Vol. 14 / N. 1 / January - December 2024

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

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

EDITORIAL BOARD

ABOUTSCIENCE

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

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. 14 is published in 2024

ARCHIVES OF PHYSIOTHERAPY

Contents Volume 14 Number 1 January-December 2024

Arch Physioter 2024; 14: 1-10
ISSN 2057-0082 | DOI: 10.33393/aop.2024.2916
REVIEW

Pragmatism in manual therapy trials for knee osteoarthritis: a systematic review

Kyle R. Adams[1,2], **Ayodeji O. Famuyide**[1,3], **Jodi L. Young**[1], **C. Daniel Maddox**[4,5], **Daniel I. Rhon**[1,6]

[1]Physical Therapy Program, Bellin College, Green Bay, Wisconsin - USA
[2]Department of Physical Therapy, Baylor University, Waco, Texas - USA
[3]Greater Baton Rouge Physical Therapy, Baton Rouge, Louisiana - USA
[4]Department of Physical Therapy, Ivester College of Health Sciences, Brenau University, Gainesville, Georgia - USA
[5]Upstream Rehab Institute, Smyrna, Georgia - USA
[6]Department of Rehabilitation Medicine, School of Medicine, Uniformed Services University of Health Sciences, Bethesda, Maryland - USA

ABSTRACT

Introduction: Manual therapy is an often-utilized intervention for the management of knee osteoarthritis (OA). The interpretation of results presented by these trials can be affected by how well the study designs align applicability to real-world clinical settings.
Aim: To examine the existing body of clinical trials investigating manual therapy for knee OA to determine where they fall on the efficacy-effectiveness spectrum.
Methods: This systematic review has been guided and informed by the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines. Randomized controlled trials that investigated manual therapy treatments for adults with knee OA were retrieved via searches of multiple databases to identify trials published prior to April 2023. The Rating of Included Trials on the Efficacy-Effectiveness Spectrum (RITES) tool was used to objectively rate the efficacy-effectiveness nature of each trial design. The Cochrane Risk of Bias 2.0 assessment tool (RoB-2) was used to assess the risk of bias across five domains.
Results: Of the 36 trials, a higher percentage of trials had a greater emphasis on efficacy within all four domains: participant characteristics (75.0%), trial setting (77.8%), flexibility of intervention (58.3%), and clinical relevance of experimental and comparison intervention (47.2%). In addition, 13.9% of the trials had low risk of bias, 41.7% had high risk of bias, and 44.4% had some concerns regarding bias.
Conclusions: While many trials support manual therapy as effective for the management of knee OA, a greater focus on study designs with an emphasis on effectiveness would improve the applicability and generalizability of future trials.

Keywords: Effectiveness, Efficacy, Knee osteoarthritis, Manual therapy, Mobilization, Systematic review

What's already known about this topic?

- *Despite clinical trials revealing substantial treatment effects favoring manual therapy for the management of knee osteoarthritis (OA), challenges still remain with implementation and translation of this work into clinical practice.*
- *One reason is that the majority of the research in this field is based on more explanatory or more pragmatic trial designs.*

What does this study add?

- *We conducted a systematic review of 36 trials that assessed treatment effects for manual therapy interventions to determine where the trails fall on the efficacy-effectiveness spectrum.*
- *Of the 36 trials, a majority had a greater emphasis on efficacy for all four domains: participant characteristics (75%), trial setting (77.8%), flexibility of intervention (58.3%) and clinical relevance of experimental and comparison intervention (47.2%).*

Received: June 29, 2023
Accepted: January 9, 2024
Published online: February 26, 2024

This article includes supplementary information

Corresponding author:
Kyle R. Adams
email: Kyle_Adams@Baylor.edu

Introduction

Knee osteoarthritis (OA) is the most common form of arthritis and a leading cause of disability in older adults, with symptomatic OA continuing to rise partly due to the global obesity epidemic and aging population (1-4). Knee OA has become a significant burden to society because of its chronic nature and high cost of treatment, with estimated costs in the

United States greater than $27 billion annually (5,6). Several nonpharmacological interventions have demonstrated effectiveness, the most promising being exercise therapy (7). Manual therapy has also proven effective for reducing pain and improving function in individuals with knee OA (5,8-12). Assessing the context in which these interventions are assessed is valuable to better understand their applicability and generalizability to real-world clinical practice. In addition to difficulty associated with blinding subjects, therapists, and assessors in these types of nonpharmacological trials, another challenge is that trials vary with respect to their study design, which can make it hard to determine their real-world clinical applicability (5,9,13,14).

The various components of a clinical trial design have characteristics that make them more explanatory or more pragmatic (15). Some trials are more explanatory in nature, meaning they are carried out under ideal and controlled circumstances to demonstrate if an intervention can achieve a desired result (16,17). When this occurs, a study is said to have high focus on efficacy and internal validity; however, the results may be less generalizable as the study parameters do not always reflect real-world practice (e.g., very selective inclusion criteria and no presence of comorbidities) (18). Other study designs are more pragmatic, with the goal of assessing the effectiveness of an intervention across various settings, people, and times in a way that would more closely reflect delivery in real-world settings (17). Trials with a pragmatic design tend to have higher external validity, leading to improved applicability in real-life situations (15). It is important to note that trials are rarely fully explanatory or pragmatic, but instead fall along a spectrum (15). These differences in design structure require readers to not only focus on the results of the study but also consider participant characteristics, trial setting, flexibility of interventions, and clinical relevance of experimental and comparison interventions in order to understand how applicable the results are for their clinical practice (19).

Several meta-analyses suggest manual therapy has value for the management of knee OA, at minimum in the short term (5,9,20). As an intervention that physical therapists continue to utilize and that patients perceive as beneficial (21), manual therapy may have the ability to provide a window of opportunity to enable active intervention approaches, such as exercise (21,22). To better understand their applicability and generalizability in real-world clinical practice, it is important to understand where manual therapy trials fall along the explanatory-pragmatic spectrum (5,9,23,24). The Rating of Included Trials on the Efficacy-Effectiveness Spectrum (RITES) tool was developed to enable the assessment of published trials along this spectrum (19), but has not yet been used to assess knee OA trials. The objective of this review was to determine where trials investigating manual therapy for knee OA fall on the explanatory-pragmatic spectrum in order to better understand optimal applicability, generalization, and implementation of this intervention.

Methods

The systemic review was conducted according to the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) (25,26). The review protocol was prospectively registered in the PROSPERO database (CRD42022327706). There were no patients involved in this review.

Search strategy

A literature search was performed using PubMed, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Cochrane Central Register of Controlled Trials (CENTRAL), and Embase to identify trials published prior to April 2023. In addition to these databases, the authors performed manual searches by cross-referencing trials included in related systematic reviews to capture all relevant studies in order to maximize the quality of this review. The systematic reviews that were examined consisted of any related to the eligibility criteria for this review.

Search strategies were developed using medical subject headings (MeSH) and keywords pertaining to the knee, OA, manual therapy, randomized controlled trials, and adult/young adult. Medical librarians assisted with the searches (Supplementary material, Appendix A). The primary search methods used were appropriate to each database, which included MeSH terms, CINAHL headings, subject headings, and keywords and their synonyms. Truncation and wildcards were used to account for different spellings and alternative words that may be used to describe our keywords (e.g., arthr* to identify arthritis or arthrosis). The Boolean operators “AND” and “OR” were used to combine search terms. Filters for the English language and the time frame of 1975 to April 2023 were used.

Study selection

Randomized clinical trials where the primary focus was assessing the effect of manual therapy interventions for adult patients with knee OA were included. Full text of all trials had to be available in the English language. Animal trials, trials that included subjects with diagnoses other than knee OA in any compartment, or trials where subjects had any surgery in the past 6 months or had undergone a knee arthroplasty in the involved knee at any time were excluded (Supplementary material, Appendix B).

Because the label of manual therapy can be broad and extensive (e.g., includes massage, lymphatic drainage, passive range of motion) (27), we deliberately limited the definition of manual therapy for this review as a treatment primarily consisting of joint mobilizations or manipulations performed by a healthcare provider, even if it was part of a multimodal intervention as long as the effect of the manual therapy intervention was being assessed. Trials including other forms of manual therapy (e.g., massage, soft tissue mobilization, lymphatic massage/drainage, cupping, dry needling, acupuncture, acupressure, and stretching) in the absence of joint mobilization or manipulation were excluded. Trials assessing manual therapy as part of a group of interventions where the effect of manual therapy was not assessed (e.g., a trial where everyone received manual therapy as part of standard care and the purpose of the trial was to assess the effect of

other interventions, such as exercise, education, medications, etc.) were also excluded from the review. The eligibility requirements for this review were chosen to maximize the relevance and overall quality of this review.

Data management

Covidence data management software (Veritas Health Innovation Ltd, Melbourne, Australia) was used for study screening, full-text review, and data extraction (28).

Data extraction

Two reviewers independently screened all titles and abstracts to determine eligibility for full-text review. Any disagreements were discussed for resolution, and a third reviewer was consulted for final disposition, as necessary. Upon completion of title and abstract screenings, the remaining full-text trials were screened by the same two reviewers using the predetermined eligibility criteria. Reasons for exclusion were documented within Covidence (Supplementary material, Appendix C).

Data extracted included total number of subjects, mean age in years, mean body mass index, proportion of males and females, and the year the trial was published. In addition, the RITES tool was used to rate the efficacy-effectiveness nature of each study. Descriptors of maximal efficacy and maximal efficiency are provided in Table 1 (19). The RITES tool is used to rate the efficacy-effectiveness nature of trials by assessing four different domains (participant characteristics, trial settings, flexibility of interventions, and clinical relevance of experimental and comparison interventions) using a 5-point Likert scale, with 1 indicating a strong emphasis on efficacy (more explanatory), and 5 indicating a strong emphasis on effectiveness (more pragmatic) (19). The two reviewers independently scored each study using the RITES tool and consulted with a third reviewer when there was a lack of consensus.

Risk of bias assessment

The Cochrane Collaboration Risk of Bias 2.0 assessment tool (RoB-2) was used to assess the risk of bias across five separate domains: randomization process, deviations from the intended interventions, missing outcome data, measurement of the outcome, and selection of the reported results (29,30). Each domain was rated as having low risk, high risk, or some concerns regarding the risk of bias for that trial. The two reviewers independently scored each trial using RoB-2 to determine the potential risk for bias when looking at the results. In the event of a difference in opinion, consensus was reached by consulting with a third reviewer.

Efficacy-effectiveness spectrum

The RITES tool was used to assess where the components of each trial fell along the efficacy-effectiveness continuum (14). It was developed for post hoc assessment of trials in a systematic review based on efficacy-effectiveness continuum along four domains: participant characteristics, trial settings, flexibility of intervention(s), and clinical relevance of experimental and comparison intervention(s). A Likert scale ranging from 1 (strong emphasis on efficacy) to 5 (strong emphasis on effectiveness) is used in scoring. A rating of not applicable (N/A) may be given when information for a domain is unavailable. Trials typically cannot be completely categorized as explanatory or pragmatic as a whole, but instead rated along a continuum. In addition, different components of a trial design may fall in different places along the efficacy-effectiveness continuum. Thus, each domain is scored independently, without putting forth an overall score for a trial.

Data synthesis and analysis

Interrater reliability between reviewers was calculated for title and abstract and full-text screening using Cohen's kappa. Levels of agreement were defined as <0 = no agreement, 0-0.20 = slight agreement, 0.21-0.40 = fair agreement, 0.41-0.60 = moderate agreement, 0.61-0.80 = substantial agreement, and 0.81-1.0 = almost perfect agreement (31). Descriptive statistics were calculated for RITES tool scores that included the count and percentage within each of the four domains. For each RITES domain, the results were separated into three different groups, those that had more emphasis on efficacy (scores of 1-2), those with more emphasis on effectiveness (scores of 4-5), and those that were balanced or neutral (scores of 3). In addition, the four domain scores from each trial were averaged together to determine if the individual trial design, with all domain scores considered together, leaned more toward efficacy or efficiency. For RoB-2, count data and percentage were calculated for all trials based on ratings of low risk, high risk, or some concerns regarding the risk of bias. Interrater reliability between reviewers was assessed for all four domains of the RITES tool and final RoB-2 scores. Finally, all trials were classified as being positive or null based on the primary outcome and then assessed for associations with trial design emphasis on efficacy vs. effectiveness using contingency tables and Fisher's exact test. IBM SPSS Statistics (version 28; Chicago, IL) was used for all analyses.

Deviations from prospective protocol registration

There were no deviations from the prospective protocol registration.

Results

Search results

The initial search yielded 2,656 citations, and after removing 1,584 duplicates, 1,074 titles and abstracts required screening. After title and abstract screening and full-text review, 36 trials (13,14,32-65) were included in the final review (Fig. 1). Specific details regarding each trial that was excluded can be found in Supplementary material, Appendix C. Features of the 36 trials are included in Supplementary material, Appendices D and E.

TABLE 1 - RITES maximal efficacy and maximal efficiency descriptors

Domain	Descriptor for efficacy orientation	Example	Descriptor for effectiveness orientation	Example
1. Participant characteristics	Participants are a homogeneous population and are not similar to those seen in usual care. Inclusion/exclusion criteria are often intentionally chosen to increase chances of compliance or successful outcomes, as compared with treatments provided in usual care.	Inclusion criteria consisted of individuals with bilateral primary knee OA based on the clinical criteria of the American College of Rheumatology as long as knee radiographs revealed Stage II-III OA, as per the Kellgren and Lawrence staging scale (35). Individuals were excluded if they had vascular and cardiovascular disease, obesity, acute or chronic pain in the spine, hip, or ankle, and those using anti-inflammatory drugs beyond simple pain relievers (35,40).	The participants are similar to those who would receive the experimental intervention as a part of usual care. This means that their age, severity of illness, and comorbidities are similar to those patients who would be candidates for the intervention in a usual care setting.	Inclusion criteria consisted of individuals with knee OA based on the clinical criteria of the American College of Rheumatology. In addition, individuals need to be able to walk in their daily living with or with aid, and need to be able to communicate and follow orders (47). Individuals were excluded if they had rheumatoid gout or other systemic joint disease; cerebrovascular conditions such as acute or recurrent myocardial infarction, unstable angina pectoris; neurological conditions such as Parkinson's disease; undergone knee surgery, corticosteroid injection to the knee within previous 30 days; and an injury or accident at the back and lower limbs within previous 6 weeks (47).
2. Trial setting	The setting used maximizes the ability to carry out the trial and identify an intervention effect if there is one. This typically means that very few settings are used, and they may be more specialized than the setting in which the experimental intervention would be delivered in usual care.	Interventions were provided in the outpatient physical therapy department of a single military medical center (13).	Multiple settings are used, often consisting of subsettings that are typical for usual care for the studied intervention (e.g., primary care, specialized care).	Interventions were provided at clinical sites in Pittsburgh, Pennsylvania, Salt Lake City, Utah, and San Antonio, Texas (42).
3. Flexibility of intervention(s)	There is a strict protocol that limits the flexibility of experimental and comparison interventions that are delivered. This is done to achieve improved intervention adherence. Cointerventions are often prohibited.	All participants received a standardized treatment, with one group receiving a standardized joint mobilization intervention, one group receiving a different standardized joint mobilization intervention, and one group receiving a standardized exercise intervention (48).	Experimental and comparison interventions are provided with flexibility that is identical to that in usual care. Cointerventions are often permitted, as is the case in usual care.	All participants were randomized into one of four groups. The treatments they received consisted of various combinations of manual therapy, exercise, and/or booster sessions, which allowed for some flexibility of interventions, as per the patient presentation (32).
4. Clinical relevance of experimental and comparison intervention(s)	The experimental and/or comparison interventions are not a clinically relevant or best current treatment. Examples would be the use of a placebo, a no treatment control, the use of subclinical doses, or trial durations that are less than that seen in usual care.	All participants were assigned to a treatment. One group received manual therapy interventions consisting of joint mobilization, muscle stretching, and soft tissue mobilization. The other group received a placebo intervention consisting of the therapist placing both hands on the knee without any pressure for 10 minutes (40).	Both the experimental and comparison interventions are likely to be considered best practice, and the duration of interventions is similar to that in usual care.	All participants were assigned to a treatment. One group received manual therapy for the knee, lumbar spine, hip, and ankle, as required, and performed a standardized exercise routine in the clinic and at home. The other group received the standardized exercise routine in the clinic and at home. All treatments could be considered best practice (13).

Based on a 5-point Likert scale the evidence deriving from each domain should be rated:
1 = strong emphasis on efficacy
2 = rather strong emphasis on efficacy
3 = balanced emphasis on both efficacy and effectiveness
4 = rather strong emphasis on effectiveness
5 = strong emphasis on effectiveness
OA = osteoarthritis; RITES = Rating of Included Trials on the Efficacy-Effectiveness Spectrum.
Table adapted from Wieland et al (19).

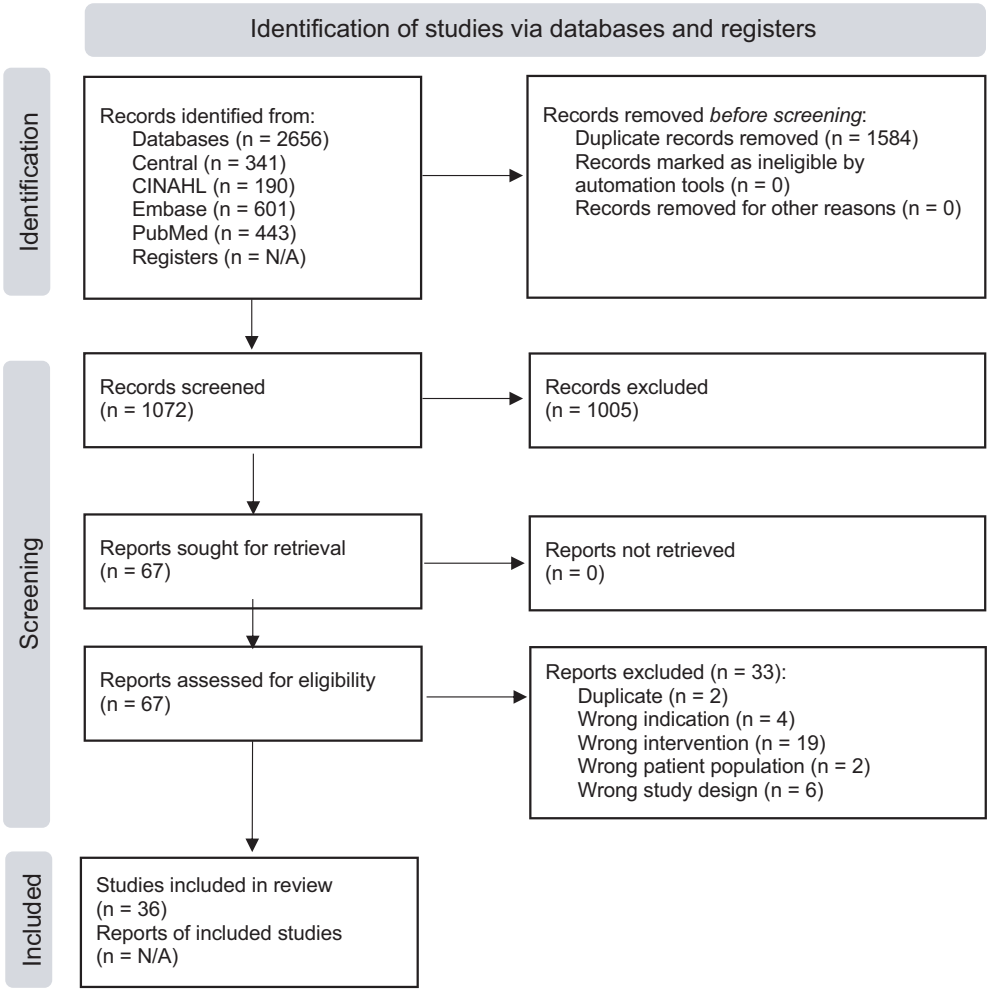


FIGURE 1 - PRISMA 2020 flow diagram for new systematic reviews that included searches of databases and registers only. PRISMA = Preferred Reporting Items for Systematic Review and Meta-Analysis.

RITES domain scores

Overall RITES scores by domain are provided in Figure 2. A higher percentage of trials had a greater emphasis on efficacy within all four domains: participant characteristics (75.0%; n = 27) (13,34-42,44,45,48-50,52-55,57-62,64,65), trial setting (77.8%; n = 28) (13,14,32-36,40,43-49,51-55,57-60,62-65), flexibility of intervention (58.3%; n = 21) (33-36,38,40,43,44, 46-48,50-52,54,55,58,60,62,63,65), and clinical relevance of experimental and comparison intervention (47.2%; n = 17) (14,34-36,38,40,43,47,48,51,52,54,55,57,61,62,64). In addition, when the RITES scores for all four domains of each trial were averaged, 29 trials were more oriented toward efficacy (mean [SD] of 2.2 [0.4] and range 1 to 3) (13,14,34-36,38,40, 43-55,57-65), whereas five trials were more oriented toward effectiveness (mean [SD] of 3.9 [0.5] and range 3 to 5) (32,37,41,42,56). The remaining two trials had a mean score of 3.0 across the four domains, indicating a balanced emphasis between efficacy and effectiveness (33,39). Despite this overall emphasis on efficacy, 20 of the 36 trials had at least one domain with a score greater on the effectiveness spectrum (13,14,32,33,37-39,41,42,44,46,47,49,51,56,57,60,61,63,65) (Tab. 2).

For the participant characteristics domain, 27 trials (75.0%) had scores that emphasized efficacy (13,34-42,44,

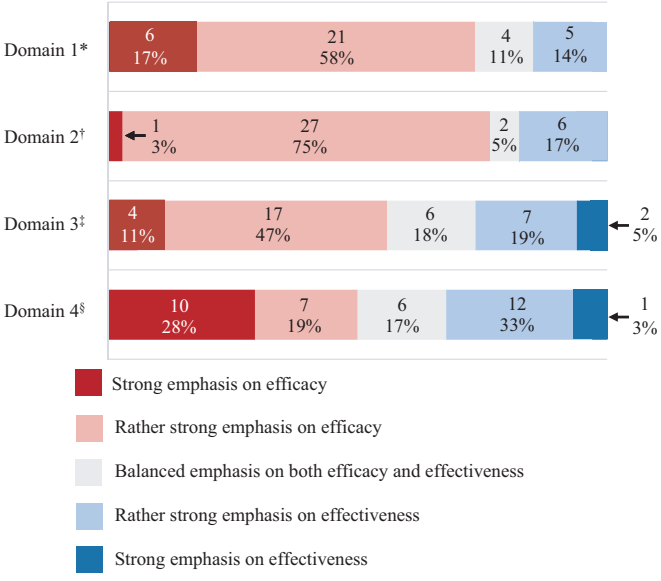


FIGURE 2 - Cumulative RITES scores (percentage and count). RITES = Rating of Included Trials on the Efficacy-Effectiveness Spectrum. *Participant characteristics; †Clinical trial setting; ‡Flexibility of intervention(s); §Clinical relevance of experimental and comparison intervention(s).

TABLE 2 - Individual trial RITES scores

Primary author and year of publication	RITES scores*			
	D1†	D2‡	D3§	D4‖
Abbott et al (32)	4	2	5	4
Ali et al (33)	4	2	2	4
Alkhawajah and Alshami (34)	2	2	2	1
Altinbilek et al (35)	1	2	2	2
Bhagat et al (36)	2	2	2	1
Bove et al (37)	5	5	4	5
Courtney et al (38)	2	4	1	1
Crossley et al (39)	1	4	4	3
Cruz-Montecinos et al (40)	1	2	2	1
Deyle et al (14)	3	2	4	2
Deyle et al (13)	2	1	4	4
Dwyer et al (41)	2	4	4	4
Fitzgerald et al (42)	2	4	4	4
Forestier et al (43)	3	2	2	1
Jeyakumar et al (44)	2	2	2	4
Jin et al (45)	2	2	3	3
Kaya Mutlu et al (46)	3	2	2	4
Kornkamon and Wanitcha (47)	4	2	2	1
Lalit et al (48)	2	2	1	2
Lizis et al (49)	2	2	3	4
Mahmooda et al (50)	1	3	2	3
Moss et al (51)	4	2	2	1
Narang and Ganvir (52)	2	2	2	2
Nigam et al (53)	1	2	3	3
Pollard et al (54)	2	2	2	1
Pozsgai et al (55)	2	2	2	1
Pryymachenko et al (56)	3	3	5	4
Rao et al (57)	2	2	4	2
Razek and Shenouda (58)	2	2	2	3
Reza et al (59)	2	2	3	3
Sharma (60)	2	2	2	4
Sit et al (61)	2	4	3	1
Syed and Wani (62)	2	2	2	2
Taj et al (63)	4	2	1	4
Tucker et al (64)	2	2	3	2
Witwit et al (65)	1	2	1	5

RITES, Rating of Included Trials on the Efficacy-Effectiveness Spectrum.
*RITES scoring, based on a 5-point Likert scale: 1 = strong emphasis on efficacy; 2 = rather strong emphasis on efficacy; 3 = balanced emphasis on both efficacy and effectiveness; 4 = rather strong emphasis on effectiveness; 5 = strong emphasis on effectiveness; N/A = information not available.
† RITES Domain 1: participant characteristics.
‡ RITES Domain 2: trial setting.
§ RITES Domain 3: flexibility of intervention(s).
‖ RITES Domain 4: clinical relevance of experimental and comparison intervention(s).

45,48-50,52-55,57-62,64), five trials (13.9%) emphasized effectiveness (32,33,47,51,63), and four trials (11.1%) had a balanced emphasis between efficacy and effectiveness (14,43,46,56). In the trial setting domain, 28 trials (77.8%) had scores that emphasized efficacy (13,14,32-36,40,43-49,51-55,57-60,62-65), six trials (16.7%) emphasized effectiveness (37-39,41,42,61), and two trials (5.5%) had a balanced emphasis between efficacy and effectiveness (50,56). The flexibility of intervention(s) domain had 21 trials (58.3%) that emphasized efficacy (33-36,38,40,43,44,46-48,50-52,54,55,58,60,62,63,65), nine trials (25.0%) emphasizing effectiveness (13,14,32,37,39,41,42,56,57), and six trials (16.7%) that exhibited a balanced emphasis between efficacy and effectiveness (45,49,53,59,61,64). Finally, the clinical relevance of experimental and comparison intervention(s) domain had 17 trials (47.2%) that emphasized efficacy (14,34-36,38,40,43,47,48,51,52,54,55,57,61,62,64), 13 trials (36.1%) emphasized effectiveness (13,32,33,37,41,42,44,46,49,56,60,63,65), and six trials (16.7%) had a balanced emphasis between efficacy and effectiveness (39,45,50,53,58,59).

Of the 36 trials, only seven had null findings (41,42,45,48,63-65). Fisher's exact test revealed no statistically significant relationship between where studies fell on the efficacy-effectiveness spectrum and a positive outcome of the primary outcome ($p = 0.27$).

Risk of bias for included trials

Five of the included trials (13.9%) had low risk of bias (13,36,54,59,65), 15 trials (41.7%) had high risk of bias (44-52,55,57,58,62-64), and 16 trials (44.4%) had some concerns for risk of bias (14,32-35,37-43,53,56,60,61) (Supplementary material, Appendix F). The most common cause for bias included measurement of the outcome (44,45,47,49,51,52,55,57,58,64), and the least amount of bias was in the selection of reported outcome (13,14,32-65). When comparing risk of bias across the trials, all five of those with low risk of bias also had an emphasis on efficacy (13,36,54,59,65).

Rater agreement

Interrater reliability was $\kappa = 0.25$ (fair agreement) for title and abstract screening and $\kappa = 0.31$ (fair agreement) for full-text screening. Interrater reliability between reviewers for the participants' characteristics domain was $\kappa = 0.45$ (fair agreement), $\kappa = 0.39$ (fair agreement) for trial settings, $\kappa = 0.34$ (fair agreement) for flexibility of interventions, and $\kappa = 0.32$ (fair agreement) for clinical relevance of experimental and comparison interventions. Interrater reliability between the reviewers for RoB-2 was $\kappa = 0.04$ (slight agreement). These values were related to initial agreement when reviewing the trials. It is important to note that consensus was reached on all initial ratings, and a third reviewer needed to be consulted for only three trials (8.3%).

Discussion

This systematic review assessed existing manual therapy trials for knee OA to determine where the current body of

evidence falls on the efficacy-effectiveness spectrum. The findings suggest that the majority of trials trend toward efficacy in all four domains of the RITES tool, especially for participant characteristics and clinical trial settings. While a previous systematic review has looked at a similar question in trials involving manual therapy for low back pain (66), this is the first known review assessing trials for knee OA.

Participant characteristics

A large percentage of trials (75.0%) were higher on the explanatory end of the spectrum in the participant characteristics domain, with the primary reason being related to their exclusion criteria (13,34-42,44,45,48-50,52-55,57-65). Patients were most commonly excluded from trials due to the presence of other diagnoses or comorbidities, and while this could confound treatment effect, it is a more accurate representation of patients seeking care for knee OA. For example, vascular and cardiovascular disease, obesity, acute or chronic pain in the spine, hip, or ankle, and those using anti-inflammatory drugs beyond simple pain relievers are common presentations for individuals with knee OA (35,40). Excluding these individuals could result in conclusions that may not be relevant to the types of patients seen in most clinics (16,18). To achieve a more pragmatic rating would have required a study population that included patients with diagnoses, comorbidities, symptom durations, and age ranges similar to common knee OA patients that seek care (16,18).

