous and will certainly rely upon more expert guidance and resources.

The use of the SCAT5 is reported in a number of studies in this review, and further research will likely include the use of the new SCAT6 (69). Further reliability and repeatability studies are warranted for this new assessment tool to ensure the most accurate clinical assessments are utilised, pitch and courtside. Clinical testing protocols such as the Cogstate Brief Battery (CBB) or King Devick (K-D) test have previously been developed for concussion injury assessment and monitoring, and played a role in clinical decision making. In this review, one study found that these testing approaches were unable to consistently provide accurate results for concussion assessment and monitoring, and the authors suggest that these assessment tools should not be used in isolation (64). Likewise, with the SCAT6, clinical judgment and expert assessment are required, and to date, no objective test offers a definitive diagnosis of concussion (27). A further study in the assessment and monitoring theme highlighted that delayed presentation to the clinic (>7 days) and greater symptom severity following a concussion resulted in protracted clinical recovery (55). It would be sensible to suggest that rapid, expert assessment, clinical impression, and treatment are commenced as quickly as possible after a suspected injury for the best health outcomes. Research into blood-based biomarkers for concussion diagnostics has been reported in the 2022 concussion consensus statement (27,70). With the evolution of healthcare technology, identifying, monitoring, and treating concussive injuries may become easier; however, we must ensure that scientific rigor backs any chosen management strategies before widespread implementation of such technologies occurs (71).

Novel treatment interventions

Novel treatment choices remain understudied; however, two studies in the review reported positive results for the use of a head/ neck cooling device for post-concussion symptoms. One study reported the effect of face cooling therapy on the parasympathetic system response. The head and neck cooling studies reported an earlier reduction in post-concussion symptoms within the earlier stages of treatment (36,50). The overall recovery trajectory for both the intervention arm and control group was similar, however, with similar symptom resolution by 1 month after commencing treatment. These results are preliminary and should be interpreted as such. Face cooling was found to have no significant effects on parasympathetic parameters in concussed versus control group athletes, yet there were notable blunted responses in HR and HRV for those with a previous concussion history. This adds to the scientific base on altered physiological parameters following concussive injuries (72). Whilst these results are promising for new concussion recovery or monitoring technology development, before widespread manufacture and use of such face or head-cooling technologies, we would expect the intervention to follow the Ideal, Development, Exploration, Assessment, and Long-term study (IDEAL) framework before conclusions are drawn (73).

The remaining study reported on the effects of hyperbaric oxygen therapy (HB02) for the treatment following concussion; however, the authors did not explore treatment efficacy (44). The results were suggestive of some clinical benefit; however, they were inconclusive due to the study design limitations. No adverse effects were reported, and future research may be warranted to explore HB02 as an alternative or complementary post-concussion treatment option.

Individual accounts of concussion management in specific populations

This review found a variety of case reports charting interventions and recovery, encompassing complex medical cases and situations where young athletes’ mental health was impacted significantly. In the cases reporting on poor mental health, or persisting symptoms, inter-disciplinary team referral and collaboration were critical in the success of ongoing management, especially the involvement of psychological interventions (45,47,53,62). For those youth athletes with a mental health complaint in conjunction with or resulting from their post-concussive symptoms, the outcomes were generally poor (47,62). Whilst the majority of studies reported on in this review yielded clinical success, it was apparent that athletes requiring specialist psychological support experienced a more complex and incomplete recovery journey. Two studies in this section of the review discussed normative return to play approaches for youth athletes, adding little to the knowledge base for management approaches but potentially offering clinicians a reference guide for management (41,60). Three studies reported on specialist medical cases including athletes requiring: steroid medication for persistent post-concussive headaches (48); ophthalmology review and input for a suspected cranial nerve palsy (46), and medical input plus academic protections for an athlete with poor blood glucose control and concussion injury (42). Two studies reported on chiropractic interventions and may again provide a guide for chiropractors in practice (43,63). Such case reports in this theme provide narrative accounts of the management of specific individuals and offer little to the research base regarding generalisable management approaches.

Strengths and Limitations

The strengths of this scoping review are the comprehensive nature of the review process, where a robust cross-database search was completed (eight databases) and manual screening of included manuscripts. There are various limitations to this work. Despite this and using pragmatic search terms, it is possible that relevant articles were missed. We suggest that any further such primary data as to the graded return-to-play and outcomes for concussed youth athletes be potentially recorded and not readily available to the research community and clinicians. A significant proportion of studies found in this review were case reports, and as such, we accept the risk of participant selection and reporting bias. Furthermore, these cases may not be truly representative of the wider management of concussion youth sport, as published case reports

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

usually pertain to particularly interesting or challenging presentations, which was the case in this review. Three case reports included in the review discussed novel treatment choices, including hyperbaric O_2 therapy and head/ neck cooling devices. We accept that these devices may not be readily available in certain settings, due to financial or logistical restraints, and therefore would not be generalisable treatment choices. The evidence for the efficacy of these devices has also not yet been demonstrated. There was also a lack of studies focusing on female-only youth athletes, so sports that predominantly females participate in have been under-represented. No studies were found that reported on management approaches for para-sporting youth athletes. We excluded non-English language studies, which also represents a limitation of this study. Notably, the majority of studies found were from the North American continent, where health care management may differ, and thus, selected management approaches may not be widely generalisable.

Implications for clinicians and researchers

Our scoping review provides a foundation of all primary data sources reporting on the outcome of specified treatment interventions for the concussed youth athlete. We would hope that our work founds the basis for routine and evidence-based management approaches for such athletes, and where we have identified primitive research fields and literature gaps, we would hope for future researchers to target these.

Conclusion

This review, within the constraints of the selected study parameters, synthesises the available primary research data on medical management approaches for concussed youth athletes worldwide. The review highlights the benefits of sub-threshold activity and exercises for concussive symptom resolution and reducing the risk of developing PCS, and may guide physiotherapists and other healthcare practitioners to prescribe threshold exercise to injured youth athletes. Limited data exists for alternative treatment strategies, and we did not find any data sources reporting on primary and secondary medical care management approaches. Nor a particularly expansive literature base evidencing physiotherapy interventions. Despite a focus on graded return to sport (GRTS) protocols in recent guidance publications, no study reported on a monitored GRTS intervention and outcomes (17,18,19,21). Limited data was presented on return-to-learn strategies for the concussed youth athlete, and was often only presented in case reports, where return to school failed. The published primary research data that reports outcomes of medical management approaches for the concussed youth athlete remains limited to exercise and activity recommendations and sub-threshold exercise programmes. Few data exist for alternative management approaches, particularly in settings where resources would not allow for exercise threshold testing and monitoring. Case reports and series saturated this review, and they add little to the generalizable knowledge base for youth concussion management approaches. Research and data published from outside the USA and Canada remain significantly limited. Concussed female youth athletes remain understudied in isolation. Large-scale and multicenter cohort injury data reporting would improve future transparency of youth adolescent athlete concussion monitoring, rehabilitation, and outcome data would provide a vital insight into injury management and help guide future research into concussion management strategies.

Acknowledgements

We express our gratitude to Maria King, specialist librarian, Edinburgh Napier University, for her assistance with the search parameters.

Disclosures

Conflict of interest: The authors report no conflict of interest.

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors contributions: All authors made a substantial contribution to this scoping review. The study was designed by TM, DFH, and ML. TM and SV collected the data. DFH provided support as reviewer three. All authors were involved in analysing and interpreting the data. TM drafted the initial manuscript, and contributions from all authors were provided to guide and produce the final draft.

Data availability statement: The search strategy and study population, concept, and context (PCC) documents are available as supplementary files.

Study registration: Registered on the open science framework – CrossRef

References

1. Russell ER, Mackay DF, Stewart K, et al. Association of field position and career length with risk of neurodegenerative disease in male former professional soccer players. JAMA Neurol. 2021;78(9):1057-1063. CrossRef PubMed
2. Russell ER, Mackay DF, Lyall D, et al. Neurodegenerative disease risk among former international rugby union players. J Neurol Neurosurg Psychiatry. 2022;93(12):1262-1268. CrossRef PubMed
3. Kelly KD, Lissel HL, Rowe BH, et al. Sport and recreation-related head injuries treated in the emergency department. Clin J Sport Med. 2001;11(2):77-81. CrossRef PubMed
4. McCrory P, Feddermann-Demont N, Dvořák J, et al. What is the definition of sports-related concussion: a systematic review. Br J Sports Med. 2017;51(11):877-887. CrossRef PubMed
5. Giza CC, Kutcher JS. An introduction to sports concussions. Continuum (MinneapMinn). 2014;20 (6 Sports Neurology): 1545-1551. CrossRef PubMed
6. Brody DL. Concussion care manual. Oxford University Press; 2019:177. CrossRef
7. Bey T, Ostick B. Second impact syndrome. West J EmergMed. 2009;10(1):6-10. PubMed
8. Cassidy JD, Carroll LJ, Peloso PM, et al. WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. Incidence, risk factors and prevention of mild traumatic brain injury: results of the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. J RehabilMed. 2004;36(43)(suppl):28-60. CrossRef PubMed
9. Veliz P, McCabe SE, Eckner JT, et al. Prevalence of concussion among US adolescents and correlated factors. JAMA. 2017;318(12):1180-1182. CrossRef PubMed

10. Meehan WP III, Mannix R. Pediatric concussions in United States emergency departments in the years 2002 to 2006. J Pediatr. 2010;157(6):889-893. CrossRef PubMed
11. Billette JM, Janz T. Injuries in Canada: Insights from the canadian community healthy survey. Statistics Canada. 2021. Online (Accessed October 2024).
12. McKeever CK, Schatz P. Current issues in the identification, assessment, and management of concussions in sports-related injuries. Appl Neuropsychol. 2003;10(1):4-11. CrossRef PubMed
13. Giza CC, Hovda DA. The neurometabolic cascade of concussion. J Athl Train. 2001;36(3):228-235. PubMed
14. Shrey DW, Griesbach GS, Giza CC. The pathophysiology of concussions in youth. Phys Med Rehabil Clin N Am. 2011;22(4): 577-602, vii. CrossRef PubMed
15. McGroarty NK, Brown SM, Mulcahey MK. Sport-related concussion in female athletes: a systematic review. Orthop J Sports Med. 2020;8(7):2325967120932306. CrossRef PubMed
16. Ono KE, Burns TG, Bearden DJ, et al. Sex-based differences as a predictor of recovery trajectories in young athletes after a sports-related concussion. Am J Sports Med. 2016;44(3):748-752. CrossRef PubMed
17. UK Government. UK Concussion Guidelines for non-elite (grassroots) sport. Landmark concussion guidance for grassroots sport published - GOV.UK.2023. Online
18. AUS GOV – concussionguidance. Concussion and Brain Health Position Statement 2024. Online (Accessed October 2024).
19. Accident Compensation Corporation. SPORT CONCUSSION IN NEW ZEALAND: NATIONAL GUIDELINES. 2024. Online (Accessed October 2024).
20. McArdle D, DeMartini A, Kim S, et al. 'It isn't my area.' Coaches' awareness of concussion protocols in Scottish youth football. Scott Aff. 2021;30(1):53-73. CrossRef
21. Sport Scotland. If in Doubt, Sit Them Out. Online (Accessed October 2024)
22. Murray IR, Murray AD, Robson J. Sports concussion: time for a culture change. Clin J Sport Med. 2015;25(2):75-77. CrossRef PubMed
23. Parry K. World Cup concussion rules may be putting players' lives at risk. The Conversation. 2022. Online (Accessed October 2024).
24. Patricios JS, Schneider KJ, Dvorak J, et al. Consensus statement on concussion in sport: the 6th International Conference on Concussion in Sport-Amsterdam, October 2022. Br J Sports Med. 2023;57(11):695-711. CrossRef PubMed
25. Piedade SR, Hutchinson MR, Ferreira DM, et al. The management of concussion in sport is not standardized. A systematic review. J Safety Res. 2021;76:262-268. CrossRef PubMed
26. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med. 2018;169(7):467-473. CrossRef PubMed
27. World Health Organization. Adolescent health. World Health Organization. Online (Accessed October 2024).
28. Melnyk BM, Fineout-Overholt E. Evidence-based practice in nursing & healthcare: a guide to best practice (Fifth edition). Wolters Kluwer. 2023 (Adapted). Online (Accessed October 2024).
29. Neely LM, Smulligan KL, Wingerson MJ, et al. The association between sleep and physical activity with persisting postconcussion symptoms among adolescent athletes. PM R. 2023;15(9):1122-1129. CrossRef PubMed
30. Kelleher E, Taylor-Linzey E, Ferrigno L, et al. A community return-to-play mTBI clinic: results of a pilot program and survey of high school athletes. J PediatrSurg. 2014;49(2):341-344. CrossRef PubMed
31. Zynda AJ, Worrall HM, Sabatino MJ, et al. Continued play following adolescent sport-related concussion: prospective data from the North Texas Concussion Registry (ConTex). Appl Neuropsychol Child. 2022;11(4):740-751. CrossRef PubMed
32. Elbin RJ, Sufrinko A, Schatz P, et al. Removal from play after concussion and recovery time. Pediatrics. 2016;138(3):e20160910. CrossRef PubMed
33. Seehusen CN, Wilson JC, Walker GA, et al. More physical activity after concussion is associated with faster return to play among adolescents. Int J Environ Res Public Health. 2021;18(14):7373. CrossRef PubMed
34. Howell DR, Seehusen CN, Walker GA, et al. Neuromuscular training after concussion to improve motor and psychosocial outcomes: a feasibility trial. PhysTher Sport. 2021;52:132-139. CrossRef PubMed
35. Worts PR, Mason JR, Burkhart SO, et al. The acute, systemic effects of aerobic exercise in recently concussed adolescent student-athletes: preliminary findings. Eur J ApplPhysiol. 2022;122(6):1441-1457. CrossRef PubMed
36. Congeni J, Murray T, Kline P, et al. Preliminary safety and efficacy of head and neck cooling therapy after concussion in adolescent athletes: a randomized pilot trial. Clin J Sport Med. 2022;32(4):341-347. CrossRef PubMed
37. Leddy JJ, Haider MN, Ellis MJ, et al. Early subthreshold aerobic exercise for sport-related concussion: a randomized clinical trial. JAMA Pediatr. 2019;173(4):319-325. CrossRef PubMed
38. Howell DR, Seehusen CN, Carry PM, et al. An 8-week neuromuscular training program after concussion reduces 1-year subsequent injury risk: a randomized clinical trial. Am J Sports Med. 2022;50(4):1120-1129. CrossRef PubMed
39. Leddy JJ, Hinds AL, Miecznikowski J, et al. Safety and prognostic utility of provocative exercise testing in acutely concussed adolescents: a randomized trial. Clin J Sport Med. 2018;28(1):13-20. CrossRef PubMed
40. Leddy JJ, Master CL, Mannix R, et al. Early-targeted heart rate aerobic exercise versus placebo stretching for sport-related concussion in adolescents: a randomised controlled trial. Lancet Child Adolesc Health. 2021;5(11):792-799. CrossRef PubMed
41. Gunter KB, Shields CJ, Ott SD, et al. Rehabilitation of an adolescent equestrian athlete with a history of multiple concussions: a case report describing an adapted Return-to-Sport protocol. J Orthop Sports Phys Ther. 2018;48(12):934-942. CrossRef PubMed
42. Lunardini DJ, Berg CL, Perriello VA Jr, et al. Persistent postconcussion symptoms in a young diabetic athlete: a case review. Athl Train Sports Health Care. 2009;1(1):37-40. CrossRef
43. Olson H, Tunning M, Lindholm S. Chiropractic management of three young athletes with concussion. Chiropractic Journal of Australia (Online). 2018;46(1):29-47. CrossRef
44. Roby PR, Lynall RC, Cools MJ, et al. Hyperbaric oxygen therapy to treat acute sport-related traumatic brain injuries: a case series. Int J AthlTher Train. 2021;26(3):140-144. CrossRef
45. Sherry NS, Feder A, Pan R, et al. Transitioning concussion care to mental health care: a case study of an elite athlete. Case Stud Sport Exerc Physiol. 2021;5(1):135-144. CrossRef
46. Stiller-Ostrowski JL. Fourth cranial nerve palsy in a collegiate lacrosse player: a case report. J Athl Train. 2010;45(4):407-410. CrossRef PubMed
47. Tucker Price S, DeCastro AO, et al. Post-concussion syndrome and pre-existing mood disorder: A case report. Int J PsychiatryMed. 2019;54(3):163-171. CrossRef PubMed
48. Bramley H, Melinosky C, Silvis M, et al. Pediatric post-traumatic headache: two cases using steroids as abortive therapy. Pediatr Emerg Care. 2012;28(10):1081-1084. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

49. Chrisman SPD, Whitlock KB, Mendoza JA, et al. Pilot randomized controlled trial of an exercise program requiring minimal in-person visits for youth with persistent sport-related concussion. Front Neurol. 2019;10:623. CrossRef PubMed
50. Smith MA, McNinch NL, Chaney D, et al. Reduced Concussion Symptom Burden in Early Adolescent Athletes Using a Head–Neck Cooling Device. Clin J Sport Med. 2022;10-1097. CrossRef PubMed
51. Haider MN, Chizuk HM, Johnson BD, et al. Parasympathetic responses to face cooling in adolescents with sport-related concussion and after clinical recovery. Neurotrauma Rep. 2025;6(1):93-105. CrossRef PubMed
52. Miller SM, Zynda AJ, Sabatino MJ, et al. A pilot randomized controlled trial of docosahexaenoic acid for the treatment of sport-related concussion in adolescents. ClinPediatr (Phila). 2022;61(11):785-794. CrossRef PubMed
53. McNally KA, Shiplett KM, Davidson AM, et al. Case study: neuropsychology-informed brief intervention for persisting symptoms after pediatric concussion in a multi-disciplinary clinic setting. J PediatrNeuropsychol. 2024;10(4):1-8. CrossRef
54. Lishchynsky JT, Rutschmann TD, Toomey CM, et al. The association between moderate and vigorous physical activity and time to medical clearance to return to play following sport-related concussion in youth ice hockey players. Front Neurol. 2019;10:588. CrossRef PubMed
55. Emery CA, Warriyar KV, Black AM, et al. Factors associa zted with clinical recovery after concussion in youth ice hockey players. Orthop J Sports Med. 2021;9(5):23259671211013370. CrossRef PubMed
56. Gagnon I, Grilli L, Friedman D, et al. A pilot study of active rehabilitation for adolescents who are slow to recover from sport-related concussion. Scand J Med Sci Sports. 2016;26(3):299-306. CrossRef PubMed
57. Olson A, Ellis MJ, Selci E, et al. Delayed symptom onset following pediatric sport-related concussion. Front Neurol. 2020;11:220. CrossRef PubMed
58. Chan C, Iverson GL, Purtzki J, et al. Safety of active rehabilitation for persistent symptoms after pediatric sport-related concussion: a randomised controlled trial. Arch PhysMedRehabil. 2018;99(2):242-249. CrossRef PubMed
59. Micay R, Richards D, Hutchison MG. Feasibility of a post-acute structured aerobic exercise intervention following sport concussion in symptomatic adolescents: a randomized controlled study. BMJ Open Sport ExercMed. 2018;4(1):e000404. CrossRef PubMed
60. Marsh AM, Fraser D, Marsh JP. Management of concussion in the pediatric patient. J Pediatr Health Care. 2013;27(6):499-504. CrossRef PubMed
61. Morissette MP, Cordingley DM, Ellis MJ, et al. Evaluation of early submaximal exercise tolerance in adolescents with symptomatic sport-related concussion. Med Sci Sports Exerc. 2020;52(4):820-826. CrossRef PubMed
62. Waldmeier-Wilhelm S, Studer M, Weisstanner C, et al. Mild traumatic brain injury: striking postconcussion symptoms due to inadequate management. Pediatr Emerg Care. 2019;35(6):e99-e103. CrossRef PubMed
63. Hunt BB, Holt K, Cade A, et al. Improvement in concussion symptoms of headache, poor concentration and photophobia in a 13-year-old male receiving chiropractic care: a case report. J Clin Chirop Pediat. 2018;1389. Online (Accessed October 2024)
64. Cosgrave C, Fuller C, Kung S, et al. A comparison of clinical assessment with common diagnostic tools for monitoring concussion recovery in adolescent rugby union players. PhysTher Sport. 2023;61:165-171. CrossRef PubMed
65. Register-Mihalik JK, Guskiewicz KM, McLeod TCV, et al. Knowledge, attitude, and concussion-reporting behaviors among high school athletes: a preliminary study. J Athl Train. 2013;48(5):645-653. CrossRef PubMed
66. Feiss RS, Lutz M, Moody JR, et al. A systematic z review of coach and parent knowledge of concussion. J Concussion. 2020;4. CrossRef
67. Zonfrillo MR, Master CL, Grady MF, et al. Pediatric providers' self-reported knowledge, practices, and attitudes about concussion. Pediatrics. 2012;130(6):1120-1125. CrossRef PubMed
68. Jervis CG, Adams SA, Fawkner S, et al. Concussion education in medical students studying in Scotland: an assessment of knowledge and future needs. Brain Inj. 2022;36(9):1196-1203. CrossRef PubMed
69. Echemendia RJ, Brett BL, Broglio S, et al. Introducing the Sport Concussion Assessment Tool 6 (SCAT6). Br J Sports Med. 2023;57(11):619-621. CrossRef PubMed
70. Tabor JB, Brett BL, Nelson L, et al. Role of biomarkers and emerging technologies in defining and assessing neurobiological recovery after sport-related concussion: a systematic review. Br J Sports Med. 2023;57(12):789-797. CrossRef PubMed
71. McKeever T, Hamilton DF. Is it scientifically valid, sensible, or safe to use biomarkers to diagnose concussion? BJSM Blog. 2023. Online (Accessed October 2024).
72. Kamins J, Bigler E, Covassin T, et al. What is the physiological time to recovery after a concussion? A systematic review. Br J Sports Med. 2017;51(12):935-940. CrossRef PubMed
73. Beard D, Hamilton D, Davies L, et al. Evidence-based evaluation of practice and innovation in physical therapy using the IDEAL-physio framework. PhysTher. 2018;98(2):108-121. CrossRef PubMed

Arch Physioter 2025; 15: 239-249
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3331
ORIGINAL RESEARCH ARTICLE

Biochemical and clinical effects of McKenzie therapy versus muscle endurance exercises in chronic low-back pain

Mistura Iyabo Olaoye[1,2], Raphael Okonji[3], Adekola Ademoyegun[1,2], Tadesse Gebrye[4], Gillian Yeowell[4], Francis Fatoye[4], Chidozie Mbada[4]

[1]Department of Physiotherapy, Osun State University Teaching Hospital, Osogbo - Nigeria
[2]Department of Medical Rehabilitation, Obafemi Awolowo University, Ile-Ife - Nigeria
[3]Department of Biochemistry, Obafemi Awolowo University, Ile-Ife - Nigeria
[4]Department of Health Professions, Manchester Metropolitan University - United Kingdom

ABSTRACT

Background and objective: Apart from mechanical dysfunction, low back pain (LBP) is also associated with underlying inflammatory and muscle-related biochemical changes. An increase in certain biomarkers, such as IL-10, a key anti-inflammatory cytokine, provides a positive objective indicator of underlying physiological responses to interventions in LBP beyond subjective clinical measures. This study assessed the effects of McKenzie Extension Protocol (MEP), Static Back Extension Endurance (SBEE), and Dynamic Back Extension Endurance (DBEE) on selected clinical outcomes and biomarkers of muscle status [creatine kinase (CK)] and inflammation (IL-4 and IL-10) in LBP.

Methods: A randomized controlled trial involving 76 patients with chronic LBP who were randomly assigned to MEP, SBEE, or DBEE groups was conducted. MEP involved a specific sequence of lumbosacral repeated movements in extension. SBEE involved five different back extensor muscle endurance protocols of increasing difficulty level. DBEE was a dynamic replica of the SBEE. Pain, CK, IL-4, and IL-10 were the primary outcomes. Functional disability and health-related quality of life were the secondary outcomes. Assessments were conducted at baseline, 3rd, and 6th week of the study.

Results: MEP and SBEE caused significant effects in all clinical and biochemical variables ($p < 0.05$) except IL-4 and IL-10 ($p > 0.05$). DBEE yielded no significant effects on IL-4 and IL-10 ($p > 0.05$). MEP had a significantly higher effect on pain ($p < 0.05$). SBEE had a greater impact on IL-4 ($p < 0.05$) and IL-10 ($p < 0.05$) at week 3. SBEE led to a higher impact on IL-4 ($p < 0.05$) and IL-10 ($p < 0.05$) at week 6. All interventions had comparable effects on other clinical parameters at week 6 ($p > 0.05$).

Conclusion: MEP reduced pain more, while SBEE led to higher changes in IL-4 and IL-10 inflammatory biomarker levels. Serum CK levels rose in all groups without indicating muscle damage. The results suggest that these exercises show potential benefits in modulating inflammation and enhancing muscle status, potentially supporting tissue repair and reducing chronic LBP, and therefore should be incorporated as part of strategies targeting underlying inflammatory processes in the management of chronic LBP.

Keywords: Biomarkers, Creatine kinase, Endurance exercise, Interleukin, Low-back pain, McKenzie protocol

What is already known about this topic:

- *There is considerable stress on muscles during endurance exercise, and the biochemical changes during exercise are associated with clinical outcomes in musculoskeletal pain. However, data on the biochemical changes to endurance exercise in LBP is limited.*

What does the study add:

- *Endurance exercises in the form of McKenzie extension protocol (MEP) and static-and-dynamic back extension had beneficial clinical effects, did not cause muscle damage, and positively modulated inflammatory markers in chronic LBP.*

Received: October 5, 2024
Accepted: August 22, 2025
Published online: September 17, 2025

This article includes supplementary material.

Clinical Trial Protocol number: Pan African Clinical Trial Registry (PACTR202208757153267).

Corresponding author:
Adekola Ademoyegun
email: aademoyegun@gmail.com

Introduction

Low back pain (LBP) is a significant healthcare concern globally, causing more disability than any other medical condition (1,2). The management of LBP typically involves pharmacological and/or non-pharmacological means (3). The pharmacological approach is usually more suitable for acute LBP (4), while physiotherapy, as a non-pharmacological or conservative approach, is reported to be effective in chronic or long-term LBP (5). However, there are discrepancies in reports on the strength of evidence for various physiotherapy

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

interventions for LBP (6,7). Thus, there are numerous exercise programmes available for patients with chronic LBP (8-12), but there is still no consensus on the most effective and beneficial exercise programme (13).

One of the most common structured specific exercises for chronic LBP is the McKenzie protocol (14). Currently, clinicians and researchers continue to rely on reported clinical effects rather than side effects when implementing exercise-based treatments for LBP. However, gaining a deeper understanding of biomarkers represents a positive step toward comprehending the effects and potential side effects of these interventions (15). Biomarkers are becoming essential in physiotherapy to assess the safety and effectiveness of treatments for LBP. They offer objective measures of biological processes, allowing for precise monitoring of how patients respond to treatment. Additionally, biomarkers help predict patient outcomes, identify potential side effects, and guide the development of therapeutic interventions (15-20).

Endurance exercise imposes a significant toll on muscles, making them more susceptible to injuries. However, there is a lack of research on using biomarkers to measure treatment responses to endurance exercise in the context of LBP. Biomarkers can serve as indicators of disease progression, pathogenesis, and treatment response (21), providing insights into the biochemical changes and clinical outcomes, including pain intensity, fatigue, and quality of life following exercise (22,23). Pinto et al. (24) in a systematic review found a correlation between elevated levels of pro-inflammatory biomarkers (CRP, IL-6, TNF-α) and non-specific LBP while noting decreased levels of the anti-inflammatory biomarker IL-10. The review indicates that effective treatments reduce levels of pro-inflammatory biomarkers, while potentially increasing levels of the anti-inflammatory biomarker IL-10. Another review affirms a moderate correlation between the severity of LBP and elevated levels of CRP and IL-6 (25). However, the benefits and safety of McKenzie therapy and back extension endurance exercises for treating chronic LBP have not been studied at the molecular level, particularly regarding muscle status and changes in cytokine levels. Thus, this study aimed to assess the effects of MEP, Static Back Extension Endurance (SBEE), and Dynamic Back Extension Endurance (DBEE) on selected clinical outcomes (pain, functional disability, and general health status); and biomarkers of muscle status [creatine kinase (CK)] and inflammation [anti-inflammatory cytokines i.e. interleukin-4 (IL-4) and interleukin-10 (IL-10)] in patients with chronic LBP.

Materials and Methods

The Consolidated Standards of Reporting Trials (CONSORT) for reporting of RCTs was followed in this study. This study was retrospectively registered with the Pan African Clinical Trial Registry (PACTR202208757153267). The health research ethics committee of the Institute of Public Health, Obafemi Awolowo University, Ile-Ife, Nigeria, approved the study (HREC NO: IPH/OAU/12/1570). Patients with chronic LBP from Osun State University Teaching Hospital, and State Specialist Hospital, Osogbo, Nigeria were recruited for this study. Eligible participants were patients aged 18 and older who had been experiencing back pain for 12 weeks or more, with or without radiculopathy, and did not have any obvious physical deformities. Patients with LBP resulting from serious spinal conditions such as fractures, tumors, and inflammatory diseases were excluded from the study. Additionally, pregnant individuals, those who had undergone back surgery, those with a reported history of cardiovascular disease that contraindicated exercise or with high blood pressure (>140/90 mmHg), those with previous experience of the McKenzie therapy, and those who had a directional preference for flexion or no directional preference based on the McKenzie assessment were also excluded. Sample size estimation for this study was based on the sample size equation by Chan (26) – $M = c \times \pi1 (1 - \pi1) + \pi2 (1 - \pi2)/(\pi1 - \pi2)^2$, where c = 7.9 for 80% power, and π1 and π2 areproportion estimates derived from Chan's formula indicating the effect size (π1 = 0.25 and π2 = 0.65). Therefore, n = 7.9 * 0.25 (1 – 0.25) + 0.65 (1 – 0.65)/(0.25 – 0.65) = 20.49. Typically, adding 10-15% to the calculated sample size helps account for potential dropouts or non-responses during a study. This adjustment ensures that the study maintains sufficient power and statistical validity, even if some participants drop out without having underpowered results and conclusions. Using a 15% adjustment, 91 patients were assessed for eligibility, nine were excluded, and 28 patients were randomized into MEP, while 27 patients were randomized into SBEE and DBEE protocols, equalling 82 patients. However, only 76 completed the study. The number of participants lost to follow-up or discontinued intervention was three, two, and one in the MEP, SBEE, and DBEE protocols, respectively. Their data were removed from the final analysis. The baseline socio-demographic, clinical, and biochemical variables between the participants lost to follow-up and those who completed the study were comparable. The CONSORT flow diagram for this study is presented in Figure 1.

Instrument

The following instruments were used in this study:

(i) The quadruple visual analogue scale (QVAS): A 10-point numeric rating scale was used to measure pain intensity in patients with LBP. A Yoruba translation of the same questionnaire was administered to the Yoruba-speaking group among the participants (27). The Yoruba version of QVAS has adequate concurrent validity (r = 0.896) and good reliability (ICC = 0.622) in patients with LBP (27). In this study, pain intensity was considered a primary outcome while disability (activity limitation and participation restriction) and health-related quality of life were the secondary outcomes.

(ii) Oswestry Disability Index (ODI) for chronic LBP patients was used to measure work disabilities and other functional limitations associated with LBP. A Yoruba translation of the same questionnaire was administered to the Yoruba-speaking group among the participants (28). The psychometric properties of the ODI Yoruba version have been tested and found adequate (28). The Yoruba version of ODI showed excellent test-retest reliability (ICC = 0.89) and internal consistency (Cronbach's alpha = 0.81) (28).

(iii) Roland Morris's LBP disability questionnaire was used to assess functional disabilities. A Yoruba translation of the same questionnaire was administered to the Yoruba-speaking group among the participants (29). The Yoruba

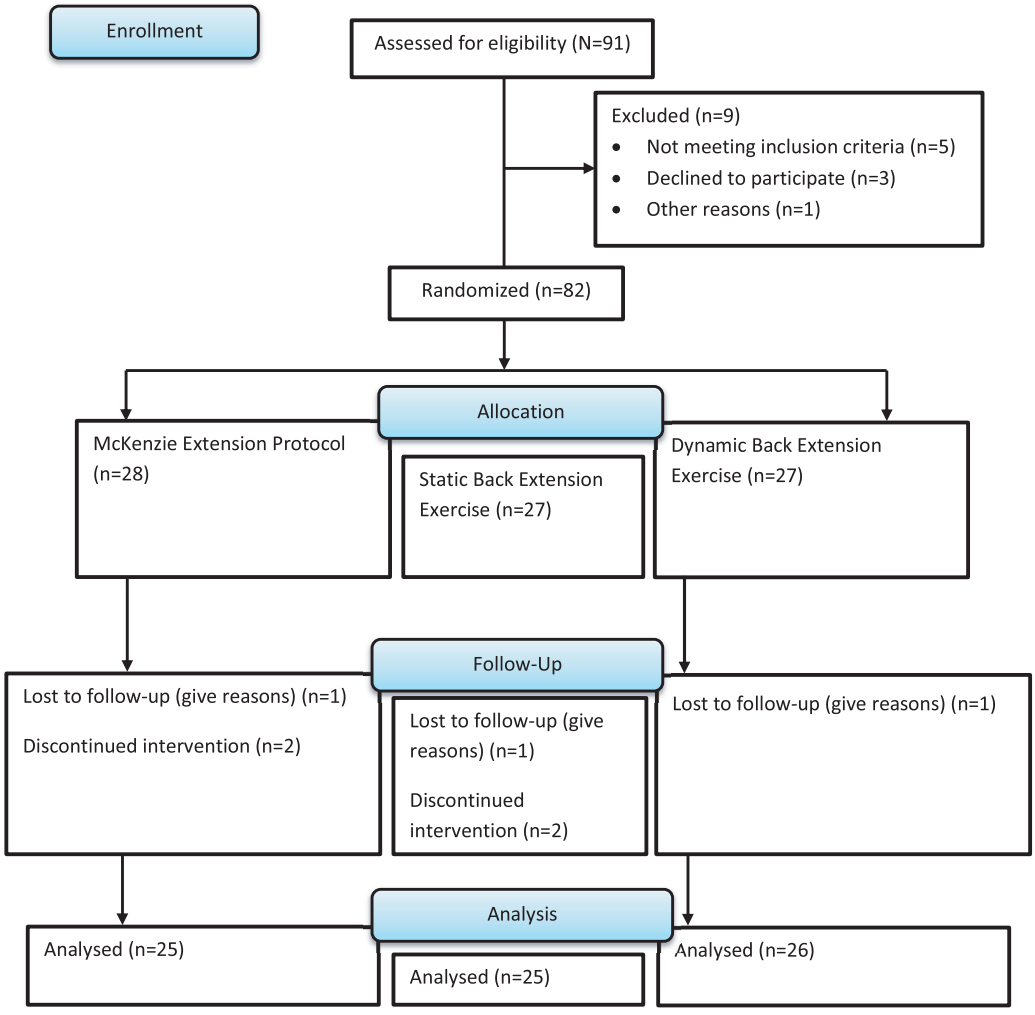


FIGURE 1 - CONSORT flow diagram for the study.

version has excellent reliability (ICC = 0.99) and validity (r = 0.82) among patients with LBP (29).

(iv) The SF-12 General Health Status Questionnaire was used to assess the general health status of the participants. A Yoruba translation of the same questionnaire was administered to the Yoruba-speaking group among the participants (30). Allscale and domain scores of the Yoruba version of SF-12 have acceptable concurrent validity (r = 0.879 – 0.938) and reliability (ICC = 0.775 – 0.949) (30).

In the present study, pain intensity was considered the primary clinical outcome, while functional disability (activity limitation and participation restriction) and general health status were the secondary clinical outcomes.

Procedure

The research team provided an explanation of the study to all participants who gave consent. To ensure inclusivity, language experts translated the informed consent form into Yoruba. Participants were recruited consecutively and then randomly assigned to one of three treatment groups. To maintain objectivity and minimize bias, a research assistant kept track of the number of participants invited to take part, the number who declined, and the number of screened patients who were ineligible, including the reasons for their ineligibility or refusal to participate. Volunteer participants who met the eligibility criteria were randomly placed in the McKenzie exercise protocol (MEP), SBEE, or DBEE exercise (DBEE) group by the same assistant who was not involved in assessing or treating the participants. Another assistant who was blinded to participant allocation was involved in the outcomes assessment. Random permuted blocks were used to ensure equal group sizes, following the method proposed by Pocock (31). A block size of six was chosen (i.e., MMDDSS, MMDSSD, MSDMSD, and other possible restricted permutations). The block permutations were computer-generated using a factorial equation formula: 6!6!, which is 1 × 2 × 3 × 4 × 5 × 6 = 7201 × 2 × 3 × 4 × 5 × 6 = 720. Some of the printout of 720 restricted computer-generated block permutations was sequentially numbered, cut, and placed in sealed envelopes. A block-permuted sequence was randomly drawn from the envelope, and accordingly, consecutive patients were assigned to either the M (MEP), S (SBEE), or D (DBEE) groups. The process of drawing block permuted sequences

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

and randomization was repeated as new participants were recruited. The physiotherapist, clinical outcomes assessor, and biochemical analysts were blinded to group allocation.

Pre-treatment Screening

All participants underwent an assessment to determine their eligibility for the study. The assessment used the McKenzie Institute's Lumbar Spine Assessment Algorithm (MILSAA), which is a defined method for classifying spinal-related disorders based on historical pain behaviour and the pain response to specific movements and activities. Participants were evaluated for their directional preference for certain movements, including those done while standing and lying down, in both front-to-back and side-to-side directions. This involved repeating each movement 5-10 times while observing their symptomatic and mechanical responses. After the movement testing, participants were asked standardized questions to assess the movements' effects on their pain (32,33). Those who did not respond well to certain movements or showed no response to repeated movements were excluded from the study. Only those who responded well to specific movements were considered eligible. Information regarding age, gender, education, occupation, marital status, onset and history of back pain, as well as previous interventions, was recorded for each participant. Blood samples were collected at baseline, 3, and 6 weeks to analyze biomarker patterns. Variations in biochemical profiles were compared with the exercise intensity throughout the study. Participants attended twelve treatment sessions, with two sessions per week. The clinic-based exercise interventions were supervised by a physiotherapist. Participants' adherence to the exercises was monitored through regular check-ins by the principal investigator (MIO) and logs in the diary, which were reviewed at sessions.

Laboratory Analysis

4 mL of whole blood was collected using BD Vacutainer™ tubes containing EDTA K2, in a vacuum puncture needle BD Eclipse™ about 10 minutes before starting each training session. CK was assayed using spectrophotometric methods with kits from Randox, while anti-inflammatory markers like interleukin-4 and -10 were assayed using Enzyme-Linked Immune Sorbent Assay (ELISA), with commercially prepared kits from DIAPRO Italy. CK, IL-4, and IL-10 were considered the primary biomarkers in this study. An experienced medical laboratory scientist conducted the biochemical analysis at the Medical Laboratory Complex of the Osun State University Teaching Hospital, Osogbo, Nigeria.

The McKenzie Protocol

The MEP is a classification-treatment-based method. Directional preference for extension was first assessed among the participants. This involved a course of specific lumbosacral repeated movements in extension that caused the symptoms to centralize, decrease, or abolish. The determination of the direction preference for extension was followed by the main MEP activities, including" Extension Lying Prone", "Extension In Prone", and" Extension In Standing". The details of the MEP are reported in an earlier publication (34).

Static Back Extensors Endurance Exercises

The static back extensors endurance exercises (SBEE) consisted of five different exercises that increased difficulty levels. During the exercises, participants adjusted the positions of their upper and lower limbs. The participants started the exercise training programme with the first exercise position and then progressed to the next exercises at their own pace once they could hold a given position for 10 seconds. After reaching the fifth progression, they continued with the fifth progression until the end of the exercise programme. The exercise period lasted from 30 to 45 minutes. The details of the SBEE are reported in an earlier publication (34).

Dynamic Back Extensors Endurance Exercise

The dynamic back extensors endurance exercise (DBEE) consisted of five different exercises. It closely resembled the static back extensors endurance exercise in terms of exercise positions, progressions, and duration. However, instead of maintaining a static posture in a prone lying position and holding the positions of the upper and lower limbs suspended in the air for 10 seconds during all five exercise progressions, the participant was asked to move the trunk and the suspended limbs 10 times. Each exercise was repeated 9 times. After 10 repetitions, the participants were instructed to rest for 30 seconds to 1 minute. The number of trunk movements in the exercise position gradually increased to 20 times to provide a more intense training stimulus. The details of the DBEE are reported in an earlier publication (34).

After the initial assessment, when all participants joined the study, two more assessments were conducted in the third and sixth weeks. During these reassessments, participants filled out questionnaires to measure the outcomes. Additionally, all participants were given a set of instructions on back care education, which included a 9-item guide on how to stand, sit, lift, and perform other daily activities. They were instructed to perform these exercises at home twice daily. The interventions were illustrated in Supplementary Material 1.

Data Analysis

Descriptive statistics of frequency, percentage, mean, and standard deviation were used to summarize data. Analysis of Variance (ANOVA) was used for across-group analysis. Within-group analysis was done using repeated-measured ANOVA (comparing baseline, third, and sixth weeks). Post Hoc Analysis (Least Significant Difference (LSD) was used to determine the trend of differences in the groups. Kruskal–Wallis test was used to compare physical functions across the three groups. Data were analyzed using the Statistical Program for Social Sciences for Windows version 20 (SPSS Inc., Chicago, Illinois, United States). Data was analysed using SPSS version 20.0. The alpha level was set at $p < 0.05$.

Results

The general characteristics and baseline parameters of the participants in each group are presented in Table 1. There was

no significant difference in age across the three groups ($p > 0.05$), but significant differences were found in height, weight, and BMI ($p < 0.05$) across groups. Among clinical variables, ODI, SF, VT, and CK scores were also comparable across groups ($p > 0.05$).

There were significant differences in the pain intensity ($F = 3.255$; $p = 0.044$; $\eta^2 = 0.12$), IL-4 ($F = 3.255$; $p = 0.044$; $\eta^2 = 0.08$), and IL-10 ($F = 5.664$; $p = 0.005$; $\eta^2 = 0.13$) across the groups at the third week of intervention (Table 2). However, there was no significant difference in other clinical parameters across the groups in the third week of intervention ($p > 0.05$). The LSD post hoc analysis showed that the pain intensity was significantly lower in DBEE compared to MEP ($p = 0.002$) and SBEE ($p = 0.011$). However, there was no significant decrease between MEP and SBEE groups ($p = 0.126$). In the third week of intervention, analysis of the biomarkers IL-4 and IL-10 showed significant differences between the treatment groups (MEP, SBEE, DBEE), with MEP showing significantly lower levels of IL-10 compared to SBEE and DBEE ($p < 0.05$ for both) (Table 2).

Table 3 presents comparisons of clinical parameters by treatment groups at the sixth week of intervention. There were significant differences in the serum levels of IL-4 ($F = 9.219$; $p = 0.001$; $\eta^2 = 0.20$) and IL-10 ($F = 7.851$; $p = 0.001$; $\eta^2 = 0.17$) across the groups. Similarly, the results of the LSD post hoc analysis showed that in the sixth week of intervention, analysis of the biomarkers IL-4 and IL-10 showed significant differences between the treatment groups (MEP, SBEE, DBEE), with MEP showing significantly lower levels of IL-10 compared to SBEE and DBEE ($p < 0.05$ for both).

The results of within-group comparisons of MEP, SBEE, and DBEE interventions on clinical parameters across the 3 time points of the study are presented in Tables 4, 5, and 6, respectively. The results showed that there were significant differences in pain severity ($F = 96.017$; $p = 0.001$; $\eta^2 = 0.92$) and CK ($F = 32.638$; $p = 0.001$; $\eta^2 = 0.57$) among patients in the MEP group across baseline, third, and sixth week of intervention (Table 4). Similarly, there were significant differences in pain severity ($F = 88.733$; $p = 0.001$; $\eta^2 = 0.78$) and CK ($F = 27.235$; $p = 0.001$; $\eta^2 = 0.53$) among patients in SBEE group (Table 5) and DBEE group (pain severity: [$F = 91.093$; $p = 0.001$; $\eta^2 = 0.78$]; CK: [$F = 26.790$; $p = 0.001$; $\eta^2 = 0.51$]) across baseline, third, and sixth week of intervention (Table 6).

TABLE 1 - One-way ANOVA comparison of general and baseline clinical characteristics across treatment groups

Variable	MEP (n = 25) x ± SD	SBEEE (n = 25) x ± SD	DBEEE (n = 26) x ± SD	F-ratio	p-value
General Characteristics					
Age (yrs)	50.0 ± 7.5	50.5 ± 7.8	50.0 ± 8.8	0.057	0.945
Height (m)	1.67 ± 0.04	1.64 ± 0.06	1.62 ± 0.06	7.615	0.001*
Weight (Kg)	68.9 ± 7.77	64.8 ± 9.36	58.3 ± 5.74	12.238	0.001*
BMI (Kg/m^2)	24.5 ± 2.32	23.9 ± 3.30	22.4 ± 2.89	3.443	0.037*
Clinical Characteristics					
ODI	24.8 ± 8.15	22.1 ± 7.22	20.0 ± 6.8	2.687	0.075
RMLDQ	15.0 ± 3.34^a	11.4 ± 3.33^b	10.8 ± 2.85^b	13.383	0.001*
QVAS	65.6 ± 13.9^a	54.8 ± 13.8^b	54.8 ± 13.9^b	5.012	0.009*
PF	24.7 ± 20.5^a	35.9 ± 27.7^b	42.9 ± 22.9^b	3.662	0.031*
RLP	90.2 ± 16.0^a	91.6 ± 15.3^b	79.6 ± 17.6^b	3.946	0.024*
RLE	87.4 ± 17.1^a	94.4 ± 13.1^b	80.2 ± 17.7^b	4.662	0.013*
BP	55.4 ± 18.4^a	92.2 ± 7.6^b	87.5 ± 5.7^b	69.170	0.001*
VT	54.4 ± 27.4	59.2 ± 21.2	65.0 ± 21.5	1.242	0.295
GH	38.2 ± 23.4^a	61.0 ± 24.2^b	72.8 ± 22.1^b	14.280	0.001*
SF	87.0 ± 12.8	89.0 ± 11.6	82.3 ± 11.6	1.893	0.158
MH	60.0 ± 13.9^a	37.3 ± 13.5^a	33.3 ± 12.1^b	29.407	0.001*
PHD	52.1 ± 10.5^a	70.1 ± 9.2^b	70.6 ± 7.1^b	33.709	0.001*
MHD	72.2 ± 8.6^a	69.9 ± 7.4^a	65.2 ± 6.9^b	5.177	0.008*
CK	69.2 ± 39.3	91.7 ± 55.4	70.4 ± 47.3	1.772	0.177
IL-4	55.9 ± 29.1^a	82.5 ± 37.6^b	70.8 ± 24.9^a	4.672	0.012*
IL-10	40.8 ± 15.7^a	54.0 ± 27.9^b	37.9 ± 24.9^c	3.346	0.041*

Alpha level was set at $p < 0.05$; X = Mean; SD = Standard deviation; MEG = McKenzie Exercise Group; SBEEEG = Static back extension endurance exercise group; DBEEEG = Dynamic back extension endurance exercise group; ODI = Oswestry Disability Index; RMLDQ = Roland Morris low back pain disability questionnaire; QVAS = Quadruple visual analogue scale; PF =Physical functioning; RLP = Role limitation physical; RLE = Role limitation emotion; BP=Bodily pain; VT = Vitality; GH = General health; SF = Social Functioning; MH = Mental health; MHD = Mental health domain; PHD = Physical health domain; CK = creatinekinase; IL-4 = Interleukin 4; IL-10 = Interleukin 10; *Superscripts (^{a,b,c}).* For a particular variable, mean values with different superscripts are significantly ($p < 0.05$) different; * indicates significant difference.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - Comparisons of Clinical Parameters by Treatment Groups at the Third Week of Intervention

Variable	MEG (n = 25) x ± SD	SBEEEG (n = 25) x ± SD	DBEEEG (n = 26) x ± SD	F-ratio	p-value	η^2
ODI	14.6 ± 6.4	13.6 ± 6.1	12.9 ± 4.8	0.516	0.599	0.01
RMLDQ	5.2 ± 3.4	5.1 ± 2.7	4.7 ± 2.5	0.140	0.870	0.00
QVAS	38.8 ± 8.6[a]	35.2 ± 9.2[a]	31.8 ± 5.5[b]	5.045	0.009*	0.12
PF	50.7 ± 21.8	56.0 ± 19.2	56.3 ± 20.7	0.582	0.561	0.01
RLP	79.0 ± 17.5	72.0 ± 14.3	78.1 ± 17.3	1.343	0.268	0.03
RLE	70.6 ± 13.1	70.6 ± 13.1	67.9 ± 9.8	0.396	0.675	0.01
BP	74.4 ± 15.0	78.4 ± 13.4	74.7 ± 12.7	0.637	0.532	0.01
VT	72.8 ± 19.4	80.8 ± 22.7	85.8 ± 16.1	2.784	0.069	0.07
GH	61.0 ± 24.2	65.4 ± 18.3	66.4 ± 21.8	0.443	0.644	0.01
SF	81.0 ± 10.9	77.0 ± 6.9	76.0 ± 5.1	2.639	0.078	0.06
MH	68.8 ± 9.8	68.0 ± 7.5	69.2 ± 14.2	0.076	0.927	0.00
PHD	66.3 ± 10.7	67.9 ± 9.2	68.9 ± 9.5	0.473	0.625	0.01
MHD	73.3 ± 8.0	74.1 ± 7.8	74.7 ± 4.9	0.253	0.777	0.00
CK	134.2 ±40.7	153.6 ± 29.7	145.2 ± 38.2	1.864	0.162	0.04
IL-4	63.7 ± 29.1[a]	82.9 ± 19.8[b]	70.7 ± 30.2[b]	3.255	0.044*	0.08
IL-10	38.3 ± 14.4[a]	61.3 ± 34.8[b]	42.4 ± 24.3[b]	5.664	0.005*	0.13

Alpha level was set at $p < 0.05$; X = Mean; SD = Standard deviation; MEG = McKenzie Exercise Group; SBEEEG = Static back extension endurance exercise group; DBEEEG = Dynamic back extension endurance exercise group; ODI = Oswestry Disability Index; RMLDQ = Roland Morris low back pain disability questionnaire; QVAS = Quadruple visual analog scale; PF = Physical functioning; RLP = Role limitation physical; RLE = Role limitation emotion; BP = Bodily pain; VT = Vitality; GH = General health; SF = Social Functioning; MH = Mental health; MHD = Mental health domain; PHD = Physical health domain; CK = creatine kinase; IL-4 = Interleukin 4; IL-10 = Interleukin 10; *Superscripts* (^{a,b,c}). For a particular variable, mean values with different superscripts are significantly ($p < 0.05$) different; * indicates significant difference.

TABLE 3 - Comparisons of the participants' clinical parameters by treatment groups at the sixth week of the intervention

Variable	MEG (n = 25) x ± SD	SBEEEG (n = 25) x ± SD	DBEEEG (n = 26) x ± SD	F-ratio	p-value	η^2
ODI	8.7 ± 5.3	7.9 ± 4.0	10.4 ± 4.2	1.969	0.147	0.05
RMLDQ	3.6 ± 3.0	2.7 ± 1.3	3.8 ± 1.9	1.867	0.162	0.04
QVAS	25.8 ± 9.6	24.9 ± 6.2	26.1 ± 7.5	0.160	0.852	0.00
PF	70.6 ± 21.7	65.3 ± 20.4	77.1 ± 16.9	2.164	0.122	0.05
RLP	70.6 ± 13.1	73.4 ± 15.3	70.8 ± 13.3	0.309	0.735	0.00
RLE	70.6 ± 13.1	72.0 ± 14.2	67.9 ± 9.8	0.662	0.519	0.01
BP	82.8 ± 15.1	85.8 ± 8.4	83.8 ± 13.8	0.392	0.677	0.01
VT	81.6 ± 27.6	84.8 ± 26.0	91.7 ± 10.1	1.242	0.295	0.03
GH	72.8 ± 22.1	78.2 ± 14.7	76.4 ± 14.8	0.835	0.438	0.02
SF	80.0 ± 10.2	78.0 ± 8.2	77.1 ± 7.1	0.735	0.483	0.02
MH	69.6 ± 11.7	71.1 ± 6.4	71.4 ± 6.7	0.304	0.739	0.00
PHD	74.2. ± 10.9	75.7 ± 7.5	77.4 ± 8.4	0.761	0.471	0.02
MHD	75.5 ± 6.9	77.2 ± 7.1	77.0 ± 4.1	0.582	0.561	0.01
CK	133.8 ± 40.3	154.7 ± 21.7	140.7 ± 40.5	2.263	0.111	0.05
IL-4	58.8 ± 22.6[a]	84.2 ± 21.0[b]	64.1 ± 22.7[b]	9.219	0.001*	0.20
IL-10	38.0 ± 15.2[a]	60.7 ± 24.4[b]	42.5 ± 23.6[b]	7.851	0.001*	0.17

Alpha level was set at $p < 0.05$; X = Mean; SD = Standard deviation; MEG = McKenzie Exercise Group; SBEEEG = Static back extension endurance exercise group; DBEEEG = Dynamic back extension endurance exercise group; ODI = Oswestry Disability Index; RMLDQ = Roland Morris low back pain disability questionnaire; QVAS = Quadruple visual analogue scale; PF = Physical functioning; RLP = Role limitation physical; RLE = Role limitation emotion; BP = Bodily pain; VT = Vitality; GH = General health; SF = Social Functioning; MH = Mental health; MHD = Mental health domain; PHD = Physical health domain; CK = creatine kinase; IL-4 = Interleukin 4; IL-10 = Interleukin 10; *Superscripts* (^{a,b,c}). For a particular variable, mean values with different superscripts are significantly ($p < 0.05$) different; * indicates significant difference.

TABLE 4 - Within-group comparison of the effect of McKenzie exercise protocol across the three time points of the study (n = 25)

Variable	Baseline x ± SD	3rd week x ± SD	6th week x ± SD	F-ratio	p-value	η^2
ODI	24.8± 8.2[a]	14.6 ± 6.4[b]	8.8 ± 5.3[c]	53.206	0.001*	0.68
RMLDQ	15.0 ± 3.3[a]	5.2 ± 3.2[b]	3.6 ± 3.0[c]	144.25	0.001*	0.85
QVAS	65.6 ± 13.9[a]	38.8 ± 8.6[b]	25.8 ± 9.6[c]	96.017	0.001*	0.92
PF	24.7 ± 20.5[a]	50.7 ± 21.8[b]	70.7 ± 21.7[c]	36.151	0.001*	0.60
RLP	90.2 ± 16.0[a]	79.0 ± 17.5[b]	70.6 ± 13.0[c]	12.823	0.001*	0.34
RLE	87.4 ± 17.1[a]	70.6 ± 13.1[b]	70.6 ± 13.1[b]	14.961	0.001*	0.38
BP	55.4 ± 18.4[a]	74.4 ± 15.0[b]	82.8 ± 15.1[c]	16.894	0.001*	0.41
VT	54.4 ± 27.4[a]	72.8 ± 19.0[b]	81.6 ± 27.6[c]	9.864	0.001*	0.29
GH	38.2 ± 23.4[a]	61.2 ± 24.2[b]	72.8 ± 22.1[c]	20.937	0.001*	0.46
SF	87.0 ± 12.7[a]	81.0 ± 10.9[b]	80.0 ± 10.2[b]	4.016	0.040*	0.14
MH	60.0 ± 13.9[a]	68.8± 9.8[b]	69.6 ± 11.7[b]	4.972	0.013*	0.17
PHD	52.1 ± 10.5[a]	66.3 ± 10.7[b]	74.2 ± 10.9[c]	33.436	0.001*	0.58
MHD	72.2 ± 8.6	73.3 ± 8.0	75.5 ± 6.9	1.381	0.261	0.05
CK	69.2 ± 39.3[a]	134.1 ± 40.7[b]	133.8 ± 40.3[b]	32.638	0.001*	0.57
IL-4	55.9 ± 29.2	63.8 ± 29.1	58.8 ± 22.6	0.702	0.478	0.02
IL-10	40.8 ± 15.7	38.3 ± 14.4	38.0 ± 15.2	0.765	0.468	0.03

Alpha level was set at $p < 0.05$; X = Mean; SD = Standard deviation; ODI = Oswestry Disability Index; RMLDQ = Roland Morris low back pain disability questionnaire; QVAS = Quadruple visual analogue scale; PF = Physical functioning; RLP = Role limitation physical; RLE = Role limitation emotion; BP = Bodily pain; VT = Vitality; GH = General health; SF = Social Functioning; MH = Mental health; MHD = Mental health domain; PHD = Physical health domain; CK = creatine kinase; IL-4 = Interleukin 4; IL-10 = Interleukin 10; *Superscripts ([a,b,c]).* For a particular variable, mean values with different superscripts are significantly ($p < 0.05$) different; * indicates significant difference.

TABLE 5 - Within-group Comparison of the effect of SBEE exercise across the three time points of the study (n = 25)

Variable	Baseline x ± SD	3rd week x ± SD	6th week x ± SD	F-ratio	p-value	η^2
ODI	22.1± 7.2[a]	13.6 ± 6.1[b]	7.9 ± 4.0[c]	70.082	0.001*	0.74
RMLDQ	11.4 ± 3.3[a]	5.1 ± 2.7[b]	2.7 ± 1.3[c]	82.659	0.001*	0.77
QVAS	54.8 ± 13.8[a]	35.2 ± 8.6[b]	24.9 ± 6.2[c]	88.733	0.001*	0.78
PF	35.9 ± 27.7[a]	56.0 ± 19.2[b]	65.3 ± 20.4[c]	17.223	0.001*	0.41
RLP	91.6 ± 15.3[a]	72.0 ± 14.3[b]	73.4 ± 15.3[b]	13.855	0.001*	0.36
RLE	94.4 ± 13.1[a]	70.6 ± 13.1[b]	72.0 ± 14.3[b]	26.000	0.001*	0.52
BP	92.2 ± 7.6[a]	78.4 ± 13.4[b]	85.8 ± 8.4[c]	12.913	0.001*	0.35
VT	59.2 ± 21.2[a]	80.8 ± 22.7[b]	84.8 ± 26.0[c]	11.946	0.001*	0.33
GH	61.0 ± 24.2[a]	65.4 ± 18.3[a]	78.2 ± 13.4[b]	5.172	0.021*	0.17
SF	89.0 ± 12.7[a]	77.0 ± 6.9[b]	78.0 ± 8.3[b]	11.955	0.001*	0.33
MH	35.6 ± 10.3[a]	68.3± 7.4[b]	71.1 ± 6.4[c]	135.540	0.001*	0.85
PHD	70.2 ± 9.2[a]	67.9 ± 9.2[a]	75.7 ± 7.5[b]	6.915	0.003*	0.22
MHD	69.9 ± 7.5[a]	74.3 ± 7.9[b]	77.2 ± 7.2[c]	7.584	0.001*	0.24
CK	91.7 ± 55.4[a]	153.6 ±26.9[b]	154.7 ±21.7[b]	27.235	0.001*	0.53
IL-4	82.5 ± 37.6	82.9 ± 19.8	84.3 ± 21.0	0.032	0.926	0.00
IL-10	54.0 ± 27.9	61.3 ± 34.9	60.7 ± 24.4	0.816	0.424	0.03

Alpha level was set at $p < 0.05$; X = Mean; SD = Standard deviation; ODI = Oswestry Disability Index; RMLDQ = Roland Morris low back pain disability questionnaire; QVAS = Quadruple visual analogue scale; PF = Physical functioning; RLP = Role limitation physical; RLE=Role limitation emotion; BP=Bodily pain; VT = Vitality; GH = General health; SF = Social Functioning; MH = Mental health; MHD = Mental health domain; PHD = Physical health domain; CK = creatine kinase; IL-4 = Interleukin 4; IL-10 = Interleukin 10; *Superscripts ([a,b,c]).* For a particular variable, mean values with different superscripts are significantly ($p < 0.05$) different; * indicates significant difference.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 6 - Within-group comparison of the effect of DBEE exercise across the three time points of the study (n = 26)

Variable	Baseline x ± SD	3rd week x ± SD	6th week x ± SD	F-ratio	p-value	η^2
ODI	20.0 ± 6.8[a]	12.9 ± 4.8[b]	10.4 ± 4.2[c]	94.671	0.001*	0.79
RMLDQ	10.8 ± 2.8[a]	4.8 ± 2.5[b]	3.9 ± 1.9[b]	97.830	0.001*	0.79
QVAS	54.8 ± 13.9[a]	31.8 ± 5.6[b]	26.1 ± 6.7[c]	91.093	0.001*	0.78
PF	42.9 ± 22.9[a]	56.3 ± 20.7[b]	77.1 ± 16.9[c]	28.338	0.001*	0.55
RLP	79.6 ± 17.6	78.1 ± 17.3	70.8 ± 13.3	3.060	0.065	0.11
RLE	80.2 ± 17.7[a]	68.0 ± 10.1[b]	68.0 ± 10.1[b]	8.482	0.003*	0.27
BP	87.5 ± 5.7[a]	74.8 ± 12.7[b]	83.3 ± 13.9[c]	8.712	0.001*	0.27
VT	65.0 ± 21.5[a]	85.8 ± 16.2[b]	91.7 ± 10.1[c]	24.939	0.001*	0.52
GH	72.3 ± 22.4	66.7 ± 22.2	78.3 ± 14.8	3.235	0.054	0.12
SF	82.9 ± 11.6[a]	76.0 ± 5.1[b]	77.1 ± 7.1[b]	5.204	0.024*	0.18
MH	33.3 ± 12.1[a]	69.2 ± 14.2[b]	71.4 ± 6.6[c]	67.463	0.001*	0.74
PHD	70.6 ± 7.1[a]	68.9 ± 9.5[a]	77.4 ± 8.4[b]	8.421	0.001*	0.26
MHD	65.2 ± 6.9[a]	74.8 ± 5.0[b]	76.9 ± 4.2[c]	24.868	0.001*	0.53
CK	70.3 ± 47.4[a]	145.2 ±38.2[b]	140.7 ±40.5[c]	26.790	0.001*	0.51
IL-4	70.8 ± 24.9	70.7 ± 30.2	64.1 ± 22.6	1.023	0.367	0.03
IL-10	37.9 ± 24.9	42.4 ± 24.3	42.5 ± 23.6	0.656	0.511	0.02

Alpha level was set at $p < 0.05$; X = Mean; SD = Standard deviation; ODI = Oswestry Disability Index; RMLDQ = Roland Morris low back pain disability questionnaire; QVAS = Quadruple visual analogue scale; PF =Physical functioning; RLP = Role limitation physical; RLE=Role limitation emotion; BP = Bodily pain; VT = Vitality; GH = General health; SF = Social Functioning; MH = Mental health; MHD = Mental health domain; PHD = Physical health domain; CK = creatine kinase; IL-4 = Interleukin 4; IL-10 = Interleukin 10; *Superscripts ([a,b,c]).* For a particular variable, mean values with different superscripts are significantly ($p < 0.05$) different; * indicates significant difference.