Trial setting

The majority of trials (77.8%) had an emphasis on efficacy in the trial setting domain due to the trials being carried out in settings that were dissimilar from common practice (13,14,32-36,40,43-49,51-55,57-60,62-65). These included specialized clinics, specialized trial or academic centers, and military clinics and settings, and also used a limited number of clinicians who were often specifically trained for the interventions being assessed. While this may enable researchers to better determine the effect of the interventions without compromising internal validity, it limits external validity (15,18). To achieve more pragmatic trial settings, researchers should strive to use a broad array of clinics and clinicians that better mimic typical medical providers and healthcare settings (15,16).

Flexibility of interventions

The flexibility of interventions domain had an emphasis on efficacy. The majority of clinical trials (58.3%) required strict manual therapy protocols with little flexibility or prohibited cointerventions (33-36,38,40,43,44,46-48,50-52,54,55,58,60,62,63,65). Some reasons for a strict protocol include the ability to better attribute the treatment effect to the intervention being assessed, rather than an influence from other confounders. Even efforts to control or improve intervention adherence may lead to different results than can be expected in real-world settings (15,16). On the other hand, nine of the trials had a more pragmatic emphasis (13,14,32,37,39,41,42,56,57), which was accomplished by allowing more flexibility with the interventions between the trial populations. This approach allowed clinicians to manage patients based on their perceived needs with greater flexibility.

Clinical relevance of experimental and comparison intervention

Clinical relevance of experimental and comparison of interventions slightly favored efficacy (47.2%) (14,34-36,38,40,43,47,48,51,52,54,55,57,61,62,64) compared to those with designs more focused on effectiveness (36.1%) (13,32,33,37,41,42,44,46,49,56,60,63,65), and those that had a balanced emphasis on efficacy and effectiveness (16.7%) (39,45,50,53,58,59). Trials with a more explanatory design were less likely to have one of the treatment arms considered clinically relevant or best practice, such as using controls, placebo, or sham interventions, all of which provide a less than desirable comparison when considering generalizability to real-world settings (19). Treatment duration may also have been much shorter than the duration of treatments used in real-world practice (19). On the other hand, trials emphasizing a pragmatic approach use flexibility with interventions that mimic typical practice, and they use comparison groups that are often considered to represent best practice or usual care (19).

Outcomes relative to efficacy-effectiveness spectrum

When examining trial outcomes relative to where studies fell on the efficacy-effectiveness spectrum, there were no significant associations. This means that trial design has no bearing on whether a study showed a treatment effect. However, definitive conclusions cannot be made because there were only seven null trials out of the 36 trials (41,42,45,48,63-65).

Clinical implications

The majority of trials investigating manual therapy for knee OA were on the explanatory end of the spectrum across all four RITES domains. This is similar to what was reported by Maddox et al (66) for individuals with low back pain, except their review found a rather strong emphasis toward the pragmatic end of the spectrum with the domain related to clinical relevance of experimental and comparison interventions.

The role of explanatory trials is to analyze the mechanism of interventions under controlled circumstances (19). In this instance, explanatory trials help to determine if manual therapy is an effective treatment for knee OA. However, they lack the ability to generalize the results because the settings are often not representative of real-world clinical practice. That is where pragmatic trials provide their value by providing clinicians with the ability to know if manual therapy can be beneficial for patients with knee OA in real-world settings (15-17). The result of this review demonstrates lack of generalizability with the majority of studies examining manual therapy for knee OA.

Recommendation for future research

While the current body of literature demonstrates potential benefits when using manual therapy for individuals with knee OA, many of those recommendations come from trials

that are more explanatory than pragmatic, making them less generalizable (5,9). Additional pragmatic studies examining manual therapy for knee OA in real-world scenarios and across a variety of settings and clinicians would help improve the applicability and implementation of these interventions. For example, the study design could include patient population with some comorbidities, especially those commonly associated with knee OA (diabetes and obesity); multiple/diverse trial settings or general clinical practice settings, not specialty treatment clinics; and flexibility of interventions, allowing cross-treatments whenever/if needed while ensuring that the methodology of interest is systematically and objectively directed toward best practice. Finally, it is worth noting that manual therapy may not be unique here, and these findings may be very similar to what is observed for trial designs of other interventions for knee OA.

Limitations

This review had the primary goal of assessing where this body of evidence falls on the efficacy-effectiveness spectrum, with no intention to examine the effectiveness of manual therapy for knee OA. Therefore, conclusions should not be inferred regarding pooled treatment effects or the value of manual therapy interventions for knee OA.

Conclusions

Thirty-six manual therapy trial designs for knee OA were assessed for their fit along the explanatory-pragmatic spectrum. The majority of trial designs were more explanatory, making the results less generalizable across patient populations, clinical settings, and compared to other commonly used interventions. When examining the effectiveness of manual therapy for the treatment of knee OA, more pragmatic study designs would help improve implementation and applicability of research results. This can be achieved by using a more diverse patient population, a larger number of clinics, intervention protocols that are more pragmatic, and comparison treatments that represent best practice or usual care. All of these will help improve the ability to generalize findings from manual therapy trials for knee OA.

Abbreviations

κ, kappa; CINAHL, Cumulative Index to Nursing and Allied Health Literature; IL, Illinois; n, number; OA, osteoarthritis; PRISMA, Preferred Reporting Items for Systematic Review and Meta-Analysis; RITES, Rating of Included Trials on the Efficacy-Effectiveness Spectrum; RoB-2, Cochrane Risk of Bias 2.0.

Acknowledgments

We would like to thank Cindy Reinl (Bellin Health Systems librarian) and Kenneth L. Carriveau, Jr. and Rachel Blume (Baylor University librarians) for their assistance with performing thorough literature searches for this systematic review. We would also like to thank Mareli Klopper for her mentoring during the completion of this systematic review.

Disclosures

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

Financial support: Authors have not received any financial support for this study.

Authors' contributions: KRA contributed to the design of the systematic review, data collection and data analysis, and all drafts and revisions of the manuscript. AOF contributed to the design of the systematic review, data collection and data analysis, and all drafts and revisions of the manuscript. JLY contributed to the design of the systematic review, data collection and data analysis, to all drafts and revisions of the manuscript, and had final approval of the version to be published. CDM contributed to the design of the systematic review, contributed to data collection and data analysis, and contributed to drafts and revisions of the manuscript. DIR contributed to the design of the systematic review, data collection and data analysis, to all drafts and revisions of the manuscript, and had final approval of the version to be published.

Data availability statement: All data generated or analyzed during this study are included in this published article and its supplementary information section.

Disclaimer: The view(s) expressed herein are those of the author(s) and do not reflect the official policy or position of the Uniformed Services University, U.S. Department of Defense, or the U.S. Government.

References

1. Paradowski PT, Lohmander LS, Englund M. Osteoarthritis of the knee after meniscal resection: long term radiographic evaluation of disease progression. Osteoarthritis Cartilage. 2016;24(5):794-800. CrossRef PubMed
2. McAlindon TE, Bannuru RR, Sullivan MC, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis Cartilage. 2014;22(3):363-388. CrossRef PubMed
3. Mora JC, Przkora R, Cruz-Almeida Y. Knee osteoarthritis: pathophysiology and current treatment modalities. J Pain Res. 2018;11:2189-2196. CrossRef PubMed
4. Zhang Y, Jordan JM. Epidemiology of osteoarthritis. Clin Geriatr Med. 2010;26(3):355-369. CrossRef PubMed
5. Anwer S, Alghadir A, Zafar H, Brismée JM. Effects of orthopaedic manual therapy in knee osteoarthritis: a systematic review and meta-analysis. Physiotherapy. 2018;104(3):264-276. CrossRef PubMed
6. Arthritis Foundation [Internet]. Arthritis by the numbers; 2020. Available from: Online (Accessed June 2023)
7. Mo L, Jiang B, Mei T, Zhou D. Exercise therapy for knee osteoarthritis: a systematic review and network meta-analysis. Orthop J Sports Med. 2023;11(5):23259671231172773. CrossRef PubMed
8. Deyle GD, Allen CS, Allison SC, et al. Physical therapy versus glucocorticoid injection for osteoarthritis of the knee. N Engl J Med. 2020;382(15):1420-1429. CrossRef PubMed
9. Tsokanos A, Livieratou E, Billis E, et al. The efficacy of manual therapy in patients with knee osteoarthritis: a systematic review. Medicina (Kaunas). 2021;57(7):696. CrossRef PubMed
10. Runge N, Aina A, May S. The benefits of adding manual therapy to exercise therapy for improving pain and function in patients with knee or hip osteoarthritis: a systematic review with meta-analysis. J Orthop Sports Phys Ther. 2022;52(10):675-A13. CrossRef PubMed
11. Feng T, Wang X, Jin Z, et al. Effectiveness and safety of manual therapy for knee osteoarthritis: an overview of

systematic reviews and meta-analyses. Front Public Health. 2023;11:1081238. CrossRef PubMed
12. Rhon DI, Kim M, Asche CV, Allison SC, Allen CS, Deyle GD. Cost-effectiveness of physical therapy vs intra-articular glucocorticoid injection for knee osteoarthritis: a secondary analysis from a randomized clinical trial. JAMA Netw Open. 2022;5(1):e2142709. CrossRef PubMed
13. Deyle GD, Henderson NE, Matekel RL, Ryder MG, Garber MB, Allison SC. Effectiveness of manual physical therapy and exercise in osteoarthritis of the knee. A randomized, controlled trial. Ann Intern Med. 2000;132(3):173-181. CrossRef PubMed
14. Deyle GD, Allison SC, Matekel RL, et al. Physical therapy treatment effectiveness for osteoarthritis of the knee: a randomized comparison of supervised clinical exercise and manual therapy procedures versus a home exercise program. Phys Ther. 2005;85(12):1301-1317. CrossRef PubMed
15. Patsopoulos NA. A pragmatic view on pragmatic trials. Dialogues Clin Neurosci. 2011;13(2):217-224. CrossRef PubMed
16. Merali Z, Wilson JR. Explanatory versus pragmatic trials: an essential concept in study design and interpretation. Clin Spine Surg. 2017;30(9):404-406. CrossRef PubMed
17. Khorsan R, Crawford C. How to assess the external validity and model validity of therapeutic trials: a conceptual approach to systematic review methodology. Evid Based Complement Alternat Med. 2014;2014:694804. CrossRef PubMed
18. Patino CM, Ferreira JC. Internal and external validity: can you apply research study results to your patients? J Bras Pneumol. 2018;44(3):183. CrossRef PubMed
19. Wieland LS, Berman BM, Altman DG, et al. Rating of Included Trials on the Efficacy-Effectiveness Spectrum: development of a new tool for systematic reviews. J Clin Epidemiol. 2017;84:95-104. CrossRef PubMed
20. Rhon DI, Flynn TW, Shepherd MH, Abbott JH. Leveraging the short-term benefits of manual therapy which includes exercise over exercise therapy alone appears justified for knee osteoarthritis. J Orthop Sports Phys Ther. 2023;53(1):49-50. CrossRef PubMed
21. Ferreira RM, Martins PN, Pimenta N, Gonçalves RS. Physical therapists' choices, views and agreements regarding non-pharmacological and non-surgical interventions for knee osteoarthritis patients: a mixed-methods study. Mediterr J Rheumatol. 2023;34(2):188-219. CrossRef PubMed
22. Rhon DI, Deyle GD. Manual therapy: always a passive treatment? J Orthop Sports Phys Ther. 2021;51(10):474-477. CrossRef PubMed
23. Collins CK, Masaracchio M, Brismée JM. The future of orthopedic manual therapy: what are we missing? J Man Manip Ther. 2017;25(4):169-171. CrossRef PubMed
24. French HP, Brennan A, White B, Cusack T. Manual therapy for osteoarthritis of the hip or knee—a systematic review. Man Ther. 2011;16(2):109-117. CrossRef PubMed
25. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg. 2010;8(5):336-341. CrossRef PubMed
26. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ. 2021;372:n160. Online CrossRef PubMed. Accessed October 15, 2022.
27. Kolasinski SL, Neogi T, Hochberg MC, et al. 2019 American College of Rheumatology/Arthritis Foundation Guideline for the management of osteoarthritis of the hand, hip, and knee. Arthritis Care Res (Hoboken). 2020;72(2):149-162. CrossRef PubMed
28. Covidence 2020. Better systematic review management. Online. Accessed June 2023.
29. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898. Online CrossRef PubMed. Accessed March 28, 2022.
30. Higgins JPT, Altman DG, Gøtzsche PC, et al; Cochrane Bias Methods Group; Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343(Oct18 2):d5928. Online CrossRef PubMed. Accessed March 28, 2022.
31. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-174. CrossRef PubMed
32. Abbott JH, Chapple CM, Fitzgerald GK, et al. The incremental effects of manual therapy or booster sessions in addition to exercise therapy for knee osteoarthritis: a randomized clinical trial. J Orthop Sports Phys Ther. 2015;45(12):975-983. CrossRef PubMed
33. Ali SS, Ahmed SI, Khan M, Soomro RR. Comparing the effects of manual therapy versus electrophysical agents in the management of knee osteoarthritis. Pak J Pharm Sci. 2014;27(4)(suppl):1103-1106. PubMed
34. Alkhawajah HA, Alshami AM. The effect of mobilization with movement on pain and function in patients with knee osteoarthritis: a randomized double-blind controlled trial. BMC Musculoskelet Disord. 2019;20(1):452. Online CrossRef PubMed
35. Altınbilek T, Murat S, Yumuşakhuylu Y, İçağasıoğlu A. Osteopathic manipulative treatment improves function and relieves pain in knee osteoarthritis: a single-blind, randomized-controlled trial. Turk J Phys Med Rehabil. 2018;64(2):114-120. CrossRef PubMed
36. Bhagat M, Neelapala YVR, Gangavelli R. Immediate effects of Mulligan's techniques on pain and functional mobility in individuals with knee osteoarthritis: a randomized control trial. Physiother Res Int. 2020;25(1):e1812. CrossRef PubMed
37. Bove AM, Smith KJ, Bise CG, et al. Exercise, manual therapy, and booster sessions in knee osteoarthritis: cost-effectiveness analysis from a multicenter randomized controlled trial. Phys Ther. 2018;98(1):16-27. CrossRef PubMed
38. Courtney CA, Steffen AD, Fernández-de-Las-Peñas C, Kim J, Chmell SJ. Joint mobilization enhances mechanisms of conditioned pain modulation in individuals with osteoarthritis of the knee. J Orthop Sports Phys Ther. 2016;46(3):168-176. CrossRef PubMed
39. Crossley KM, Vicenzino B, Lentzos J, et al. Exercise, education, manual-therapy and taping compared to education for patellofemoral osteoarthritis: a blinded, randomised clinical trial. Osteoarthritis Cartilage. 2015;23(9):1457-1464. CrossRef PubMed
40. Cruz-Montecinos C, Flores-Cartes R, Montt-Rodriguez A, Pozo E, Besoaín-Saldaña A, Horment-Lara G. Changes in co-contraction during stair descent after manual therapy protocol in knee osteoarthritis: a pilot, single-blind, randomized study. J Bodyw Mov Ther. 2016;20(4):740-747. CrossRef PubMed
41. Dwyer L, Parkin-Smith GF, Brantingham JW, et al. Manual and manipulative therapy in addition to rehabilitation for osteoarthritis of the knee: assessor-blind randomized pilot trial. J Manipulative Physiol Ther. 2015;38(1):1-21.e2. CrossRef PubMed
42. Fitzgerald GK, Fritz JM, Childs JD, et al. Exercise, manual therapy, and use of booster sessions in physical therapy for knee osteoarthritis: a multi-center, factorial randomized clinical trial. Osteoarthritis Cartilage. 2016;24(8):1340-1349. CrossRef PubMed

43. Forestier R, Genty C, Waller B, et al. Crenobalneotherapy (spa therapy) in patients with knee and generalized osteoarthritis: a post-hoc subgroup analysis of a large multicentre randomized trial. Ann Phys Rehabil Med. 2014;57(4):213-227. CrossRef PubMed
44. Jeyakumar S, Alagesan J, Ramachandran A. A comparative study on the efficacy of Maitland's mobilisation and Mulligan's mobilisation in sub-acute osteoarthritis knee. Biomedicine (Taipei). 2017;37(4):518-520.
45. Jin L, Ma B, Liu X, Teng W. A randomized clinical trial assessment of nonsteroidal anti-inflammatory drugs and Chinese bone setting manipulation therapy in knee osteoarthritis. Int J Clin Exp Med. 2017;10(3):5106-5115.
46. Kaya Mutlu E, Ercin E, Razak Ozdıncler A, Ones N. A comparison of two manual physical therapy approaches and electrotherapy modalities for patients with knee osteoarthritis: a randomized three arm clinical trial. Physiother Theory Pract. 2018;34(8):600-612. CrossRef PubMed
47. Kornkamon C, Wanitcha K. Immediate effects of self-manual therapy and supervised manual therapy in individuals with knee osteoarthritis. Indian J Public Health Res Dev. 2019;10(11): 2992-2998. CrossRef
48. Lalit SY, Suhas MB, Amita M. Effect of manual therapy techniques on knee proprioception in patients with osteo-arthritis of knee. Indian J Physiother Occup Ther. 2012;6(3):285. Online
49. Lizis P, Manko G, Kobza W, Para B. Manual therapy with cryotherapy versus kinesiotherapy with cryotherapy for knee osteoarthritis: a randomized controlled trial. Altern Ther Health Med. 2019;25(4):40-45. PubMed
50. Mahmooda S, Ishaq I, Safdar M, Sabir M, Tahir A, Irshad S. Effects of mulligan's mobilization with movements versus myofascial release in addition to usual care on pain and range in knee osteoarthritis. Rawal Med J. 2020;45(2):353-357.
51. Moss P, Sluka K, Wright A. The initial effects of knee joint mobilization on osteoarthritic hyperalgesia. Man Ther. 2007; 12(2):109-118. CrossRef PubMed
52. Narang S, Ganvir S. Efficacy of Kaltenbohn mobilization on patients with osteoarthritis of knee joint. Indian J Physiother Occup Ther. 2014;8(3):162. Online
53. Nigam A, Satpute KH, Hall TM. Long term efficacy of mobilisation with movement on pain and functional status in patients with knee osteoarthritis: a randomised clinical trial. Clin Rehabil. 2021;35(1):80-89. CrossRef PubMed
54. Pollard H, Ward G, Hoskins W, Hardy K. The effect of a manual therapy knee protocol on osteoarthritic knee pain: a randomised controlled trial. J Can Chiropr Assoc. 2008;52(4):229-242. PubMed
55. Pozsgai M, Péter IA, Farkas N, Than P, Nusser N. End-range Maitland mobilization decreasing pain sensitivity in knee osteoarthritis: randomised, controlled clinical trial. Eur J Phys Rehabil Med. 2022;58(3):442-541. CrossRef
56. Pryymachenko Y, Wilson R, Sharma S, Pathak A, Abbott JH. Are manual therapy or booster sessions worthwhile in addition to exercise therapy for knee osteoarthritis: economic evaluation and 2-year follow-up of a randomized controlled trial. Musculoskelet Sci Pract. 2021;56:102439. Online CrossRef PubMed
57. Rao RV, Balthillaya G, Prabhu A, Kamath A. Immediate effects of Maitland mobilization versus Mulligan mobilization with movement in osteoarthritis knee: a randomized crossover trial. J Bodyw Mov Ther. 2018;22(3):572-579. CrossRef PubMed
58. Razek RA, Shenouda MM. Efficacy of Mulligan's mobilization with movement on pain, disability, and range of motion in patients with knee osteoarthritis: a randomized controlled pilot study. Indian J Physiother Occup Ther. 2014;7(1):242-247. Online
59. Reza MK, Shaphe MA, Qasheesh M, Shah MN, Alghadir AH, Iqbal A. Efficacy of specified manual therapies in combination with a supervised exercise protocol for managing pain intensity and functional disability in patients with knee osteoarthritis. J Pain Res. 2021;14:127-138. CrossRef PubMed
60. Sharma SS. A randomized comparison of effectiveness of clinical exercises and manual therapy procedures versus clinical exercises alone in the treatment of osteoarthritis of knee. Indian J Physiother Occup Ther. 2013;7(3):198. Online
61. Sit RWS, Chan KKW, Zou D, et al. Clinic-based patellar mobilization therapy for knee osteoarthritis: a randomized clinical trial [Internet]. Vol. 16. Ann Fam Med. 2018;16(6):521-529. CrossRef PubMed
62. Syed S, Wani S. Effect of two different manual therapy protocols on osteoarthritic knee pain & functional disability: a comparison study. Rom J Phys Ther/Revista Romana de Kinetoterapie. 2014;20(34). Online. Accessed June 2023.
63. Taj S, Anwar K, Arshad H, Khalid M, Ali MQ, Hussain E. Effectiveness of Maitland mobilization versus pain relief phenomena for pain, range of motion and disability in early knee osteoarthritis. Pak J Med Health Sci. 2023;17(1). CrossRef
64. Tucker M, Brantingham JW, Myburg C. Relative effectiveness of a non-steroidal anti-inflammatory medication (Meloxicam) versus manipulation in the treatment of osteo-arthritis of the knee. Eur J Chiropractic. 2003;50:163-183.
65. Witwit RT, Shadmehr A, Mir SM, Fereydounnia S, Jalaei S. Comparison of non-thrust manipulation vs muscle energy techniques in management of patients with knee osteoarthritis: a randomized clinical trial. Neuroquantology. 2022; 20(67):6843-6859.
66. Maddox CD, Subialka JA, Young JL, Rhon DI. Over half of clinical trials of mobilization and manipulation for patients with low back pain may have limited real-world applicability: a systematic review of 132 clinical trials. J Orthop Sports Phys Ther. 2022;52(8):532-545. CrossRef PubMed

Arch Physioter 2024; 14: 11-19
ISSN 2057-0082 | DOI: 10.33393/aop.2024.2702
ORIGINAL RESEARCH ARTICLE

The psychometric properties of the modified fear of falling avoidance behavior questionnaire in Parkinson's disease and older adults

Merrill R. Landers[1], Ash M. Haller[2], Arturo Aldaco,[3] Billy La[4], Adetayo A. Babarinde[5], John V. Rider[6], Jason K. Longhurst[7]

[1]Department of Physical Therapy, School of Integrated Health Sciences, University of Nevada, Las Vegas, Nevada - USA
[2]Sequoia Hospital, Dignity Health, Redwood City, California - USA
[3]Sunrise Hospital and Medical Center, Las Vegas, Nevada - USA
[4]FYZICAL Balance and Therapy Centers – Buffalo, Las Vegas, Nevada - USA
[5]School of Integrated Health Sciences, University of Nevada, Las Vegas, Nevada - USA
[6]School of Occupational Therapy, Touro University Nevada, Henderson, Nevada - USA
[7]Department of Physical Therapy and Athletic Training, Doisy College of Health Sciences, Saint Louis University, St. Louis, Missouri - USA

ABSTRACT

Introduction: The Fear of Falling Avoidance Behavior Questionnaire (FFABQ) has good psychometric properties. However, we have recently modified the FFABQ (mFFABQ) to improve the clarity of the questions and Likert responses. This study aimed to examine the reliability and validity of this modified version in older adults and people with Parkinson's disease (PD).
Methods: A total of 88 participants, 39 with PD (age = 72.2 ± 9.5; 29 males, 10 females) and 49 older adults (age = 72.8 ± 5.0; 13 males, 36 females), answered the mFFABQ twice, separated by 1 week, for test-retest reliability. Construct validity was evaluated through correlational analyses with fall history, Activities-Specific Balance Confidence Scale (ABC), Berg Balance Scale (BBS), Timed Up and Go, 30-Second Sit to Stand, Sensory Organization Test, Zung Anxiety Scale, Beck Depression Inventory, Consequences of Falling Questionnaire (CoFQ), and average daily activity levels using an activity monitor.
Results: The mFFABQ had good overall test-retest reliability (intraclass correlational coefficient [ICC] = 0.822; older adult ICC = 0.781, PD ICC = 0.806). The mFFABQ correlated with fall history (r = −0.430) and exhibited high correlation with the ABC (rho = −0.804) and moderate correlations with CoFQ (rho = 0.582) and BBS (rho = −0.595). The mFFABQ also correlated with time stepping (rho = −0.298) and number of steps (rho = −0.358).
Conclusion: These results provide supportive evidence for the reliability and validity of the mFFABQ in older adults and people with PD, which supports its suitability as a clinical and research tool for the assessment of fear of falling avoidance behavior.

Keywords: Balance, Balance confidence, Falls, Gait, Postural instability, Reliability

What is known about the topic?

- *Fear of falling avoidance behavior is common in older adults and people with Parkinson's disease and, because of its negative downstream consequences, it is important that therapists have a way to reliably assess its impact.*

What does the study add:

- *This study adds evidence for the reliability and validity of the mFFABQ. Because it is clinically feasible and has sound psychometric properties, it is suitable for both the clinic and research lab.*

Received: October 27, 2023
Accepted: April 4, 2024
Published online: May 2, 2024

Corresponding author:
Merrill R. Landers
email: merrill.landers@unlv.edu

Introduction

Falls are common in older adults and people with Parkinson's disease (PD), with prevalence estimates of 26.5% (1) and 35%-90% (2), respectively. After a fall or near fall, fear of falling (FOF) can develop, which can lead to FOF avoidance behavior (FFAB). Interestingly, FFAB can also occur in

those who have not had a recent fall (3). FFAB can be protective (adaptive) in that it may limit the occurrence of falls in the short term (4). That is, people with FOF may avoid risky tasks that threaten their balance, thereby protecting them from a future fall. However, excessive FFAB (maladaptive) may lead to a disproportionate amount of avoidance behavior, which reduces physical activity and increases sedentary behaviors (4). As a result of decreased activity, other downstream consequences may emerge, including physical deconditioning, weakened balance systems, poor bone health, social isolation, loneliness, and depression (4-6). Physical decline ultimately magnifies the consequences of avoidance behavior, leading to worsening balance function, thereby creating a vicious cycle (4,6).

The Fear of Falling Avoidance Behavior Questionnaire (FFABQ) was created as a tool for researchers and clinicians to quickly and reliably assess avoidance behavior (7). It has been shown to have sound psychometric properties (7) and to be associated with future falls in older adults (8). High FFABQ scores (high avoidance behavior) have also been shown to be related to emotional regulation and depression (9) and vision impairment in community-dwelling older adults (10). Furthermore, the FFABQ has contributed to research involving self-efficacy in older adults regarding fall prevention (11). Although the FFABQ has sound psychometrics, our experience suggests that the Likert responses (completely disagree, disagree, unsure, agree, completely agree) are unclear for some and may not match the sentence stem "Due to my FOF, I avoid [insert activity]." For this reason, there was a necessity to reexamine the language of the FFABQ and make improvements.

A modified version (mFFABQ) was subsequently created to improve the clarity of the Likert responses with the question stem. Specifically, the mFFABQ uses a different Likert response for each of the 14 items, which shifts the focus from agreement to quantification. In the mFFABQ, the item stem is the same, but the updated Likert responses provide a more quantitative focus for each activity: never (0% of the time), rarely (25% of the time), sometimes (50% of the time), often (75% of the time), and always (100% of the time). We believe that the updated mFFABQ responses align more clearly with the stem. The first aim of this study was to examine the test-retest reliability and minimal detectable change (MDC) of the mFFABQ to determine if it had properties similar to the original. The second aim of this study was to provide evidence of the construct validity of the mFFABQ in people with PD and older adults. Specifically, we hypothesized the following:

1. The mFFABQ would be strongly correlated with the original FFABQ (criterion-related validity).
2. Participants with PD, who theoretically have more balance and gait impairment, will have higher mFFABQ scores than older adult participants (known-groups validity).
3. Fallers will have higher mFFABQ scores than non-fallers (known-groups validity).
4. Measures of closely related constructs (e.g., balance confidence, balance performance, mobility and motor function, anxiety, depression, catastrophization) would moderately correlate with the mFFABQ as they should share some variance; strong correlations would not be expected since they are not measuring the same constructs.
5. mFFABQ scores would be predictive of sedentary behavior and fall history and will have suitable cut points for clinical decision-making (predictive validity).

Methods

Study design

This study utilized a cross-sectional design for test-retest reliability wherein participants completed the mFFABQ twice, separated by approximately 1 week. One week was deemed a suitable wash-out period for remembering specific mFFABQ items, but not too long that there was a maturation effect (e.g., worsening or improving condition) or history effect (e.g., fall). All physical performance measures and additional questionnaires were administered during the in-person assessments at the Gait and Balance Research Laboratory at the University of Nevada, Las Vegas by members of the research team, except for the second administration of the mFFABQ. Participants with PD (n = 39) also completed the original FFABQ to allow for comparison of the two questionnaires. Participants wore activity monitors between the two assessments to collect data about their level of physical activity (e.g., time stepping, step count, time sitting/lying). The second mFFABQ was completed at home and returned at the same time as the activity monitor. Construct validity was examined by comparing the mFFABQ to the self-perceived balance confidence, balance, mobility, postural control, affective function, and physical activity levels. Known-groups and convergent validity were analyzed using these same measures.

Sample size estimation

The sample size was estimated using confidence intervals (CIs) for the intraclass correlation module in PASS 20.0.6 (NCSS, LLC.; Kaysville, Utah, USA). Based on data from the original FFABQ reliability study, a sample of 59 participants was needed for Aim 1 (reliability) (7). This estimation was based on a two-way mixed-effects analysis of variance (ANOVA) model (intraclass correlational coefficient [ICC] (3,1)) with each participant measured twice, a two-sided 95% CI with a width of 0.178, and an ICC of 0.815. For Aim 2 (validity), a sample size of 46 would achieve 80% power to detect a Pearson correlation coefficient of 0.40 for convergent validity analyses using a two-sided hypothesis test with a significance level of 0.05.

Participants

Inclusion criteria for both groups were the following: 60 to 90 years old; willingness to participate in one, 60-minute testing session; and willingness to wear an activity monitor for 1 week. Additional inclusion criteria for PD participants were that they had been diagnosed with PD by a neurologist. Participants were excluded if they were unable to read or speak English, exhibited evidence of dementia (Montreal Cognitive Assessment [MoCA] <18 or Mini-Mental State Exam <25) (13), or were unable to stand unassisted for

10 minutes. Participants were recruited from local PD support groups, senior centers, community events, and community centers through print advertisements and snowball recruitment. The study protocol was approved by the University of Nevada, Las Vegas Institutional Review Board. Data were collected from 2014 to 2023.

Measures

To examine construct validity, the mFFABQ was compared to the following:

1. **Self-perceived balance confidence:** The Activities-Specific Balance Confidence Scale (ABC) (14)
2. **Balance, mobility, and postural control:** Berg Balance Scale (BBS) (15), 30-Second Sit to Stand (30STS) (16), Timed Up and Go (TUG) (17), 2-minute step test (2MST) (18), and computerized dynamic posturography – Sensory Organization Test (SOT) (19)
3. **Affective function:** Zung Anxiety Scale (ZAS) (20), Beck Depression Inventory (BDI) (21), and Consequences of Falling Questionnaire (CoFQ) (22)
4. **Physical activity levels:** Average daily activity levels (i.e., time sitting/lying, time standing, time stepping, and number of steps) using a physical activity monitor

These measures and questionnaires were chosen for known-groups and convergent validity, permitting inferences regarding the validity of the mFFABQ.

mFFABQ. The mFFABQ is a 14-item self-report questionnaire with a 5-point Likert scale to measure FFAB (7). Item scores were summed to form a total score ranging from 0 to 56, with higher scores indicating more FFAB.

Self-perceived balance confidence

The ABC is a 16-item self-report measure that evaluates balance confidence during various activities of daily living (14). Evidence for the reliability and validity of the scale has been provided for older adults with and without PD (23,24).

Balance, mobility, and postural control

BBS. The BBS (25) was used as a performance-based balance scale with 14 functional balance tasks (25). It has good evidence for reliability (26) and validity (27) in predicting the risk of falls, multiple falls, and injurious falls in older adults with and without PD (28,29).

30STS. The 30STS was used to measure lower body strength in older adults (16). Evidence suggests excellent reliability in people with PD (30).

TUG. The TUG test was used as a measure of functional mobility in older adults (17). Evidence suggests good reliability and validity in older adults with and without PD (31).

2MST. The 2MST was used to assess aerobic capacity. Evidence suggests good reliability in older adults (18) and is strongly correlated with the Six-Minute Walk Test (32).

SOT. Bertec Balance Computerized Dynamic Posturography (Bertec®, Model 80P-0019, 2500 Citygate Drive, Columbus, OH) was used to calculate a composite balance score based on sway over six conditions (33). Evidence suggests good reliability in older adults (19) and has been shown to be a sensitive tool for identifying fall risk in people with PD (34,35).

Affective function

ZAS. The ZAS, a 20-item, self-report questionnaire, was used to measure anxiety (20). Scores range from 20 to 80 with a higher score indicating a higher level of anxiety (20).

BDI. The BDI, a 21-item self-administered questionnaire, was used to measure symptoms of depression (36). The overall score ranges from 0 to 63, with a higher score suggesting a higher level of depression. The BDI demonstrates high internal consistency in psychiatric and nonpsychiatric populations (37).

CoFQ. The CoFQ, a 12-item, self-report questionnaire, was used to measure catastrophization related to falling (22). The total score ranges from 12 to 48, with a higher score suggesting more catastrophization about falling. It has two subscales, damage to identity and loss of functional independence. Evidence suggests excellent internal reliability and moderate test-retest reliability in older adults (22).

Physical activity levels

Activity monitor. Physical activity levels were measured using ActivPAL activity monitors (PAL Technologies Ltd., Glasgow, United Kingdom) over a 7-day period. Any devices returned with less than 5 days of data were excluded from the analyses. Data extracted included the number of hours per day that the participant was sitting, lying down, biking, or standing. In addition, total steps and time stepping per day were collected.

Data analysis

Data were analyzed using SPSS version 28.0 (IBM SPSS Statistics for Windows, IBM Corp, Armonk, NY, USA) with $\alpha = 0.05$. For Aim 1 (reliability), a two-way mixed-effects ANOVA model ICC (3,1) was used for the two mFFABQ measurements. The MDCs were calculated based on the Standard Error of Measurement (SEM) using the test-retest reliability statistic (ICC value) where r_{xx} = test-retest reliability: $\text{SEM} = \text{baseline standard deviation} \times \sqrt{1 - r_{xx}}$ (38). Once SEM was determined, the MDC at the 95% confidence level was calculated by multiplying the SEM by 1.96 (representing 95% of the area under the curve of a normal distribution) and 1.41 (the square root of 2, to control for possible error associated with calculating the coefficient from two time points). Aim 2 (validity) of the study was to provide evidence for the criterion-related validity of the mFFABQ relative to the original FFABQ; these were compared using Spearman's rho. Additionally, construct validity for the mFFABQ was conducted using known-groups and convergent validity analyses. Known-group analysis was used to determine if there were differences between those with PD and healthy older adults on the mFFABQ. In addition, differences were explored based on fall history, which included fallers or non-fallers in

the previous year (fall status), fallers and non-fallers in the previous month (fall recency), and injurious fallers and non-injurious fallers in the previous year (fall injury) on mFFABQ scores via t-tests. Convergent validity was evaluated by comparing the mFFABQ to measures of similar constructs using Spearman's rho. Since there was likely a nonlinear relationship with falls over time (inverted U curve) (4), the ratio of the number of falls (falls in the last year, last month, and injurious falls) per average daily steps taken was compared to the mFFABQ using Spearman's rho. To determine the optimal cut point for the mFFABQ on sedentary behavior (step-defined sedentary lifestyle index of <5,000 steps per day) (39) and fall history (one or more falls in the last year), the area under the receiver operating characteristic (ROC) curve was calculated and the Youden Index (maximum vertical distance or difference between the ROC curve and the diagonal or chance line) was used to optimize the mFFABQ's ability, given both sensitivity and specificity.

Results

Participants

Ninety-one participants were recruited for the study, 3 participants were excluded due to missing data points, and 3 were excluded due to dementia. A total of 39 participants (age = 72.2 ± 9.5; 29 males, 10 females) diagnosed with PD (Hoehn and Yahr [HY] (12) median and mode = 3.0; frequencies – HY Stage 1 [n = 10], HY Stage 1.5 [n = 1], HY Stage 2 [n = 6], HY Stage 2.5 [n = 1], HY Stage 3 [n = 20], HY Stage 4 [n = 1]) and 49 healthy older adults (age = 72.9 ± 5.0; 13 males, 36 females) participated (Tab. 1).

TABLE 1 - Means with standard deviations, medians with ranges (specified), and proportions for the overall sample and those with PD and older adults

		Overall (n = 88)	**People with PD (n = 39, 44.3%)**	**Older adults (n = 49, 55.7%)**
Demographics	Age	72.6 ± 7.3	72.2 ± 9.5	72.9 ± 5.0
	Sex	42 males (47.7%) 46 females (52.3%)	29 males (74.4%) 10 females (25.6%)	13 males (26.5%) 36 females (73.5%)
PD-specific characteristics	MDS-UPDRS overall	Not applicable	66.2 ± 31.7	Not applicable
	MDS-UPDRS Part 1: mental		13.4 ± 7.3	
	MDS-UPDRS Part 2: activities of daily living		17.3 ± 10.8	
	MDS-UPDRS Part 3: motor		30.2 ± 16.8	
	Freezing of gait from MDS-UPDRS, item 2.13		No freezing = 24 Slight = 6 Mild = 3 Moderate = 2 Severe = 1	
	PDQ39		25.1 ± 15.8	
Balance confidence	ABC	75.2 ± 24.5	64.2 ± 25.9	84.2 ± 19.2
Balance, mobility, and postural control	Falls in the last year (#)	6.8 ± 37.7 Median = 0 (0-350)	14.5 ± 56.1 Median = 0 (0-350)	0.7 ± 1.0 Median = 0 (0-3)
	Falls in the last month (#)	0.7 ± 3.3 Median = 0 (0-30)	1.4 ± 5.0 Median = 0 (0-30)	1.1 ± 0.3 Median = 0 (0-1)
	Injurious falls in the last year (#)	0.4 ± 1.1 Median = 0 (0-8)	0.6 ± 1.4 Median = 0 (0-8)	1.2 ± 0.7 Median = 0 (0-3)
	BBS (scale points)	49.8 ± 9.5	44.4 ± 12.1	54.1 ± 2.4
	30STS (#)	10.2 ±5.5	8.5 ± 7.1	11.6 ± 3.4
	TUG (seconds)	10.9 ± 5.0	13.4 ± 6.3	8.9 ± 2.3
	2MST (#)	65.8 ± 34.0	50.0 ± 37.2	78.4 ± 25.2
	SOT composite (equilibrium score)	62.4 ± 18.0	65.3 ± 26.6	62.2 ± 17.7
Affective function	ZAS (scale points)	41.4 ± 10.5	43.7 ± 11.1	39.7 ± 9.7
	BDI (scale points)	8.8 ± 8.0	12.2 ± 9.3	6.1 ± 5.5
	CoFQ (scale points)	24.7 ± 6.6	27.4 ± 5.1	22.6 ± 7.0

		Overall (n = 88)	People with PD (n = 39, 44.3%)	Older adults (n = 49, 55.7%)
Physical activity levels	Time sitting/lying per day (hours)	18.8 ± 2.4	19.5 ± 2.2	18.2 ± 2.3
	Time standing per day (hours)	3.7 ± 1.9	3.6 ± 1.8	3.9 ± 2.0
	Time stepping per day (hours)	1.3 ± 0.7	0.9 ± 0.6	1.6 ± 0.7
	Number of daily steps (steps)	6,131.6 ± 3,696.5 Median = 5,924 (11-18,457)	4,471.4 ± 2,964.5 Median = 3,966 (11-10,536)	7,533.6 ± 3,699.1 Median = 6,999 (1,973-18,457)

2MST = Two-Minute Step Test; 30STS = 30-Second Sit to Stand; ABC = Activities-Specific Balance Confidence Scale; BBS = Berg Balance Scale; BDI = Beck Depression Inventory; CoFQ = Consequences of Falling Questionnaire; MDS-UPDRS = Movement Disorders Society Unified Parkinson's Disease Rating Scale; PD = Parkinson's disease; PDQ39 = Parkinson's Disease Questionnaire – 39; SOT = Sensory Organization Test; TUG = Timed Up and Go; ZAS = Zung Anxiety Scale.

Reliability

The mFFABQ demonstrated good overall reliability, ICC(3,1) = 0.822 (95% CI: 0.739-0.881) for all participants, including those with mild cognitive impairment. The mFFABQ demonstrated good overall test-retest reliability for older adults and people with PD, ICC (3,1) = 0.781 (95% CI: 0.636-0.871) and 0.806 (95% CI: 0.658-0.894), respectively. The 95% MDC was 14.8 scale points for the overall sample and 12.2 and 17.7 scale points for older adults and people with PD, respectively.

Criterion-related validity

The correlation between the mFFABQ (average of the two scores) and the original FFABQ was rho = 0.874, p < 0.001.

Known-groups validity

Participants with PD had higher mFFABQ scores than older adults, p < 0.001 (Tab. 2). Participants who reported at least one fall in the previous year ("fallers") during the in-person assessment had higher mFFABQ scores than non-fallers, p < 0.001 (Tab. 2). Participants who had experienced a fall in the previous month ("recent faller") had higher mFFABQ scores than nonrecent fallers, p = 0.208 (Tab. 2). There was no difference between those who had experienced a fall injury in the previous year and those who had not, p = 0.471 (Tab. 2).

Convergent validity

For all participants, the mFFABQ was significantly correlated with fall history (fall status rho = −0.430, p < 0.001) and fall recency (rho = −0.235, p = 0.031) but not with fall injuries (rho = 0.173, p = 0.113). The correlations between the mFFABQ and the ratio of steps per day and falls (fall-to-step) were as follows: falls/year/step (rho = 0.630, p < 0.001), falls/month/steps (rho = 0.209, p = 0.189), and injurious falls/year/steps (rho = 0.172, p = 0.282). The mFFABQ also correlated with the ABC (rho = −0.804, p < 0.001), BBS (rho = −0.595, p < 0.001), TUG (rho = 0.560, p < 0.001), and 30STS (rho = −0.386, p < 0.001). The mFFABQ correlated with the ZAS (rho = 0.428, p < 0.001), BDI (rho = 0.606, p < 0.001), and CoFQ (rho = 0.582, p < 0.001) including damage to identity (rho = 0.608, p < 0.001) and loss of functional independence (rho = 0.497, p < 0.001) subscales of the CoFQ. For physical activity, the mFFABQ did not correlate with sitting/lying (rho = 0.129, p = 0.248) or standing time (rho = −0.072, p = 0.520); however, it did correlate with time stepping (rho = −0.298, p = 0.007) and number of steps (rho = −0.358, p < 0.001) in a direction consistent with the hypotheses and the construct.

TABLE 2 - Known-groups validity comparisons on the mFFABQ

Comparison	Group	mFFABQ	SEM	Cohen's D with 95% CI (Hedges correction)	Statistic	p-Value
Diagnosis	People with PD (n = 38, 44.7%)	16.4	2.3	0.777 (0.335 to 1.215)	t = 3.596	<0.001
	Healthy older adults (n = 47, 55.3%)	7.6	1.3			
Fall status	Fallers (n = 44, 51.8%)	16.4	2.1	0.917 (0.471 to 1.359)	t = 4.264	<0.001
	Non-fallers (n = 41, 48.2%)	6.3	1.1			
Fall recency	Recent fallers (n = 15, 17.6%)	15.1	2.4	0.357 (−0.199 to 0.912)	t = 1.268	0.208
	Nonrecent faller (n = 70, 82.4%)	10.8	1.5			
Fall injury	Fall injury (n = 20, 23.5%)	13.3	2.1	−0.184 (−0.681 to 0.314)	t = −0.725	0.471
	Nonfall injury (n = 65, 76.5%)	11.0	1.6			

CI, confidence interval; mFFABQ = modified Fear of Falling Avoidance Behavior Questionnaire; PD = Parkinson's disease; SEM = standard error of the mean.

Predictive validity

The area under the ROC curve was 0.720 (95% CI: 0.613-0.828) with an optimal cut point of 11.5 on the mFFABQ (scores range from 0 to 56) for predicting sedentary behavior (<5,000 steps per day) (39) (Fig. 1). The sensitivity and specificity of the 11.5 cut point were 0.667 and 0.702, respectively. The area under the ROC curve for fall history (one or more falls in the last year) was 0.723 (95% CI: 0.618-0.827) and the optimal cut point was 13.5, with a sensitivity of 0.551 and specificity of 0.810 (Fig. 2).

Discussion

An important objective of the original FFABQ was to create a reliable, clinically feasible, and accessible tool to assess FFAB (7). The changes made in the mFFABQ were implemented to improve clarity and, thus, reliability and validity. Our results provide evidence that the mFFABQ has acceptable reliability for the overall sample (ICC = 0.822), older adults (ICC = 0.781), and people with PD (ICC = 0.806). Although these reliability coefficients are solidly in the "good reliability" range (40), they represent modest improvements over the original FFABQ (overall ICC = 0.812) and for people with neurological conditions (ICC = 0.751) (7). The reliability coefficients for the mFFABQ were consistent with the Brazilian Portuguese FFABQ in older adults (ICC = 0.810) (41) but were lower than the Turkish FFABQ in older adults (ICC = 0.999) (42). Based on these data, we recommend that both the FFABQ and mFFABQ are suitable for clinical or research use. Still, we favor the mFFABQ because the Likert options are more quantitative and, based on our experience, make more theoretical sense to the research team and some of the participants than the original. While we did not keep track, there were more clarifying questions from participants about the FFABQ than the mFFABQ. Additionally, the modest improvements in the reliability coefficients support our recommendation to use the mFFABQ. Also, there have been no studies reporting evidence for the validity and reliability of the original FFABQ in a PD population; thus, the evidence reported in this study for the mFFABQ supports our recommendation for its use. In addition, with an MDC of 15 scale points (14.8), which is consistent with the original, a therapist or researcher can be confident that a change in score beyond this value would be indicative of an increase or decrease beyond error at 95% of confidence in FFAB.

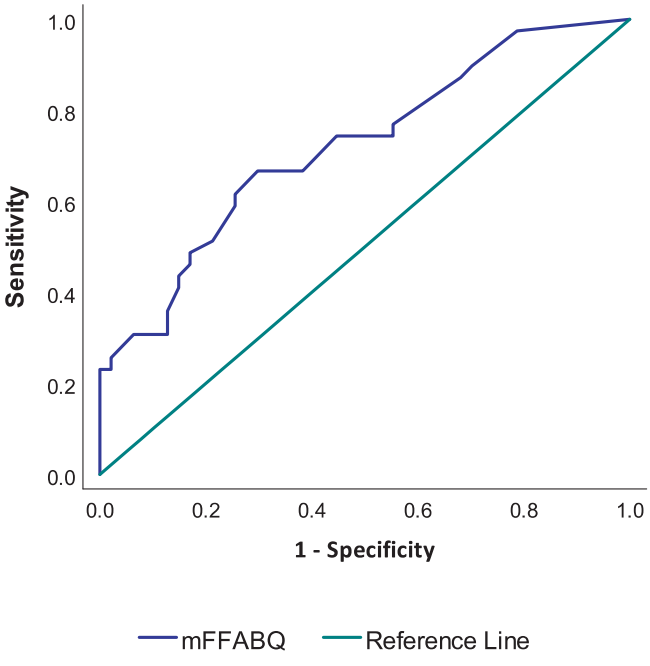


FIGURE 1 - Receiver operating characteristic (ROC) curve for the modified Fear of Falling Avoidance Behavior Questionnaire (mFFABQ) on sedentary activity (<5,000 steps per day).

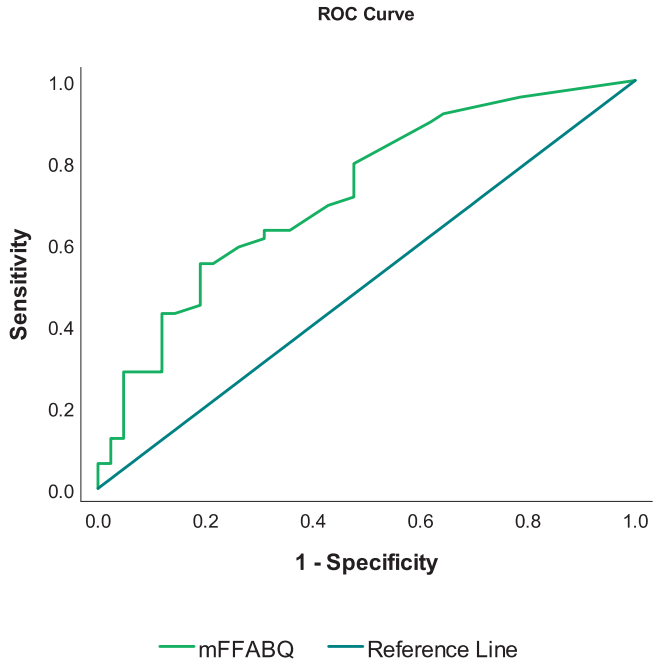


FIGURE 2 - Receiver operating characteristic (ROC) curve for the modified Fear of Falling Avoidance Behavior Questionnaire (mFFABQ) on fall history (one or more falls in the last year).

The criterion-related validity of the mFFABQ was supported by its strong correlation with the original FFABQ as initially hypothesized. In the original study of the FFABQ, the results supported the notion that the FFABQ measured FFAB rather than balance confidence, self-efficacy, or fear (7). The results of the present study are consistent with the original study and further support the validity of these instruments. Likewise, as hypothesized, the results of the known-groups analyses support the validity of the mFFABQ. That is, those with gait and balance dysfunction inherent to their disease (known group [PD]) would have higher mFFABQ scores (i.e., more FFAB) than healthy older adults, who would logically have less gait and balance dysfunction and, subsequently, lower mFFABQ scores. This was indeed the case, and these differences were also observed among other known groups, including fallers and recent fallers. These results add evidence

about the validity of the mFFABQ in these populations in discriminating between two known groups that would logically differ in the construct of the instrument.

Similar to the original FFABQ (7), the mFFABQ correlated with performance-based balance measures (i.e., BBS, TUG, and 30STS), which supports the convergent validity of the mFFABQ. The correlations between these performance-based tools make theoretical sense because they are in the same gait and balance domain. It was hypothesized that the correlations would be moderate. If the correlations would have been strong, then that would suggest that they were measuring the same construct or significantly overlapping constructs. The strongest correlation was found with the ABC. This is logical because balance self-confidence is a closer construct to FOF and, subsequently, FFAB. The directionality and strength of these correlations are consistent with the results from the Brazilian Portuguese FFABQ translation, reliability, and validity study (41) but were lower than the Turkish FFABQ translation, reliability, and validity, which reported stronger correlations (42). Despite being in the same domain, fall history was not consistently or strongly correlated with the mFFABQ, reinforcing the idea that falls are nonlinear over time or as a disease progresses (inverted U curve) (4). It is theorized that individuals with high FOF and FFAB, triggered by decreased balance capability/confidence, limit or recalibrate their exposure to risky balance conditions and, thus, are less likely to fall (4). Moreover, the correlations between the mFFABQ and measures in the gait and balance domain suggest that individuals with increased FFAB are likely to demonstrate impaired balance with functional activities (4). This supports the notion that increased FFAB may decrease fall frequency but does not decrease postural instability (43).

In additional support of the construct of validity of the mFFABQ, there were also moderate, positive correlations with the following scales in the affective domain: ZAS (anxiety), BDI (depression), and CoFQ (catastrophization). These results are consistent with our original hypotheses and are also consistent with other studies (5,44). While many consider anxiety and fear to be related (both deal with the idea of danger or threat) but different constructs (fear is seen as a reaction to a specific, observable threat, while anxiety is worry about a future threat that has not happened or may never happen), they are clearly interrelated and it is not surprising that there are moderate correlations between these constructs and FFAB. However, the cause-and-effect direction is not known from our study and, subsequently, it is possible that this relationship could be bidirectional, with FOF triggering a generalized anxiety disorder or, alternatively, someone with an anxiety disorder could be more susceptible to developing fear in other aspects of their life. The relationship could also be more complex with mediator and moderator effects. Likewise, the relationship with depression makes theoretical sense and could also be bidirectional. One theory regarding this is that FFAB may have downstream consequences such as social isolation and loneliness, which could, in turn, trigger or exacerbate depression (4). Furthermore, as hypothesized, there was a moderate correlation with the CoFQ (catastrophization), particularly the damage to identity subsection. This is consistent with research suggesting that the FFAB is more strongly correlated with damage to identity (i.e., the immediate consequences of pain, shame, and embarrassment) than loss of functional independence (i.e., enduring consequences of injury and disability) in people with PD (44). From a clinical context, it is important to holistically consider the associations of the mFFABQ with constructs in the affective domains and to collaborate with other members of the healthcare team with expertise in this area. Because therapists frequently encounter FOF and FFAB during gait and balance treatment, it is important that therapists become well-versed in these areas to mitigate the consequences and optimize care.

As hypothesized, convergent validity of the mFFABQ was also supported by moderate correlations with time stepping and the number of steps taken on average per day. Because avoidance behavior likely affects activity levels, these correlations support the notion that those with high avoidance behavior exhibit more sedentary behavior (less time stepping and fewer steps per day). As functional balance declines, a person is likely to cope through increased sedentary behaviors and avoiding activities that challenge balance (45). These results are consistent with other studies in the literature for people with PD and support the notion of a vicious cycle of FFAB (4-6). In addition, predictive validity was supported by the ROC analyses, which suggest that the mFFABQ is predictive of sedentary behavior (i.e., less than 5,000 steps per day) at a cut-off score of 11.5 (AUC = 0.720) and also falls within the last year (AUC = 0.723 with a 13.5 mFFABQ cut point). These results are consistent with FFAB predicting future falls in older adults (8). From a clinical perspective, because FFAB may have several negative downstream consequences, including a vicious cycle (4,6), the mFFABQ may be a helpful clinical tool in a comprehensive examination for clinical decision-making related to sedentary behavior, activity limitation and participation restriction, fall and balance behavior, and outcomes of different treatment approaches to mitigate the downstream consequences of FFAB. These treatment approaches may include high-intensity multimodal exercise with balance training (46) and cognitive behavioral therapy (4,47,48).

One of the limitations of this study was that many participants had low FFAB, especially in the older adult group. Recruiting people with high FFAB is challenging because their FFAB makes them less likely to leave their homes and travel to an urban campus, which would likely entail significant walking, physical performance tests, and other factors that would feed into their FOF. Thus, the results of this study may not be fully generalizable to typical clinical populations that are most likely to be seen and evaluated for gait and balance problems. Future research in this area should consider conducting assessments in participant homes to remove some barriers to participation for individuals with high FFAB. Another limitation was that the sample size estimation was for the overall sample and, subsequently, the subgroups may not have been sufficiently powered. Therefore, the subgroup analyses should be interpreted with some caution. However, psychometrically, the subgroup analyses were actually quite strong so this may only be a minor concern. Another limitation was that the sex proportion of the participants was

different in the participants with PD (males > females) and older adults (females > males). Neither of these proportions are inconsistent with expectations of clinical research in these populations (i.e., there are more males with PD, more females volunteer for research studies); however, this does limit the generalizability. Another limitation is that two scales that also measure the construct of FFAB, the Falls Efficacy Scale – International (49) and Survey of Activities and Fear of Falling in the Elderly (50), were not included in this study. They would have added support to the construct validity of the mFFABQ. Lastly, only participants with PD completed both the original and modified versions of the FFABQ; therefore, the correlational data reported in this article should be interpreted with caution and should not be generalized to older adults.

Conclusions

The results of this study provide sound psychometric support for the use of the mFFABQ as a clinical or research measure for FFAB in older adults and people with PD. Similar to the original FFABQ, the mFFABQ exhibited good reliability and demonstrated good evidence of its validity in the measurement of the construct of FFAB. These results also demonstrated a modest improvement in psychometric properties relative to the original and, therefore, it is recommended that clinicians and researchers use the updated, mFFABQ version. However, the original FFABQ remains a suitable measure. Researchers and clinicians should adhere to a single scale and not intermix them.

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: Conceptualization (MRL), data curation (MRL, AMH, AA, BL), formal analysis (MRL, AMH, AA, BL, AAB), investigation/methodology (MRL, AMH, AA, BL, AAB), project administration (MRL, JVR, JKL), writing – original draft (MRL, AMH, AA, BL), writing – review and editing (MRL, AMH, AA, BL, AAB, JVR, JKL).

Data availability statement: Data for this study are not currently available as this was not stipulated in the IRB-approved protocol.

Ethics approval statement: Ethical approval was received from the University of Nevada, Las Vegas Biomedical Institutional Review Board. All authors have adhered to the ethical standards stipulated in the IRB-approved trial.

Patient consent statement: All participants consented to participate by signing an IRB-approved consent form.