Discussion

This study aimed to evaluate the impact of MEP, SBEE, and DBEE on various clinical outcomes (such as pain, functional disability, and general health status) as well as on biomarkers of muscle status (CK) and inflammation (anti-inflammatory cytokines, i.e., IL-4 and IL-10) in patients with chronic LBP. To our knowledge, this is the first study to provide empirical evidence of the biochemical effects of endurance exercises in the form of MEP and static-and-dynamic back extension in patients with chronic LBP. The average age of the patients in the study was 50.0 ± 7.5 years, falling within the 40-80 age brackets where LBP is commonly observed (2). The findings from within the study group indicated that MEP, SBEE, and DBEE had significant effects on pain intensity, activity limitation, participation restriction, and general health status. These results align with earlier studies showing evidence for the use of the McKenzie protocol (35-37) and back extensor exercises (34, 38, 39). Evidence has linked the therapeutic efficacy of the McKenzie protocol and back extension exercises in LBP to include lessening of pressure on sensitive tissue (40), reduction of load on the spinal disc (40, 41), increase in the height of the spine (42), and reduction in posterior protrusions in some intervertebral discs (43).

Regarding clinical outcomes, the study results showed that there were no significant differences in the effects of MEP, SBEE, and DBEE at the 3rd and 6th week of the study, except for pain intensity, which was significantly lower in DBEE compared to MEP and SBEE at the 3rd week. This indicates that all three interventions were equally effective and had similar effects on pain intensity, activity limitation, participation restriction, and general health status in patients with chronic LBP. The possible reason for this similar effect may be attributed to the mechanism of action of these exercise interventions, which are reported to be complementary. Although McKenzie's protocol works through centralization of pain by positively altering spinal mechanics, and back extension exercise works by improving the spinal muscle chain and spinal stability, both protocols involve spinal extension movements (44-46).

Similarly, there were no significant differences in the scores of IL-4 and IL-10 over time in patients receiving the SBEE and DBEE. The analysis of the biochemical effects of these interventions is an expanding area in physiotherapy research and could help improve patient care. The clinical outcomes align somewhat with expectations for similar interventions, but the results from blood samples are new. Biochemical markers have been explored as objective measures for risk assessment, diagnosis, or evaluation, as well as surrogate endpoints in chronic pain. Researchers have investigated biomarkers within pain pathways, such as inflammatory markers, molecular receptors detecting metabolites that influence pain, and neurotransmitters. These biomarkers are involved in the conduction, synaptic transmission, or modulation of pain response and are considered potential

therapeutic targets (47). Certain biochemical markers can provide diagnostic, prognostic, or intervention efficacy-related information (48). Additionally, they can indicate adaptation to regular training or changes occurring during or after exercise (48). However, there is a need to consolidate current studies assessing the effects of exercise on these biomarkers to evaluate their translational value in chronic pain. It is believed that the nature, volume, and intensity of exercise can produce varying effects on several biochemical markers (49).

Creatine kinase has been used as an indicator of training intensity and a diagnostic marker of overtraining. This suggests that the values fall within the normal range, indicating no harm to muscle status and no evidence of muscle damage (200-395 U/L for males and 200-207 U/L for females) (50). The findings on IL-4 and IL-10 support the argument submitted that training enhances the transcription of genes involved in the switch from a pro- to an anti-inflammatory macrophage phenotype (20). It is therefore expected to have more anti-inflammatory cytokines as a result of the training exercises. Anti-inflammatory cytokines are a series of immunoregulatory molecules that control the pro-inflammatory cytokine response. IL-4 and IL-10 are among the anti-inflammatory cytokines, but IL-10 is the most important anti-inflammatory cytokine found within the human immune system (51). Even while there is no significant rise in the cytokine levels, there is still some increase, which may be more at some points, showing some increase in the deposits of anti-inflammatory cytokines, therefore setting the stage for muscle recovery.

This is the first study to investigate the biochemical effect of MEP and endurance exercises in patients with chronic LBP. Hence, there is an apparent dearth of similar studies that directly compare findings. However, some emerging evidence indicates that non-pharmacologic interventions may positively modulate inflammatory mechanisms in chronic LBP (52). In their systematic review, Puerto Valencia et al. found that non-pharmacologic interventions such as yoga, acupressure, osteopathy, isokinetic training, neuro-emotional technique, etc., resulted in the decrease of pro-inflammatory markers (e.g., IL-6) and increase in the anti-inflammatory markers (e.g., IL-4) (52) in patients with chronic LBP. Biomarkers play a crucial role in monitoring patient responses, predicting outcomes, and guiding treatment decisions, which enhances the precision and effectiveness of interventions for patients with LBP (51). For instance, CK levels can indicate muscle damage, while cytokine profiles can help tailor anti-inflammatory treatments, leading to a more personalized and effective management approach for LBP (50,53,54). Additionally, IL-4 and IL-10 are important biomarkers as well. Elevated IL-10 levels suggest reduced inflammation and pain, whereas lower levels may indicate persistent inflammation (53, 54). Similarly, high IL-4 levels can signal a positive treatment response by promoting anti-inflammatory effects, while low IL-4 levels could suggest inadequate anti-inflammatory responses, leading to ongoing discomfort (53, 54).

As per earlier reports, this study found that the MEP and back extension endurance exercises are beneficial in managing lower back pain (LBP). Additionally, testing for serum biomarkers to assess muscle status, potential muscle damage, and inflammatory responses resulting from these interventions is an innovative approach. The study's results showed that while there was an increase in the level of serum CK following MEP, SBEE, and DBEE, the changes were not significant enough to indicate substantial muscle damage.

Clinical Implications of the Study

Creatine kinase is mainly a marker for muscle damage, while IL-4 can have both pro-inflammatory and anti-inflammatory effects, generally promoting anti-inflammatory responses by encouraging T helper type 2 (TH2) cell development and playing a crucial role in the immune response (55). IL-10 is predominantly anti-inflammatory, which is crucial for limiting inflammatory responses by inhibiting pro-inflammatory cytokine production (56). At weeks three and six, CK levels were elevated across all groups, indicating ongoing muscle activity or damage due to the exercise interventions, with no significant differences between groups. IL-4 and IL-10 levels were significantly higher in the SBEE group compared to the MEP group, suggesting that static exercise promotes a stronger and more effective anti-inflammatory response. Across all three exercise interventions (within-group comparison), CK levels significantly increased from baseline to week three and remained elevated at week six, indicating substantial muscle activity or potential damage. IL-4 and IL-10 levels remained relatively stable throughout the intervention periods, which may imply consistent anti-inflammatory responses for all exercise types. However, key differences include the slight increase in IL-4 levels at week three for the MEP group, which returned closer to baseline by week six, while IL-4 levels in the SBEE and DBEE groups remained stable. Additionally, IL-10 levels showed a slight increase at week three for the SBEE groups but remained stable, indicating similar anti-inflammatory responses across all interventions. This suggests that the interventions are effective and safe. Moreover, the minimal increase in IL-4 and IL-10 may indicate muscle healing. Therefore, understanding these biochemical processes can lead to more relevant treatment options and reduce the likelihood of failed treatments, providing significant benefits for biomarker development. Using serum biomarkers has the potential to identify more tailored treatments with improved efficacy in specific patient populations.

Limitations of the study

The clinical outcomes and biochemical changes of the different interventions were only assessed over a relatively short period of six weeks, which may be different at long-term follow-up. The non-blinding of therapist-to-treatment allocation may have also introduced performance bias. It must be stated that this potential bias was minimized by blinding the therapist to group allocation and the assessment of clinical and biochemical outcomes. The patients in this study had similar general characteristics and some baseline clinical parameters, including CK scores. Baseline characteristics are important indicators of how patients will respond to treatment in clinical trials for lower back pain (57). It is known that having comparable baseline measures in clinical

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

trials decreases the likelihood of factors other than the intervention affecting the outcomes. However, Friedman et al. (58) argued that for many measurements, baseline data may not accurately reflect the participant's true condition at the time of the baseline, as investigators typically perform the baseline assessment close to the time of the intervention. Therefore, the results obtained at different points in this study could have been largely influenced by the effects of the various treatment regimens. The lack of a placebo or control group in this study may also serve as a limitation. Also, the findings from this study are limited only to patients with chronic LBP who have a directional preference for extension. Finally, the significant differences in some baseline characteristics of the groups may have moderated the results, potentially influencing the observed outcomes.

Conclusion

MEP reduced pain more, while SBEE led to higher changes in IL-4 and IL-10 inflammatory biomarker levels. Serum CK levels rose in all groups without indicating muscle damage or significant inflammation. The significance of this study is that MEP showed a greater efficacy in reducing movement-evoked pain and central sensitization, while SBEE may offer broader anti-inflammatory benefits.

Acknowledgments

In memory of Dr. Kamil Lasisi, Head of Musculoskeletal Physiotherapy at Osun State University Teaching Hospital, Osogbo, Nigeria, whose immense efforts and influence on this work still resonate.

Disclosures

Conflict of interest: The authors declare no conflict of interest.

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author's contributor role: Conceptualization: CM, MIS; Data collection, curation, analysis, and interpretation: MIS, RO, AA, TG, GY, FF, CM; Manuscript writing, review, and final approval: MIS, RO, AA, TG, GY, FF, CM

Data availability statement: The data presented in this study are available on request from the corresponding author.

References

1. Deyo RA, Mirza SK, Martin BI. Back pain prevalence and visit rates: estimates from U.S. national surveys, 2002. *Spine*. 2006;31(23):2724-2727. CrossRef
2. Hoy D, March L, Brooks P, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis*. 2014;73(6):968-974. CrossRef
3. Stochkendahl MJ, Kjaer P, Hartvigsen J, et al. National Clinical Guidelines for non-surgical treatment of patients with recent onset low back pain or lumbar radiculopathy. *Eur Spine J*. 2018;27(1):60-75. CrossRef
4. National Institute for Health and Care Excellence. *Low back pain and sciatica in over 16s: assessment and management. NICE guideline.* [NG59] NICE; 2016.
5. Koes BW, van Tulder M, Lin CWC, et al. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. Eur Spine J. 2010;19:2075-94. CrossRef
6. Ferreira ML, Smeets RJ, Kamper SJ, et al. Can we explain heterogeneity among randomized clinical trials of exercise for chronic back pain? A meta-regression analysis of randomized controlled trials. *Phys Ther*. 2010;90(10):1383-1403. CrossRef
7. Agency for Healthcare Research and Quality (AHRQ). Complementary and alternative therapies for back pain II. 10. (Evidence Report, Technology Assessment; Band 194), 2010. PubMed PMCID: PMC4781408.
8. Cost B. 13 working group. European guidelines for the management of chronic non-specific low back pain. *Eur Spine J*. 2006;15(S2):S192-S300. CrossRef
9. Furlan AD, Imamura M, Dryden T, et al. Massage for low-back pain. *Cochrane Database Syst Rev*. 2008;4:CD001929.
10. Lee SW, Nguyen D, Mack D, et al. Conservative management of low back pain. *HCA Healthc J Med*. 2021;2(5):319-328. CrossRef
11. Hayden JA, Ellis J, Ogilvie R, et al. Exercise therapy for chronic low back pain. *Cochrane Database Syst Rev*. 2021;9:CD009790.
12. Gordon R, Bloxham S. A systematic review of the effects of exercise and physical activity on non-specific chronic low back pain. *Healthcare (Basel)*. 2016;4(2):22. CrossRef
13. van Middelkoop M, Rubinstein SM, Verhagen AP, et al. Exercise therapy for chronic non-specific low-back pain. *Best Pract Res Clin Rheumatol*. 2010;24(2):193-204. CrossRef
14. Mbada CE, Ayanniyi O, Ogunlade SO, et al. Influence of McKenzie protocol and two modes of endurance exercises on health-related quality of life of patients with long-term mechanical low-back pain. *Pan Afr Med J*. 2014;17(suppl 1):5. CrossRef
15. Aydın E, Turan Y. Biochemical markers for osteoarthritis: is there any promising candidate? *Meandros Med Dent J*. 2016;17(1):27-34. CrossRef
16. Alonso-Sal A, Alonso-Perez JL, Sosa-Reina MD, et al. Effectiveness of physical activity in the management of non-specific low back pain: a systematic review. *Rev Med (São Paulo)*. 2024;60(12):2065. CrossRef
17. Morris P, Ali K, Merritt M, et al. A systematic review of the role of inflammatory biomarkers in acute, subacute and chronic non-specific low back pain. *BMC Musculoskelet Disord*. 2020;21(1):142. CrossRef
18. Chazaud B. Inflammation during skeletal muscle regeneration and tissue remodeling: application to exercise-induced muscle damage management. *Immunol Cell Biol*. 2016;94(2):140-145. CrossRef
19. Hyldahl RD, Hubal MJ. Lengthening our perspective: morphological, cellular, and molecular responses to eccentric exercise. *Muscle Nerve*. 2014;49(2):155-170. CrossRef
20. Gordon PM, Liu D, Sartor MA, et al. Resistance exercise training influences skeletal muscle immune activation: a microarray analysis. *J Appl Physiol*. 2012;112(3):443-453. CrossRef
21. Garcia-Cosamalon J, del Valle ME, Calavia MG, et al. Intervertebral disc, sensory nerves and neurotrophins: who is who in discogenic pain? *J Anat*. 2010;217(1):1-15. CrossRef
22. Matousek RH, Dobkin PL, Pruessner J. Cortisol as a marker for improvement in mindfulness-based stress reduction. *Complement Ther Clin Pract*. 2010;16(1):13-19. CrossRef
23. Licciardone JC, Kearns CMs, Hodge LM, et al. Associations of cytokine concentrations with key osteopathic lesions and clinical outcomes in patients with non-specific chronic low back pain: results from the OSTEOPATHIC Trial. *J Am Osteopath Assoc*. 2012;112:596-605. CrossRef

24. Pinto EM, Neves JR, Laranjeira M, et al. The importance of inflammatory biomarkers in non-specific acute and chronic low back pain: a systematic review. *Eur Spine J*. 2023;32(9): 3230-3244. CrossRef
25. Van den Berg R, Jongbloed EM, De Schepper EIT, et al. The association between pro-inflammatory biomarkers and non-specific low back pain: a systematic review. *Spine J*. 2018;18(11):2140-2151. CrossRef
26. Chan YH. Randomised controlled trials (RCTs)-sample size: the magic number? *Singapore Med J*. 2003;44(4):172-174.
27. Mbada CE, Akindele FF, Fatoye CT, et al. Translation and psychometric evaluation of the Yoruba Version of Quadruple Visual Analogue Scale. *Niger J Health Sci*. 2018;18(2):63-69. CrossRef
28. Mbada CE, Oguntoyinbo OE, Fasuyi FO, et al. Cross-cultural adaptation and psychometric evaluation of the Yoruba version of Oswestry disability index. *PLoS One*. 2020;15(1):e0221138. CrossRef
29. Mbada CE, Idowu OA, Ogunjimi R, et al. Cross-cultural adaptation, reliability and validity of the Yoruba version of the Roland Morris Disability Questionnaire. *Spine*. 2017;42(7):1. CrossRef
30. Mbada CE, Awokoya AS, Oyewole OO, et al. Translation, cross-cultural adaptation and psychometric evaluation of Yoruba version of the short-form 12 health survey. *Ann Ig*. 2021;33(3):254-267.
31. Pocock SJ. Allocation of patients to treatment in clinical trials. *Biometrics*. 1979;35(1):183-197. CrossRef
32. Clare H, Adams R, Maher CG. A systematic review of efficacy of McKenzie therapy for spinal pain. *Aust J Physiother*. 2004;50(4):209-216. CrossRef
33. de Oliveira IO, Pinto LLS, de Oliveira MA, et al. McKenzie method for low back pain. *Rev Dor*. 2016;17:4. CrossRef
34. Mbada CE, Olaoye MI, Dada OO, et al. Comparative efficacy of clinic-based and telerehabilitation application of mckenzie therapy in chronic low-back pain. *Int J Telerehabil*. 2019;11(1):41-58. CrossRef
35. Nwuga G, Nwuga V. Relative therapeutic efficacy of the Williams and McKenzie protocols in back pain management. *Physiother Pract*. 1985;1(2):99-105. CrossRef
36. Cherkin DC, Deyo RA, Battiè M, et al. Comparison of physiotherapy, chiropractic manipulation, and provision of an educational booklet for the treatment of patients with low back pain. *N Engl J Med*. 1998;339(15):1021-1029. CrossRef
37. Machado LAC, De Souza MVS, Ferreira PH, et al. The McKenzie method for low back pain. A systematic review of the literature with a meta-analysis approach. *Spine*. 2006;31(9):E254-E262. CrossRef
38. Yudurum SG, Snijders CJ. On the form of the human spine and some aspects of its mechanical behaviour. *Acta Orthop Belg*. 1969;35:584.
39. Chok B, Lee R, Latimer J, et al. Endurance training of the trunk extensor muscles in people with subacute low back pain. *Phys Ther*. 1999;79(11):1032-1042. CrossRef
40. Adams MA, May S, Freeman BJ, et al. Effects of backward bending on lumbar intervertebral discs. Relevance to physiotherapy treatments for low back pain. *Spine*. 2000;25(4): 431-437. CrossRef
41. Adams MA, McNally DS, Dolan P.' Stress' distribution inside intervertebral discs: the effects of age and degeneration. *J Bone Joint Surg Br*. 1996;78(6):965-972. CrossRef
42. Magnusson ML, Aleksiev AR, Spratt KF, et al. Hyperextension and spine height changes. *Spine*. 1996;21(22):2670-2675. CrossRef
43. Fennell AJ, Jones AP, Hukins DW. Migration of the nucleus pulposus within the intervertebral disc during flexion and extension of the spine. *Spine*. 1996;21(23):2753-2757. CrossRef
44. Ameh V, Egbuchunam CU, Ibeachu C. Effect of McKenzie protocol and dynamic back endurance exercise on health-related quality of life of patients with long-term mechanical low back pain. *Pan Afr Med J*. 2017;17(1):5.
45. Gill G, Sharma R, Patel M. Efficacy and effectiveness of McKenzie exercises in chronic low back pain management. A comprehensive review. *Int J Health Sci Res*. 2024;14(9): 213-222. CrossRef
46. Bharath KH, Kumar S, Reddy P. Effectiveness of McKenzie and core strengthening exercise in a patient with chronic low back pain. A case study. *Int J Res Rev.* 2024;11(9):371-372. CrossRef
47. DeLeo JA, Yezierski RP. The role of neuroinflammation and neuroimmune activation in persistent pain. *Pain*. 2001;90(1):1-6. CrossRef
48. Califf RM. Biomarker definitions and their applications. Exp Biol Med (Maywood). 2018;243(3):213-221. CrossRef
49. Sanchis-Gomar F, Lippi G. Physical activity-an important preanalytical variable. *Biochem Med (Zagreb)*. 2014;24(1):68-79. CrossRef
50. Tamaki T, Uchiyama S, Tamura T, et al. Changes in muscle oxygenation during weight-lifting exercise. *Eur J Appl Physiol Occup Physiol*. 1994;68(6):465-469. CrossRef
51. WHO International Programme on Chemical Safety Biomarkers in Risk Assessment. Validity and Validation. 2001. Online (Accessed October 2024)
52. Puerto Valencia LM, He Y, Wippert PM. The changes of blood-based inflammatory biomarkers after non-pharmacologic interventions for chronic low back pain: a systematic review. *BMC Musculoskelet Disord*. 2024;25(1):209. CrossRef
53. Khan AN, Jacobsen HE, Khan J, et al. Inflammatory biomarkers of low back pain and disc degeneration: a review. Ann N Y Acad Sci. 2017;1410(1):68-84. CrossRef
54. Banimostafavi ES, Fakhar M, Abediankenari S, et al. Determining serum levels of IL-10 and IL-17 in patients with low back pain caused by lumbar disc degeneration. *Infect Disord Drug Targets*. 2021;21(5):67-71. CrossRef
55. Biedermann T, Röcken M. Pro-and anti-inflammatory effects of IL-4: from studies in mice to therapy of autoimmune diseases in humans. Animal Models of T cell-mediated skin diseases; 2005:235-242.
56. Vincent JL. Pro- and anti-inflammatory biomarkers. In: Molnar Z, Ostermann M, Shankar-Hari M, eds. *Management of Dysregulated Immune Response in the Critically Ill. Lessons from the ICU.* Springer; 2023, CrossRef
57. Underwood MR, Morton V, Farrin A. Do baseline characteristics predict response to treatment for low back pain? Secondary analysis of the UK BEAM dataset [ISRCTN 32683578]. *Rheumatology (Oxford)*. 2007;46(8):1297-1302. CrossRef
58. Friedman BW, Chilstrom M, Bijur PE, et al. Diagnostic testing and treatment of low back pain in US emergency departments. A national perspective. *Spine*. 2010;35(24):E1406. CrossRef

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 250-264
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3428
REVIEW

Metric properties of the Tonic Stretch Reflex Threshold (TSRT) as a measure of spasticity: a systematic review with meta-analysis

Diego Longo[1], Marco Baccini[2], Giulio Cherubini[1], Aurora Caporaso[3], Mario De Marco[1], Angela Politi[2], Guido Santini[1], Francesca Cecchi[1,2], Maria Angela Bagni[1]

[1]Department of Experimental and Clinical Medicine, University of Florence, Florence - Italy
[2]IRCCS Fondazione Don Gnocchi Onlus, Florence - Italy
[3]Bachelor's degree programme in Physiotherapy, University of Florence, Florence - Italy

ABSTRACT

Introduction: Spasticity is a common symptom after brain injury, often interfering with functional recovery and rehabilitation. The Tonic Stretch Reflex Threshold (TSRT) was proposed as an objective neurophysiological assessment of spasticity that could overcome the limitations of clinical scales. This systematic review aimed to appraise the current evidence on the metric properties of TSRT.

Methods: Electronic databases (MEDLINE, CINAHL, Scopus, Web of Science, and EMBASE) were screened from inception to June 30, 2025, for studies reporting data on reliability, validity, and/or responsiveness of TSRT in adults with stroke. Two reviewers independently selected the studies, assessed the methodological quality, and extracted relevant data. When possible, pooled estimates for each property were computed.

Results: Of the 9804 titles retrieved, 17 were eventually included, to which 2 articles from cross-references were added. We found insufficient values for both intra-rater (two studies, ICC = 0.548, 0.330-0.710) and inter-rater (three studies, ICC = 0.687, 0.511-0.808) reliability, with high measurement error. Data on validity were found in 14 articles, with conflicting results on the association of TSRT with clinical scales of spasticity and motricity, but good ability to discriminate among relevant groups. Only one study investigated responsiveness with an external anchor, finding that TSRT measurements failed to accurately detect improved participants.

Conclusion: Despite the potential of TSRT as a measure of spasticity, its metric properties, particularly reliability, are not fully supported. Future research should prioritize improving its reliability and investigating its validity and responsiveness with neurophysiological measures rather than relying solely on clinical scales.

Keywords: Adult, Muscle spasticity, Reproducibility of results, Stretch reflex, Stroke

What is already known about this topic?

- *Clinical scales for spasticity have limited reliability and validity. The Tonic Stretch Reflex Threshold (TSRT) has been proposed as an objective measure of spasticity for use in both clinical and research settings.*

What the study adds?

- *The reliability of the TSRT measurement is currently not fully supported, and this might affect its validity and responsiveness. Future research should focus on improving reliability through the use of better instrumentation.*

Introduction

Spasticity is one common and disabling symptom after a first motor neuron injury (1), with prevalence ranging from 25% to 38% in stroke survivors (2-4). It interferes with functional recovery and rehabilitation processes, can lead to secondary complications such as muscle retractions, weakness, and pain (5), and impairs activities of daily living and sleep (6).

The pathophysiological mechanisms of post-stroke spasticity are still debated, and several neuromusculoskeletal computational models have been used to study them. Most models focus on reflex gains and thresholds as the main parameters related to its severity, including both neural (alpha and gamma motor neuron firing, afferent input mainly from muscle spindles) and non-neural (viscoelastic elements of the muscular-tendon unit) components (7). The neural component is generally accepted as being caused primarily by supraspinal

Received: December 10, 2024
Accepted: August 21, 2025
Published online: September 19, 2025

This article includes supplementary material

Corresponding author:
Marco Baccini
email: marco.baccini@unifi.it

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

dysregulation of the spinal reflex loop, to which growing evidence indicates that imbalance of the dorsal and medial reticulo-spinal tracts contributes (8). The uncontrasted excitatory input through the medial reticulo-spinal tract would increase the intrinsic excitability of motoneurons, facilitating persistent inward currents, i.e., depolarizing currents that tend to remain activated and are associated with subthreshold depolarization of spinal motor neurons (8).

There is still no complete agreement on the definition of spasticity. Early in 1980 Lance (9) defined spasticity as "A motor disorder caused by a speed-reflex increase of the stretch tonic reflex with exaggerated tendon response, resulting from a hyperexcitability of the stretch reflex as a component of the syndrome of the first motoneuron". Later, Pandyan et al. (10) extended the concept to all positive symptoms of an upper motor neuron lesion redefining it as "disordered sensorimotor control, presenting as intermittent, or sustained involuntary involvement of muscles". More recently, Dressler et al. (11) suggested a simple, operational definition as "involuntary muscle hyper-activity triggered by rapid passive joint movements," which differentiates it from other symptoms such as rigidity, spasm or dystonia, and from hypertonia, which may also exist at rest. In 2017, 37 experts from 12 European countries participating in two consensus meeting, based on a Delphi approach (12), concluded that the term "hyper-resistance," instead of hypertonia or spasticity, should be preferred to describe the abnormal neuromuscular response during passive stretch, to which both neural and non-neural components contribute, and that the term "spasticity" should be used with care, and only referring, in agreement with Lance (9), to velocity dependent stretch hyperreflexia as part of the neural contributions to hyper-resistance.

The lack of a common definition is at least partly behind the variety of methods for assessing spasticity. In a recent review, 60 different tools were identified (13), including 33 clinical, 18 biomechanical, and eight neurophysiological measures. The Ashworth scale (AS) and its modified version (MAS) are the most used clinical measures of spasticity in the adult population. However, besides showing questionable metric properties (14,15), they assess resistance to passive movement without distinguishing between neural and non-neural components of resistance (13,14), and show low correlation with neurophysiological tests like H-reflex-based measurements (16,17). The poor validity of clinical scales, including the AS, has even been suggested as one of the causes of the failure of many clinical trials in neurology (18).

A direct measurement of spasticity has been proposed in the context of indirect, referent control theory of motor control (19,20). A main mechanism in posture and movement control is the ability to set and reset the spatial threshold of reflexes, named lambda (λ). In the healthy nervous system, the threshold regulatory range is defined by the task-specific ability to activate or relax muscles at each joint position within the biomechanical range of motion (21). Thus, the range of threshold regulation is larger than the biomechanical range, whereas, after neurological injury, the range of possible λ-shifts is limited. Spasticity is the result of a reduction in the central regulatory range of λ thresholds: when the upper limit of λ range (R+) of a muscle is within the biomechanical range of motion, an externally imposed stretching will find an active muscular resistance beginning at R+ and increasing with increasing lengthening. The λ threshold decreases as the rate of muscle stretch increases, and its equivalent parameter at the joint level is called stretch reflex threshold (SRT) (19), which represents the joint angle at which motoneurons of the involved muscles begin to be recruited. This parameter depends on stretch velocity and is commonly referred to as Dynamic Stretch Reflex Threshold (DSRT), whereas the Tonic Stretch Reflex Threshold (TSRT) represents a specific value of DSRT for zero velocity. In particular, when the TSRT is within the biomechanical range and the subject cannot shift it, it discriminates the joint configurations in which the muscles are spastic from those in which they are not (22,23).

Another parameter detected through TSRT measurement is coefficient μ, a time-dimensional parameter that characterizes the velocity sensitivity of TSRT (19). Theoretically, it correlates with the index of sensitivity, changing from dynamic to static muscle spindle afferents, and may be partly controlled by dynamic γ-motor neurons' activity (19). In a velocity/angle plot, μ represents the slope of the dynamic activation area boundary of the muscle and determines a temporal variation in the entry of the muscle into the activation area (19). This parameter is as essential as TSRT, because it is crucial to understand the actual dependence on the velocity of the observed hypertonia.

Surface electromyography (sEMG) combined with a joint angle detection system is used to detect DSRTs (24). The assessment procedure consists of applying several passive stretches to the target muscles at different speeds. For each speed, DSRT is identified as the angle at which reflex EMG activity appears. A linear regression through the DSRTs obtained from repeated stretches is conducted to compute the TSRT (24). In the original procedure, stretches are applied manually, and measures are taken with the Montreal Spasticity Measure instrumentation, which comprises an electrogoniometer, a two-channel surface electromyograph, and dedicated software (24). A complete description on how the TSRT is calculated can be found in a recent critical review on this topic (25), which, however, highlights that the methodologies used by different researchers vary greatly in terms of equipment, number, velocity and mode of application of stretches (manually or with a torque motor), time interval between stretches, criteria for defining the onset of reflex muscle activity, analyses conducted to calculate TSRT and μ. All these features considerably impact the measurement and might explain some inconsistencies in the findings of different research groups.

The aim of this review is to critically appraise, summarize, and meta-analyze the current literature on the metric properties of TSRT as a measure of spasticity in adults with upper motor neuron lesions.

Methods

This review was conducted according to COSMIN (Consensus-Based Standards for the Selection of Health Measurement Instruments) guidelines (26). It was registered prospectively in the PROSPERO database (PROSPERO 2023 CRD42023412289) as a more general systematic review on instrumental methods for assessing spasticity.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Data Sources and Searches

We searched five databases (MEDLINE, EMBASE, CINAHL, SCOPUS, and WOS) for literature published in English on April 20, 2023, with no initial date limit. The search strategy followed the COSMIN guidelines (27) and contained the following words: (1) spasticity, (2) tone, (3) cerebral palsy, (4) stroke, (5) spinal cord injury, (6) upper motor neuron, (7) measure, (8) evaluation, and (9) assessment. A combination of logic Boolean operators was implemented to finalize the existing COSMIN search filters for each database (Supplementary material). The reference list of all relevant articles was also searched for other eligible publications. An updated search was conducted on June 30, 2025, including only the term Tonic Stretch Reflex Threshold or TSRT.

Study Selection

We included studies that investigated any metric property (28) of the TSRT measurement in adults with spasticity, with no restrictions regarding spasticity etiology. We included only studies where the TSRT was measured following the procedure described by Calota et al. (24), so excluding the articles where this term was used with a different meaning, e.g., to simply indicate the joint angle of onset of reflex muscle contraction triggered by a passive stretch, even if instrumentally measured. Subjects with cerebral palsy outcomes were included if they were assessed when older than 18 years. Studies reporting data on metric properties were included even if this evaluation was not the primary aim. Two authors independently performed the study selection. After removing duplicates, titles, abstracts, and, when needed, full texts were analyzed for eligibility. Disagreements were solved by a third researcher.

Data Extraction

Two authors independently extracted the following data from the articles: authors, year of publication, population, comparator tool, when applicable, muscles evaluated, metric properties evaluated, statistics conducted, and values found for each property, as detailed below. In case of disagreement, consensus was reached with the intervention of a third author.

Reliability

Intraclass correlation coefficients (ICC) were considered the most appropriate index, and the statistics used for their computation were also recorded, when provided; other correlation coefficients were also considered appropriate if the study assessed and excluded significant differences between the two measurements. Standard Error of the Measurement (SEM), Minimal Detectable Change (MDC), Coefficient of Variation, and Bland-Altman's Limits of Agreement were considered appropriate indices of measurement error.

Criterion validity

If the study used a reference measure for spasticity, the correlation with the TSRT was noted. If the study dichotomized the threshold measurement result (e.g., spasticity present/absent), diagnostic accuracy values (sensitivity, specificity, area under the receiver operating characteristic curve, AUC) were collected.

Construct validity

All data about the a priori hypotheses tested were collected, including convergent (e.g., correlation between TSRT and motor impairment severity) and discriminative validity, i.e., the ability of the TSRT to discriminate among relevant subgroups.

Responsiveness

Any index of responsiveness, both internal (e.g., effect size) or external (e.g., correlation with an anchor measure of change), was collected. When provided, the Minimal Clinically Important Difference (MCID) computed in the study was noted.

Quality Assessment

Two authors independently assessed the risk of bias of the included articles, and disagreements were resolved by consensus, with the intervention of a third author. For reliability studies, the Quality Appraisal for Reliability (QAREL) checklist was used (29), which consists of 11 items that can be answered "yes" if the item is satisfied, "no" if the item is not satisfied, "unclear" if the information is insufficient, or "not applicable." Although the COSMIN checklists on risk of bias (30) also cover reliability, the QAREL checklist was preferred for reliability because it lists all possible sources of error in more detail. The COSMIN checklists (30) were used instead for validity (boxes 8 and 9) and responsiveness (box 10). Each COSMIN item may be rated as "very good," "adequate," "doubtful," or "inadequate," when applicable, and the lowest score assigned to any item of a box is taken to establish the risk of bias.

Data Synthesis and Analysis

Pooled estimations were obtained after Fisher's z transformation of correlation coefficients to normalize their distribution (31), including the estimation of the prediction interval (PI), as recommended (32). All meta-analyses were based on random effects models to account for the heterogeneity of estimates (33). Heterogeneity and inconsistency were assessed using the Q statistic and the I^2 statistic, respectively (34).

When missing, SEM and MDC_{95} were computed provided that authors reported the necessary data, by applying the formulas $SEM = DS \times \sqrt{(1 - ICC)}$ and $MDC = SEM \times 1{,}96 \times \sqrt{2}$.

Convergent validity was assessed through the correlation of TSRT and μ with clinical scales of spasticity and motricity. When relevant indices were not reported in the article, they were computed provided that the authors reported the necessary data. When data were provided in graphical form, values were extracted from graphs using Microsoft PowerPoint 2021. If this was not possible, the authors were contacted and asked to provide the requested values or raw data for computation.

Results

Out of 9804 titles retrieved from the databases, 17 (35-51) were ultimately included (see Fig. 1 for details on the selection process). Cross-referencing led to the inclusion of two other studies (19,52). Table 1 shows the characteristics of the included studies.

All studies enrolled individuals with stroke outcomes, more often in the chronic phase (at least six months after the event) (19,35,37,42,43,45,49,51,52). Four studies enrolled participants in the subacute phase (less than 6 months) (44,46, 48, 50), and five studies had a mixed population (36,38-41), whereas in one study (47), this information is not reported. In one study (44), the sample also included two patients with spinal cord injury and one with traumatic brain injury. Sample size ranged from 4 to 55 participants. TSRT was measured at the elbow flexors (19,35-37,39-50,52), elbow extensors (19,36,39), and ankle plantar flexors. (38,51)

Four articles (35,37,38,46) evaluated TSRT reliability, whereas in 14 articles (19,35,36,39,41-48,51,52) data about TSRT validity were reported or retrievable. One article (46) studied both internal and external responsiveness of TSRT, using the upper limb motor section of the Fugl-Meyer Assessment Scale (53) (FMA-UL) as an anchor for the latter.

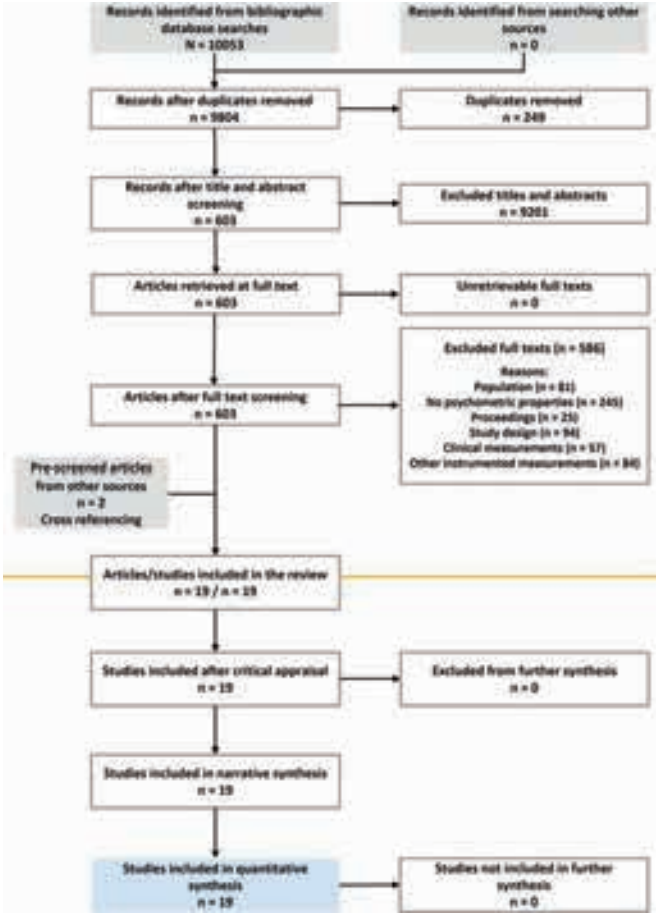


FIGURE 1 - Flow-chart of the process of study selection.

However, we were able to extract data about internal responsiveness from three other studies (40,49,50).

Reliability and measurement error

Data on intra-rater and inter-rater reliability were reported by two (35,37) and three (35,38,46) articles, respectively. Ferreira et al. (37) did not compute any reliability coefficient, so we analyzed their data to calculate the $ICC_{2,1}$, including only participants with stroke who completed all the assessments. Only one of these studies (46) provided the measurement error, but we computed it from the raw data reported in the other three articles (35,37,38). All studies obtained a QAREL score <50% (Fig. 2), indicating a possible high risk of bias.

Intra-rater reliability

The intra-rater reliability of biceps brachii TSRT measurement in a sample of subjects with stroke outcome was assessed by Calota et al. (35) (15-18 participants assessed twice, with a 2-7 days delay, by three different raters) and by Ferreira et al. (37) (14 participants assessed three times, with 2-5 days delay, by the same rater). We meta-analyzed the data of these studies (Fig. 3, top), finding a pooled ICC = 0.548 (95% CI: 0.330, 0.710; 95% PI: 0.330, 0.710; sample size = 64), with no heterogeneity and high consistency (Q = 0,778, p = 0.855; I^2 = 0,00%). The pooled MDC_{95} was 41°.

Inter-rater reliability

TSRT inter-rater reliability was evaluated for biceps brachii (35,46) and triceps surae (38), involving two (46), three (35), or nine (38) raters. Since Calota et al. (35) reported data measured by three raters at two different occasions, data from four comparisons were meta-analyzed. The pooled ICC was 0.687 (95% CI: 0.511, 0,808; 95% PI: 0.346, 0.868; I^2 = 50,41%; Q = 6,049, p = 0.109; sample size = 123) (Fig. 3, bottom). Including only data on biceps brachii resulted in a lower value of pooled ICC (0.618; 95% CI: 0.471, 0,732; 95% PI: 0.471, 0,732; sample size = 95), with no heterogeneity (Q = 0,454, p = 0,796) and inconsistency (I^2 = 0.00%). The MDC_{95} was 14.55° for triceps surae and 35.16° for biceps brachii.

We classified the overall rating of inter-rater and intra-rater reliability and measurement error of TSRT as insufficient for using the measure both at the group and individual levels. We rated the level of evidence as low (risk of bias and imprecision).

Validity

Convergent and discriminative validity of TSRT was evaluated by analyzing the correlation with measures of similar constructs (clinical scales of spasticity) or of a construct that might theoretically be associated (motor function), and comparing values measured in different groups (healthy, subacute or chronic stroke, Parkinson's disease), respectively. The quality of ratings of included studies is shown in Table 2. Usually, the authors did not state their a-priori hypotheses about the expected correlations.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 1 - Characteristics of included studies

Study	Participants, number, and condition (M/F)	Time since event, months (mean (SD))	Severity of paresis/ spasticity[§]	Age, years (mean (SD))	Muscle evaluated	Property evaluated[#]	Devices	Number of passive stretches	Algorithm for muscle activity onset	TSRT value[§]
Levin et al., 2000 (19)	11 CS (7/4)	47.3 (22.4)	FMA-UL: 37.3 (19.5); CSI: 10.6 (3.9)	41.2 (12.1)	EF-EE[‡]	Convergent validity (CSI, FMA-UL)	sEMG, manipulandum coupled to a torque motor and a digital resolver	56	>2 SD over baseline for 50ms at least	118.0 (29.2)
Calota et al., 2008 (35)	20 CS (16/4)	77.9 (111.7)	MAS: 1.7 (1.3)	62.9 (12.9)	EF	Reliability Convergent validity (MAS)	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	135.4 (24.1)
Kim et al., 2011 (36)	9 SS, 6CS (7/8)	6.7 (9.1)	MAS: 1.4 (0.4)	63.5 (15.6)	EF	Convergent validity (MAS)	sEMG, electrogoniometer	About 50	>2 SD over baseline	123.6 (4.6)
Ferreira et al., 2013 (37)	14 CS (11/3)	53.9 (38.4)	NR	60.21 (14.40)	EF	Reliability	sEMG, electrogoniometer	30	Modified double threshold protocol	188.7 (37.6)
Blanchette et al., 2016 (38)	2 SS, 26 CS (21/7)	47.9 (43.8)	NR	57.4 (9.7)	PF	Reliability	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	66.0 (13.1),
Turpin et al., 2017 (39)	2 SS, 11 CS (12/1)	62.3 (69.2)	FMA-UL: 36.6 (13.9); CSI: 9.1 (2.6)	56.1 (9.1)	EF-EE	Convergent validity (CSI, FMA-UL)	sEMG, manipulandum coupled to a torque motor and optical encoder	15	>3SD over baseline	105.3 (20.5)
Vieira et al., 2017 (40)	4 SS, 12 CS (12/4)	7.6 (2.5)	NR	63.9 (8.5)	EF	Internal responsiveness*	sEMG, electrogoniometer	30	>2SD over baseline for 200ms at least	69.7 (19.3)
Levin et al., 2018 (41)	13 SS, 20 CS (26/7)	20.4 (23.6)	FMA-UL: 22.2 (15.8); MAS: 1,4 (0.4)	52.2 (8.7)	EF	Convergent validity (MAS)	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	85.1 (13.5)
Subramanian et al., 2018 (42)	12 CS (10/2)	54.7 (71.4)	FMA-UL: 39.2 (12.0); CSI: 9.5 (2.8)	58.5 (9.5)	EF	Convergent validity (CSI, FMA-UL)*	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	134.8 (15.1)
Marques et al., 2019 (43)	10 CS (8/2)	35.4 (29.2)	MAS: 1.4 (0.4)	68.2 (11.6)	EF	Convergent validity (MAS)	sEMG, electrogoniometer	30	>2SD over baseline for 200ms at least	49.8 (57.8)
Zhang et al., 2019 (44)	14 SS, 1 TBI, 1 SCI (14/2)	1,7 (0.3)	MAS: 1.7 (0.6)	54 (10)	EF	Convergent validity (MAS)	sEMG, IMU	15-20	>3 SD over baseline	49.9 (21.7)
Alves et al., 2021 (45)	22 CS (16/6)	71.4 (139.3)	MAS: 1.4 (0.6)	56.1 (11.3)	EF	Convergent validity (MAS)	sEMG, electrogoniometer	30	>2 SD over baseline for 200ms at least	41.4 (33.7)

(*Continued*)

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

TABLE 1 - (*Continued*)

Study	Participants, number, and condition (M/F)	Time since event, months (mean (SD))	Severity of paresis/ spasticity[§]	Age, years (mean (SD))	Muscle evaluated	Property evaluated[#]	Devices	Number of passive stretches	Algorithm for muscle activity onset	TSRT value[§]
Frenkel-Toledo et al., 2021 (46)	55 SS (37/18)	2.2 (1.4)	FMA-UL: 29.6 (15.9); MAS: 1.4 (0.4)	54.3 (10.7)	EF	Reliability Convergent validity (MAS) Responsiveness	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	100.6 (20.3)
Wang et al., 2021 (47)	16 Stroke (11/5)	NR	NR	51.6 (13.6)	EF	Convergent validity (MAS)*	Exoskeletal device coupled to sEMG, torque sensor, and angular position sensor	9	NR	66.69 (19.29)
Frenkel-Toledo et al., 2022 (48)	45 SS (25/16)	2,2 (1,3)	FMA-UL: 32.3 (12.2)	53.66 (11.27	EF	Convergent validity (FMA-UL, MAS)	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	107.4 (25.7)
Marques et al., 2022 (49)	10 CS (8/2)	36 (20.8)	NR	59.7 (9.07)	EF	Internal responsiveness*	sEMG, electrogoniometer	6 at least	>2SD over baseline for 200ms at least	35.2 (28.3)
Levin et al., 2023 (50)	35 SS (25/10)	2.2 (1.4)	FMA-UL: 32.6 (13.1); MAS: 1.4 (0.4)	53.7 (11.3)	EF	Internal responsiveness*	MSM (sEMG, electrogoniometer)	20 at least	>2 SD over baseline for 25 ms at least	101.0 (20.2)
Longo et al., 2023 (51)	10 CS (6/4)	134.4 (124.3)	MAS: 2.6 (0.7)	52.1 (15.0)	PF	Convergent validity (MAS)*	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	88.2 (13.8)
Mullick et al., 2013 (52)	10 CS (6/4)	58.3 (19.5)	CSI 8.5 (2.2)	65.3 (12.2)	EF-EE	Convergent validity (CSI)	sEMG, manipulandum coupled to a torque motor and precision gauges	30	>2SD over baseline for 50 ms at least	133.7 (16.2)

[§]When multiple assessments were conducted (test-retest, pre-post intervention), data from the first assessment are reported, pooling data of different groups or different raters when needed.
[#]The asterisk indicates that the estimation of metric properties was not an aim of the study.
[‡]Elbow extensors assessed in 4 participants.
M/W = men/women (ratio); SS = subacute stroke; CS = chronic stroke; TBI = traumatic brain injury; SCI = spinal cord injury; FMA-UL = Fugl-Meyer assessment scale, upper limb motor section; CSI = Composite Spasticity Index; MAS = Modified Ashworth Scale; EF = elbow flexors; EE = elbow extensors; PF = plantar flexors; sEMG = surface Electromiography; IMU = inertial measurement units; MSM = Montreal Spasticity Measure; NR = not reported; SD = Standard Deviation; TSRT=Tonic Stretch Reflex Threshold.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

FIGURE 2 - Quality ratings of studies on TSRT reliability.

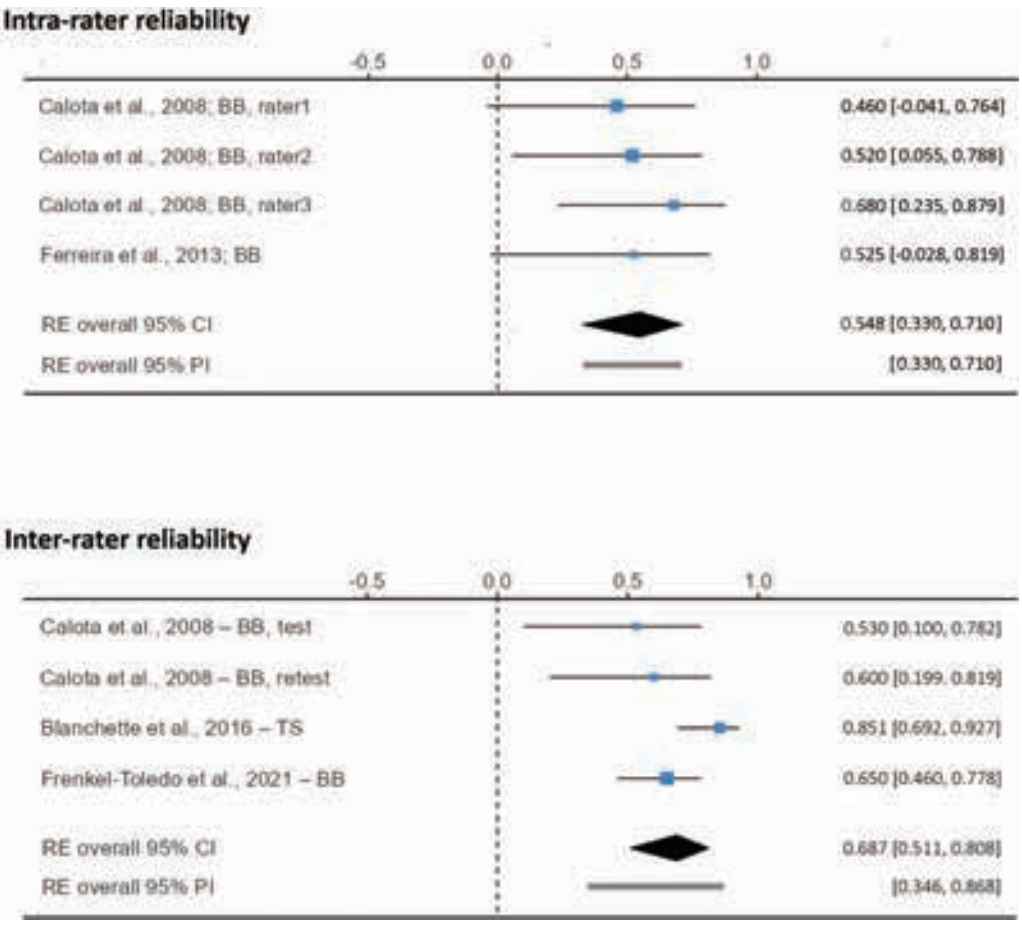


FIGURE 3 - Meta-analysis of studies on intra-rater reliability (top) and inter-rater reliability (bottom) of TSRT. *(BB = biceps brachii).*

We could retrieve data on the association of TSRT with spasticity clinical scales from 14 articles that used the Modified Ashworth Scale (MAS) (35,36,41,43-48,51) or the Composite Spasticity Index (CSI) (19,39,42,52). CSI (54) sums up three features, namely resistance to passive movement, clonus, and exaggerated tendon reflexes. Only few studies reported the Spearman correlation coefficient (19,35,43,45,46,48), so we computed it from the raw data of the other articles (36,39,41,42,44,47,51,52), in two cases by extracting values from graphs (44,47), and in one case from data provided on request by the authors (39). Figure 4 shows the meta-analysis conducted aggregating data from all studies, resulting in a pooled coefficient of −0.305 (95% CI: −0.500, −0.081; 95% PI: −0.822, 0.487; sample size = 283) with high heterogeneity and inconsistency (Q = 66.862, $p < 0.001$; I^2 = 74.57%). Pooling the data separately for each comparator, heterogeneity persists with MAS (rho = −0.327; 95% CI: −0.557, −0.050; PI: −0.853, 0.528; Q = 60.060, $p < 0.001$; I^2 = 81.68%; sample size = 238), not with CSI (rho = −0.240; 95% CI: −0.537, 0.011; PI: −0.670, 0.311; Q = 6.717, p = 0.243; I^2 = 25.56%; sample size = 45).

We estimated the association of biceps brachii TSRT with motor impairment as measured by the FMA-UL from data presented in five studies (19,39,41,42,48). Only one

TABLE 2 - Quality appraisal of studies on TSRT construct validity and responsiveness

	Convergent validity (COSMIN box 9a)					Discriminative validity (COSMIN box 9b)			
Study	**1**	**2**	**3**	**4**	**Risk-of-bias score**	**5**	**6**	**7**	**Risk-of-bias score**
Levin et al., 2000 (19) – CSI	VG	D	VG	D*	doubtful				
Levin et al., 2000 (19) – FMA-UL	VG	A	VG	D*	doubtful				
Calota et al., 2008 (35) – MAS	VG	D	VG	D*	doubtful				
Kim et al., 2011 (36) – MAS	VG	D	VG	D*	doubtful				
Turpin et al., 2017 (39) – CSI	VG	A	VG	D*	doubtful				
Turpin et al., 2017 (39) – FMA-UL	VG	A	VG	D*	doubtful				
Levin et al., 2018 (41) – MAS	VG	D	VG	D*	doubtful				
Levin et al., 2018 (41) – FMA-UL	VG	A	VG	D*	doubtful				
Levin et al., 2018 (41) – chronic vs subacute patients						VG	VG	D*	doubtful
Subramanian et al., 2018 (42) – CSI	VG	A	NA^	D*	doubtful				
Subramanian et al., 2018 (42) – FMA-UL	VG	A	NA^	D*	doubtful				
Mar+ques et al., 2019 (43) – MAS	VG	D	VG	D*	doubtful				
Zhang et al., 2019 (44) – MAS	VG	D	A	D*	doubtful				
Zhang et al., 2019 (44) – healthy vs spastic participants						VG	VG	D*	doubtful
Alves et al., 2021 (45) – MAS	VG	D	VG	D*	doubtful				
Frenkel-Toledo et al., 2021 (46) – MAS	VG	D	VG	D*	doubtful				
Frenkel-Toledo et al., 2021 (46) –FMA-UL	VG	A	VG	D*	doubtful				
Wang et al., 2021 (47) – MAS	VG	D	NA^	D*	doubtful				
Frenkel-Toledo et al., 2022 (49) – MAS	VG	D	VG	D*	doubtful				
Frenkel-Toledo et al., 2022 (49) –FMA-UL	VG	A	VG	D*	doubtful				
Longo et al., 2023 (51)	VG	D	NA^	D*	doubtful				
Mullick et al., 2013 (52) – CSI	VG	A	VG	D*	doubtful				
Mullick et al., 2013 (52) – UPDRS	VG	VG	VG	D*	doubtful				
Mullick et al., 2013 (52) – healthy, stroke vs Parkinson						VG	VG	D*	doubtful

	Responsiveness (COSMIN box 10a and 10d)						
Study	**10a_1**	**10a_2**	**10a_3**	**10d_11**	**10d_12**	**10d_13**	**Risk-of-bias score**
Frenkel-Toledo et al., 2021 (46)	VG	NA	D**	VG	VG	VG	doubtful

1: Is it clear what the comparator instrument(s) measure(s)? 2: Were the measurement properties of the comparator instrument(s) sufficient? 3: Was the statistical method appropriate for the hypotheses to be tested? 4: Were there any other important flaws in the design or statistical methods of the study? 5: Was an adequate description provided of important characteristics of the subgroups? 6: Was the statistical method appropriate for the hypotheses to be tested? 7, 10a_3, 10d_13: Were there any other important flaws in the design or statistical methods of the study? 10a_1: For continuous scores: were correlations between change scores, or the AUC calculated? 10a_2: For dichotomous scales: were sensitivity and specificity determined? 10d_11: Was an adequate description provided of the intervention given? 10d_12: Were statistical methods appropriate for the before-after comparison being made? VG = very good; AD = adequate; D = doubtful; IN = inadequate; N/A = not applicable.
* Blinding to the comparator or to groups.
** Suitability of the anchor.
^Not applicable because the study did not aim to evaluate the TSRT metric properties.

study (19) also measured triceps brachii TSRT and reported a significant positive correlation with motor impairment for both muscles, whereas one study reported a non-significant negative correlation (48); two studies (39,41) just reported the absence of significant correlation, and one study (42) did not investigate such an association. We computed the Spearman correlation from the raw data presented in the articles (41,42) or provided on request by the authors (39). The meta-analysis (Fig. 5, top) showed a pooled correlation coefficient = 0.117 (95% CI: −0.140, 0.360; PI: −0.293, 0.492; Q = 6.970; p = 0.223; I^2 = 28.26%; sample size = 109).

Only two studies also investigated the association between FMA-UL and μ of biceps brachii (19,41), with conflicting results, i.e., non-significant negative correlation (41), and significant positive correlation (19). A pooled rho = 0.119 (95% CI: −0.391, 0.574; PI: −0.575, 0.714; Q = 3.439; p = 0.064; I^2 = 50.9%; sample size = 44) resulted from the meta-analysis (Fig. 5, bottom).

We classified the overall rating of TSRT and μ convergent validity with clinical scales of motor impairment or spasticity as indeterminate, with a very low level of evidence (imprecision, risk of bias, and very serious inconsistency).

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

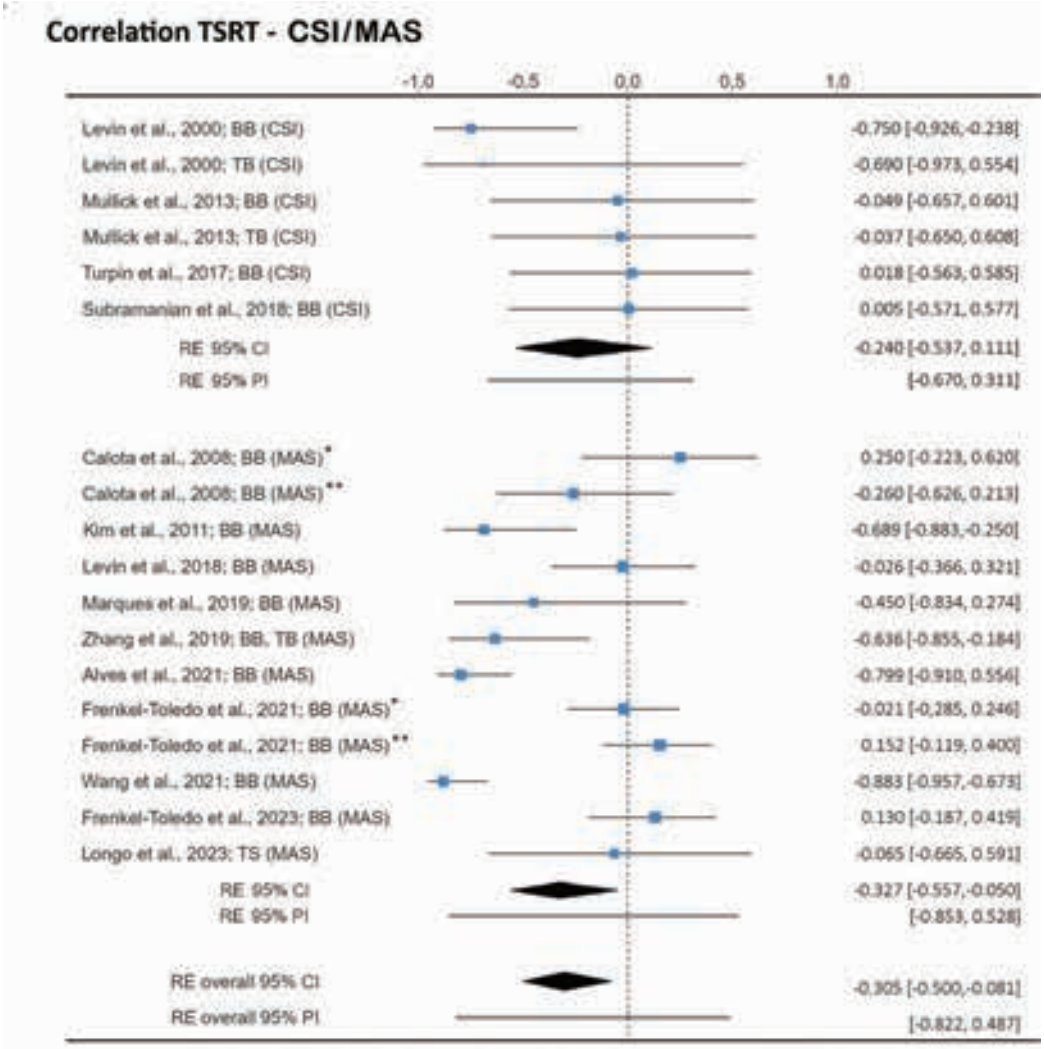


FIGURE 4 - Meta-analysis of studies on TSRT convergent validity: correlation with clinical scales of spasticity. (*BB = biceps brachii; TB = triceps brachii; MAS = Modified Ashworth Scale; CSI = Composite Spasticity Index; *measured at first assessment (test); **measured at second assessment (retest)*).

As for discriminative validity, significant differences were found between TSRT values measured in healthy and spastic subjects (44,52), and between TSRT and μ values measured in participants with stroke and Parkinson's disease (52). Conversely, no differences were found between persons with subacute and chronic stroke (41). We classified the overall rating of TSRT discriminative validity as positive, with a low level of evidence (imprecision, risk of bias).

Responsiveness

We found only one study (46) that comprehensively evaluated this property in 44 participants who received 10 sessions of upper limb intervention combining Virtual Reality and real/sham transcranial direct current stimulation. Contrary to MAS, TSRT changed significantly in the whole group, detecting a medium treatment effect (mean change: 8.79 ± 24.96; Effect Size = 0.40; Standardized Response Mean = 0.35). External responsiveness was examined using a receiver-operating characteristic curve approach and classifying participants as improved or unimproved according to a change in FMA-UL above or below, respectively, the MDC of this scale. A cut-off value of 6.8° TSRT change was found with near significance (Area Under Curve, AUC = 0.671, p = 0.056). The risk of bias of this study was rated as doubtful due to the doubtful suitability of the chosen anchor (Table 2).

We retrieved data on internal responsiveness (Effect Size, ES) in three other studies not specifically designed to evaluate this psychometric property. We found large ESs in a study (40) enrolling a small number of persons with stroke outcomes who were treated with biofeedback (N = 6, ES = 1.17) or conventional physical therapy (N = 6, ES = 0.86), and in another study (49) where 10 participants with chronic stroke practiced 15 sessions of training based on serious game and virtual reality (ES = 0.92). A third study (50) reported an ES = 0.570 in a group of 46 participants with chronic stroke who were engaged in 10 sessions of personalized reaching training. Considering the very different rehabilitation interventions, which probably have different effectiveness, we decided not to aggregate the data from these studies.