References

1. Salari N, Darvishi N, Ahmadipanah M, Shohaimi S, Mohammadi M. Global prevalence of falls in the older adults: a comprehensive systematic review and meta-analysis. J Orthop Surg Res. 2022;17(1):334. CrossRef PubMed
2. Allen NE, Schwarzel AK, Canning CG. Recurrent falls in Parkinson's disease: a systematic review. Parkinsons Dis. 2013;2013:906274. CrossRef PubMed
3. Lee S, Oh E, Hong GS. Comparison of factors associated with fear of falling between older adults with and without a fall history. Int J Environ Res Public Health. 2018;15(5):982. CrossRef PubMed
4. Landers MR, Nilsson MH. A theoretical framework for addressing fear of falling avoidance behavior in Parkinson's disease. Physiother Theory Pract. 2023;39(5):895-911. CrossRef PubMed
5. Landers MR, Lopker M, Newman M, Gourlie R, Sorensen S, Vong R. A cross-sectional analysis of the characteristics of individuals with Parkinson disease who avoid activities and participation due to fear of falling. J Neurol Phys Ther. 2017;41(1):31-42. CrossRef PubMed
6. Landers MR, Jacobson KM, Matsunami NE, McCarl HE, Regis MT, Longhurst JK. A vicious cycle of fear of falling avoidance behavior in Parkinson's disease: a path analysis. Clin Park Relat Disord. 2021;4:100089. CrossRef PubMed
7. Landers MR, Durand C, Powell DS, Dibble LE, Young DL. Development of a scale to assess avoidance behavior due to a fear of falling: the Fear of Falling Avoidance Behavior Questionnaire. Phys Ther. 2011;91(8):1253-1265. CrossRef PubMed
8. Landers MR, Oscar S, Sasaoka J, Vaughn K. Balance confidence and fear of falling avoidance behavior are most predictive of falling in older adults: prospective analysis. Phys Ther. 2016;96(4):433-442. CrossRef PubMed
9. Scarlett L, Baikie E, Chan SWY. Fear of falling and emotional regulation in older adults. Aging Ment Health. 2019;23(12):1684-1690. CrossRef PubMed
10. Lee SP, Hsu YW, Andrew L, Davis T, Johnson C. Fear of falling avoidance behavior affects the inter-relationship between vision impairment and diminished mobility in community-dwelling older adults. Physiother Theory Pract. 2022;38(5):686-694. CrossRef PubMed
11. Soh SL, Tan CW, Thomas JI, et al. Falls efficacy: extending the understanding of self-efficacy in older adults towards managing falls. J Frailty Sarcopenia Falls. 2021;6(3):131-138. CrossRef PubMed
12. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. Neurology. 1967;17(5):427-442. CrossRef PubMed
13. Hoops S, Nazem S, Siderowf AD, et al. Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease. Neurology. 2009;73(21):1738-1745. CrossRef PubMed
14. Powell LE, Myers AM. The Activities-specific Balance Confidence (ABC) scale. J Gerontol A Biol Sci Med Sci. 1995;50A(1):M28-M34. CrossRef PubMed
15. Berg K, Wood-Dauphinee S, Willliams JI, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. Physiother Can. 1989;41(6):304-311. CrossRef
16. Rikli RE, Jones CJ. Development and validation of a functional fitness test for community-residing older adults. J Aging Phys Act. 1999;7(2):127-129. CrossRef
17. 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
18. Rikli RE, Jones CJ. Development and validation of criterion-referenced clinically relevant fitness standards for maintaining physical independence in later years. Gerontologist. 2013;53(2):255-267. CrossRef PubMed
19. Ford-Smith CD, Wyman JF, Elswick RK Jr, Fernandez T, Newton RA. Test-retest reliability of the sensory organization test in noninstitutionalized older adults. Arch Phys Med Rehabil. 1995;76(1):77-81. CrossRef PubMed
20. Zung WW. A rating instrument for anxiety disorders. Psychosomatics. 1971;12(6):371-379. CrossRef PubMed
21. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry. 1961;4(6):561-571. CrossRef PubMed

22. Yardley L, Smith H. A prospective study of the relationship between feared consequences of falling and avoidance of activity in community-living older people. Gerontologist. 2002; 42(1):17-23. CrossRef PubMed
23. Talley KM, Wyman JF, Gross CR. Psychometric properties of the activities-specific balance confidence scale and the survey of activities and fear of falling in older women. J Am Geriatr Soc. 2008;56(2):328-333. CrossRef PubMed
24. Jonasson SB, Nilsson MH, Lexell J. Psychometric properties of four fear of falling rating scales in people with Parkinson's disease. BMC Geriatr. 2014;14(1):66. CrossRef PubMed
25. Donoghue D, Stokes EK; Physiotherapy Research and Older People (PROP) group. How much change is true change? The minimum detectable change of the Berg Balance Scale in elderly people. J Rehabil Med. 2009;41(5):343-346. CrossRef PubMed
26. Meseguer-Henarejos AB, Rubio-Aparicio M, López-Pina JA, Carles-Hernández R, Gómez-Conesa A. Characteristics that affect score reliability in the Berg Balance Scale: a meta-analytic reliability generalization study. Eur J Phys Rehabil Med. 2019; 55(5):570-584. CrossRef PubMed
27. Park SH, Lee YS. The diagnostic accuracy of the Berg Balance Scale in predicting falls. West J Nurs Res. 2017;39(11):1502-1525. CrossRef PubMed
28. Muir SW, Berg K, Chesworth B, Speechley M. Use of the Berg Balance Scale for predicting multiple falls in community-dwelling elderly people: a prospective study. Phys Ther. 2008;88(4):449-459. CrossRef PubMed
29. Qutubuddin AA, Pegg PO, Cifu DX, Brown R, McNamee S, Carne W. Validating the Berg Balance Scale for patients with Parkinson's disease: a key to rehabilitation evaluation. Arch Phys Med Rehabil. 2005;86(4):789-792. CrossRef PubMed
30. Petersen C, Steffen T, Paly E, Dvorak L, Nelson R. Reliability and minimal detectable change for Sit-to-Stand Tests and the Functional Gait Assessment for individuals with Parkinson disease. J Geriatr Phys Ther. 2017;40(4):223-226. CrossRef PubMed
31. Christopher A, Kraft E, Olenick H, Kiesling R, Doty A. The reliability and validity of the Timed Up and Go as a clinical tool in individuals with and without disabilities across a lifespan: a systematic review. Disabil Rehabil. 2021;43(13):1799-1813. CrossRef PubMed
32. Berlanga LA, Matos-Duarte M, Abdalla P, Alves E, Mota J, Bohn L. Validity of the two-minute step test for healthy older adults. Geriatr Nurs. 2023;51:415-421. CrossRef PubMed
33. Grove CR, Whitney SL, Hetzel SJ, Heiderscheit BC, Pyle GM. Validation of a next-generation sensory organization test in adults with and without vestibular dysfunction. J Vestib Res. 2021;31(1):33-45. CrossRef PubMed
34. Freeman L, Gera G, Horak FB, Blackinton MT, Besch M, King L. Instrumented test of sensory integration for balance: a validation study. J Geriatr Phys Ther. 2018;41(2):77-84. CrossRef PubMed
35. Rossi-Izquierdo M, Basta D, Rubio-Rodríguez JP, et al. Is posturography able to identify fallers in patients with Parkinson's disease? Gait Posture. 2014;40(1):53-57. CrossRef PubMed
36. Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clin Psychol Rev. 1988;8(1):77-100. CrossRef
37. Wang YP, Gorenstein C. Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. Br J Psychiatry. 2013;35(4):416-431. CrossRef PubMed
38. Beaton DE, Bombardier C, Katz JN, Wright JG. A taxonomy for responsiveness. J Clin Epidemiol. 2001;54(12):1204-1217. CrossRef PubMed
39. Tudor-Locke C, Craig CL, Thyfault JP, Spence JC. A step-defined sedentary lifestyle index: <5000 steps/day. Appl Physiol Nutr Metab. 2013;38(2):100-114. CrossRef PubMed
40. 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
41. Nunes ACL, Jesus-Moraleida FR, Lima POP, Landers MR. Translation, cross-cultural adaptation, and measurement properties of the Brazilian-Portuguese version of the fear of falling avoidance behavior questionnaire in older adults. J Aging Phys Act. 2022;30(6):1014-1023. CrossRef PubMed
42. Acaröz Candan S, Demircioğlu A, Şahin ÜK. Cross-cultural adaptation and psychometric properties of the Fear of Falling Avoidance Behavior Questionnaire in Turkish community-dwelling older adults. Disabil Rehabil. 2021;43(23):3387-3394. CrossRef PubMed
43. Longhurst JK, Rider JV, Eckard K, et al. Factors predicting fear of falling avoidance behavior in parkinsonisms. NeuroRehabilitation. 2022;50(1):65-73. CrossRef PubMed
44. Rider JV, Longhurst JK, Lekhak N, Navalta JW, Young DL, Landers MR. Psychological factors associated with fear of falling avoidance behavior in Parkinson's disease: the role of depression, anxiety, and catastrophizing. J Geriatr Psychiatry Neurol. 2023 May;36(3):215-224. CrossRef PubMed
45. Nilsson MH, Drake AM, Hagell P. Assessment of fall-related self-efficacy and activity avoidance in people with Parkinson's disease. BMC Geriatr. 2010;10(1):78. CrossRef PubMed
46. Landers MR, Navalta JW, Murtishaw AS, Kinney JW, Pirio Richardson S. A high-intensity exercise boot camp for persons with Parkinson disease: a phase II, pragmatic, randomized clinical trial of feasibility, safety, signal of efficacy, and disease mechanisms. J Neurol Phys Ther. 2019;43(1):12-25. CrossRef PubMed
47. Tennstedt S, Howland J, Lachman M, Peterson E, Kasten L, Jette A. A randomized, controlled trial of a group intervention to reduce fear of falling and associated activity restriction in older adults. J Gerontol B Psychol Sci Soc Sci. 1998;53(6):384-392. CrossRef PubMed
48. Dorresteijn TA, Zijlstra GA, Ambergen AW, Delbaere K, Vlaeyen JW, Kempen GI. Effectiveness of a home-based cognitive behavioral program to manage concerns about falls in community-dwelling, frail older people: results of a randomized controlled trial. BMC Geriatr. 2016;16(1):2. CrossRef PubMed
49. Yardley L, Beyer N, Hauer K, Kempen G, Piot-Ziegler C, Todd C. Development and initial validation of the Falls Efficacy Scale-International (FES-I). Age Ageing. 2005;34(6):614-619. CrossRef PubMed
50. Lachman ME, Howland J, Tennstedt S, Jette A, Assmann S, Peterson EW. Fear of falling and activity restriction: the survey of activities and fear of falling in the elderly (SAFE). J Gerontol B Psychol Sci Soc Sci. 1998;53(1):43-50. CrossRef PubMed

Arch Physioter 2024; 14: 20-28
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3023
REVIEW

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 - The Netherlands

ABSTRACT

Introduction: Worldwide many countries provide direct access in physiotherapy. The aim of this scoping review was to synthesize the available evidence on the quality of primary care musculoskeletal physiotherapy from different perspectives.
Methods: Systematic searches were conducted in three databases up to September 2022. Studies were included when regarding assessment of at least one of the following perspectives: patient (quality of Life, patient satisfaction, pain, functioning, adverse events), provider (treatment compliance, responsibility, liability, status, prestige, job satisfaction), and society (number of referrals, amount of medical imaging, medication use, number of sessions needed for rehabilitation, and overall costs and cost-effectiveness). Selection and methodological quality assessment of systematic reviews were performed. Data extraction and analysis were performed separately for systematic reviews and individual primary studies.
Results: Five systematic reviews as well as 17 primary studies were included. From a patient perspective, no significant effect of direct access was found for pain and a tendency in favour of direct access was found for quality of life, functioning, and well-being. Concerning providers, higher treatment compliance was found in direct access to physiotherapy and decision-making was more accurate. From a societal perspective, significant differences in favour of direct access physiotherapy were found for waiting time, prescribed medication, and medical imaging. In addition, there was a tendency towards lower health care costs.
Conclusions: Emerging evidence suggests that direct access physiotherapy could provide at least equal quality of care for patients and better opportunities for providers and the society on selected outcomes.

Keywords: Direct access, Physiotherapy, Quality of care, Scoping review

What is already known

- *Direct access physiotherapy has proven to be a valid strategy in primary musculoskeletal care.*

What this study adds

- *The article brings together the results of previous systematic reviews and additionally includes those of recent randomized controlled trials. The review suggests that direct access physiotherapy could provide at least equal quality of care for patients and better opportunities for providers and society compared to physiotherapy on referral.*

Received: January 19, 2024
Accepted: June 5, 2024
Published online: July 1, 2024

Corresponding author:
Erik Cattrysse
email: Erik.cattrysse@vub.be

Introduction

Musculoskeletal (MSK) disorders are among the top ten leading causes of years lived with disability (YLD) (1,2). Mounting evidence suggests that the quality of care offered in primary care settings treating these MSK disorders is sub-optimal and that often inaccurate diagnoses are made by

primary care physicians (3,4). Physiotherapy is a frequently recommended treatment option for the management of MSK disorders (5). Responses to the latest World Confederation for Physical Therapy survey reveal direct access (DA) is available in 48 countries and there is no restriction on private practice in 77 countries (6). In countries where physiotherapists receive professional autonomy through DA, evidence suggests several benefits including a more valid diagnosis as compared to primary care physicians, better outcomes for patients, and more efficient use of resources, while maintaining high patient satisfaction (5,7).

Further benefits can be linked to DA physiotherapy, such as shorter waiting times, reduced health care costs including physician fees, medical imaging expenses, and medication costs (5,8), increased prestige for physiotherapists (9,10), and decreased workload for primary care physicians (10). However, cost reduction may be restricted to direct costs and general workload for the Physical Therapist (PT) may not necessarily be reduced (11). Also, potential disadvantages to this model of health care have been described, for instance, potential erosion of a strong patient-doctor relationship (12,13) or a robust physiotherapy-doctor connection (13), as well as concerns about overconsumption of physiotherapy services (14).

This scoping review aimed to identify, appraise, and synthesize existing literature to assess the impact of DA on primary care physiotherapy for patients presenting with various MSK disorders. The impact of DA will focus on outcomes from the perspectives of the patient, the provider, and society.

Methods

The reporting of this scoping review conforms to the PRISMA Extension for Scoping Reviews (PRISMA-ScR) guidelines (15).

MEDLINE (PubMed) and Web of Science were searched from 1990 until March 2024. The electronic search strategy used in these searches is listed in Table 1.

TABLE 1 - Search strategy used in MEDLINE (PubMed) and Web of Science

(“referral and consultation”[MeSH Terms] OR “direct access” OR “self-referred” OR “self-referral” OR “primary care”) AND (physical therapy modalities[MeSH Terms] OR modality physical therapy[MeSH Terms] OR “physical therapy” OR “physiotherapy” OR “physical therapist” OR “physiotherapist” OR “rehabilitation”[MeSH Terms]) AND (“quality of life”[MeSH Terms] OR “assessment”, “outcomes”[MeSH Terms] OR “pain”[MeSH Terms] OR “back pain” OR “neck” OR “musculoskeletal subjective reporting” OR “discomfort” OR “injuries”[MeSH Terms] OR “trauma” OR “disability” OR “activities” OR “recovery” OR “safety” OR “sick leave”[MeSH Terms] OR “patient satisfaction”[MeSH Terms] OR “disability” OR “disability leave” OR “disability leaves” OR “illness days” OR “cost-effectiveness” OR “economic evaluation” OR “cost analysis” OR “analyses cost”[MeSH Terms] OR “cost” OR “cost projection analysis”)

Titles and abstracts were independently reviewed by two reviewers up to March 2024, applying the following inclusion criteria: availability of quantitative data of at least one group that received physiotherapy through DA or direct allocation without consulting a physician and assessment of at least one of the perspectives for the patient (quality of life [QoL], well-being, satisfaction, pain, functioning, or adverse events), physiotherapists (treatment compliance, responsibility, liability, status, prestige, or job satisfaction), and society (number of referrals with and without a DA setting, amount of medical imaging, medication use, number of sessions needed for rehabilitation, and overall costs and cost-effectiveness). Articles written in English, Dutch, or French were considered. Papers not complying with the inclusion criteria were excluded. Randomized clinical trials were selected and studies retrieved by the above search string which were published after the latest systematic review were added to this scoping review.

The Risk of Bias in Systematic Reviews (ROBIS) tool was developed specifically to be used by guideline developers, authors of overviews of systematic reviews (“reviews of reviews”), and review authors who might want to assess or avoid risk of bias in their reviews (16). The ROBIS tool was utilized by two reviewers independently to assess the risk of bias in the included systematic reviews. Discrepancies between the two reviewers were resolved by discussion and if disagreement persisted, a third reviewer made the final decision.

Results

The flow of studies through the review is presented in Fig. 1.

Systematic reviews

Five systematic reviews were included in this scoping review. Table 2 summarizes the authors and dates of the primary studies included in each of these reviews. From this overview, it can be concluded that, overall, 56 individual studies were covered. Each of the reviews employed its own methodological quality evaluation protocol regarding the included studies (see Tab. 3). Regarding the assessment based on the ROBIS tool, one review showed an overall low risk of bias (17). On each of the different domains, at least one study scored low risk of bias and all domains were scored at high risk in at least one study. Data collection was scored as unclear for two studies. Overall bias in one study was considered unclear and in the three remaining there was a high risk.

Patient perspective

The low risk of bias review by Babatunde et al (17) found no significant differences in *pain reduction* between DA physiotherapy and care supervised by a general practitioner (GP). Similar findings were reported by Piscitelli et al (18) and Demont et al (19), although small differences in favour of DA physiotherapy were noticed ($p = 0.76$) (18). Ojha et al (20) reported a significant but small result for pain reduction in favour of DA physiotherapy ($p = 0.011$).

Babatunde et al (17) reported no significant results for *QoL and function* while Ojha et al (20) and Piscitelli et al (18) showed better outcomes in terms of QoL and function in

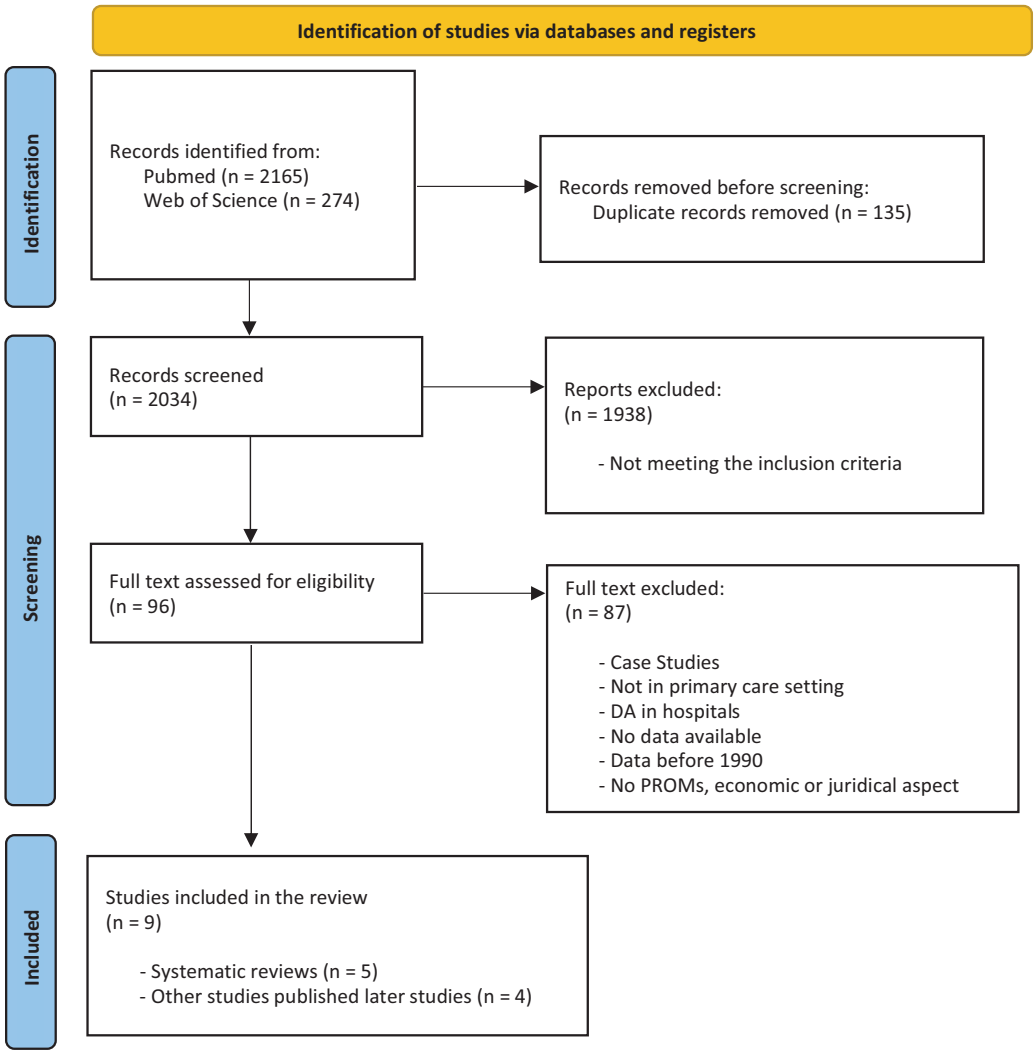


FIGURE 1 - PRISMA flow chart of the study selection.

favour of DA physiotherapy (p = 0.04 and p = 0.03, respectively). According to Ojha et al (20) and Demont et al (19), patients reported 5.0-21.5% greater satisfaction in the DA groups than in physician referral groups (p < 0.01). Additionally, Gallotti et al (21) reported equal to superior QoL and a tendency to higher patient satisfaction in DA groups.

Physiotherapist perspective

Piscitelli et al (18) and Demont et al (19) both showed significantly higher *treatment compliance* in DA physiotherapy compared to GP referral physiotherapy (p = 0.004). No other results were reported on aspects of the physiotherapists' perspective.

Societal perspective

Piscitelli et al (18) and Demont et al (19) concur that DA physiotherapy can reduce the *waiting time* for primary consults by 4 to 63 days (p < 0.001). Similarly, Gallotti et al (21) reported shorter waiting times as well as improved management accuracy regarding the type of access to PT (i.e. by GP referral, access by consultant, or DA physiotherapy).

Four out of five systematic reviews showed consistent results regarding the amount of *prescribed medication* and *medication use* (17-20). DA physiotherapy led to 11.9-65.0% less prescribed medication (p < 0.01) and reduced pharmacological costs by $42-710 (p < 0.01) (17,20).

Ojha et al (20) reported significantly fewer *physiotherapy visits* in a DA setting, with a range of 1.1-13.4 visits (p < 0.01). Demont et al (19) found no consensus about the number of physiotherapy visits, with either two to three fewer physiotherapy visits needed in a DA setting (p = 0.001) or no significant difference found. Babatunde et al (17) and Piscitelli et al (18) reported that DA physiotherapy led to 2.0-21.5% fewer *follow-up visits* with the primary care physician (p < 0.05). Gallotti et al. (21) indicate a shorter time to discharge in DA. Demont et al (19) reported that 17% fewer patients required a primary care physician visit in a DA physiotherapy setting (p = 0.0113).

TABLE 2 - Overview of the included studies in the five systematic reviews on direct access physiotherapy

Ojha et al 2014 (n = 8)	Piscitelli et al 2018 (n = 12)	Demont et al 2019 (n = 18)	Babatunde et al 2020 (n = 26)	Gallotti et al 2023 (n = 28)
Hackett et al 1993	Gentle et al 1984	Overman et al 1988	Greenfield et al 1975	Daker-White et al 1999
Mitchell et al 1997	Hackett et al 1993	Holdsworth et al 2004	Mitchell et al 1997	Oldmeadow et al 2007
Holdsworth et al 2004	Mitchell et al 1997	Moore et al 2005	Overman et al 1988	Sephton et al 2010
Moore et al 2005	Holdsworth et al 2004	Holdsworth et al 2006	Ferguson et al 1999	Ludvigsson et al 2012
Holdsworth et al 2007	Holdsworth et al 2006	Holdsworth et al 2007	Moore et al 2005	Phillips et al 2012
Webster et al 2008	Holdsworth et al 2007	Brooks et al 2008	Holdsworth 2007/2008	Kooijman et al 2013
Leemrijse et al 2008	Leemrijse et al 2008	Leemrijse et al 2008	Bossonnaulth et al 2010	Salisbury et al 2013
Pendergast et al 2012	Brooks et al 2008	Webster et al 2008	Pendergast et al 2012	Mallett et al 2014
	Webster et al 2008	Ludvigsson et al 2012	Phillips et al 2012	O'farrell et al 2014
	Ludvigsson et al 2012	Pendergast et al 2012	Chetty et al 2012	Samsson et al 2014
	Pendergast et al 2012	Mallett et al 2014	Ludvigsson et al 2012	Bornhöft et al 2015
	Badke et al 2014	Swinkels et al 2014	McCallum et al 2012	Samsson et al 2015
	Bishop et al 2017	Bomhöft et al 2015	McGill et al 2013	Bird et al 2016
		Mintken et al 2015	Badke et al 2014	Kerridge-Weeks et al 2016
		Goodwin et al 2016	Mallett et al 2014	Samsson et al 2016
		Bishop et al 2017	Swinkels et al 2014	Bishop et al 2017
		Bomhöft et al 2019	Bornhöft et al 2015	Chang et al 2018
		Downie et al 2019	Mintken et al 2015	Bornhöft et al 2019
			Ojha et al 2015	Caffrey et al 2019
			Boissonnaulth et al 2016	Downie et al 2019
			Goodwin et al 2016	Lankhorst et al 2020
			Harland et al 2016	Ojha et al 2020
			Pearson et al 2016	Oostendorp et al 2020
			Bishop et al 2017	Peterson et al 2021
			Mant et al 2017	Ho-Henrikson et al 2022
			Denninger et al 2018	Lyons et al 2022
				Szymanek et al 2022

Shaded sections refer to primary studies that have been analysed also in previous systematic reviews.

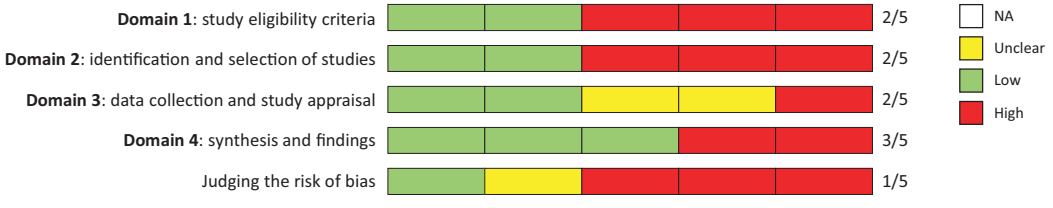


FIGURE 2 - Risk of bias assessment of included systematic reviews (n = 5).

Four systematic reviews concluded that DA physiotherapy could lead to 6.3-70.0% fewer X-rays and other *medical imaging* ($p < 0.001$) (17,18, 20,21). Babatunde et al (17) and Piscitelli et al (18) also showed lower *overall health costs* ($p < 0.01$) up to 20%. Ojha et al (20) and Demont et al (19) showed decreased costs in a DA physiotherapy setting compared to a GP referral setting ($p < 0.05$). This was further supported by Gallotti et al (21).

Babatunde et al (17) and Gallotti et al (21) reported less *work-related absence* and sick leave in DA physiotherapy. Ojha et al (20) reported an average of 17.4 days less work absence in a DA physiotherapy. Piscitelli et al (18) did not find a consensus for the *return-to-work* rate. They found either no difference in return-to-work rate or 14.1% less lost time from work and daily duties ($p < 0.05$).

TABLE 3 - Summary of systematic reviews (n = 5)

First author (year of publication)	Patient perspective	Physiotherapist perspective	Societal perspective
Ojha et al (2014) (20)	Significantly less average pain was reported in DA. (VAS decreased from 5.7 to 2.7 vs. from 5.7 before treatment to 3.2 after treatment; p = 0.011.) Higher improvement in function in DA (79% vs. 60%; p = 0.04). Higher levels of satisfaction in DA (79%-93.2% vs. 74%-84.1%; p < 0.01). DA PT is not related to any adverse events (p>0.05).	/	Lower costs for radiological examinations in DA PT (5.1%-13.6%; p = NR)/ (4.00 vs. 7.43; p < 0.01). Significantly fewer medication prescribed in DA PT (32.2%-48% vs. 44.1%-84%; p < 0.01). Significantly less PT visits in DA PT (5.9-20.2 vs. 7.0-33.6; p < 0.01). Decreased cost in DA (£9.55-$14.83 vs. £47.94-$63.65; p < 0.01). Mean number of days of work missed was significantly less in DA PT (10.4 vs. 27.9 days: p = NR).
Piscitelli et al (2018) (18)	No significant difference between GP referral and DA PT in terms of pain reduction (64.6% vs. 66.6%; p = 0.76). DA PT has better clinical outcomes in terms of function and health-related QoL (self-reported questionnaires: 2.4 ± 2.8 vs. 4.1 ± 4.6; p = 0.03). DA PT is not related to any adverse events (p>0.05).	DA PT provides greater treatment compliance (76%-79% vs. 58%-69%; p = 0.004).	DA PT reduces waiting time and improves accessibility of PT (5-6 vs. 9-69 days; p < 0.001). DA PT leads to less medical imaging ($44 ± 190 vs. $175 ± 541; p < 0.01)/(47 vs. 242 patients; p < 0.001). DA PT leads to less prescribed medication ($36-163 vs. $78-873; p < 0.01)/ (62-79 vs. 219-276 patients; p < 0.001).
Demont et al (2021) (19)	No significant difference between DA and physician-led usual medical care for pain but tendencies to positive effects (p≥0.05). Better clinical outcomes for function in DA PT (RMQ function score 4.1 in GP referral vs. 2.4 in DA PT; p = 0.03). No significant difference in QoL, but tendencies to positive effects (p≥0.05). Significantly higher satisfaction levels in DA compared to physician-led usual medical care (74.7% vs. 53.2%; p = 0.002). DA PT is not related to any adverse events (p>0.05).	Significantly higher treatment compliance in DA (76%-79% vs. 58%-69%; p = 0.004).	Significantly shorter waiting time with DA PT (4-5 vs. 9-31 days; p < 0.001). Significantly less X-ray imaging described in DA PT (7.3% vs. 13.6%; p < 0.001). Significantly less medication prescribed in DA PT (32.2% vs. 44.1%; p < 0.01). Fewer or the same number of PT visits in DA PT (two to three fewer sessions; p = 0.001). Significant fewer patients required a primary care physician visit in DA (54% vs. 71%; p = 0.0113). Average cost saving per episode is significantly lower in DA PT (average saving = £36.42/patient/episode of care; p = 0.016)/(average costs £66.31 vs. £89.99; p < 0.05).
Babatunde et al (2020) (17)	No significant differences between patient characteristics in DA/self-referral vs. medical triage (p≥0.05) regarding age and gender. No significant differences in pain, functional outcomes, and QoL for patients who assessed MSK care via DA compared to GP-led care. No reports of adverse effects or missed red flags.		Less work-related absence and sick leave in DA PT. 10%-20% less total cost for DA PT compared to GP-led care. >65% less prescribed analgesics and NSAID. >70% less radiology. 2%-10% lower follow-up consults.
Gallotti et al (2023) (21)	Better management accuracy, less waiting time, and equal to higher QoL.	Higher rate of presence; shorter time to discharge (+6%).	Higher cost-effectiveness (€−441.9 per episode of care) Imaging −28%; medication −41.2%; referral −20.7% Less sick leave (5% less prescriptions, −37 less sick leave days prescribed)

DA = direct access; GP = general practitioner; MSK = musculoskeletal; NSAID = nonsteroidal anti-inflammatory care; PT = physiotherapy; QoL = quality of life; RMQ = Roland-Morris questionnaire; VAS = visual analogue scale.

Primary studies

Four primary studies, subsequent to the randomized controlled trial (RCT) conducted by Gallotti et al (21), contribute pertinent and insightful information to this scoping. Among these, one is a pilot RCT and three are retrospective cohort studies, which are elucidated further in the subsequent section.

Reddington et al (22) conducted a pilot RCT, employing qualitative analysis to examine patient expectations and experiences concerning accelerated access to physiotherapy. They engaged participants diagnosed with sciatica (n = 33) in individual interviews (n = 46) recruited from 14 National Health Service (NHS) primary care general practices and a physiotherapy service provider in the UK. Their findings indicate that expedited access to physiotherapy holds merit in terms of perceived recovery enhancement and/or mitigation of further physical and psychological decline. Negative patient expectations of physiotherapy predominantly stemmed from prior experiences of unfruitful physiotherapy. Based on their overarching study outcomes, the authors advocate for an individualized patient-centric approach alongside expedited access to physiotherapy for sciatica patients.

Crowell et al (23) conducted a retrospective cohort study to assess adherence to the low back pain Healthcare Effectiveness Data and Information Set (HEDIS) within the United States Military Health System Data Repository. They compared military personnel suffering from acute back pain regarding interventions administered in a DA physical therapy clinic vs. a general primary care clinic. Results indicate that in the physical therapy clinic, 96.7% of encounters did not entail imaging orders within the initial 28 days of symptom onset, compared to 82.0% in the primary care clinic ($p < 0.001$). The authors conclude that PTs operating in a DA setting are notably more inclined than primary care providers to adhere to low back pain imaging guidelines, particularly in young, athletic patients.

Wood et al (24) conducted another retrospective cohort study comprising a substantial qualitative analysis based on patient free-text reports concerning experiences with first contact physiotherapists (FCP) for MSK issues. Of the participants (n = 498), 73% reported being “extremely likely” to recommend the FCP service to friends and family, while 22% reported “likely” to recommend it. Conversely, only 1% would not recommend the service. Most respondents highlighted the communication skills of the FCP, emphasizing the importance of clear and understandable information provision. Additionally, respondents valued a diagnosis and treatment plan, as well as consultation with knowledgeable specialists. Self-management skills and shared decision-making were also perceived as valuable components. A small proportion of respondents reported unresolved conditions or dissatisfaction due to delays in treatment. Respondents appreciated being treated with respect and empathy, often comparing FCP consultations favourably to those with GPs.

The study by Halfpap et al (25) aimed to evaluate health care utilization and associated outcomes for Active Duty Service Members (ADSM) receiving services at an acute spine pain clinic (ASPC) during its initial 5 years of operation at a large military treatment facility in the United States. The most common chief complaint among 1,215 ADSM patients was acute lumbar spine pain (73%), followed by cervical spine pain (15%), with thoracic spine pain representing the fewest cases (12%). On average, patients attended 3.5 physical therapy visits (range 1-13), with the majority (61.1%) utilizing three or fewer visits. A review of medical records for 100 randomly selected patients within 12 months of their initial evaluation indicated reduced medication use, imaging, and referrals to surgical services. The authors concluded that the DA physiotherapy approach demonstrates potential benefits in terms of rapid access to treatment and education for patients with acute spine pain, facilitated by PTs in military treatment facilities.

In summary, it can be stated that the studies of Reddington et al (22) and Wood et al (24) indicate that from a patient’s perspective, several advantages are experienced especially regarding perceived recovery enhancement and high communication and information skills of PTs as well as the shared decision-making and self-management approach. However, also a small number of negative experiences are reported related to delayed referral for further treatment in case of unresolved conditions.

From a more societal perspective, the study from Crowell et al (23) indicates a better adherence to actual treatment guidelines in case of DA physiotherapy and resulting reduced medication use, imaging, and referrals to surgical services according to Halfpap et al (25).

Table 4 provides a summary of the individual studies presenting information such as the study and country, design and aim of the study, study setting, population and sample size, intervention, and outcome measures.

TABLE 4 - Summary of primary studies (n = 4)

Individual studies				
Study and country	**Design, aim of the study**	**Study setting, population, sample size**	**Intervention**	**Outcome measures**
Crowell et al (2022) (23) United States	Retrospective data analysis cohort study	To compare rates of compliance with the National Committee for Quality Assurance – Healthcare Effectiveness Data and Information Set (HEDIS) recommendations for diagnostic imaging in low back pain between physical therapists and primary care providers in young, athletic patients	Analysis of 1,845 Military Health System Data Repository (MDR) data	In the physical therapy clinic, 96.7% of encounters did not have imaging ordered within the first 28 days of onset of symptoms, compared with 82.0% in the primary care clinic ($p < 0.001$).

(*Cont*.)

TABLE 4 - (*Continued*)

Individual studies				
Study and country	**Design, aim of the study**	**Study setting, population, sample size**	**Intervention**	**Outcome measures**
Halfpap et al (2022) (25) United States	Retrospective trial on low back pain in military	To compare DA PT in (acute) low back pain to random retrospective sample	1,215 patients compared to 100 randomly selected patients' medical records	Medication: 26% in PT vs 20% in non-PT and 47.4%-72% in the ED Radiographs: 7% in PT vs. 28% in non-PT vs. 26.1 in the ED Complex Imaging: 1% in PT vs. 12% in non-PT vs. 8.2% in the ED
Reddington et al (2022) (22) UK	Descriptive nested qualitative study via semi-structured patient interview	To explore sciatica patients' experiences with DA	80 patients with sciatica	This study suggests that accelerated access to physiotherapy has value in terms of aiding perceived recovery and/or halting further physical and psychological decline
Wood et al (2022) (24) UK	Online survey	Patient-reported experience and outcomes for DA	680 reported questionnaires and 785 free-text responses	Approximately 70% of participants reported no need for consulting other health care professionals

DA = direct access; ED = emergency department; PT = physiotherapy.

Discussion

The purpose of this scoping review was to analyse the impact of DA on the quality of MSK primary physiotherapy care from the perspectives of the patient, the provider, and society. No differences were found for pain reduction, QoL, functioning, and well-being although some systematic reviews indicated a tendency in favour of DA physiotherapy. Higher treatment compliance and a more accurate decision-making were found concerning the providers' perspective and, finally, differences in favour of DA were found for waiting time, prescribed medication, and medical imaging. Less work-related absence and a clear reduction in health care costs were reported in some studies. Although some of the included reviews had a high risk of bias, their findings are in full agreement with the recent low-bias review of Babatunde et al (17).

Proponents of the DA system argue many advantages of this system. For patients, the most important advantage is that the physiotherapist becomes much more accessible, and patients lose less valuable time in the diagnostic process (26). Other advantages were expected to be found in terms of QoL, functioning, and patient satisfaction. From a clinical aspect, DA physiotherapy performs at least as well as physiotherapy by referral. However, all primary studies had relatively short follow-up of about 1 month, and the added benefits of DA may not be captured in the absence of a long-term follow-up.

For physiotherapists, an advantage could be found in the prestige of their profession and the greater responsibility involving an interesting challenge for physiotherapists in the functional evaluation of the patient (9,10). Moreover, several other health professions, such as chiropractors and osteopaths, are directly accessible, while their training courses are less focused on the diagnosis and screening of red and yellow flags (8). Red flags may indicate the presence of a serious underlying cause explaining the current symptoms. However, this must be put in perspective, as the evidential value for red flags has proved insufficient to state that they are excellent predictors of serious underlying disorders (12). Yellow flags identify underlying patient characteristics that could potentially lead to a slower recovery process or ending up in chronicity. It is very important to identify the presence of yellow flags to avoid non-response to the treatment of patients (27).

Furthermore, DA physiotherapy could be a step forward in the physiotherapists' autonomy and development as a diagnostician. However, it should not become compulsory as some physiotherapists may not consider themselves competent or do not support it. In addition, it would be useful for older physiotherapists who did not follow the most modern training, especially in clinical reasoning and diagnosis, to be given further training in the field of diagnosis. However, no information emerged from this review regarding the responsibility, liability, status, prestige, and job satisfaction of the physiotherapists.

Evidence suggests greater treatment compliance of patients and fewer missed appointments in a DA physiotherapy setting, allowing the physiotherapist to spend their time optimally (21). Some diagnoses made by a GP do not provide added value in the clinical reasoning of the rehabilitation plan. Some diagnoses made by a GP are actually superfluous, as the pattern of symptoms may not have a clear pathophysiological foundation (28,29). The best example of this is the well-known non-specific low back pain phenomenon. In many cases, no underlying pathophysiological mechanism can be found. Treatment by the physiotherapist is then based on the pattern of symptoms and not on the prescribed medical diagnosis (29). Some studies (13) reported that the decision-making ability of physiotherapists is great, but they do not consistently recognize the need for immediate referral. Physiotherapists with MSK specialization were more likely to make correct decisions for patients with MSK conditions and critical medical conditions (13). But overall, the reported

results may indicate that training might have to be adapted for this purpose.

Consensus on the benefits of the society perspective has not been found in all areas. DA physiotherapy can decrease the workload and pressure of GPs because a large number of patients with MSK disorders proceed directly to a physiotherapist. In this way, GPs have more time to focus on their other patients. It can also lead to less waiting time, less medical imaging, and fewer prescription and use of medication (17-21). Some studies (17, 19-21) reported less work-related absence and sick leave in DA physiotherapy but others did not find a consensus for the return-to-work rate (18). And, in general, there is evidence of an important cost reduction in health care for MSK disorders (18-21).

As a final argument in favour of DA physiotherapy, it can be noticed that there is no reporting of countries in which DA physiotherapy has been introduced where it was subsequently rejected. This shows that the advantages at least counterbalance, but presumably overweigh the possible disadvantages (30).

Limitations of this scoping review

The information gathering was restricted to English, Dutch, and French and two different databases (MEDLINE and Web of Science), possibly causing some relevant articles to be missed. A specific additional search in the PEDro database for systematic reviews did not reveal any additional publications.

Recommendations for future research

Different types of studies, preferably high-quality RCTs, should be conducted, focusing on various perspectives that remain unanswered or unclear. Economic evaluations could be performed from the societal perspective, and several options exist for designing studies from both the patients' and therapists' perspectives.

Conclusions

This scoping review suggests that DA physiotherapy can offer multiple advantages over GP referral physiotherapy. Although no significant effects were found for pain and QoL, strong evidence from one unbiased study, and supported by some lower quality evidence, indicates that DA does not result in a significant decrease in functional outcome. As such DA physiotherapy seems to be as beneficial to the patient as physiotherapy by referral. Moreover, evidence indicates that it does reduce the use of medical imaging and leads to less prescription and use of medication, resulting in costs. The small significant differences in favour of DA physiotherapy from the patients' perspective, combined with no loss in terms of pain reduction, suggest at least an equal level of care quality. Moreover, it seems that DA physiotherapy does not have any adverse effects on patients. This, coupled with predominantly positive benefits from DA physiotherapy from a societal perspective, suggests that the advantages of DA physiotherapy are more situated in the societal domain.

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: All authors actively participated in the preparation of this manuscript. EC: data curation, formal analysis, investigation, methodology, project administration, supervision, validation, writing original draft, writing review and editing. JVdB: data curation, formal analysis, investigation, supervision writing original draft. RP and AT: data curation, formal analysis, writing original draft. AS: formal analysis, validation. EVT: conceptualization, methodology, supervision, writing review and editing.

Data availability statement: Additional material may be requested directly from the responsible author. The information generated and analysed during the current study are publicly available by means of the mentioned search engines and publications in the reference list

References

1. Murray CJ, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2197-2223. CrossRef PubMed
2. Vos T, Barber RM, Bell B, et al; Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2015;386(9995):743-800. CrossRef PubMed
3. Davenport TE, Sebelski CA. The physical therapist as a diagnostician: how do we, should we, and could we use information about pathology in our practice? Phys Ther. 2011;91(11):1694-1695. CrossRef PubMed
4. Williams CM, Maher CG, Hancock MJ, et al. Low back pain and best practice care: a survey of general practice physicians. Arch Intern Med. 2010;170(3):271-277. CrossRef PubMed
5. Foster NE, Hartvigsen J, Croft PR. Taking responsibility for the early assessment and treatment of patients with musculoskeletal pain: a review and critical analysis. Arthritis Res Ther. 2012;14(1):205. CrossRef PubMed
6. World Physiotherapy. Country profiles reveal global state of the physical therapy profession 2019. Online. Accessed January 2024.
7. Bürge E, Monnin D, Berchtold A, Allet L. Cost-effectiveness of physical therapy only and of usual care for various health conditions: systematic review. Phys Ther. 2016;96(6):774-786. CrossRef PubMed
8. Mitchell JM, de Lissovoy G. A comparison of resource use and cost in direct access versus physician referral episodes of physical therapy. Phys Ther. 1997;77(1):10-18. CrossRef PubMed
9. American Physical Therapy Association (APTA). Vision Statement for the Physical Therapy Profession. 2022. Online. Accessed January 2024.
10. Leemrijse CJ, Swinkels ICS, Veenhof C. Direct access to physical therapy in the Netherlands: results from the first year in community-based physical therapy. Phys Ther. 2008;88(8):936-946. CrossRef PubMed
11. Bishop A, Ogollah RO, Jowett S, et al; STEMS study team. STEMS pilot trial: a pilot cluster randomised controlled trial to investigate the addition of patient direct access to physiotherapy to usual GP-led primary care for adults with musculoskeletal pain. BMJ Open. 2017;7(3):e012987. CrossRef PubMed

12. Galliker G, Scherer DE, Trippolini MA, Rasmussen-Barr E, LoMartire R, Wertli MM. Low back pain in the emergency department: prevalence of serious spinal pathologies and diagnostic accuracy of red flags. Am J Med. 2020;133(1):60-72.e14. CrossRef PubMed
13. Jette DU, Ardleigh K, Chandler K, McShea L. Decision-making ability of physical therapists: physical therapy intervention or medical referral. Phys Ther. 2006;86(12):1619-1629. CrossRef PubMed
14. Brooks G, Dripchak S, Vanbeveren P, Allaben S. Is a prescriptive or an open referral related to physical therapy outcomes in patients with lumbar spine-related problems? J Orthop Sports Phys Ther. 2008;38(3):109-115. CrossRef PubMed
15. 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
16. Whiting P, Savović J, Higgins JPT, et al; ROBIS group. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol. 2016;69:225-234. CrossRef PubMed
17. Babatunde OO, Bishop A, Cottrell E, et al. A systematic review and evidence synthesis of non-medical triage, self-referral and direct access services for patients with musculoskeletal pain. PLoS One. 2020;15(7):e0235364. CrossRef PubMed
18. Piscitelli D, Furmanek MP, Meroni R, De Caro W, Pellicciari L. Direct access in physical therapy: a systematic review. Clin Ter. 2018;169(5):e249-e260. PubMed
19. Demont A, Bourmaud A, Kechichian A, Desmeules F. The impact of direct access physiotherapy compared to primary care physician led usual care for patients with musculoskeletal disorders: a systematic review of the literature. Disabil Rehabil. 2021;43(12):1637-1648. CrossRef PubMed
20. Ojha HA, Snyder RS, Davenport TE. Direct access compared with referred physical therapy episodes of care: a systematic review. Phys Ther. 2014;94(1):14-30. CrossRef PubMed
21. Gallotti M, Campagnola B, Cocchieri A, Mourad F, Heick JD, Maselli F. Effectiveness and consequences of direct access in physiotherapy: a systematic review. J Clin Med. 2023;12(18):5832. CrossRef PubMed
22. Reddington M, Baxter S, Walters SJ, Cole A. Patients' expectations and experiences of accelerated access to physiotherapy for sciatica: qualitative findings from a mixed-methods study. Physiotherapy. 2022;115:102-110. CrossRef PubMed
23. Crowell MS, Mason JS, McGinniss JH. Musculoskeletal imaging for low back pain in direct access physical therapy compared to primary care: an observational study. Int J Sports Phys Ther. 2022;17(2):237-246. CrossRef PubMed
24. Wood L, Bishop A, Goodwin R, Stynes S. Patient satisfaction with the first contact physiotherapy service: results from the national evaluation survey. Musculoskeletal Care. 2022;20(2):363-370. CrossRef PubMed
25. Halfpap J, Riebel L, Tognoni A, Coller M, Sheu RG, Rosenthal MD. Improving access and decreasing healthcare utilization for patients with acute spine pain: five-year results of a direct access clinic. Mil Med. 2022 Mar 12:usac064. CrossRef PubMed
26. GOV-UK. Our health, our care, our say: a new direction for community services. 2006. Online. Accessed January 2024.
27. Stewart J, Kempenaar L, Lauchlan D. Rethinking yellow flags. Man Ther. 2011;16(2):196-198. CrossRef PubMed
28. Casazza BA. Diagnosis and treatment of acute low back pain. Am Fam Physician. 2012;85(4):343-350. PubMed
29. Pinnington MA, Miller J, Stanley I. An evaluation of prompt access to physiotherapy in the management of low back pain in primary care. Fam Pract. 2004;21(4):372-380. CrossRef PubMed
30. Deyle GD. Direct access physical therapy and diagnostic responsibility: the risk-to-benefit ratio. J Orthop Sports Phys Ther. 2006;36(9):632-634. CrossRef PubMed

Arch Physioter 2024; 14: 29-46
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3063
REVIEW

Hip microinstability and its association with femoroacetabular impingement: A scoping review

Rahel Caliesch[1,2], David Beckwée[1], Jan Taeymans[3,4], Joseph M. Schwab[5], Thomas Renaud[6], Quentin Brossard[7], Roger Hilfiker[8]

[1]Department of Physiotherapy, Human Physiology and Anatomy, Faculty of Physical Education and Physiotherapy, Rehabilitation Research (RERE) Research Group, Vrije Universiteit Brussel, Brussels - Belgium
[2]School of Health Sciences, University of Applied Sciences and Arts Western Switzerland, Valais, Sion - Switzerland
[3]Division of Physiotherapy, Department of Health Professions, University of Applied Sciences Bern, Bern - Switzerland
[4]Faculty of Physical Education and Physiotherapy, Vrije Universiteit Brussel, Brussels - Belgium
[5]Department of Orthopaedic Surgery and Traumatology, HFR Fribourg – Cantonal Hospital, University of Fribourg, Fribourg - Switzerland
[6]Institute for Physiotherapy Research, Brügg, Bern - Switzerland
[7]Medical Center Wankdorf - Department of Sport Physiotherapy, Bern - Switzerland
[8]Physiotherapie Tschopp & Hilfiker, Glis - Switzerland

ABSTRACT

Introduction: Hip microinstability has become a recognized cause of non-arthritic hip pain and disability in young patients. However, its pathophysiology remains unclear. We want to (1) present an overview of the evidence of hip microinstability and of its association with femoroacetabular impingement (FAI), (2) map out the type of evidence available, and (3) make recommendations for future research.
Methods: A deductive analysis and extraction method was used to extract information. In addition, diagnostic accuracy statistics were extracted or calculated.
Results: Of the 2,808 identified records, 123 were eligible for inclusion. Different definitions for microinstability exist. A standardized terminology and clear diagnostic criteria are lacking. FAI and microinstability may be associated and may aggravate each other. Conservative treatment strategies for FAI and microinstability are similar. The reported prevalence of microinstability in combination with FAI ranges from 21% to 42% in adults undergoing hip arthroscopy or magnetic resonance arthrography (MRA) of the hip.
Conclusion: Hip microinstability and FAI may be associated, occur together, or exacerbate each other. To better address this topic, a standardized terminology for microinstability is essential. Achieving consensus on physical examination and diagnosis is also necessary. Initial efforts to establish uniform diagnostic criteria have been made, but further work is needed. Specifically, randomized controlled trials are required to evaluate the effectiveness of training programmes aimed at reducing symptoms in individuals with microinstability, with or without FAI. Such studies will enable clinicians to manage microinstability with greater confidence within this context.

Keywords: Femoroacetabular impingement (FAI), Hip impingement, Hip microinstability, Scoping review

What is already known about this topic:

- *Hip microinstability became increasingly recognized as a cause of non-arthritic hip pain and disability in young and active people, just as it is the case with femoroacetabular impingement (FAI). There is no consensus on the diagnostic criteria for hip microinstability. Treatments very similar to those for FAI are proposed. However, studies on the efficacy of conservative treatment for hip microinstability are lacking.*

What does the study add:

- *This study clarifies the concept of hip microinstability and elucidates the relationship between microinstability and FAI. An overview of the evidence on the definition, diagnosis, aetiology, prevalence, and treatments of hip microinstability, and of its broader association with FAI are presented.*

Received: March 13, 2024
Accepted: July 8, 2024
Published online: August 1, 2024

This article includes supplementary material

Corresponding author:
Rahel Caliesch
email: rahel.caliesch@hevs.ch

Introduction

Hip microinstability is a relatively new diagnosis and not yet well established (1). Recently, this condition has received increasing interest as a medical picture responsible for non-arthritic hip pain and disability, particularly in young and active people (2,3). Another condition found in young and active people is FAI syndrome (4).

At the International Hip-related Pain Research Network meeting in Zürich, Switzerland, in 2018 (4), three categories for hip-related pain were proposed: (1) FAI syndrome, (2) acetabular dysplasia and/or hip instability, and (3) other conditions causing hip-related pain (including labrum, cartilage, and ligamentum teres lesions without a specific bony morphology). However, there is no consensus on the diagnostic criteria for hip microinstability (4,5).

The hip was believed to be a stable joint through his bony architecture (6). As the understanding of hip mechanics improved, it appeared that some hips are not as stable as thought (7). Hip stability is ensured by the bony, as well as the soft-tissue and muscle structures (1). Thus, bone abnormalities constitute anatomic risk factors for microinstability (8). On imaging, many patients with hip microinstability showed signs of dysplasia, but also of FAI morphologies (3,9). Thus, FAI and microinstability may not be mutually exclusive and may coexist.

The management of hip microinstability lacks clear establishment (6). Researchers propose surgical and conservative therapy strategies, which include strengthening exercises for the hip and core muscles, as well as activity modification (6,8). Thus, these treatments closely resemble those for FAI. However, researchers lack high-level studies on the efficacy of conservative treatment for hip microinstability (6). The effect of conservative care for FAI has been investigated in four randomized controlled trials (RCTs) (10-13). The surgical treatments aim to correct the underlying deformity in each case. However, they may differ substantially between the two diagnoses.

To the best of our knowledge, no systematic or scoping reviews highlighting microinstability in the context of hip impingement have been published to date.

There is a clear need to investigate the relation between microinstability and FAI, especially in the context of diagnosis and understanding of nonsurgical treatments of these two conditions.

This scoping review aims to: (1) present an overview of the evidence on the definition, diagnosis, aetiology, prevalence, and potential treatments of hip microinstability, and of its broader association with FAI, (2) map out the type of evidence available, and (3) make specific recommendations for future research.

Methods

Protocol and registration

The protocol for this scoping review was published on osf.io (DOI: CrossRef). The review was conducted according to recommendations of the JBI (formerly known as Joanna Briggs Institute) group (14,15). The authors wrote the manuscript according to the extension of the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* for Scoping Reviews (PRISMA-ScR) checklist (16).

Eligibility criteria

All types of study design were included in the review. No language or publication date restrictions were applied. Articles about hip microinstability and people with or suspected of having pincer or cam morphology were included. Studies about instability after total hip arthroplasty, hip dislocation, traumatic instability, developmental dysplasia; studies including infants/toddlers; studies comparing different surgical techniques; as well as cadaveric and animal studies were excluded.

Search

Electronic searches were performed in MEDLINE (Ovid), CINAHL, and EMBASE from inception up to 12 July 2023. Reference lists of included articles were also screened for additional articles.

The full search strategy for each database is presented in the Supplementary Appendix A1.

Selection of sources of evidence

Reviewer 1 and reviewer 2 screened half of the abstracts and full texts, reviewer 1 and reviewer 3 screened the other half of the abstracts and full texts. A fourth reviewer was contacted in case of disagreement to determine a final decision. The flow diagram of study selection is shown in Figure 1.

Data charting process and data items

A data charting form (Supplementary Appendix A2) was developed and used to extract general source information (type of evidence, author, publication year) as well as key messages from each study on the following topics: definition, diagnosis, aetiology, prevalence, and treatment. Data extraction was not linear, but an iterative process. A deductive analysis and extraction method was used to extract contextual information from each study, extracted as text and grouped into separate sheets in Excel. Quantitative data (description of sample, group differences, etc.) was extracted in a further Excel form. For the diagnostic tests, true-positive, false-positive, false-negative, and true-negative frequencies were extracted or calculated.

Synthesis of results

A cross-table presenting the design of the studies and the topics covered was created, with an overlayed heat map (Fig. 2). A thematic construction was used to provide an overview of key concepts regarding definition, aetiology, diagnosis, treatment, and prevalence. Key messages were analysed chronologically. For the diagnostic tests, sensitivity, specificity, as well as positive and negative diagnostic likelihood ratios were extracted or calculated and presented with forest plots (Fig. 3 and Supplementary Tab. A2).

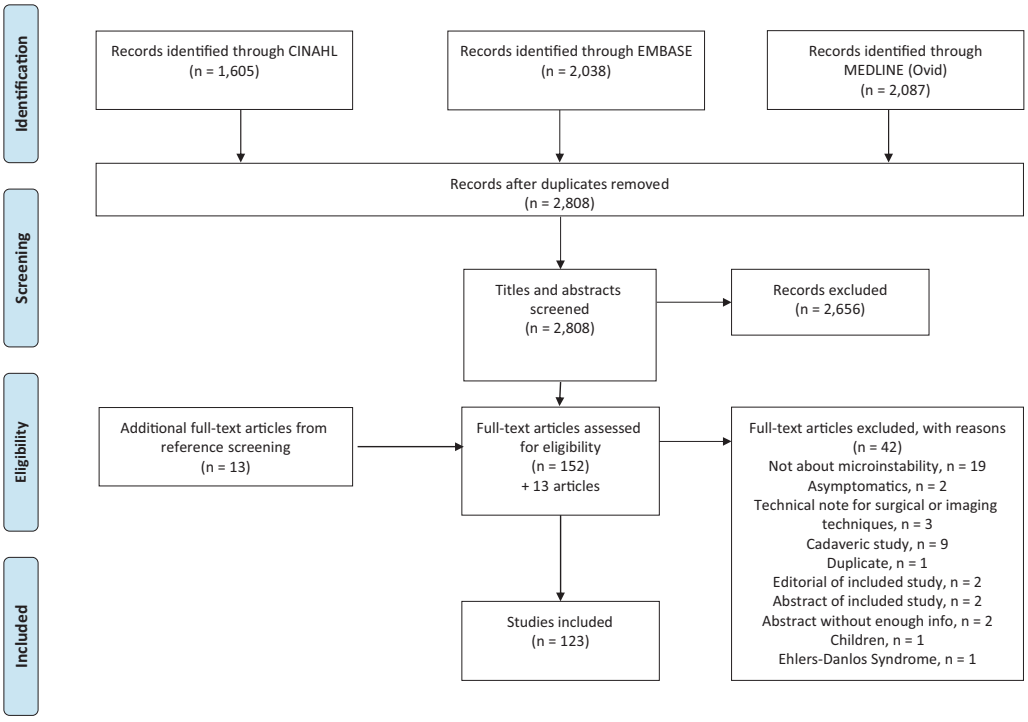


FIGURE 1 - Flow diagram.