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

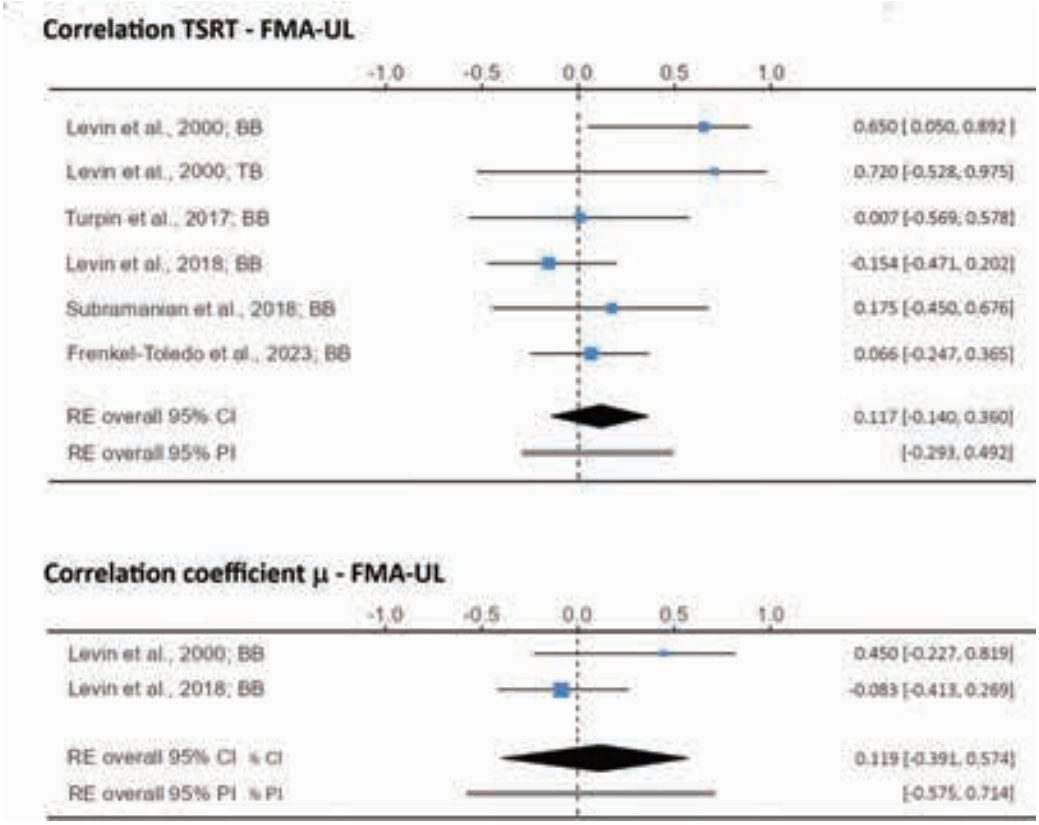


FIGURE 5 - Meta-analysis of studies on convergent validity of TSRT (top) and coefficient μ (bottom): correlation with Fugl-Meyer Assessment Scale–motricity section. *(BB = biceps brachii; TB = triceps brachii).*

We classified the overall rating of TSRT responsiveness as insufficient, with a very low level of evidence (imprecision, risk of bias).

Discussion

This systematic review investigated the psychometric properties of TSRT as a potential objective measure of spasticity, finding only four articles on reliability and up to 14 articles reporting data on its convergent validity, even if many of them do not explicitly refer to TSRT validity. This is surprising since reliability is a prerequisite for exploring criterion and construct validity of any assessment tool (28), but it is rather common in this field. A narrative review on new tools for assessing spasticity (55), i.e., tools that were developed from 2016 to 2022, identified four new clinical scales, eight medical imaging methods, and three spasticity assessment devices. From the data presented, it appears that none of them have been subjected to a thorough evaluation of metric properties, and in particular, reliability. In a recent systematic review (56) aimed at evaluating the quality of clinical assessment tools for spasticity, the Tardieu scale was the only clinician-reported instrument recommended, since it is consistent with the definition of spasticity as a speed-dependent increase of stretch reflexes. The authors, however, did not meta-analyze data on Tardieu scale metric properties and pointed out that its reliability has not been fully established.

We will first discuss the results concerning TSRT reliability because, in our opinion, they may partly explain the conflicting results regarding validity.

Reliability

The data presented show that TSRT reliability, either intra- or inter-rater, is not sufficiently supported. Although different criteria are used for ICC interpretation, the values found in the present review should be considered low when referring either to the COSMIN criteria (30) (ICC = 0.70 at least), or to the more conservative criteria proposed by Fitzpatrick et al. (57) (ICC = 0.70 or = 0.90 at least for use in assessing groups or individual performance, respectively). The insufficient reliability is confirmed by the high measurement error for both biceps brachii [over 35° out of a total range of about 150° (58)] and triceps surae [14.55° out of a total passive range of approximately 74° (58)]. For biceps brachii, MDC_{95} is far higher than the MCID of TSRT reported by Frenkel-Toledo et al. (46), so it should be rated as negative also according to COSMIN criteria (30).

One source of error might be the data collection procedure and, in particular, the assessment of joint angles through electrogoniometry. Very few studies evaluated the reliability of electrogoniometry, with results that raise serious doubts about this property. Two studies assessed the reliability of range of motion measurement at the ankle (59,60) in small samples (<20) of healthy subjects. The first

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

study (59) only stated that no significant differences were found between the repeated measurements, whereas the other (60) reported high test-retest reliability (ICC = 0.987; SEM = 3,63°), but computing the average measure ICC, which overestimates reliability. Only one non-recent study (61) evaluated the reliability of electrogoniometry at the elbow in 23 female healthy volunteers, finding very low ICCs (ranging 0.00-0.43).

The main issue of this kind of assessment is the difficulty of avoiding small shifts of the goniometers' arms and axis during limb mobilization that might generate artefacts in data recording. In order to reduce this source of error, it would likely be better to replace the electrogoniometer with more reliable motion analysis instruments. Marker-based motion capture systems, for example, are considered a gold standard in human biomechanics assessment (62). Such systems are hardly feasible as a routine evaluation procedure in clinical practice, but a recent study reports that markerless systems could be quite reliable, with significant advantages in terms of usability and rapidity of procedures (63). Indeed, some authors (19,39,52) measured the TSRT with different devices (a horizontal manipulandum coupled to a torque motor, with an angular position sensor placed on the axis of the manipulandum to measure joint angles) that could reduce the random variability of the measurement. Unfortunately, none of them investigated the reliability of this alternative procedure.

sEMG has been more widely investigated in terms of reliability at both elbow (64-67) and ankle muscles (68-70), and showed high test-retest reproducibility in different conditions. However, a key point for TSRT measurement is the detection of the onset of muscle activity from EMG signals, which depends on the criteria adopted. Silva et al. (71) compared the performance of four different criteria – two of which were employed in two studies (36,38) included in the present review – and found that the TSRTs measured by each method differed significantly from one another. However, this should affect the validity of the measurements, rather than their reliability.

Validity

The lack of consistent correlations between TSRT and clinical measures of spasticity is somewhat surprising. However, some authors underline the difficulty in distinguishing between spasticity and spastic dystonia, the latter considered due to the secondary soft tissue changes that occur in upper motor neuron syndrome (72-74) because both features may co-occur in these patients (72). A clinical evaluation of the resistance to passive stretch could hardly discriminate between these intrinsic components of spasticity that are strictly related to the phase of recovery after stroke. For this reason, we could expect different results when enrolling chronic or subacute subjects, but such a difference was not evidenced by the results of our review. We found conflicting results between studies that enrolled only chronic patients (19,42,52) and between studies that enrolled only subacute patients (44,46). In both cases, we find studies reporting strong negative correlation (19,44) or no correlation (42,46, 52). Therefore, the phase after stroke does not seem to be a relevant factor for this association. Indeed, recent findings show that architectural changes in muscles are driven by both reduced mobility and muscle overactivity, which increase passive resistance and promote progression to fibrosis, and may also increase stretch sensitivity and reflex mechanisms (75).

In a very recent article, Piscitelli et al. (76) reported data about the correlation between elbow flexors TSRT measured in subjects with acute, subacute, or chronic stroke, and clinical scales of motor impairment (FMA-UL) and spasticity (MAS). The authors found a significant, negative correlation between TSRT and MAS in the total sample (N = 247) and in the subgroup of acute/early subacute participants (N = 158), and a significant, positive correlation between TSRT and FMA-UL in the subgroup of chronic participants (N = 33). We decided not to include these data in our review because in this article, the authors pooled data from eight studies conducted in their laboratory, five of which are unpublished studies and therefore cannot be assessed for risk of bias. Moreover, we already included two (35,39) of the three published studies reported in this article in our meta-analysis on the correlation of TSRT with clinical scales, both enrolling only, or overwhelmingly, participants with chronic stroke. The first one (35) found conflicting results on correlation with MAS at two different assessments, and the other (39) found no correlation with CSI nor with FMA-UL. We could not include in the meta-analysis the third study that enrolled participants with subacute stroke because the published article (50) reports no data on correlation with clinical scales and only provides data on changes in TSRT after a period of training, which was included in our analysis of TSRT internal responsiveness. Nevertheless, the pooled correlation coefficients found in the present meta-analysis are about the same as those reported by Piscitelli et al. (76) for the whole sample, both for MAS (−0.327 vs −0.32) and FMA-UL (0.117 vs 0.08), although in our analyses the high heterogeneity resulted in wide prediction intervals.

We believe that insufficient TSRT reliability may be a factor behind the high inter-study variability, but the lack of reliability of clinical scales (MAS and CSI) might also contribute. Indeed, poor to moderate reliability was found for both the original and the modified AS (14,77-79). Not sufficient data are available for CSI reliability; instead, it was evaluated only in two studies (80,81) in very limited samples of patients who had a stroke (N = 10), with different results. The authors reported ICC = 0.87 (80) and ICC = 0.97 (81), but did not compute the measurement error. From the data presented in the two articles, we estimated an MDC_{95} of 3.4 and 0.67, respectively, i.e., quite different values.

Conflicting results might also depend on the method used to measure TSRT, which varied among studies. However, this does not seem to be the case, as conflicting data were also reported in studies that adopted the same method. For example, two studies on participants with chronic stroke (19,52) adopted same device, same algorithm to detect EMG activity onset, and same clinical scale as a comparator (CSI). However, results ranged from a strong, negative correlation (19) to no correlation (52).

Similar considerations can be made for the conflicting results on the association between TSRT and motor scales; in this case, problems mainly concern the reliability of TSRT, since the reliability of the FMA scale is fairly well established (82,83). However, the lack of correlation with motor scales is not necessarily an unexpected finding; in fact, TSRT and FMA assess quite different aspects of motor control, and we may find people with severe paresis who have either high spasticity or no spasticity. Thus, the conflicting findings might also be related to the very limited sample sizes, ranging from 4 to 33 participants, who are likely not fully representative of the population with stroke outcomes. It is worth noting that one study (39) found that motor impairments as measured by FMA-UL were negatively associated with the ability to modulate stretch reflex thresholds between active and passive stretch rather than to passive-state values. This is an interesting finding, because during active stretch, inhibitory circuits responsible for reciprocal inhibition may be involved (84), as well as central inhibition of motor neurons (85), and both may be related to impairment of active movement (72). In our opinion, this finding could support the validity of TSRT measurements more than the correlation between motor scales and TSRT during passive stretch, but it needs to be confirmed in studies with adequate power.

Given the lack of reliable and valid clinical scales for spasticity, and the different constructs assessed by motor function scales, not necessarily correlated with spasticity as defined by Lance (9), we would have expected to find studies comparing TSRT to other neurophysiological measures, such as H-reflex-related measures. The H-reflex is a neurophysiological technique for evaluating spasticity that offers insights into spinal cord excitability and sensorimotor integration and can be employed to investigate different spinal mechanisms involved in motor control, such as presynaptic and reciprocal inhibition (86). Indeed, we found two articles (87,88) that assessed the association between some parameters detected by H-reflex test and a parameter somewhat related to TSRT, named Reflex Threshold Angle (RTA), measured in plantar flexors of patients with post-stroke spasticity. Both studies found that dysfunctions in reciprocal and presynaptic inhibition, as measured by H-reflex, correlated very weakly with RTA. However, RTA is defined as the angle at which a simultaneous increase in torque and EMG activity occurs during a passive stretch (87). It seems fully comparable with DSRT rather than with TSRT, so these studies were excluded from the present review. Future research should investigate whether TSRT, which specifies where the spastic zone begins along the muscle length, shows a stronger association with H-reflex parameters.

Conversely, the findings on between-groups differences, although obtained on still very limited samples, strongly support the discriminative validity of the measure, in particular with regards to differences between spasticity (participants who had a stroke) and rigidity (participants with Parkinson's disease). It should be noted that, in group comparisons, less precision of measurements is required, so insufficient reliability has less impact on the results. This indirectly supports the hypothesis that reliability problems affect the results of correlations with clinical scales.

Responsiveness

Responsiveness of TSRT has not been largely investigated, but the single study that comprehensively evaluated this property (46) reported rather negative results. Although the estimation of responsiveness is still a matter of debate, and several different methods are employed (89), only the use of an external anchor makes it possible to evaluate whether an instrument is able to detect important changes (90). In this respect, Frenkel-Toledo et al. (46) found that TSRT measurements failed in detecting participants who had changed, since the AUC was not significantly different from 0.50. However, two factors should be considered when interpreting their results. First, the sample size (N = 44) is too low for evaluating responsiveness (30). Most important, the authors used changes in FMA-UL as an external anchor, a questionable choice considering the low correlation between the two measures, as highlighted above. Future research on this property should enrol an adequate number of subjects and choose different anchors. We believe that subjective anchors such as the Global Rating of Perceived Change questionnaire are hardly applicable to this symptom, and instrumental measures of spasticity should be preferred. Conceivably, H-reflex-related measures are better suited to detect true changes in spasticity and could therefore be a more appropriate anchor in responsiveness studies.

Limitations

Some limitations impose caution in interpretating the results, primarily the low quality of many studies, the small sample size, and the high heterogeneity of many meta-analytical estimates. A more direct limitation might be the failure to retrieve all relevant titles, since about 10% of the articles were not found through database searches. However, this can largely be explained by the fact that in these studies, the evaluation of the TSRT metric properties was not an objective or was not explicitly stated. Indeed, 1 out of 2 articles included by cross-referencing (52) did not conduct any analyses of metric properties, and we re-analyzed the data reported in this article to compute correlations between TSRT and clinical scales, whereas the other article (19) is a study on motor control where the authors just mention results on correlation between TSRT and clinical scales in the main text. Nevertheless, we can wonder whether useful data, particularly on the association with clinical scales, might be found in other published studies.

Conclusions

Despite the potential of TSRT as an objective measure of spasticity, its metric properties are currently not fully supported, particularly with regard to reliability. Likely, the high measurement error contributes at least in part to the inconsistent results regarding convergent validity with clinical measures of spasticity. In order to reduce the random error, the use of alternative instruments to the electrogoniometer is recommended, as already implemented by some authors. Future research should prioritize improving the reliability of

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TSRT by using alternative instrumentation. Once reliability is established, the validity and responsiveness of the measure should be evaluated using neurophysiological measures as a reference standard rather than relying solely on clinical spasticity scales, because the reliability and validity of the latter are also questionable.

Disclosures

Conflict of Interest: The authors declare no conflict of interest.

Financial support: The funders played no role in the design, conduct, or reporting of this study.

Author's contributor role: DL: study design, studies selection process, quality appraisal, writing – draft; MB: study design, quality appraisal, data analysis, writing – editing; GC: studies selection process, writing – editing; AC: studies selection process, writing – editing; MDM: quality appraisal, writing – draft; AP: data extraction, writing – draft; GS: data extraction, writing – editing; FC: study design, writing – editing; MAB: study design, writing – editing.

Data Availability Statement: The data presented in this study are not publicly available but are available on request from the corresponding author.

References

1. Gallichio JE. Pharmacologic management of spasticity following stroke. Phys Ther. 2004;84(10):973-981. CrossRef
2. Schinwelski M, Sławek J. Prevalence of spasticity following stroke and its impact on quality of life with emphasis on disability in activities of daily living. Systematic review. Neurol Neurochir Pol. 2010;44(4):404-411. CrossRef
3. Zeng H, Chen J, Guo Y, et al. Prevalence and risk factors for spasticity after stroke: a systematic review and meta-analysis. Front Neurol. 2021;11:616097. CrossRef
4. Cheng H, Fang X, Liao L, et al. Prevalence and factors influencing the occurrence of spasticity in stroke patients: a retrospective study. Neurol Res. 2023;45(2):166-172. CrossRef
5. Kheder A, Nair KP. Spasticity: pathophysiology, evaluation and management. Pract Neurol. 2012;12(5):289-298. CrossRef
6. Bethoux F, Marrie RA. A cross-sectional study of the impact of spasticity on daily activities in multiple sclerosis. Patient. 2016;9(6):537-546. CrossRef
7. da Silva VA, da Silva RL, Withers JW, et al. Neuromusculoskeletal modeling of spasticity: A scoping review. PLoS One. 2025;20(5): e0320153. CrossRef
8. Li S, Chen YT, Francisco GE, Zhou P, Rymer WZ. A Unifying Pathophysiological Account for Post-stroke Spasticity and Disordered Motor Control. Front Neurol. 2019;10:468. CrossRef
9. Lance JW. The control of muscle tone, reflexes, and movement: Robert Wartenberg Lecture. Neurology. 1980;30(12): 1303-1313. CrossRef
10. Pandyan AD, Gregoric M, Barnes MP, et al. Spasticity: clinical perceptions, neurological realities and meaningful measurement. Disabil Rehabil. 2005;27(1-2):2-6. CrossRef
11. Dressler D, Bhidayasiri R, Bohlega S, et al. Defining spasticity: a new approach considering current movement disorders terminology and botulinum toxin therapy. J Neurol. 2018;265(4): 856-862. CrossRef
12. van den Noort JC, Bar-On L, Aertbeliën E, et al. European consensus on the concepts and measurement of the pathophysiological neuromuscular responses to passive muscle stretch. Eur J Neurol. 2017;24(7):981-e38. CrossRef
13. Aloraini SM, Alyosuf EY, Aloraini LI, Aldaihan MM. Assessment of spasticity: an overview of systematic reviews. Phys Ther Rev. 2022;27(4):290-303. CrossRef
14. Fleuren JF, Voerman GE, Erren-Wolters CV, et al. Stop using the Ashworth Scale for the assessment of spasticity. J Neurol Neurosurg Psychiatry. 2010;81(1):46-52. CrossRef
15. Pandyan AD, Johnson GR, Price CI, et al. A review of the properties and limitations of the Ashworth and modified Ashworth Scales as measures of spasticity. Clin Rehabil. 1999;13(5):373-383. CrossRef
16. Stokic DS, Yablon SA, Hayes A. Comparison of clinical and neurophysiologic responses to intrathecal baclofen bolus administration in moderate-to-severe spasticity after acquired brain injury. Arch Phys Med Rehabil. 2005;86(9):1801-1806. CrossRef
17. Martens G, Deltombe T, Foidart-Dessalle M, et al. Clinical and electrophysiological investigation of spastic muscle overactivity in patients with disorders of consciousness following severe brain injury. Clin Neurophysiol. 2019;130(2):207-213. CrossRef
18. Hobart JC, Cano SJ, Zajicek JP, et al. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations [published correction appears in Lancet Neurol 2008 7: 25]. Lancet Neurol. 2007;6(12):1094-1105. CrossRef
19. Levin MF, Selles RW, Verheul MH, et al. Deficits in the coordination of agonist and antagonist muscles in stroke patients: implications for normal motor control. Brain Res. 2000;853(2): 352-369. CrossRef
20. Feldman AG. Indirect, referent control of motor actions underlies directional tuning of neurons. J Neurophysiol. 2019;121(3): 823-841. CrossRef
21. Feldman AG, Levin MF. Spatial control of reflexes, posture and movement in normal conditions and after neurological lesions. J Hum Kinet. 2016;52(1):21-34. CrossRef
22. Latash ML, Levin MF, Scholz JP, et al. Motor control theories and their applications. Medicina (Kaunas). 2010;46(6):382-392. CrossRef
23. Feldman AG. The nature of voluntary control of motor actions. In: Latash ML, Lestienne F, eds. Motor Control and Learning. Springer; 2006. CrossRef
24. Calota A, Levin MF. Tonic stretch reflex threshold as a measure of spasticity: implications for clinical practice. Top Stroke Rehabil. 2009;16(3):177-188. CrossRef
25. Levin MF, Piscitelli D, Khayat J. Tonic stretch reflex threshold as a measure of disordered motor control and spasticity - a critical review. Clin Neurophysiol. 2024;165:138-150. CrossRef
26. Elsman EBM, Mokkink LB, Terwee CB, et al. Guideline for reporting systematic reviews of outcome measurement instruments (OMIs): PRISMA-COSMIN for OMIs 2024. Qual Life Res. 2024;33(8):2029-2046. CrossRef
27. Terwee CB, Jansma EP, Riphagen II, et al. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res. 2009;18(8):1115-1123. CrossRef
28. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737-745. CrossRef
29. Lucas NP, Macaskill P, Irwig L, et al . The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). J Clin Epidemiol. 2010;63(8):854-861. CrossRef
30. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Qual Life Res. 2018;27(5):1171-1179. CrossRef

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

31. Fisher RA. 014: On the "probable error" of a coefficient of correlation deduced from a small sample. Metron. 1921;1:3-32.
32. IntHout J, Ioannidis JP, Rovers MM, et al. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open. 2016;6(7):e010247. CrossRef
33. Der Simonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-188. CrossRef
34. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539-1558. CrossRef
35. Calota A, Feldman AG, Levin MF. Spasticity measurement based on tonic stretch reflex threshold in stroke using a portable device. Clin Neurophysiol. 2008;119(10):2329-2337. CrossRef
36. Kim KS, Seo JH, Song CG. Portable measurement system for the objective evaluation of the spasticity of hemiplegic patients based on the tonic stretch reflex threshold. Med Eng Phys. 2011;33(1):62-69. CrossRef
37. Ferreira JF, Moreira V, Lima CS, et al. Development, test and validation of a mechatronic device for spasticity quantification. IJARS. 2013;10(5). CrossRef
38. Blanchette AK, Mullick AA, Moïn-Darbari K, et al. Tonic stretch reflex threshold as a measure of ankle plantar-flexor spasticity after stroke. Phys Ther. 2016;96(5):687-695. CrossRef
39. Turpin NA, Feldman AG, Levin MF. Stretch-reflex threshold modulation during active elbow movements in post-stroke survivors with spasticity. Clin Neurophysiol. 2017;128(10):1891-1897. CrossRef
40. Vieira D, Silva MB, Melo MC, et al. Effect of myofeedback on the threshold of the stretch reflex response of post-stroke spastic patients. Disabil Rehabil. 2017;39(5):458-467. CrossRef
41. Levin MF, Solomon JM, Shah A, et al. Activation of elbow extensors during passive stretch of flexors in patients with post-stroke spasticity. Clin Neurophysiol. 2018;129(10):2065-2074. CrossRef
42. Subramanian SK, Feldman AG, Levin MF. Spasticity may obscure motor learning ability after stroke. J Neurophysiol. 2018;119(1):5-20. CrossRef
43. Marques IA, Silva MB, Silva AN, et al. Measurement of post-stroke spasticity based on tonic stretch reflex threshold: implications of stretch velocity for clinical practice. Disabil Rehabil. 2019;41(2):219-225. CrossRef
44. Zhang X, Tang X, Zhu X, et al. A regression-based framework for quantitative assessment of muscle spasticity using combined emg and inertial data from wearable sensors. Front Neurosci. 2019;13:398. CrossRef
45. Alves CM, Rezende AR, Marques IA, et al. SpES: a new portable device for objective assessment of hypertonia in clinical practice. Comput Biol Med. 2021;134:104486. CrossRef
46. Frenkel-Toledo S, Solomon JM, Shah A, et al. Tonic stretch reflex threshold as a measure of spasticity after stroke: Reliability, minimal detectable change and responsiveness. Clin Neurophysiol. 2021;132(6):1226-1233. CrossRef
47. Wang C, Peng L, Hou ZG, et al. The assessment of upper-limb spasticity based on a multi-layer process using a portable measurement system. IEEE Trans Neural Syst Rehabil Eng. 2021;29:2242-2251. CrossRef
48. Frenkel-Toledo S, Levin MF, Berman S, et al. Shared and distinct voxel-based lesion-symptom mappings for spasticity and impaired movement in the hemiparetic upper limb. Sci Rep. 2022;12(1):10169. CrossRef
49. Marques IA, Alves CM, Rezende AR, et al. Virtual reality and serious game therapy for post-stroke individuals: a preliminary study with humanized rehabilitation approach protocol. Complement Ther Clin Pract. 2022;49:101681. CrossRef
50. Levin MF, Berman S, Weiss N, et al. ENHANCE proof-of-concept three-arm randomized trial: effects of reaching training of the hemiparetic upper limb restricted to the spasticity-free elbow range. Sci Rep. 2023;13(1):22934. CrossRef
51. Longo D, Santini G, Cherubini G, et al. The muscle shortening maneuver in individuals with stroke: a consideration-of-concept randomized pilot trial. Top Stroke Rehabil. 2023;30(8):807-819. CrossRef
52. Mullick AA, Musampa NK, Feldman AG, et al. Stretch reflex spatial threshold measure discriminates between spasticity and rigidity. Clin Neurophysiol. 2013;124(4):740-751. CrossRef
53. Fugl-Meyer AR, Jääskö L, Leyman I, et al. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. J Rehabil Med. 1975;7(1):13-31. CrossRef
54. Levin MF, Hui-Chan CW. Relief of hemiparetic spasticity by TENS is associated with improvement in reflex and voluntary motor functions. Evok Pot. 1992;85(2):131-142. CrossRef
55. He J, Luo A, Yu J, et al. Quantitative assessment of spasticity: a narrative review of novel approaches and technologies. Front Neurol. 2023;14:1121323. CrossRef
56. Gal O, Baude M, Deltombe T, et al. Clinical outcome assessments for spasticity: review, critique, and recommendations. Mov Disord. 2025;40(1):22-43. CrossRef
57. Fitzpatrick R, Davey C, Buxton MJ, et al. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess. 1998;2(14):i-74. CrossRef
58. Neumann DA. Kinesiology of the Musculoskeletal System. 3rd ed. Mosby; 2016.
59. Rome K, Cowieson F. A reliability study of the universal goniometer, fluid goniometer, and electrogoniometer for the measurement of ankle dorsiflexion. Foot Ankle Int. 1996;17(1):28-32. CrossRef
60. Bronner S, Agraharasamakulam S, Ojofeitimi S. Reliability and validity of electrogoniometry measurement of lower extremity movement. J Med Eng Technol. 2010;34(3):232-242. CrossRef
61. Goodwin J, Clark C, Deakes J, et al. Clinical methods of goniometry: a comparative study. Disabil Rehabil. 1992;14(1):10-15. CrossRef
62. Robertson DGE, Caldwell GE, Hamil J. Research Methods in Biomechanics. 2nd ed. Human Kinetics; 2013.
63. Das K, de Paula Oliveira T, Newell J. Comparison of markerless and marker-based motion capture systems using 95% functional limits of agreement in a linear mixed-effects modelling framework. Sci Rep. 2023;13(1):22880. CrossRef
64. Gaudet G, Raison M, Maso FD, et al. Intra- and intersession reliability of surface electromyography on muscles actuating the forearm during maximum voluntary contractions. J Appl Biomech. 2016;32(6):558-570. CrossRef
65. Lee J, Jung MY, Kim SH. Reliability of spike and turn variables of surface EMG during isometric voluntary contractions of the biceps brachii muscle. J Electromyogr Kinesiol. 2011;21(1):119-127. CrossRef
66. Carius D, Kugler P, Kuhwald HM, et al. Absolute and relative intrasession reliability of surface EMG variables for voluntary precise forearm movements. J Electromyogr Kinesiol. 2015;25(6):860-869. CrossRef
67. Howatson G, Glaister M, Brouner J, et al. The reliability of electromechanical delay and torque during isometric and concentric isokinetic contractions. J Electromyogr Kinesiol. 2009;19(5):975-979. CrossRef
68. Gupta A, Mudie KL, Clothier PJ. The reliability of determining the onset of medial gastrocnemius muscle activity during a stretch-shorten-cycle action. J Electromyogr Kinesiol. 2014;24(5):588-592. CrossRef
69. Gallina A, Pollock CL, Vieira TM, et al. Between-day reliability of triceps surae responses to standing perturbations in people

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

post-stroke and healthy controls: a high-density surface EMG investigation. Gait Posture. 2016;44:103-109. CrossRef
70. Cè E, Rampichini S, Limonta E, et al. Fatigue effects on the electromechanical delay components during the relaxation phase after isometric contraction. Acta Physiol (Oxf). 2014;211(1): 82-96. CrossRef
71. Silva MB, Silva AN, Naves EL, et al. An improved approach for measuring the tonic stretch reflex response of spastic muscles. Comput Biol Med. 2017;80:166-174. CrossRef
72. Sheean G, McGuire JR. Spastic hypertonia and movement disorders: pathophysiology, clinical presentation, and quantification. PM R. 2009;1(9):827-833. CrossRef
73. Marinelli L, Currà A, Trompetto C, et al. Spasticity and spastic dystonia: the two faces of velocity-dependent hypertonia. J Electromyogr Kinesiol. 2017;37:84-89. CrossRef
74. Trompetto C, Marinelli L, Mori L, et al. Pathophysiology of spasticity: implications for neurorehabilitation. BioMed Res Int. 2014;354906.
75. Raghavan P. Muscle physiology in spasticity and muscle stiffness. Toxicon. 2025;259:108350. CrossRef
76. Piscitelli D, Khayat J, Feldman AG, et al. Clinical relevance of the tonic stretch reflex threshold and μ as measures of upper limb spasticity and motor impairment after stroke. Neurorehabil Neural Repair. 2025;39(5):386-399. CrossRef
77. Ansari NN, Naghdi S, Arab TK, et al. The interrater and intrarater reliability of the Modified Ashworth Scale in the assessment of muscle spasticity: limb and muscle group effect. NeuroRehabilitation. 2008;23(3):231-237. CrossRef
78. Blackburn M, van Vliet P, Mockett SP. Reliability of measurements obtained with the modified Ashworth scale in the lower extremities of people with stroke. Phys Ther. 2002;82(1):25-34. CrossRef
79. Haas BM, Bergström E, Jamous A, et al. The inter rater reliability of the original and of the modified Ashworth scale for the assessment of spasticity in patients with spinal cord injury. Spinal Cord. 1996;34(9):560-564. CrossRef
80. Levin MF, Hui-Chan C. Are H and stretch reflexes in hemiparesis reproducible and correlated with spasticity? J Neurol. 1993;240(2):63-71. CrossRef
81. Ng SS, Hui-Chan CW. The timed up & go test: its reliability and association with lower-limb impairments and locomotor capacities in people with chronic stroke. Arch Phys Med Rehabil. 2005;86(8):1641-1647. CrossRef
82. Gladstone DJ, Danells CJ, Black SE. The fugl-meyer assessment of motor recovery after stroke: a critical review of its measurement properties. Neurorehabil Neural Repair. 2002;16(3): 232-240. CrossRef
83. Hochleitner I, Pellicciari L, Castagnoli C, et al. Intra- and inter-rater reliability of the Italian Fugl-Meyer assessment of upper and lower extremity. Disabil Rehabil. 2023;45(18):2989-2999. CrossRef
84. Morita H, Crone C, Christenhuis D, et al. Modulation of presynaptic inhibition and disynaptic reciprocal Ia inhibition during voluntary movement in spasticity. Brain. 2001;124(4):826-837. CrossRef
85. Nielsen JB, Crone C, Hultborn H. The spinal pathophysiology of spasticity--from a basic science point of view. Acta Physiol (Oxf). 2007;189(2):171-180. CrossRef
86. Knikou M. The H-reflex as a probe: pathways and pitfalls. J Neurosci Methods. 2008;171(1):1-12. CrossRef
87. Allison SC, Abraham LD, Allison SC. Correlation of quantitative measures with the modified Ashworth scale in the assessment of plantar flexor spasticity in patients with traumatic brain injury. J Neurol. 1995;242(10):699-706. CrossRef
88. Annaswamy T, Mallempati S, Allison SC, et al. Measurement of plantarflexor spasticity in traumatic brain injury: correlational study of resistance torque compared with the modified Ashworth scale. Am J Phys Med Rehabil. 2007;86(5):404-411. CrossRef
89. Husted JA, Cook RJ, Farewell VT, et al. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol. 2000;53(5):459-468. CrossRef
90. Beaton DE, Bombardier C, Katz JN, et al. Looking for important change/differences in studies of responsiveness. OMERACT MCID Working Group. Outcome Measures in Rheumatology. Minimal Clinically Important Difference. J Rheumatol. 2001; 28:400-405.

Arch Physioter 2025; 15: 265-276
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3314
ORIGINAL RESEARCH ARTICLE

Usability of Myosuit exosuit and effects of device-mediated rehabilitation on chronic stroke survivors: a non-randomized pilot study

Stefano Doronzio[1,2], Silvia Campagnini[1], Andrea Mannini[1], Chiara Basla[3], Tommaso Ciapetti[1], Michele Piazzini[1], Angela Politi[1], Chiara Pedrini[1], Erika Guolo[1], Serena Malloggi[1], Teresa Barretta[1], Robert Riener[3,4], Francesca Cecchi[1,2]

[1]IRCCS Fondazione Don Carlo Gnocchi, Firenze - Italy
[2]University of Florence, Department of Experimental and Clinical Medicine, Firenze - Italy
[3]Sensory-Motor Systems (SMS) Lab, Institute of Robotics and Intelligent Systems (IRIS), ETH Zürich, Zürich - Switzerland
[4]Spinal Cord Injury Center, Balgrist University Hospital, Medical Faculty, University of Zürich, Zürich - Switzerland

ABSTRACT

Objectives: The aim of this study is to assess the usability of the Myosuit within a chronic stroke survivors' rehabilitation program and to explore its therapeutic and assistive role on gait, stair negotiation, sit-to-stand transfers, and balance.
Methods: Ten chronic stroke survivors with gait impairments were enrolled. The System Usability Scale (SUS) was the primary outcome of the study; secondary outcomes were the Stroke Self-efficacy Questionnaire (SSEQ), the Short Physical Performance Battery, the 10-meter Walking Test (10mWT), the 2-minute Walking Test (2minWT), and the Stair Climbing Test. Tests were carried out before (T0) and after (T1) the training sessions, with and without the exoskeleton.
Results: The SUS rated poor-to-ok in 30% of the participants, good in 40%, and excellent to best imaginable in 30%. Comparing T1 vs T0, all the functional tests, except stair descending, showed statistically significant improvements without the exoskeleton, and SSEQ did not change significantly. T1 vs T0 comparisons with the exoskeleton showed improvements in all functional tasks, statistically significant for all, except for 2minWT and 10mWT.
Conclusions: This study confirmed the feasibility of a Myosuit-mediated treatment in a sample of chronic stroke survivors. Despite the usability of the wearable robot being generally positively perceived, it varied among users. Furthermore, the Myosuit exhibited both therapeutic and assistive potential in the sample.

Keywords: Exoskeleton, Exosuit, Neurorehabilitation, Stroke, Usability

What is already known?

- *Previous studies have demonstrated the potential of exoskeletons to improve mobility and function in individuals with gait impairments, including chronic stroke survivors. The System Usability Scale (SUS) is widely used to assess the usability of such devices, with prior research exploring both the therapeutic and assistive benefits of exoskeletons. While exoskeletons can provide immediate support during activities, their long-term impact on functional recovery has been an ongoing area of investigation.*

What this manuscript adds?

- *This study specifically evaluated the usability and role of the Myosuit, a soft wearable exoskeleton, in a rehabilitation program for chronic stroke survivors' therapy and assistance. It found that while the Myosuit had varied usability ratings, it demonstrated both therapeutic and assistive benefits. Participants showed significant functional improvements in walking, balance, and transfers after training, even without the exoskeleton, confirming its therapeutic potential. When using the Myosuit, most tasks also improved, though not all gains were statistically significant.*

Received: September 26, 2024
Accepted: September 30, 2025
Published online: October 24, 2025

Corresponding author:
Silvia Campagnini
email: scampagnin@ricres.org

Introduction

Stroke can affect motor abilities and performance in Activities of Daily Living (ADL) in individuals who survive the acute event (1–3). Physical consequences of stroke often include reduced mobility and balance, asymmetry of gait, spasticity, and muscle weakness in the more affected side (4).

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

It is estimated that at 3 months after stroke, about 70% of affected individuals walk at reduced speed and capacity (5). These factors contribute to reduced long-term physical activity and increased dependence on caregivers (4). In addition, the increased risk of falls and the energy cost of walking negatively affect mobility, participation, and, in the most severe cases, even the perception of personal identity (6, 7). Also, self-efficacy, i.e., the individual's confidence in enacting actions to achieve results (8), is positively associated with objective measures of mobility (9,10). Thus, it is possible to hypothesize that the use of an assistive robot, possibly improving ambulation performance, would also improve participants' self-efficacy concerning such activities.

Rehabilitation plays a key role in restoring autonomy in movement, helping participants reintegrate into society (10), and improving perceived self-efficacy (12). In this context, task-oriented and intensive programs have proven effective in motor recovery (13,14). In particular, robotic devices have been introduced into clinical practice to expand therapeutic approaches and support therapists in rehabilitation programs, showing promising results (15). Among robots designed for gait training, exoskeletons can assist participants during walking by transmitting mechanical power from body-worn actuators to the users' lower limbs (16,17). Some subtypes of such devices, such as exosuits, have been developed to provide the necessary assistance for a functional gait while stimulating users' active participation (16,18). Compared to rigid exoskeletons, exosuits are lightweight, minimally bulky, and do not impose biomechanical constraints. Because of these features, they are often considered both rehabilitative and assistive devices, i.e., they can be used for therapeutic purposes in clinical settings, and for assistive purposes in community settings (19,20).

Among lower limb exosuits, the Myosuit (MyoSwiss AG, Zürich) is designed to augment the motor performance of people with lower limb disability during walking, sit-to-stand and stand-to-sit transitions, and stairs negotiation (20). A previous study has shown Myosuit-mediated training to be safe and well-tolerated by a small, mixed population of subjects with lower limb motor disorders (19). Additionally, preliminary positive effects were observed on gait performance, increasing the average walking speed and endurance. However, although previous studies have shown promising results on the use of the device, they used heterogeneous methodologies and involved multiple clinical conditions (19,21), thereby warranting further investigation on stroke survivors in particular, to validate the usability and self-efficacy on a comprehensive set of mobility tasks. Additionally, no previous studies have evaluated the clinical outcomes of participants using the Myosuit for both rehabilitative and assistive purposes.

In light of these considerations, the primary objective of this trial was to verify the usability of the Myosuit in rehabilitation programs for a specific population: chronic (>6 months) stroke survivors. The secondary objectives of the study were to investigate potential changes in self-efficacy and to preliminarily explore the therapeutic and assistive role of the Myosuit on all motor tasks involving the lower limbs (i.e., walking, sit-to-stand and stand-to-sit transitions, stair ascent and descent, balance control), which have never been explored before. These preliminary results will constitute a determining element for the assessment of the feasibility of such a protocol for rehabilitative or assistive purposes.

Methods

Study design and setting

This study is a non-profit, non-controlled, interventional usability trial. The enrollment was carried out in the Neuromotor Rehabilitation Unit of the IRCCS Fondazione Don Carlo Gnocchi Onlus, Firenze, between May 2022 and September 2023. The study received approval from the local Ethics Committee under the number 20946_SPE in April 2022 and was registered on ClinicalTrails.gov under the number NCT05579197. The CONSORT reporting guidelines for pilot and feasibility trials were followed for this study reporting (22).

Participants

According to the eligibility criteria, researchers consecutively enrolled 10 individuals with chronic stroke at the outpatient neurorehabilitation clinic. Participants were informed of the study by a researcher, who provided their willingness to participate by signing the written informed consent form.

The inclusion criteria were the following:

- Chronic stroke (time from event >6 months);
- Age ≥18 years;
- Clinical stability was assessed by the administration of the Modified Early Warning Score scale (MEWS) (23);
- Being able to stand up from a chair without bending more than 45° to the right or left during movement;
- Being able to walk over 10 meters without assistance (the use of conventional aids other than knee orthoses was allowed, e.g., cane/ ankle-foot orthoses);
- Height between 150 cm and 195 cm;
- Weight between 45 kg and 110 kg;
- Functional Ambulation Category (FAC) ≥3 (24);
- Modified Ashworth Scale (MAS) ≤2 (25);
- Hospital Anxiety and Depression Scale (HADS) in the normal range (>10/21 for each subscale) (26);
- Montreal Cognitive Assessment (MoCA) >1 (using the equivalent score by Italian normative adjustments from Santangelo et al. (27));
- Signed informed consent.

The exclusion criteria were:

- Neurological and psychiatric conditions preceding the vascular event, active or with outcomes;
- Severe bilateral hearing loss or severe visual impairment;
- Functional Reach Test <15.24 cm, following normative value and user manual contraindication (28);
- Non-reducible retraction in flexion of the hip or knee greater than 10°;
- Varus or valgus of the knee greater than 10° (criteria required by the user manual of the device);
- Pregnancy;
- Previous stroke.

The device

The robotic device used was the Myosuit (20), a CE-marked medical exosuit to improve motor functions of the lower limbs in individuals affected by various neuromotor conditions (MyoSwiss AG, Zürich). The Myosuit resembles an exo-muscle providing active assistance to both hip and knee extension and supporting hip flexion via passive components. Its assistance is based on the analysis of inertial measurement units (IMUs)-derived kinetic and kinematic data to identify the synergistic sequences involved in supporting the body against gravity in ADL.

Specifically, through cables powered by two mechanical actuators located in the backpack, the Myosuit provides active assistance for hip and knee extension during the stance phase of gait. Additionally, two passive elastic bands located anteriorly to the hip support hip flexion in the swing phase. The overall weight of the system is 4.6 kg.

In this study, two assistive modes were used to tune the assistive profile of the hip and knee extension provided by the robot: the concentric and isometric modes. In both, the level of support could be adjusted for each leg independently on a 6-point scale, each level corresponding to a portion of the maximum available assistance (230 N), ranging from 0% to 100%, and with 20% intervals. The concentric mode assists the user when moving against gravity. For this mode, the user's locomotion data obtained from the IMUs is embedded in the device, where the support is synchronized with the kinematic profile of the task. The isometric mode provides a selected level of support by keeping the tension of the Myosuit cables constant, simulating an isotonic muscle contraction. Thus, the concentric mode is advised and was used for walking and stair climbing, while the isometric mode is for balance and stair descending. For sit-to-stand transitions, both modalities are contemplated and were tested in this study.

Lastly, the robot is equipped with strings connecting the middle third of the tibia to the toe to correct the foot drop. This feature was used according to a visual clinical assessment of the gait and the users' preferences.

Experimental protocol

Before starting the study, the interdisciplinary team conducted a series of unstructured tests on both healthy and post-stroke participants. These preliminary tests aimed to further train clinicians in using the device, ensure careful attention to wearability, and align the familiarization and tuning phases for selecting assistive parameters and modalities across different tasks.

During the experiments, the study participants underwent one *enrollment*, one *baseline assessment* (T0), seven *treatments*, and one *final assessment* (T1) session, for a total of 10 sessions for each participant (Figure 1). In the following subsections, details for each session are provided.

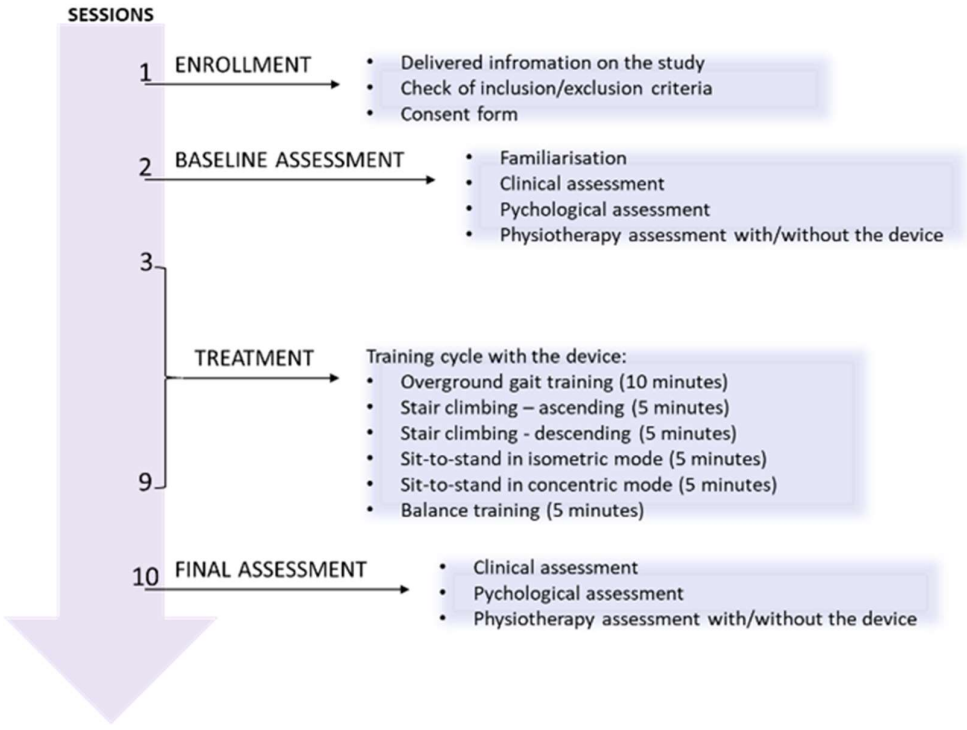


FIGURE 1 - Timeline of experiments. Details on the assessments are displayed in blue squares, and the treatment is in purple.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Session 1. Enrollment

In the enrollment session, a multidisciplinary team consisting of a physiatrist, a psychologist, and a physical therapist provided the study participants with all information on the investigation objectives and methods, verified that the eligibility criteria were met, and collected the signed consent form.

Both baseline and final assessment sessions were characterized by a comprehensive assessment of the participants, including physiotherapy, psychological, and clinical assessments.

The following tests were administered for the physiotherapy assessment (T0 and T1):

- Short Physical Performance Battery (SPPB) (29), a short lower-limb performance assessment composed of:
  - A Balance Test (BT) where a rater assesses the person's ability to stand unassisted without the use of a cane or walker in different conditions (i.e., Side-by-side, semi-tandem, and tandem stand);
  - The 4-meter Walking Test (4mWT) is used to assess the self-selected walking speed in meters per second over a short distance.
  - The 5-times Sit to Stand (5-tStS) test measures the time required to stand up and sit on a chair to quantify functional lower extremity strength during transitional movements.
- 10-meter Walking Test (10mWT) (30), to assess the maximal walking speed in meters per second over a short distance;
- 2-minute Walking Test (2minWT) (31), a measurement of endurance that assesses walking distance over two minutes;
- Stair Climbing Test (SCT) (32), a measure of the ability to ascend and descend a flight of stairs.

The following scales were administered for the psychological assessment:

- Stroke Self-Efficacy Questionnaire (8, 33), to assess stroke survivors' self-efficacy, intended as how confident the patient believes him/herself capable of organizing and executing the courses of action required to attain 13 tasks related to Activities of Daily Living despite the stroke (8);
- HADS, for the assessment of the depression and anxiety levels of hospitalized patients.
- MoCA, only at T0, is used to screen the global cognitive status of patients.

Lastly, both in T0 and T1, to complete the verification of eligibility criteria and to conduct a comprehensive multidimensional assessment, the clinical assessment included the following measures:

- MEWS is designed to identify early deteriorating patients with a composite score of physiological parameters to aid decision-making for escalating patient care;
- FAC, a 6-point assessment tool for functional walking, assessing the ambulation capacity of individuals by gauging the degree of human assistance needed during ambulation, regardless of whether they use a personal assistive device;
- MAS (34), to evaluate the level of spasticity of patients' lower limb muscle groups;
- MRC (35), to clinically assess the muscle strength on a 0-5 rating scale;
- Numerical Rating Scale (NRS) (36), to evaluate pain on a 0-10 analog scale;
- Modified Barthel Index (mBI) (37), for the assessment of the functional status of the patients;
- National Institutes of Health Stroke Society (NIHSS) (38), for an assessment of stroke severity.

At T0, after the assessment battery without the exoskeleton was completed and before carrying out the tests with the exoskeleton, each participant underwent 20 minutes of familiarization with the device.

This sub-session aimed to set the most appropriate torque profile for each of the motor tasks to be carried out in the protocol. Also, familiarization helped participants gain confidence with the robot and understand how to use it. This phase was carried out using a standardized approach developed by the clinical team. The intensity level of both concentric and isometric modalities of the robot was tuned, starting from symmetric configurations between the two legs and eventually switching to asymmetric ones based on the clinical features and comfort level of the users.

Sessions 3 to 9. Treatment

After the initial assessment, each participant performed seven circuit training sessions, wearing the Myosuit, 3 times per week. Each session was composed as follows: 10 minutes of overground gait training, 5 minutes of stair climbing, 5 minutes of stair descent, 5 minutes of sit-to-stand in isometric mode, 5 minutes of sit-to-stand in concentric mode, and 5 minutes of balance training. Moreover, a 5-minute break was taken between tasks to ensure the patient received the necessary rest, for a total of approximately 1 hour per session.

To avoid consistently training the same functions at the end of the session and introducing a potential bias from intra-session fatigue, the sequence of tasks was kept constant across the seven training sessions, but with a rotation: the first task in one session moved to the end in the following session (e.g., session one began with walking, session 2 with stair climbing, while walking was performed last). This procedure allowed each participant to train each motor task at different times during the training session. Finally, between tasks, participants could rest as much as needed based on subjective perception of exertion fatigue, generally around 5 minutes (Fig. 2).

Given participants' motor changes throughout the rehabilitation sessions, physical therapists were fine-tuning the levels of support based on visual analysis of motor tasks and users' feedback.

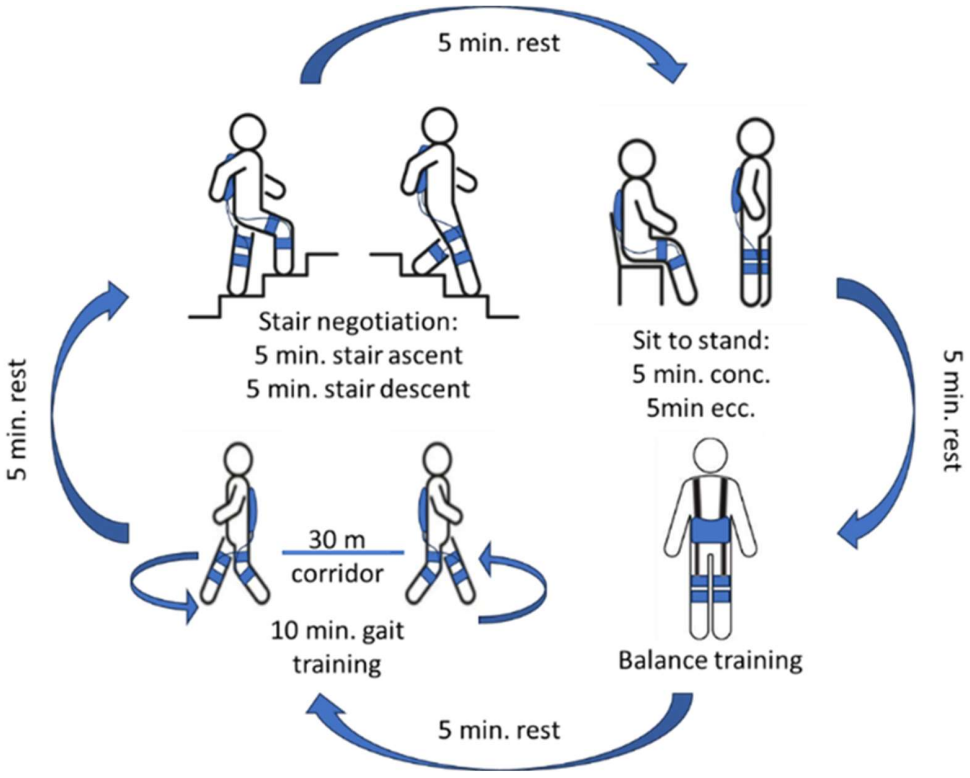


FIGURE 2 - Representation of the circuit performed in the training sessions.

Outcomes

The primary outcome was the System Usability Scale (SUS) (39), administered at T1, for the evaluation of the usability of the device once the protocol was completed. The SUS score ranges from 0 (worst) to 100 (best), and using such a score, we obtained the usability categories. These consisted of 6 different levels corresponding to different total score ranges of the SUS, namely: worst imaginable, SUS between 0 and 25; poor, SUS between 25.1 and 39; ok, SUS between 39.1 and 52; good, SUS between 52.1 and 73; excellent, SUS between 73.1 and 85; best imaginable, SUS between 85.1 and 100 rating (40). Secondary outcomes included the SSEQ, the SPPB score, the 2minWT (m), 10mWT (s), and the SCT (s).

Additionally, researchers recorded qualitative clinical notes regarding potential adverse events occurring during the trial. An adverse event was defined as any undesirable effect (e.g., red skin, muscle soreness) observed during or immediately after the treatment that appears to be caused by the device.

Data analysis

Statistical analyses were conducted using SPSS Statistics (IBM Corp. Released in 2021. IBM SPSS Statistics for Windows, Version 28.0. Armonk, NY: IBM Corp).

Concerning the primary objective, both the total score and the usability categories (40) were evaluated at T1 for the SUS. Furthermore, the median values of each SUS item were graphically investigated both overall and among subgroups generated by positive (SUS total score between 52.1 and 100) and negative (SUS total score between 0 and 52) SUS usability categories.

Concerning the secondary objectives, the measures of outcome considered were the SSEQ, the SPPB, the 10mWT, the 2mWT, and the descend and ascend SCT. For SPPB, the analysis was performed on both its total score and the score of each exercise. Additionally, metrics were obtained for the SPPB measures acquired with the device on, both in concentric and isometric modes.

Statistical analyses were focused on the description of the sample and comparison of both before/after the treatment and with/without the device for the main measures involved.

Descriptive statistics were performed through median (interquartile range, IQR) or absolute and percentage frequencies for continuous and categorical variables, respectively. Paired associations were performed using the Wilcoxon and McNemar tests for continuous and categorical variables, respectively. For all the tests, a p-value <0.05 was considered statistically significant. Given the numerosity of the sample, non-parametric tests were selected for analysis without checking for normality of the distributions.

Lastly, additional analyses involved the association between usability and baseline participants' characteristics. Specifically, Spearman's correlations were performed to investigate the linear association between baseline participants' characteristics and the SUS total score. Additionally, Mann-Whitney tests were conducted to investigate differences in baseline characteristics in the two groups determined by positive or negative SUS usability categories.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Results

Participants

During the trial, 24 outpatients with chronic stroke were consecutively screened for eligibility in the study. A total of 10 people were not enrolled in the study. Specifically, one refused to participate in the study. And the remaining nine did not fit the inclusion criteria. Of the remaining, four dropped out. Specifically, one was affected by a stable aortic dissection and showed signs of hypoperfusion and altered blood pressure response to physical exercise, even without the Myosuit, so clinicians decided to interrupt the protocol; one could not attend three weekly sessions because no family member could accompany him, and two withdrew consent for personal reasons unrelated to the intervention itself.

Participants were seven males and three females; six had an ischemic and four had a hemorrhagic stroke; six had a right-sided lesion, and four had a left-sided lesion; the median age (median (IQR)) was 61.5 (20.5); the median (IQR) time from the event was 36.5 (61) months.

Besides occasional discomfort, corrected by wearability adjustments, no side effects were experienced during the treatment. A description of the clinical, psychological, and physiotherapy characteristics of the sample is presented in Table 1.

Primary objective

Concerning the primary objective, the median (IQR) SUS total score was 55 (36.3). Its usability categories resulted in poor to ok by 30% of the participants, good by 40%, and excellent to best imaginable by 30%. By dividing patients into two groups with negative and positive SUS usability categories, two sub-samples of 3 and 7 participants were obtained,

TABLE 1 - Description and paired association (T0-T1) of the main clinical, psychological, and physiotherapy characteristics of the sample

Variables	Baseline assessmentsession (T0)		Final assessmentsession (T1)		p-value
	Median (IQR) or Frequency (%)	N	Median (IQR) or Frequency (%)	N	
Demographics and stroke history					
Age	61.5 (20.5)	10	–	–	–
Time from event (months)	36.5 (61.0)	10	–	–	–
Sex	Male: 7 (70%) Female: 3 (30%)	10	–	–	–
Education (years)	9.0 (8.0)	10	–	–	–
Aetiology	Ischaemic: 6 (60%) Hemorrhagic: 4 (40%)	10	–	–	–
Side of the lesion	Right: 6 (60%) Left: 4 (40%) Bilateral: 0 (0%)	10	–	–	–
Clinical assessment					
NIHSS	3.0 (4.0)	10	3.0 (3.0)	10	0.279
mBI	96.0 (7.0)	10	98.0 (5.0)	10	0.168
MAS_ LL	3.0 (3.0)	8	2.5 (2.0)	10	0.221
FAC	5.0 (1.0)	8	4.0 (1.0)	10	0.564
NRS	0.0 (0.0)	8	0.0 (2.0)	10	0.180
MoCA (adjusted score)	25.5 (2.4)	10	–	–	–
HADS_ depression	2.5 (4.0)	10	2.5 (5.0)	10	0.582
HADS_ anxiety	3.0 (7.0)	10	4.0 (6.0)	10	0.359
SSEQ	28.0 (10.0)	10	29.5 (17.0)	10	0.592
SUS	–	–	55 (36.3)	10	–
SUS usability categories	–	–	Worst imaginable: 0 (0%) Poor. 1 (10%) Ok: 2 (20%) Good: 4 (40%) Excellent: 2 (20%) Best imaginable: 1 (10%)	10	–

Abbreviations: FAC: Functional Ambulation Category, HADS: Hospital Anxiety Depression Scale, IQR: Interquartile Range, MAS_LL: Modified Ashworth Scale_ Lower Limbs, mBI: Modified Barthel Index, MoCA: Montreal Cognitive Assessment, NIHSS: National Institute of Health Stroke Scale, NRS: Numerical Rating Scale, SSEQ: Stroke Self-Efficacy Questionnaire, SUS: System Usability Scale.

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

respectively. Figure 3 reports the median SUS item values over the 10 participants and the two subsamples of 7 with positive ratings and 3 with negative ratings. Item 9, "*I felt very confident using the system*", revealed very positive ratings from both groups and thus overall. On the contrary, item 4, "*I think I would need the support of a technical person to be able to use this system,*" received overall negative ratings. Conflicting ratings were especially encountered for items 1, "*I think that I would like to use the system frequently*", and 10, "*I needed to learn a lot of things before I could get going with this system*", with positive results for the group with positive SUS usability categories.

Secondary objectives

Clinical assessment demonstrated no statistically significant changes when comparing T1 vs T0 without the exoskeleton; on the other hand, all the functional tests, except stair descending, showed statistically significant improvements. The stair descending component of the SCT showed a statistically significant longer time after the training (median (IQR); 7.6 s (6.7) to 8.0 s (5.8); p = 0.022). Also, T1 vs T0 comparisons with the exoskeleton showed statistically significant differences in all functional tasks, except for the 2minWT and the 10mWT. In detail, the SPPB in concentric mode increased from 8.0 (3.0) to 9.5 (4.0), p = 0.037; the SPPB in isometric mode from 8.0 (3.0) to 9.5 (4.0), p = 0.024; the 2minWT (m) from 99 (65.8) to 109 (65.1), p = 0.028; the 10mWT (s) from 10.8 (8.9) to 10.9 (8.5), p = 0.028; the SCT up from 8.3 (5.4) to 7.6 (3.3), p = 0.028; the SCT down from 8.3 (7.6) to 7.3 (5.3), p = 0.022. The assistive effect is not supported by statistically significant changes when comparing with and without the exoskeleton, both in T0 and T1. All changes across conditions and evaluation time points in functional tasks are reported in Table 2.

Correlation analyses between participants' usability and baseline characteristics resulted in non-significant associations for all the variables (Table 3).

Discussion

In the current study, we showed that an exosuit-mediated treatment for chronic stroke survivors, specifically using the Myosuit device, was usable in an outpatient rehabilitation setting, with no reported adverse events. Despite mostly positive usability scores, the exosuit was found to be differently usable from one user to another, ranging from poor to best imaginable across study participants. In detail, seven participants rated the Myosuit higher than the usability threshold of 68 points. These values are in line with the results of a previous study investigating the usability of the Myosuit, where five out of seven participants scored in the SUS questionnaire more than the good usability threshold (41). In their study, Basla et al. discussed the possibility of a selection bias favoring a positive outcome on the SUS due to the high interest in the technology of the sample. This bias is less impactful in the current study, where, out of 10 participants, only two were previously included in robotics trials. Furthermore, the participants in our study did not purchase the exoskeleton directly but were contacted by the research team and involved in the

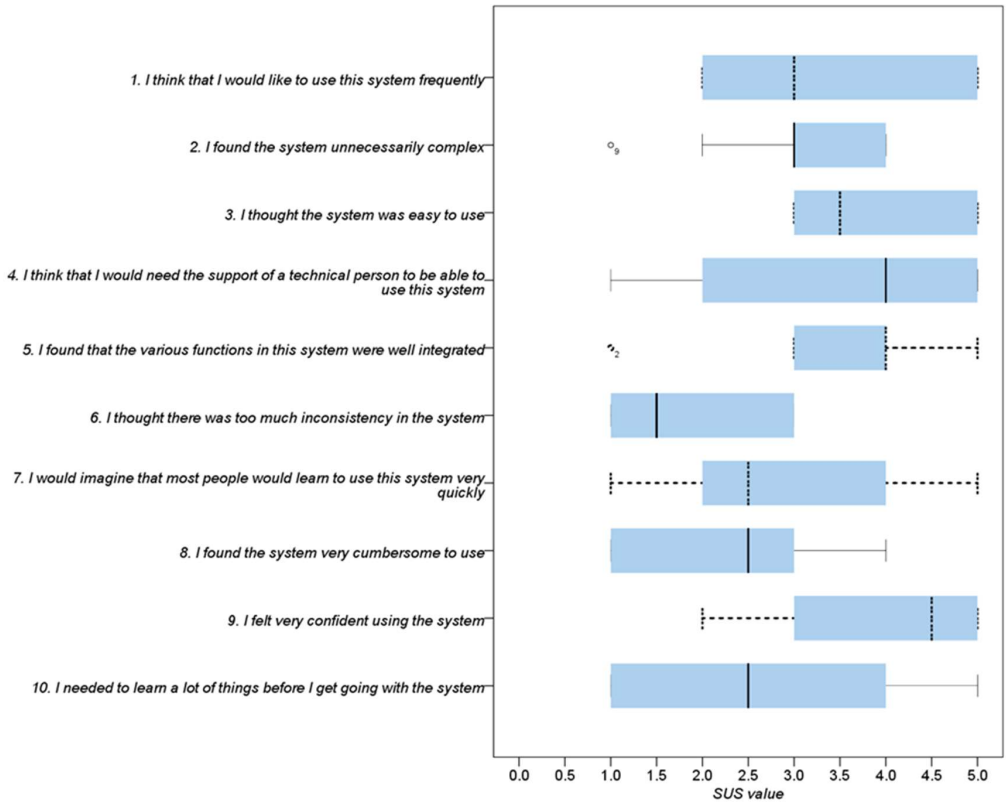


FIGURE 3 - Median scores of individual items on the System Usability Scale (SUS).

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - Functional task scores in T0 and T1, without and with the exoskeleton

Variables	Baseline assessment session (T0)			Final assessment session (T1)			p-value without (T0/T1)	p-value with (T0/T1)
	Without device	With device	p-value baseline (with/without)	Without device	With device	p-value final (with/without)		
SPPB _ concentric mode	7.0 (2.0)	7.5 (3.0)	0.655	8.5 (4.0)	9.0 (4.0)	0.317	**0.016**	**0.037**
SPPB _ isometric mode	7.0 (2.0)	8.0 (3.0)	0.317	8.5 (4.0)	9.5 (4.0)	0.458	**0.016**	**0.024**
BT	0: 0 (0%) 1: 0 (0%) 2: 2 (20%) 3: 4 (40%) 4: 4 (40%)	0: 0 (0%) 1: 0 (0%) 2: 2 (20%) 3: 3 (30%) 4: 5 (50%)	0.317	0: 0 (0%) 1: 0 (0%) 2: 1 (10%) 3: 3 (30%) 4: 6 (60%)	0: 0 (0%) 1: 0 (0%) 2: 0 (0%) 3: 2 (20%) 4: 8 (80%)	0.083	0.257	0.102
4mWT	0: 0 (0%) 1: 2 (20%) 2: 1 (10%) 3: 5 (50%) 4: 2 (20%)	0: 0 (0%) 1: 2 (20%) 2: 1 (10%) 3: 4 (40%) 4: 3 (30%)	0.564	0: 0 (0%) 1: 2(20%) 2: 1 (10%) 3: 4 (40%) 4: 3 (30%)	0: 0 (0%) 1: 2 (20%) 2: 1 (10%) 3: 2 (20%) 4: 5 (50%)	0.317	0.314	0.157
5-tStS _isometric mode	0: 0 (0%) 1: 6 (60%) 2: 3 (30%) 3: 1 (10%) 4: 0 (0%)	0: 0 (0%) 1: 6 (60%) 2: 2 (20%) 3: 2 (20%) 4: 0 (0%)	0.655	0: 0 (0%) 1: 2 (20%) 2: 2 (20%) 3: 6 (60%) 4: 0 (0%)	0: 0 (0%) 1: 4 (40%) 2: 2 (20%) 3: 3 (30%) 4: 1 (10%)	0.180	**0.014**	0.129
5-tStS _ concentric mode	0: 0 (0%) 1: 6 (60%) 2: 3 (30%) 3: 1 (10%) 4: 0 (0%)	0: 0 (0%) 1: 7 (70%) 2: 2 (20%) 3: 1 (10%) 4: 0 (0%)	0.655	0: 0 (0%) 1: 2 (20%) 2: 2 (20%) 3: 6 (60%) 4: 0 (0%)	0: 0 (0%) 1: 6 (60%) 2: 2 (20%) 3: 2 (20%) 4: 0 (0%)	**0.023**	**0.014**	0.480
2minWT (m)	100 (59.7)	99 (65.8)	0.508	111 (56.7)	109 (65.1)	0.683	**0.028**	0.203
10mWT (s)	10.9 (6.8)	10.8 (8.9)	0.799	10.1 (6.1)	10.9 (8.5)	0.114	**0.028**	0.074
SCT_up	8.4 (4.7)	8.3 (5.4)	0.359	8.2 (3.7)	7.6 (3.3)	0.646	**0.028**	**0.028**
SCT_down	7.6 (6.7)	8.3 (7.6)	0.508	8.0 (5.8)	7.3 (5.3)	0.878	**0.022**	**0.037**

Abbreviations: 10mWT: Ten Meters Walking Test, 2minWT: Two Minutes Walking Test, 4mWT: Four Meters Walking Test, 5-tStS: 5-times Sit to Stand, BT: Balance Test, SCT: Stair Climb Test, SPPB: Short Physical Performance Battery.

project. A difference to be highlighted between the study by Basla et al. and ours concerns the eligibility criteria; in the previous study, they were broader, including individuals with a variety of diseases with functional gait limitations, whereas our results are focused on a relatively larger sample of chronic stroke survivors.

In our study, the median SUS score was found to be 55, which corresponds to good usability categories overall, indicating that the system is usable, though it can possibly be improved to reach better usability rates. The item that received the most negative feedback was item 6, '*I thought there was too much inconsistency in the system*,' which had a median score of 1.5 (Figure 3). The low score for this item indicates that users perceive a lack of consistency in the Myosuit system, which may cause frustration and reduced trust in the device. To improve this, efforts could be focused on enhancing the user interface to provide clearer feedback and more predictable responses. Streamlining the transition between different modes of assistance and reducing the variability in system behavior could help users feel more in control and ultimately increase the overall usability and user satisfaction in this aspect. One notable aspect is the IQR of 36.3, reflecting high variability in the responses. While some participants found the device usable, others encountered difficulties. This disparity is further highlighted by the substantial difference between the minimum (37.5) and maximum scores (92.5), underscoring the varying perceptions of usability among the participants.

In our sample, the correlation analysis did not show any significant association between participants' features and the SUS total score. Although descriptive differences were noted among patients—specifically, the three participants who provided an overall positive SUS score were older, more

TABLE 3 - Correlation analysis between SUS total score and baseline participants' characteristics

Variables	Correlation coefficient	p-value
Age	0.298	0.404
Time from the event (months)	0.262	0.464
Education (years)	0.440	0.203
MoCA (adjusted score)	−0.189	0.601
SSEQ- (T0)	−0.328	0.355
NIHSS	−0.227	0.528
mBI	−0.040	0.912
SPPB_ without device	−0.032	0.931
SPPB_with device_isometric mode	0.032	0.931
SPPB_with device_concentric mode	−0.056	0.879
2minWT_without device (m)	0.250	0.486
2minWT_with device (m)	0.354	0.316
10mWT_without device (s)	−0.140	0.699
10mWT_with device (s)	−0.213	0.554

Abbreviations: 10mWT: Ten Meters Walking Test, 2minWT: Two Minutes Walking Test, mBI: modified Barthel index, SSEQ: Stroke Self Efficacy Questionnaire, MoCA: Montreal Cognitive Assessment, NIHSS: National Institute of Stroke Society, SPPB: Short Physical Performance Battery, SUS: System Usability Scale.

educated, and presented a lower SSEQ compared to those with negative ratings—these observations were not subjected to statistical analyses. Given the limited literature on the relationship between SUS scores and patient characteristics, as well as the limited sample size of this pilot study, we encourage future studies with larger samples to rigorously investigate these potential relationships. The association between psychological and social factors and functional performance following stroke has been confirmed in previous studies (42), and the role of self-efficacy as a movement recovery moderator is acknowledged (43). In our study, overall self-efficacy was very low compared to a previous study by Jones et al. (8); even if we excluded persons experiencing moderate to severe depressive symptoms, differences both in participants' psychological profiles and objective performances may explain this discrepancy. Unfortunately, in the previous article, the functional profile of the participants was not assessed, thus impeding an in-depth comparison with our results. Assessing self-efficacy is also relevant in rehabilitation studies to verify whether the clinical benefit is directly due to the experimental treatment (44) or whether self-efficacy changes play a relevant role in the outcome. In this pilot, we did not see a statistically significant improvement in the SSEQ after the treatment. This is well explained by the size of the reported functional improvements, which are significant but small and have little impact on the tasks explored by the SSEQ, particularly mobility. On the other hand, this result suggests a direct effect of the Myosuit training on measured functional outcomes, independent from potential changes in patient motivation or self-confidence.

Concerning secondary outcomes, comparing T1 vs T0, we observed an increase in performance in almost all functional tasks conducted without the device (Table 2), suggesting a therapeutic effect of the rehabilitation training with Myosuit. The SPPB increased after the treatment without the exoskeleton by a mean (standard deviation, std) of 1.3 (1.1) points. This improvement, despite its relatively large dispersion, exceeded most of the values of the Standard Error of Measurement (SEM = 0.68-1.42), the Minimal Detectable Change (MDC = 0.54-2.9), and the Minimal Clinically Important Difference (MCID = 1) reported in previous studies (45,46). However, the significance of our findings remains to be clarified, given the limited sample size and the fact that the psychometric properties of SPPB are not validated on a population of stroke survivors but on community-dwelling older adults (46,47). Moreover, it is worth mentioning that when looking at the differences between the SPPB items, only the sit-to-stand and stand-to-sit show statistically significant improvements, meaning that the improvement in the total test score is dependent on the changes observed in the 5-tStS and not on the 4mWT and BT. Regarding the 2minWT, our study shows a statistically significant increase with a mean (std) of 9.0 (10.0) meters in the T1-T0 comparison without the exoskeleton, getting very close to the MDC95% for the 2minWT, found by Hiengkaew et al. in 2012 in 61 chronic stroke survivors, equal to 13.4 m or 23% change from baseline (48). These results need insights into the future, adequately powered randomized controlled studies, to determine if the recorded positive changes can be attributed to this device specifically.

When considering the therapeutic effects in the 10mWT, our results showed a mean (std) improvement of 0.09 (0.12) m/s without the exosuit at T1 compared to T0. This improvement is lower than the substantial meaningful changes reported by Perera et al., 2006 (47) (0.14 m/s) and by Tilson et al, 2010 (49) (0.16 m/s). One aspect that has potentially affected the motor recovery of training is the limited number of treatment sessions, as literature proposes longer treatments for chronic stroke (50).

Descending stairs was the only ambulatory task that did not show a statistically significant improvement at T1 without the exoskeleton. This could be due to the device applying constant torque throughout the entire task of descending the stairs. Doing so provided support in bearing the weight of the body during the stance phase but restricted the movements of the wearer during the swing phase. A dedicated modality for stair descending was not available on the device (the isometric modality is currently used), but it could be desirable.

Regarding the comparison of T1 vs T0 with the exoskeleton, significant differences were observed in SPPB scores in both concentric and isometric modes, as well as in SCT_up and SCT_down. In all these tasks, the observed effect size remains borderline in terms of clinical significance. The other tasks that without exoskeleton showed a difference between times, the 5-tStS (both isometric and concentric), the 2minWT, and the 10mWT, do not show statistical significance with the exoskeleton. This could be attributable to the weight of the device masking the induced rehabilitation

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

effects. Regarding assistive effects, the baseline comparison between conditions (exo vs no-exo) shows no statistically significant difference. However, at the post-treatment evaluation, a significant difference appears in the 5-tStS concentric mode, likely due to the improved human-machine interaction over time.

In relation to these results, it is worth highlighting that the current study was a usability study, with a limited sample size and no control group. However, the positive preliminary results we found on some secondary outcomes, particularly walking speed, seem promising and deserve to be investigated in future power-sized controlled clinical trials. Although our sample included only chronic stroke individuals with inclusion criteria limiting clinical variability, we cannot ensure the exclusion of strokes (e.g., certain pontine or midbrain strokes) that may cause lower limb weakness and balance impairment as separate dysfunctions. Furthermore, the order of execution of the functional tests might lead to some limitations. Indeed, these were always carried out without the exoskeleton before, potentially introducing a fatigue bias in the advantage of the without-exosuit condition. Also, we tested the Myosuit only in an outpatient-like research setting. Although our protocol tried to emulate everyday life scenarios, testing the device in a laboratory environment is a major limitation to an ecological investigation of the robot's assistive potential. Lastly, this study lacks T1 data for dropouts, potentially introducing a selection bias and limiting longitudinal analyses, both for the SUS and clinical effects. Even if more details are missing, future clinical trial designs should consider the dropout rate obtained in this pilot trial.

In light of these limitations, we can conclude that the Myosuit device, in the selected cohort of post-stroke individuals, showed encouraging results in terms of potential therapeutic applicability. These results are promising for the potential further applicability of this device since previous studies have primarily highlighted its assistive benefits.

Conclusions

The results of this study support the usability and feasibility of a Myosuit-mediated rehabilitation program and present no factors that would hinder the repeatability of the treatment on a larger sample of stroke survivors. Also, the Myosuit has shown preliminary evidence of effectiveness, but the extent of the observed clinical improvement remains uncertain.

Future investigations should then consider the inclusion of a larger sample and a controlled study design to demonstrate the role of the experimental treatment against evidence-supported therapy. Ultimately, as the study did not include subacute stroke survivors, exploring the usability of the Myosuit in an inpatient setting at an earlier stage of the disease remains a priority for understanding the rehabilitation potential of the robot.

Acknowledgements

We would like to thank all patients who made themselves available for the study; the MyoSwiss team for the support provided during the investigation; Benedetta Basagni, Alessandro Maselli, Mathias Bossaerts, and Leonardo Pellicciari for their valuable role in the early phases of the study.

Disclosures

Conflict of interest: RR is the author of the European patent application 16207252.4 filed by ETH Zürich and licensed by MyoSwiss AG and is a minor shareholder at MyoSwiss AG. All other authors have no competing interests.

Ethics approval and consent to participate: The study was approved by the local Ethics Committee (20946_SPE, April 2022) and registered on ClinicalTrails.gov (NCT05579197). Written informed consent was obtained from all subjects before the study.

Financial support: Ricerca Corrente – Italian Ministry of Health.

Authors' contributions: FC, RR, SC, AM, SD, and CB were involved in the study design; TB and EG contributed to the clinical protocol design; SD, TC, and MP participated in the physiotherapy protocol design; CP and SM were involved in the psychological protocol design; SC performed statistical analyses; SD edited the manuscript. All authors were involved in the revision of the manuscript.

Data availability statement: Pseudonymized data will be available, upon request to the corresponding author, for research purposes.

References

1. Adams HP Jr, Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke*. 1993;24(1):35-41. CrossRef PubMed
2. Lai SM, Studenski S, Duncan PW, et al. Persisting consequences of stroke measured by the Stroke Impact Scale. *Stroke*. 2002;33(7):1840-1844. CrossRef PubMed
3. Sveen U, Bautz-Holter E, Sødring KM, et al. Association between impairments, self-care ability and social activities 1 year after stroke. *Disabil Rehabil*. 1999;21(8):372-377. CrossRef PubMed
4. Farrell JW III, Merkas J, Pilutti LA. The effect of exercise training on gait, balance, and physical fitness asymmetries in persons with chronic neurological conditions: a systematic review of randomized controlled trials. *Front Physiol*. 2020;11:585765. CrossRef PubMed
5. Jørgensen HS, Nakayama H, Raaschou HO, et al. Recovery of walking function in stroke patients: the Copenhagen Stroke Study. *Arch Phys Med Rehabil*. 1995;76(1):27-32. CrossRef PubMed
6. Opara JA, Jaracz K. Quality of life of post-stroke patients and their caregivers. *J Med Life*. 2010;3(3):216-220. PubMed
7. Vestling M, Ramel E, Iwarsson S. Quality of life after stroke: well-being, life satisfaction, and subjective aspects of work. *Scand J Occup Ther*. 2005;12(2):89-95. CrossRef PubMed
8. Jones F, Partridge C, Reid F. The Stroke Self-Efficacy Questionnaire: measuring individual confidence in functional performance after stroke. *J Clin Nurs*. 2008;17(7B):244-252. CrossRef PubMed
9. Brands I, Köhler S, Stapert S, et al. Influence of self-efficacy and coping on quality of life and social participation after acquired brain injury: a 1-year follow-up study. *Arch Phys Med Rehabil*. 2014;95(12):2327-2334. CrossRef PubMed

10. Torrisi M, De Cola MC, Buda A, et al. Self-efficacy, poststroke depression, and rehabilitation outcomes: is there a correlation? *J Stroke Cerebrovasc Dis*. 2018;27(11):3208-3211. CrossRef PubMed
11. Tedla JS, Dixit S, Gular K, et al. Robotic-assisted gait training effect on function and gait speed in subacute and chronic stroke population: a systematic review and meta-analysis of randomized controlled trials. *Eur Neurol*. 2019;81(3-4): 103-111. CrossRef PubMed
12. Best KL, Miller WC, Huston G, et al. Pilot study of a peer-led wheelchair training program to improve self-efficacy using a manual wheelchair: a randomized controlled trial. *Arch Phys Med Rehabil*. 2016;97(1):37-44. CrossRef PubMed
13. French B, Thomas LH, Leathley MJ, et al. Repetitive task training for improving functional ability after stroke. *Cochrane Database Syst Rev*. 2007;(4):CD006073. PubMed
14. Perry SB. Stroke rehabilitation: guidelines for exercise and training to optimize motor skill. *J Neurol Phys Ther*. 2004;28(2):101. CrossRef
15. Mehrholz J, Thomas S, Werner C, et al. Electromechanical-assisted training for walking after stroke. *Cochrane Database Syst Rev*. 2017;5(5):CD006185. PubMed
16. Awad LN, Bae J, O'Donnell K, et al. A soft robotic exosuit improves walking in patients after stroke. *Sci Transl Med*. 2017;9(400):eaai9084. CrossRef PubMed
17. Louie DR, Eng JJ. Powered robotic exoskeletons in post-stroke rehabilitation of gait: a scoping review. *J Neuroeng Rehabil*. 2016;13(1):53. CrossRef PubMed
18. Lee HJ, Lee SH, Seo K, et al. Training for walking efficiency with a wearable hip-assist robot in patients with stroke: a pilot randomized controlled trial. *Stroke*. 2019;50(12):3545-3552. CrossRef PubMed
19. Haufe FL, Schmidt K, Duarte JE, et al. Activity-based training with the Myosuit: a safety and feasibility study across diverse gait disorders. *J Neuroeng Rehabil*. 2020;17(1):135. CrossRef PubMed
20. Schmidt K, Duarte JE, Grimmer M, et al. The myosuit: bi-articular anti-gravity exosuit that reduces hip extensor activity in sitting transfers. *Front Neurorobot*. 2017;11:57. CrossRef PubMed
21. Just IA, Fries D, Loewe S, et al. Movement therapy in advanced heart failure assisted by a lightweight wearable robot: a feasibility pilot study. *ESC Heart Fail*. 2022;9(3):1643-1650. CrossRef PubMed
22. Eldridge SM, Chan CL, Campbell MJ, et al.; PAFS consensus group. CONSORT 2010 statement: extension to randomized pilot and feasibility trials. *BMJ*. 2016;355:i5239. CrossRef PubMed
23. Subbe CP, Kruger M, Rutherford P, et al. Validation of a modified Early Warning Score in medical admissions. *QJM*. 2001;94(10):521-526. CrossRef PubMed
24. Collen FM, Wade DT, Bradshaw CM. Mobility after stroke: reliability of measures of impairment and disability. *Int Disabil Stud*. 1990;12(1):6-9. CrossRef PubMed
25. Harb A, Kishner S. *Modified Ashworth Scale. StatPearls.* (Internet) StatPearls Publishing; 2024, Available from Online Accessed February 13, 2024.
26. Breeman S, Cotton S, Fielding S, et al. Normative data for the Hospital Anxiety and Depression Scale. *Qual Life Res*. 2015;24(2):391-398. CrossRef PubMed
27. Santangelo G, Siciliano M, Pedone R, et al. Normative data for the Montreal Cognitive Assessment in an Italian population sample. *Neurol Sci*. 2015;36(4):585-591. CrossRef PubMed
28. Acar M, Karatas GK. The effect of arm sling on balance in patients with hemiplegia. *Gait Posture*. 2010;32(4):641-644. CrossRef PubMed
29. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol*. 1994;49(2):M85-M94. CrossRef PubMed
30. Cheng DK, Nelson M, Brooks D, et al. Validation of stroke-specific protocols for the 10-meter walk test and 6-minute walk test conducted using 15-meter and 30-meter walkways. *Top Stroke Rehabil*. 2020;27(4):251-261. CrossRef PubMed
31. Kosak M, Smith T. Comparison of the 2-, 6-, and 12-minute walk tests in patients with stroke. *J Rehabil Res Dev*. 2005;42(1):103-107. CrossRef PubMed
32. Modai G, Sharon B, Bar-Haim S, et al. Stair climbing test post-stroke: feasibility, convergent validity and metabolic, cardiac, and respiratory responses. *Top Stroke Rehabil*. 2015;22(4):281-288. CrossRef PubMed
33. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. *J Usability Stud*. 2009;4:114-123.
34. Morris S. Ashworth And Tardieu Scales: their clinical relevance for measuring spasticity in adult and paediatric neurological populations. *Phys Ther Rev*. 2002;7(1):53-62. CrossRef
35. Vanhoutte EK, Faber CG, van Nes SI, et al.; PeriNomS Study Group. Modifying the Medical Research Council grading system through Rasch analyses. *Brain*. 2012;135(Pt 5):1639-1649. CrossRef PubMed
36. Krebs EE, Carey TS, Weinberger M. Accuracy of the pain numeric rating scale as a screening test in primary care. *J Gen Intern Med*. 2007;22(10):1453-1458. CrossRef PubMed
37. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. *J Clin Epidemiol*. 1989;42(8):703-709. CrossRef PubMed
38. Kwah LK, Diong J. National Institutes of Health Stroke Scale (NIHSS). *J Physiother*. 2014;60(1):61. CrossRef PubMed
39. Brooke J. SUS: A quick and dirty usability scale. Usability Eval Ind. 1995;189.
40. Bangor A, Kortum PTP, Miller, TJ. An empirical evaluation of the system usability scale. *Int J Hum Comput Interact*. 2008;24:574. CrossRef
41. Basla C, Hungerbühler I, Meyer JT, et al. Usability of an exosuit in domestic and community environments. *J Neuroeng Rehabil*. 2022;19(1):131. CrossRef PubMed
42. Robinson-Smith G. Self-efficacy and quality of life after stroke. *J Neurosci Nurs*. 2002;34(2):91-98. CrossRef
43. Ma HI, Hung PH, Lin SH, et al. Role of self-efficacy in the predictive relationship of motor ability to functional performance after task-related training in stroke: a secondary analysis of longitudinal data. *Arch Phys Med Rehabil*. 2021;102(8):1588-1594. CrossRef PubMed
44. Young CA, Mills RJ, Woolmore J, et al. The unidimensional self-efficacy scale for MS (USE-MS): developing a patient based and patient reported outcome. *Mult Scler*. 2012;18(9):1326-1333. CrossRef PubMed
45. Mangione KK, Craik RL, McCormick AA, et al. Detectable changes in physical performance measures in elderly African Americans. *Phys Ther*. 2010;90(6):921-927. CrossRef PubMed
46. Olsen CF, Bergland A. "Reliability of the Norwegian version of the short physical performance battery in older people with and without dementia". *BMC Geriatr*. 2017;17(1):124. CrossRef PubMed
47. Perera S, Mody SH, Woodman RC, et al. Meaningful change and responsiveness in common physical performance measures in older adults. *J Am Geriatr Soc*. 2006;54(5):743-749. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

48. Hiengkaew V, Jitaree K, Chaiyawat P. Minimal detectable changes of the Berg Balance Scale, Fugl-Meyer Assessment Scale, Timed "Up & Go" Test, gait speeds, and 2-minute walk test in individuals with chronic stroke with different degrees of ankle plantarflexor tone. *Arch Phys Med Rehabil*. 2012;93(7):1201-1208. CrossRef PubMed
49. Tilson JK, Sullivan KJ, Cen SY, et al.; Locomotor Experience Applied Post Stroke (LEAPS) Investigative Team. Meaningful gait speed improvement during the first 60 days poststroke: minimal clinically important difference. *Phys Ther*. 2010;90(2):196-208. CrossRef PubMed
50. Clark B, Whitall J, Kwakkel G, et al. The effect of time spent in rehabilitation on activity limitation and impairment after stroke. *Cochrane Database Syst Rev*. 2021;10(10):CD012612. PubMed

Arch Physioter 2025; 15: 277-285
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3456
ORIGINAL RESEARCH ARTICLE

Development and feasibility testing of a video-based exercise program for improving upper limb motor function among stroke survivors in a low-income setting

Hafsat Maina Ali[1,2], Surajo Kamilu Sulaiman[3], Jibrin Sammani Usman[1], Muhammad Aliyu Abba[1], Aminu Alhassan Ibrahim[3], Dauda Salihu[4], Auwal Abdullahi[1], Umar Muhammad Bello[5]

[1]Department of Physiotherapy, Faculty of Allied Health Sciences, College of Health Sciences, Bayero University Kano, Kano - Nigeria
[2]Department of Physiotherapy, Federal Medical Centre Jabi, Abuja - Nigeria
[3]Physiotherapy Department, Faculty of Applied Science, Tishk International University, Erbil, Kurdistan Region - Iraq
[4]Department of Psychiatric, Mental Health & Community Health, College of Nursing, Qassim University, Buraydah - Saudi Arabia
[5]Department of Physiotherapy and Paramedicine, Glasgow Caledonian University - UK

ABSTRACT

Background: Rehabilitation after stroke often falls short of providing optimal benefits to patients due to challenges related to restricted access, financial burden, and lack of adherence. Video-based intervention may help to overcome these barriers.
Purpose: To develop and test the feasibility of video-based telerehabilitation (TR) exercises for facilitating upper extremity motor function among stroke survivors.
Method: The study was conducted in two phases: video development and feasibility testing. The research team developed the video. Feasibility was assessed using a single-group pre-test–post-test design, focusing on adherence, acceptability, satisfaction, and motor function. A total of 30 stroke survivors were purposively recruited. Each participant received a 3-minute-42-second video-based TR program to be performed three times weekly over four weeks. Adherence and acceptability were monitored through participant-maintained diaries. Satisfaction and motor function were measured using the Client Satisfaction Questionnaire-8 (CSQ-8) and the Wolf Motor Function Test (WMFT), respectively.
Results: Twenty-nine participants (15 males) completed the study. The participants showed a strong commitment to the intervention with a 97% adherence rate. They accepted and reported the intervention as satisfactory. A remarkable score of satisfaction corroborates this based on the CSQ-8 (Mean [SD] = 28.6 [4.90]). A significant improvement in the WMFT was observed ($p < 0.001$) with a large effect size ($d = 1.14$) post-intervention. The mean difference was 8.8, with a 95% confidence interval ranging from 4.75 to 12.9.
Conclusion: The video-based TR is feasible and can be deployed as a supplementary intervention. Future evaluation of the intervention is warranted to establish its effectiveness.

Keywords: Acceptability, Adherence, Satisfaction, Stroke, Telerehabilitation, Video-based exercises

What is already known about this topic?

- *Previous studies have demonstrated the effectiveness of video-based interventions in enhancing motor function recovery among stroke survivors, especially in high-income countries with abundant rehabilitation resources.*

What does the study add?

- *The present study adds to the existing literature by establishing the feasibility of video-based home exercises for facilitating upper extremity function among stroke survivors, particularly in low-resource settings like Nigeria.*

Received: January 5, 2025
Accepted: September 29, 2025
Published online: October 31, 2025

This article includes supplementary materials

Trial registration: The study was registered with the Pan African Clinical Trials Registry (PACTR 202403746569623).

Corresponding author:
Surajo Kamilu Sulaiman
email: surajo.sulaiman@tiu.edu.iq

Introduction

Stroke is a clinical syndrome characterized by rapidly developing signs of focal or global disturbance of cerebral functions that last for more than 24 hours (1). Stroke has emerged as a significant health challenge in Nigeria (2,3). Studies indicate that stroke contributes to about 80% of neurological disabilities, highlighting a critical need for effective rehabilitation services (4,5). The burden of stroke is compounded by the fact that many stroke survivors experience

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

limited access to appropriate rehabilitation facilities. Stroke is one of the leading causes of long-term disability in adults and presents a global epidemic as well as a major public health concern (4,6). The high burden of stroke, combined with inadequate rehabilitation services, especially in low-resource settings, necessitates the development and evaluation of accessible and affordable rehabilitation strategies, such as telerehabilitation (TR) (5). Low-resource or low-income settings refer to areas with inadequate healthcare infrastructure and limited access to rehabilitation services, impacting patient recovery and quality of life (7).

Telerehabilitation (TR) utilizes electronic communications to deliver healthcare services, aiming to reduce healthcare expenses and ensure equitable accessibility for patients who are geographically isolated or facing financial hardships (8). Access to medical and rehabilitation services in low-resource settings is often limited due to factors such as a shortage of skilled professionals, remote geographical locations, and sociocultural barriers (9,10). The high prevalence of stroke and its burden on traditional face-to-face rehabilitation underscore the necessity of exploring and evaluating innovative approaches, including the implementation of TR-based intervention strategies (11). Thus, a video-based TR could offer a practical solution to help overcome the challenges associated with low-resource settings.

Studies from high-income countries have demonstrated the efficacy of video-based TR in improving motor function and quality of life among stroke survivors (12,13). Conversely, there is a paucity of evidence on the use of TR to facilitate upper limb functional recovery among stroke survivors in low-income countries, particularly in Nigeria (5,14-16). The use of TR in Nigeria is still developing, partly due to cultural barriers, financial factors, resource limitations, geographical isolation, compliance with rehabilitation, and lack of awareness (15). The use of mobile technology in Nigeria is on the rise. Approximately 70% of the population has access to mobile phones, and internet penetration has reached about 50% of the population (17). These statistics suggest a growing infrastructure that can support TR.

Notwithstanding, some studies in Nigeria have shown the effectiveness of TR in improving clinical outcomes such as motor and functional abilities in low back pain and osteoarthritis (14,16). Moreover, a study similar to the current one examined the development and feasibility of a Yoruba language video-based home exercise program for stroke survivors in Nigeria (5). Although the study demonstrated the feasibility and acceptability of the program, the language used may limit the video's applicability to only Yoruba speakers. An English-based video would likely have a greater outreach than a Yoruba-based video, as English is the adopted national language in Nigeria (18).

Despite the current emerging evidence for the benefits of video-based TR, it remains an under-researched area, especially in low-resource settings, and its implementation in Nigeria seems to be low due to several mitigating factors (15). Hence, the objective of this study was to develop and evaluate the feasibility of an English-language, video-based TR program for stroke survivors with upper extremity impairment in Nigeria, using a single-group pre-test–post-test design.

Methods

The study was conducted in two phases: video development and feasibility testing.

Video development

The development of the video-based TR required a focus group to select the appropriate exercises and craft a simple video. A videographer was involved in developing the video to ensure clarity and quality following the selection of the exercises.

Focus group

We conducted one session of a focus group that lasted for one hour and thirty minutes in a physiotherapy department. The participants of the focus group were selected purposively to ensure the inclusion of information-rich cases in the group (19). The participants included five physiotherapists and a videographer. Semi-structured questions were used to guide the focus group discussion around the video content. Specifically, the exercises to be selected, the length, and the quality of the video. During the focus group session, the context of the discussion was provided, ground rules were set, and the purpose was explained to the participants. The discussion was anchored by the first author with the help of a research assistant.

Focus group outcome

The selected exercises were based on the recommendations of the National Stroke Foundation Clinical Guidelines (20). The guidelines included techniques such as constraint-induced movement therapy, mobility-task training, and repetitive task-specific training (for example, shoulder flexion and extension, elbow flexion and extension, forearm supination and pronation, and hand slide movement).

These exercises were designed for individuals with mild to moderate upper limb impairments, operationalized as a score of two on the Modified Ranking Scale (mRS). This classification was used to ensure that stroke survivors retained sufficient functional mobility to follow and perform the prescribed movements. Exercises were adapted to meet individual needs while maintaining a consistent level of intensity across the stroke survivors.

An agreement was made about the concept of video, including duration, the individual to be featured in the video, clarity of demonstration, audiovisual quality, and cultural appropriateness.

Video production

A script was created based on the insight gathered from the focus group discussion. During the video shoot, attention was paid to factors such as lighting and sound. Adjustments to the video were made based on additional insights or recommendations from the focus group. The videographer

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

recorded the exercise program into a video of 14 MB with a duration of three minutes and forty-two seconds. The quality of the video is based on Motion Picture Expert Group-4 (MPEG-4) with a resolution of 360 x 640.

Testing and iteration

The video-based TR was sent to three participants via WhatsApp Messenger to gather feedback on its effectiveness and appeal. The feedback from the three participants was used to make further improvements to the video content. Final adjustments were made, and the video-based TR was finalized for distribution. The final video began with a brief introduction and included five short exercises. The stroke survivors were required to adhere to the exercises prescribed in the video.

Instruction manual

An instruction manual was developed to accompany the video-based TR. The manual contained written information about the video-based exercise program, including a diagrammatic illustration of the prescribed exercises. The exercises were described in plain English. To ensure comprehension among participants with lower literacy levels, the written instructions were crafted using clear and straightforward sentences, avoiding complex medical terminology (supplementary file).

Feasibility testing

We used the CONSORT 2010 statement extension to randomized pilot and feasibility trials to report the feasibility testing (22).

Study design

We employed a single-group pretest-posttest feasibility design to test for adherence, acceptability, satisfaction, and preliminary effectiveness of the video-based TR exercise. The methodology of this study was adapted from the work of Odetunde et al. (5).

Participants

The participants were screened using predetermined eligibility criteria. The inclusion criteria were: an episode of a cerebrovascular accident leading to one-sided hemiplegia, mRS <2, Adults (25-65 years), Individuals capable of initiating wrist and finger extension, stable cognitive functioning (Montreal Cognitive Assessment, MoCA ≥26), and having a caregiver during the TR session. Additionally, possessing a phone with internet access and having basic functional literacy, such as practical skills needed to read, write, and do math for real-life purposes, and performing simple tasks to understand short tests in stroke survivors with upper extremity impairment were considered (23). The exclusion criteria included: a major active, coexistent neurological or psychiatric disease (including dementia), history of serious chronic obstructive pulmonary disease or oxygen dependence, previous upper limb injury or condition that limits upper limb range of motion, and significant visual impairment. Visual impairment was assessed through self-report and observable functional limitations during screening; no standardized visual assessment tool was used.

Intervention

The intervention was a three-minute forty-two-second video-based TR for the upper extremities, which comprised five exercises: shoulder flexion and extension, elbow flexion and extension, forearm supination and pronation, hand slide movement, and constraint-induced movement therapy. The exercise dose for shoulder flexion and extension, elbow flexion and extension, forearm supination and pronation, and hand slide movement consisted of 20 repetitions in two sets. For the constraint-induced movement therapy, the unaffected hand was constrained for six hours, enabling the affected hand to perform 200 repetitions of intensive training for 2 hours per day of activities of daily living.

The video-based TR, including the instruction manual, was distributed to the stroke survivors. The manual contained information on how to perform the exercises demonstrated in the video. The patients were asked to watch the video. The video clip was played three times, where the participants observed the displayed movement and replicated it using their affected upper extremity. The participants found it easy to perform the movement without encountering any difficulty during the trial. The participants were asked to complete their home exercise programs three times a week for four weeks with no follow-up or access to a physiotherapy clinic. After this period, follow-up outcomes were measured by an independent assessor.

All participants were requested to keep a diary to record their training. For the prevention and management of adverse events, participants can call the researcher at any time if they have any questions or problems. If participants experienced an intolerable muscle pain or injury, they were advised to visit a hospital and were examined for their condition by the principal investigator.

Outcomes

The outcomes of the study included feasibility aspects (adherence, acceptability, and compliance, which were evaluated based on participant-maintained diaries), satisfaction with the intervention (measured with the Client Satisfaction Questionnaire [CSQ-8]), and motor function (assessed with the Wolf Motor Function Test [WMFT]). Adherence to the intervention involves recording withdrawals and intervention dropout rates, together with the number of therapy sessions completed. Adherence was evaluated through self-reported diaries. Patients were instructed to document the frequency of daily sessions and the duration of home exercises completed during each session. Acceptability and compliance were evaluated based on diary completeness, participant feedback, and expressed willingness to continue the exercise program. The self-reported diaries were verified and monitored by the assessor. However, there was no objective monitoring.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

The WMFT is a valid and reliable instrument used in the assessment of upper extremity motor function of stroke patients (24). The WMFT has been validated for use with patients ranging from acute to chronic stroke stages (25). The participants' motor function was evaluated at baseline and after four weeks of intervention. The WMFT consists of three parts (time, functional ability, and strength), including 15 function-based tasks and two strength-based tasks. Items 1-6 involve timed functional tasks, 7-14 are measures of strength, and the remaining 9 items consist of movement quality when completing various tasks. It uses a 6-point ordinal scale of 0-5 with a maximum score of 75, where lower scores indicate a lower functioning level.

The CSQ-8 is a valid instrument used to evaluate the satisfaction of patients (26). The CSQ-8 scores range from 8 to 32, with higher values indicating higher satisfaction. Items are scored on a Likert scale from 1 (low satisfaction) to 4 (high satisfaction) with different descriptors for each response point. The CSQ-8 is advantageous because it is a concise, 8-item tool that has shown strong reliability and validity in a variety of settings (27). The participants' satisfaction was assessed in the fourth week.

The demographic characteristics and clinical variables of the participants were assessed at the beginning of the study using a biographic information sheet. The initial outcomes were evaluated before the start of the study. Subsequent follow-up outcomes were assessed by an independent evaluator who was not involved in patient recruitment or management.

Sample size estimation

The sample size was estimated using the G*Power software (28). We employed an alpha level of 0.05, a power of 80%, and an effect size of 0.5 (5). The estimated sample size was 21 participants. However, to accommodate a projected attrition rate of 42%, the final enrollment was adjusted to 30 participants. We selected the participants using a purposeful sampling technique to ensure the inclusion of only participants who met the eligibility criteria.

Ethical consideration

This study was approved by the Health Research Ethics Committee of the College of Health Sciences, Bayero University Kano, Nigeria (BUK/PHYSIO/03/0001). Additional clearance was also obtained from the Ethical Committee of the Federal Medical Center Jabi, Nigeria (FMCABJ/HREC/2023/085). The study was conducted at the Physiotherapy Department of the Federal Medical Center, Jabi, Nigeria. Before enrollment, all participants were provided with a written information sheet developed by the research team, which contained detailed information about the study's objectives, procedures, potential risks and benefits, confidentiality measures, and their right to withdraw at any time during the study. Written informed consent was given by the participants before the commencement of the study, after the details of the study had been thoroughly explained to them (via the information sheet). The study was registered prospectively with the Pan African Clinical Trials Registry (PACTR20240374656923) on the 28th of March 2024.

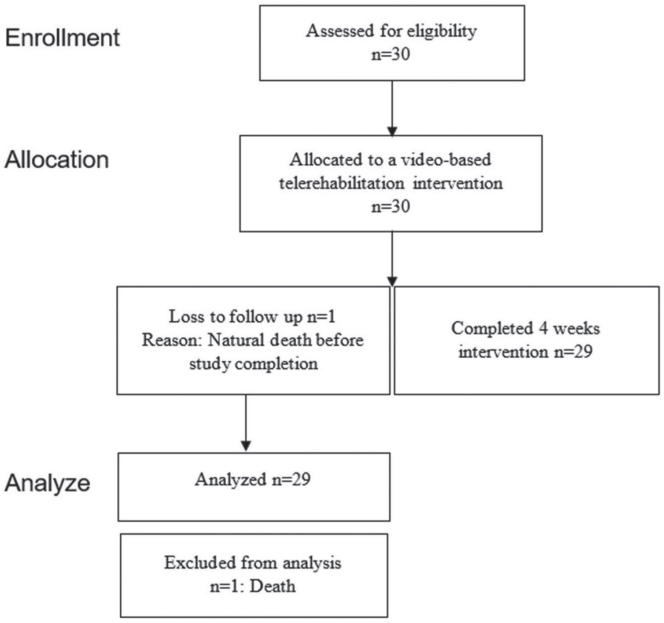


FIGURE 1 - Study flow diagram.

Data collection procedure

The participants were recruited from the Physiotherapy Department of the Federal Medical Center, Jabi, Nigeria, where they were admitted to the rehabilitation ward and the Physiotherapy Outpatient Clinic. Before discharge, eligible participants were introduced to the video-based TR intervention and were informed of the study's purpose and procedures. Thereafter, informed consent was obtained from each patient, and confidentiality was ensured by storing the data in password-protected files on a secure computer with restricted access. A study assessor conducted the baseline evaluation, which included demographics, clinical variables, and motor function assessments. Following completion of the baseline evaluation, the video was sent to the participants via WhatsApp messenger, and the manual explaining the exercises with the aid of a diagram and the importance of keeping their routine for the following 4 weeks was attached. The study participants were expected to adhere to and comply with the exercises shown in the video. Participants were required to indicate whether they had fulfilled their exercise requirement in each session. Adherence was monitored through participant-maintained diaries. Acceptability was evaluated based on diary completeness, participant feedback, and expressed willingness to continue the exercise program. The participants' compliance was supervised, and regular follow-ups were conducted via WhatsApp Messenger once a week. An attrition rate was calculated based on the turnout of the study participants post-intervention. The participants were reassessed in the 4th week by the assessor upon completion of the intervention.

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Statistical analysis

Information gathered was coded and entered into the computer with the statistical package for social sciences SPSS version 20 (SPSS Inc., Chicago, IL, USA). A probability level of 0.05 was used to indicate statistical significance. Descriptive statistics of mean, frequencies, and percentages were used to summarize the data (socio-demographic and clinical variables). A paired t-test was used to analyze the differences in WMFT mean scores between pre- and post-intervention. Effect size was calculated using the formula $d = t/\sqrt{N}$ (29). Where t is the t-score and N is the total sample size. Based on Cohen's d criteria, the effect size values were interpreted as trivial (<0.2), small (≥0.2 and <0.5), moderate (≥0.5 and <0.8), and large (≥0.8) (30).

Results

Development of the video

The video script was developed based on the outcomes of the focus group, drawing from the recommendations of the National Stroke Foundation Clinical Guidelines. The videographer also contributed to editing, ensuring video clarity, appropriate lighting, sound, and overall user-friendliness**.** The videographer recorded the exercise program using a 2778 by 1284-pixel resolution at 458ppi. The video size is 14 MB, the duration is three minutes and 42 seconds, and the video quality is MPEG-4 resolution at 360 x 640. The physiotherapist in the video gave consent for it to be shared publicly. The video can be accessed online via the following YouTube link: Online

Socio-demographic and clinical variables

A total of 15 (51.7%) males and 14 (48.3%) females participated in the study (one participant did not complete the study). The mean age ± SD of the participants combined was 52.2 ± 11.3 years. The mean age ± SD for males and females was 49.6 ± 12.6 and 54.9 ± 9.3 years, respectively. The participant's cognitive function was good, with a mean ± SD MOCA score of 27.8 ± 2.31. More than half (58.6%) of the participants were married. Moreover, 48.3% of the participants were employed. Most of the participants had attained at least a college-level education (41.4%), with a small fraction having no formal education (3.4%). More than three-quarters of the patients have an ischemic stroke (69.0%) with right-sided impairment (65.5%). The mean time since the onset of stroke was 6.07 ± 7.01 months. In terms of severity, the majority of the participants have mild to moderate impairment. Ten of the participants were classified as having mild hand paresis, 16 moderate, and three severe. This classification was done based on participants' performance during the WMFT test and clinical observation, focusing on their ability to complete timed and functional tasks. There were six left-handers. Ten participants had a left-hemispheric stroke, and 19 had a right-hemispheric stroke (Table 1). Figure 2 provides a visualization of the time since the stroke.

TABLE 1 - Socio-demographic and clinical variables

Variable	Category	n (%)	Mean and SD
Age (Years)			
	All participants		52.21± 11.309
	Male		49.67 ± 12.68
	Female		54.93 ± 9.32
Gender			
	Male	15 (51.7)	
	Female	14 (48.3)	
Marital Status			
	Single	2 (6.9)	
	Married	17 (58.6)	
	Divorced	6 (20.7)	
	Widowed	4 (13.8)	
Education Level			
	Informal	1 (3.4)	
	Elementary	2 (6.9)	
	Middle school	4 (13.8)	
	High school	10 (34.5)	
	College and above	12 (41.4)	
Employment			
	Unemployed	7 (24.1)	
	Employed	14 (48.3)	
	Retiree	8 (27.6)	
Type of Stroke			
	Ischemic	20 (69.0)	
	hemorrhagic	9 (31.0)	
Affected Side			
	Right	19(65.5)	
	Left	10(34.5)	
Functional Dominance			
	Right	23 (79.3)	
	Left	6 (20.7)	
Stroke Severity			
	Mild	10 (34.5)	
	Moderate	16 (55.2)	
	Severe	3 (10.3)	
MOCA	All participants		27.83 ± 2.316
Time Since Stroke Onset (Months)	All participants		6.07 ± 7.015
Co-Morbidities			
	Nil	2 (6.9)	
	Hypertension and Diabetes	9 (31.0)	
	Hypertension only	18 (62.1)	
	Diabetics	0	

MoCA: Montreal Cognitive Assessment, n = frequency, % = percentage, SD = standard deviation.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

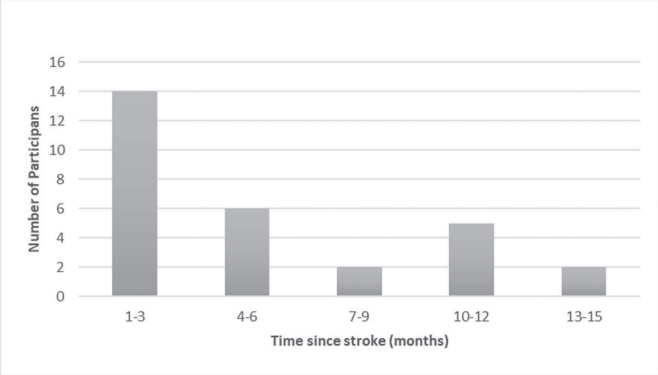


FIGURE 2 - Time since stroke.

Feasibility aspects

Acceptability

Acceptability refers to how well the intervention is received by the participants and whether they find it appropriate and suitable for their needs and circumstances (31). By the end of the study, the participants found the intervention satisfactory. They diligently followed their prescribed exercise program and even expressed interest in receiving additional video exercises and willingness to engage with and continue using the video-based TR, thus indicating high acceptability (100%).

Adherence

Adherence is simply the degree to which the participants actively choose to follow the recommendations of the intervention (32). The participants followed the intervention willingly, and no participant withdrew from the study. However, we recorded one mortality from unrelated medical issues. The participant who died was not included in the overall analysis. Thus, we recorded high adherence to the intervention with 29 participants completing the study (96.6%).

Compliance

Compliance is similar to adherence; however, the participants passively implement the intervention requirements (32). Weekly check-ins were conducted via WhatsApp to remind participants to complete their exercises, address any questions, and encourage adherence to the program. Participants received a diary with prompts to help track the frequency and duration of their exercises. The stroke survivors had a thorough understanding of the study; hence, there was maximum cooperation during the study.

Satisfaction

The participants expressed their satisfaction with the intervention and showed a remarkable score on the CSQ-8 with a mean ± SD value of 28.6 ± 4.90.

Comparison of the pre- and post-intervention mean scores of WMFT

Table 2 shows the change in motor function among the participants. A paired *t*-test revealed statistically significant differences in WMFT mean scores ($p < 0.001$, $d = 1.14$), with a large effect size 4 weeks post-intervention. The mean change was 8.8 points.

TABLE 2 - Comparison between pre-intervention and post-intervention mean scores of WMFT

Outcome	Pre-intervention	Post-intervention	Difference (95%CI)	t-value (df)	*p*-value	Effect size (*d*)
	Mean ± SD	Mean ± SD				
WMFT	41.5 ± 14.7	50.3 ± 14.2	8.8 (4.75–12.9)	4.44 (28)	0.0001*	1.14

WMFT, Wolf Motor Function Test; df, degree of freedom; CI, confidence interval; SD, Standard Deviation. *$p < 0.001$

Discussion

This study aimed to develop and evaluate the feasibility of an English-language, video-based TR program for stroke survivors with upper extremity impairment in Nigeria. Recommended evidence-based exercises were selected, and scripting of the video contents was done and extrapolated to the intended population. The results of the study demonstrated that home rehabilitation guided by a video-based therapy program could be an excellent alternative to conventional hospital-based therapy. It is also safe; completing therapy at home did not lead to increased stroke-related complications. Moreover, the study showed a significant preliminary improvement in the upper limb motor function. Prior studies have demonstrated the efficacy of video-based intervention in enhancing stroke recovery and indicated its practicability and acceptability for individuals with mild to moderate impairment resulting from stroke (5,12,33). However, those studies differ from the present study because the current study employed the English language medium, which is a national language in Nigeria. Moreover, the present study focuses solely on upper extremity motor function.

Ninety-six percent of the enrolled participants completed the intervention. This high adherence rate could be attributed to several possible factors, including the convenience of at-home participation and the engaging nature of the program. The video-based exercise appears to be accessible and affordable to stroke survivors, for instance, by reducing the cost of transportation and other related issues associated with face-to-face rehabilitation, which is hindered by several specific barriers. These include limited availability of rehabilitation facilities, high costs associated with in-person visits, and long waiting times for appointments. Additionally, the lack of trained rehabilitation professionals in many regions further complicates access to required services (34). Participants' satisfaction in the current study was exceptionally high, consistent with findings from previous video-based TR studies (5,35).

Following the intervention, there was a substantial preliminary improvement in the WMFT, with a large effect size

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

(1.14). Moreover, the change was clinically significant, with an 8.8-point increase, exceeding the 1.2-point minimal clinically important difference reported for WMFT in acute stroke (36). Based on the preexisting literature (37,38), the magnitude of change demonstrated in the present study exceeds previously reported potential practice effects. The findings of this study suggest that the improvement observed in stroke survivors is likely attributable to factors such as family support, motivation, the opportunity to engage in therapy at their preferred pace, early rehabilitation, and the familiar environment in which the therapy took place 12,39-41).

Several studies have demonstrated that a video-based therapy program conducted at home proves to be a highly effective substitute for traditional therapy conducted in hospitals (5,12,42). The findings from these studies affirm the feasibility and applicability of video-based therapeutic rehabilitation, thus corroborating the findings of the present study. Conversely, some conflicting findings are discernible in the literature. For instance, a study reported no significant improvement in WMFT following a video-based home exercise program among stroke survivors (43). This could be due to the likelihood that patients with chronic stroke may not demonstrate a significant gain in motor function recovery (44).

To our knowledge, this is the first attempt to test the feasibility and usefulness of an English language medium video-based TR program for stroke survivors with upper extremity impairment in Nigeria. However, the study has limitations. Firstly, we employed a pretest-posttest single-group feasibility design; hence, our findings should be interpreted as preliminary. Secondly, the absence of a control group and lack of randomization restrict the ability to draw causal inferences about the effectiveness of the intervention. Thirdly, the study sample was purposely selected and consisted largely of individuals with relatively high levels of education and digital literacy, which may have facilitated their engagement with the video-based TR program. This may have positively influenced their ability to engage with the intervention, thus limiting the generalizability of the findings to broader populations, particularly in underserved or low-literacy communities. Fourthly, adherence and compliance were based on self-reported data, which may introduce potential recall and social desirability biases. Finally, the scalability of the intervention in rural areas may be limited by practical constraints such as poor mobile connectivity and unreliable electricity (45). Hence, future studies should employ a rigorous study design, such as a randomized controlled trial, to robustly assess the impact of the video-based TR program.

Implications for physiotherapy practice

The results of this study provided valuable information regarding the feasibility and preliminary effectiveness of video-based therapeutic exercises. The findings suggest that video-based TR can be an effective, accessible, and probably cost-efficient supplement to traditional physiotherapy for upper-extremity rehabilitation among stroke survivors. Hence, physiotherapists should consider alternative approaches such as the use of TR to reach out to patients, for example, in resource-poor settings or during health crises. Moreover, the use of a common language while developing TR could facilitate the understanding and assimilation of the information. However, despite the promising nature of this video-based TR, further studies are warranted to establish its broader clinical efficacy and integration into standard physiotherapy practice.

Conclusion

This study developed and evaluated the feasibility and preliminary effectiveness of an English-language video-based TR program for stroke survivors with upper extremity impairment in Nigeria. Although the findings are preliminary, they suggest potential clinical benefits. Hence, further investigation through a fully powered, randomized controlled trial is warranted to rigorously evaluate the efficacy and generalizability of this intervention, particularly among diverse patient populations in settings with greater technological barriers.

Acknowledgments

The authors wish to extend their appreciation to the Department of Physiotherapy, Federal Medical Centre Jabi, and the participants of the study for facilitating the smooth execution of this research.

Disclosures

Conflict of interest: The authors declare no conflict of interest.

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contributions: HMA, SKS, and UMB conceptualized the study; HMA collected data while SKS supervised the data collection. HMA, DS, and SKS analyzed the data. HMA drafted the initial manuscript. All authors provided critical feedback and contributed substantially to revising the manuscript, and have agreed to the present version of the manuscript.

Data availability statement: Data will be made available upon a valid request.

Ethical approval statement: We obtained ethical approval from the College of Health Sciences, Bayero University, Kano and Federal Medical Center, Jabi, Abuja, Nigeria.

Patient consent statement: We obtained informed consent from the participants prior to the study.

References

1. Sharma J, Abdullah U. Recent advancements to improve balance in stroke rehabilitation. Int J Adv Res (Indore). 2019; 7(3):365-372. CrossRef
2. Adeloye D, Ezejimofor M, Auta A, et al. Estimating morbidity due to stroke in Nigeria: a systematic review and meta-analysis. J Neurol Sci. 2019;402:136-144. CrossRef PubMed
3. Feigin VL, Stark BA, Johnson CO, et al.; GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol. 2021; 20(10):795-820. CrossRef PubMed
4. Mohan U, Babu SK, Kumar KV, et al. Effectiveness of mirror therapy on lower extremity motor recovery, balance and mobility in patients with acute stroke: a randomized sham-controlled pilot trial. Ann Indian Acad Neurol. 2013;16(4):634-639. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

5. Odetunde MO, Binuyo OT, Maruf FA, et al. Development and feasibility testing of video home based telerehabilitation for stroke survivors in resource limited settings. Int J Telerehabil. 2020;12(2):125-136. CrossRef PubMed
6. Johnston SC, Mendis S, Mathers CD. Global variation in stroke burden and mortality: estimates from monitoring, surveillance, and modelling. Lancet Neurol. 2009;8(4):345-354. CrossRef PubMed
7. Maanoosi M. Stroke rehabilitation in low resource countries: time to provide an organised service. *South Sudan Med J.* 2024;17(1):27-31. CrossRef
8. Rogante M, Grigioni M, Cordella D, et al. Ten years of telerehabilitation: a literature overview of technologies and clinical applications. NeuroRehabilitation. 2010;27(4):287-304. CrossRef PubMed
9. Surya N, Someshwar HP. Low-cost telerehabilitation in low- and middle-income countries (LMICs): overcoming barriers to access and improving healthcare delivery. NeuroRehabilitation. 2025;56(1):30-36. CrossRef PubMed
10. Mohamad IS, Defi IR. Telerehabilitation in low- and middle-income countries. In: Wang TC, ed. Telehealth and Telemedicine. IntechOpen; 2022, CrossRef.
11. Calvaresi D, Schumacher M, Marinoni M, et al. Agent-based systems for telerehabilitation: strengths, limitations and future challenges. In: Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). LNAI; 2017:3-24. Lecture Notes in Computer Science; vol 10685., CrossRef.
12. Redzuan NS, Engkasan JP, Mazlan M, et al. Effectiveness of a video-based therapy program at home after acute stroke: a randomized controlled trial. Arch Phys Med Rehabil. 2012;93(12):2177-2183. CrossRef PubMed
13. Prajjwal P, Chandrasekar KK, Battula P, et al. The efficacy of virtual reality-based rehabilitation in improving motor function in patients with stroke: a systematic review and meta-analysis. Ann Med Surg (Lond). 2024;86(9):5425-5438. CrossRef PubMed
14. Mbada CE, Olaoye MI, Dada OO, et al. Comparative efficacy of clinic-based and telerehabilitation application of McKenzie therapy in chronic low-back pain. Int J Telerehabil. 2019;11(1):41-58. CrossRef PubMed
15. Odole AC, Odunaiya NA, Ojo OD, et al. Tele-physiotherapy in Nigeria: perceived challenges by physiotherapists to its implementation. Int J Telemed Clin Pract. 2015;1(2):186. CrossRef
16. Odole AC, Ojo OD. Is telephysiotherapy an option for improved quality of life in patients with osteoarthritis of the knee? Int J Telemed Appl. 2014;2014:903816. CrossRef PubMed
17. GSMA. *The Role of Mobile Technology in Driving the Digital Economy in Nigeria.* 2024. Online
18. Adepoju T. Literacy instructors' perception on the use of English as the lingua franca in Nigeria. Lwati A J Contemp Res. 2010;7(3):20-26. CrossRef
19. Palinkas LA, Horwitz SM, Green CA, et al. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health. 2015;42(5):533-544. CrossRef PubMed
20. Wright L, Hill KM, Bernhardt J, et al.; National Stroke Foundation Stroke Guidelines Expert Working Group. Stroke management: updated recommendations for treatment along the care continuum. Intern Med J. 2012;42(5):562-569. CrossRef PubMed
21. National Stroke Foundation. Rehabilitation. Living Clinical Guidelines for Stroke Management. 2023. Online (Accessed January 2025)
22. Eldridge SM, Chan CL, Campbell MJ, et al.; PAFS consensus group. CONSORT 2010 statement: extension to randomized pilot and feasibility trials. BMJ. 2016;355:i5239. CrossRef PubMed
23. Hahn EA, Magasi SR, Carlozzi NE, et al. Health and functional literacy in physical rehabilitation patients. Health Lit Res Pract. 2017;1(2):e71-e85. CrossRef PubMed
24. Wolf SL, Thompson PA, Morris DM, et al. The EXCITE trial: attributes of the Wolf Motor Function Test in patients with subacute stroke. Neurorehabil Neural Repair. 2005;19(3):194-205. CrossRef PubMed
25. Winstein CJ, Miller JP, Blanton S, et al. Methods for a multisite randomized trial to investigate the effect of constraint-induced movement therapy in improving upper extremity function among adults recovering from a cerebrovascular stroke. Neurorehabil Neural Repair. 2003;17(3):137-152. CrossRef PubMed
26. Kelly PJ, Kyngdon F, Ingram I, et al. The Client Satisfaction Questionnaire-8: psychometric properties in a cross-sectional survey of people attending residential substance abuse treatment. Drug Alcohol Rev. 2018;37(1):79-86. CrossRef PubMed
27. Attkisson CC, Zwick R. The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. Eval Program Plann. 1982;5(3):233-237. CrossRef PubMed
28. Kang H. Sample size determination and power analysis using the G*Power software. J Educ Eval Health Prof. 2021;18:17. CrossRef PubMed
29. Lenhard W, Lenhard A. Computation of effect sizes. Psychometrica. 2016;2:5-9. CrossRef
30. AlWahaibi ISH, Al-Hadabi DAMY, Al-Kharusi HAT. Cohen's criteria for interpreting practical significance indicators: a critical study. Cypriot J Educ Sci. 2020;15(2):246-258. CrossRef
31. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res. 2017;17(1):88. CrossRef PubMed
32. Petkova V. Compliance, adherence and concordance - effect on public health. World J Pharm Pharm Sci. 2020;9(3):1-4. CrossRef
33. Mbada C, Oluwatobi S, Awosika H, et al. Development and feasibility testing of a smartphone video-based exercise program for patients with knee osteoarthritis. Digit Med. 2022;8(1):9. CrossRef
34. Ferreira V, Agnihotram RV, Bergdahl A, et al. Maximizing patient adherence to prehabilitation: what do the patients say? Support Care Cancer. 2018;26(8):2717-2723. CrossRef PubMed
35. Denny MC, Vahidy F, Vu KYT, Sharrief AZ, Savitz SI. Video-based educational intervention associated with improved stroke literacy, selfefficacy, and patient satisfaction. *PLoS One*. 2017;12(3):1-12. CrossRef
36. Fritz SL, Blanton S, Uswatte G, et al. Minimal detectable change scores for the Wolf Motor Function Test. Neurorehabil Neural Repair. 2009;23(7):662-667. CrossRef PubMed
37. Lang CE, Edwards DF, Birkenmeier RL, et al. Estimating minimal clinically important differences of upper extremity measures early after stroke. Arch Phys Med Rehabil. 2008;89(9):1693. CrossRef
38. Langan J, Delave K, Phillips L, et al. Home-based telerehabilitation shows improved upper limb function in adults with chronic stroke: a pilot study. J Rehabil Med. 2013;45(2):217-220. CrossRef PubMed

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

39. Gbiri CA, Akinpelu AO. Quality of life of Nigerian stroke survivors during first 12 months post-stroke. Hong Kong Physiother J. 2012;30(1):18-24. CrossRef
40. Li W, Zhu G, Lu Y, et al. The relationship between rehabilitation motivation and upper limb motor function in stroke patients. Front Neurol. 2024;15:1390811. CrossRef PubMed
41. Sikuka HM, Lupenga J, Nkhata L. Predictors of upper limb motor recovery in stroke survivors: a pre-post test study design. BMJ Open. 2024;14(11):e081936. CrossRef PubMed
42. Denny MC, Vahidy F, Vu KYT, et al. Video-based educational intervention associated with improved stroke literacy, self-efficacy, and patient satisfaction. PLoS One. 2017;12(3): e0171952. CrossRef PubMed
43. Emmerson KB, Harding KE, Taylor NF. Home exercise programmes supported by video and automated reminders compared with standard paper-based home exercise programmes in patients with stroke: a randomized controlled trial. Clin Rehabil. 2017;31(8):1068-1077. CrossRef PubMed
44. Li S. Stroke recovery is a journey: prediction and potentials of motor recovery after a stroke from a practical perspective. Life (Basel). 2023;13(10):1-13. CrossRef PubMed
45. Tahmasebi F. The digital divide: a qualitative study of technology access in rural communities. AI Tech Behav Soc Sci. 2023;1(2):33-39. CrossRef

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 286-296
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3590
ORIGINAL RESEARCH ARTICLE

Physical therapists' perspectives and practices on weight management for chronic pain patients with obesity: a cross-sectional survey

Melanie Liechti[1,2,3], Massimo Menegon[2], Alexander P. Schurz[1,2,4], Matteo Vanroose[1,3], Ron Clijsen[1,2,5,6], Anneleen Malfliet[3,7,8], David Faeh[2], Heiner Baur[2], Jan Taeymans[1,2], Nathanael Lutz[1,2]

[1]Department of Movement and Sport Sciences, Faculty of Physical Education and Physiotherapy, Vrije Universiteit Brussel, Brussels - Belgium
[2]School of Health Professions, Bern University of Applied Sciences, Bern - Switzerland
[3]Pain in Motion Research Group (PAIN), Department of Physiotherapy, Human Physiology and Anatomy (KIMA), Faculty of Physical Education & Physiotherapy, Vrije Universiteit Brussel, Brussel - Belgium
[4]Faculty of Medicine, University of Bern, Bern - Switzerland
[5]Rehabilitation and Exercise Science Laboratory RESLab, Department of Business Economics, Health, and Social Care, University of Applied Sciences and Arts of Southern Switzerland, Landquart/Manno - Switzerland
[6]International University of Applied Sciences THIM, Landquart - Switzerland
[7]Research Foundation – Flanders (FWO), Brussels - Belgium
[8]Chronic pain rehabilitation, Department of Physical Medicine and Physiotherapy, University Hospital Brussels - Belgium

ABSTRACT

Introduction: Overweight and obesity management in primary care gained importance due to its global rise and association with cardiometabolic diseases and chronic musculoskeletal disorders (MSD). Physical therapists are well-positioned to incorporate weight management in their practice. This survey evaluates attitudes, barriers and clinical practices of physical therapists regarding weight management for individuals with chronic pain and comorbid overweight or obesity.
Methods: German-speaking physical therapists in Switzerland volunteered in an online survey. Descriptive statistics summarized physical therapists' attitudes, barriers and practices. Logistic regression analyses identified factors associated with recommending and implementing weight reduction interventions.
Results: Of the 581 respondents, 92.1% acknowledged the importance of educating patients on obesity-related health risks, and 81.6% recommended weight reduction. Two-thirds offered weight reduction interventions in physical therapy, although fewer (57.3%) received corresponding training during entry-level education. Most weight reduction interventions focused on movement, endurance, and strength, whilst a minority addressed nutrition, sleep or stress. The odds of offering weight reduction interventions were higher among those with entry-level education (OR: 2.1, 95% CI: 1.4-3.3), further education (OR: 1.9, 95% CI: 1.1-3.3), and for those who perceived weight reduction interventions (OR: 6.5, 95% CI: 3.5-12.3) and counseling on other lifestyle factors (OR: 2.9, 95% CI: 1.2-7.3) as within their professional scope.
Conclusions: While physical therapists showed positive attitudes toward weight reduction interventions for individuals with chronic pain and overweight or obesity, their education and implementation remain inconsistent, primarily focusing on improving physical activity. Clear competencies and structured training are needed to integrate evidence-based weight management into clinical practice.