Total number of articles	**Diagnosis** *70*	**Aetiology** *49*	**Treatment** *42*	**Definition** *16*	**Prevalence** *9*
Review Article	14	17	12	9	
Expert Article	3	6	4	2	
Editorial/Commentary	4	4	6	1	
Retrospective Cohort Study	5	2	2	1	2
Prospective Diagnostic Accuracy Study	6				4
Retrospective Case Series Study	4	1	5		
Retrospective Case Control Study	6	2	1		
Prospective Cohort Study	4	3	1		1
Systematic Review	5	2	2		
Technical Note	1	3	2		
Case Report	2	1	2	1	
Scientific Meeting	1	1	1		
Retrospective Validation Study	2				1
Infographic	1	1	1		
Prospective Case Control Study		2			
Prospective Diagnostic Reliability Study	1			1	
Prospective Case Series Study	1	1			
Retrospective Diagnostic Accuracy Study	1				1
Cross-Sectional Study	1	1			
Consensus Paper	2				
Book Chapter	1			1	
Meta Analysis			1		
Scoping Review			1		
Prospective Descriptive Laboratory Study		1			
Prospective Diagnostic Case Series Study	1				
Retrospective Diagnostic Accuracy and Prospective Reliability Study	1				
Retrospective Cohort and Prospective Diagnostic Validity Study	1				
Retrospective Nested Case Control Study	1				
Retrospective Reliability Study	1				
Laboratory Model and Retrospective Cohort Study		1			
Intervention Study			1		

FIGURE 2 - Heat map of topic by study type.

Results

Selection of sources of evidence

The electronic search yielded a total of 2,808 records, after removal of duplicates. After additional reference screening, 123 articles met the inclusion criteria (Fig. 1). General study description, together with information about which topic is addressed in each study are listed in Supplementary Tab. A1.

Characteristics of sources of evidence

A heat map (Fig. 2) gives an overview of the different study types by topic. There were 31 types of evidence.

Synthesis of results

Definition

The most frequently cited definition of hip microinstability is the one by Shu and Safran (17): "Hip instability can be defined as extra-physiologic hip motion that causes pain with or without the symptom of hip joint unsteadiness. The cause can be traumatic or atraumatic and is related to both bony and soft-tissue abnormality. Gross instability is caused by trauma or iatrogenic injury. Subtle microinstability, from microtraumatic or atraumatic causes, is an evolving concept." Notably, the authors did not make any distinction between hip instability and microinstability definition, except for the causes leading to instability.

Another frequently quoted definition is the one by Cerezal et al (18): "Hip microinstability is the inability to keep the femoral head centred within the acetabular fossa, without complete luxation or marked subluxation of the joint. Microinstability is laxity with the presence of symptoms. Asymptomatic hip joint laxity is not microinstability."

Table 1 summarizes all studies presenting a definition of microinstability, with the respective sources.

TABLE 1 - Studies defining or citing a definition of hip microinstability

First author Year	Title	Definition of microinstability	Citation of definition
Shu 2011 (17)	Hip instability: anatomic and clinical considerations of traumatic and atraumatic instability	"Hip instability is uncommon because of the substantial conformity of the osseous femoral head and acetabulum. It can be **defined as extra-physiologic hip motion that causes pain with or without the symptom of hip joint unsteadiness**. The cause can be traumatic or atraumatic, and is related to both bony and soft tissue abnormality. Gross instability caused by trauma or iatrogenic injury has been shown to improve with surgical correction of the underlying deficiency. **Subtle microinstability, particularly from microtraumatic or atraumatic causes, is an evolving concept** with early surgical treatment results that are promising."	
Cerezal 2012 (18)	Emerging topics on the hip: ligamentum teres and hip microinstability	"Hip microinstability is **the inability to keep the femoral head centered within the acetabular fossa**, without complete luxation or marked subluxation of the joint. Hip laxity is not equivalent to microinstability. The difference is the presence of symptoms associated with laxity when we classified as microinstability. **Only when symptoms are present in the context of laxity can it be classified as microinstability.** An asymptomatic patient that is able to subluxate a joint has laxity, but not microinstability. Patients with microinstability often have laxity in both hips; only the symptomatic is classified as having microinstability."	
Kalisvaart 2015 (8)	Microinstability of the hip – it does exist: etiology, diagnosis and treatment	"Hip instability is generally **defined as extraphysiologic hip motion that causes pain with or without symptoms of hip joint unsteadiness**" "Symptomatic hip microinstability, however, has not received as much attention [as dislocation and traumatic subluxation], as it is more poorly defined, has a less dramatic clinical presentation, lacks consistent objective evaluative criteria, and it has only recently emerged as a significant cause of pain and disability in younger patients and athletes."	Shu 2011 (17)
Suter 2015 (19)	MR findings associated with positive distraction of the hip joint achieved by axial traction	"Atraumatic instability of the hip, also known as microinstability, is **defined by two elements. The first element is laxity of the hip joint with the inability to keep the femoral head centered within the acetabular fossa**, typically without complete luxation or marked subluxation of the joint. **The second element is the presence of symptoms, such as pain or unsteadiness**."	Cerezal 2012 (18), Shu 2011 (17)
Bolia 2016 (20)	Microinstability of the hip: a previously unrecognized pathology	"Unlike other joints in the anatomy, hip instability is generally defined **as extra-physiologic hip motion that causes pain with or without symptoms of hip joint instability**. This entity is not well defined, as no objective criteria has been proposed to characterise hip microinstability."	Shu 2011 (17), Kalisvaart 2015 (8)

(*Continued*)

TABLE 1 - (*Continued*)

First author Year	Title	Definition of microinstability	Citation of definition
Dangin 2016 (1)	Microinstability of the hip: a review	"It is generally defined as a **painful supra-physiological mobility of the hip, associating architectural and functional abnormalities that impair hip stability**. **Microinstability is distinguished from hyperlaxity by its painful nature**, and from traumatic (macro-) instability by its progressive onset and chronicisation following repeated microtrauma concerning at-risk patients. The typical patient is a young female adult with sports activity requiring suppleness and extensive ranges of motion, such as dancing or gymnastics. Microinstability is difficult to identify and thus probably underestimated." "Microinstability is represented by excessive femoral head movement within the acetabulum."	Jackson 2016 (21), Domb 2013 (22), Kalisvaart 2015 (8), Cerezal 2012 (18)
Harris 2016 (23)	Microinstability of the hip and the splits radiograph	"The spectrum of hip instability ranges from subtle microinstability to traumatic dislocation. Microinstability may be either a cause or an effect of several other hip pathologies." "Dance, gymnastics, figure skating, yoga, and cheerleading are among the sports and activities that may predispose to **microinstability (symptomatic)** over simple **hyperlaxity** or **hypermobility (asymptomatic)**."	
d'Hemecourt 2019 (24)	Can dynamic ultrasonography of the hip reliably assess anterior femoral head translation?	"Hip microinstability is **defined as painful supraphysiological mobility of the hip with associated architectural and functional abnormalities that impair joint stability.**"	Bolia 2016 (20), Dangin 2016 (1), Kalisvaart 2015 (8), Jackson 2016 (21)
Harris 2019 (25)	Hypermobile hip syndrome	"Hypermobile hip syndrome may be **defined as a triad of symptoms (patient's unwanted or undesired subjective complaints), signs (physical examination abnormalities with excessive motion that provoke the inciting symptoms), and imaging findings (plain radiographs, magnetic resonance imaging [MRI], computed tomography [CT], or ultrasound) consistent with instability**. A patient with hypermobile hip syndrome may exhibit a constellation of symptom severity, from microinstability to frank dislocation." "The **key distinction between laxity and instability is the absence (former) or presence (latter) of symptoms**. Thus, 'microinstability', by definition, mandates the presence of symptoms."	Harris 2016 (23) Bellabarba 1998 (26)
Safran 2019 (6)	Microinstability of the hip – gaining acceptance	"Microinstability of the hip is **defined as extraphysiologic hip motion that causes pain with or without symptoms of hip joint unsteadiness** and may be the result of bony deficiency and/or soft-tissue damage or loss."	Shu 2011 (17)
Mascarenhas 2020 (27)	Hip, pelvis and sacro-iliac joints	"The concept of microinstability is based on **symptomatic hip laxity without marked subluxation**. Aetiology may be either (1) traumatic (single or repetitive trauma) or (2) atraumatic (generalised laxity or developmental dysplasia of the hip (DDH)). Patients may feel hip unsteadiness, snapping, and/or pain during sports. Diagnosis is problematic, due to no established criteria."	Cerezal 2012 (18)
Parvaresh 2021 (28)	Hip instability in the athlete: anatomy, etiology, and management	"The concept of hip microinstability emerged more recently as **a clinical entity characterised by extraphysiologic motion resulting in hip pain or dysfunction with or without gross symptomatic instability**. A diagnosis of instability may be challenging, because there are no objective criteria that are universally accepted for microinstability."	Bolia 2016 (20), Safran 2019 (6), Kalisvaart 2015 (8)
Vera 2021 (29)	Hip instability in ballet dancers: a narrative review	"The difference between laxity and instability is the absence or presence of symptoms, respectively. Hip instability may present across a diverse spectrum from microinstability to frank dislocation. Thus, 'microinstability,' by definition, **mandates the presence of symptoms.**" "The nebulous term 'microinstability' may be **better termed 'the hypermobile hip syndrome'.** Hypermobile hip syndrome may be **defined as a triad of symptoms (unwanted or undesired subjective complaints), signs (physical examination abnormalities with excessive motion that provoke the inciting symptoms), and imaging findings (plain radiographs, magnetic resonance imaging [MRI], computed tomography [CT], or ultrasound) consistent with instability.**"	Mitchell 2016 (30), Harris 2016 (23), Harris 2015 (31), Kalisvaart 2015 (8)

(*Continued*)

TABLE 1 - (*Continued*)

First author Year	Title	Definition of microinstability	Citation of definition
Rosinsky 2022 (32)	Editorial commentary: hip joint laxity, microinstability, or instability require precise definition: no matter what you call it, it's here to stay!	"The more common, and often interchangeable terms, are instability and microinstability … 'instability' has the advantage of conveying the significant impact the condition has on a patient's life. On the other hand, 'microinstability' may more accurately reflect the vague clinical presentation that we often encounter in the average hip patient with instability. Most patients do not complain of symptoms commonly seen in other joints with 'instability', complaints such as giving way, subluxation, and recurrent dislocations. In the hip, the symptoms are generally less tangible, and hence, the term 'microinstability' may be more appropriate."	Kalisvaart 2015 (8)
Martin 2022 (33)	Pre- and intraoperative decision-making challenges in hip arthroscopy for femoroacetabular impingement	"Hip instability or microinstability, **defined as extraphysiologic hip movement causing pain**, is now widely recognised as a **cause of morbidity and dysfunction, particularly in young patients and athletes, and can co-exist in patients with FAI**."	Kalisvaart 2015 (8), Shu 2011 (17)
Wong 2022 (34)	Physical examination of the hip: assessment of femoroacetabular impingement, labral pathology, and microinstability	"Microinstability of the hip is **defined as supraphysiologic hip motion that causes pain or discomfort with or without subjective unsteadiness of the joint**, and it is believed to be caused by soft tissue injury or loss and/or bony deficiency related to developmental dysplasia of the hip, connective tissue disorders, trauma, idiopathic causes, and iatrogenic causes."	Kalisvaart 2015 (8)

FAI = femoroacetabular impingement; MR = magnetic resonance.

Diagnosis

Diagnosis of microinstability is rather straightforward if significant bony abnormalities or underlying connective tissue disorders are present that can explain the instability (8). However, diagnosis can be much more challenging in the case of idiopathic microinstability (6,8,35). There is no imaging modality, diagnostic or physical test alone that can be used to make a definitive diagnosis (6,8,19,35). The diagnosis is more a pattern recognition of several clinical, radiological, and intraoperative signs. However, an international expert panel has developed a diagnostic tool comprising 34 criteria, which are categorized into "history", "examination", and "imaging" and hold diagnostic value (36).

Patients' history. Patients' history may provide helpful information for the diagnosis of microinstability (37). Patients may describe a painful pop (20,26,38), feeling of instability (1,38-40), pain, "hip giving away", apprehension, snapping, clicking and catching, with or without hip impingement symptoms (6,8,20,23,25,28,29,36,38,40-44). The main symptom is mild to severe hip or groin pain and instability, with the typical "C-sign" pain location by making a "C" with the thumb and hand and placing it at the front and side of the hip (6,8,23,25,29,41,45) or pain located in the inguinal fold (1,46). Patients report activity-related pain, especially after end of range motion (29,44). Pain onset is either atraumatic and progressive, or after an acute trauma (44). Sports or other activities can sustain the symptoms and lead to persistent, constant pain (1,8,23,25). After subluxation or dislocation, the capsuloligamentous structures may lose their stabilizing function and lead to microinstability (23).

People with connective tissue disorders (e.g. Ehlers-Danlos, Down syndrome) are at greater risk of microinstability (8,20,23,25,29,36). Clinicians need to ask about previous injuries and especially previous surgery (e.g. for dysplasia, cam osteoplasty, pincer resection, capsulotomy, labral tear, ligamentum teres tear, iliopsoas surgery), as there is possibility of iatrogenic instability (8,25,36). The probability of microinstability is higher in females (36,47) and in athletes involved in sports that require a large range of motion (ROM), such as gymnastics, dancing or martial arts (25). No study has been done to see what outcomes measure is best to identify hip microinstability. The two validated outcomes measures for non-arthritic hip pain in active patients are the International Hip Outcome Tool (iHOT) (48) and the Copenhagen Hip and Groin Outcome Score (HAGOS) (49).

Clinical examination. Sixteen studies were found reporting the diagnostic accuracy of different tests and radiological signs. Seven studies evaluated seven clinical tests (5,39,50-54) and eight studies reported on six radiological signs (2,3,55-60). One study presented an intraoperative pull test (7). Figure 3 depicts an overview of diagnostic accuracy.

People with hip microinstability show a higher prevalence of the following signs and symptoms: generalized ligamentous laxity (Beighton's Physical Examination Criteria) (1,6,8,20,23,25,26,28,29,34,36,42-44,61), antalgic or abnormal gait patterns, or Trendelenburg sign (23,25,26,29,38,42).

Other tests to diagnose hip microinstability, such as increased ranges of motion (often increased rotation) (34,53,54,61), the log roll test (external rotation recoil or hip dial test) (1,6,8,19,20,23,25,28,29,34,36,38,39,43,44,50,62), easy distraction of the joint (with apprehension) (6,19,20,25,28,

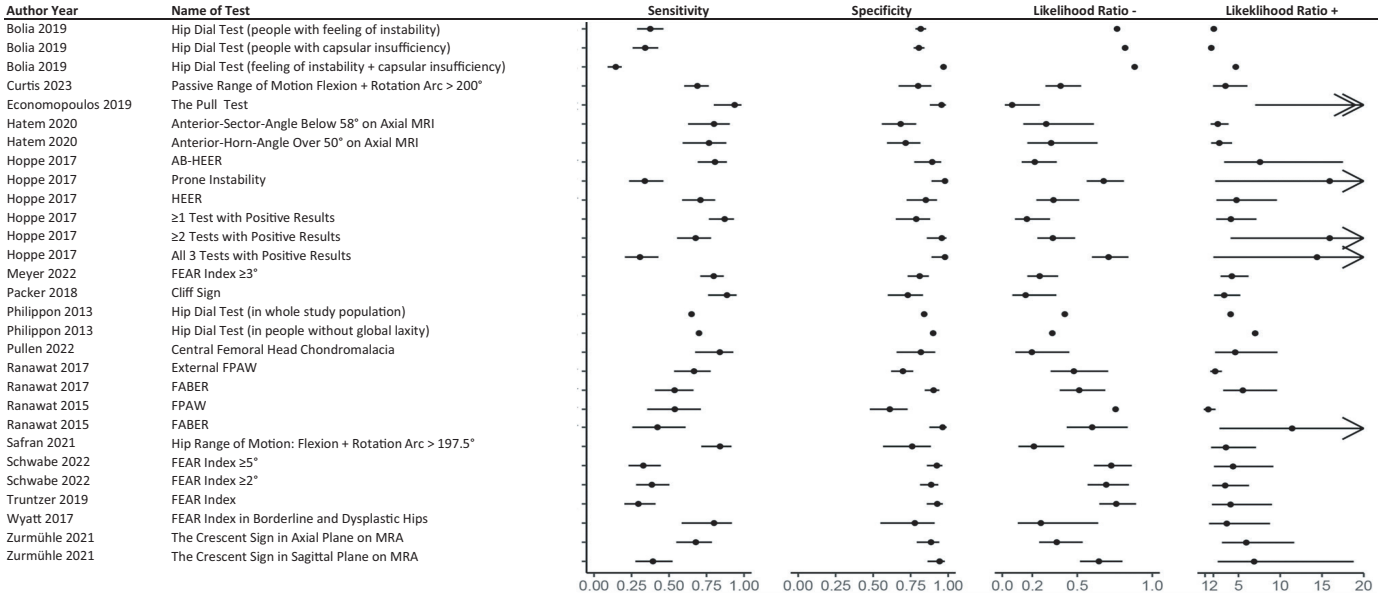


FIGURE 3 - Forest plots of diagnostic accuracy for clinical tests and radiological signs to detect hip microinstability. AB-HEER = abduction hyperextension external rotation; AI = acetabular index; AP = anterior posterior; CT = computed tomography; DDH = developmental dysplasia of the hip; FABER = flexion abduction external rotation; FEAR = femoro-epiphyseal acetabular roof; FPAW = foot progression angle walking; HEER = hyperextension external rotation; LCEA = lateral centre-edge angle; MRA = magnetic resonance arthrography; MRI = magnetic resonance imaging.

29,34,38,44,63), internal rotation with over pressure (IROP) (63), pain with flexion, adduction, and internal rotation (FADIR), and posterior apprehension test (FADIR plus posterior force applied) have also been proposed (34,44). However, they usually lack the specificity to rule in microinstability (1,6,8,26,29). However, the hip dial test seems to be highly specific for the diagnosis of anterior capsular insufficiency in patients with FAI syndrome reporting a feeling of instability (39). The ROM threshold of hip flexion + rotation arc of ≥200° may help identify microinstability (53,54) (Fig. 3 and Supplementary Tab. A2).

Additional special tests that have been described are: the posterior impingement test (hyperextension-external rotation [HEER], anterior apprehension) (1,5,6,8,19,20,28, 29,34,36,38,42-44,47,61), FABER (flexion, abduction, and external rotation test with increased amount of external rotation compared with the unaffected side) (29,51,52,64), FPAW (foot progression angle walking test) (51,52), hip pivot shift (25), and the posterior relocation test (65).

The abduction-hyperextension-external rotation (AB-HEER), the prone external rotation (5,6,8,34,37,44), and the HEER (anterior apprehension) test are well studied (5,33,44) and show small to substantial shifts in probability of having hip instability, especially when combined (Fig. 3 and Supplementary Tab. A2) (5,66,67). The Prone Apprehension Relocation Test (PART) is proposed to diagnose an anterior acetabular undercoverage, which may lead to anterior instability (68). Interrater reliability has been shown to be excellent (kappa 0.81, 95% confidence interval [CI] 0.69-0.93) (69). But this test was not validated against a gold standard that confirms hip instability.

Altered muscle activation patterns are typical in patients with microinstability of the hip (37,41). A useful clinical sign for hip microinstability is a reactive spasm of the secondary stabilizing muscles, such as the iliopsoas or the iliotibial (IT) band, which may be tender on palpation (43). Weakness may be present in those muscles, as well as in the abductor muscles (20,25,29,37). The strength of the core, pelvic, hip, and lower extremity should be assessed (23).

Imaging

Radiography: X-ray/computed tomography (CT). Proposed X-ray views are anteroposterior pelvis view, standing false-profile view, supine Dunn (45°, 90°) or frog-leg lateral view, and hip splits view (8,23,29).

Subluxation is observed with manual traction on an anterior-posterior (AP) traction view (6,26,30,36,38) or in a splits position (6,20,30,62). Subluxation is influenced by any dysplastic changes, larger alpha angle, and smaller femoral neck shaft angle (20,30).

Radiographic images are to be screened for significant acetabular and femoral abnormalities, such as dysplasia and FAI (decreased centre-edge angle [CEA] or lateral centre-edge angle [LCEA] <20-25°, Tönnis angle, acetabular inclination [AI] >13°, aspherical femoral head, higher alpha angle, coxa valga, coxa vara, anteversion of the femoral neck, retroversion of the acetabulum), all of which are regarded as risk factors for microinstability (1,6,8,19,27,29,33,36,38,42,43,47, 59,62). Cam and pincer morphologies would create a levering effect and posterior translation (19,25). The hypermobile hip

crevasse and anterior vertical chondro-osseous lesions can be observed (25), as well as a broken Shenton line, a positive crescent sign, and a distal femoral neck sclerosis (6,57,62,70).

Three variables were associated with instability in borderline hip dysplasia (LCEA 20-25°) (71): AI, anterior centre-edge angle (ACEA), and maximum alpha angle. Odds ratio estimates and 95% CI limits were 1.50 (1.28-1.76), 0.92 (0.86-0.99), and 0.94 (0.90-0.98), respectively.

Several imaging markers are signs of hip instability and should be used in the context of each other: borderline acetabular dysplasia, increased femoral anteversion (>15°), a laterally oriented femoro-epiphyseal acetabular roof (FEAR) index, and anterior wall deficiency (2,6,33,36,44,45,56,58-60,62,72-75). They predicted worse outcomes (iHOT12) of hip femoral osteoplasty with or without labral repair for FAI in female patients (72). However, the optimal cut-off for the FEAR index remains to be established (60,74). A vacuum sign and a femoral head cliff sign are also described as diagnostic tools for instability (3,6,28,33,36,44,62).

A new score was developed for the prediction of instability in people with borderline dysplastic hips (BDH) (71): The Borderline Hip Instability Score (BHIS), considering four radiological and clinical signs (AI, ACEA, maximum alpha angle, and internal rotation in 90 degrees of flexion), demonstrated excellent predictive (discriminatory) ability with an area under the receiver operating characteristic (ROC) curve of 0.89 in the study population. In a population for external validation, the BHIS maintained an excellent area under the ROC curve of 0.92.

A FEAR index ≥4° is able to detect patients at risk of failure of arthroscopy for cam impingement combined with mild to moderate hip dysplasia, with 96% specificity (76).

People with FAI syndrome show a hip translation between neutral and FABER positions in CT images of a mean of 0.84 mm, mainly in the posterior inferior medial direction (77). Femoral anteversion must be considered; if there is more than 10-25° of femoral anteversion, FAI may arise, which is an additional factor for instability (1).

Magnetic resonance imaging (MRI)/Magnetic resonance arthrography (MRA). MRI or MRA helps to identify an increased femoral anteversion (44) or hip capsule laxity (6,20,28,44). Muscle problems, iliopsoas and IT band tendonitis, labral tears, and chondral or ligamentum lesions can be observed (1,8,28,33,37,43,44,78). A traction view can demonstrate the vacuum sign, indicating abnormal distraction (8), with larger or easier widening of the hip joint during traction, suggesting hip laxity (27). Patients suspected to have hip microinstability may also have a thickened iliofemoral ligament with irregularities on the undersurface of the anterior capsule, and an increased capsular volume (27). Other findings associated with positive joint distraction were higher alpha angle, higher neck-shaft angle, smaller acetabular depth, and hypertrophy of the ligamentum teres (19,27,78). Widening of the anterior joint recess (>5 mm) and thinning of the anterior capsular (<3 mm), as well as accumulation of contrast in the posterior-inferior joint in ≥2 planes (6,33) can be seen. Increased intracapsular volume and anterosuperior capsular changes were found in iatrogenic instability after arthroscopy (79).

Dynamic ultrasound. Dynamic ultrasound showed excellent intra- and inter-rater reliability to measure anterior femoral head translation in participants with no hip pathology or functional limitation (intra-rater Intraclass Correlation Coefficients [ICCs] from 0.794 to 0.945, inter-rater ICCs from 0.725 to 0.846) (24). However, in order to achieve good clinical results and outcomes it is important to clarify whether that technique is truly valid for symptomatic patients with microinstability, and whether the magnitude of instability can be precisely measured and integrated into a treatment algorithm (80). In patients with hip pain and clinical suspicion of either instability or impingement, the inter-rater reliability to measure anterior femoral head translation of ≥2 mm provoked by either the figure of 4 or AB-HEER manoeuvres for the diagnosis of microinstability was substantial (kappa 0.606, 95%CI 0.221-0.991) (81).

Intraoperative testing. Anterior capsular insufficiency is shown in patients with FAI without generalized laxity or dysplasia (39). Widening of the anterior joint recess (>5 mm) and anterior capsular thinning (<3 mm) lateral to the zona orbicularis are associated with capsular laxity (6,64,82). However, the correlation of anterior joint recess width (>5 mm) with hip laxity is not yet proven (82). Hip laxity can be confirmed with displacement of the hip with minimal amount of traction force (6,7,33,36,40,64,82-85) or if there is no hip reduction after release of negative intra-articular pressure and traction prior to the start of hip arthroscopy (82,83). Labrum separation, chondral damage, and ligamentum teres tears or hypertrophy can be seen intraoperatively (6,85), with typical inside-out chondral wear of the acetabulum and central femoral head wear (33,85). In patients with FAI who have labral hypertrophy, the hypertrophy is a significant clinical indicator of subtle hip dysplasia and hip microinstability; hence there can be an overlap of FAI and dysplasia characteristics (86).

One study (87) proposed the "Divot" sign as a useful arthroscopic sign of hip microinstability. Of 690 cases of primary hip arthroscopy, 14 hips had a "Divot" sign, and all had risk factors for hip microinstability.

Miscellaneous. To complete the physical examination, before any further investigations are made, an intra-articular hip injection of local anaesthetics can help to confirm a diagnosis of intra-articular pathology (8,44). However, no differentiation can be made between hip microinstability and FAI. In both conditions synovial inflammation has been found (88). Despite this, synovitis scores were lower in the hip microinstability group compared with the FAI group, which also had cartilage damage (88). The presence of synovial inflammation in both groups supports an inflammatory component in the pathogenesis of non-arthritic hip pathology (88).

In patients with femoral head chondromalacia undergoing hip arthroscopy for FAI and/or instability, central head chondromalacia was associated with 84% sensitivity, 82% specificity, 81% positive predictive value, and 84% negative predictive value for a diagnosis of microinstability (89). Hip microinstability was defined as patients with symptoms of intra-articular hip pain with concomitant intraoperative laxity of the symptomatic joint.

As hip microinstability leads to an excessive translation of the femur in the acetabulum, changes in the dynamic loading

of the hip can be observed (41). The magnitude of acceleration during gate cycle shows that the axial, anteroposterior, and mediolateral accelerations differ significantly in people with hip instability compared with healthy asymptomatic controls (41). The axial and mediolateral acceleration values were higher, and the anteroposterior acceleration was lower in the microinstability group compared with the FAI group.

Prevalence

Eleven articles presented or enabled calculation of the prevalence of microinstability with or without signs of hip impingement (2,3,5,7,51,52,56,59,60,90,91) (Supplementary Tab. A3). The prevalence of microinstability with FAI was in the range of 21%-42%, in adults undergoing hip arthroscopy or MRA of the hip. The prevalence of instability without FAI was in the range 27% (in patients with unilateral hip or groin pain) to 57% (in patients with suspicion of microinstability who underwent hip arthroscopy), except in a sample with borderline acetabular dysplasia, where it was 62.9%. Population size range was 39-953 hips investigated. All subjects were people with hip pain or who had undergone hip arthroscopy.

Aetiology

A total of 49 articles reported on contributing or risk factors for the development of hip microinstability. In general, it is stated that hip microinstability is a multifactorial disorder. It can be a cause for, or a consequence of, multiple pathological conditions of the hip. These may be osseous, chondrolabral, capsuloligamentous, musculotendinous, or neuromuscular dysfunctions of the kinetic chain (23,25,29).

Aetiologies are classified into six categories: (i) significant bony abnormalities, such as developmental dysplasia of the hip (DDH), (ii) connective tissue disorders, (iii) post-traumatic, (iv) microtraumatic, (v) iatrogenic, and (vi) idiopathic (6,8,28). In the absence of significant bony abnormalities, the pathology originates primarily in the supporting soft tissue (26).

FAI may induce instability in the following four ways (23): (1) excessive acetabular anteversion, resulting in posterior acetabular rim impingement and anterior hip instability; (2) excessive acetabular retroversion, resulting in anterior impingement and posterior instability; (3) excessive femoral anteversion, resulting in posterior acetabular rim impingement and anterior hip instability; and (4) excessive femoral retroversion, resulting in anterior impingement and posterior instability. The combination of borderline dysplasia and FAI with an increased femoral anteversion leads to worse instability in extension (92).

FAI may lead to instability and, vice versa, the excessive femoral head translation relative to the acetabulum may contribute to the FAI pathomechanism, with a potential mechanical overloading of the hip structures, leading to pain (67,93), central femoral head wear, and subluxation (89). In addition, an increasingly frequent indication for revision arthroscopy for FAI is capsular complication and subsequent hip instability (61,94,95).

Table 2 gives an overview of all mentioned risk factors or contributing factors for hip microinstability.

TABLE 2 - Risk factors or contributing factors for hip microinstability

Risk factors or contributing factors for hip microinstability	Details
Overuse	Microtrauma caused by repetitive axial loading (with external rotation or abduction) with motion to or beyond the limits, such as in hockey, golf, football, ballet or gymnastics, leads to repeated injury or elongation of the capsule and to labral tears (8,18-20,22,23,28,33,37,38,42,96). This increases the forces on the other static stabilizers. Injury of the ilio-, pubo- and ischiofemoral ligament, and the ligamentum teres may contribute to microinstability (8,20,29,37,38,42,44,97). Disruption of the soft tissue affects the stability because of loss of coupling force (98). Labral tears may induce loss of suction seal effect and worsen instability through subluxation (42,44,63). In addition, the labrum has a nociceptive and a proprioceptive function. When injured, the altered sensory information may affect joint stability (20). The labrum is constantly stressed in the dancer's hip and the hip capsule is frequently thinner (29). With a torn labrum and a thin capsule, the hip may show instability (29,33).
FAI	Cam or pincer morphology can also induce microinstability, by excessive acetabular anteversion (advanced posterior bone contact and anterior instability), acetabular retroversion (advanced anterior bone contact and posterior instability), excessive femoral anteversion (posterior cam effect and anterior instability), or excessive femoral retroversion (anterior cam effect and posterior instability), and thus increase the risk of subluxation (1,23,63,67,92,98,99).

(*Continued*)

TABLE 2 - (*Continued*)

Risk factors or contributing factors for hip microinstability	Details
	The osseous impingement at end of motion positions may lead to levering of the femoral head out of the socket (8,30) and to posterior chondral and capsular-labral junction injury ("contrecoup" injury), and therefore, to secondary subluxation or dislocation (1,6,17,19,28,30,99-101), especially in the athlete with FAI, where the functional ROM required is often greater than the limited physiological motion allowed by the cam and/or rim impingement lesions (23,99,101). In end of range movements, a FAI occurs that leads to subluxation, even without pincer or cam morphology, seen in ballet dancers (29,30,102). Women show greater subluxation than men during the "grand écart facial" position, with increasing subluxation with larger alpha angles and smaller neck-shaft angles (30). Increased flexion and internal rotation may lead to impingement between the cam and the anterior acetabulum and levering of the femoral head posteriorly (23,99), with posterior acetabular rim fracture and posterior capsulolabral tear, analogous to a posterior bony Bankart lesion of the shoulder (19,101). In addition, the repetitive abutment of the femur head-neck junction against the acetabulum may lead to trauma of the anterior labrum and stretch of the capsule and capsular ligaments (6). This increases the movement of the femoral head and may result in subluxation (38,103). Furthermore, there is risk of a primary anteroinferior impingement through abutment of the prominence of the medial femoral metaphysis and/or anteroinferior border of the acetabulum in extension and internal rotation (104). A posterior extra-articular ischiofemoral impingement can cause secondary anterior instability of the femur in extension (104). These patients show anteroinferior abrasion of the cartilage with rupture and degeneration of the labrum, similar to a posteroinferior contrecoup lesion that can be seen with anterior pincer impingement (104). In extreme end of range motions, for example, in ballet dancers, an insufficient femoral version leads to a posterior impingement of the femoral neck on the acetabulum that results in anterior subluxation (29). FAI causes migration of the femoral head, thus the relation of the head and the acetabulum alters. This increases shear forces and leads to microinstability (105).
Hip arthroscopy	Hip arthroscopy may lead to microinstability (20). Excessive resection of the acetabular bone while managing pincer morphology may induce subluxation or migration of the femoral head out of the acetabulum (23,28,97). Also, overcorrection of cam morphologies can lead to instability (23,28,98,106,107). Overzealous capsulotomy without repair after arthroscopy for FAI or capsulectomy can result in iatrogenic instability (6,8,20,22,23,28,61,94,95,98,100,108-111). Excessive labrum or ligamentum teres resection or psoas tenotomy may also be an iatrogenic cause (1,23,28). In general, previous arthroscopy may lead to increased distractibility of the hip joint compared with the native hip (84).
PAO	Acetabular retroversion and high to normal femoral version treated with anteverting PAO can lead to anterior instability (92).
Special osseous morphologies	Lack of acetabular coverage/dysplasia or borderline dysplasia may lead to atraumatic instability (1,6,8,19,28,29, 33,38,42,44,63,112). An increased femoral anteversion and a coxa valga will contribute to further instability, even more in case of a borderline hip, while a decreased femoral version would contribute to increasing impingement (1,33,113). Extra-articular bone impingement, especially between the greater trochanter and pelvis (1). A coxa vara demonstrates ischiofemoral/greater trochanter impingement, particularly with abduction/side splits in ballet dancers, with subluxation of the femoral head (29).
Ligamentous laxity, soft-tissue disorders, capsular laxity/thin capsule	Soft-tissue disorders (e.g. Ehlers-Danlos syndrome), ligamentous and capsular laxity, or a thin capsule may result in atraumatic instability (1,8,19,23,26,28,29,33,38,42,44,114). Abnormal joint forces are the result of capsular laxity that may lead to labral injuries and femoral neck impingement at high flexion "secondary impingement" (18).

(*Continued*)

TABLE 2 - (*Continued*)

Risk factors or contributing factors for hip microinstability	Details
Legg-Calvé-Perthes disease	Legg-Calvé-Perthes disease leads to significant impingement that develops secondary dysplasia and thus instability (38).
Increased translational motion in the hip joint	Instability leads to an increased movement of the femoral head in the hip joint potentially causing cartilage wear, degenerative changes, and capsular stress. It also places the labrum at risk of shear injury and microtrauma, further contributing to pathological articular changes (45,97).
Deep hip muscle weakness	Weakness of deep hip muscles results in instability and overactivation of secondary movers. This may result in an anterior gliding of the femoral head and exaggerate anterior joint loading (44,93). FAI morphologies may enhance this loading and result in labrum alterations (93).
Ligamentum teres tears	There is a possible interrelationship between FAI, labral tears, and ligamentum teres pathology (115,116). Trauma, overuse at end of range motion, FAI, and other osseous risk factors for instability, such as borderline dysplasia, may result in ligamentum teres injury (112,115). Ligamentum teres tears contribute to microinstability and damage of the labrum and the cartilage with sporting activities (18,117,118). People with complete tears are more likely to exhibit capsular laxity (115). In patients with chondrolabral dysfunctions associated with FAI, approximately 90% had a partially or completely torn ligamentum teres and they were 3.6 times more likely to have capsular laxity (116). Thus, torn ligamentum teres may lead to microinstability (116). Of 20 subjects with complete ligamentum teres ruptures all had labral pathology and evidence of FAI, with 19 cam and 1 pincer. Of these, five out of nine subjects contacted for follow-up noted instability (117).

FAI, femoroacetabular impingement; PAO = periacetabular osteotomy; ROM, range of motion.

Treatment

There is consensus that the first-line treatment for hip microinstability is conservative management based on modifiable factors. Strengthening, sensorimotor training, activity modification, and education are proposed. In addition, non-steroidal anti-inflammatory drugs (NSAIDs) or corticosteroid injections can be used. There is a lack of RCTs evaluating the effectiveness of different treatment modalities in patients with hip microinstability. Surgical management is indicated if conservative treatment of 3-6 months fails and symptoms last for at least 6 months (6,8,25,28,33,43,119,120). Surgical procedures are performed either by arthroscopy or open surgery. They target redirectional osteotomies, capsular and labral management, and address intra-articular bony pathology with acetabuloplasty for pincer and femoral osteoplasty for cam morphology. It is essential to determine why the hip is unstable before considering surgery of the capsule, bones, or soft tissue. Additionally, intraoperative hip testing and re-testing can help uncover additional causes of impingement or instability once the primary causes are addressed. Table 3 gives an overview of treatment options for hip microinstability.

A retrospective case series study showed that two-thirds of patients with microinstability were able to avoid surgery and had improved clinical outcome scores after hip and core strengthening exercises two times a week for 6 weeks plus home exercises (121).

Different non-RCT intervention studies showed clinical improvement after surgical intervention. In a pre-post cohort study of 25 patients without dysplasia undergoing periacetabular osteotomy (PAO) because of hip pain and/or instability after failure of arthroscopy (128), 18 patients (72%) reached the minimal clinically important improvement in the modified Harris hip score (mHHS) and in the iHOT-33 at 6 months follow-up. Favourable and significant pre-post improvements were also shown for patient-reported outcomes (mHHS, visual analogue scale [VAS] for pain, Hip Outcome Score – Sport-Specific Score, Non-Arthritic Hip Score) in 65 high-level athletes after primary arthroscopy in the setting of borderline dysplasia and hip microinstability (131). In addition, high rates of return to sports were achieved (80.7%). In 140 patients undergoing hip arthroscopy for FAIS with a standard post-operative rehabilitation protocol, 19 patients had hip instability (FEAR index ≥2°), whereas 121 patients did not (FEAR index <2°). Both groups had similar improvement in 2-year outcomes (132). Another 32 females with atraumatic microinstability, with anterior labral and cartilage pathology, were treated with arthroscopy and capsular plication without any bony resections (133). There was significant clinical pre-post improvement in pain and function. However, in a retrospective case series study of 27 hips with microinstability treated with combined arthroscopy and open capsular plication in the absence of acetabular dysplasia or severe femoral anteversion, 45% had reoperation (arthroscopy, femoral osteotomy, or PAO) and persistent symptoms (129).

Poor surgical prognostic factors for patients with dysplastic hip microinstability are a broken Shenton's line, a femoral

TABLE 3 - Treatment options for hip microinstability

Treatment options	In detail	References
Conservative management		
Strengthening	Iliopsoas, hip abductors, adductors, external rotators, gluteus maximus, core muscles, low back, iliocapsularis, rectus femoris, TFL, hamstrings	(1,6,8,20,23,29,33,43-45,93,121)
Stretching	Iliotibial band, hamstrings, rectus femoris, abdominal muscles	(1,23)
Sensorimotor training	Neuromuscular rehabilitation to address functional deficits	(28,29,93)
Activity modification	Education, relative rest, activity modification (avoidance of provocative manoeuvres), adaptive sport activities	(1,6,8,20,25,28,29,33,38,43,44)
Medication/injection	Non-steroidal anti-inflammatory drugs, oral non-opioids, corticosteroid injections in conjunction with local anaesthetic	(1,6,8,18,20,23,25,28,29,37,38,43,44)
Physical therapy	Multimodal rehabilitation exercises	(18,23,25,28,37,38,120,122)
Surgical management		
Capsular management	Capsular closure after arthroscopy, repair, suture, reduction of capsular volume by plication or capsulorrhaphy, as well as capsular management in the setting of revision arthroscopy	(1,6,8,18,20,23,25,29,33,37,38,42-46, 61,96,97,120,123-126)
Iliofemoral ligament	Ligament repair, length restoration	(42,97,123)
Ligamentum teres	Reconstruction	(1,23,29,75,120)
Labrum	Labral repair, refixation, reconstruction, debridement, graft (to recreate suction seal effect)	(1,6,8,20,25,33,37,38,44,127)
Osteotomy, osteoplasty	Periacetabular osteotomy, femoral osteotomy, acetabuloplasty, femoral osteoplasty, address FAI morphologies	(1,6,8,20,23,25,28,29,33,37,43,75,92,128-130)

FAI = femoroacetabular impingement; TFL = tensor fasciae latae.

neck-shaft angle >140°, a lateral CEA <19°, and a body mass index (BMI) >23 kg/m^2 (28).

Two case series studies showed substantial improvement in function for patients after revision surgery with capsular repair, who had iatrogenic hip microinstability after a first arthroscopy (134,135).

Cam and pincer morphology, as well as hip dysplasia may lead to labrum and adjacent acetabular cartilage damage (136,137). In a cohort of 75 patients, 55% failed conservative treatment and needed surgical procedure (118). The best predictor for failure of conservative treatment was a tear of the ligamentum teres (118). The authors claimed that people with a torn ligamentum teres develop subtle hip instability.

Special case borderline dysplasia and FAI. Borderline dysplasia might lead either to instability or to impingement of the hip (113). Decision-making for the optimal surgical treatment in case of borderline dysplasia is extremely difficult (130), especially if there is excessive femoral anteversion (138). No clinical standards exist to decide if there is significant structural instability, or FAI and microinstability, or no instability (130). Instability related to acetabular dysplasia or retroversion would be treated with PAO, while FAI (with or without instability) could be treated with arthroscopy, via capsulotomy during PAO, or with an open surgical hip dislocation (130). Arthroscopy could potentially replace PAO for soft-tissue related instability and FAI in patients with borderline dysplasia (139). Modern PAO, however, is done with additional arthroscopy, such as acetabular resection or femoral head-neck offset decompression, to address bony morphologies leading to FAI (113,130).

Discussion

This scoping review included 123 studies and collected evidence on five main topics: definition, diagnosis, aetiology, prevalence, and treatment of hip microinstability. There are numerous types of evidence reporting on the concept of hip microinstability and its context with FAI. The main findings of this review are described below.

Different definitions for hip microinstability exist. A standardized terminology should be established (80). Supraphysiological motion or excessive motion of the femoral head is mentioned, but the term “hip microinstability” should be used only when the centre of rotation of the femoral head is not stable in the acetabulum, that is, when there is excessive femoral head movement within the acetabulum (1,18). However, there is no objective quantification and cut-off for excessive movement. A classification system should be established to facilitate future clinical studies (32,80).

Diagnosis is a puzzle of history, clinical examination, radiographic and intraoperative signs. An international expert panel published a consensus study for the diagnosis of microinstability (36). They propose a diagnostic tool in a tabular format with 34 criteria deemed to have diagnostic value.

Another international expert consensus conference showed strong agreement on eight operating room criteria to confirm hip microinstability (85). The experts propose using this list as a basis for further research to build a scoring or weighting system for the diagnosis of hip instability. Data relating to the items should be recorded prospectively, so that the relative importance of the items to symptoms and treatment response could be stratified.

Hip microinstability and FAI may be associated, they can occur in combination, and they may aggravate each other (1). Cam and pincer seem to predispose the hip joint to instability by a multifactorial mechanism, consisting of abnormal osseous morphologies, weakened static stabilizers, and dynamic factors (140).

The static and dynamic stabilizers of the hip joint are well described, for example, the role of the capsule or the deep hip muscles. However, there is no data regarding dynamic or static hip instability and its contributors. There are some mechanical factors that may lead to a dynamic instability, such as cam and pincer morphology or femoral retroversion (141). Other mechanical factors, such as hip dysplasia and femoral anteversion, may lead to static instability (141). This topic requires further exploration.

Symptomatic hip microinstability with additional FAI morphologies is present in 21%-42% of adults undergoing hip arthroscopy or MRA of the hip. The main symptom in both conditions is pain. Both conditions require symptoms to be diagnosed, but not all patients with radiographic signs of instability or FAI morphologies are symptomatic (91,142,143). Therefore, it should be investigated why some people develop symptoms and others do not.

Researchers suggest the same conservative treatment strategies for hip microinstability and FAI patients. They conducted several RCTs to show the effectiveness of conservative versus surgical treatment in FAI (10-13). However, they do not explain the treatment propositions in detail, and the frequencies of treatment vary greatly. Consequently, researchers lack a clear description of an appropriate non-operative treatment for hip microinstability and FAI. Casartelli et al (93) proposed active physical therapy aimed at improving hip neuromuscular function. If the passive stability mechanisms are inadequate, the muscular system needs to augment stability (144). To enhance joint stability, deep hip stabilizing muscles should be retrained, following the same rationale as strengthening the local muscles before the more superficial ones at the spine and shoulder (144). There is evidence that the local stabilizing muscles can improve function, reduce pain, and restore normal feedforward response in other joints, such as the knee, the lumbar and the cervical spine (145-147). The dynamic stability of the hip joint needs to be improved and the anterior gliding of the femoral head minimized in people with hip microinstability. Hence, hip flexor and abductor muscles and the deep hip external rotators need to be strengthened (144). Feedforward mechanisms are needed for normal postural activity, and they can be trained by repeated voluntary activation of a muscle (146). Attention and motor imagery are important for improved motor performance and greater transfer to task performance (146).

The surgical treatment of hip microinstability differs considerably between pure instability and pure FAI. Often there is a combination of both problems, especially in the case of BDH. Intra-articular pathologies, such as ligamentum teres tears, pincer or cam morphologies, should be addressed, because, if not treated, they may further create hip instability (75). If the instability part is overseen and surgical intervention is only made to correct the bony impingement, the risk of increasing the instability is high. To differentiate whether a BDH has instability or impingement characteristics they propose using the FEAR index. However, there is no absolute consensus for the cut-off value of the FEAR index (73). Hence, an in-depth analysis of the situation before choosing the surgical intervention is crucial. There is large consensus that the capsule should be repaired after arthroscopic surgery for FAI, to avoid iatrogenic microinstability.

Limitations and strengths

The difficulty of clear diagnostic criteria and definition of hip microinstability may have led to under- or over-inclusion of papers in this review. The scientific rigour of the included studies was not investigated, therefore there is no grading of evidence. Overall, there is a lack of high-quality RCTs for the management of patients with hip microinstability.

A sensitive search was performed, resulting in a large number of papers being included in this study. This allowed a comprehensive overview of the topic and resulted in sensibilization of the association between hip microinstability and FAI.

Conclusions and implications for research and practice

Microinstability of the hip lacks consistent objective evaluative criteria. A standardized terminology should be established. Furthermore, consensus is necessary regarding physical examination, diagnostic criteria, and a classification system of hip microinstability. Only with consistent, quantitative, and valid diagnostic criteria can clinicians and researchers start to examine target populations and build high-quality research projects with clear research questions. Hip microinstability and FAI may be associated; they can occur in combination and may aggravate each other. There is a lack of evidence regarding the feasibility and effectiveness of effective training in reducing symptoms in people with hip microinstability with or without FAI. We need RCTs in this population with targeted training to assess the effectiveness of the interventions under evaluation. Furthermore, we need larger studies on sports performance and long-term outcomes for athletes. Further research is necessary to enable clinicians to confidently manage hip microinstability, also in the context of FAI.

Disclosures

Conflict of interest: The authors have no conflicts of interest to disclose.

Funding and financial support: The authors have no funding and financial support to disclose.

Authors' contributions: CR contributed to the conception of the work, undertook the literature review, prepared figures and tables, authored and reviewed drafts of the article, and approved the final draft. DB authored or reviewed drafts of the article and approved the final draft. JT authored or reviewed drafts of the article and approved the final draft. JMS authored or reviewed drafts of the article and approved the final draft. TR undertook the literature review, prepared figures, reviewed drafts of the article, and approved the final draft. QB undertook the literature review, prepared figures, reviewed drafts of the article, and approved the final draft. RH contributed to the conception of the work, undertook the literature review, prepared figures and tables, authored or reviewed drafts of the article, and approved the final draft.

Data Availability Statement: Data sharing not applicable: No new data was created or analysed in this study and more data is available as supplementary material to this article.

References

1. Dangin A, Tardy N, Wettstein M, May O, Bonin N. Micro-instability of the hip: a review. Orthop Traumatol Surg Res. 2016;102(8S):S301-S309. CrossRef PubMed
2. Truntzer JN, Hoppe DJ, Shapiro LM, Safran MR. Can the FEAR index be used to predict microinstability in patients undergoing hip arthroscopic surgery? Am J Sports Med. 2019;47(13):3158-3165. CrossRef PubMed
3. Packer JD, Cowan JB, Rebolledo BJ, et al. The Cliff Sign: a new radiographic sign of hip instability. Orthop J Sports Med. 2018;6(11):2325967118807176. CrossRef PubMed
4. Reiman MP, Agricola R, Kemp JL, et al. Consensus recommendations on the classification, definition and diagnostic criteria of hip-related pain in young and middle-aged active adults from the International Hip-related Pain Research Network, Zurich 2018. Br J Sports Med. 2020;54(11):631-641. CrossRef PubMed
5. Hoppe DJ, Truntzer JN, Shapiro LM, Abrams GD, Safran MR. Diagnostic accuracy of 3 physical examination tests in the assessment of hip microinstability. Orthop J Sports Med. 2017;5(11):2325967117740121. CrossRef PubMed
6. Safran MR. Microinstability of the hip – gaining acceptance. J Am Acad Orthop Surg. 2019;27(1):12-22. CrossRef PubMed
7. Economopoulos KJ, Kweon CY, Gee AO, Morris ST, Hassebrock JD, Chhabra A. The pull test: a dynamic test to confirm hip microinstability. Arthrosc Sports Med Rehabil. 2019;1(1):e67-e74. CrossRef PubMed
8. Kalisvaart MM, Safran MR. Microinstability of the hip – it does exist: etiology, diagnosis and treatment. J Hip Preserv Surg. 2015;2(2):123-135. CrossRef PubMed
9. Duncan ST, Bogunovic L, Baca G, Schoenecker PL, Clohisy JC. Are there sex-dependent differences in acetabular dysplasia characteristics? Clin Orthop Relat Res. 2015;473(4):1432-1439. CrossRef PubMed
10. Griffin DR, Dickenson EJ, Wall PDH, et al; FASHIoN Study Group. Hip arthroscopy versus best conservative care for the treatment of femoroacetabular impingement syndrome (UK FASHIoN): a multicentre randomised controlled trial. Lancet. 2018;391(10136):2225-2235. CrossRef PubMed
11. Hunter DJ, Eyles J, Murphy NJ, et al. Multi-centre randomised controlled trial comparing arthroscopic hip surgery to physiotherapist-led care for femoroacetabular impingement (FAI) syndrome on hip cartilage metabolism: the Australian FASHIoN trial. BMC Musculoskelet Disord. 2021;22(1):697. CrossRef PubMed
12. 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
13. Palmer AJR, Ayyar Gupta V, Fernquest S, et al; FAIT Study Group. Arthroscopic hip surgery compared with physiotherapy and activity modification for the treatment of symptomatic femoroacetabular impingement: multicentre randomised controlled trial. BMJ. 2019;364:l185. CrossRef PubMed
14. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Implement. 2021;19(1):3-10. CrossRef PubMed
15. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol. 2018;18(1):143. CrossRef PubMed
16. 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
17. Shu B, Safran MR. Hip instability: anatomic and clinical considerations of traumatic and atraumatic instability. Clin Sports Med. 2011;30(2):349-367. CrossRef PubMed
18. Cerezal L, Arnaiz J, Canga A, et al. Emerging topics on the hip: ligamentum teres and hip microinstability. Eur J Radiol. 2012;81(12):3745-3754. CrossRef PubMed
19. Suter A, Dietrich TJ, Maier M, Dora C, Pfirrmann CW. MR findings associated with positive distraction of the hip joint achieved by axial traction. Skeletal Radiol. 2015;44(6):787-795. CrossRef PubMed
20. Bolia I, Chahla J, Locks R, Briggs K, Philippon MJ. Microinstability of the hip: a previously unrecognized pathology. Muscles Ligaments Tendons J. 2016;6(3):354-360. CrossRef PubMed
21. Jackson TJ, Peterson AB, Akeda M, et al. Biomechanical effects of capsular shift in the treatment of hip microinstability: creation and testing of a novel hip instability model. Am J Sports Med. 2016;44(3):689-695. CrossRef PubMed
22. Domb BG, Philippon MJ, Giordano BD. Arthroscopic capsulotomy, capsular repair, and capsular plication of the hip: relation to atraumatic instability. Arthroscopy. 2013;29(1):162-173. CrossRef PubMed
23. Harris JD, Gerrie BJ, Lintner DM, Varner KE, McCulloch PC. Microinstability of the hip and the splits radiograph. Orthopedics. 2016;39(1):e169-e175. CrossRef PubMed
24. d'Hemecourt PA, Sugimoto D, McKee-Proctor M, et al. Can dynamic ultrasonography of the hip reliably assess anterior femoral head translation? Clin Orthop Relat Res. 2019;477(5):1086-1098. CrossRef PubMed
25. Harris JD. Hypermobile hip syndrome. Oper Tech Sports Med. 2019;27(3):108-118. CrossRef
26. Bellabarba C, Sheinkop MB, Kuo KN. Idiopathic hip instability. An unrecognized cause of coxa saltans in the adult. Clin Orthop Relat Res. 1998;(355):261-271. PubMed
27. Mascarenhas VV, Castro MO, Afonso PD. Hip, pelvis and sacro-iliac joints. In: Vanhoenacker FM, Maas M, Gielen JL, eds. Imaging of orthopedic sports injuries. Medical radiology. Springer, Cham 2020; 353-422. CrossRef
28. Parvaresh KC, Rasio J, Azua E, Nho SJ. Hip instability in the athlete: anatomy, etiology, and management. Clin Sports Med. 2021;40(2):289-300. CrossRef PubMed
29. Vera AM, Nho SJ, Mather RC, Wuerz TH, Harris JD. Hip instability in ballet dancers: a narrative review. J Dance Med Sci. 2021;25(3):176-190. CrossRef PubMed

30. Mitchell RJ, Gerrie BJ, McCulloch PC, et al. Radiographic evidence of hip microinstability in elite ballet. Arthroscopy. 2016;32(6):1038-1044.e1. CrossRef PubMed
31. Harris JD, Slikker W, Abrams GD, Nho SJ. Atraumatic instability and surgical technique. In: Nho S, Leunig M, Larson C, Bedi A, Kelly B, eds. Hip arthroscopy and hip joint preservation surgery. New York: Springer 2015; 1001-1014. CrossRef
32. Rosinsky PJ. Editorial commentary: hip joint laxity, microinstability, or instability require precise definition: no matter what you call it, it's here to stay! Arthroscopy. 2022;38(10):2850-2851. CrossRef PubMed
33. Martin H, Robinson PG, Maempel JF, et al. Pre- and intraoperative decision-making challenges in hip arthroscopy for femoroacetabular impingement. Bone Joint J. 2022;104-B(5):532-540. CrossRef PubMed
34. Wong SE, Cogan CJ, Zhang AL. Physical examination of the hip: assessment of femoroacetabular impingement, labral pathology, and microinstability. Curr Rev Musculoskelet Med. 2022; 15(2):38-52. CrossRef PubMed
35. Jean PO, Safran MR, Ayeni OR. Hip microinstability: fact or fiction? Knee Surg Sports Traumatol Arthrosc. 2023;31(1):1-3. CrossRef PubMed
36. Khanduja V, Darby N, O'Donnell J, et al; International Microinstability Expert Panel. Diagnosing hip microinstability: an international consensus study using the Delphi methodology. Knee Surg Sports Traumatol Arthrosc. 2023;31(1):40-49. CrossRef PubMed
37. Shindle MK, Ranawat AS, Kelly BT. Diagnosis and management of traumatic and atraumatic hip instability in the athletic patient. Clin Sports Med. 2006;25(2):309-326, ix-x. ix-x. CrossRef PubMed
38. Boykin RE, Anz AW, Bushnell BD, Kocher MS, Stubbs AJ, Philippon MJ. Hip instability. J Am Acad Orthop Surg. 2011; 19(6):340-349. CrossRef PubMed
39. Bolia I, Briggs KK, Philippon MJ. High specificity of the hip dial test to diagnose anterior capsular insufficiency in patients with FAI-related microinstability of the hip. Arthrosc - J Arthrosc Relat Surg. 2019;35:e29-e. CrossRef
40. Parvaresh KC. Editorial commentary: axial stress examination under anesthesia provides a highly reliable test for measurement of hip distraction. Arthroscopy. 2022;38(5):1478-1479. CrossRef PubMed
41. Neira A, Amenabar T, Cristi-Sánchez I, et al. Evaluation of atraumatic hip instability measured by triaxial accelerometry during walking. J Hip Preserv Surg. 2019;6(2):134-139. CrossRef PubMed
42. Philippon MJ, Zehms CT, Briggs KK, Manchester DJ, Kuppersmith DA. Hip instability in the athlete. Oper Tech Sports Med. 2007; 15(4):189-194. CrossRef
43. Smith MV, Sekiya JK. Hip instability. Sports Med Arthrosc Rev. 2010;18(2):108-112. CrossRef PubMed
44. Curtis DM, Murray IR, Money AJ, Pullen WM, Safran MR. Hip microinstability: Understanding a newly defined hip pathology in young athletes. Arthroscopy. 2022;38(2):211-213. CrossRef PubMed
45. Lall AC, Walsh JP, Maldonado DR, et al. Teamwork in hip preservation: the ISHA 2019 annual scientific meeting. J Hip Preserv Surg. 2020;7(suppl 1):2-21. CrossRef PubMed
46. Cohen D, Jean PO, Patel M, et al. Hip microinstability diagnosis and management: a systematic review. Knee Surg Sports Traumatol Arthrosc. 2023;31(1):16-32. CrossRef PubMed
47. Curtis DM, Pullen WM, Murray IR, Money AJ, Segovia N, Safran MR. Female gender, decreased lateral center edge angle and a positive hyperextension-external rotation test are associated with ease of hip distractability at time of hip arthroscopy. Knee Surg Sports Traumatol Arthrosc. 2022;30(6):2188-2194. CrossRef PubMed
48. Mohtadi NG, Griffin DR, Pedersen ME, et al. The development and validation of a self-administered quality-of-life outcome measure for young, active patients with symptomatic hip disease: the International Hip Outcome Tool (iHOT-33). Arthroscopy. 2012;28(5):595-605; quiz 606-10.e1. CrossRef PubMed
49. Thorborg K, Hölmich P, Christensen R, Petersen J, Roos EM. The Copenhagen Hip and Groin Outcome Score (HAGOS): development and validation according to the COSMIN checklist. Br J Sports Med. 2011;45(6):478-491. CrossRef PubMed
50. Philippon MJ, Briggs KK, Goljan P, Peixoto LP. The hip dial test to diagnose symptomatic hip instability. Arthroscopy. 2013; 29(10):e123. CrossRef
51. Ranawat AS, Gaudiani MA, Slullitel PA, Satalich J, Rebolledo BJ. Foot progression angle walking test: a dynamic diagnostic assessment for femoroacetabular impingement and hip instability. Orthop J Sports Med. 2017;5(1):2325967116679641. CrossRef PubMed
52. Ranawat AS, Park C, Licatesi T, Rebolledo BJ, Satalich J. Foot progression angle walking test – an effective dynamic test for the diagnosis of femoroacetabular impingement and hip instability. Orthop J Sports Med. 2015;3(7_suppl 2):2325967115S00140. CrossRef
53. Curtis DM, Pullen WM, Hopkins JN, et al. Can hip passive range of motion predict hip microinstability? A comparative study. Orthop J Sports Med. 2023;11(6):23259671231169978. CrossRef PubMed
54. Safran MR, Hopkins J, Elisman K, Finlay A, Segovia NA. Can hip microinstability be predicted by hip range of motion. J ISAKOS. 2021;6(6):494.
55. Hatem MA, Helal A, Khoury AN, Martin HD. Anteroinferior hip instability in flexion during dynamic arthroscopic examination is associated with abnormal anterior acetabular horn. Orthop J Sports Med. 2020;8(12):2325967120965564. CrossRef PubMed
56. Wyatt M, Weidner J, Pfluger D, Beck M. The Femoro-Epiphyseal Acetabular Roof (FEAR) index: a new measurement associated with instability in borderline hip dysplasia? Clin Orthop Relat Res. 2017;475(3):861-869. CrossRef PubMed
57. Zurmühle CA, Kuner V, McInnes J, Pfluger DH, Beck M. The crescent sign – a predictor of hip instability in magnetic resonance arthrography. J Hip Preserv Surg. 2021;8(2):164-171. CrossRef PubMed
58. Meyer AM, Schaver AL, Cohen BH, Glass NA, Willey MC, Westermann RW. FEAR index in predicting treatment among patients with femoroacetabular impingement and hip dysplasia and the relationship of femoral version. J Hip Preserv Surg. 2022;9(2):84-89. CrossRef PubMed
59. Schwabe M, Pascual-Garrido C, Clohisy J, Graesser E, Nepple J. Acetabular dysplasia: three-dimensional deformity predictors of the diagnosis of symptomatic instability treated with periacetabular osteotomy. Orthop J Sports Med. 2020;8(7_suppl 6):2325967120S00426. CrossRef
60. Schwabe MT, Clohisy JC, A Graesser E, Pascual-Garrido C, Nepple JJ. External validation of the FEAR index in borderline acetabular dysplasia. Orthop J Sports Med. 2022;10(8): 23259671221113837. CrossRef PubMed
61. Larson CM. Editorial commentary: restoration of hip capsular tension is more important than repair construct configuration. Arthroscopy. 2022;38(12):3149-3151. CrossRef PubMed
62. Woodward RM, Vesey RM, Bacon CJ, White SG, Brick MJ, Blankenbaker DG. Microinstability of the hip: a systematic