Keywords: Chronic pain, Lifestyle, Obesity, Physical therapy, Weight management

What is already known about this topic:

- *Physical therapists recognize the importance of addressing lifestyle factors in their clinical practice.*
- *The implementation of weight management strategies remains limited across different healthcare systems and countries.*

What does the study add:

- *This study highlights Swiss physical therapists' role in promoting weight management as part of chronic pain treatment. It identifies key implementation barriers and emphasizes the need for education, systemic support, and development of competency standards.*

Received: May 9, 2025
Accepted: October 29, 2025
Published online: December 1, 2025

This article includes supplementary material

Corresponding author:
Melanie Liechti
email: melanie.liechti@bfh.ch

Introduction

The prevalence of overweight and obesity is high globally, affecting 38% and 14% of adults, respectively, in 2020 (1). Overweight and obesity contribute to various non-communicable diseases, such as metabolic diseases, cardiovascular diseases, and musculoskeletal disorders (MSD), like chronic low back pain (CLBP) or osteoarthritis (OA) (1-4). The causes

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

of obesity are multifactorial, reflecting complex and not yet fully understood interactions between genetic, environmental, socio-cultural, and psychological factors (5,6). Obesity is recognized as a chronic condition that may occur with or without abnormalities in adipose tissue function or distribution (5,6). According to the most recent definitions, the diagnostic criteria of overweight and obesity should combine Body Mass Index (BMI) with at least one additional measure of adiposity, such as waist-to-hip ratio or direct measurement of body fat (6).

International MSD guidelines recommend a conservative multimodal treatment approach, with first-line treatments including supervised exercise therapy and cognitive-behavioral therapy (7,8), as well as lifestyle modifications, including weight management (9-12). Interestingly, weight reduction interventions, including the promotion of exercise and a healthy diet, show promising results in terms of pain relief, improved function, and enhanced quality of life for the treatment of CLBP and knee OA (13-15). Weight reductions may reduce pain and improve function through a range of biopsychosocial mechanisms. Biologically, it reduces low-grade inflammation and improves insulin sensitivity (16), and psychologically, it reduces somatization and results in improved psychological well-being and management of comorbidities (17). Socially, multimodal treatments can help to reduce pain-related fear of movement (18), fostering greater confidence in physical activity and engagement in daily and social activities, which are often reduced in individuals with obesity and chronic pain (19).

Since conservative care for patients with MSD is frequently provided by physical therapists, the implementation of weight reduction interventions within physical therapy practice should be explored. Physical therapists are trained within a holistic, biopsychosocial framework, which makes them ideally placed to support individuals with chronic MSD and overweight or obesity in weight mmanagement (12,20,21). Various lifestyle factors, including a healthy diet, smoking cessation, weight control, and stress management, are already incorporated into physical therapy curricula in the United States, United Kingdom, Canada, Australia, New Zealand, and Ireland (22). Therefore, integrating these factors into curricula in other countries seems both feasible and relevant. However, to ensure effective implementation, it is essential to explore physical therapists' attitudes, barriers and practices regarding these factors. Surveys across different countries have highlighted a general awareness among physical therapists of the importance of addressing lifestyle factors within their professional role (23-26). These surveys showed that physical therapists mainly assessed patients' physical activity levels (78%), while dietary status (36-55%) and smoking status (35-46%) were evaluated less often (23,24,27). Moreover, results revealed that physical therapists' knowledge and experience in managing lifestyle behaviors, especially weight management, remain limited (23,25,28).

Although lifestyle factors are integrated into physical therapy curricula, their implementation into clinical practice remains limited across countries (23,25,28), likely due to variations in educational structures and healthcare systems. Therefore, it is essential to evaluate physical therapists' attitudes and practices on a local level to tailor future educational programs to better suit each healthcare system. Hence, this cross-sectional survey explored Swiss physical therapists' attitudes, barriers, and practices regarding weight management in patients with chronic pain and overweight or obesity.

The research questions for this survey were:

(i) What are the attitudes, barriers, and practices of Swiss physical therapists regarding weight reduction in individuals with chronic pain and overweight or obesity?
(ii) Which factors are associated with the recommendation and implementation of weight reduction interventions by physical therapists?

Materials and methods

Study design

A cross-sectional, anonymous online survey was conducted using Limesurvey (29) between December 2023 and March 2024. The study followed the Strengthening the Reporting of Observational Studies in Epidemiology Statement (STROBE) checklist (30). Because no health-related data were collected, this study did not fall under the Swiss Human Research Act (31,32). Informed consent of all participants was obtained. On the first page of the survey, participants were informed about the survey, and by clicking the button to start the survey, they implied their consent.

Participants and recruitment

German-speaking physical therapists working clinically in Switzerland and students currently working in an internship were eligible to participate.

According to the analysis of the Swiss Federal State Office, 61.8% of Swiss residents speak German (33), and approximately 13,577 physical therapists are working in Switzerland (34), resulting in a target population of 8,391 participants. The required sample size for the online survey was calculated using the R package "presize" (35), with a significance level of 0.05, a proportion of 0.62, and a 95% confidence level, yielding a result of 360 participants. A Swiss federal register including all practicing physical therapists is lacking. Therefore, a snowball sampling approach was followed in order to reach as many participants as possible (32). Responsible individuals from physical therapy schools, local physical therapy associations, and hospitals were asked to spread the survey link through their networks. In addition, lists of physical therapy practices were screened for email addresses on the website of the Swiss Physiotherapy Association, and the survey link was distributed on social media. Every invitation via email or social media included a request to distribute the link among co-workers or other physical therapists in their network.

Outcome measure

A self-administered questionnaire (Supplement 1) was developed using surveys from previous studies on physical therapists and healthcare professionals (20, 24-28, 36-40) along with literature on survey design (32).

In total, the survey consisted of 44 questions, beginning with the inclusion question, whether the participant

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

is currently working as a physical therapist in Switzerland. Subsequently, demographic data were collected, followed by questions assessing attitudes, beliefs, and barriers regarding weight reduction in individuals with chronic pain and overweight or obesity. The survey also explored methods for assessing nutritional status, overweight or obesity, and investigated attitudes and knowledge regarding nutrition management. Of the 44 survey questions, only a subset of 32 questions was used to address the research questions.

Face and content validity of the questionnaire were assessed by different experts, including experienced researchers (n = 3) with expertise in the fields of public health, epidemiology, nutrition, prevention, and knowledge about the Swiss Bachelor's physical therapy education program. A pilot test to evaluate the usability of the survey was performed, including six physical therapists (target population). Modifications of the survey were made according to the feedback and suggestions of the participants.

Statistical analysis

Respondents who did not answer any question beyond the demographics were excluded from analysis. Some participants might not complete all sections of the survey, either due to branching or because they stopped the survey before completion, resulting in missing values. A potential relationship between missing values and demographic data was examined using logistic regression. Descriptive statistics were used to summarize participants' characteristics and to describe continuous variables of the survey questions. Absolute and relative frequency distributions were calculated for categorical variables. Open-ended questions were analyzed using qualitative content analysis (32), which was performed with the software Taguette (41).

Two logistic regression analyses were conducted to assess different factors associated with physical therapists' recommendations for weight reduction and factors associated with the implementation of weight reduction interventions in therapy (Supplement 2). Correlation coefficients were calculated to identify potential multicollinearity, while a stepwise backward elimination approach was used to identify the most important predictor variables, controlling for gender and work experience.

All statistical analyses were calculated using R software, version 4.4.3 (R Core Team, 2025) (42), and 95% confidence intervals (CI) were calculated to provide an estimate of the precision of the parameter estimates for the regression analysis.

Results

A total of 744 responses were collected, of which 163 were excluded, resulting in 581 responses included in the analysis. Exclusion reasons were: Currently not working as a physical therapist in Switzerland (n = 14) and incomplete demographics or no clinical questions answered (n = 149). Of the 581 responses, 80 showed missing values. Logistic regression analyses revealed that the work setting was statistically significantly associated with missing data. The probabilities of missing values across the different work settings were as follows: outpatient (14.1%), inpatient (19.8%), and both settings (12.8%). Therefore, pairwise deletion was applied in the regression analysis, excluding missing data from variables with missing values. Since similar answer patterns across respondents were found, dropouts seemed not to introduce meaningful bias, and therefore their data were included in the descriptive analysis. Supplements 3 and 4 depict patterns of missing values and discuss potential sources of survey discontinuation.

Most respondents were female (81.2%), with an average work experience of 14.1 years (range: 0-43 years), primarily working in an outpatient practice (68.2%). Table 1 shows details of demographic characteristics.

TABLE 1 - Participant characteristics of 581 participants

Characteristic	**n (%)**	
Gender		
Male	108 (18.6)	
Female	472 (81.2)	
Other	1 (0.2)	
Age (y)		
20-24	43 (7.4)	
25-34	233 (40.1)	
35-44	123 (21.2)	
45-54	101 (17.4)	
55-64	76 (13.1)	
65+	5 (0.8)	
Degree*		
Bachelor of Science in training	45 (7.7)	
Bachelor of Science	318 (54.7)	
Master of Science	92 (15.8)	
PhD	7 (1.2)	
Diploma	81 (13.9)	
Other degree	60 (10.3)	
Work setting*		
Stationary	201 (34.6)	
Outpatient	539 (92.8)	
Other	29 (5.0)	
Work environment*		
Hospital	156 (26.9)	
Practice	396 (68.2)	
Rehabilitation center	48 (8.3)	
Care home	40 (6.9)	
Other	30 (5.2)	
Characteristic	**Mean (sd)**	**Range**
Work percentage	70.8 (22.9)	10-100
Clinical experience (y)	14.1 (11.2)	0-43

* The sum might be over 100% as multiple answer selections were possible.
Abbreviations: sd: standard deviation; y: years.

Responses regarding physical therapists' attitudes toward weight management in individuals with chronic pain and overweight or obesity, with a focus on weight reduction, are shown in Figure 1. Counseling on nutrition, exercise, and other

lifestyle factors (e.g., smoking, sleep, and stress) as part of a weight reduction intervention was perceived as important by 95.3%, 98.6%, and 97.6% of respondents, respectively. The lowest agreement was found regarding the perception that weight reduction interventions fall within the scope of physical therapists (56%) and that they feel confident in implementing such interventions in clinical practice (44%).

The perceived barriers faced by respondents when performing weight reduction interventions are displayed in Figure 2. More categories were found during the content analysis of 46 respondents who described other barriers. Those revealed a lack of patient interest (n = 14, 30.4%), weight reduction interventions having no priority in physical therapy (n = 8, 17.4%), insufficient interdisciplinary collaboration

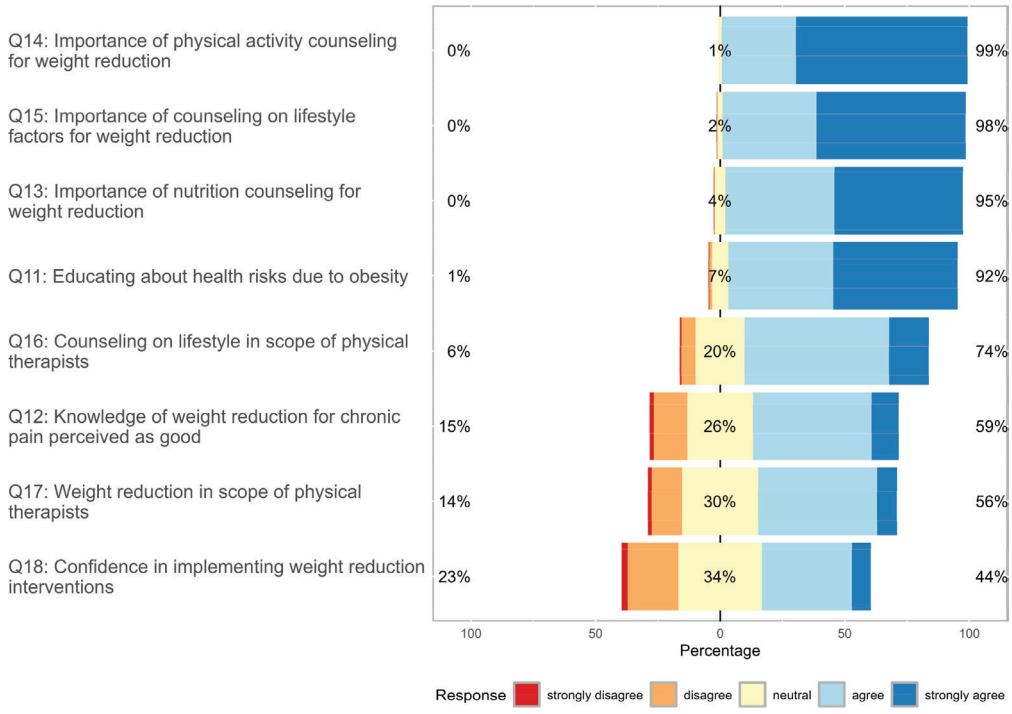


FIGURE 1 - Relative frequencies of physical therapists' attitudes regarding weight reduction in chronic pain patients with obesity. In total, 581 participants answered this question. Q: Question. Questions are displayed in descending order of the level of agreement. The percentages on the left include both "strongly disagree" and "disagree" responses, the percentages on the right include both "agree"and "strongly agree" responses, and the percentage in the middle represents only "neutral" responses.

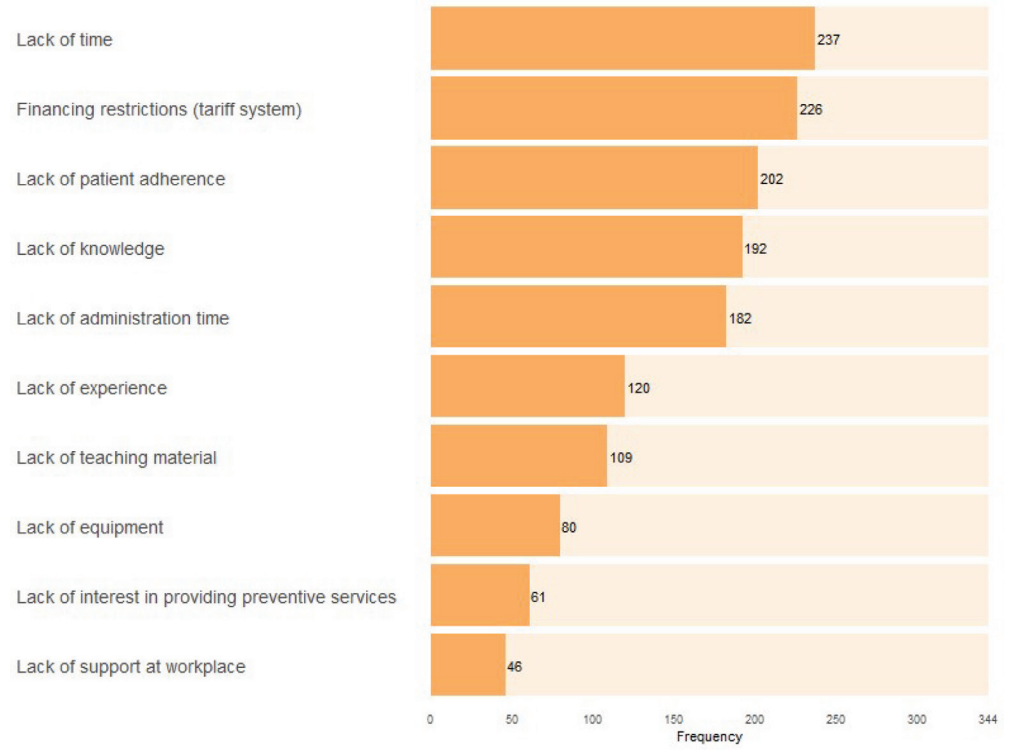


FIGURE 2 - Absolute frequencies of barriers regarding the implementation of weight reduction interventions in physical therapy. In total, 344 participants answered this question (only those which provide weight reduction interventions). Multiple answer selection was possible. Absolute frequencies are displayed for the selected barriers (number).

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

(n = 8, 17.4%), limited resources in physical therapy (e.g., time constraints and lack of interprofessional agreements, n = 6, 13%), and it was perceived as outside the scope of a physical therapist's practice (n = 5, 10.9%).

Concerning practical implementations in physical therapy, a total of 429 out of 526 respondents (81.6%) recommend weight reduction for individuals with chronic pain and overweight or obesity, while 344 out of 526 respondents (65.4%) offer interventions for weight reduction in such patients within their therapy setting (Table 2). All participants were asked how they assessed overweight or obesity in patients (open question). Of the 521 respondents, the majority reported using BMI (n = 287, 55.1%) or observation (n = 276, 53%), and few took circumference measurements (n = 32, 6.1%) or performed bioimpedance analysis (n = 31, 6%). The respondents who already implemented weight reduction interventions in individuals with chronic pain and overweight or obesity were asked which interventions they use in practice (open question). Out of 343 responses, the most frequent answer was a recommendation for physical activity

TABLE 2 - Frequencies of questions about practices and education

Question	Number of respondents	Answers	n (%)
Practices			
Q 24: Assessment methods to capture overweight and obesity in practice*	521	BMI	287 (55.1)
		Observation	276 (53)
		Weight/Height	75 (14.4)
		Anamnesis	60 (11.5)
		Diagnosis available	58 (11.1)
		Circumferences	32 (6.1)
		Bioimpedance analysis	31 (6)
Q 25: Recommendation of weight reduction as part of treatment or in general	526	Yes:	429 (81.6)
Q 26: Implementation of weight reduction interventions within physical therapy	526	Yes:	344 (65.4)
Q 27: Reasons for not implementing weight reduction interventions within physical therapy*	178**	Lack of competence	64 (36)
		Out of scope	38 (21.3)
		Lack of time	29 (16.3)
		No priority	25 (14)
		Lack of interdisciplinary setting	19 (10.7)
Q 28: Interventions to facilitate weight reduction within physical therapy*	343**	Move more	200 (58.3)
		Endurance training	95 (27.7)
		Strength training	84 (24.5)
		Diet advice	71 (20.7)
		Refer to dietician	55 (16)
		Education about diet, training and obesity	36 (10.5)
		Lifestyle change	34 (9.9)
Education			
Q 37: Lifestyle factors covered in entry-level education	501	Yes:	287 (57.3%)
Q 38: Quality rating of question 37.	287**	Very good:	6 (2.1)
		Good:	52 (18.1)
		Neutral:	84 (29.3)
		Fair:	117 (40.8)
		Poor:	28 (9.8)
Q 41: Completion of further education on lifestyle interventions, weight reduction or nutritional management	501	Yes:	107 (21.4)
Q 43: Willingness to participate in further training	394**	Yes:	301 (76.4)

*For content analysis results, only the most frequent categories are listed. Results may exceed 100% as several categories could be named; **Question not shown to all participants due to branching.

(n = 200, 58.3%), including general advice to move more during the day, to implement training, or enhance physical activity to increase energy expenditure. Advising endurance and strength training was reported 95 times (27.7%) and 84 times (24.5%), respectively. Nutritional advice was implemented by 71 respondents (20.7%), while 36 respondents (10.5%) reported giving general information on diet, training, and its association with overweight. Fifty-five of the 343 respondents (16%) recommended a dietitian for dietary changes to reach weight reduction, and only 34 of the respondents (9.9%) suggested other lifestyle changes (e.g., sleep, stress, smoking, social connections).

Further, self-reported reasons for not implementing weight reduction interventions can be found in Table 2.

Regarding the education of the respondents, a total of 287 out of 501 (57.3%) reported that lifestyle factors were covered during their entry-level education (Table 2). Most respondents rated the quality of the content as neutral (n = 84, 29.3%) and fair (n = 117, 40.8%). Further, 107 out of 501 respondents (21.4%) completed further training in lifestyle interventions, weight reduction, or nutrition management. Of the 394 respondents who did not complete further education, 301 (76.4%) would be willing to do so in the future. An infographic summarizing the main descriptive results of the study is provided in Supplement 5.

Table 3 depicts the multiple logistic regression models for recommending and implementing weight reduction interventions by physical therapists in Switzerland. Perceiving weight reduction interventions as within the scope of physical therapists was associated with higher odds of recommending weight reduction to individuals suffering from chronic pain and overweight or obesity (OR: 3.6, 95% CI: 1.9-6.6, Table 3A), compared to those who were neutral or disagreed with this perception. Additionally, respondents who completed further education on lifestyle interventions, weight reduction, or nutrition management were 2.1 times more likely to recommend weight reduction to patients (OR: 2.1, 95% CI: 1.1-4.5), compared to those without further education.

Respondents who agreed that weight reduction interventions are within the scope of physical therapists had 6.5 times higher odds (95% CI: 3.5-12.3), and those who were neutral to this statement had 2.7 times higher odds (95% CI: 1.4-5.1) of offering weight reduction interventions, compared to those who disagreed with this statement (Table 3B). In addition, respondents who agreed that counseling on lifestyle factors falls within their scope of practice reported an OR of 2.9 (95% CI: 1.2-7.3), and respondents who were neutral showed an OR of 3.3 (95% CI: 1.3-8.9), compared to their peers who disagreed with this statement. Moreover, receiving training about lifestyle factors during entry-level

TABLE 3 - Logistic regression model of factors associated with (A) the recommendation of weight reduction, (B) offering weight reduction interventions to individuals with chronic pain and overweight or obesity

	A			
Variable	**Model[a]**		**Model[b]**	
	OR (95% CI)	***p*-value**	**OR (95% CI)**	***p*-value**
(Intercept)	1.7 (0.9-3.1)	0.080	1.7 (1.0-2.9)	0.038*
Weight reduction intervention in the scope (agree)	3.6 (1.9-6.6)	<0.001*	3.6 (1.9-6.6)	<0.001*
Weight reduction intervention in the scope (neutral)	1.8 (1.0-3.4)	0.060	1.8 (1.0-3.4)	0.059
Further education (yes)	2.1 (1.1-4.5)	0.029*	2.1 (1.1-4.4)	0.028*
Gender (m)	1.0 (0.6-1.9)	0.984		
Work experience	1.0 (1.0-1.0)	0.991		
	B			
Variable	**Model[a]**		**Model[b]**	
	OR (95% CI)	**p-value**	**OR (95% CI)**	**p-value**
(Intercept)	0.1 (0.0-0.3)	<0.001*	0.1 (0.0-0.3)	<0.001*
Weight reduction intervention in the scope (agree)	6.5 (3.5-12.3)	<0.001*	6.6 (3.6-12.6)	<0.001*
Weight reduction intervention in the scope (neutral)	2.7 (1.4-5.1)	0.002*	2.7 (1.5-5.1)	0.002*
Lifestyle counseling in the scope (agree)	2.9 (1.2-7.3)	0.015*	2.9 (1.3-7.2)	0.015*
Lifestyle counseling in the scope (neutral)	3.3 (1.3-8.9)	0.012*	3.3 (1.3-8.7)	0.013*
Lifestyle topic during entry-level education (yes)	2.1 (1.4-3.3)	0.001*	2.2 (1.5-3.4)	<0.001*
Further education (yes)	1.9 (1.1-3.3)	0.018*	1.9 (1.1-3.2)	0.022*
Gender (m)	1.4 (0.8-2.4)	0.262		
Work experience	1.0 (1.0-1.0)	0.451		

Model[a]: adjusted model; Model[b]: unadjusted model; *: Indicates statistical significance at p < 0.05; The intercept of regression A contains the factors weight reduction intervention in the scope (disagree), further education (no), gender (female) and work experience (0 years), the intercept of regression B contains the factors weight reduction intervention in the scope (disagree), lifestyle counseling in the scope (disagree), lifestyle topic during entry-level education (no), further education (no), gender (female) and work experience (0 years); Knowledge about weight reduction interventions, highest degree and work setting were excluded during the stepwise backward elimination.
Abbreviations: m: male; OR: odds ratio; 95% CI: 95% confidence interval.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

education, and further education on lifestyle interventions, weight reduction, or nutrition management was associated with offering weight reduction interventions by physical therapists (OR: 2.1, 95% CI: 1.4-3.3 and OR: 1.9, 95% CI: 1.1-3.3, respectively). Controlling for gender and work experience did not substantially influence the main effects (Table 3 Model[b]).

Discussion

The aim of this cross-sectional survey was to assess attitudes, barriers, and practices faced by physical therapists working in the German-speaking part of Switzerland regarding weight management in individuals with chronic pain and overweight or obesity.

Lifestyle interventions, including nutritional and exercise therapy, build the cornerstone of overweight and obesity care, with the overarching goal of health gain (i.e., improved physical function, quality of life, metabolic markers and mental health) (5,43). Weight reduction represents only one marker of treatment effectiveness and is often associated with intensive therapy programs, which are typically followed by weight regain after completion (5,44,45). Consequently, strategies for long-term management are warranted to support sustained improvements in overall health (5,6,43). Additional treatments, such as pharmacotherapy and bariatric surgery, may be considered as part of obesity management and have demonstrated greater weight reduction compared to lifestyle interventions or placebo (5,46,47). However, their use depends on individual patient characteristics and clinical context, and both approaches carry a risk of adverse events (5). Weight regain is commonly observed after discontinuation of pharmacotherapy (48) or after bariatric surgery (49), and neither approach replaces the need for sustained lifestyle modification as the foundation of obesity care (5,47). A meta-analysis including 52 studies showed an average loss of 8 kg of FFM and lean body mass (LBM) within the first year after bariatric surgery, accounting for 21% and 22% of total weight loss, respectively (50). Additionally, nutritional deficiencies were observed following different surgery procedures (i.e., sleeve gastrectomy, Roux-en-Y gastric bypass, gastric band), due to surgery-related malabsorption, pre-existing deficits, and limited access or adherence to supplementation (51). These findings emphasize the importance of promoting general healthy lifestyle behaviors, including balanced nutrition, regular physical activity, stress management, and regenerating sleep, which provide benefits beyond weight loss by improving overall health, metabolic function, and quality of life, while minimizing the risk of nutritional deficits and loss of lean mass (5). Recent evidence demonstrates that physical therapists can play a meaningful role in this interdisciplinary framework. For instance, Allison et al., showed that physical therapists can safely and effectively deliver structured weight reduction interventions, including exercise programs combined with meal replacements, in patients with knee OA and overweight or obesity (52). However, external validity is limited, as the translation of such intensive, resource-supported interventions into routine clinical practice remains challenging due to constraints in time, staffing, and variability in professional training. Nevertheless, physical therapists are often the first point of contact for individuals with chronic musculoskeletal pain and coexisting overweight or obesity (7,10). Their expertise in exercise prescription, movement behavior change, and functional rehabilitation positions them to support sustainable lifestyle modification as part of a multidisciplinary team (38). Within this context, the present survey provides insights into how physical therapists in the German-speaking region of Switzerland perceive their role in weight management for individuals with chronic pain and overweight or obesity. The findings indicate that physical therapists in this region generally acknowledge the importance of educating patients about the health risks associated with overweight and obesity (agreement over 90%) and perceive lifestyle counseling (74%) and weight reduction interventions (56%) as within physical therapists' scope. These findings align with existing surveys from the United States, Canada, and Australia, which indicated that 67 to 81.5% of surveyed physical therapists believe that weight management is within their professional scope when treating individuals with overweight or obesity (25,26). Moreover, education on obesity and its associated health implications, and counseling regarding a healthy lifestyle have been recognized in 61.2 to 90% of surveyed physical therapists in those countries (26,28). However, confidence in implementing weight management into clinical practice when working with individuals with overweight or obesity varied among surveyed physical therapists. Only 20.4% of respondents in the United States and 49% in Australia reported feeling confident, which is comparable to the 44% reported in the current survey in Switzerland (28,39).

Physical therapists in Switzerland and their international counterparts face similar barriers to implement weight management, including systemic barriers like the lack of access to equipment and time constraints, patient barriers including a lack of motivation to adhere to the therapy or that weight management is not perceived as a necessary goal, and therapist barriers including limited training and knowledge to manage overweight or obesity, lack of experience, perceiving weight reduction is not within the scope of physical therapists, and lack of interest in providing preventive services (24,26,27,39). Although lifestyle factors, including weight management, are integrated into modern physical therapy curricula in several countries (as mentioned in the introduction) (22), barriers remain regarding the translation of this knowledge into clinical practice.

Physical therapists in the German-speaking part of Switzerland mainly assessed overweight or obesity using BMI (55.1%) and observation (53%). This can be problematic, as observation is not an objective measurement and BMI does not account for body fat distribution (e.g., waist circumference or waist-to-hip ratio), which a recent expert commission report recommends using for the assessment of obesity (6,53). Moreover, a survey of 851 Canadian physical therapists found no clear consensus on their role in adiposity assessment, with 39% of respondents expressing neutrality and 18% disagreeing that physical therapists should assess adiposity in their patients, indicating uncertainty in this area (26).

In this present study, implemented interventions were primarily focused on increasing physical activity (58.3%), despite 57.3% receiving training in lifestyle factors during their entry-level education and 21.4% completing further education in this field. Comparable research findings from surveys of physical therapists have shown that the strongest agreement regarding physical therapy roles in obesity management was observed in areas such as exercise programs (90-96%), cardiorespiratory training programs (95%), or equipment description (76%) (24,26). About 19.6 to 40% of the volunteering physical therapists in these studies provided nutritional advice (24,26,28), similar to our survey results where 20.7% of respondents provided advice on diet. These findings indicate that weight reduction strategies among physical therapists focus mainly on physical activity, which is not recommended as a standalone approach in current guidelines (43), highlighting the need for interventions that also integrate dietary and behavioral components.

Regression analyses showed that physical therapists were more likely to implement weight reduction interventions in clinical practice if they had received entry-level or further education on healthy lifestyle factors and perceived lifestyle counseling and weight reduction as part of their professional scope. This observation may suggest that, beyond entry-level education and further education, the conviction that weight reduction interventions and lifestyle counseling belong to the scope of physical therapists is crucial for greater engagement and implementation into clinical practice. These findings underscore a gap between awareness and action in physical therapy practice, as observed in both our survey and similar studies conducted in other countries (27,39). While respondents acknowledged the importance of weight reduction and provided recommendations to patients, this did not automatically lead to the implementation of weight reduction interventions (27,39). This observation suggests that, alongside the required education and practical competencies, the personal conviction of physical therapists is essential for the implementation of weight reduction interventions in practice (25).

In our analysis, clinical experience was not associated with the recommendation or implementation of weight reduction interventions. However, a survey conducted in the United States in 2009, including 345 physical therapists, reported that older therapists with longer practical experience had less knowledge about obesity compared to younger therapists (28). Another survey conducted in Australia among 65 physical therapists reported that the provision of dietary advice was significantly associated with weight management training during professional entry-level education (25). A total of 57.3% of physical therapists in our survey reported receiving entry-level education on lifestyle factors, suggesting a shift toward more lifestyle-focused curricula. Unlike the findings from the 2009 U.S. survey, where older physical therapists demonstrated less knowledge of obesity, our results suggest that increased education improves the likelihood of physical therapists implementing weight management interventions into clinical practice, similar to the findings in the Australian survey.

However, evidence-based implementation of weight reduction interventions, including lifestyle changes, diet, and increased physical activity (54), by physical therapists remains limited, as indicated by the responses to open-ended questions from this survey and existing literature. This limitation is mainly attributed to the focus on physical activity-related interventions and inadequate assessment of obesity (24,26). Despite the growing relevance of counseling and coaching on lifestyle-focused modifications by physical therapists, knowledge and skills for effective implementation remain challenging (55,56). Results from a randomized controlled trial including 80 physical therapists, randomly allocated to an intervention or control group, demonstrated an improvement in physical therapists' confidence, knowledge, and skills in weight management for people with OA in the intervention group after an online self-directed training program on biopsychosocial elements of obesity and weight management (57). Moreover, studies from the United States and Canada support the effect of a weight reduction program on improvement in pain and physical function, including different health care professionals like physicians, exercise specialists, dieticians, and nurses (15,58,59). Thus, showing promising evidence for the potential role of physical therapists in lifestyle and weight management strategies, though an interdisciplinary approach may be necessary. An advanced practice role of physical therapists incorporating weight management alongside the management of different chronic MSD could represent an important evidence-based therapy approach in the future. To integrate weight management into daily physical therapy practice, education and healthcare policies must evolve, as barriers and inconsistencies in entry-level curricula persist across countries (22,26,27,39,60). Entry-level education should incorporate training in lifestyle counseling, with a specific focus on weight management, to ensure that future physical therapists develop the necessary skills and confidence to address overweight, obesity, and related conditions like chronic MSD (20,61). Moreover, healthcare systems should recognize and support the role of physical therapists in lifestyle interventions by addressing key barriers such as financial constraints and limited administrative time.

Strengths and limitations

One strength of this study is the large sample size, which exceeds the required 360 participants, aligning with findings in the literature (62). The snowball sampling method enabled fast and low-cost data collection. However, due to this non-probability sampling method, a potential risk of sampling bias cannot be totally excluded. Since respondents are likely to refer the survey to other individuals within their social networks, the sample may not be fully representative of all physical therapists. Furthermore, as the survey was conducted in the German-speaking part of Switzerland only, the generalizability is limited to this region, as the prevalence of overweight and obesity is higher among German-speaking residents compared to French- or Italian-speaking residents (63). Additionally, cross-sectional designs prevent causal conclusions from the logistic regression analyses.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

The survey did not capture the detailed educational status of students (e.g., second, third or fourth year), which could not be controlled in the analysis. However, excluding students (n = 45) from the analysis did not change overall patterns and only resulted in negligible differences. Nevertheless, more accurate quantification of professional experience and educational status may enhance future research in this field. Furthermore, self-reported outcomes might be affected by recall bias, particularly in questions related to entry-level education, where memory-related imprecisions could result in biased responses. Overall, the results need to be interpreted with caution, as the survey was not fully validated at all stages. In addition, the survey focused on lifestyle interventions and primarily weight reduction within physical therapy care, and did not fully reflect the modern paradigm of overweight and obesity management, which emphasizes the improvement in overall health, function and quality of life in addition to weight change (5,43).

Conclusions

This cross-sectional survey examined the attitudes, perceived barriers, and current practices of physical therapists in the German-speaking region of Switzerland regarding weight management for individuals with chronic pain and overweight or obesity. Of the surveyed physical therapists, 74% believed that lifestyle counseling for individuals with chronic pain and overweight or obesity was within their scope. Furthermore, 56% viewed weight reduction interventions as part of their role, and 65.4% reported implementing weight reduction interventions in clinical practice, predominantly through physical activity-related approaches. Common barriers, such as financing restrictions, limited administrative time, and lack of interdisciplinary collaboration, indicate the need for systemic changes to facilitate the integration of weight management into physical therapy practice. Along with entry-level and further education, respondents' perception that lifestyle counseling and weight reduction interventions fall within the scope of physical therapists emerged as the strongest predictors of the implementation of weight reduction interventions in clinical practice. To date, the extent of knowledge and competence level remains unclear, highlighting the need for competency standards in weight management for physical therapists and evidence-based education programs to develop management guidelines and evaluate the effectiveness of weight reduction interventions for individuals with both chronic pain and comorbid overweight or obesity in the long term.

Acknowledgments

The authors would like to thank all participants and researchers for their contribution to this study.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used ChatGPT to improve clarity and language refinement. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Institutional review board statement

The study was conducted according to the Declaration of Helsinki (64).

Disclosures

Conflict of interest: The authors declare no conflict of interest

Financial support: This work was supported by the Swiss National Science Foundation (SNF) (ref. nr. 204659) in the context of the PhD funding of ML and APS and the PhD funding of MV from the Research Foundation Flanders (ref.nr. G0894322). Otherwise, no external funding was received.

Author contributions: ML: conceptualization, methodology, software, data collection, formal analysis, data curation, writing original draft, visualization, supervision of project progress; MM: conceptualization, methodology, software, data collection, writing reviewing & editing; APS: conceptualization, validation, writing reviewing & editing; MV: writing reviewing & editing; JT: conceptualization, methodology, supervision of project progress, project administration, validation, writing reviewing & editing; HB: data collection, resources, writing reviewing & editing; RC: data collection, resources, writing reviewing & editing; DF: validation, writing reviewing & editing; AM: writing reviewing & editing; NL: analysis, writing reviewing & editing

Data availability statement: Data is available from the corresponding author upon reasonable request. The data is not publicly available due to consent limitations, as participants did not agree to public data sharing.

References

1. World Obesity Federation. World Obesity Atlas 2024. 2024. Online (Accessed May 2025)
2. Peiris WL, Cicuttini FM, Hussain SM, et al. Is adiposity associated with back and lower limb pain? A systematic review. *PLoS One*. 2021;16(9):e0256720. CrossRef PubMed
3. Narouze S, Souzdalnitski D. Obesity and chronic pain: opportunities for better patient care. *Pain Manag*. 2015;5(4):217-219. CrossRef PubMed
4. Boutari C, DeMarsilis A, Mantzoros CS. Obesity and diabetes. *Diabetes Res Clin Pract*. 2023;202:110773. CrossRef PubMed
5. Lingvay I, Cohen RV, Roux CWL, et al. Obesity in adults. *Lancet*. 2024;404(10456):972-987. CrossRef PubMed
6. Rubino F, Cummings DE, Eckel RH, et al. Definition and diagnostic criteria of clinical obesity. *Lancet Diabetes Endocrinol*. 2025;13(3):221-262. CrossRef PubMed
7. Nicol V, Verdaguer C, Daste C, et al. Chronic low back pain: a narrative review of recent international guidelines for diagnosis and conservative treatment. *J Clin Med*. 2023;12(4):1685. CrossRef PubMed
8. Airaksinen O, Brox JI, Cedraschi C, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *Eur Spine Journal*. Mar 2006;15 Suppl 2(Suppl 2): S192-300. CrossRef
9. Moseng T, Vliet Vlieland TPM, Battista S, et al. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis: 2023 update. *Ann Rheum Dis*. 2024;83(6):730-740. CrossRef PubMed

10. De la Corte-Rodriguez H, Roman-Belmonte JM, Resino-Luis C, et al. The role of physical exercise in chronic musculoskeletal pain: best medicine-a narrative review. *Healthcare (Basel)*. 2024;12(2):242. CrossRef PubMed
11. Flynn DM. Chronic musculoskeletal pain: nonpharmacologic, noninvasive treatments. *Am Fam Physician*. 2020;102(8): 465-477. PubMed
12. Dean E, Söderlund A. What is the role of lifestyle behaviour change associated with non-communicable disease risk in managing musculoskeletal health conditions with special reference to chronic pain? *BMC Musculoskelet Disord*. 2015;16(1):87. CrossRef PubMed
13. Cooper L, Ryan CG, Ells LJ, et al. Weight loss interventions for adults with overweight/obesity and chronic musculoskeletal pain: a mixed methods systematic review. *Obes Rev*. 2018;19(7):989-1007. CrossRef PubMed
14. Chen LH, Weber K, Mehrabkhani S, et al. The effectiveness of weight loss programs for low back pain: a systematic review. *BMC Musculoskelet Disord*. 2022;23(1):488. CrossRef PubMed
15. Messier SP, Mihalko SL, Legault C, et al. Effects of intensive diet and exercise on knee joint loads, inflammation, and clinical outcomes among overweight and obese adults with knee osteoarthritis: the IDEA randomized clinical trial. *JAMA*. 2013;310(12):1263-1273. CrossRef PubMed
16. Bosch-Sierra N, Grau-Del Valle C, Hermenejildo J, et al. The impact of weight loss on inflammation, oxidative stress, and mitochondrial function in subjects with obesity. *Antioxidants*. 2024;13(7):870. CrossRef PubMed
17. Zhu B, Gostoli S, Benasi G, et al. The role of psychological well-being in weight loss: new insights from a comprehensive lifestyle intervention. *Int J Clin Health Psychol*. 2022;22(1):100279. CrossRef PubMed
18. Martins C, Sayegh S, Faundez A, et al. Effectiveness of a group-based rehabilitation program combining education with multimodal exercises in the treatment of patients with nonspecific chronic low back pain: a retrospective uncontrolled study. *Biology (Basel)*. 2022;11(10):1508. CrossRef PubMed
19. Allen SA, Dal Grande E, Abernethy AP, et al. Two colliding epidemics - obesity is independently associated with chronic pain interfering with activities of daily living in adults 18 years and over; a cross-sectional, population-based study. *BMC Public Health*. 2016;16(1):1034. CrossRef PubMed
20. Tatta J, Nijs J, Elma Ö, et al. The critical role of nutrition care to improve pain management: a global call to action for physical therapist practice. *Phys Ther*. 2022;102(4):pzab296. CrossRef PubMed
21. World Physiotherapy. *Description of physical therapy policy statement*. 2019. Online (Accessed May 2025)
22. Bodner ME, Rhodes RE, Miller WC, et al. Benchmarking curriculum content in entry-level health professional education with special reference to health promotion practice in physical therapy: a multi-institutional international study. *Adv Health Sci Educ Theory Pract*. 2013;18(4):645-657. CrossRef PubMed
23. Alodaibi FA, Alotaibi MA, Almohiza MA, et al. Physical therapists' role in health and wellness promotion for people with musculoskeletal disorders: a cross-sectional description study conducted in Saudi Arabia. *J Multidiscip Healthc*. 2022;15: 567-576. CrossRef PubMed
24. O'Donoghue G, Cunningham C, Murphy F, et al. Assessment and management of risk factors for the prevention of lifestyle-related disease: a cross-sectional survey of current activities, barriers and perceived training needs of primary care physiotherapists in the Republic of Ireland. *Physiotherapy*. 2014;100(2):116-122. CrossRef PubMed
25. Snodgrass SJ, Carter AE, Guest M, et al. Weight management including dietary and physical activity advice provided by Australian physiotherapists: a pilot cross-sectional survey. *Physiother Theory Pract*. 2014;30(6):409-420. CrossRef PubMed
26. You L, Sadler G, Majumdar S, et al. Physiotherapists' perceptions of their role in the rehabilitation management of individuals with obesity. *Physiother Can*. 2012;64(2):168-175. CrossRef PubMed
27. Abaraogu UO, Ogaga MO, Odidika E, et al. Promotion of healthy nutrition in clinical practice: A cross-sectional survey of practices and barriers among physiotherapists in south-east Nigeria. *Hong Kong Physiother J*. 2016;35:21-29. CrossRef PubMed
28. Sack S, Radler DR, Mairella KK, et al. Physical therapists' attitudes, knowledge, and practice approaches regarding people who are obese. *Phys Ther*. 2009;89(8):804-815. CrossRef PubMed
29. LimeSurvey GmbH. LimeSurvey. An open source survey tool, LimeSurvey GmbH, Hamburg, Germany. Online (Accessed May 2025)
30. Vandenbroucke JP, von Elm E, Altman DG, et al.; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Epidemiology*. 2007;18(6):805-835. CrossRef PubMed
31. Duc M, Mittaz Hager AG, Zemp D, et al. Current practices of physiotherapists in Switzerland regarding fall risk-assessment for community-dwelling older adults: a national cross-sectional survey. *F1000 Res*. 2023;11:513. CrossRef PubMed
32. Ritschl Valentin, Weigl Roman, Tanja S. Wissenschaftliches Arbeiten und Schreiben - Verstehen, Anwenden, Nutzen für die Praxis. Springer Verlag; 2016. CrossRef
33. Federal Statistical Office. Main languages, 1970-2022 Permanent resident population. Online (Accessed May 2025)
34. Zukünftige ambulante Grundversorgung: Einstellungen und Präferenzen von Medizinal- und Gesundheitsfachpersonen ausgewählter Berufsgruppen (Obsan Bericht 06/2023). Schweizerisches Gesundheitsobservatorium; 2023. access date: 18.11.24. Online
35. Haynes A, Lenz A, Stalder O, et al. Presize: an R-package for precision-based sample size calculation in clinical research. *J Open Source Softw*. 2021;6(60):3118. CrossRef
36. Campbell K, Crawford D. Management of obesity: attitudes and practices of Australian dietitians. *Int J Obes (Lond)*. 2000;24(6):701-710. CrossRef PubMed
37. Laws R; Counterweight Project Team. Current approaches to obesity management in UK Primary Care: the Counterweight Programme. *J Hum Nutr Diet*. 2004;17(3):183-190. CrossRef PubMed
38. Berner P, Bezner JR, Morris D, et al. Nutrition in physical therapist practice: tools and strategies to act now. *Phys Ther*. 2021;101(5): pzab061. CrossRef PubMed
39. Stevens J, Orrock P, Verco B, et al. Overweight and obesity management in musculoskeletal primary care: a survey of Australian health practitioners exploring the elephant in the room. *Aust J Prim Health*. 2022;28(6):573-579. CrossRef PubMed
40. Foster GD, Wadden TA, Makris AP, et al. Primary care physicians' attitudes about obesity and its treatment. *Obes Res*. 2003;11(10): 1168-1177. CrossRef PubMed
41. Rampin R, Rampin V. Taguette: open-source qualitative data analysis. *J Open Source Softw*. 2021;6(68):3522. CrossRef
42. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2025. Online (Accessed May 2025)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

43. (NICE) NIfHaCE. Overweight and obesity management - NICE guideline. Guideline. 2025. 14.01. Online (Accessed May 2025)
44. Goessl CL, Befort CA, Pathak RD, et al. Chronic pain and weight regain in a lifestyle modification trial. *Obes Sci Pract*. 2021;7(2):192-198. CrossRef PubMed
45. Machado AM, Guimarães NS, Bocardi VB, et al. Understanding weight regain after a nutritional weight loss intervention: systematic review and meta-analysis. *Clin Nutr ESPEN*. 2022;49:138-153. CrossRef PubMed
46. Wong HJ, Sim B, Teo YH, et al. Efficacy of GLP-1 receptor agonists on weight loss, BMI, and waist circumference for patients with obesity or overweight: a systematic review, meta-analysis, and meta-regression of 47 randomized controlled trials. *Diabetes Care*. 2025;48(2):292-300. CrossRef PubMed
47. Perdomo CM, Cohen RV, Sumithran P, et al. Contemporary medical, device, and surgical therapies for obesity in adults. *Lancet*. 2023;401(10382):1116-1130. CrossRef PubMed
48. Wilding JPH, Batterham RL, Davies M, et al.; STEP 1 Study Group. Weight regain and cardiometabolic effects after withdrawal of semaglutide: the STEP 1 trial extension. *Diabetes Obes Metab*. 2022;24(8):1553-1564. CrossRef PubMed
49. El Ansari W, Elhag W. Weight regain and insufficient weight loss after bariatric surgery: definitions, prevalence, mechanisms, predictors, prevention and management strategies, and knowledge gaps-a scoping review. *Obes Surg*. 2021;31(4):1755-1766. CrossRef PubMed
50. Nuijten MAH, Eijsvogels TMH, Monpellier VM, et al. The magnitude and progress of lean body mass, fat-free mass, and skeletal muscle mass loss following bariatric surgery: a systematic review and meta-analysis. *Obes Rev*. 2022;23(1):e13370. CrossRef PubMed
51. O'Kane M, Parretti HM, Pinkney J, et al. British Obesity and Metabolic Surgery Society Guidelines on perioperative and postoperative biochemical monitoring and micronutrient replacement for patients undergoing bariatric surgery-2020 update. *Obes Rev*. 2020;21(11):e13087. CrossRef PubMed
52. Allison K, Jones S, Hinman RS, et al. Alternative models to support weight loss in chronic musculoskeletal conditions: effectiveness of a physiotherapist-delivered intensive diet programme for knee osteoarthritis, the POWER randomized controlled trial. *Br J Sports Med*. 2024;58(10):538-547. CrossRef PubMed
53. Dario AB, Ferreira ML, Refshauge K, et al. Are obesity and body fat distribution associated with low back pain in women? A population-based study of 1128 Spanish twins. *Eur Spine J*. 2016;25(4):1188-1195. CrossRef PubMed
54. Bray GA, Frühbeck G, Ryan DH, et al. Management of obesity. *Lancet*. 2016;387(10031):1947-1956. CrossRef PubMed
55. Bezner JR. Promoting health and wellness: implications for physical therapist practice. *Phys Ther*. 2015;95(10):1433-1444. CrossRef PubMed
56. Rethorn ZD, Pettitt CD. What is the effect of health coaching delivered by physical therapists? A systematic review of randomized controlled trials. *Phys Ther*. 2019;99(10):1354-1370. CrossRef PubMed
57. Allison K, Jones S, Hinman RS, et al. Effects of an online education program on physical therapists' confidence in weight management for people with osteoarthritis: a randomized controlled trial. *Arthritis Care Res (Hoboken)*. 2023;75(4):835-847. CrossRef PubMed
58. Roffey DM, Ashdown LC, Dornan HD, et al. Pilot evaluation of a multidisciplinary, medically supervised, nonsurgical weight loss program on the severity of low back pain in obese adults. *Spine J*. 2011;11(3):197-204. CrossRef PubMed
59. Foy CG, Lewis CE, Hairston KG, et al.; Look AHEAD Research Group. Intensive lifestyle intervention improves physical function among obese adults with knee pain: findings from the Look AHEAD trial. *Obesity (Silver Spring)*. 2011;19(1):83-93. CrossRef PubMed
60. Bouma SE, van Beek JFE, Diercks RL, et al. Barriers and facilitators perceived by healthcare professionals for implementing lifestyle interventions in patients with osteoarthritis: a scoping review. *BMJ Open*. 2022;12(2):e056831. CrossRef PubMed
61. Dean E, de Andrade AD, O'Donoghue G, et al. The Second Physical Therapy Summit on Global Health: developing an action plan to promote health in daily practice and reduce the burden of non-communicable diseases. *Physiother Theory Pract*. 2014;30(4):261-275. CrossRef PubMed
62. Adam AM. Sample size determination in survey research. *J Sci Res Rep*. 2020;26 (5):90–97. CrossRef
63. Bundesamt für Statistik BFS. Übergewicht und Adipositas BFS Bericht 2022. 2024; Access date 18.11.25: Online
64. World Medical Association. WMA Declaration of Helsinki – Ethical principles for medical research involving human subjects. Updated 30.04.25. Online (Accessed May 2025)

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Arch Physioter 2025; 15: 297-313
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3533
REVIEW

Is resistance training with external loads superior to unloaded exercise in the management of chronic low back pain? A systematic review and meta-analysis

Marco Ranzani[1], **Andrea Pozzi**[2], **Daniele Fornasari**[2], **Diego Ristori**[2], **Marco Testa**[2]

[1] Casa di Cura Igea – Department of Neurorehabilitation Sciences, Milan - Italy
[2] Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genova, Campus of Savona, Savona - Italy

ABSTRACT

Introduction: Chronic non-specific low back pain is a leading cause of disability worldwide. While resistance training using external loads is common in rehabilitation, its added value over unloaded exercise remains uncertain, particularly across physical and psychological variables.

Method: This systematic review and meta-analysis, registered on PROSPERO (CRD42022366975), included randomized controlled trials comparing externally loaded resistance training to unloaded exercise in adults with chronic non-specific low back pain. Primary outcomes were pain intensity and disability. Secondary outcomes included back muscle endurance, maximal strength, fear-avoidance beliefs, and pain catastrophizing. Random-effects meta-analyses were conducted, stratified by follow-up duration.

Results: Thirteen randomized trials (778 participants) were included. At follow-up periods beyond seven weeks, externally loaded resistance training showed a small but statistically significant reduction in pain compared to unloaded exercise (mean difference = –0.52 on a 0-10 scale; 95% confidence interval [–0.92, –0.08]). No significant differences were found at short-term or post-washout follow-ups. Effects on disability were inconsistent and highly variable. Resistance training was associated with improvements in back muscle endurance and suggested a possible effect on long-term maximal strength, although wide prediction intervals prevent definitive conclusions. No meaningful differences were found for psychological variables, and pain catastrophizing was assessed in only one trial, limiting conclusions.

Conclusion: Externally loaded resistance training is safe and feasible for chronic non-specific low back pain, but its effects on pain, disability and psychosocial outcomes are comparable to unloaded exercise. In line with the multifactorial nature of chronic pain, improvements appear driven more by exposure, adherence and therapeutic context than by load intensity alone. Exercise prescription should therefore remain individualized and embedded within a biopsychosocial framework.

Keywords: Low back pain, Chronic, Exercise therapy, Resistance training, Weight training, Load

What's already known about this topic?

- *RT can enhance both strength and endurance. The specific role of load as a variable in RT for managing chronic NS-LBP, particularly its impact on pain and disability, remains not fully established.*

What does the study add?

- *While loaded exercises induce greater neuromuscular adaptations, they do not provide better improvements in pain or disability compared to unloaded exercises in individuals with chronic NS-LBP. Exercise volume and adherence may play a more significant role in symptom management.*

Received: April 16, 2025
Accepted: November 3, 2025
Published online: December 5, 2025

This article includes supplementary material

Corresponding author:
Marco Ranzani
email: m.ranzani@casadicuraigea.it

Introduction

Low back pain (LBP) is the most prevalent musculoskeletal disorder (1), affecting approximately 540 million people worldwide at any given time (2). When no specific patho-anatomical cause can be identified, LBP is classified as non-specific (NS-LBP) (3). Among all musculoskeletal conditions, chronic NS-LBP represents the leading cause of disability, as measured

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

in disability-adjusted life-years (4), posing a significant burden on healthcare systems and affected individuals.

Traditionally conceptualized through a biomechanical lens, chronic NS-LBP is now widely recognized as a multifactorial condition in which pain and disability arise from dynamic interactions among biological, psychological, and social factors (5). This paradigm shift has supported the adoption of the biopsychosocial (BPS) model as the best practice for its understanding and management (6).

Exercise is widely recommended as the first-line treatment for chronic NS-LBP, as highlighted by the National Institute for Health and Care Excellence (NICE) guidelines (7).

Beyond its physical effects, exercise can be conceptualized as a multidimensional therapeutic intervention, capable of modulating psychological and behavioral processes (8). Evidence suggests that physical exercise may positively influence factors such as anxiety, mood, fear-avoidance, and pain-related beliefs (9,10), reinforcing its role within the BPS model.

Despite its recognized value, the optimal characteristics of exercise interventions for chronic NS-LBP remain unclear. Current literature shows limited differences in effectiveness between various exercise modalities, with generally modest improvements in pain and disability (11,12). This suggests a need to move beyond broad exercise categories and instead focus on specific intervention components that may drive better outcomes.

One such component is the type and intensity of resistance, or load, used during training. Resistance training (RT) is a form of physical exercise involving internal (e.g., body weight) or external (e.g., free weights, machines, bands) resistance to stimulate skeletal muscle contractions, aiming to enhance strength, power, muscular endurance, and muscle mass (13,14). Although RT is commonly applied in rehabilitation settings, the optimal dosage and resistance parameters for chronic musculoskeletal conditions remain poorly defined (15-17). In research, the term "load" typically refers to external resistance, distinguishing "loaded" from "unloaded" exercises (18).

Loaded exercises may induce adaptations across multiple, interrelated domains. They have been associated with greater gains in muscle strength (19), soft tissue capacity (20), cartilage turnover (21), as well as enhanced neurochemical responses which have been linked to increases in pain thresholds (22,23). Incorporating external load may serve as a graded exposure stimulus, helping patients confront fears, rebuild confidence in movement, and challenge beliefs related to fragility or harm (7). Such mechanisms may be particularly valuable for targeting maladaptive responses such as fear-avoidance, kinesiophobia, and pain catastrophizing, which are known to contribute to the persistence of chronic NS-LBP (24).

Although the addition of external load to an exercise program may influence the adaptations described above, its specific contribution to clinical outcomes remains poorly understood (25,26). Available studies are hindered by inconsistent terminology, insufficiently described loading protocols and inadequate control conditions lacking active comparators (17,27). Therefore, it remains unclear to what extent the inclusion of external resistance affects the multifaceted nature of chronic NS-LBP.

This systematic review and meta-analysis consequently seeks to address the following question: "To what extent does the use of external load in resistance training influence symptoms, function, and psychosocial factors in chronic NS-LBP management?"

The primary objective of this systematic review and meta-analysis was to assess the effects of RT with external resistance compared to unloaded exercises on pain and disability. Secondary outcomes included physical performance metrics (muscle endurance, maximal strength) and psychosocial variables (fear-avoidance, kinesiophobia, and pain catastrophizing).

Methods

This study followed the guidelines outlined in the 'Cochrane Handbook for Systematic Reviews of Interventions' (28) and was structured in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) Statement (29). The PRISMA checklist is reported in 'Appendix A'.

Eligibility criteria

The eligibility criteria were developed based on a research question structured according to the Population–Intervention–Comparison–Outcome (PICO) framework. The population of interest comprised adults (≥18 years) diagnosed with chronic NS-LBP. Eligible Interventions included RT programs that incorporated external loads, either alone or in combination with internal resistance (bodyweight). To isolate the effect of external loads in RT, studies were excluded if the experimental group received multimodal interventions involving additional exercise modalities (e.g., aerobic training, motor control exercises) or therapies (e.g., manual therapy).

For the Comparison group, only exercise interventions without any form of external resistance—hereafter referred to as unloaded exercises (UE)—were considered eligible. Eligible studies had to report on at least one of the primary Outcomes.

Only randomized controlled trials (RCTs) were included. Studies were excluded if they included individuals with specific LBP (e.g., fractures, radicular pain, radiculopathies, spinal stenosis, or axial spondylarthritis) or with a history of spinal surgery. Finally, while studies not available in English or Italian were excluded, no restrictions were applied regarding publication date or methodological quality.

Information sources and search strategy

Two independent reviewers (MR and DF) each developed and conducted a separate literature search, working in a blinded manner without any prior agreement on which databases to search. The search strategies were based on the previously defined PICO framework, although no terms were included for the Comparison component to reduce the risk of missing relevant studies.

As recommended by the Cochrane Handbook (28), three core electronic databases were searched: PubMed, Cochrane

Library, and EMBASE. Filters for "Randomized Controlled Trial" in PubMed and "Trial" in the Cochrane Library were applied to restrict the results to study types relevant to the review's objective. To enhance comprehensiveness, two multidisciplinary databases (Scopus and Web of Science), a physical therapy-specific database (The Physiotherapy Evidence Database - PEDro) and gray literature sources (Google Scholar) were also queried.

The searches were first conducted in September 2022 and subsequently updated in October 2024 and March 2025. Full search strategies, including the specific databases, search terms, and filters used, are provided in 'Appendix B'.

Finally, a planned citation search was also performed by screening the reference lists of previously published systematic reviews investigating the effects of exercise on LBP (11,12,25,27,30-35) to enhance the comprehensiveness of the search strategy.

Selection process

The search results were imported into Rayyan software (Online), where duplicate records were manually removed before screening (36).

Two independent reviewers (MR and DF) conducted a blinded, independent screening of titles and abstracts. Studies meeting the eligibility criteria at this stage were exported to an Excel file, and their full texts were retrieved. The reviewers then independently assessed the full texts in a blinded manner. Any discrepancies during the screening process were resolved through discussion between MR and DF or, if necessary, with the involvement of a third reviewer (AP). Inter-rater agreement was not calculated.

Data collection process and data items

Two independent reviewers (MR and DF) extracted data from each included study. To standardize the process, a structured synoptic Excel spreadsheet was developed, in line with the recommendations of the Cochrane Handbook of Systematic Reviews (28).

The extracted data included: (i) Study characteristics: lead author, publication year, sample size, duration, adverse events and follow-up. (ii) Population details: sex, age, height, weight, BMI, baseline pain, baseline disability, duration of symptoms and baseline level of activity. (iii) Intervention characteristics: exercise type and name, type of resistance, total duration of the intervention, weekly frequency, session duration, number of sets and repetitions, baseline and peak intensity and progression parameters. (iv) Outcomes: mean and standard deviations at every follow-up, as reported in the studies or, when not available, derived from other summary statistics using the conversion methods recommended in the Cochrane Handbook (28).

Outcome measures

Pain intensity and disability were the primary outcomes of the systematic review. Pain intensity was assessed based on the Visual Analog Scale (VAS) and Numeric Pain Rating Scale (NPRS), while disability was evaluated using the Oswestry Disability Index (ODI) and Roland and Morris Disability Questionnaire (RMDQ).

Secondary outcomes included measures of physical performance and psychosocial variables. For physical performance, eligible measures comprised back muscle endurance, as assessed by the Biering-Sørensen test (BST), and maximal strength, expressed in terms of muscle force (newtons, N) or joint torque (newton-meters, N·m). Psychosocial outcomes included fear-avoidance beliefs related to physical activity (Physical Activity subscale of the Fear-Avoidance Beliefs Questionnaire, FABQ-PA), kinesiophobia (Tampa Scale of Kinesiophobia, TSK) or pain catastrophizing (Pain Catastrophizing Scale, PCS). The FABQ-PA subscale was selected as a relevant measure of fear-avoidance beliefs due to its specific focus on movement and exercise (37), its established construct validity (37,38), and its documented association with disability and less favorable clinical outcomes in individuals with chronic NS-LBP (39,40).

Study risk of bias assessment and certainty of evidence grading

Two independent reviewers (MR and DF) conducted a blinded risk of bias assessment for each primary and secondary outcome using the revised version of the Cochrane Risk of Bias Assessment Tool (RoB2) (41). Risk-of-bias plots were generated using the Robvis tool (42). The Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) system was used to assess the certainty of evidence at each follow-up for each outcome. The grading process was conducted using 'GRADEpro' (43).

Effect measures

The effect size used for the meta-analysis was calculated as the standardized mean difference (SMD) with a 95% confidence interval (95%CI) for outcomes reported in non-convertible units or assessed using different rating instruments. When all studies employed the same measurement scale for a given outcome, the mean difference (MD) with a 95%CI was used instead. Pain intensity scores reported on 100-point scales were rescaled to 0–10, and BST times were converted from seconds to minutes to ensure comparability across studies.

To explore treatment effects within each study, we also calculated within-group changes using pre- and post-intervention data (SMDs with 95%CI), where available. In addition, we computed between-group differences at post-intervention using SMDs based on post-treatment means and standard deviations (SDs). These results are presented descriptively in 'Appendix D', along with indications of statistical significance, to facilitate interpretation of individual study findings.

Unit-of-analysis issues in trials with shared comparison groups were addressed by splitting the shared group's sample size evenly, while retaining the original means, SDs, and participant counts, in accordance with the Cochrane Handbook guidelines (28).

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Synthesis methods

To generalize the findings beyond the included studies, an unconditional inference model was employed. The restricted maximum likelihood estimation method (REML) was used to estimate between-study variance in the random-effects model. Post-intervention results were categorized based on the duration of the exercise program. 'Short training programs' (STP) were defined as those with follow-up data collected within 6 weeks or less, while 'extended training programs' (ETP) were defined as those with follow-up data collected after at least 7 weeks. This cutoff was based on evidence suggesting that programs lasting at least 7 weeks may be more effective for managing chronic pain (44). Additionally, results were analyzed, where applicable, at the final follow-up to assess long-term differences between loaded and unloaded interventions. For inclusion in the 'Post-Washout' (PW) follow-up analysis, data had to be collected at least 6 months after the study's initiation, following a washout period (i.e., a treatment-free interval to minimize carryover effects). When multiple eligible follow-ups were reported, data from the most distant follow-up were used. Heterogeneity was assessed using Cochran's Q test (χ^2), I^2 statistic, and tau^2. These metrics were reported descriptively, without applying fixed thresholds for interpretation. In addition, 95% prediction intervals (95%PI) were calculated and used as the primary indicator of the expected dispersion of true effects across comparable future settings (45,46). Meta-analyses were conducted using R version 4.2.3 for Mac OS (47), with the 'meta' package and its "metagen" function (48). The complete analysis output (RStudio report) is publicly available on the Open Science Framework (OSF) at Online.

Post-hoc sub-group and sensitivity analysis

To explore potential sources of heterogeneity in the primary outcomes, post hoc subgroup analyses were performed based on: (i) methodological quality, classified using the RoB2 tool as high or low risk of bias; (ii) baseline pain intensity, categorized according to the Pain Monitoring Model (49) as low (<2/10), acceptable (2–5/10), or high (>5/10); and (iii) the type of resistance used in the intervention, comparing programs using only external resistance versus those combining internal and external resistance. Results of these analyses are presented in 'Appendix F'.

Sensitivity analyses were conducted to assess the robustness of the findings and to examine the influence of individual studies on overall estimates and heterogeneity. These analyses were performed only for meta-analyses that included data from at least five RCTs, to ensure sufficient statistical power and stable estimation of heterogeneity parameters (50). Following Viechtbauer (2010), a leave-one-out approach was applied, iteratively excluding each study from the model and assessing changes in the pooled effect size and heterogeneity parameters (I^2, tau^2, and Q) (50). A study was considered influential if its removal resulted in: (a) a substantial reduction in heterogeneity (≥25% reduction in tau^2 or I^2), or (b) a meaningful change in the direction or magnitude of the pooled effect. Studies meeting one or both of these criteria were excluded in subsequent analyses to evaluate whether they had a disproportionate impact on the overall results. This approach allowed for a cautious and transparent assessment of the stability of the conclusions in the presence of statistical heterogeneity.

Reporting bias assessment

Publication bias was assessed using contour-enhanced funnel plots, which illustrate the relationship between the size of the studies and effect sizes (51). For meta-analyses with at least 10 comparisons (52), Egger's regression test was used as a quantitative measure of reporting bias. The "funnel.meta" and "metabias" (47) functions in RStudio (47) were used to evaluate the reporting bias.

Results

Registration and Protocol

The review protocol was registered in PROSPERO (CRD42022366975), but several amendments were made during the review process. Due to the limited number of available studies and incomplete reporting of intervention data, the planned meta-regression analyses, as specified in the original protocol, could not be performed. Consequently, modifications were made to both the title and statistical analyses in this manuscript.

Study selection

The search strategy identified a total of 1710 records, of which 1609 were retrieved from databases and 101 from citation search. After removing 555 duplicates using Rayyan, 1155 unique records remained. These were screened based on titles and abstracts, resulting in the exclusion of 1082 records for not meeting the inclusion criteria. Overall, 73 records were identified for retrieval. Of those, one (53) could not be retrieved despite attempts to contact the authors. The remaining 72 full-texts were screened for eligibility. Among them, 12 studies identified from database searches (54-65) and 3 from citation searching (66-68) met the eligibility criteria, totaling 15 papers included in the review (Fig. 1). However, Michaelson et al. (62) and Mannion et al. (68) reported supplementary data from the same participant samples described in Aasa et al. (54) and Mannion et al. (67), respectively. These reports were therefore integrated into the original studies, resulting in a final count of 13 unique RCTs included in the meta-analysis.

A total of 57 full-text articles were excluded following the eligibility assessment. Four records were excluded as they were conference abstracts (69–72). An additional four studies were excluded due to the absence of participant randomization; three were published studies (73-75), and one was an unpublished thesis (Costa K. Effects of a trunk strengthening program on pain perception, strength, and flexibility in patients with non-specific low back pain. Doctor of Physiotherapy thesis, Bond University; 2010. Online. Last accessed on 22/09/2025). Four studies were excluded because they enrolled participants with specific forms of LBP (76-79). Eighteen studies were excluded for failing to incorporate external resistance into their exercise protocols

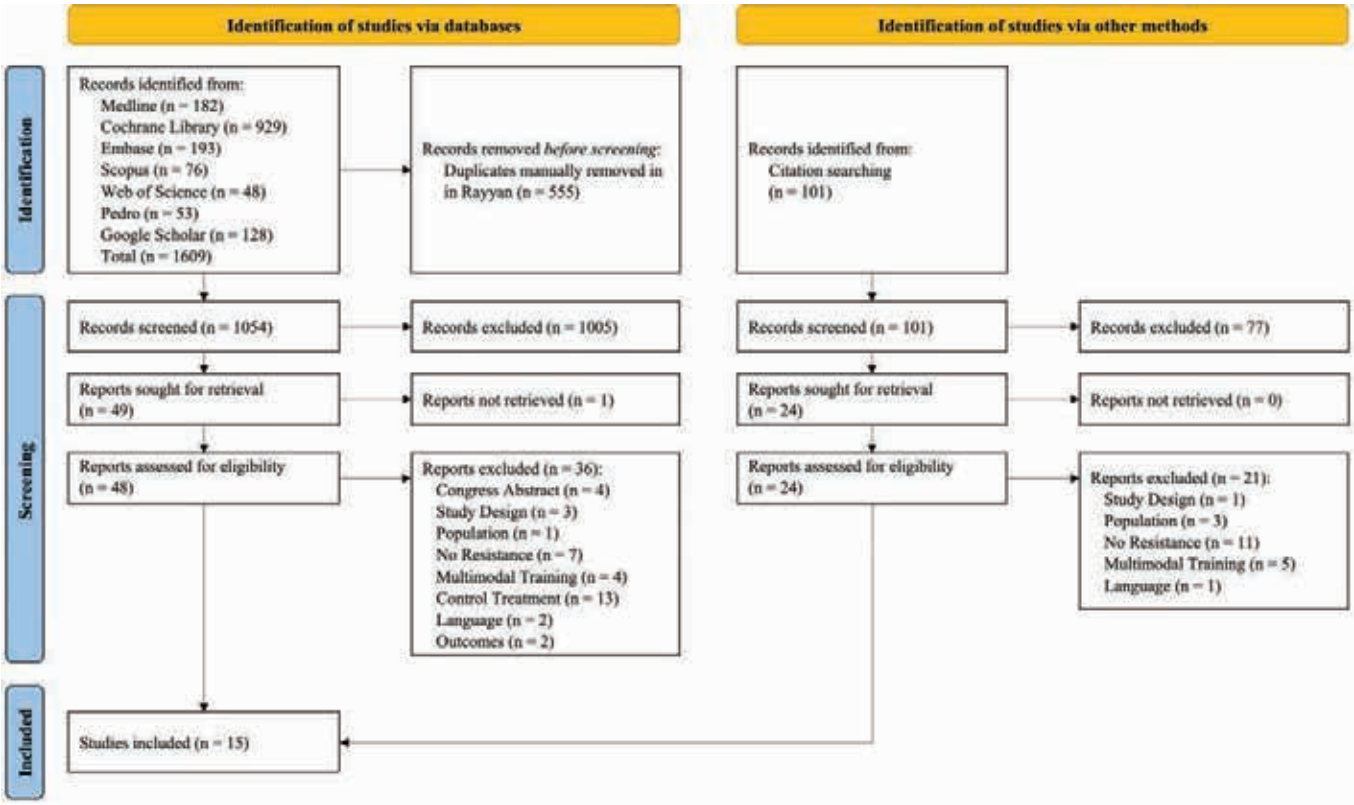


FIGURE 1 - Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) flow diagram of study selection.

(80-97), and nine were excluded because they did not isolate resistance training from other concurrent interventions (98-106). Thirteen studies were excluded for employing either passive control conditions or applying resistance training to both study groups (107-119). Finally, three studies were excluded due to language restrictions (120-122), and two were excluded because the outcomes assessed were not relevant to the review question (123-124).

Description of Study Populations

This systematic review included 13 studies published between 1999 and 2022 involving a total of 778 subjects with chronic NS-LBP. Among them, 395 individuals participated in an RT program incorporating external loads. Participants in the Resistance Training Group (RTG) had an average age of 39.3 years (range: 20.2-61.5 years) and a mean symptom duration of 3.96 years (range: 3.9 months to 13 years). The Unloaded Exercise Group (UEG) had a similar average age of 39.3 years (range: 20.8-57.2 years) and a mean symptom duration of 3.21 years (range: 4.1 months to 9.7 years). The average height and weight were 168.5 cm (range: 159-183 cm) and 74.4 kg (range: 61.7-88.4 kg), respectively, in the RTG, and 167.5 cm (range: 156-182 cm) and 72.3 kg (range: 60.3-86.2 kg), respectively, in the UEG. Perceived pain levels during activities averaged 5.4/10 (range: 2.9-8) in the RTG and 5.7/10 (range: 2.7-7.4) in the UEG. Gender distribution was reported in 10 of the 13 included RCTs. In RTG, 198 men (61.9%) and 122 women (38.1%) were enrolled, while in UEG, 218 men (64.1%) and 122 women (35.9%) were included. Detailed participants' characteristics are presented in Table 1.

Description of Exercise Interventions

Both RTG and UEG followed equivalent training periods and weekly frequencies. Training duration typically lasted for eight weeks, ranging from three (64) to sixteen weeks (59). Participants trained an average of three sessions per week, ranging from one (65) to seven sessions per week (66), including both supervised and unsupervised formats. The median number of sessions across studies was 12 (interquartile range: 12-16) and was comparable between groups.

Training intensity and Resistance modalities

The RTG program intensity was quantified using various metrics. Five studies used an estimated percentage of One Repetition Maximum (1RM) (54,55,58,59,66). One study employed repetition in reserve (RIR) (57). One study used an estimated percentage of maximal voluntary isometric contraction (66). Resistive modalities included: free weights (54,57,59,63,66), weight-stack machines (58,59, 63,65,67), elastic bands (55,56) and isokinetic machines (60,61,64). Seven studies used multi-joint exercises (54-57,59,63,66), while six targeted the lumbopelvic muscles (58,60,61,64, 65,67).

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 1 - Characteristics of the included participants

Study	G	n	Male/ Female, n	Age, Years	Weight, Kg	Height, Cm	Baseline pain	Baseline Disability	Type of LBP	Adverse events*, n	Duration of Symptoms, Years
Aasa et al., (54)	RTG	35	15/20	42 ± 10	74 ± 13	174 ± 8	43 (0-100) ± 24	7 (0-24) ± 4	C	2 (related)	5.98 ± 5.96
	UEG	35	16/19	42 ± 11	78 ± 15	172 ± 10	47 (0-100) ± 28	7 (0-24) ± 5	C	0	6.52 ± 3.64
Cai et al., (66)	RTG (a)	28	14/14	28.9 ± 5.3	61.7 ± 12.6		3.5 (0-10) ± 1		C	1 (unrelated)	1.19 ± 0.67
	RTG (b)	28	14/14	26.1 ± 4.1	61.7 ± 10.8		3.4 (0-10) ± 0.9		C	1 (unrelated)	1.24 ± 0.63
	UEG	28	14/14	26.9 ± 6.4	60.3 ± 12.1		3.6 (0-10) ± 1.1		C	2 (unrelated)	1.32 ± 0.52
Calatayud et al., (55)	RTG	42		52 ± 11	76 ± 19	164 ± 10	6.2 (0-10) ± 2	4.3 (0-24) ± 2	C		
	UEG	43		50 ± 12	72 ± 14	165 ± 7	6.3 (0-10) ± 2	5.1 (0-24) ± 3	C		
Castro et al., (56)	RTG	12	0/12	61.5 ± 7.5	73.2 ± 13.3	159± 13	6.75 (0-10) ± nr	13.8 (0-24) ± nr	C		
	UEG	13	0/12	57.2 ± 10	67.3 ± 11	156 ± 7	6.1 (0-10) ± nr	13.2 (0-24) ± nr	C		
Gibbs et al. (57)	RTG	32	19/13	35.6 ± 12.4			5.6 (0-10) ± 2.1	18.8 (0-100) ± 7.8	C	1 (related)	4.42± 3.68
	UEG	32	17/15	33.5 ± 11.9			5.8 (0-10) ± 2.2	15.4 (0-100) ± 8.7	C	1 (related)	4.79 ± 5.62
Helmhout et al. (58)	RTG	71	69/3	37 ± 11	85 ± 12	183± 8		8.3 (0-24) ± 4.8	SA, C	1 (related)	
	UEG	56	54/2	35 ± 11	86 ± 11	182 ± 7		7.9 (0-24) ± 4.4	SA, C	0	
Kell et al., (59)	RTG	9	6/3	40.1 ± 8.7	88.4 ± 22.4	174 ± 8	5.4 (0-10) ± 0.9	40.4 (0-100) ± 2.4	C		2.29 (range 0.5 - 8)
	UEG	9	5/4	36.7 ± 8.9	81.7 ± 11.5	173 ± 10	5.1 (0-10) ± 0.8	39.8 (0-100) ± 2.3	C		2.29 (range 0.5 - 8)
Mannion et al., (67)	RTG	49	22/27	43.7 ± 10.1	70.3 ± 13.4	172 ± 9	4.2 (0-10) ± 1.8	8 (0-24) ± 5.1	C	0	13 ± 10
	UEG	50	23/27	45.2 ± 9.7	68 ± 12.3	170 ± 11	4.1 (0-10) ± 1.8	7.7 (0-24) ± 4.7	C	2 (related)	9.7 ± 9.1
Nambi et al., (60)	RTG	20	20/0	21.1 ± 1.4	66.8 ± 1.5	167 ± 14	7.2 (0-10) ± 0.4		C		0.325 ± 0.06
	UEG (a)	20	20/0	22.1 ± 1.3	66.8 ± 1.5	166 ± 15	7.3 (0-10) ± 0.3		C		0.35 ± 0.05

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

Study	G	n	Male/ Female, n	Age, Years	Weight, Kg	Height, Cm	Baseline pain	Baseline Disability	Type of LBP	Adverse events*, n	Duration of Symptoms, Years
	UEG (b)	20	20/0	21.4 ± 1.4	66.4 ± 1.4	165 ± 13	7.4 (0-10) ± 0.6		C		0.367 ± 0.06
Nambi et al., (61)	RTG	15	15/0	20.2 ± 1.6	65.6 ± 1.4	168 ± 14	7.3 (0-10) ± 0.5		C		0.342 ± 0.05
	UEG (a)	15	15/0	21.3 ± 1.2	66.2 ± 1.4	165 ± 18	7.1 (0-10) ± 0.6		C		0.342 ± 0.03
	UEG (b)	15	15/0	20.8 ± 1.6	65.5 ± 1.5	168 ± 15	7.3 (0-10) ± 0.6		C		0.358 ± 0.04
Santos et al., (63)	RTG (a)	14		49.1 ± 7.1	78.2 ± 17.8	161± 7	7 (0-10) ± 1.7	14.8 (0-24) ± 3.2	C		6.2 ± 5.4
	RTG (b)	15		43.3 ± 12.1	79.1 ± 12.4	163± 7	8 (0-10) ± 1.3	16.8 (0-24) ± 5	C		3.7 ± 2.4
	UEG	14		46.7 ± 15.3	86.2 ± 11.4	160 ± 6	6.53 (0-10) ± 1.2	16.1 (0-24) ± 5.4	C		5 ± 3.88
Sertpoyraz et al., (64)	RTG	20	4/16	38.8 ± 7.8			4.9 (0-10) ± 0.9	16.6 (0-100) ± 8.1	C		2.4 ± 2.1
	UEG	20	5/15	38.3 ± 7.4			5.4 (0-10) ± 1.3	18.8 (0-100) ± 7.8	C		3.8 ± 4.3
Smith et al., (65)	RTG (a)	16					30.1 (0-100) ± 17.2	39.2 (0-100) ± 14.7	C		
	RTG (b)	17					28.7 (0-100) ± 17.4	35.7 (0-100) ± 12.6	C		
	UEG	13					26.8 (0-100) ± 9	32.7 (0-100) ± 5.9	C		

Data are presented as Mean ± Standard Deviation (SD), unless otherwise stated.
* Total number of adverse events of any type. Parentheses indicate whether events were reported as “related” or “unrelated” to the intervention.
Legend: G = group; RTG = Resistance Training Group; UEG = Unloaded Exercise Group; n = number; Cm = centimeters; BMI = Body Mass Index; SA = Subacute; C = Chronic; nr = not reported;

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Adverse events related to the exercise intervention

Reporting of adverse events was heterogeneous and predominantly descriptive. Overall, the RTG did not demonstrate a higher incidence of events compared with the UEG. Aasa et al. (54) and Helmhout et al. (58) each reported one withdrawal in the RTG due to symptom aggravation, with no corresponding events in the UEG. In contrast, Mannion et al. reported two dropouts in the UEG due to LBP flare-ups, with none in the RTG (67). Gibbs et al. observed one transient flare-up in each group, which was managed through temporary reduction of load or repetitions (57). Cai et al. documented three musculoskeletal injuries, equally distributed across groups and deemed unrelated to the training interventions (66). The remaining RCTs did not provide explicit information regarding adverse events associated with the interventions. Taken together, adverse events were infrequent, mild in nature, and evenly distributed between groups. The number of adverse events reported in each group is summarized in Table 1.

Progressive overload strategies

While studies using isokinetic machines (60,61,64) did not incorporate progressive overload, all other RT programs implemented progressive overload methods.

Aasa et al. (54), Helmhout et al. (58), and Smith et al. (65) increased the load by 2.5 kg or 5% when participants exceeded the target repetitions. Cai et al. (66) and Kell et al. (59) adjusted load based on planned 10RM reassessments.

Gibbs et al. (57) focused on progressing movement complexity for four weeks before increasing intensity to 1RM in the final week. Castro et al. (56) and Calatayud et al. (55) progressively increased elastic resistance every two weeks, starting from 20RM and reaching 10RM. Santos et al. (63) applied linear periodization, gradually increasing intensity while decreasing training volume. Further details on the training programs used in the RTG and the UEG are presented in 'Appendix C'.

Results of individual studies

The results of the individual studies categorized into STP, ETP and PW follow-ups, along with the calculated within-group improvements, are presented in Appendix D. Notably, none of the included studies reported post-intervention scores for kinesiophobia, and only one study (57) provided data on pain catastrophizing.

Short Training Program (STP) Follow-up

At the STP follow-up, improvements in pain intensity (59,60,61,63), disability (59,63), and muscle endurance (63) were generally greater in the RTG than in the UEG. Most studies reported within-group improvements in pain and disability, though gains in muscle endurance and strength were smaller. One UE program (63) demonstrated a significant improvement in VAS scores despite a slight decline in BST performance.

Extended Training Program (ETP) Follow-up

At the ETP follow-up, pain intensity (66,59,65) and disability (54,55,57-59) improved more in the RTG, whereas muscle endurance and strength showed similar changes in both groups. Calatayud et al. (55) reported a mean BST performance change of 44.39 seconds, although pain and disability improvements were only a minimal clinically important difference (MCID) (125). Gibbs et al. (57) found reductions in pain catastrophizing on the PCS (0-52) in both groups at post-intervention. RTG improved from 16.0± 10.3 to 8.5 ± 9.6. UEG improved from 18.0 ± 11.7 to 8.9 ± 8.5.

Post-Washout (PW) Follow-up

At the PW follow-ups, primary outcomes generally continued to improve. Aasa et al. (54) reported a slight worsening in pain intensity and disability, but muscle endurance and strength improved. Helmhout et al. (58) found that the UEG experienced a slight decline in maximal strength, with no change in disability at 8 and 14 months. Regarding PCS scores, Gibbs et al. (57) reported a slight increase in pain catastrophizing in both groups: RTG from 8.5 ± 9.6 to 9.8 ± 10.1. UEG from 8.9 ± 8.5 to 10.8 ± 10.5.

Between-Group Comparisons

Seven studies (54-58,64,66) found no significant differences between RTG and UEG after treatment. Five studies (59,60, 63-65) concluded that RT led to superior outcomes. Conversely, Nambi et al. (61) reported that UE was superior to RT.

Risk of bias in studies

The detailed ROB2 assessment is provided in Appendix E, including a summary and traffic light plots for each outcome. None of the ROB2 domains were rated as having a high risk of bias. However, five RCTs (58-60,65,66) were judged to have a high overall risk of bias due to the cumulative impact of multiple domains rated as "Some Concerns."

Primary outcomes meta-analyses

Data on pain intensity were reported in 12 RCTs, while data on disability were available from 10 RCTs. One study (56) was excluded from the quantitative synthesis due to the unavailability of summary statistics, despite attempts to contact the authors.

Pain intensity

At the STP follow-up (6 RCTs, n = 282), the MD was -0.48/10 (95%CI: [-1.15, 0.18]; p = 0.15), indicating no clear average between-group difference. The 95%PI (-3.25 to 2.44) suggests a wide range of possible effects, including the possibility of both meaningful pain reduction and no benefit. Heterogeneity statistics were I^2 = 97% and tau^2 = 1.02 (Fig. 2 – Forest Plot 1a). At ETP follow-up (7 RCTs, n = 474), the pooled MD was −0.52/10 (95%CI: [−0.92, −0.08]; p = 0.022), suggesting a small average reduction in pain favoring RTG. The 95%PI (–1.88-0.83) includes the null, indicating that future studies may still find no effect in some settings. Heterogeneity statistics were I^2 = 67% and tau^2 = 0.28 (Fig. 2 – Forest Plot 1b). At PW follow-up (5 RCTs, n = 345), the pooled MD was −0.36/10 (95%CI: [−1.18, 0.47]; p = 0.40), again showing no clear average difference. The 95%PI (–3.25-2.54) indicates a broad distribution of

possible effects across studies, consistent with the variability captured by I^2 = 97% and tau² = 1.09 (Fig. 2 – Forest Plot 1c).

Disability

At STP follow-up (3 RCTs, n = 206), the pooled SMD was −2.04 (95%CI: [−3.92, −0.16]; p = 0.033), suggesting a large average reduction in disability favoring the RTG. The 95%PI (−4.43-0.35) indicates that future studies could observe effects ranging from very large improvements favoring the RTG to moderate effects favoring the UEG. Heterogeneity statistics were I^2 = 93% and tau² = 4.18 (Fig. 2 – Forest Plot 2a). At the ETP follow-up (7 RCTs, n = 505), the pooled SMD was –0.40 (95%CI: [-1.00, 0.20]; p = 0.19), indicating that the average improvement in disability was similar between groups. The 95%PI (–1.68-0.88) suggests that future studies may find results ranging from large improvements favoring either the RTG or the UEG. Heterogeneity statistics were I^2 = 82% and tau² = 0.63 (Fig. 2 – Forest Plot 2b). At the PW follow-up (4 RCTs, n = 331), the pooled SMD was 0.06 (95%CI: [−0.16, 0.28]; p = 0.59), suggesting that the degree of improvement in disability was comparable between groups. The 95%PI (–0.43 to 0.55) indicates that future studies are likely to observe small to moderate effects in either direction. Heterogeneity statistics were I^2 = 0% and tau² = 0 (Fig. 2 – Forest Plot 2c).

Post hoc Subgroup Analyses

Subgroup effects were observed only in the STP meta-analysis for disability, where greater treatment effects

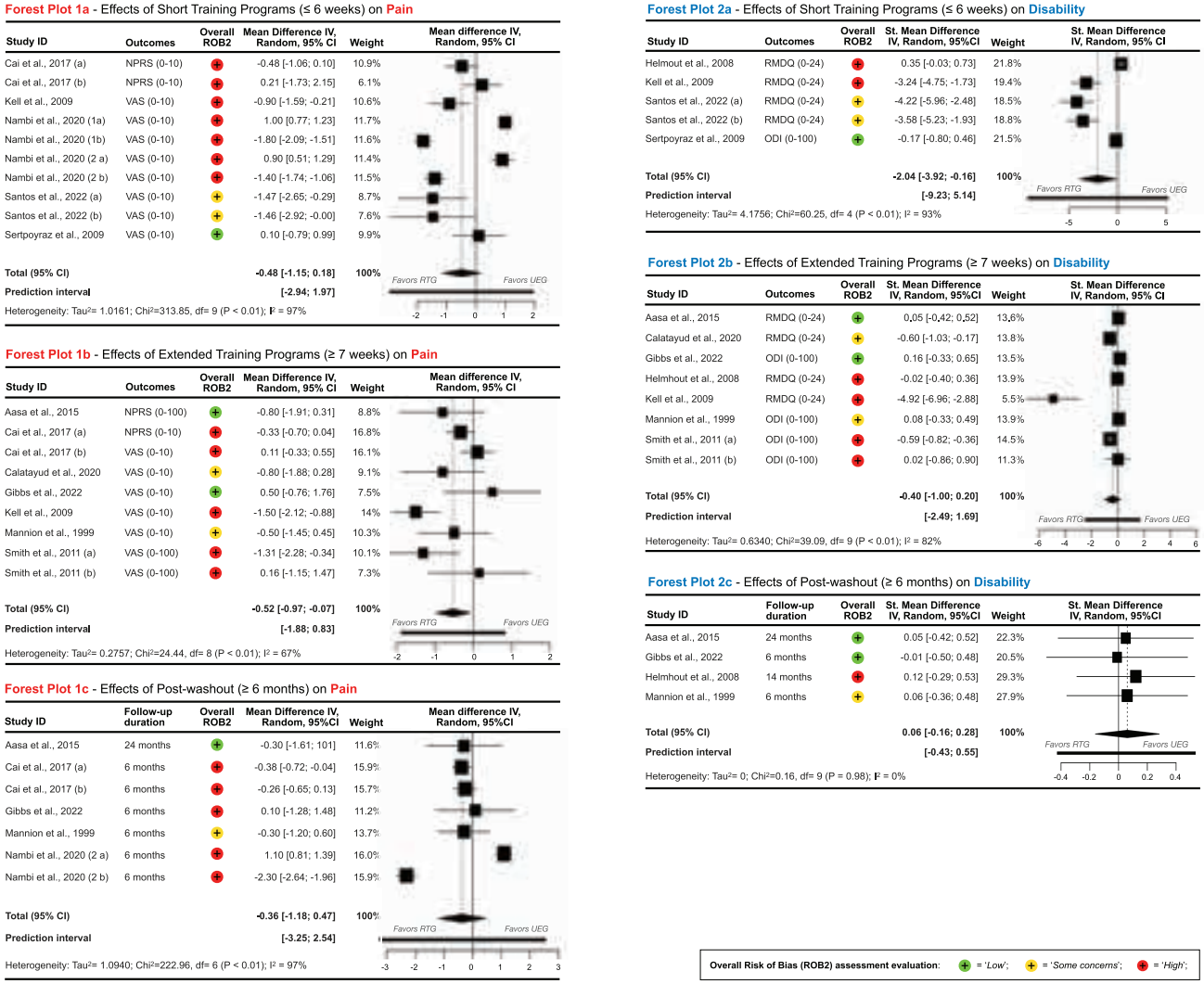


FIGURE 2 - Primary Outcomes Forest Plots: Effects of Loaded and Unloaded Exercise on Pain and Disability .
ROB2: Risk of Bias Tool 2; CI: Confidence Interval; PI: Prediction Interval; LL: Lower Limit; UL: Upper Limit; IV: Inverse Variance; W: Weight; MD: Mean Difference; SMD: Standardized Mean Difference; VAS: Visual Analog Scale; ODI: Oswestry Disability Index; RMDQ: Roland and Morris Disability Questionnaire; RTG: Resistance Training Group; UEG: Unloaded Exercise group

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

were found among participants with high baseline pain (Q = 59.55; df = 2; p < 0.0001), and in interventions combining internal and external resistance (Q = 49.02; df = 1; p < 0.0001). Full subgroup results are reported in Appendix F.

Sensitivity analyses were performed only for pain STP, ETP and PW follow-ups and the disability ETP follow-up, as these were the only meta-analyses including data from more than five RCTs. In the pain STP analysis, the studies by Nambi et al. (60,61) were identified as influential, as their exclusion resulted in a statistically significant pooled effect (MD = −0.53/10; 95%CI: [−1.06, −0.05]; p = 0.0478) and a marked reduction in heterogeneity statistics (I^2 = 58%, tau^2 = 0.24). At the PW follow-up, Nambi et al. (61) were again identified as influential due to the reduction of the I^2 and tau^2 statistics to zero and the emergence of a significant pooled effect (MD = −0.31/10; 95%CI: [−0.55, −0.07]; p = 0.01). The study by Kell et al. (59) was influential in both the ETP pain and disability analysis. In the pain ETP analysis, its exclusion reduced tau^2 from 0.276 to 0.091 and I^2 from 67.3% to 37.8%, and the pooled effect became non-significant (MD = −0.33/10; 95% CI: [−0.69, 0.02]; p = 0.687). In the disability ETP follow-up, tau^2 dropped from 0.634 to 0.081 and I^2 from 82.1% to 68.9%, although the effect direction and significance remained unchanged (SMD = −0.169; 95%CI: [−0.44, 0.10]; p = 0.223). These findings suggest that the identified trials contributed disproportionately to between-study heterogeneity and, in some cases, to the statistical significance of the pooled effects.

Secondary outcomes meta-analyses

Although meta-analyses were planned for all secondary outcomes, pooling for the PCS was not feasible, as data were available from only one study (57). Consequently, quantitative syntheses were conducted only for back muscle endurance (4 RCTs), maximal strength (3 RCTs), and FABQ-PA (2 RCTs).

Back Muscle Endurance

At STP follow-up (2 RCTs, n = 56), the pooled MD was 0.50 minutes (95%CI: [0.26, 0.74]; p < 0.001), indicating a statistically significant improvement in the performance of the BST, favoring the RTG. The 95%PI (–1.80 to 2.80) suggests that future studies could observe large effects favoring either treatment. Heterogeneity statistics were I^2 = 38% and tau^2 = 0.018 (Fig. 3 – Forest Plot 3a). At ETP follow-up (3 RCTs, n = 173), the pooled MD was 0.47 minutes (95%CI: [0.09, 0.85]; p = 0.015), again indicating a statistically significant improvement in the BST performance favoring the RTG. The very wide 95%PI (–3.80 to 4.73) reflects the considerable uncertainty about the generalizability of the pooled effect. Heterogeneity statistics were I^2 = 69% and tau^2 = 0.075 (Fig. 3 – Forest Plot 3b). At PW follow-up, Aasa et al. (54) also reported a significant effect on the BST favoring the RTG (MD = 0.35 minutes; 95%CI: [0.03, 0.67]).

Maximal strength

At STP follow-up (2 RCTs, n = 150), the pooled SMD was 0.00 (95%CI: [–0.30, 0.30]; p = 0.998), indicating that the average change in maximal strength was similar between the RTG and UEG. The 95%PI (–1.96-1.96) suggests that future studies may observe large effects on maximal strength, favoring either the RTG or the UEG. Heterogeneity statistics were I^2 = 0% and tau^2 = 0 (Fig. 3 – Forest Plot 4a). At the ETP follow-up (2 RCTs, n = 176), the pooled SMD was 0.22 (95%CI: [–0.70, 0.26]; p = 0.38), again suggesting that average changes were comparable between RTG and UEG. Heterogeneity statistics were I^2 = 20% and tau^2 = 0.033 (Fig. 3 – Forest Plot 4b). At PW follow-up (2 RCTs, n = 176), the pooled SMD was 0.40 (95%CI: [0.08, 0.71]; p = 0.013), suggesting a moderate and statistically significant effect favoring the RTG. Heterogeneity statistics were I^2 = 0% and tau^2 = 0.0 (Fig. 3 – Forest Plot 4c). The 95%PI was not calculated for ETP and PW due to the limited number of comparisons included.

Fear-avoidance beliefs related to physical activity

At ETP follow-up (2 RCTs, 2 comparisons), the pooled MD was –0.09/24 (95%CI: [–0.88, 0.71]; p = 0.82), indicating that average FABQ-PA scores were similar between RTG and UEG. Heterogeneity statistics were I^2 = 0% and tau^2 = 0 (Fig. 3 – Forest Plot 5a).

At PW follow-up (2 RCTs, 2 comparisons), the pooled MD was –0.24/24 (95%CI: [–1.58, 1.09]; p = 0.72), again suggesting no clear difference in FABQ-PA scores between groups. Heterogeneity statistics were I^2 = 0% and tau^2 = 0 (Fig. 3 – Forest Plot 5b). The 95%PI could not be estimated at either follow-up due to the restricted number of comparisons available.

Reporting biases

Reporting bias was assessed using funnel plots (Appendix G). The plots showed no significant signs of bias (126). Due to the limited number of comparisons, Egger's test was applied only to the STP meta-analysis for pain, yielding an intercept of 0.20 (t = −0.57, p = 0.58), suggesting no funnel plot asymmetry.

Quality of the evidence (GRADE Assessment)

The overall certainty of the evidence for primary outcomes was rated as 'very low', due to methodological limitations and inconsistency across studies. For back muscle endurance, certainty was rated "moderate" at both STP and ETP follow-ups, suggesting probable benefits of externally loaded resistance training over unloaded interventions. At the PW follow-up, the certainty decreased to "very low" due to the limited number of trials and imprecision. For maximal strength, the certainty was consistently rated "low" across all follow-ups, primarily due to the small number of included studies and risk of bias. For fear-avoidance beliefs related to physical activity, the evidence was judged as "low" certainty at both ETP and PW timepoints, reflecting imprecision and methodological concerns. For pain catastrophizing, certainty was rated as "very low" at both time points, as only one study provided data, limiting confidence in the estimate. A detailed evidence quality assessment is provided in Appendix H.

Discussion

Main findings

This systematic review evaluated current evidence on the effectiveness of RT interventions involving external loads in

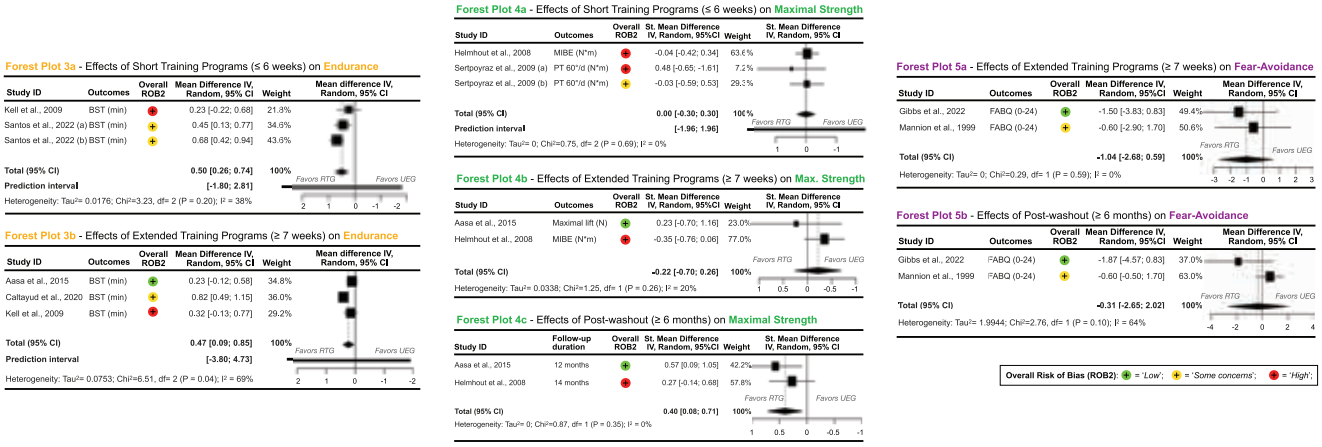


FIGURE 3 - Secondary Outcomes Forest Plots: Effects of Loaded and Unloaded Exercise on Muscle Endurance, Maximal Strength and Fear-Avoidance Beliefs related to Physical Activity.
ROB2: Risk of Bias Tool 2; CI: Confidence Interval; PI: Prediction Interval; LL: Lower Limit; UL: Upper Limit; IV: Inverse Variance; W: Weight; MD: Mean Difference; SMD: Standardized Mean Difference; BST: Biering-Sørensen Test; MIBE: Maximal Isometric Back Extension; PT: Peak Torque; n: newton; nm: newton meters; RTG: Resistance Training Group; UEG: Unloaded Exercise group.

the management of chronic NS-LBP. Thirteen studies were included, comprising 395 participants enrolled in exercise programs utilizing various external load modalities such as free weights, elastic bands, weight stack machines, and isokinetic devices.

Across studies, both RTG and UEG were associated with reductions in pain intensity and disability, but between-group differences were generally modest and inconsistent across follow-ups. At ETP follow-up, the RTG showed a small but statistically significant advantage over the UEG in improving pain intensity (MD = –0.52/10; 95%CI: [–0.92, –0.08]). However, the magnitude of this effect did not exceed the MCID (125), questioning its clinical relevance. For disability, only the STP meta-analysis indicated superiority of the RTG (SMD = −2.04; 95%CI: [−3.92, −0.16]), though this estimate was accompanied by very high inconsistency and was strongly influenced by small-sample trials. Importantly, the 95% PI for both outcomes was wide at every follow-up, suggesting that in future comparable settings, true effects may range from clinically meaningful benefit to no added value of external load. Taken together, these findings indicate that even when significant, the average between-group differences remain small, and the variability across contexts makes strong generalizations premature.

For secondary outcomes, signals in favor of RTG were observed for physical performance measures, although these results should be interpreted cautiously. Back muscle endurance improved more consistently with RTG, with significant effects reported at both STP (MD = 0.50 min; 95%CI: [0.26, 0.74]) and ETP (MD = 0.47 min; 95%CI: [0.09, 0.85]).

The very wide 95%PI calculated for these analyses largely reflects the small number of contributing RCTs. At PW, one trial (54) showed maintenance of these benefits (MD = 0.35 min; 95%CI: [0.03, 0.67]) one year after baseline. For maximal strength, the effects of RT became significant only at PW follow-up (SMD = 0.40; 95%CI: [0.08, 0.71]), potentially reflecting a delayed recovery of physical capacity consistent with the Fitness-Fatigue model (127). Yet, for both ETP and PW follow-ups, 95%PI could not be calculated due to the limited number of comparisons.

In contrast, psychosocial variables were rarely reported. Two RCTs assessed the FABQ-PA (57,67) and only one the PCS (57), without showing any clear between-group difference. The underreporting of psychosocial outcomes likely reflects the historical dominance of biomechanical perspectives in LBP research (5,6). This narrow emphasis limits our ability to determine to what extent training variables, such as load, exert effects beyond neuromuscular adaptations.

Interpretations and Clinical Implications

This meta-analysis examined whether adding external load to RT could enhance the effectiveness of exercise-based treatment for chronic NS-LBP, compared to unloaded approaches. While our findings suggest that external loads may promote measurable neuromuscular adaptations, these benefits do not consistently translate into superior outcomes in pain or disability, nor into clear advantages on psychosocial variables such as fear-avoidance or catastrophizing.

This dissociation reinforces the notion that improvements in physical capacity alone are not sufficient, on their own, to drive meaningful change in chronic NS-LBP—a condition shaped by complex and interacting biological, psychological, and social processes (5,6). The modest and inconsistent effects observed here are in line with accumulating evidence that challenges the assumption of a linear relationship between biomechanical gain and symptom relief (24,128).

Rather than being determined by tissue status or physical capacity, changes in pain and disability may often reflect modifications in pain-related beliefs, behaviors, and emotional responses (129). Although external load may serve as a means of graded exposure or a catalyst for behavioral

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

re-signification (130,131), current data suggest that its effectiveness is not inherently superior. Simply increasing load, without integrating a therapeutic narrative, may be insufficient to influence the multidimensional experience of pain (132,133). It is probable that its effectiveness depends on being embedded within a broader psychological or contextual framework.

From a clinical perspective, this implies that the value of load should be seen less as a mechanical input, but as a potential behavioral signal—the meaning of which depends on patient interpretation, context, and clinical communication (133). In this sense, the delivery and contextual framing of exercise may be more impactful than its intensity or volume.

Given the current heterogeneity in exercise protocols and outcomes, and the limited added value associated with higher intensities, we believe that providing specific dosage prescriptions (e.g., frequency, sets, repetitions) falls beyond the scope of this meta-analysis. However, our subgroup analyses also suggest that exposure consistency and volume, rather than training intensity, may play a more central role in symptom modulation (134). This supports growing interest in low-intensity, high-frequency strategies, such as "exercise snacks," which may provide similar benefits while enhancing adherence, reducing perceived threat, and improving safety (135).

Ultimately, load can be a useful tool—but not a universally necessary one. Its clinical relevance depends on how it is integrated into a person-centred, biopsychosocial framework that acknowledges not only tissue adaptation, but also individual experience, meaning, and the therapeutic alliance.

Study limitations

While interpreting the findings of this review, several limitations should be considered.

The search and selection strategy introduced some limitations. Restricting eligibility to studies published in English and Italian may have introduced language bias, potentially excluding relevant evidence in other languages. The use of an RCT filter increased the specificity of the search but may also have reduced its sensitivity, with the risk of omitting relevant trials. Notably, no eligible studies published between 2022 and March 2025 were identified. However, similar or even longer publication gaps had occurred in earlier periods, suggesting that this pattern was more likely related to the specificity of the inclusion criteria. In addition, although independent screening was conducted by two reviewers, inter-rater agreement was not formally calculated, which is acknowledged as a methodological limitation of the study selection process.

The characteristics of the included populations also limit the generalizability of the findings. The age range of included participants (20.2 to 52 years) limits the applicability of these results to younger and older individuals with chronic NS-LBP. In addition, sex distribution was imbalanced, with men comprising the majority of participants (62–64% across groups). This imbalance may in part reflect the specific populations investigated in certain trials rather than systematic recruitment bias: for instance, Helmhout et al. (58) enrolled military personnel, and Nambi et al. (60,61) studied male soccer players—both predominantly male populations. While this contextualizes the skewed sex distribution, it nonetheless constrains the transferability of our findings, particularly to women with chronic NS-LBP.

Considerable clinical and methodological heterogeneity was observed across studies, particularly in terms of exercise type, loading modalities, intensity, progression strategies, and baseline symptom severity. This variability, together with the lack of standardized intervention protocols, may have influenced treatment effects and contributed to statistical inconsistency. In addition, although available data suggest a favorable safety profile for externally loaded RT, with event rates comparable to unloaded exercise and mostly mild, transient adverse events, harm reporting was inconsistent and largely descriptive, preventing reliable estimation of event incidence.

Additionally, psychosocial variables such as fear-avoidance, catastrophizing, and kinesiophobia were rarely reported, limiting insights into the cognitive-affective impact of externally loaded exercise. Lastly, subgroup and sensitivity analyses were conducted post hoc and were not based on predefined hypotheses; thus, their findings should be interpreted with caution. Finally, some deviations from the original protocol, including the omission of planned meta-regression analyses due to insufficient data, limited the possibility of exploring potential dose–response relationships.

Implications for future research

To better support person-centered care, future research should investigate how RT variables interact not only with physical adaptations but also with psychological and behavioral responses. It remains unclear whether higher training loads confer meaningful benefits beyond strength gains, particularly when not embedded within a therapeutic narrative that addresses pain-related fear or perceptions of fragility. Similarly, increased exercise variety or task complexity may act as a form of graded exposure, helping to reshape pain-related beliefs and improve self-efficacy, rather than solely enhancing motor control. Total RT volume—including frequency, intensity, and duration—should be examined in relation to adherence, perceived safety, and emotional responses, not merely physiological outcomes. Future trials should incorporate validated psychosocial variables to clarify how exercise influences the full spectrum of pain experience and whether individualized interventions can optimize outcomes through mechanisms beyond tissue-level adaptations.

Conclusion

This review highlights that both loaded and unloaded resistance training can lead to reductions in pain and disability among individuals with chronic NS-LBP, although the magnitude and consistency of these effects remain modest. Notably, symptom improvement does not appear to be solely dependent on load intensity. While high-load RT (85-100% 1RM) may be safe and effective when appropriately progressed—even in individuals with elevated pain levels—clinical outcomes likely depend more on how exercise is

structured, delivered, and interpreted than on mechanical intensity alone. Variables such as total training volume, consistency of exposure, and therapeutic framing may play a more substantial role in driving meaningful improvements. From a biopsychosocial perspective, exercise should not be prescribed merely to restore physical capacity, but also as a behavioral intervention aimed at challenging maladaptive beliefs, reducing fear, and fostering movement confidence. Accordingly, exercise programs should be individualized, progressive, and tailored to the patients' needs. Clinicians may consider incorporating both internal and external resistance, along with multi-joint exercises and adequate variation, in order to promote neuromuscular adaptation, enhance psychological engagement, and support long-term adherence.

Declaration of generative AI in scientific writing

During the preparation of this manuscript, the authors used ChatGPT (OpenAI, version GPT-4o) for language editing and improving the clarity of some sections. The tool was used solely to improve readability and language clarity. All outputs were reviewed, and the authors take full responsibility for the final content.

Disclosures

Conflict of interest: The authors declare that they have no conflicts of interest.

Financial support: No funding was awarded for this project.

Data availability statement: All data extracted and analyzed in this systematic review are provided in the appendices of this manuscript. Their inclusion ensures transparency and facilitates the reproducibility of our findings. Additional information can be obtained by contacting the corresponding author.

Ethical compliance: Ethical approval was not required for this study, as it involved the collection and synthesis of data from previously conducted clinical trials in which informed consent had already been obtained by the original investigators.

Author contribution role: All authors contributed equally to the conceptualization and study methodology. MR and DF played equal roles in the data collection and curation. MR conducted the analyses. MR, DR and AP worked on data interpretation and drafted the manuscript. MR edited the manuscript, and MT reviewed the manuscript.

Reference

1. GBD 2021 Low Back Pain Collaborators. Global, regional, and national burden of low back pain, 1990-2020, its attributable risk factors, and projection to 2050: a systematic analysis of the Global Burden of Disease Study 2021. Lancet Rheumatol 2023;22;5(6):e316-e329. CrossRef
2. Hartvigsen J, Hancock MJ, Kongsted A, et al.; Lancet Low Back Pain Series Working Group. What low back pain is and why we need to pay attention. Lancet. 2018;391(10137):2356-2367. CrossRef PubMed
3. Maher C, Underwood M, Buchbinder R. Non-specific low back pain. Lancet. 2017;389(10070):736-747. CrossRef PubMed
4. Wu A, March L, Zheng X, et al. Global low back pain prevalence and years lived with disability from 1990 to 2017: estimates from the Global Burden of Disease Study 2017. Ann Transl Med. 2020;8(6):299. CrossRef PubMed
5. Smith BE, Hendrick P, Bateman M, et al. Musculoskeletal pain and exercise-challenging existing paradigms and introducing new. Br J Sports Med. 2019;53(14):907-912. CrossRef PubMed
6. Bunzli S, Smith A, Schütze R, et al. Making sense of low back pain and pain-related fear. J Orthop Sports Phys Ther. 2017;47(9):628-636. CrossRef PubMed
7. National Institute for Health and Care Excellence. Low back pain and sciatica in over 16s: assessment and management. NICE Clinical Guideline 59. 2016 (updated December 2020). Online (Accessed April 2025)
8. Liu R, Menhas R, Saqib ZA. Does physical activity influence health behavior, mental health, and psychological resilience under the moderating role of quality of life? Front Psychol. 2024;15:1349880. CrossRef PubMed
9. Jadhakhan F, Sobeih R, Falla D. Effects of exercise/physical activity on fear of movement in people with spine-related pain: a systematic review. Front Psychol. 2023;14:1213199. CrossRef PubMed
10. Marshall PWM, Schabrun S, Knox MF. Physical activity and the mediating effect of fear, depression, anxiety, and catastrophizing on pain related disability in people with chronic low back pain. PLoS One. 2017;12(7):e0180788. CrossRef PubMed
11. Owen PJ, Miller CT, Mundell NL, et al. Which specific modes of exercise training are most effective for treating low back pain? Network meta-analysis. Br J Sports Med. 2020;54(21):1279-1287. CrossRef PubMed
12. Searle A, Spink M, Ho A, et al. Exercise interventions for the treatment of chronic low back pain: a systematic review and meta-analysis of randomized controlled trials. Clin Rehabil. 2015;29(12):1155-1167. CrossRef PubMed
13. Kraemer WJ, Ratamess NA. Fundamentals of resistance training: progression and exercise prescription. Med Sci Sports Exerc. 2004;36(4):674-688. CrossRef PubMed
14. Ogawa M, Hashimoto Y, Mochizuki Y, et al. Effects of free weight and body mass-based resistance training on thigh muscle size, strength and intramuscular fat in healthy young and middle-aged individuals. Exp Physiol. 2023;108(7):975-985. CrossRef PubMed
15. Herold F, Müller P, Gronwald T, et al. Dose-response matters! – A perspective on the exercise prescription in exercise – cognition research. Front Psychol. 2019;10:2338. CrossRef PubMed
16. Hansford HJ, Wewege MA, Cashin AG, et al. If exercise is medicine, why don't we know the dose? An overview of systematic reviews assessing reporting quality of exercise interventions in health and disease. Br J Sports Med. 2022;56(12):692-700. CrossRef PubMed
17. Fernández-Rodríguez R, Álvarez-Bueno C, Cavero-Redondo I, et al. Best exercise options for reducing pain and disability in adults with chronic low back pain: pilates, strength, core-based, and mind-body. A network meta-analysis. J Orthop Sports Phys Ther. 2022;52(8):505-521. CrossRef PubMed
18. Schoenfeld BJ, Grgic J, Ogborn D, et al. Strength and hypertrophy adaptations between low- vs. high-load resistance training: a systematic review and meta-analysis. J Strength Cond Res. 2017;31(12):3508-3523. CrossRef PubMed
19. Weakley J, Schoenfeld BJ, Ljungberg J, et al. Physiological responses and adaptations to lower load resistance training: implications for health and performance. Sports Med Open. 2023;9(1):28. CrossRef PubMed
20. Cook JL, Docking SI. "Rehabilitation will increase the 'capacity' of your …insert musculoskeletal tissue here…." Defining 'tissue capacity': a core concept for clinicians. Br J Sports Med. 2015;49(23):1484-1485. CrossRef PubMed
21. Maestroni L, Read P, Bishop C, et al. The benefits of strength training on musculoskeletal system health: practical applications

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

for interdisciplinary care. Sports Med. 2020;50(8):1431-1450. CrossRef PubMed
22. Pacheco-Barrios K, Carolyna Gianlorenço A, Machado R, et al. Exercise-induced pain threshold modulation in healthy subjects: a systematic review and meta-analysis. Princ Pract Clin Res. 2020;6(3):11-28. CrossRef PubMed
23. Westcott WL. Resistance training is medicine: effects of strength training on health. Curr Sports Med Rep. 2012;11(4):209-216. CrossRef PubMed
24. Vlaeyen JW, de Jong J, Geilen M, et al. Graded exposure in vivo in the treatment of pain-related fear: a replicated single-case experimental design in four patients with chronic low back pain. Behav Res Ther. 2001;39(2):151-166. CrossRef PubMed
25. Steiger F, Wirth B, de Bruin ED, et al. Is a positive clinical outcome after exercise therapy for chronic non-specific low back pain contingent upon a corresponding improvement in the targeted aspect(s) of performance? A systematic review. Eur Spine J. 2012;21(4):575-598. CrossRef PubMed
26. Steele J, Fisher J, Perrin C, et al. Does change in isolated lumbar extensor muscle function correlate with good clinical outcome? A secondary analysis of data on change in isolated lumbar extension strength, pain, and disability in chronic low back pain. Disabil Rehabil. 2019;41(11):1287-1295. CrossRef PubMed
27. Tataryn N, Simas V, Catterall T, et al. Posterior-chain resistance training compared to general exercise and walking programmes for the treatment of chronic low back pain in the general population: a systematic review and meta-analysis. Sports Med Open. 2021;7(1):17. CrossRef PubMed
28. Higgins JPT, Thomas J, Chandler J, et al. eds. Cochrane Handbook for Systematic Reviews of Interventions version 6.5 (updated August 2024). Cochrane, 2024. Online (Accessed April 2025)
29. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372(71):n71. CrossRef PubMed
30. Chang WD, Lin HY, Lai PT. Core strength training for patients with chronic low back pain. J Phys Ther Sci. 2015;27(3):619-622. CrossRef PubMed
31. Ishak NA, Zahari Z, Justine M. Effectiveness of strengthening exercises for the elderly with low back pain to improve symptoms and functions: a systematic review. Scientifica (Cairo). 2016;2016:3230427. CrossRef PubMed
32. Mayer J, Mooney V, Dagenais S. Evidence-informed management of chronic low back pain with lumbar extensor strengthening exercises. Spine J. 2008;8(1):96-113. CrossRef PubMed
33. Smith BE, Littlewood C, May S. An update of stabilization exercises for low back pain: a systematic review with meta-analysis. BMC Musculoskelet Disord. 2014;15(1):416. CrossRef PubMed
34. Steele J, Bruce-Low S, Smith D. A review of the clinical value of isolated lumbar extension resistance training for chronic low back pain. PM R. 2015;7(2):169-187. CrossRef PubMed
35. Wewege MA, Booth J, Parmenter BJ. Aerobic vs. resistance exercise for chronic non-specific low back pain: A systematic review and meta-analysis. J Back Musculoskelet Rehabil. 2018;31(5):889-899. CrossRef PubMed
36. Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan-a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210. CrossRef PubMed
37. Waddell G, Newton M, Henderson I, et al. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain. 1993;52(2):157-168. CrossRef PubMed
38. Calley DQ, Jackson S, Collins H, et al. Identifying patient fear-avoidance beliefs by physical therapists managing patients with low back pain. J Orthop Sports Phys Ther. 2010;40(12):774-783. CrossRef PubMed
39. Trinderup JS, Fisker A, Juhl CB, et al. Fear avoidance beliefs as a predictor for long-term sick leave, disability and pain in patients with chronic low back pain. BMC Musculoskelet Disord. 2018;19(1):431. CrossRef PubMed
40. Osuka S, Koshino Y, Watanabe K, et al. Fear-avoidance beliefs associated with non-specific chronic low back pain in college athletes. J Pain Res. 2024;17:285-292. CrossRef PubMed
41. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomized trials. BMJ. 2019;366:l4898. CrossRef PubMed
42. McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): an R package and shiny web app for visualizing risk-of-bias assessments. Res Synth Methods. 2021;12(1):55-61. CrossRef PubMed
43. GRADEpro GDT: GRADEpro Guideline Development Tool [Software]. McMaster University and Evidence Prime, 2023. Online (Accessed April 2025)
44. De la Corte-Rodriguez H, Roman-Belmonte JM, Resino-Luis C, et al. The role of physical exercise in chronic musculoskeletal pain: best medicine-a narrative review. Healthcare (Basel). 2024;12(2):242. CrossRef PubMed
45. Rücker G, Schwarzer G, Carpenter JR, et al. Undue reliance on I(2) in assessing heterogeneity may mislead. BMC Med Res Methodol. 2008;8(1):79. CrossRef PubMed
46. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. BMJ. 2011;342(feb10 2):d549. CrossRef PubMed
47. R Core Team. (2022). R: A Language and Environment for Statistical Computing (Version 4.2.0) [Computer software]
48. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health. 2019;22(4):153-160. CrossRef PubMed
49. Silbernagel KG, Thomeé R, Eriksson BI, et al. Continued sports activity, using a pain-monitoring model, during rehabilitation in patients with Achilles tendinopathy: a randomized controlled study. Am J Sports Med. 2007;35(6):897-906. CrossRef PubMed
50. Viechtbauer W, Cheung MW. Outlier and influence diagnostics for meta-analysis. Res Synth Methods. 2010;1(2):112-125. CrossRef PubMed
51. Peters JL, Sutton AJ, Jones DR, et al. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. J Clin Epidemiol. 2008;61(10):991-996. CrossRef PubMed
52. Harrer M, Cuijpers P, Furukawa T, et al. Doing meta-analysis with R: a hands-on guide. Chapman and Hall/CRC; 2021, CrossRef.
53. Nambi G, Basuodan RM, Alwhaibi RM, et al. Clinical and endocrinological responses to different exercise training methods in chronic low back pain: a randomized controlled trial. Endocr Metab Immune Disord Drug Targets. 2023;23(6):801-810. CrossRef PubMed
54. Aasa B, Berglund L, Michaelson P, et al. Individualized low-load motor control exercises and education versus a high-load lifting exercise and education to improve activity, pain intensity, and physical performance in patients with low back pain: a randomized controlled trial. J Orthop Sports Phys Ther 2015;45(2):77-85, B1-4. CrossRef
55. Calatayud J, Guzmán-González B, Andersen LL, et al. Effectiveness of a group-based progressive strength training in primary care to improve the recurrence of low back pain exacerbations and function: a randomised trial. Int J Environ Res Public Health. 2020;17(22):8326. CrossRef PubMed
56. Castro JB, Lima VP, Mello DB, et al. Effects of Pilates with and without elastic resistance on health variables in

postmenopausal women with low back pain. Pain Manag. 2022;12(4):509-520. CrossRef PubMed
57. Gibbs MT, Morrison NM, Raftry S, et al. Does a powerlifting inspired exercise programme better compliment pain education compared to bodyweight exercise for people with chronic low back pain? A multicentre, single-blind, randomized controlled trial. Clin Rehabil. 2022;36(9):1199-1213. CrossRef PubMed
58. Helmhout PH, Harts CC, Viechtbauer W, et al. Isolated lumbar extensor strengthening versus regular physical therapy in an army working population with nonacute low back pain: a randomized controlled trial. Arch Phys Med Rehabil. 2008;89(9):1675-1685. CrossRef PubMed
59. Kell RT, Asmundson GJ. A comparison of two forms of periodized exercise rehabilitation programs in the management of chronic non-specific low-back pain. J Strength Cond Res. 2009;23(2):513-523. CrossRef PubMed
60. Nambi G, Abdelbasset WK, Alqahtani BA, et al. Isokinetic back training is more effective than core stabilization training on pain intensity and sports performances in football players with chronic low back pain: A randomized controlled trial. Medicine (Baltimore). 2020;99(21):e20418. CrossRef PubMed
61. Nambi G, Abdelbasset WK, Elsayed SH, et al. Comparative effects of isokinetic training and virtual reality training on sports performances in university football players with chronic low back pain-randomized controlled study. Evid Based Complement Alternat Med. 2020;2020(1):2981273. CrossRef PubMed
62. Michaelson P, Holmberg D, Aasa B, et al. High load lifting exercise and low load motor control exercises as interventions for patients with mechanical low back pain: a randomized controlled trial with 24-month follow-up. J Rehabil Med. 2016;48(5):456-463. CrossRef PubMed
63. Santos AOBD, Castro JBP, Nunes RAM, et al. Effects of two training programs on health variables in adults with chronic low back pain: a randomized clinical trial. Pain Manag. 2022;12(4):447-459. CrossRef PubMed
64. Sertpoyraz F, Eyigor S, Karapolat H, et al. Comparison of isokinetic exercise versus standard exercise training in patients with chronic low back pain: a randomized controlled study. Clin Rehabil. 2009;23(3):238-247. CrossRef PubMed
65. Smith D, Bissell G, Bruce-Low S, et al. The effect of lumbar extension training with and without pelvic stabilization on lumbar strength and low back pain. J Back Musculoskelet Rehabil. 2011;24(4):241-249. CrossRef PubMed
66. Cai C, Yang Y, Kong PW. Comparison of lower limb and back exercises for runners with chronic low back pain. Med Sci Sports Exerc. 2017;49(12):2374-2384. CrossRef PubMed
67. Mannion AF, Müntener M, Taimela S, et al. A randomized clinical trial of three active therapies for chronic low back pain. Spine. 1999;24(23):2435-2448. CrossRef PubMed
68. Mannion AF, Müntener M, Taimela S, et al. Comparison of three active therapies for chronic low back pain: results of a randomized clinical trial with one-year follow-up. Rheumatology (Oxford). 2001;40(7):772-778. CrossRef PubMed
69. Saner J, Luomajoki H, Sieben JM, et al. Movement control exercise versus general exercise to reduce disability in patients with low back pain: randomized controlled multicentre study. Physiotherapy. 2015;101(S1):e917-e918. CrossRef
70. Thomas A. Comparative analysis of motor control stability and strengthening program in treatment of chronic low back pain among male weight lifters. Physiotherapy. 2015;101(S1):e1512. CrossRef
71. Verbrugghe J, Agten A, Stevens S, et al. Effects of high intensity training on pain, disability, exercise capacity and muscle strength in persons with non-specific chronic low back pain: preliminary RCT results. Ann Phys Rehabil Med. 2018;61(suppl):e17. CrossRef
72. Vincent HK, Conrad B, Seay A, et al. Low back strength gain contributes to walking improvement in obese older adults with chronic low back pain. PM R. 2013;5(9S):S142. CrossRef
73. Agbonhalor EI, Subulade A. Effects of 10-weeks strength training program on pain intensity, muscle endurance and kinesophobia in patients with non-specific low back pain. Turkish Journal of Kinesiology. 2020;6(1):40-48. CrossRef
74. Günay S, Yildirim Y, Karadibak D. The effect of the muscle endurance training on the chronic low back pain. Turk J Physiother Rehabil. 2014;25(4):28-34. Online
75. Welch N, Moran K, Antony J, et al. The effects of a free-weight-based resistance training intervention on pain, squat biomechanics and MRI-defined lumbar fat infiltration and functional cross-sectional area in those with chronic low back. BMJ Open Sport Exerc Med. 2015;1(1):e000050. CrossRef PubMed
76. Liu-Ambrose TY, Khan KM, Eng JJ, et al. Both resistance and agility training reduce back pain and improve health-related quality of life in older women with low bone mass. Osteoporos Int. 2005;16(11):1321-1329. CrossRef PubMed
77. Manniche C, Lundberg E, Christensen I, et al. Intensive dynamic back exercises for chronic low back pain: a clinical trial. Pain. 1991;47(1):53-63. CrossRef PubMed
78. Miller ER, Schenk RJ, Karnes JL, et al. A Comparison of the McKenzie approach to a specific spine stabilization program for chronic low back pain. J Man Manip Ther. 2005;13(2):103-112. CrossRef
79. Torstensen TA, Ljunggren AE, Meen HD, et al. Efficiency and costs of medical exercise therapy, conventional physiotherapy, and self-exercise in patients with chronic low back pain. A pragmatic, randomized, single-blinded, controlled trial with 1-year follow-up. Spine. 1998;23(23):2616-2624. CrossRef PubMed
80. Andrusaitis SF, Brech GC, Vitale GF, et al. Trunk stabilization among women with chronic lower back pain: a randomized, controlled, and blinded pilot study. Clinics (Sao Paulo). 2011;66(9):1645-1650. CrossRef PubMed
81. França FR, Burke TN, Caffaro RR, et al. Effects of muscular stretching and segmental stabilization on functional disability and pain in patients with chronic low back pain: a randomized, controlled trial. J Manipulative Physiol Ther. 2012;35(4):279-285. CrossRef PubMed
82. Goldby LJ, Moore AP, Doust J, et al. A randomized controlled trial investigating the efficiency of musculoskeletal physiotherapy on chronic low back disorder. Spine. 2006;31(10):1083-1093. CrossRef PubMed
83. Highland KB, Schoomaker A, Rojas W, et al. Benefits of the restorative exercise and strength training for operational resilience and excellence yoga program for chronic low back pain in service members: a pilot randomized controlled trial. Arch Phys Med Rehabil. 2018;99(1):91-98. CrossRef PubMed
84. Kofotolis N, Kellis E, Vlachopoulos SP, et al. Effects of Pilates and trunk strengthening exercises on health-related quality of life in women with chronic low back pain. J Back Musculoskelet Rehabil. 2016;29(4):649-659. CrossRef PubMed
85. Koumantakis GA, Watson PJ, Oldham JA. Supplementation of general endurance exercise with stabilization training versus general exercise only. Physiological and functional outcomes of a randomized controlled trial of patients with recurrent low back pain. Clin Biomech (Bristol). 2005;20(5):474-482. CrossRef PubMed
86. Kumar S, Sharma VP, Shukla R, et al. Comparative efficacy of two multimodal treatments on male and female sub-groups with low back pain (part II). J Back Musculoskelet Rehabil. 2010;23(1):1-9. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

87. Lee JS, Kang SJ. The effects of strength exercise and walking on lumbar function, pain level, and body composition in chronic back pain patients. J Exerc Rehabil. 2016;12(5):463-470. CrossRef PubMed
88. Macedo LG, Latimer J, Maher CG, et al. Effect of motor control exercises versus graded activity in patients with chronic non-specific low back pain: a randomized controlled trial. Phys Ther. 2012;92(3):363-377. CrossRef PubMed
89. Marshall PW, Kennedy S, Brooks C, et al. Pilates exercise or stationary cycling for chronic non-specific low back pain: does it matter? a randomized controlled trial with 6-month follow-up. Spine. 2013;38(15):E952-E959. CrossRef PubMed
90. Moon HJ, Choi KH, Kim DH, et al. Effect of lumbar stabilization and dynamic lumbar strengthening exercises in patients with chronic low back pain. Ann Rehabil Med. 2013;37(1):110-117. CrossRef PubMed
91. Rasmussen-Barr E, Ang B, Arvidsson I, et al. Graded exercise for recurrent low-back pain: a randomized, controlled trial with 6-, 12-, and 36-month follow-ups. Spine. 2009;34(3):221-228. CrossRef PubMed
92. Rhee HS, Kim YH, Sung PS. A randomized controlled trial to determine the effect of spinal stabilization exercise intervention based on pain level and standing balance differences in patients with low back pain. Med Sci Monit. 2012;18(3):CR174-CR181. CrossRef PubMed
93. Roche G, Ponthieux A, Parot-Shinkel E, et al. Comparison of a functional restoration program with active individual physical therapy for patients with chronic low back pain: a randomized controlled trial. Arch Phys Med Rehabil. 2007;88(10):1229-1235. CrossRef PubMed
94. Shamsi M, Sarrafzadeh J, Jamshidi A, et al. The effect of core stability and general exercise on abdominal muscle thickness in non-specific chronic low back pain using ultrasound imaging. Physiother Theory Pract. 2016;32(4):277-283. CrossRef PubMed
95. Shnayderman I, Katz-Leurer M. An aerobic walking programme versus muscle strengthening programme for chronic low back pain: a randomized controlled trial. Clin Rehabil. 2013;27(3):207-214. CrossRef PubMed
96. Suh JH, Kim H, Jung GP, et al. The effect of lumbar stabilization and walking exercises on chronic low back pain: a randomized controlled trial. Medicine (Baltimore). 2019;98(26):e16173. CrossRef PubMed
97. Unsgaard-Tøndel M, Fladmark AM, Salvesen Ø, et al. Motor control exercises, sling exercises, and general exercises for patients with chronic low back pain: a randomized controlled trial with 1-year follow-up. Phys Ther. 2010;90(10):1426-1440. CrossRef PubMed
98. Bae CR, Jin Y, Yoon BC, et al. Effects of assisted sit-up exercise compared to core stabilization exercise on patients with non-specific low back pain: a randomized controlled trial. J Back Musculoskelet Rehabil. 2018;31(5):871-880. CrossRef PubMed
99. Danneels LA, Vanderstraeten GG, Cambier DC, et al. Effects of three different training modalities on the cross sectional area of the lumbar multifidus muscle in patients with chronic low back pain. Br J Sports Med. 2001;35(3):186-191. CrossRef PubMed
100. Gatti R, Faccendini S, Tettamanti A, et al. Efficacy of trunk balance exercises for individuals with chronic low back pain: a randomized clinical trial. J Orthop Sports Phys Ther. 2011;41(8):542-552. CrossRef PubMed
101. Hurley DA, Tully MA, Lonsdale C, et al. Supervised walking in comparison with fitness training for chronic back pain in physiotherapy: results of the SWIFT single-blinded randomized controlled trial (ISRCTN17592092). Pain. 2015;156(1):131-147. CrossRef PubMed
102. Inani SB, Selkar SP. Effect of core stabilization exercises versus conventional exercises on pain and functional status in patients with non-specific low back pain: a randomized clinical trial. J Back Musculoskelet Rehabil. 2013;26(1):37-43. CrossRef PubMed
103. Kim YS, Lee NJ. The effect of progressive resistance exercise and obeisance exercise on functional disorder, back pain, and strength in middle-aged female patients with chronic low back pain for wellness. J Korea Entertain Indust Assoc. 2017;11(4):237-247. CrossRef
104. Oliveira CT, Kanas M, Wajchenberg M. Treatment of non-specific chronic low back pain: resistance training with or without using weights? Rev Bras Med Esporte. 2021;27(6):603-609. CrossRef
105. Tagliaferri SD, Miller CT, Ford JJ, et al. Randomized trial of general strength and conditioning versus motor control and manual therapy for chronic low back pain on physical and self-report outcomes. J Clin Med. 2020;9(6):1726. CrossRef PubMed
106. Verbrugghe J, Agten A, Stevens S, et al. High intensity training to treat chronic non-specific low back pain: effectiveness of various exercise modes. J Clin Med. 2020;9(8):2401. CrossRef PubMed
107. Atalay E, Akova B, Gür H, et al. Effect of upper-extremity strengthening exercises on the lumbar strength, disability and pain of patients with chronic low back pain: a randomized controlled study. J Sports Sci Med. 2017;16(4):595-603. CrossRef PubMed
108. Bates NA, Huffman A, Goodyear E, et al. Physical clinical care and artificial-intelligence-guided core resistance training improve endurance and patient-reported outcomes in subjects with lower back pain. Clin Biomech (Bristol). 2023;103:105902. CrossRef PubMed
109. Bruce-Low S, Smith D, Burnet S, et al. One lumbar extension training session per week is sufficient for strength gains and reductions in pain in patients with chronic low back pain ergonomics. Ergonomics. 2012;55(4):500-507. CrossRef PubMed
110. Cortell-Tormo JM, Sánchez PT, Chulvi-Medrano I, et al. Effects of functional resistance training on fitness and quality of life in females with chronic non-specific low-back pain. J Back Musculoskelet Rehabil. 2018;31(1):95-105. CrossRef PubMed
111. Harts CC, Helmhout PH, de Bie RA, et al. A high-intensity lumbar extensor strengthening program is little better than a low-intensity program or a waiting list control group for chronic low back pain: a randomized clinical trial. Aust J Physiother. 2008;54(1):23-31. CrossRef PubMed
112. Helmhout PH, Harts CC, Staal JB, et al. Comparison of a high-intensity and a low-intensity lumbar extensor training program as minimal intervention treatment in low back pain: a randomized trial. Eur Spine J. 2004;13(6):537-547. CrossRef PubMed
113. Iversen VM, Vasseljen O, Mork PJ, et al. Resistance band training or general exercise in multidisciplinary rehabilitation of low back pain? A randomized trial. Scand J Med Sci Sports. 2018;28(9):2074-2083. CrossRef PubMed
114. Jackson JK, Shepherd TR, Kell RT. The influence of periodized resistance training on recreationally active males with chronic non-specific low back pain. J Strength Cond Res. 2011;25(1):242-251. CrossRef PubMed
115. Kell RT, Risi AD, Barden JM. The response of persons with chronic non-specific low back pain to three different volumes of periodized musculoskeletal rehabilitation. J Strength Cond Res. 2011;25(4):1052-1064. CrossRef PubMed
116. Rittweger J, Just K, Kautzsch K, et al. Treatment of chronic lower back pain with lumbar extension and whole-body vibration exercise: a randomized controlled trial. Spine. 2002;27(17):1829-1834. CrossRef PubMed

117. Steele J, Bruce-Low S, Smith D, et al. A randomized controlled trial of limited range of motion lumbar extension exercise in chronic low back pain. Spine. 2013;38(15):1245-1252. CrossRef PubMed
118. Vincent HK, George SZ, Seay AN, et al. Resistance exercise, disability, and pain catastrophizing in obese adults with back pain. Med Sci Sports Exerc. 2014;46(9):1693-1701. CrossRef PubMed
119. Vincent HK, Vincent KR, Seay AN, et al. Back strength predicts walking improvement in obese, older adults with chronic low back pain. PM R. 2014;6(5):418-426. CrossRef PubMed
120. Raoul T, Malferiot J, Barizien N, et al. Effects of a muscle strengthening program designed for spine extensors in triathletes with chronic back pain. Randomized controlled trial in 67 athletes. J Traumatol Sport. 2019;36(3):183-193. CrossRef
121. Yaghoubi Z, Kahrizi S, Parnianpour M, et al. Short effects of two common stabilization exercise on back and abdominal muscle recruitment and lumbar curvature in non-specific chronic low back pain patients: a crossover clinical trial study. Koomesh. 2014;15(4):e152887. Online
122. Yi T, Lee JH, Lee YJ, et al. Comparisons of spinal stabilization exercise and lumbar extensor strengthening exercise in chronic low back pain. Ann Rehabil Med. 2008;32(5):570-575. online
123. Berglund L, Aasa B, Michaelson P, et al. Effects of low-load motor control exercises and a high-load lifting exercise on lumbar multifidus thickness: a randomized controlled trial. Spine. 2017;42(15):E876-E882. CrossRef PubMed
124. Käser L, Mannion AF, Rhyner A, et al. Active therapy for chronic low back pain: part 2. Effects on paraspinal muscle cross-sectional area, fiber type size, and distribution. Spine. 2001;26(8):909-919. CrossRef PubMed
125. Kovacs FM, Abraira V, Royuela A, et al. Minimal clinically important change for pain intensity and disability in patients with non-specific low back pain. Spine. 2007;32(25):2915-2920. CrossRef PubMed
126. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomized controlled trials. BMJ. 2011;343(jul22 1): d4002. CrossRef PubMed
127. Chiu L, Barnes JL. The fitness-fatigue model revisited: implications for planning short- and long-term training. Strength Condit J. 2003;25(6):42-45. CrossRef
128. Wernli K, Tan JS, O'Sullivan P, et al. Does movement change when low back pain changes? A systematic review. J Orthop Sports Phys Ther. 2020;50(12):664-670. CrossRef PubMed
129. Raja SN, Carr DB, Cohen M, et al. The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises. Pain. 2020;161(9):1976-1982. CrossRef PubMed
130. Booth J, Moseley GL, Schiltenwolf M, et al. Exercise for chronic musculoskeletal pain: a biopsychosocial approach. Musculoskeletal Care. 2017;15(4):413-421. CrossRef PubMed
131. Vlaeyen JWS, Crombez G. Behavioral conceptualization and treatment of chronic pain. Annu Rev Clin Psychol. 2020;16(1): 187-212. CrossRef PubMed
132. Darlow B, Dowell A, Baxter GD, et al. The enduring impact of what clinicians say to people with low back pain. Ann Fam Med. 2013;11(6):527-534. CrossRef PubMed
133. Main CJ, George SZ. Psychologically informed practice for management of low back pain: future directions in practice and research. Phys Ther. 2011;91(5):820-824. CrossRef PubMed
134. Neason C, Miller CT, Tagliaferri SD, et al. Exercise prescription variables predict reductions in pain intensity in adults with chronic low back pain: secondary analysis of a randomized controlled trial. BMJ Open Sport Exerc Med. 2024;10(1):e001744. CrossRef PubMed
135. Fyfe JJ, Hamilton DL, Daly RM. Minimal-dose resistance training for improving muscle mass, strength, and function: a narrative review of current evidence and practical considerations. Sports Med. 2022;52(3):463-479. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 314-325
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3612
REVIEW

Postoperative education for patients after hip/knee arthroplasty: a systematic review with meta-analysis of randomized controlled trials

Leonardo Piano[1,2], Chiara Brossa[2], Leandro Francesco Gioia[2], Adele Occhionero[2,3], Andrea Dell'Isola[4], Marco Trucco[2,5]

[1] Unit of Rehabilitation and Functional Recovery, Fondazione dei Santi Lorenzo e Teobaldo, Rodello - Italy
[2] School of Physiotherapy, University of Turin School of Medicine, Turin - Italy
[3] Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genova, Campus Universitario di Savona, Savona - Italy
[4] Faculty of Medicine, Department of Clinical Science, Lund; Orthopedics, Clinical Epidemiology Unit, Lund University, Lund - Sweden
[5] Presidio Sanitario San Camillo di Torino, Turin - Italy

ABSTRACT

Introduction: Guidelines often recommend patient education in the management of hip or knee replacement, despite evidence of its effectiveness is scarce. The aim of this study was to assess the effectiveness of postoperative patient education on pain, function, quality of life, and psychosocial issues.

Methods: We performed a systematic review and meta-analysis of RCTs investigating the effects of patient education (alone or in combination with other treatments) compared to usual care delivered after hip/knee replacement. Risk of bias was assessed using the Cochrane Risk of Bias 2.0. Certainty of evidence was established according to the GRADE approach.

Results: We included five trials. We found that patient education likely results in little to no effects compared to usual care in the medium-term on pain (SMD −0.09, 95% CI −0.41-0.22) and quality of life (MD 0.11, 95% CI 0.00-0.22). Similarly, it likely results in little to no effect on anxiety in both short- (MD 1.59, 95% CI −3.16 to −0.02) and medium-term (MD −1.51, 95% CI −3.07-0.05), as well as on depression in the short- (SMD −0.22, 95% CI −0.58-0.15) and medium-term (SMD −0.22, 95% CI −0.55-0.12). In contrast, usual care may improve long-term physical function (SMD 0.64, 95% CI −0.03-1.3).

Conclusion: Postoperative patient education provided no to small benefit on pain, physical function, quality of life or psychosocial issues compared to usual care. Its role may be more impactful when patient-tailored and integrated into multimodal rehabilitation strategies for people after hip/knee arthroplasty.

Keywords: Hip arthroplasty, Knee arthroplasty, Patient education, Pain, Physical function, Quality of life

What is already known about this topic:

- *Patient education is often encompassed in the management of patients following hip or knee arthroplasty, but no systematic reviews have investigated the effects when delivered after surgery.*

What does the study add:

- *Patient education following hip or knee arthroplasty results in no to little effect on pain, physical function, quality of life or psychosocial issues (e.g., depression, anxiety) compared to usual care. There is a need for high-quality trials investigating the effects of postoperative education following HA/KA.*

Received: June 2, 2025
Accepted: November 20, 2025
Published online: December 16, 2025

This article includes supplementary material

PROSPERO: CRD42024542032

Corresponding author:
Leonardo Piano
email: leonardo.piano@fondazionesanlorenzo.it

Introduction

Hip and knee arthroplasty (THA/TKA, respectively) are two of the most performed orthopaedic surgeries, usually reserved for end-stage osteoarthritis, when conservative approaches have failed (1). The burden of THA/TKA is continuously rising due to the progressive ageing of the population, obesity, and sedentary lifestyle, which leads to an increased utilization of these procedures (1-3).

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

Despite most patients achieving good clinical outcomes, about 10% of people following THA and 20% of people following TKA report poor outcomes and subjective dissatisfaction due to persistent pain, inability to perform functional activities (e.g., negotiating stairs), and difficulty with a full recovery (4-7).

Beyond the core outcome measures (pain, physical function and quality of life), psychosocial factors (such as anxiety, catastrophizing attitude, wrong beliefs and fear of movement) are often reported among patients in the perioperative period of joint replacement, and seem to be relevant predictors of (poor) outcomes following surgery (8-12). Patient education, defined as "the process of enabling individuals to make informed decisions about their personal health-related behavior" (13), is often recommended by guidelines not only to address psychosocial issues but also to enhance understanding of the postoperative pathway, reduce pain, promote functional recovery, and optimize quality of life (14-16). Patient education is considered capable of improving knowledge about the clinical condition and future perspectives, helping patients following HA/KA to cope with their "new" condition (17,18).

Most systematic reviews explored the role of patient education when delivered before surgery, with conflicting and inconclusive evidence about its effectiveness in improving patients' outcomes after surgery (18-21).

Delivering the education after the surgery may have some advantages as the patient(s) could experience the "new" health condition and receive information and advice more directly related to the topics covered by the educational intervention (e.g., how to cope with postsurgical pain and/or mobility deficits) (22).

Furthermore, cognitive performance—including memory—may be temporarily affected by multiple factors, such as anesthesia (even regional anesthesia), anxiety about the surgery, or the unfamiliar hospital environment, potentially limiting the patient's ability to retain information provided during the preoperative period (23).

To the best of our knowledge, no systematic synthesis of evidence regarding the effects of patient education administered following THA/TKA exists.

This systematic review aimed to assess the effectiveness of patient education on pain, function, quality of life, and psychosocial issues delivered after hip or knee arthroplasty.

Methods

We prospectively registered the study protocol on the PROSPERO database (CRD42024542032) and reported this systematic review following the PRISMA 2020 guidelines (24).

Eligibility criteria

Population

We included randomized controlled trials (RCTs) on adults (i.e., people > 18 years old) following THA or TKA (including unicompartmental) due to osteoarthritis. We excluded studies on people following THA/TKA due to causes other than osteoarthritis, such as avascular necrosis, fractures, or inflammatory diseases. We only considered full-text articles written in English or Italian for final eligibility.

Intervention

We included RCTs focusing on patient education, defined as *"a systematic learning experience in which a combination of methods is generally used, such as the provision of information and advice and behavior modification techniques, which influence the way the patient experiences his illness and/or his knowledge"* (25).

We considered any type of patient education (e.g., face-to-face, app, pamphlet), as long as provided after surgery. We excluded trials focusing on the preoperative or intraoperative stages. We included trials delivering patient education both as a standalone and as a co-intervention with other treatments (e.g., usual care, exercise)

Comparator

We included trials comparing the effectiveness of patient education to (i) usual care (e.g., standard physiotherapy), (ii) no treatment/wait and see.

Outcomes

We included trials investigating at least one of the following primary outcome measures:

- Pain (measured with e.g., visual analogue scale, numeric rating scale, Hip disability and Osteoarthritis Outcome Score (HOOS) (26) subscale "pain", Knee Injury and Osteoarthritis Outcome Score (KOOS) subscale "pain" (27), Western Ontario and McMaster Universities Arthritis Index (WOMAC) subscale "pain" (28);

When multiple outcome measures for pain were reported, we extracted the following measures according to the recommendation of Saadat et al (29), who prioritized: pain overall, pain on stairs, pain at night, pain on walking, pain at rest, WOMAC pain, WOMAC global, and Lequesne index.

- Physical function (e.g., HOOS/KOOS, WOMAC, Six minutes Walking Test (6MWT), Timed Up and Go (TUG) test) (26-28,30,31);
- Quality of Life (e.g., 36-Item Short-Form Health Survey, Short-Form Health Survey 12) (32,33)

We were also interested in the following secondary outcome measures:

- Anxiety (e.g., Beck Anxiety Inventory) (34)
- Depression (e.g., Beck Depression Inventory)
- Kinesiophobia (e.g., Tampa Scale of Kinesiophobia) (35)
- Self-efficacy (e.g., General Self-Efficacy Scale) (36)

We collected outcomes at the following time points:

- short-term – (closest to 1 month after randomization)
- medium-term – (closest to 3 months after randomization)
- long-term – (closest to 12 months after randomization)

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Information sources

We searched the Medline (via PubMed), CENTRAL, CINAHL, EMBASE, and Scopus databases from inception to March 15, 2024. We also searched Google Scholar (first 100 records) and Clinicaltrials.gov.

Search strategy

We built the search strategies for each database, adhering to the PRISMA-S guidelines (37). All the search strategies were presented in Appendix A.

Selection process

Duplicated entries were removed through the Deduplicator function of Systematic Review Accelerator software (38). Two independent reviewers (CB and LG) screened the titles and abstracts of the retrieved studies by using the Rayyan software (39). Any disagreement was solved by a third reviewer (LP). We contacted the corresponding author if full texts were not available or in case of unclear information.

Data collection process

Two independent reviewers (CB and LG) independently extracted mean scores and standard deviations (SD) for continuous outcomes; odds ratios (OR) or relative risk (RR) for dichotomous outcomes. When necessary, data from graphs were extracted through the Web Plot Digitizer software (40).

Data items

Three independent reviewers (CB, LG, and LP) extracted through an Excel file the following data: first author, year of publication, country, number of participants allocated to the education and the control group, patient characteristics (age, gender, baseline pain and duration of symptoms), characteristics of the education intervention (e.g., duration, frequency, type of education), type of comparator, and follow-up. When a study was reported through multiple reports, we collected data through different forms and combined them into a single dataset. Any missing data or unclear information was resolved by contacting the corresponding author of the studies.

Study risk of bias assessment

We used the Cochrane revised risk of bias 2.0 tool to assess risk of bias across outcomes (41). ROB 2.0 evaluates the risk of bias according to these domains: bias arising from (i) the randomization process, (ii) deviation from the intended intervention, (iii) missing outcome data, (iv) measurement of the outcome, and (v) selection of the reported results. Each domain was judged with "High risk", "Low risk", or "Some concerns" according to subsequent signaling questions. A pilot session was conducted to familiarize oneself with the tool. Consensus was reached by discussion between the 2 review authors (CB and LG). The RoB graph was created through the RobVis visualization tool (42).

Effects measures and synthesis methods

We calculated mean differences (MDs) if the same measurement instruments were used for the same outcome, and standardized mean differences (SMDs) if different instruments were used for the same outcome (e.g., NRS or VAS for pain).

If variance in outcomes were reported as confidence intervals, inter-quartile ranges, or standard errors, we converted them into standard deviation using the formulas presented in the Cochrane Handbook (Chapter 7.7.3.2) (43). When data were not fully available, the corresponding authors were contacted to obtain missing data. If SD were not reported in the original RCT and could not be obtained through author contact, we performed imputation as recommended by the Cochrane Handbook (Chapter 16.1.3.1) (44): we used the SD of the most similar study (e.g., in terms of number of randomized patients) included in that specific meta-analysis.

If studies were sufficiently homogeneous in terms of clinical context, we performed meta-analyses using a Der Simonian-Laird random-effect model. We used the Review Manager 5 software to calculate the pooled mean difference (MD) or standardized mean difference (SMD), each reported with 95% confidence intervals (95% CI) (45). We examined heterogeneity as a measure of the between-study variation and calculated it as the I^2 statistic, measuring the proportion of variation in the combined estimates due to between-study variance. An I^2 value of 0% indicates that none of the variance in the pooled estimate can be attributed to between-study variance, and an I^2 value of 100% indicates that all of the variance in the pooled estimate is attributable to between-study variance.

We referred to the thresholds suggested by the Cochrane Handbook (44). In case of clinical heterogeneity, which prevented meta-analysis, we presented results descriptively following the synthesis without meta-analysis (SWiM) guidelines (46).

Reporting bias assessment

To assess publication bias, we generated funnel plots for meta-analyses when at least 10 studies were available: we observed the symmetry of each funnel plot to assess the potential presence of publication bias. If fewer than 10 studies were available, we followed the algorithm published on the Core GRADE recommendations (47).

To assess outcome reporting bias, we compared the outcomes specified in trial protocols with the outcomes reported in the corresponding trial publications; if trial protocols were unavailable, we compared the outcomes reported in the methods and results sections of the trial publications.

Certainty assessment

We used the Grading of Recommendation, Assessment, Development and Evaluation (GRADE) approach to assess the overall certainty of evidence, based on the following domains: risk of bias, inconsistency, imprecision, indirectness, and publication bias (48). We evaluated the evidence by using the GRADEpro software (42). We downgraded by

one level if serious limitations or two levels if very serious limitations were identified in any of the domains. The criteria adopted for the assessment of the certainty were summarized in Appendix C. The certainty of evidence was classified as high, moderate, low, or very low (48). The results are presented in a summary of findings table.

Subgroup analysis

We explored statistical heterogeneity by performing subgroup analyses based on the following criteria: (a) hip vs knee arthroplasty, (b) face-to-face vs remote education, and (c) group-based vs individual education.

Clinical relevance

We considered clinically relevant the effect of educational intervention compared to the control group according to the following criteria:

- Pain intensity: an improvement of 20 points on a numeric rating scale of 0-100(50) or a SMD > 0.2 (51);
- Physical function: an improvement of 9.47 on a 0-100 scale (52) or a SMD > 0.2 (51);
- Quality of life: an improvement of 20% or an SMD > 0.2 (51).
- Psychosocial variables; an improvement of 20% or an SMD > 0.2 (51).

Results

Study selection

We retrieved 749 records: 506 from databases and registers, 100 from Google Scholar, and 143 from ClinicalTrials.gov. After removing duplicates (n = 231), we screened 518 records by title and abstract and read the full text of 41 trials. Finally, we included six trials encompassing 498 patients (36,9% men, mean age 69.0 years) (Fig. 1) (53-58). We detailed the reasons why all potentially relevant trials were not included in our review in Appendix B.

Study characteristics

We summarized the main characteristics of the included trials in Table 1. Two trials were conducted in the Netherlands (53,54), one in China (58), Taiwan (55), Turkey (56), and the United States (57).

In three trials, education was delivered through an app (53,54,58). In the other two trials, education was delivered through a video (55,56). Pellegrini et al. delivered an in-person educational intervention (57). Five trials collected data at short-term (i.e., closest to one month after randomization) (53-56,58), whereas four trials at medium-term follow-up (i.e., closest to three months after randomization) (55-58). No trials had data at long-term follow-up.

One trial was judged at low risk of bias (58). The other five trials were judged at an overall high risk of bias (Fig. 2) (53-57). We summarized the most credible findings (i.e., results with moderate to high-certainty evidence) in the evidence profile table (Table 2). All results are reported descriptively in the following paragraphs.

Results of synthesis

Pain

Three trials (n = 300) investigated the role of patient education on short-term pain. Patient education may result in a slight reduction compared to usual care (MD −5.19, 95% CI −13.01–2.63, I^2 = 61%) (53,54,56) (Fig. 3).

Three trials (n = 158) investigated the role of patient education on medium-term pain intensity: Patient education likely results in little to no difference compared to usual care (SMD −0.09, 95% CI −0.41–0.22, I^2 = 0%), (57,58) (Fig. 4).

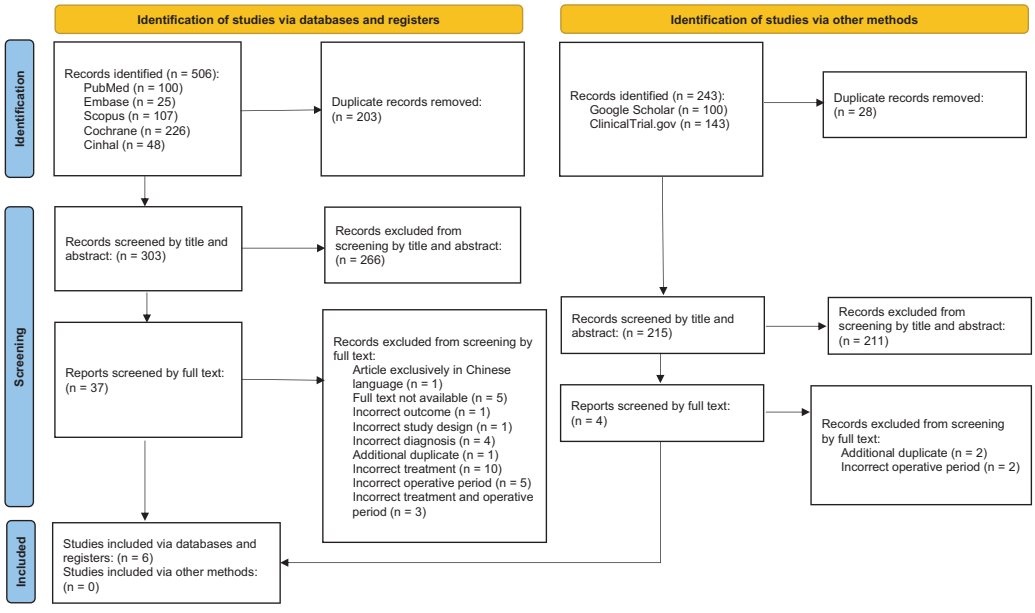


FIGURE 1 - Flow of trials through the review.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 1 - Characteristics of the included studies

Study	Type of surgery	Experimental		Control		Outcomes	Follow up (weeks)
		Sample characteristics	Type of intervention	Sample characteristics	Type of intervention		
Chang 2024 (55)	Knee	n = 26 age 73.4 (5.5) M / F (%) 6 / 20 (23.07%) Pain 52.7 (23.9) Physical function 62.6 (16.9) QoL 24.5 (17.5)	Educational Video + Pamphlet D = 4 months N sessions = NR S = 30’	n = 26 age 73.5 (4.7) M/F (%) 10 / 16 (38.46%) Pain 55.6 (19.0) Physical function 60.7 (20.8) QoL 21.6 (14.5)	Standard rehabilitation D = NR N sessions = twice daily S = NR	Physical function = KOOS (0-100) Depression = GDS-SF (0-15)	1, 6, 12, 16
Eren 2022 (56)	Hip	n = 13 age 49.61 (9.52) M / F (%) 4 / 9 (30.77%) Pain NR Physical function NR	Video-assisted education D = 12 weeks N sessions = NR S = NR	n = 18 age 55.33 (11.83) M / F (%) 6 / 12 (33.3%) Pain NR Physical function NR QoL NR	Physiotherapy D = N sessions = NR S = NR	Pain = VAS (0-10) Physical function = HHS (0-100) ADL (NEADL) Kinesiophobia (TSK)	1, 12
Pellegrini 2022 (57)	Knee	n = 24 age 67.2 (5.4) M/F (%) 8/16 (30%) Pain 54.3 (15.4) Physical function 27.1 (12.2) QoL 38.3 (18.4)	Education + standard physiotherapy D = NR N sessions = NR S = NR	n = 21 age 63.3 (8.1) M/F (%) 10/11 (47.6%) Pain 59.3 (15.4) Physical function 21.5 (10.4) QoL 42.3 (19.5)	Standard physiotherapy D = NR N sessions = NR S = NR	Pain = WOMAC/ KOOS (0-100) Physical function = KOOS (0-100)	12
Pronk 2020 (54)	Knee	n = 38 age 62.6 M/F (%) 23 / 15 (60.52%) Pain 33.0 (48.68) Physical function 47.3 (20.83) QoL 86.0 (31.49)	Paincoach App	n = 33 age 64.6 M/F (%) 19/14 (57.57%) Pain 32.0 (44.0) Physical function 48.5 (24.82) QoL 86.0 (30.32)	Usual care D = N sessions = 7/week S =	Pain = VAS (0-100) Physical function = KOOS (0-100) Quality of life = EQ-5D	4
Timmers 2019 (53)	Knee	n = 114 age 64.74 M/F (%) 40/74 (35.08%) Pain 5.20 (NR) Physical function 45.91 (12.77) QoL 0.67 (0.21)	Customized and active education via Mobile APP	n = 99 Age 65.63 M/F 39/60 (39.39%) Pain 5.54 (NR) Physical function 43.83 (10.87) QoL 0.65 (0.24)	Basic information	Pain = NRS (0-10) Physical function = KOOS (0-100) Quality of life = EQ5D	4
Wang 2023 (58)	Mixed	n = 43 age 63.0 (14.42) M/F (%) 11/32 (25.58%) Pain 2.00 (5.00) Physical function 62.92 (21.24) QoL 0.45 (0.24)	Mobile APP + usual care D = 6 weeks N sessions S = NR	n = 43 age 66.25 (18.18) M/F (%) 8/35 (18.6%) Pain 2.00 (4.00) Physical function 63.10 (19.89) QoL 0.51 (0.21)	Usual care D = 6 weeks N sessions = 15 (5/week) S = NR	Pain = NRS (0-10) Physical function = HOOS/KOOS (0-100) Quality of life = EQ5D Anxiety (HADS-A) (0-21) Depression (HADS-D) Self-efficacy	6, 10

n = number of participants randomized (number of participants completing the study); VAS = visual analogue scale; D = duration of treatment; S = duration of each session; NR = Not Reported; PT = physiotherapy; NRS = Numeric Rating Scale; HOOS = Hip disability and Osteoarthritis Outcome Score; KOOS = Knee disability and Osteoarthritis Outcome Score
Mean age is presented with standard deviation within round brackets.

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

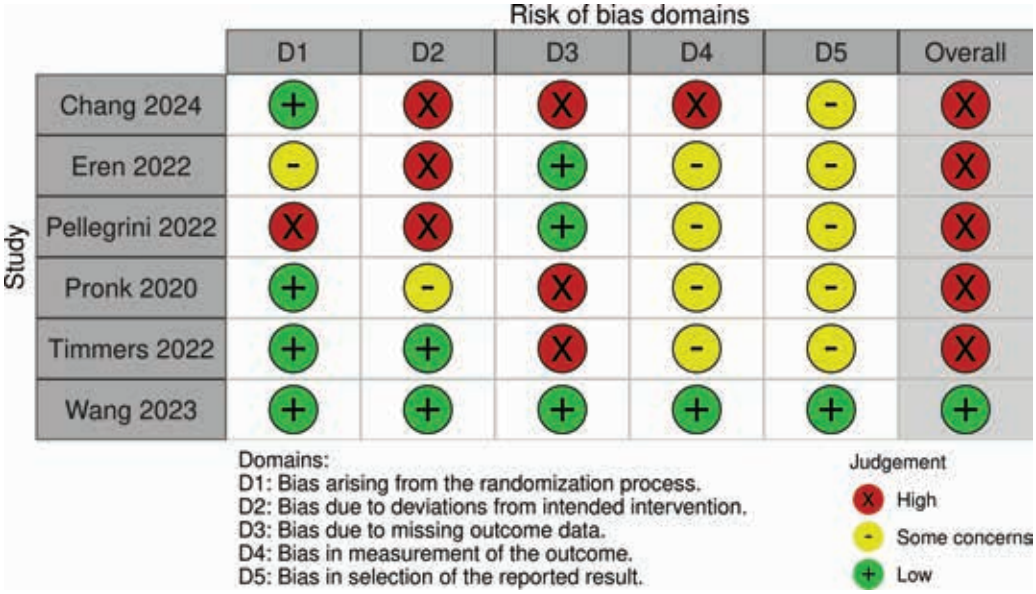


FIGURE 2 - Risk of bias table for each trial included in the review.

Physical function

All the included trials investigated the role of patient education on physical function (53-58).

Three trials (n = 321) investigated the role of patient education on short-term physical function (53-55). The evidence is very uncertain about the effect of patient education (MD −1.88, 95% CI −7.27–3.52, I^2 = 72%) (Fig. 5).

Four trials (n = 209) investigated the role of patient education on medium-term physical function (55–58). Usual care may result in an increase compared to patient education (SMD 0.64, 95%CI −0.03-1.30, I^2 = 80%) (Fig. 6).

Quality of Life

Three trials (n = 355) investigated the role of patient education on short-term quality of life (53,54,58). Patient education may result in no difference when compared to usual care (MD 0.06, 95% CI −0.01-0.13, I^2 = 70%) (Fig. 7). One trial (n = 86) reported that usual care likely results in little to no difference in medium-term quality of life compared to patient education plus usual care (MD 0.11, 95% CI 0.00-0.22) (58).

Psychosocial factors

Two trials (n = 137) investigated the role of patient education on depression (55,58). Patient education likely reduces both short-term (SMD −0.22, 95% CI −0.58-0.15, I^2 = 13%) (Fig. 8) and medium-term depression slightly (SMD -0.22, 95%CI -0.55 to 0.12, I^2 = 0%) (Fig. 9).

One trial focused on anxiety and self-efficacy (58). Patient education likely reduces short-term (MD −1.59, 95% CI −3.16 to −0.02) or medium-term anxiety (MD −1.51, 95% CI −3.07-0.05).

Subgroup analysis

Hip arthroplasty

Only one trial investigated the role of patient education compared to usual care after THA (56). Patient education may result in no difference on short-term (MD −0.04, 95% CI −1.69-1.61) or medium-term pain (MD 0.97, 95% CI −2.59-0.65).

Usual care may result in an increase in medium-term physical function (MD 20.25, 95% CI 7.46-33.04).

Knee arthroplasty

Two trials (n = 269) (53,54) investigated the role of patient education on pain and quality of life, whereas three trials (n = 321) (53-55) focused on physical function after TKA in the short-term. Patient education may result in a reduction in pain (SMD −0.33, 95% CI −0.76-0.10, I^2 = 61%) and increase in physical function (MD −1.88, 95% CI −7.27-3.52 I^2 = 72%), whereas little to no difference in quality of life (MD 0.04, 95% CI −0.04-0.13, I^2 = 80%).

One trial (n = 41) displayed small effects of patient education on medium-term pain (MD −0.60, 95% CI −2.68-1.48) (57).

Two trials (n=92) investigated the role of patient education on physical function after TKA at medium-term (55,57). Patient education may result in no effects compared to usual care (MD 0.76, 95% CI −0.59-2.10, I^2 = 80%).

Face-to-face vs remote/digital

Five trials delivered patient education through an off-site approach (e.g., video, app) (53-56,58). One trial delivered patient education with face-to-face sessions between clinicians and patients (57).

Patient education delivered through video/app probably results in little to no effect on medium-term pain (MD −0.16, 95% CI −0.86-0.54, I^2 = 23%) (Fig. 10), whereas usual care probably increases medium-term physical function (SMD 0.83, 95% CI −0.03-1.70, I^2 = 84%) (Fig. 11).

Group vs individual education

We did not perform any subgroup analysis since no studies focused on group-based education.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

TABLE 2 - Summary of findings table—evidence quality (GRADE assessment)

Comparison: Patient education versus no intervention – overall					
Outcome	**# studies**	**N**	**Effect (95% CI)**	**Reason(s) for downgrading**	**GRADE**
Pain					
1 month	3	300	MD −0.97 (−1.53 to −0.51)	RoB, Imprecision	Low
3 months	3	158	SMD −0.09 (−0.41-0.22)	Imprecision	Moderate
Physical function					
1 month	3	321	MD −1.88 (−7.27-3.52)	RoB, Inconsistency, Imprecision	Very low
3 months	4	209	SMD 0.64 (−0.03-1.30)	Imprecision	Low
Quality of Life					
1 month	3	355	MD 0.06 (−0.01-0.13)	RoB, Imprecision	Low
3 months	1	86	MD 0.11 (0.00-0.22)	Imprecision	Moderate
Anxiety					
1 month	1	86	MD −1.59 (−3.16 to −0.02)	Imprecision	Moderate
3 months	1	86	MD −1.51 (−3.07-0.05)	Imprecision	Moderate
Depression					
1 month	2	137	SMD −0.22 (−0.58, 0.15)	Imprecision	Moderate
3 months	2	137	SMD −0.22 (−0.55, 0.12)	Imprecision	Moderate
Comparison: Patient education versus no intervention – hip arthroplasty					
Outcome	**# studies**	**N**	**Effect (95% CI)**	**Reason(s) for downgrading**	**GRADE**
Pain					
1 month	1	31	MD −0.04 (−1.69-1.61)	RoB, Imprecision	Low
3 months	1	31	MD 0.97 (−2.59-0.65)	RoB, Imprecision	Low
Physical function					
3 months	1	31	MD 20.25 (7.46-33.04)	RoB, Imprecision	Low
Comparison: Patient education versus no intervention—knee arthroplasty					
Pain					
1 month	2	269	SMD −0.33 (−0.76, 0.10)	RoB, Imprecision	Low
3 months	1	41	MD −0.60 (−2.68, 1.48)	RoB, Imprecision	Low
Physical function					
1 month	3	321	MD −1.88 (−7.27, 3.52)	RoB, Imprecision	Low
3 months	2	92	MD 0.76 (−0.59, 2.10)	RoB, Imprecision	Low
Quality of life					
1 month	2	269	MD 0.04 (−0.04, 0.13)	RoB, Imprecision	Low
3 months	1	86	MD 0.11 (0.00, 0.22)	RoB, Imprecision	Low

Table 2: We presented in bold the results that achieved statistical significance. Negative values of MD/SMD indicate a superiority of patient education, whereas positive values indicate a superiority of the control intervention.
MD = Mean difference; N = Number of subjects; SMD = Standardized mean difference; RoB = Risk of Bias

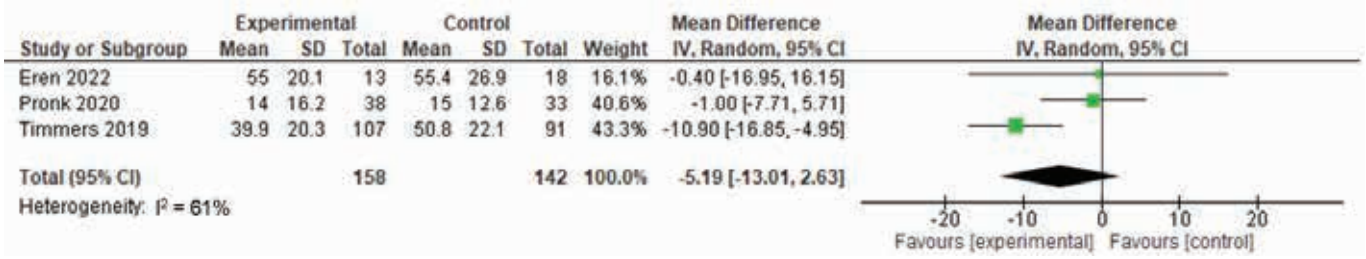


FIGURE 3 - Forest plot of the comparison of patient education versus usual care on short-term pain.

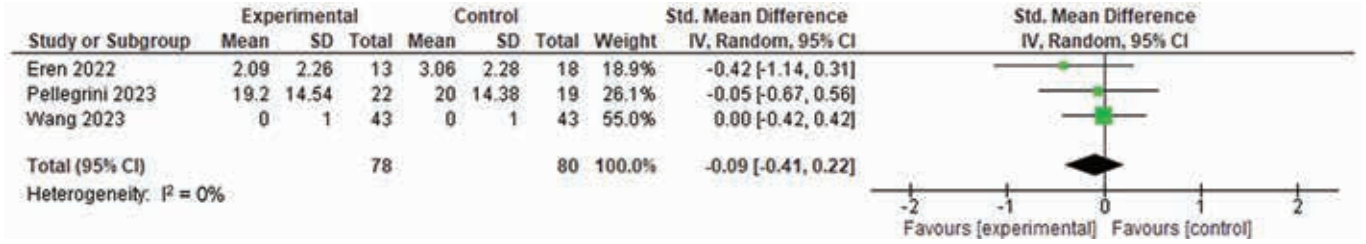


FIGURE 4 - Forest plot of the comparison of patient education versus usual care on medium-term pain.

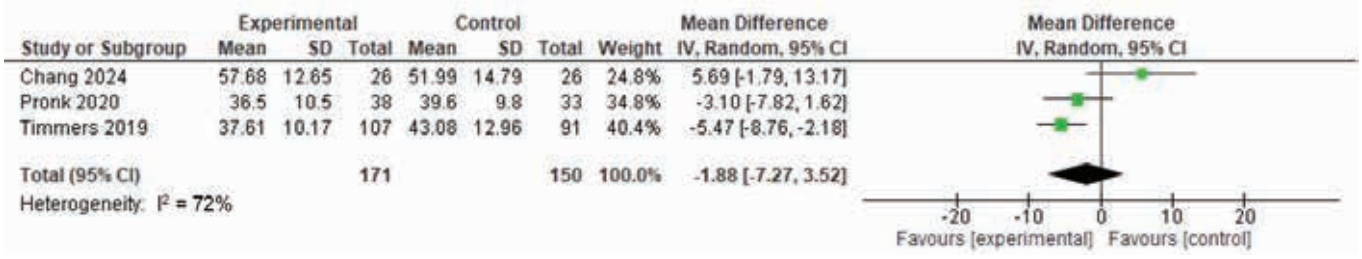


FIGURE 5 - Forest plot of the comparison of patient education versus usual care on short-term physical function.

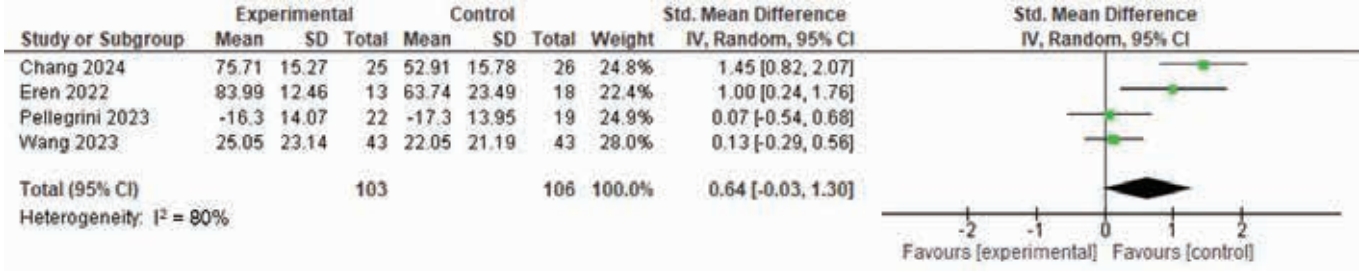


FIGURE 6 - Forest plot of the comparison of patient education versus usual care on medium-term physical function.

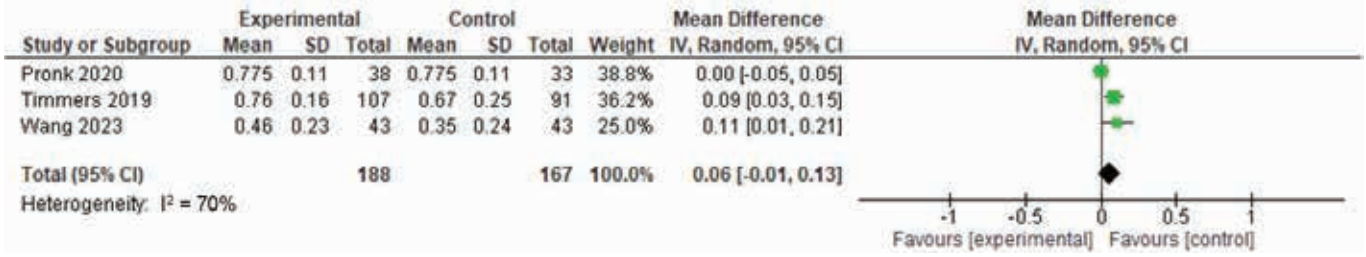


FIGURE 7 - Forest plot of the comparison of patient education versus usual care on short-term quality of life.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

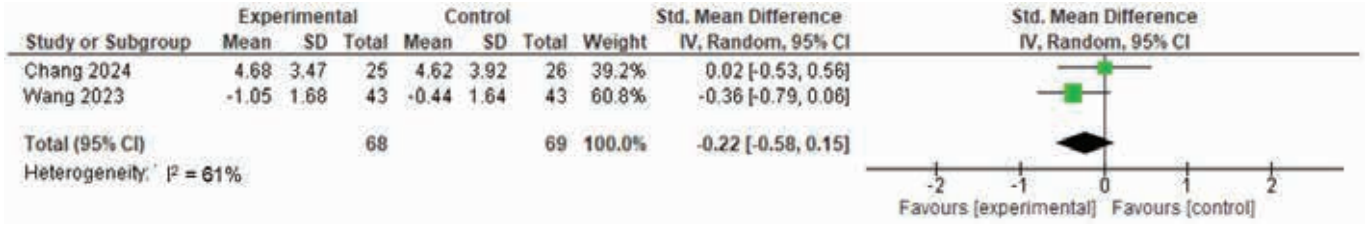


FIGURE 8 - Forest plot of the comparison of patient education versus usual care on short-term depression.

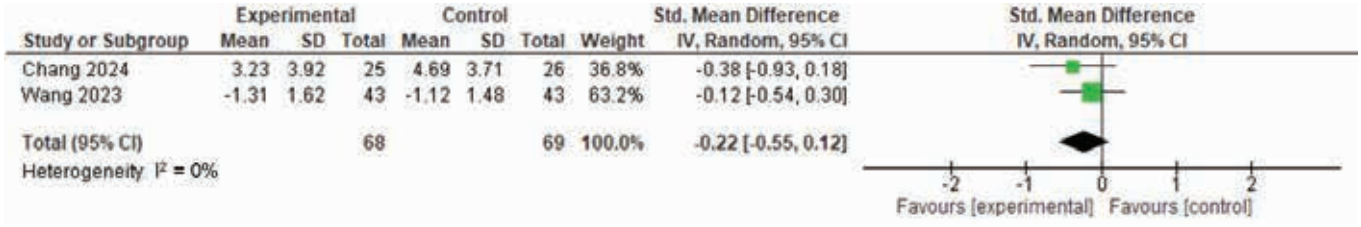


FIGURE 9 - Forest plot of the comparison of patient education versus usual care on medium-term depression.

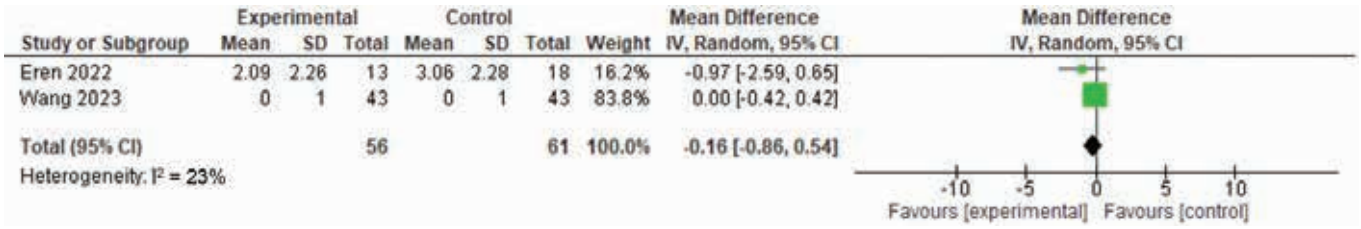


FIGURE 10 - Forest plot of the comparison of patient education delivered face-to-face vs remote on medium-term physical function.

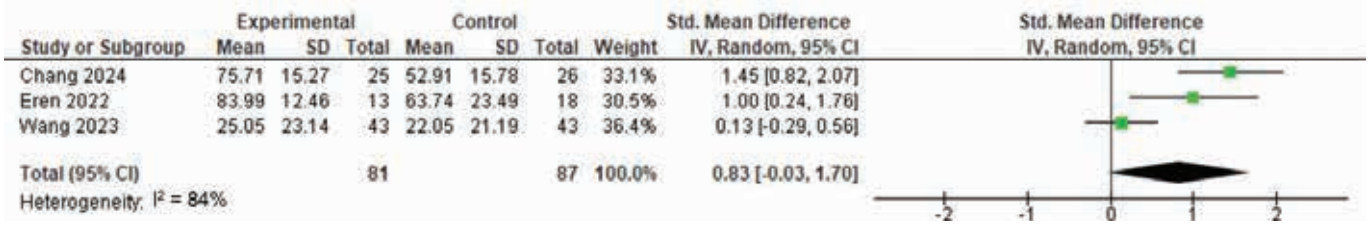


FIGURE 11 - Forest plot of the comparison of patient education delivered face-to-face vs remote on medium-term depression.

Discussion

This systematic review focused on the effects of patient education compared to usual care delivered to patients following THA/TKA on pain, physical function, and quality of life.

We found moderate certainty evidence that patient education had small or no effects on medium-term pain and quality of life, and both short- and medium-term anxiety and depression. We also found from low to very low certainty of evidence that patient education had no small or no effects compared to usual care on disability, short-term pain and quality of life.

Subgroup analyses did not modify our findings, with low-quality evidence for each group (i.e., THA vs TKA, face-to-face vs remote education, individual vs group-based education).

This is the first systematic review investigating the effects of patient education administered following THA/TKA, so we are not able to contrast our findings with other studies. Extending to other studies delivering education before joint surgery, particularly to the Cochrane by MacDonald et al., patient education does not seem to provide a relevant benefit when compared to no intervention (18). These findings should be contextualized in a scenario where the included trials had several limitations: high risk of bias and low sample sizes were common features of the included studies.

The included trials did not adequately report the content (what), the intensity (how much), the duration (how long), and the frequency (how often) of the educational intervention, preventing the opportunity to draw definitive conclusions. In addition, most of the included trials investigated

the effects of standard patient education for all THA/TKA patients: only Timmers et al (53) used an interactive APP that delivered tailored education to patients after TKA. The other trials did not employ customized methods of education: this issue did not likely address the need for a more tailored and patient-centred care advocated by several studies and initiatives (59-62). A personalized approach might be desirable (above all after surgery) to satisfy the patient's needs of knowledge about the short- and long-term recovery, discussing the capability to return to pre-surgery activities, such as activities of daily living, work and also return to sport. The return to potentially higher-demanding activities (e.g., physical work, sports activity) seems to justify the need for a tailored approach, also in the context of the rising number of arthroplasties earlier in life and in active elderly people (63).

Guidelines often recommend patient education as part of a multimodal approach, although weak evidence of effects exists (15,16): beyond the (questionable) clinical effectiveness, patient education might be recommended due to ethical implications, since it is reasonable for clinicians to maintain patient education in the context of the therapeutic relationship with patients following THA/TKA.

Limitations

Our systematic review is not free from limitations, which must be considered. The research question focused on the core outcome set for people with osteoarthritis, and on psychosocial factors (e.g., self-efficacy, anxiety) as secondary outcomes. However, patient education could better impact on other variables, focusing on patients' perspectives, such as patients' satisfaction or compliance with treatment (64,65).

Secondly, almost all trials included in our review administered education through a remote method (e.g., pre-registered video, app): this issue should be discussed since it prevents the delivery of a more interactive and tailored education which may be more "customized" and patient-centred compared to a "standardized" approach, with likely effects on patient's recovery (61,66).

Implications for practice and research

Patient education is often recommended by guidelines to improve patients' involvement, adherence to therapy and commitment (15,16), although the clinical evidence does not justify this approach as a standalone opportunity for people following joint replacement. Clinicians should consider the potential effects of patient education as a complementary intervention to usual care to promote patients' engagement and empowerment (67).

Future research is needed to address the current limitations of the existing literature. The quality of reporting is scarce, as well as the extent of published literature (all comparisons did not achieve the threshold of 400 subjects), focusing on one-to-one patient education, the approach most commonly adopted by clinicians in their clinical practice. We suggest future trialists use appropriate tools such as the Template for Intervention Description and Replication (TIDieR) to better describe the characteristics of education (58), as recommended in the context of other systematic reviews about patient education, although delivered towards a different population (i.e., low back pain patients) (69,70).

Furthermore, we hypothesize that some patients—like those with depression symptoms, anxiety, poor self-efficacy, or kinesiophobia—may benefit from educational intervention more than others. Therefore, researchers should reflect on the opportunity to address educational interventions to "well-selected" patients, planning the trial(s) on a specific population (i.e., patients with a relevant psychosocial burden).

Conclusions

Postoperative patient education provided uncertain clinical relevance compared to usual care on pain, physical function, and quality of life, but may be part of a multimodal approach for people after THA/TKA. High-quality and future studies are needed to improve the sample size and the quality of current evidence.

Disclosures

Conflict of interest: The authors declare no conflict of interest

Financial support: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contributions statement: LP and MT were responsible for the conception and design of the study.
LP, CB, LG, and AO were involved in the data collection. LP was involved in the drafting of the manuscript. AO, AD and MT provided methodological competencies.
All authors contributed to the interpretation of the data for the work and revised it critically for important intellectual content.
All the authors finally approved the manuscript. LP was responsible for the integrity of the work as a whole.
All authors have read and agreed to the published version of the manuscript.

Data availability statement: All data relevant to the study are included in the article or are available as supplementary files.

References

1. Ackerman IN, Bohensky MA, Zomer E, et al. The projected burden of primary total knee and hip replacement for osteoarthritis in Australia to the year 2030. BMC Musculoskelet Disord. 2019;20(1):90. CrossRef PubMed
2. Singh JA, Yu S, Chen L, et al. Rates of total joint replacement in the United States: future projections to 2020-2040 using the national inpatient sample. J Rheumatol. 2019;46(9):1134-1140. CrossRef PubMed
3. Hooper G, Lee AJJ, Rothwell A, et al. Current trends and projections in the utilization rates of hip and knee replacement in New Zealand from 2001 to 2026. N Z Med J. 2014;127(1401):82-93. PubMed
4. Özden VE, Osman WS, Morii T, et al. What percentage of patients are dissatisfied post-primary total hip and total knee arthroplasty? J Arthroplasty. 2025;40(2S1):S55-S56. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

5. Gunaratne R, Pratt DN, Banda J, et al. Patient dissatisfaction following total knee arthroplasty: a systematic review of the literature. J Arthroplasty. 2017;32(12):3854-3860. CrossRef PubMed
6. Hafkamp FJ, Gosens T, de Vries J, et al. Do dissatisfied patients have unrealistic expectations? A systematic review and best-evidence synthesis in knee and hip arthroplasty patients. EFORT Open Rev. 2020;5(4):226-240. CrossRef PubMed
7. Nikolajsen L, Brandsborg B, Lucht U, et al. Chronic pain following total hip arthroplasty: a nationwide questionnaire study. Acta Anaesthesiol Scand. 2006;50(4):495-500. CrossRef PubMed
8. Belford K, Gallagher N, Dempster M, et al. Psychosocial predictors of outcomes up to one year following total knee arthroplasty. Knee. 2020;27(3):1028-1034. CrossRef PubMed
9. Bletterman AN, de Geest-Vrolijk ME, Vriezekolk JE, et al. Preoperative psychosocial factors predicting patient's functional recovery after total knee or total hip arthroplasty: a systematic review. Clin Rehabil. 2018;32(4):512-525. CrossRef PubMed
10. Giusti EM, Lacerenza M, Gabrielli S, et al. Psychological factors and trajectories of postsurgical pain: A longitudinal prospective study. Pain Pract. 2022;22(2):159-170. CrossRef PubMed
11. Giusti EM, Lacerenza M, Manzoni GM, et al. Psychological and psychosocial predictors of chronic postsurgical pain: a systematic review and meta-analysis. Pain. 2021;162(1):10-30. CrossRef PubMed
12. Lopez-Olivo MA, Ingleshwar A, Landon GC, et al. Psychosocial determinants of total knee arthroplasty outcomes two years after surgery. ACR Open Rheumatol. 2020;2(10):573-581. CrossRef PubMed
13. Bellamy R. An introduction to patient education: theory and practice. Med Teach. 2004;26(4):359-365. CrossRef PubMed
14. Anger M, Valovska T, Beloeil H, et al.; PROSPECT Working Group* and the European Society of Regional Anaesthesia and Pain Therapy. PROSPECT guideline for total hip arthroplasty: a systematic review and procedure-specific postoperative pain management recommendations. Anaesthesia. 2021;76(8):1082-1097. CrossRef PubMed
15. National Institute for Health and Care Excellence (NICE). Joint replacement (primary): hip, knee and shoulder. 2020 Online. (Accessed June 2025)
16. Jette DU, Hunter SJ, Burkett L, et al.; American Physical Therapy Association. Physical therapist management of total knee arthroplasty. Phys Ther. 2020;100(9):1603-1631. CrossRef PubMed
17. Louw A, Diener I, Butler DS, et al. Preoperative education addressing postoperative pain in total joint arthroplasty: review of content and educational delivery methods. Physiother Theory Pract. 2013;29(3):175-194. CrossRef PubMed
18. McDonald S, Page MJ, Beringer K, et al. Preoperative education for hip or knee replacement. Cochrane Database Syst Rev. 2014;2014(5):CD003526. CrossRef PubMed
19. Aydin D, Klit J, Jacobsen S, et al. No major effects of preoperative education in patients undergoing hip or knee replacement-a systematic review. Dan Med J. 2015;62(7):A5106. PubMed
20. Longo UG, De Salvatore S, Rosati C, et al. The impact of preoperative education on knee and hip replacement: a systematic review. Osteology. Osteology. 2023;3(3):94-112. CrossRef
21. Moyer R, Ikert K, Long K, et al. The value of preoperative exercise and education for patients undergoing total hip and knee arthroplasty: a systematic review and meta-analysis. JBJS Rev. 2017;5(12):e2. CrossRef PubMed
22. Toronto Metropolitan University. Postoperative patient education: a systematic review. Online. (Accessed June 2025)
23. Vahabi S, Karimi A, Beiranvand S, et al. Evaluating memory dysfunction after spinal anesthesia among patients undergoing elective surgery: descriptive-analytical study. Ann Med Surg (Lond). 2021;62:168-174. CrossRef PubMed
24. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372(71):n71. CrossRef PubMed
25. van den Borne HW. The patient from receiver of information to informed decision-maker. Patient Educ Couns. 1998;34(2):89-102. CrossRef PubMed
26. Nilsdotter AK, Lohmander LS, Klässbo M, et al. Hip disability and osteoarthritis outcome score (HOOS)–validity and responsiveness in total hip replacement. BMC Musculoskelet Disord. 2003;4(1):10. CrossRef PubMed
27. Roos EM, Roos HP, Lohmander LS, et al. Knee Injury and Osteoarthritis Outcome Score (KOOS)–development of a self-administered outcome measure. J Orthop Sports Phys Ther. 1998;28(2):88-96. CrossRef PubMed
28. Salaffi F, Leardini G, Canesi B, et al.; GOnorthrosis and Quality Of Life Assessment (GOQOLA). Reliability and validity of the Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index in Italian patients with osteoarthritis of the knee. Osteoarthritis Cartilage. 2003;11(8):551-560. CrossRef PubMed
29. Saadat P, Pereira TV, Lalji R, et al. Evidence-based hierarchy of pain outcome measures for osteoarthritis clinical trials and meta-analyses. Osteoarthritis Cartilage. 2025;33(1):42-49. CrossRef PubMed
30. King LK, Hawker GA, Stanaitis I, et al. Minimal clinically important difference for improvement in six-minute walk test in persons with knee osteoarthritis after total knee arthroplasty. BMC Musculoskelet Disord. 2022;23(1):307. CrossRef PubMed
31. Yuksel E, Unver B, Kalkan S, et al. Reliability and minimal detectable change of the 2-minute walk test and Timed Up and Go test in patients with total hip arthroplasty. Hip Int. 2021;31(1):43-49. CrossRef PubMed
32. Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ. 1992;305(6846):160-164. CrossRef PubMed
33. Webster KE, Feller JA. Comparison of the short form-12 (SF-12) health status questionnaire with the SF-36 in patients with knee osteoarthritis who have replacement surgery. Knee Surg Sports Traumatol Arthrosc. 2016;24(8):2620-2626. CrossRef PubMed
34. Fydrich T, Dowdall D, Chambless DL. Reliability and validity of the Beck Anxiety Inventory. J Anxiety Disord. 1992;6(1):55-61. CrossRef
35. Alshahrani MS, Reddy RS, Tedla JS, et al. Association between kinesiophobia and knee pain intensity, joint position sense, and functional performance in individuals with bilateral knee osteoarthritis. Healthcare (Basel). 2022;10(1):120. CrossRef PubMed
36. Schwarzer R, Jerusalem M, Weinman J, et al. Generalized self-efficacy scale. Meas Health Psychol Users Portf Causal Control Beliefs Windsor; 1995.
37. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al.; PRISMA-S Group. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev. 2021;10(1):39. CrossRef PubMed
38. Guimarães NS, Ferreira AJF, Ribeiro Silva RC, et al. Deduplicating records in systematic reviews: there are free, accurate automated ways to do so. J Clin Epidemiol. 2022;152:110-115. CrossRef PubMed
39. Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan—a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210. CrossRef PubMed

40. WebPlotDigitizer - Extract data from plots, images, and maps. Online. (Accessed June 2025)
41. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomized trials. BMJ. 2019;366:l4898. CrossRef PubMed
42. McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): an R package and shiny web app for visualizing risk-of-bias assessments. Res Synth Methods. 2021;12(1):55-61. CrossRef
43. Cochrane Handbook for Systematic Reviews of Interventions. 7.7.3.2 Obtaining standard deviations from standard errors and confidence intervals for group means. Online. (Accessed June 2025)
44. Cochrane Handbook for Systematic Reviews of Interventions (current version). Online. (Accessed June 2025)
45. Cochrane. Learn. Free resources and tools for learning about health evidence. Online (Accessed June 2025)
46. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ. 2020;368:l6890. CrossRef PubMed
47. Guyatt G, Wang Y, Eachempati P, et al. Core GRADE 4: rating certainty of evidence-risk of bias, publication bias, and reasons for rating up certainty. BMJ. 2025;389:e083864. CrossRef PubMed
48. Guyatt GH, Oxman AD, Vist GE, et al.; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-926. CrossRef PubMed
49. GRADEpro. Online. (Accessed June 2025)
50. Salaffi F, Stancati A, Silvestri CA, et al. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. Eur J Pain. 2004;8(4):283-291. CrossRef PubMed
51. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Routledge; 1988, CrossRef.
52. Emara AK, Pasqualini I, Jin Y, et al.; Cleveland Clinic Arthroplasty Group. What are the diagnosis-specific thresholds of minimal clinically important difference and patient acceptable symptom state in Hip Disability and Osteoarthritis Outcome Score after primary total hip arthroplasty? J Arthroplasty. 2024;39(7):1783-1788.e2. CrossRef PubMed
53. Timmers T, Janssen L, van der Weegen W, et al. The effect of an app for day-to-day postoperative care education on patients with total knee replacement: randomized controlled trial. JMIR Mhealth Uhealth. 2019;7(10):e15323. CrossRef PubMed
54. Pronk Y, Peters MCWM, Sheombar A, et al. Effectiveness of a mobile eHealth App in guiding patients in pain control and opiate use after total knee replacement: randomized controlled trial. JMIR Mhealth Uhealth. 2020;8(3):e16415. CrossRef PubMed
55. Chang HL, Hsu MF, Wong TH, et al. Effects of a hybrid teaching program on lower limb muscle strength, knee function, and depression in older adults after total knee replacement: a randomized controlled trial. Res Gerontol Nurs. 2024;17(1):31-40. CrossRef PubMed
56. Cetinkaya Eren O, Buker N, Tonak HA, et al. The effect of video-assisted discharge education after total hip replacement surgery: a randomized controlled study. Sci Rep. 2022;12(1):3067. CrossRef PubMed
57. Pellegrini CA, Brown D, DeVivo KE, et al. Promoting physical activity via physical therapist following knee replacement: a pilot randomized controlled trial. PM R. 2023;15(8):965-975. CrossRef PubMed
58. Wang Q, Hunter S, Lee RLT, et al. The effectiveness of a mobile application-based programme for rehabilitation after total hip or knee arthroplasty: a randomized controlled trial. Int J Nurs Stud. 2023;140:104455. CrossRef PubMed
59. Better Health Channel. Patient-centred care explained. Online. (Accessed June 2025)
60. Kreuter MW, Oswald DL, Bull FC, et al. Are tailored health education materials always more effective than non-tailored materials? Health Educ Res. 2000;15(3):305-315. CrossRef PubMed
61. Ricci L, Villegente J, Loyal D, et al. Tailored patient therapeutic educational interventions: a patient-centred communication model. Health Expect. 2022;25(1):276-289. CrossRef PubMed
62. Kreuter MW, Wray RJ. Tailored and targeted health communication: strategies for enhancing information relevance. Am J Health Behav. 2003;27(1)(suppl 3):S227-S232. CrossRef PubMed
63. Keadle SK, McKinnon R, Graubard BI, et al. Prevalence and trends in physical activity among older adults in the United States: a comparison across three national surveys. Prev Med. 2016;89:37-43. CrossRef PubMed
64. Yeh MY, Wu SC, Tung TH. The relation between patient education, patient empowerment and patient satisfaction: a cross-sectional-comparison study. Appl Nurs Res. 2018;39:11-17. CrossRef PubMed
65. Losi S, Berra CCF, Fornengo R, et al. The role of patient preferences in adherence to treatment in chronic disease: a narrative review. Drug Target Insights. 2021;15:13-20. CrossRef PubMed
66. Yildiz T, Gürkan S, Gür Ö, et al. Effect of standard versus patient-targeted in-patient education on patients' anxiety about self-care after discharge from cardiovascular surgery clinics. Cardiovasc J S Afr. 2014;25(6):259-264. CrossRef PubMed
67. Paterick TE, Patel N, Tajik AJ, et al. Improving health outcomes through patient education and partnerships with patients. Proc Bayl Univ Med Cent. 2017;30(1):112-113. CrossRef PubMed
68. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014;348(mar07 3):g1687. CrossRef PubMed
69. Piano L, Ritorto V, Vigna I, et al. Individual patient education for managing acute and/or subacute low back pain: little additional benefit for pain and function compared to placebo. A systematic review with meta-analysis of randomized controlled trials. J Orthop Sports Phys Ther. 2022;52(7):432-445. CrossRef PubMed
70. Piano L, Audasso P, Benzi L, et al. Individual education for patients with chronic low back pain: likely a clinically-relevant effect for long-term disability compared to non-educational interventions. A systematic review with meta-analysis. J Orthop Sports Phys Ther. 2025;55(5):331-343. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Arch Physioter 2025; 15: 326
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3879
ERRATUM

Erratum in: Evaluation and synthesis of physiotherapy protocols for Femoroacetabular Impingement Syndrome (FAIS): a scoping review

Céline Berguerand, Noemie Rossi, Rahel Caliesch

School of Health and Sciences, HES-SO Valais-Wallis - Switzerland

In the article "Evaluation and synthesis of physiotherapy protocols for Femoroacetabular Impingement Syndrome (FAIS): a scoping review" (1), which appeared in Volume 15, Issue 1 of Archives of Physiotherapy, Figure 4 was incorrect. The correct Figure 4 is reported here:

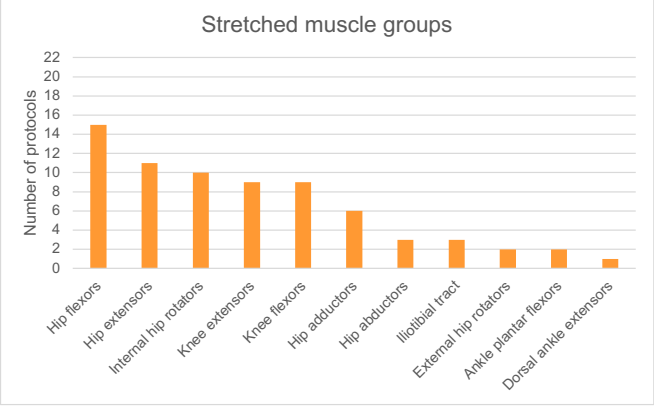


FIGURE 4 - Lower limb muscle groups cited in 22 stretching protocols.

We apologize for any inconvenience caused to the readers by these error, which does not affect the overall results or conclusions of the study.

The final version of this article is available online and includes a reference to this correction.

Reference

1. Berguerand C, Rossi N, Caliesch R. Evaluation and synthesis of physiotherapy protocols for Femoroacetabular Impingement Syndrome (FAIS): a scoping review. Archives of Physiotherapy. 2025;15(1):165-173. CrossRef

Received: December 11, 2025
Accepted: December 15, 2025
Published online: December 22, 2025

Corresponding author:
Céline Berguerand
email: berguerandceline@gmail.com

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

Arch Physioter 2025; 15: 327-336
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3722
MASTERCLASS

Mastering the use of hand-held dynamometry in clinical practice

Claire Minshull

School of Health, Leeds Beckett University, Leeds, UK; Get Back To Sport, Nottingham, UK

ABSTRACT

Measures of muscle performance, such as strength and the rate of force development (RFD), are important for function, rehabilitation, and successful aging. Clinicians seeking to use objective measurement methods of muscle performance in support of their assessments and rehabilitation programs have many affordable dynamometry options. However, substantive differences exist between devices on important characteristics such as sampling frequency, load capacity and determining force onset, which dramatically affect the ability to obtain accurate estimates of muscle performance. The assessment environment and setup also require careful consideration. Busy clinicians are often unaware of the extent to which methodological variability and inconsistencies in testing protocols can inflate measurement error and render tests insensitive to change. Where data inform treatment and return-to-play decisions (vs. motivational aid), ensuring validity and reliability is paramount, particularly given that clinicians typically assess individual, not group performance. This is because ascertaining change or difference in intra-individual performance demands a greater level of measurement precision compared to assessing performance between groups of people.

This evidence-informed Masterclass will exemplify some of the critical technical and methodological factors that intrude on measurement accuracy. It will provide readers with the knowledge: how to critically evaluate the utility of dynamometers, answering the question, which to buy and why? How to construct assessment protocols to improve quality data collection, and how to understand what constitutes real change in performance beyond "differences" caused by measurement error.

Keywords: Exercise therapy, Muscle strength, Muscle strength dynamometer, Rehabilitation

Introduction

Do you have access to a dynamometer, and are you proficient in its use? Are you aware of its sampling frequency and technical limitations? If you are considering the acquisition of a dynamometer, what are your selection criteria, and upon what evidence or rationale is your decision based?

Measures of muscle performance, such as strength and rate of force development (RFD), are important rehabilitation outcomes; are essential for successful aging and falls avoidance, and they are important indicators of performance in sports (1-3). Knowledge of muscle performance can also help to guide clinical decisions, stratify patient treatment groups and provide prognostic markers of health (4). Clinicians seeking to obtain objective measures of muscle performance in support of their assessments and rehabilitation programming have a multitude of affordable hand-held dynamometry (HHD) options to choose from. However, the assumption that generating a numerical output equates to a valid and reliable measure is misleading.

The market for HHD has expanded rapidly, offering a wide range of affordable and portable devices, often marketed with claims of high reliability and accuracy. Yet important differences in technical specifications, such as sampling frequency and force onset detection, exist. These parameters significantly influence the fidelity of performance estimates (5,6). Consideration should also be given to load capacity, particularly in settings where high force outputs are anticipated, such as during the assessment of athletes and whole-body performance, for example, the isometric mid-thigh pull (IMTP). The absence of regulatory standards or consistent validation procedures across devices compounds the challenge, leaving clinicians to navigate a landscape of devices of variable quality and unclear clinical utility.

In addition to these technical issues, methodological inconsistencies permeate clinical practice (7). Inaccurate or poorly standardized measurements risk not only being clinically uninformative but may also lead to misinformed decisions that compromise patient outcomes and quality and direction of care. Heterogenous testing protocols, even when done well, can yield variable results, and thus the establishment of "normative values" is a futile exercise.

To add to the mix, clinicians almost exclusively measure individual performances (intra-individual), such as tracking

Received: August 1, 2025
Accepted: December 23, 2025
Published online: December 31, 2025

Corresponding author:
Claire Minshull
email: minshullc@hotmail.com

© 2025 The Authors. This article is published by AboutScience and licensed under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).
Commercial use is not permitted and is subject to Publisher's permissions. Full information is available at www.aboutscience.eu

progress across sessions or measuring the difference between limbs. The precision required to detect meaningful change in individuals is substantially higher than when comparing group means (8). Without an understanding of the sources and magnitude of measurement error, tests risk being statistically and clinically insensitive to physiologic change. Thus, when "differences" between test scores are observed, the clinician cannot be sure if this represents real change or is an artefact of measurement error.

This evidence-informed Masterclass aims to equip clinicians with the conceptual and practical tools necessary to critically evaluate and apply HHD in practice. It will examine the impact of technical device characteristics, methodological choices, and operator skill on measurement quality. It will also address key questions such as: What makes a device fit-for-purpose? How can protocols be constructed to minimize error and maximize sensitivity? And, how can clinicians distinguish between true performance change and variation attributable to measurement error?

Ultimately, this manuscript aims to enhance clinicians' confidence in selecting the appropriate tool/s for their needs and provide the essential components to designing robust muscle performance assessments that can meaningfully inform patient care.

The assessment setting

What are we assessing; what do we want to achieve or understand?

A growing body of research highlights the importance of muscle performance indices, such as strength and RFD, for quality of life (9,10). Current guidelines also recommend limb symmetry indices greater than 90% for a safe return to play (RTP) (11), with evidence showing that deficits in muscle performance may increase the risk of injury (12,13). As a result, the use of dynamometry to obtain objective measures that support rehabilitation and clinical decision-making across multiple patient groups is becoming increasingly common in contemporary practice.

In clinical practice, two testing scenarios predominate: 1 - evaluation of both limbs (affected and unaffected) within a single session to calculate a Limb Symmetry Index (LSI), which is the performance of the affected/injured limb expressed as a percentage of the non-affected/uninjured limb; 2 - tracking performance metrics over time to gauge the effectiveness of rehabilitation and conditioning programmes. Each testing application (e.g., within session [intra-session] and/or between days [inter-day]) involves assessing an individual's performance and comparing results between limbs, over time, or both.

This type of intra-individual testing requires a more stringent level of measurement precision than testing a group of people (8). In empirical research, methodologies document the number of recruited participants to each experimental or control group. Calculated using data from prior and/or pilot studies, participant numbers are manipulated to achieve the required level of statistical power to detect the level of change or difference that the researchers want to see (14). These calculations take into account factors that influence measurement precision, such as variability.

In contrast, when conducting intra-individual assessments where the focus is on detecting change of difference within a single person, the measurement precision must be sufficient to distinguish true change from random variability or measurement error in that single person. Without this heightened precision, any observed differences risk being indistinguishable from normal performance fluctuations and experimental inconsistencies, undermining the clinical or experimental utility of the assessment (8).

A comprehensive review of the clinimetric properties of HHD is beyond the scope of this Masterclass; however, it is important to acknowledge that numerous studies have demonstrated its strong reliability across a range of muscle groups and clinical populations. Readers are directed to systematic reviews such as Chamorro et al. (15) and empirical studies, such as Mentiplay et al. (16), for further detailed evidence on the reliability and measurement characteristics of HHD across different testing protocols and populations.

In summary, the majority of assessments within clinical practice are to detect change in intra-individual performance over time or between limbs and to do this, precision of measurement is important.

Which indices of performance?

Muscle strength, experimentally termed peak force (PF), and RFD are two key indices of muscle performance, which offer critical insight into neuromuscular function and have been assessed within neuromuscular physiology settings for decades (17,18). Recent developments in HHDs have enabled the quick and easy procurement of PF and RFD data, with applications in both clinical and athletic or performance-focused environments. When appropriately obtained and reported, these data can help identify performance deficits, guide individualized rehabilitation programmes and RTP decisions, and quantify improvements resulting from therapeutic interventions. Moreover, they may also provide estimates of dynamic stabilization capabilities during joint loading (19-21), with deficits potentially indicating elevated injury risk or incomplete recovery (22).

Peak Force (PF)

Peak force (PF) is the maximal voluntary contractile force of a muscle in a single contraction, typically measured in newtons (N). With a robust testing configuration and reproducible methods, PF is a relatively easy metric to obtain and, by comparison to other indices of neuromuscular function, can be relatively impervious to minor deviations in things like dynamometer technical specifications (23). Participants are instructed to produce a maximal voluntary contraction (MVC) as "hard and as fast" as they can, and the focus of the measurement is on the PF output regardless of time.

Rate of Force Development (RFD)

Rate of force development (RFD) is a measure of explosive force production, and the focus of the measurement is on the rapidity of force production, often irrespective of the peak. RFD can be reported in several different ways; the

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

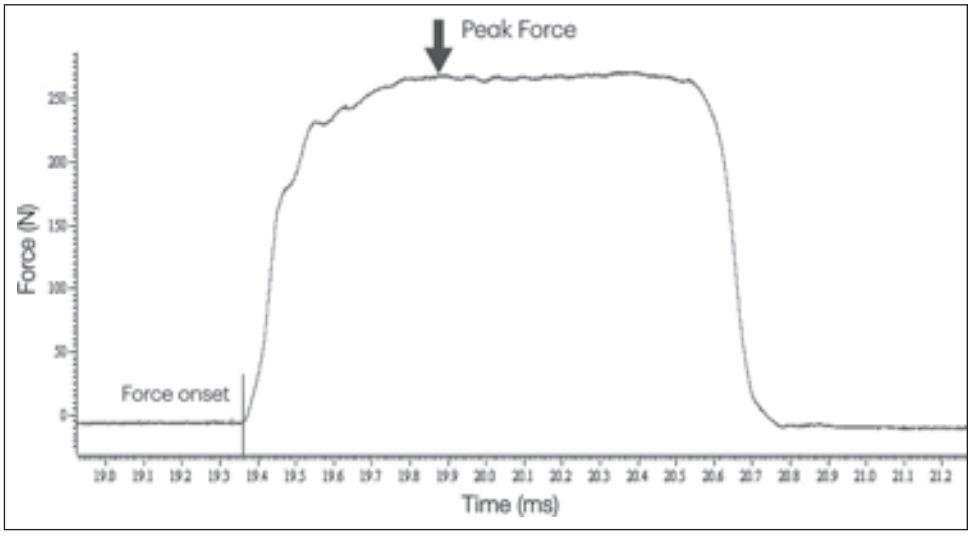


FIGURE 1 - Individual quadriceps force-time curve, 6 months post ACL reconstruction. Sampling Frequency: 2000 Hz. Unpublished data.

most common method involves calculating the gradient of the slope of the force-time (F-T) curve within specified time epochs from the onset of force, resulting in a newtons per second ($N \cdot s^{-1}$) value. For example, the average force produced over 0-50 ms and 0-100 ms captures early-phase RFD and 100-150 ms and 100-200 ms reports late-phase RFD (24,25). Alternative methods include "peak RFD," which is the steepest slope of the force-time signal, no matter where it occurs, and "zero to peak," which provides an average RFD across the whole force-time curve. The index of RFD involves rapid changes in force over extremely short time frames, and as such, minor deviations in methodology and differences in dynamometer technical specifications can heavily influence the accuracy of data (23,26).

It is beyond the scope to provide a detailed review of the rationale for each method of testing and RFD index selection; however, Figure 2 simply exemplifies the consequences of RFD index selection on performance interpretation.

Zero-to-peak, often reported with contemporary HHDs, can provide a general marker of RFD performance; however, it lacks temporal (time-based) precision. It may substantially underestimate RFD by averaging over a broader time interval, thus masking rapid changes in force output (27). Furthermore, zero-to-peak is heavily influenced by the threshold that determines force onset. Insensitive devices may use a large threshold, which requires participants to produce significant force before recording commences.

The zero-to-peak index is less suitable when rapid force production is the primary focus. In contrast, analyzing smaller specific segments of the F-T curve provides more detailed insights, but doing so requires accurate data collected over short sampling windows. This depends heavily on the characteristics of the hand-held dynamometer (HHD) and associated software, particularly the use of a high sampling frequency (see below).

In summary, clinicians seeking to accurately assess muscle performance metrics, PF and RFD should recognize that both index selection and dynamometer characteristics will influence performance interpretation.

Dynamometers and technical specifications

Selecting the right tool for the job

In many respects, the proliferation of the number of commercially available hand-held and other dynamometers is a positive development. Historically, the objective measurement of muscle force production was largely confined to research laboratories or clinical environments with access to high-cost isokinetic or custom-built dynamometry systems.

However, the increased accessibility and commercialization of devices has, in some cases, been accompanied by insufficient attention to device technical specifications and methodological rigour, for example, slow sampling frequencies and variable testing methods, respectively. Consequently, clinicians may adopt these tools without critical evaluation of their measurement properties or adequate training in their use. This practice raises important concerns regarding the reliability, validity, and overall utility of the data being collected.

Types of hand-held dynamometers (HHDs)

Many different types of dynamometers exist; the focus here is on HHDs, albeit many of the technical and methodological considerations apply to other forms of assessment.

Hand-held dynamometers are portable devices that measure muscle force production. Most new generations of HHDs have smart device app software and offer PF and some index of RFD as an output. Devices can be categorized according to the method of force application. Compression-type dynamometers assess force when the user pushes against the device, for example, during isometric contractions where the dynamometer is placed between the limb and a fixed surface, tester, or secured using strapping (Fig. 3a). Tension-type dynamometers measure force through pulling, often via straps or cables attached to the limb and anchored to a fixed point (Figs 3b and c).

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

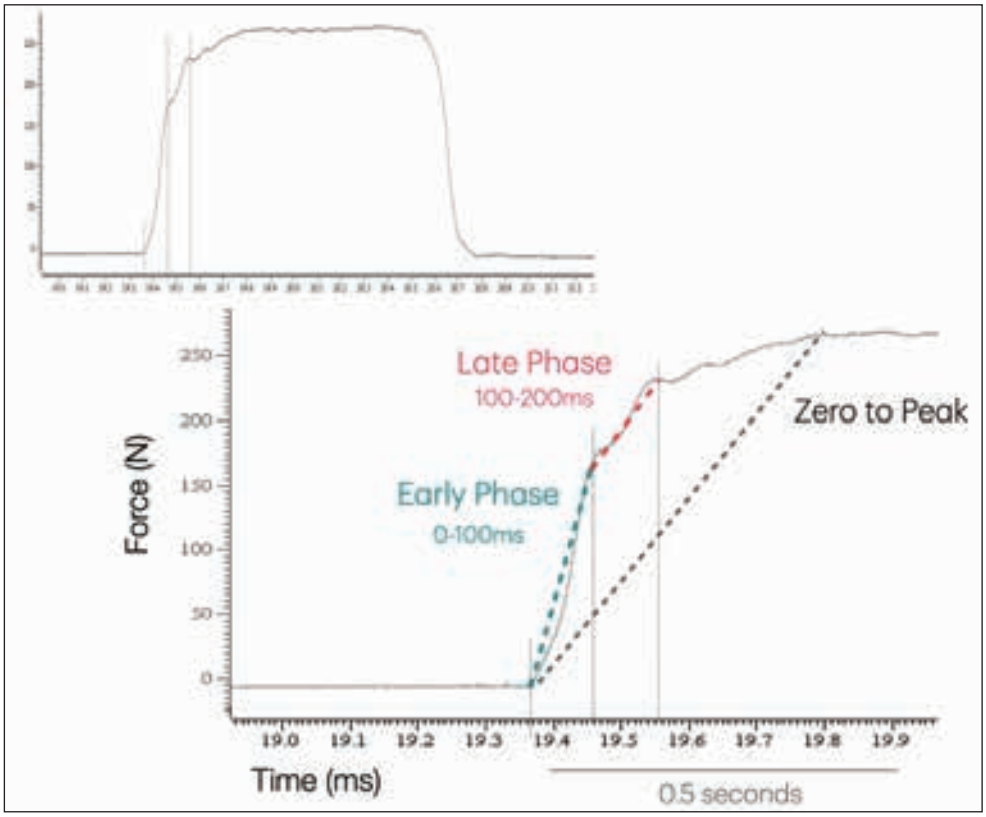


FIGURE 2 - Individual quadriceps force-time curve, 6 months post ACL reconstruction, illustrating three methods of reporting RFD. Sampling Frequency: 2000 Hz. Unpublished data.

FIGURE 3 - Assessment of quadriceps muscle force using HHD. 3a: Compression-type dynamometer secured using strapping (VALD DynaMoLite); 3b: Tension-type dynamometer secured using strapping (Kinvent Pull). Plinth-based testing is pragmatic in clinical settings; there may be some compromise in the dynamometer and strapping setup. 3c: Adaptation of the knee extensor machine to secure tension-type dynamometer (VALD DynaMoPlus).

Important dynamometer technical specifications

When selecting a measurement tool, financial outlay is important; however, priority should be given to the technical specifications of the selected tool to enable the accurate measurement of the data the clinician seeks to procure.

Sampling frequency

A critical aspect of dynamometry is sampling frequency. This refers to the number of data points of a signal collected per second, and it is measured in hertz (Hz). 1 Hz is one sample per second, 100 Hz is 100 samples per second, etc.

Sampling frequency is important in the assessment of muscle force production, especially for events that occur quickly, like RFD (5). From a rehabilitation and training perspective, different phases of RFD are driven by different physiologic events. Early phase RFD is influenced mainly by neural events, such as neural drive and late phase by muscular and morphological characteristics, such as muscle cross-sectional area (28). Should there be a need to create

specific rehabilitation and training programmes to target different parts of the F-T curve, higher sampling frequencies are required to record accurately over small sampling windows (<100 ms). Too slow and information will be lost, resulting in poor sensitivity and inaccurate representation of RFD.

To exemplify, in well-conditioned people and in non-athletes who are habituated to isometric testing, PF can occur within 300ms of force onset—one third of a second (29). In other populations, it's not dissimilar, for example, individuals with mild knee osteoarthritis (350 ms: (30)). Figure 4a shows a section of the F-T curve of a knee extensor assessment of an individual's non-injured limb prior to ACL reconstruction (time to PF 350 ms). The original data is sampled at 2000 Hz; the lines overlaid represent an estimate of what the data may look like when obtained at a sampling frequency of 10 Hz, which is the sampling frequency of the EasyForce (Meloq) HHD at the time of writing. With a sampling frequency of just 10 Hz, only three data points are obtained across a 300 ms time period, contrasted to 600 data points with a sampling frequency of 2000 Hz. This results in a different picture of the F-T curve. A low sampling frequency that captures only a few data points during rapid force production may underestimate the true change in muscle force over time and therefore distort the calculated RFD. In contrast, a high sampling frequency (e.g., 2000 Hz) provides much greater temporal resolution to more accurately represent the rapid changes in force that can define the true RFD. Users must also be aware that slow sampling frequencies, coupled with high threshold force onsets, can obscure the commencement of force production.

What is an adequate sampling frequency?

The sampling frequencies available across commercially available HHDs range from 10 Hz (EasyForce [Meloq]) to 1000 Hz (Kinvent Push/Pull [Physio Kinvent]) and recently 1200 Hz (DynamoMax [VALD Performance]). The minimal required sampling frequency will depend on the data acquisition and its importance.

The index PF is fairly resilient to sampling frequency deviations; quality data can be obtained at a sampling rate of 100 Hz with minimal differences with higher sampling rates (6). For lower sampling rates, information such as the true PF may be underrepresented during explosive contractions, and users may wish to consider maintaining the maximal contraction for 2 or 3 seconds to obtain a reliable estimate within the available sampling window.

It is widely regarded that much higher sampling frequencies are required to accurately measure RFD, especially in the early phase of the force-time curve and from explosive contractions. Several data-driven studies suggest a minimum sampling frequency of 500 Hz for PF and RFD (e.g., 6,31), with even higher rates (1000 Hz/>) recommended where high fidelity data is required. Sampling frequencies below 1000 Hz have been shown to underestimate peak and early-phase RFD due to insufficient temporal resolution to capture the steep initial rise in force (5,6). It is important to note that these findings are based on data collected using isokinetic dynamometry or wired force plates, methods that differ from most HHDs, whereby data are transmitted via Bluetooth. Wireless (Bluetooth) transmission may lead to data packet loss (32), potentially affecting the accuracy of RFD calculations. If RFD is a critical metric for clinical decision-making, practitioners would be prudent to use a device with a sampling frequency of at least 1000 Hz to ensure sufficient data quantity and quality. Figure 5 shows a simple flow chart to help guide clinicians on the required sampling frequency for their data acquisition needs.

Some manufacturers offer exportation of raw data in CSV format. For the more technically-minded clinician, this enables manual calculation of RFD across the F-T curve by using customized scripts.

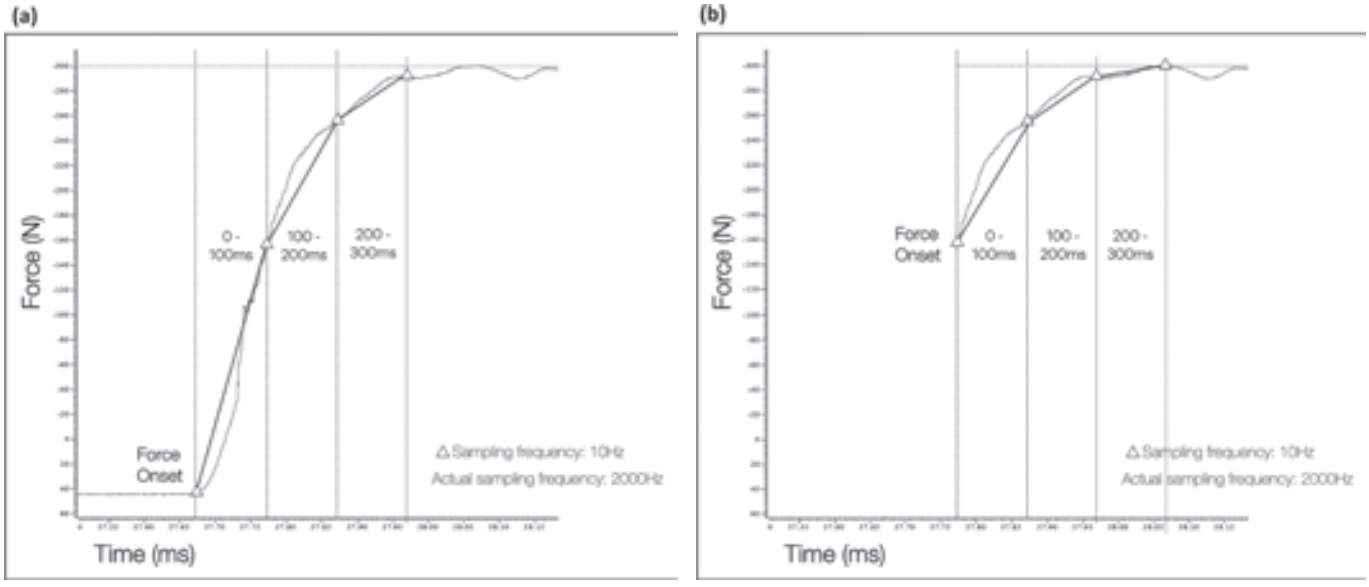


FIGURE 4 - a. Individual data from Minshull et al. (39). Quadriceps force-time curve, pre-ACL reconstruction, non-injured leg. Δ Example data captured at 10 Hz. b. Individual data from Minshull et al. (39). Quadriceps force-time curve, pre-ACL reconstruction, non-injured leg. Δ Example data captured at 10 Hz with a 20 N force onset threshold.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

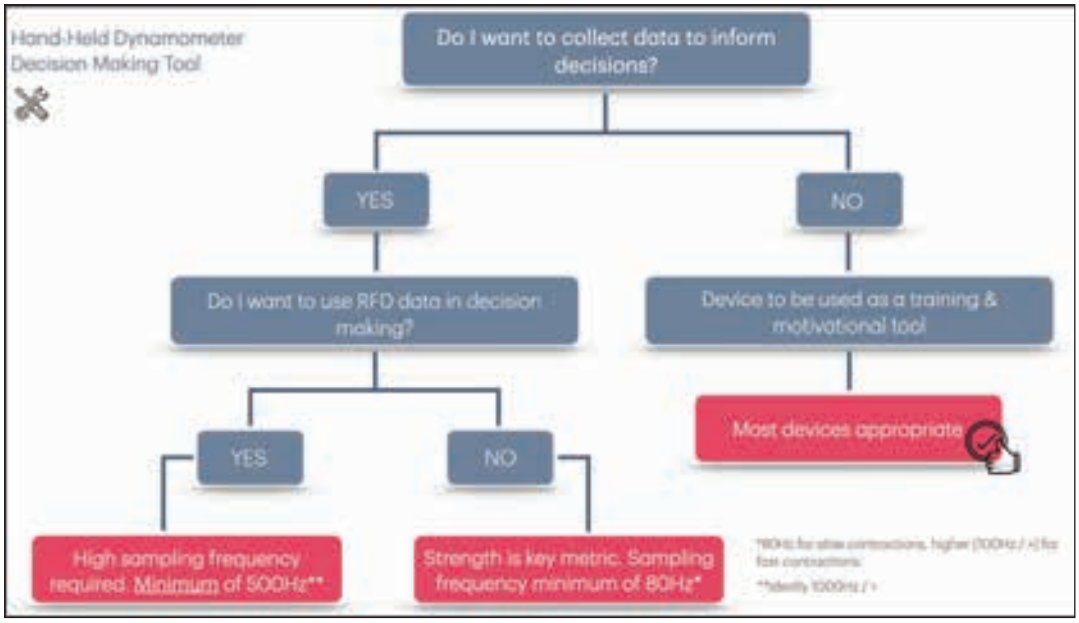


FIGURE 5 - HHD decision-making tool, reproduced with permission, Get Back To Sport 2025.

Reliability does not necessarily confer validity

A recent study (33) reported good to excellent inter- and intra-session reliability of the Microfet2 (Hoggan Scientific) in measuring RFD, supporting their recommendation of its use for assessing knee extensor RFD. The Microfet2 samples at 100 Hz, and thus the recommendation appears at odds with the literature cited above. However, reliability does not confer validity (i.e., the extent to which the test measures what it claims to measure). The slow sampling frequency captures only a small number of data points during the critical early phase of contraction. Under-sampling blurs the force-time curve, masking rapid changes in muscle force and reducing the device's ability to detect small but meaningful differences between trials (5), thereby compromising its validity. Consequently, between-trial differences may appear smaller than they actually are, leading to inflated within-trial consistency and potentially misleading conclusions about reliability.

Force Onset

The threshold for determining force-onset with HHD software isn't always adjustable by end-users; however, it is important to be aware of how this is determined. High absolute values signalling force onset may compromise accurate evaluation of RFD, particularly in smaller muscles. A 20 N onset threshold for registration of muscle force production, with a shoulder internal rotation strength of, for example, 100 N, represents 20% of the total force production. In this instance, much of the early phase of the F-T curve could be missed, precluding the accurate determination of RFD.

Some manufacturers recommend a pre-tensioning setup, whereby the participant produces a low-level contraction to remove the slack from the securing straps prior to MVC. Compliance, or slack within a measurement setting, should be removed to improve assessment accuracy; however, pretensioning the muscle prevents an accurate picture of early-phase RFD as the contraction does not start from a resting level. Furthermore, depending on the strength of contraction, the prior contraction may potentiate the preceding contraction, thereby enhancing force output compared to that elicited from a relaxed state (34).

In summary, the clinician must determine which data they want to obtain to aid their decision-making in advance of the procurement of a device. Secondarily, evaluation of the technical specifications of the tools available will determine whether or not they are fit for purpose. Readers are referred to Online for further HHD comparison resources.

Measurement error

Testing performance is important, but not as important as testing performance accurately.

Understanding Measurement Error

The greater the error, or variability of a measurement, the less confident we can be about the score representing the person's true value (35), and the less we can rely on it to help make data-informed decisions, for example, do we keep going with the rehabilitation or has that person's strength increased by the target amount?

To illustrate, Figure 6 shows a hypothetical scenario in which quadriceps muscle strength was assessed using an HHD across multiple visits during the course of a rehabilitation programme. The clinician is interested in monitoring the patient's progress over time following the implementation of a targeted muscle strengthening intervention. A central question arises: at what point does a meaningful improvement in muscle strength occur?

Although a calculated change of +8% in force output relative to baseline suggests an improvement in strength at week 3, this observation cannot be confirmed with confidence

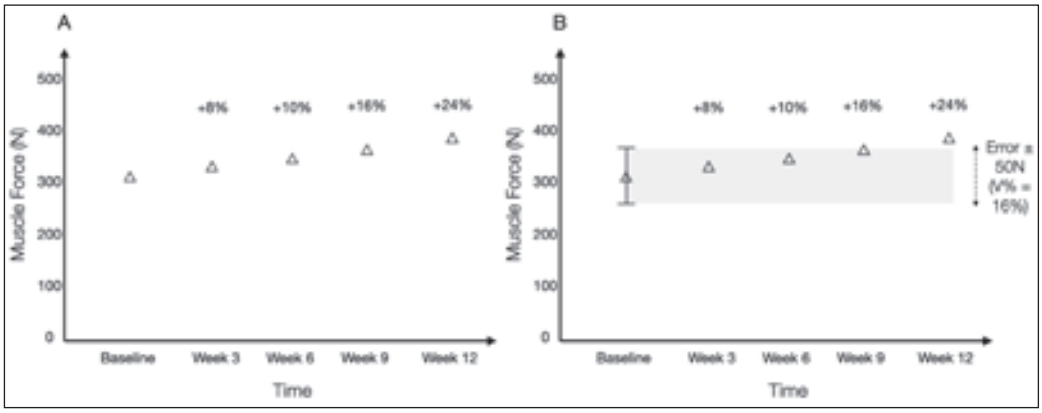


FIGURE 6 - Hypothetical quadriceps PF data collected at baseline and four additional times across a 12-week rehabilitation program. A: showing percentage change relative to baseline value; B: greyed area showing measurement error: coefficient of variation [V%] (multiplied by 1.96 to represent 95% of the variance) relative to baseline value.

without knowledge of the measurement error. For instance, if testing procedures were inconsistent, involving non-standardized participant positioning, inadequate stabilization of the dynamometer, or a lack of formal protocol, then measurement error is likely to intrude considerably upon the precision of measurement and the inference of change.

Measurement error can be calculated in a number of ways, and readers are directed to the excellent paper by Atkinson and Nevill (8) for a comprehensive discussion of the various methods within the context of assessment of the individual, including, standard error of the measurement (SEM), Bland-Altman, coefficient of variation (V%), and intra-class correlation (ICC). It is unwise to rely on a sole index to describe measurement error; however, the simply calculable index: the coefficient of variation (V%), has been used here solely to illustrate the practical impact of measurement error in this hypothetical example.

The V% is an estimate of measurement error. Calculated as the standard deviation of a dataset as a percentage of its mean, it represents a method by which to express the variability of a single measurement, regardless of the unit of measurement (35).

$$V\% = \left(\frac{\text{Standard Deviation (pooled)}}{\text{Mean}} \right) \times 100$$

Here measurement error was calculated by the experimenters to be approximately ±50 N, or ±16.6% of the baseline strength value. A sample of >40-100 measures is generally recommended to provide stable error estimates (36). Only 68% of the error is represented with standard V% calculations. To capture 95% of the variance, the sample standard deviation should be multiplied first by 1.96 before being expressed as the coefficient of variation (V%) (8).

Consequently, only changes exceeding this threshold can be interpreted as genuine improvements. Based on this criterion, meaningful gains in muscle strength are not suggested until approximately week 9, and most likely week 12 into the intervention.

This example highlights how measurement error can undermine the utility of strength assessments. It also underscores the importance of understanding and, where possible, minimizing the sources of error, or at the very least being aware of the specific error margins associated with a given test. These considerations are equally relevant for:

- Inter-limb comparisons and when aiming to achieve, for example, an LSI of 90%. In such cases, it is essential to determine whether the error margins are adequate to detect a 10% inter-limb difference and, if not, efforts should be directed toward reducing measurement error (see next section).
- Determining minimal clinical important difference (MCID). For example, the MCID of grip strength in patients treated by volar locking plate fixation for a distal radius fracture has been reported as a decrease of 6.6 kg (63.7 N), or 19.5% (37). In situations of excessive measurement error, MCID may be undetectable in the individual patient.
- Comparing individual scores to normative values, where such data are available for the relevant population. Normative values, if representative of the individual tested and obtained using comparable procedures, can be useful for categorizing individuals as above or below average, setting rehabilitation goals, and, for example, identifying those at risk of sarcopenia (38). However, confident categorization of individual performance and subsequent clinical decision-making based on such data requires an understanding of the measurement error to ascertain whether observed deviations from population norms reflect true differences in muscle performance rather than normal variability or measurement noise.

Minimizing Measurement Error

Measurement errors are typically categorized into systematic and random. Systematic errors are consistent, repeatable inaccuracies often caused by flaws in measurement instruments or procedures that cause all measurements to deviate in the same direction from the true value. Random errors are unpredictable fluctuations that arise from uncontrollable variables, such as natural biological variation or inconsistent effort, leading to measurements that scatter around the true value without a consistent pattern (8). To facilitate pragmatic improvements in measurement accuracy, it may be more constructive to understand the underlying sources from which these errors originate. As such, errors will be categorized into technical and biological sources.

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

Technical error

Technical error encompasses variability in equipment, testing protocol, and setup procedures.

Equipment

If a force of, for example, exactly 10 kg (98.1 N) is applied repeatedly to an HHD under controlled conditions, a reliable device should yield nearly identical readings across trials. While minor deviations are expected, the degree of variability should be minimal. Most commercially available HHDs are designed to meet this level of precision, and the calibration data to verify this should be accessible from the manufacturer.

Protocol

Assuming a reliable, sensitive and calibrated device, most likely the largest source of technical error will arise from the protocol and test setup, that is to say, being able to replicate exactly the same procedures each time. This includes fundamental things like:

1. minimizing extraneous movement (e.g., ensuring the person's limb is secured)
2. standardizing instructions (and motivational prompts)
3. standardizing joint and HHD position
4. ensuring maximal effort is understood and performed
5. standardizing warm-up procedures
6. removing the "slack"/compliance from the measurement system if using strapping to secure the HHD

These might appear basic; however, each can influence the variability of force output substantially and confound the utility of a test. To exemplify, variation of knee flexion angle of just 10° resulted in >20% difference in isometric quadriceps strength (39). Clinicians can minimize changes in joint positioning within the test by paying strict attention to points 3 and 6 on each assessment occasion.

Other measures, such as RFD, can vary substantially with subtle differences in verbal cuing. Instructing someone to contract as "*hard and fast*" vs., "*as fast*" as possible can underestimate explosive performance by over 30% and may double measurement error (26). Clinicians can minimize performance variation by being focused on points 2 and 4 and providing sufficient opportunity for practice. In situations where RFD and PF data are required, but separation of assessments is not possible (40), then adequate data can be procured from instructions cueing "*hard and fast*" contractions, as long as participants are appropriately practised (5,41).

Biological error

Biological error refers to the variability of 'the person' and their responses, not to be confused with actual change caused by training or deconditioning.

The assessor

Within a hand-held setup, a significant source of error is the tester themselves (42,43). Assessor strength, fatigue and experience can all contribute unnecessary and additional amounts of error into assessments. Tethering the HHD securely to something immovable (see Figs 3a and b) is recommended to eliminate this effect.

On repeated tests, using the exact same setup, even when the device is tethered, strength scores will still vary a little; people are naturally variable. However, good methodological design should minimize this.

Learning

Learning effects refer to when people become better at the test with practice. Here, the assessor will observe continually increasing scores on repeated tests.

To minimize learning effects:

– Ensure the person/patient is familiarised with the test. Ensure they understand what to do by clearly and concisely explaining the test and the requirements. Check their understanding.
– Habituate the person to the test. Allow a few practice attempts before you start the test so that the person is able to constantly deliver the type of muscle contractions required. To save time, this can be incorporated into the end stages of a warm-up.

Fatigue

The effects of acute muscle fatigue will be visible by an observed decline in performance over repeated efforts. The assessor may see reduced PF and/or RFD scores on sequential efforts associated with insufficient recovery.

To minimize acute muscle fatigue

– Provide 10-30 s rest between maximal contractions to enable recovery. Observe the data in real time and provide more rest if the force profile is declining.
– Allow approximately 2 minutes of rest after every 5 maximal contractions. Multiple maximal efforts may be required to obtain scores of true maximal performance in individuals new to testing or those being assessed following recovery from injury. Split the efforts into sets and provide adequate time for recovery in between.

Summary

Refining Testing Protocols

Designing an effective protocol for assessment of muscle performance using HHD begins with clarity on assessment and clinical intent. Protocols must be aligned with specific goals, whether it is identifying muscular asymmetries, monitoring rehabilitation progress, or evaluating readiness to return to activity. The purpose will determine the schedule of testing and which performance indices are most relevant.

Each index of performance requires distinct methodological considerations. For example, accurate assessment of RFD, particularly in the early phase, demands high sampling rates and attention to force onset, while PF may allow for slightly more flexibility in setup but still requires consistent stabilization and instruction.

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com

From a resource standpoint, selecting the appropriate device is crucial. Clinicians should prioritize dynamometers with sufficient force capacity, adequate sampling frequency, and the ability to export raw data where needed. Devices should also allow for consistent tester positioning and secure fixation.

Minimizing error is central to robust protocol design. Simple but influential methodical issues such as strap compliance, inconsistent participant positioning and variable participant effort can compromise data integrity. Strategies such as standardized patient, HHD and limb positioning, use of external fixation, habituation trials, and consistent verbal cues can help mitigate these sources of variability. Where early phase RFD is important and where software settings like force-onset thresholds are not user-adjustable, clinicians should be aware of how these values are defined, as they can influence force measurements.

In summary, effective protocol design for HHD involves aligning the method with clinical goals, selecting appropriate performance indices, choosing reliable equipment, and controlling for known error sources. Clinicians are encouraged to pilot their own protocols and estimate within-session reliability (e.g., using the coefficient of variation with ≥10 participants), not as a definitive evaluation of measurement quality, but as a practical introduction to how error may infiltrate personal testing procedures. This process is intended to promote critical reflection on protocol design and to inform subsequent refinement or more comprehensive reliability assessment, balancing clinical feasibility with measurement integrity.

Disclosures

Conflict of interest: The author declares that they have no financial relationships or affiliations with any companies or organizations that could be perceived as a potential conflict of interest in relation to this work. However, the author has received products free of charge from multiple companies for research and/or evaluation purposes. These companies had no role in the design, execution, analysis, or writing of this manuscript.

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Reid KF, Fielding RA. Skeletal muscle power: a critical determinant of physical functioning in older adults. Exerc Sport Sci Rev. 2012;40(1):4-12. CrossRef PubMed
2. Lee J. Associations between handgrip strength and disease-specific mortality including cancer, cardiovascular, and respiratory diseases in older adults: a meta-analysis. J Aging Phys Act. 2020;28(2):320-331. CrossRef PubMed
3. Turpeinen JT, Freitas TT, Rubio-Arias JÁ, et al. Contractile rate of force development after anterior cruciate ligament reconstruction—a comprehensive review and meta-analysis. Scand J Med Sci Sports. 2020;30(9):1572-1585. CrossRef PubMed
4. Gil S, Jacob Filho W, Shinjo SK, et al. HCFMUSP COVID-19 Study Group. Muscle strength and muscle mass as predictors of hospital length of stay in patients with moderate to severe COVID-19: a prospective observational study. J Cachexia Sarcopenia Muscle. 2021;12(6):1871-1878. CrossRef PubMed
5. Maffiuletti NA, Aagaard P, Blazevich AJ, et al. Rate of force development: physiological and methodological considerations. Eur J Appl Physiol. 2016;116(6):1091-1116. CrossRef PubMed
6. Thompson BJ. Influence of signal filtering and sample rate on isometric torque – time parameters using a traditional isokinetic dynamometer. J Biomech. 2019;83:235-242. CrossRef PubMed
7. Stone CA, Nolan B, Lawlor PG, et al. Hand-held dynamometry: tester strength is paramount, even in frail populations. J Rehabil Med. 2011;43(9):808-811. CrossRef PubMed
8. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998;26(4):217-238. CrossRef PubMed
9. Wang DXM, Yao J, Zirek Y, et al. Muscle mass, strength, and physical performance predicting activities of daily living: a meta-analysis. J Cachexia Sarcopenia Muscle. 2020;11(1):3-25. CrossRef PubMed
10. Khodadad Kashi S, Mirzazadeh ZS, Saatchian V. A systematic review and meta-analysis of resistance training on quality of life, depression, muscle strength, and functional exercise capacity in older adults aged 60 years or more. Biol Res Nurs. 2023;25(1):88-106. CrossRef PubMed
11. Lynch AD, Logerstedt DS, Grindem H, et al. Consensus criteria for defining 'successful outcome' after ACL injury and reconstruction: a Delaware-Oslo ACL cohort investigation. Br J Sports Med. 2015;49(5):335-342. CrossRef PubMed
12. Achenbach L, Laver L, Walter SS, et al. Decreased external rotation strength is a risk factor for overuse shoulder injury in youth elite handball athletes. Knee Surg Sports Traumatol Arthrosc. 2020;28(4):1202-1211. CrossRef PubMed
13. Namazi P, Zarei M, Hovanloo F, et al. The association between the isokinetic muscle strength and lower extremity injuries in young male football players. Phys Ther Sport. 2019;39:76-81. CrossRef PubMed
14. Serdar CC, Cihan M, Yücel D, et al. Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. Biochem Med (Zagreb). 2021;31(1):010502. CrossRef PubMed
15. Chamorro C, Arancibia M, Trigo B, et al. Absolute reliability and concurrent validity of hand-held dynamometry in shoulder rotator strength assessment: systematic review and meta-analysis. Int J Environ Res Public Health. 2021;3;18(17):9293. CrossRef
16. Mentiplay BF, Perraton LG, Bower KJ, et al. Assessment of lower limb muscle strength and power using hand-held and fixed dynamometry: a reliability and validity study. PLoS One. 2015; 28;10(10):e0140822. CrossRef
17. Norman RW, Komi PV. Electromechanical delay in skeletal muscle under normal movement conditions. Acta Physiol Scand. 1979;106(3):241-248. CrossRef PubMed
18. Ismail HM, Ranatunga KW. Isometric contractions of normal and spastic human skeletal muscle. Muscle Nerve. 1981;4(3):214-218. CrossRef PubMed
19. Buckthorpe M, Roi GS. The time has come to incorporate a greater focus on rate of force development training in the sports injury rehabilitation process. Muscles Ligaments Tendons J. 2018;7(3):435-441. CrossRef PubMed
20. Domire ZJ, Boros RL, Hashemi J. An examination of possible quadriceps force at the time of anterior cruciate ligament injury during landing: a simulation study. J Biomech. 2011;44(8):1630-1632. CrossRef PubMed
21. Minshull C, Eston R, Bailey A, et al. Repeated exercise stress impairs volitional but not magnetically evoked electromechanical delay of the knee flexors. J Sports Sci. 2012;30(2):217-225. CrossRef PubMed
22. Grindem H, Snyder-Mackler L, Moksnes H, et al. Simple decision rules can reduce reinjury risk by 84% after ACL reconstruction: the Delaware-Oslo ACL cohort study. Br J Sports Med. 2016;50(13):804-808. CrossRef PubMed

© 2025 The Authors. Published by AboutScience - www.aboutscience.eu

23. White MS, Graham MC, Janatova T, et al. Effect of sampling rate, filtering, and torque onset detection on quadriceps rate of torque development and torque steadiness. Sensors (Basel). 2024;24(13):4250. CrossRef PubMed
24. Andersen LL, Andersen JL, Zebis MK, et al. Early and late rate of force development: differential adaptive responses to resistance training? Scand J Med Sci Sports. 2010 Feb;20(1):e162-9. CrossRef
25. D'Emanuele S, Maffiuletti NA, Tarperi C, et al. Rate of force development as an indicator of neuromuscular fatigue: a scoping review. Front Hum Neurosci. 2021;15:701916. CrossRef PubMed
26. Jaafar H, Lajili H. The influence of verbal instruction on measurement reliability and explosive neuromuscular performance of the knee extensors. J Hum Kinet. 2018;65(1):21-34. CrossRef PubMed
27. Andersen LL, Aagaard P. Influence of maximal muscle strength and intrinsic muscle contractile properties on contractile rate of force development. Eur J Appl Physiol. 2006;96(1):46-52. CrossRef PubMed
28. Andersen LL, Andersen JL, Zebis MK, et al. Early and late rate of force development: differential adaptive responses to resistance training? Scand J Med Sci Sports. 2010;20(1):e162-e169. CrossRef PubMed
29. Aagaard P, Simonsen EB, Andersen JL, et al. Increased rate of force development and neural drive of human skeletal muscle following resistance training. J Appl Physiol. 2002;93(4):1318-1326. CrossRef PubMed
30. Suzuki Y, Iijima H, Nakamura M, et al. Rate of force development in the quadriceps of individuals with severe knee osteoarthritis: a preliminary cross-sectional study. PLoS One. 2022;17(1):e0262508. CrossRef PubMed
31. Dos'Santos T, Jones PA, Kelly J, et al. Effect of sampling frequency on isometric midthigh-pull kinetics. Int J Sports Physiol Perform. 2019;14(4):525-530. CrossRef PubMed
32. Tipparaju VV, Mallires KR, Wang D, et al. Mitigation of data packet loss in Bluetooth low energy-based wearable healthcare ecosystem. Biosensors (Basel). 2021;11(10):350. CrossRef PubMed
33. Yona T, Fischer AG. Assessing the knee extensor's rate of force development using a fixed hand-held dynamometer: an inter and intrasession reliability study. J Sport Rehabil. 2024;34(4):443-448. CrossRef PubMed
34. Sale DG. Postactivation potentiation: role in human performance. Exerc Sport Sci Rev. 2002;30(3):138-143. CrossRef PubMed
35. Atkinson G. What is this thing called measurement error? In: Kinanthropometry VIII: Proceedings of the 8th International Conference of the International Society for the Advancement of Kinanthropometry (ISAK). London: Taylor & Francis; 2003. P.3–13. CrossRef
36. Mokkink LB, Eekhout I, Boers M, et al. Studies on reliability and measurement error of measurements in medicine – from design to statistics explained for medical researchers. Patient Relat Outcome Meas. 2023;14(14):193-212. CrossRef PubMed
37. Kim JK, Park MG, Shin SJ. What is the minimum clinically important difference in grip strength? Clin Orthop Relat Res. 2014;472(8):2536-2541. CrossRef PubMed
38. Landi F, Calvani R, Martone AM, et al. Normative values of muscle strength across ages in a 'real world' population: results from the longevity check-up 7+ project. J Cachexia Sarcopenia Muscle. 2020;11(6):1562-1569. CrossRef PubMed
39. Minshull C, Rees D, Gleeson NP. Joint angle affects volitional and magnetically-evoked neuromuscular performance differentially. J Electromyogr Kinesiol. 2011;21(4):672-677. CrossRef PubMed
40. Hannah R, Minshull C, Buckthorpe MW, et al. Explosive neuromuscular performance of males versus females. Exp Physiol. 2012;97(5):618-629. CrossRef PubMed
41. Minshull C, Gallacher P, Roberts S, et al. Contralateral strength training attenuates muscle performance loss following anterior cruciate ligament (ACL) reconstruction: a randomized-controlled trial. Eur J Appl Physiol. 2021;121(12):3551-3559. CrossRef PubMed
42. Wadsworth C, Nielsen DH, Corcoran DS, et al. Interrater reliability of hand-held dynamometry: effects of rater gender, body weight, and grip strength. J Orthop Sports Phys Ther. 1992;16(2):74-81. CrossRef
43. Beshay N, Lamb P, Murrell G. The reliability of fixed versus hand-held dynamometer for shoulder strength testing. J Sci Med Sport. 2010;13:e34-e35. CrossRef

© 2025 The Authors. Arch Physioter - ISSN 2057-0082 - www.archivesofphysiotherapy.com