review of the imaging findings. Skeletal Radiol. 2020;49(12): 1903-1919. CrossRef PubMed
63. Berthelot JM, Brulefert K, Arnolfo P, Le Goff B, Darrieutort-Laffite C. Update on contribution of hip labral tears to hip pain: a narrative review. Joint Bone Spine. 2023;90(1):105465. CrossRef PubMed
64. Magerkurth O, Jacobson JA, Morag Y, Caoili E, Fessell D, Sekiya JK. Capsular laxity of the hip: findings at magnetic resonance arthrography. Arthroscopy. 2013;29(10):1615-1622. CrossRef PubMed
65. Maas R, Wallentine S, Gerke D, Crager S, Stewart J. Posterior hip instability relocation testing: a resident's case report. J Man Manip Ther. 2017;25(4):215-220. CrossRef PubMed
66. Reiman MP, Décary S, Mathew B, Reiman CK. Accuracy of clinical and imaging tests for the diagnosis of hip dysplasia and instability: a systematic review. J Orthop Sports Phys Ther. 2019;49(2):87-97. CrossRef PubMed
67. Gowd AK, Beck EC, Trammell AP, Edge C, Stubbs AJ. Evaluation of additional causes of hip pain in patients with femoroacetabular impingement syndrome. Front Surg. 2022;9:697488. CrossRef PubMed
68. Spiker AM, Fabricant PD, Wong AC, Suryavanshi JR, Sink EL. Radiographic and clinical characteristics associated with a positive PART (Prone Apprehension Relocation Test): a new provocative exam to elicit hip instability. J Hip Preserv Surg. 2020; 7(2):288-297. CrossRef PubMed
69. Watchmaker LE, Hetzel SJ, Sink EL, Spiker AM. Interrater reliability of the prone apprehension relocation test. Orthop J Sports Med. 2021;9(9):23259671211032229. CrossRef PubMed
70. MacDonald AA, Blankenbaker DG, Brick MJ, Hanna CM, Woodward RM. The posterior crescent sign on MRI and MR arthrography: is it a marker of hip dysplasia and instability? Skeletal Radiol. 2023;52(6):1127-1135. CrossRef PubMed
71. Graesser E, Rhea L, Pascual-Garrido C, et al. Development and external validation of a novel clinical score to quantify the presence of instability characteristics in patients with borderline acetabular dysplasia. Orthop J Sports Med. 2020;8(7_suppl 6):2325967120S00431. CrossRef
72. Horton B, West H, Marl J, Wylie J. Imaging markers of hip instability are associated with worse outcomes at two to four-year follow-up in female patients undergoing hip arthroscopy for femoroacetabular impingement. American Orthopaedic Society for Sports Medicine (AOSSM), 2020 Annual Meeting (Virtual). Orthop J Sports Med. 2020;8:1-2.
73. Cohen D, Ifabiyi M, Mathewson G, et al. The radiographic femoroepiphyseal acetabular roof index is a reliable and reproducible diagnostic tool in patients undergoing hip-preservation surgery: a systematic review. Arthroscopy. 2023;39(4):1074-1087.e1. CrossRef PubMed
74. Dusak IWS, Ciatawi K. The application of the novel femoral-epiphyseal acetabular roof (FEAR) index. Orthop J Sports Med. 2023;11(2 suppl):2325967121S00914. CrossRef
75. Atzmon R, Safran MR. Arthroscopic treatment of mild/borderline hip dysplasia with concomitant femoroacetabular impingement – literature review. Curr Rev Musculoskelet Med. 2022;15(4):300-310. CrossRef PubMed
76. Foissey C, Abid H, Martinot P, Cazor A, Thaunat M. Predictive radiological parameters of failure following surgical management of femoroacetabular impingement associated with borderline acetabular dysplasia. Orthop Traumatol Surg Res. 2023;109(4):103349. CrossRef PubMed
77. Cvetanovich GL, Beck EC, Chalmers PN, et al. Assessment of hip translation in vivo in patients with femoroacetabular impingement syndrome using 3-dimensional computed tomography. Arthrosc Sports Med Rehabil. 2020;2(2):e113-e120. CrossRef PubMed
78. Agten CA, Sutter R, Buck FM, Pfirrmann CW. Hip imaging in athletes: sports imaging series. Radiology. 2016;280(2):351-369. CrossRef PubMed
79. Featherall J, O'Neill DC, Mortensen AJ, Tomasevich KM, Metz AK, Aoki SK. Three-dimensional magnetic resonance arthrography of post-arthroscopy hip instability demonstrates increased effective intracapsular volume and anterosuperior capsular changes. Arthrosc Sports Med Rehabil. 2021;3(6):e1999-e2006. CrossRef PubMed
80. Maldonado DR. CORR Insights®: can dynamic ultrasonography of the hip reliably assess anterior femoral head translation? Clin Orthop Relat Res. 2019;477(5):1099-1100. CrossRef PubMed
81. Sahr ME, Endo Y, Sink EL, Miller TT. Dynamic ultrasound assessment of hip instability and anterior and posterior hip impingement. Skeletal Radiol. 2023;52(7):1385-1393. CrossRef PubMed
82. Packer JD, Foster MJ, Riley GM, et al. Capsular thinning on magnetic resonance arthrography is associated with intra-operative hip joint laxity in women. J Hip Preserv Surg. 2020;7(2):298-304. CrossRef PubMed
83. Curtis DM, Pullen WM, Murray IR, Money A, Safran M. The diagnosis of hip microinstability is correlated with ease of intra-operative hip distraction. Knee Surg Sports Traumatol Arthrosc. 2023;31(1):33-39. CrossRef PubMed
84. Mortensen AJ, Tomasevich KM, Ohlsen SM, O'Neill DC, Featherall J, Aoki SK. Previous arthroscopic hip surgery increases axial distractibility compared to the native contralateral hip and may suggest instability. Arthroscopy. 2022; 38(5):1466-1477. CrossRef PubMed
85. Safran MR, Murray IR, Andrade AJ, et al. Criteria for the operating room confirmation of the diagnosis of hip instability: the results of an international expert consensus conference. Arthroscopy. 2022;38(10):2837-2849.e2. CrossRef PubMed
86. Nwachukwu BU, Gaudiani MA, Marsh NA, Ranawat AS. Labral hypertrophy correlates with borderline hip dysplasia and microinstability in femoroacetabular impingement: a matched case-control analysis. Hip Int. 2019;29(2):198-203. CrossRef PubMed
87. Rosinsky PJ, Mayo BC, Kyin C, et al. The femoral head "divot" sign: a useful arthroscopic sign of hip microinstability. Orthop J Sports Med. 2020;8(5):2325967120917919. CrossRef PubMed
88. Abrams GD, Luria A, Sampson J, et al. Decreased synovial inflammation in atraumatic hip microinstability compared with femoroacetabular impingement. Arthroscopy. 2017;33(3):553-558. CrossRef PubMed
89. Pullen WM, Curtis DM, Safran MR. Central femoral head chondromalacia is associated with a diagnosis of hip instability. Arthrosc Sports Med Rehabil. 2021;4(2):e453-e457. CrossRef PubMed
90. Shibata KR, Matsuda S, Safran MR. Is there a distinct pattern to the acetabular labrum and articular cartilage damage in the non-dysplastic hip with instability? Knee Surg Sports Traumatol Arthrosc. 2017;25(1):84-93. CrossRef PubMed
91. Tibor LM, Liebert G, Sutter R, Impellizzeri FM, Leunig M. Two or more impingement and/or instability deformities are often present in patients with hip pain. Clin Orthop Relat Res. 2013;471(12):3762-3773. CrossRef PubMed
92. Westermann RW, Willey MC. Femoral version in hip arthroscopy: does it matter? Sports Med Arthrosc Rev. 2021;29(1): 28-34. CrossRef PubMed
93. Casartelli NC, Maffiuletti NA, Bizzini M, Kelly BT, Naal FD, Leunig M. The management of symptomatic femoroacetabular impingement: what is the rationale for non-surgical treatment? Br J Sports Med. 2016;50(9):511-512. CrossRef PubMed

94. Selley R, Lawton C, Dooley M, et al. Capsular complications and subsequent instability on the rise as indications for revision hip arthroscopy. Orthop J Sports Med. 2021; 9(10_suppl 5). CrossRef
95. Mortensen AJ, Metz AK, Froerer DL, Aoki SK. Hip capsular deficiency – a cause of post-surgical instability in the revision setting following hip arthroscopy for femoroacetabular impingement. Curr Rev Musculoskelet Med. 2021;14(6): 351-360. CrossRef PubMed
96. Bayer JL, Sekiya JK. Hip instability and capsular laxity. Oper Tech Orthop. 2010;20(4):237-241. CrossRef
97. Bowman KF Jr, Fox J, Sekiya JK. A clinically relevant review of hip biomechanics. Arthroscopy. 2010;26(8):1118-1129. CrossRef PubMed
98. Canham CD, Domb BG, Giordano BD. Atraumatic hip instability. JBJS Rev. 2016;4(5):17. CrossRef PubMed
99. Krych AJ, Thompson M, Larson CM, Byrd JW, Kelly BT. Is posterior hip instability associated with cam and pincer deformity? Clin Orthop Relat Res. 2012;470(12):3390-3397. CrossRef PubMed
100. Bedi A, Galano G, Walsh C, Kelly BT. Capsular management during hip arthroscopy: from femoroacetabular impingement to instability. Arthroscopy. 2011;27(12):1720-1731. CrossRef PubMed
101. Hammoud S, Bedi A, Voos JE, Mauro CS, Kelly BT. The recognition and evaluation of patterns of compensatory injury in patients with mechanical hip pain. Sports Health. 2014;6(2):108-118. CrossRef PubMed
102. Charbonnier C, Kolo FC, Duthon VB, et al. Assessment of congruence and impingement of the hip joint in professional ballet dancers: a motion capture study. Am J Sports Med. 2011;39(3):557-566. CrossRef PubMed
103. Blakey CM, Field MH, Singh PJ, Tayar R, Field RE. Secondary capsular laxity of the hip. Hip Int. 2010;20(4):497-504. CrossRef PubMed
104. Tibor LM, Ganz R, Leunig M. Anteroinferior acetabular rim damage due to femoroacetabular impingement. Clin Orthop Relat Res. 2013;471(12):3781-3787. CrossRef PubMed
105. Eijer H, Hogervorst T. Femoroacetabular impingement causes osteoarthritis of the hip by migration and micro-instability of the femoral head. Med Hypotheses. 2017;104:93-96. CrossRef PubMed
106. Hunt SA. Editorial commentary: hip cam overresection may result in inferior outcomes: the Goldilocks paradox of too little, too much, or just right? Arthroscopy. 2021;37(9):2817-2819. CrossRef PubMed
107. Jackson TJ. Editorial commentary: microinstability after cam osteochondroplasty results from over-resection – everything in moderation. Arthroscopy. 2021;37(1):171-172. CrossRef PubMed
108. Friel N, Ukwuani G, Nho SJ. Current techniques in treating femoroacetabular impingement: capsular repair and plication. Am J Orthop (Belle Mead NJ). 2017;46(1):49-54. PubMed
109. Ortiz-Declet V, Mu B, Chen AW, et al. Should the capsule be repaired or plicated after hip arthroscopy for labral tears associated with femoroacetabular impingement or instability? A systematic review. Arthroscopy. 2018;34(1):303-318. CrossRef PubMed
110. Woodward RM, Philippon MJ. Persistent or recurrent symptoms after arthroscopic surgery for femoroacetabular impingement: a review of imaging findings. J Med Imaging Radiat Oncol. 2019;63(1):15-24. CrossRef PubMed
111. Tahoun MF, Lizano-Díez X, Soler BC, Pons MT. Superior outcomes after arthroscopic treatment of femoroacetabular impingement and labral tears with closed versus open capsule. Knee Surg Sports Traumatol Arthrosc. 2023;31(10):4501-4509. CrossRef PubMed
112. Wu JY, Li W, Xu LY, Zheng G, Chen XD, Shen C. Ligamentum teres tears and increased combined anteversion are associated with hip microinstability in patients with borderline dysplasia. Arthroscopy. 2024;40(3):745-751. CrossRef PubMed
113. Nepple JJ, Fowler LM, Larson CM. Decision-making in the borderline hip. Sports Med Arthrosc Rev. 2021;29(1):15-21. CrossRef PubMed
114. Featherall J, Froerer D, Mortensen A, Aoki S, Metz A. Poster 195: inverse relationship of hip capsular thickness on magentic resonance imaging and increased axial distraction under anesthesia: further characterization of hip laxity. Orthop J Sports Med. 2022;10(7_suppl 5):2325967121S0075. CrossRef
115. Martin RL, McDonough C, Enseki K, Kohreiser D, Kivlan BR. Clinical relevance of the ligamentum teres: a literature review. Int J Sports Phys Ther. 2019;14(3):459-467. CrossRef PubMed
116. Chahla J, Soares EAM, Devitt BM, et al. Ligamentum teres tears and femoroacetabular impingement: prevalence and preoperative findings. Arthroscopy. 2016;32(7):1293-1297. CrossRef PubMed
117. Martin RL, Palmer I, Martin HD. Ligamentum teres: a functional description and potential clinical relevance. Knee Surg Sports Traumatol Arthrosc. 2012;20(6):1209-1214. CrossRef PubMed
118. Kaya M, Kano M, Sugi A, Sakakibara Y, Yamashita T. Factors contributing to the failure of conservative treatment for acetabular labrum tears. Eur Orthop Traumatol. 2014;5(3):261-265. CrossRef
119. Swarup I, Zaltz I, Robustelli S, Sink E. Outcomes of periacetabular osteotomy for borderline hip dysplasia in adolescent patients. J Hip Preserv Surg. 2020;7(2):249-255. CrossRef PubMed
120. Matthewson G, Williams A, Macciacchera M, Shahabinezhad A, Duong A, Ayeni OR. Effective management options for treatment of microinstability of the hip: a scoping review. Curr Rev Musculoskelet Med. 2023;16(1):9-18. CrossRef PubMed
121. Ejnisman L, Elisman K, Safran MR. Effectiveness of nonoperative management of hip microinstability. Am J Sports Med. 2022;50(4):1013-1019. CrossRef PubMed
122. Savic M, D'Angelo K. Exploring the role of microinstability of the hip: an atypical presentation of femoroacetabular impingement (FAI) and labral tear in a collegiate endurance athlete: a case report. J Can Chiropr Assoc. 2019;63(3):187-196. PubMed
123. Beck EC, Alter T, Mehta N, et al. Contemporary hip capsular management and closure using a suture passing device. Arthrosc Tech. 2019;8(9):e947-e952. CrossRef PubMed
124. Wylie J, Beckmann J, Maak T, Aoki S. Capsular repair for instability following hip arthroscopy for femoroacetabular impingement: preliminary outcomes and description of surgical technique. Arthroscopy. 2013;29(12):e206-e207. CrossRef
125. Charles T, Jayankura M, Laude F. Treatment of hip microinstability with arthroscopic capsular plication: a retrospective case series. Hip Pelvis. 2023;35(1):15-23. CrossRef PubMed
126. Wolff AB, Scanaliato JP. Editorial commentary: the importance of capsular closure following hip arthroscopy – leave no trace: an outdoorsman's ramblings. Arthroscopy. 2022;38(11):3041-3042. CrossRef PubMed
127. Wu ZX, Ren WX, Ren YM, Tian MQ. Arthroscopic labral debridement versus labral repair for patients with femoroacetabular impingement: a meta-analysis. Medicine (Baltimore). 2020;99(19):e20141. CrossRef PubMed
128. Brusalis CM, Peck J, Wilkin GP, et al. Periacetabular osteotomy as a salvage procedure: early outcomes in patients treated for iatrogenic hip instability. J Bone Joint Surg Am. 2020;102 (suppl 2):73-79. CrossRef PubMed

129. Nepple JJ, Ambastha C, Abu-Amer W, Clohisy JC. Mid-term outcomes of combined hip arthroscopy and limited open capsular plication in the non-dysplastic hip. Iowa Orthop J. 2021;41(1):133-139. PubMed
130. Nepple JJ. Editorial commentary: at the intersection of borderline dysplasia and femoroacetabular impingement – which way should we turn? Arthroscopy. 2020;36(4):1185-1188. CrossRef PubMed
131. Jimenez AE, Monahan PF, Miecznikowski KB, et al. Achieving successful outcomes in high-level athletes with borderline hip dysplasia undergoing hip arthroscopy with capsular plication and labral preservation: a propensity-matched controlled study. Am J Sports Med. 2021;49(9):2447-2456. CrossRef PubMed
132. Wong SE, Newhouse AC, Wichman DM, Bessa F, Williams J, Nho SJ. Patients with a high femoroepiphyseal roof with concomitant borderline hip dysplasia and femoroacetabular impingement syndrome do not demonstrate inferior outcomes following arthroscopic hip surgery. Arthroscopy. 2022; 38(5):1509-1515. CrossRef PubMed
133. Kalisvaart MM, Safran MR. Hip instability treated with arthroscopic capsular plication. Knee Surg Sports Traumatol Arthrosc. 2017;25(1):24-30. CrossRef PubMed
134. O'Neill DC, Mortensen AJ, Cannamela PC, Aoki SK. Clinical and radiographic presentation of capsular iatrogenic hip instability after previous hip arthroscopy. Am J Sports Med. 2020;48(12):2927-2932. CrossRef PubMed
135. Wylie JD, Beckmann JT, Maak TG, Aoki SK. Arthroscopic capsular repair for symptomatic hip instability after previous hip arthroscopic surgery. Am J Sports Med. 2016;44(1):39-45. CrossRef PubMed
136. Leunig M, Beaulé PE, Ganz R. The concept of femoroacetabular impingement: current status and future perspectives. Clin Orthop Relat Res. 2009;467(3):616-622. CrossRef PubMed
137. Parvizi J, Bican O, Bender B, et al. Arthroscopy for labral tears in patients with developmental dysplasia of the hip: a cautionary note. J Arthroplasty. 2009;24(6)(suppl):110-113. CrossRef PubMed
138. Curley AJ, Domb BG. Editorial commentary: evaluate for the Beighton Score and additional radiographic signs of instability prior to proceeding with hip arthroscopy in patients with combined borderline hip dysplasia and excessive femoral anteversion. Arthroscopy. 2023;39(4):978-980. CrossRef PubMed
139. Hohmann E. Editorial commentary: hip arthroscopy for femoroacetabular impingement in patients with borderline dysplasia does not result in inferior outcomes compared with outcomes in patients without dysplasia: do not fear. Arthroscopy. 2022; 38(5):1516-1518. CrossRef PubMed
140. Canham CD, Yen YM, Giordano BD. Does femoroacetabular impingement cause hip instability? A systematic review. Arthroscopy. 2016;32(1):203-208. CrossRef PubMed
141. Bedi A, Dolan M, Leunig M, Kelly BT. Static and dynamic mechanical causes of hip pain. Arthroscopy. 2011;27(2):235-251. CrossRef PubMed
142. Philippon MJ, Ho CP, Briggs KK, Stull J, LaPrade RF. Prevalence of increased alpha angles as a measure of cam-type femoroacetabular impingement in youth ice hockey players. Am J Sports Med. 2013;41(6):1357-1362. CrossRef PubMed
143. Heerey J, Agricola R, Smith A, et al. The size and prevalence of bony hip morphology do not differ between football players with and without hip and/or groin pain: findings from the force cohort. J Orthop Sports Phys Ther. 2021;51(3):115-125. CrossRef PubMed
144. Retchford TH, Crossley KM, Grimaldi A, Kemp JL, Cowan SM. Can local muscles augment stability in the hip? A narrative literature review. J Musculoskelet Neuronal Interact. 2013;13(1):1-12. PubMed
145. Cowan SM, Bennell KL, Hodges PW, Crossley KM, McConnell J. Simultaneous feedforward recruitment of the vasti in untrained postural tasks can be restored by physical therapy. J Orthop Res. 2003;21(3):553-558. CrossRef PubMed
146. Tsao H, Hodges PW. Immediate changes in feedforward postural adjustments following voluntary motor training. Exp Brain Res. 2007;181(4):537-546. CrossRef PubMed
147. Jull G, Trott P, Potter H, et al. A randomized controlled trial of exercise and manipulative therapy for cervicogenic headache. Spine. 2002;27(17):1835-1843. CrossRef PubMed

Arch Physioter 2024; 14: 47-55
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3015
ORIGINAL RESEARCH ARTICLE

Defining the glenohumeral range of motion required for overhead shoulder mobility: an observational study

Linda Dyer[1], **Jaap Swanenburg**[2], **Hermann Schwameder**[3], **Samy Bouaicha**[4]

[1]Department of Physiotherapy, Balgrist University Hospital, University of Zurich, Zurich - Switzerland
[2]Department of Chiropractic Medicine, Balgrist University Hospital, Zurich - Switzerland
[3]Department of Sport and Exercise Science, Paris Lodron University of Salzburg, Salzburg - Austria
[4]Department of Orthopedics, Balgrist University Hospital, University of Zurich, Zurich - Switzerland

ABSTRACT

Background: Recovery of overhead mobility after shoulder surgery is time-consuming and important for patient satisfaction. Overhead stretching and mobilization of the scapulothoracic and glenohumeral (GH) joints are common treatment interventions. The isolated GH range of motion (ROM) of flexion, abduction, and external rotation required to move above 120° of global shoulder flexion in the clinical setting remains unclear. This study clarified the GH ROM needed for overhead mobility.

Methods: The timely development of shoulder ROM in patients after shoulder surgery was analyzed. Passive global shoulder flexion, GH flexion, abduction, and external rotation ROM were measured using goniometry and visually at 2-week intervals starting 6-week postsurgery until the end of treatment. Receiver operating characteristic curves were used to identify the GH ROM cutoff values allowing overhead mobility.

Results: A total of 21 patients (mean age 49 years; 76% men) after rotator cuff repair (71%), Latarjet shoulder stabilization (19%), and arthroscopic biceps tenotomy (10%) were included. The ROM cutoff value that accurately allowed overhead mobility was 83° for GH flexion and abduction with the area under the curve (AUC) ranging from 0.90 to 0.93 ($p < 0.001$). The cutoff value for GH external rotation was 53% of the amount of movement on the opposite side (AUC 0.87, $p < 0.001$).

Conclusions: Global shoulder flexion above 120° needs almost full GH flexion and abduction to be executable. External rotation ROM seems less important as long as it reaches over 53% of the opposite side.

Keywords: Glenohumeral, Postoperative, Range of motion, Rehabilitation, Shoulder, Stiffness

What's already known about this topic?

- *Overhead shoulder mobility after shoulder surgery is an important treatment goal, but it requires time and a global shoulder flexion angle of over 120°. The exact relationship between global shoulder flexion and GH ROM remains unclear.*

What does the study add?

- *An observation of ROM development in patients after shoulder surgery provides the GH ROM cutoff values for global shoulder flexion above 120°. GH ROM measurements can be used to predict overhead shoulder mobility.*

Background

Arm elevation is a crucial function of the shoulder girdle. Restricted shoulder elevation impairs many daily and athletic activities, such as reaching overhead. Overhead movement requires global shoulder flexion angles over 120° (1,2). After shoulder surgery, mobility can be restricted due to different underlying mechanisms (3). Restoring arm elevation is an important goal for all shoulder treatments and plays an important role in subsequent patient satisfaction (4). Common treatments include mobilizing and stretching the shoulder into passive end range elevation (5,6). However, these treatment approaches often cause severe pain (7,8). Therefore, understanding shoulder biomechanics and the relationship between its components is necessary to treat shoulders with motion loss (9).

Global shoulder flexion is defined as the motion of the humerus relative to the thorax in the sagittal plane of

Received: December 22, 2023
Accepted: July 18, 2024
Published online: August 26, 2024

Corresponding author:
Linda Dyer
email: linda.dyer@balgrist.ch

approximately 160° (1). It is a combined motion of the scapulothoracic and glenohumeral (GH) joints. The normal function of the shoulder complex is a coordinated motion sequence of all joint elements. Impairment in one joint directly affects the whole kinematic chain (10,11). Loss of GH range of motion (ROM) alters the entire kinematics of motion. The scapular upward rotation occurs earlier during arm elevation as a compensatory strategy for limited GH ROM (11,12). The scapula pulls the clavicle into an early final retraction position close to the neck muscles. Once the final position of the scapula and clavicle is achieved, only the thoracic spine can move to gain more elevation motion (13). The abnormal movement pattern of the scapula and clavicle often continues for a longer time, even after GH mobility has restored considerably (7,11).

Improving arm elevation by optimizing the scapulothoracic substitution is important in managing restricted shoulder ROM (12,13). However, excessive compensatory movements could cause secondary problems in other joints (14). Mobilizing and stretching the shoulder into further global flexion results in greater rotation of the scapula at the acromioclavicular joint. This may induce compression of the soft tissues between the coracoid process and the clavicle, which can again lead to pain (7). Further, subacromial structures can be irritated when the arm is pushed into elevation (8). In other words, GH loss of motion can result in mechanically related shoulder pain.

Thus, sufficient GH ROM is mandatory for overhead mobility as it decreases the requirement for the scapulothoracic substitution and allows the scapula and clavicle to move around the thorax (7). An impairment-based rehabilitation approach should therefore focus on improving GH ROM (15). In particular, increasing external rotation (ER) ROM has been recommended to improve global shoulder flexion (8,14,16,17).

From a biomechanical perspective, the amount of GH ER in full global shoulder flexion is controversially discussed in the literature (18-23). However, with regard to postoperative patient satisfaction, ER, if not massively impaired, has no major influence on overall patient satisfaction (24).

The amount of GH flexion, abduction, and ER mobility required to perform an overhead arm movement in patients with restricted shoulder ROM remains unclear. However, knowledge of the relationship of GH mobility and global shoulder flexion is important to guide the rehabilitation process.

Therefore, this study aimed to investigate the passive GH cutoff value for overhead mobility (global shoulder flexion above 120°). Based on preliminary data, we hypothesized that nearly full GH flexion and abduction is required for overhead mobility, whereas GH ER is negligible.

Methods

Study design and setting

A retrospective observational study was conducted in a group of patients after a variety of shoulder surgery. Data of patients who underwent postoperative physical therapy at the Balgrist University Hospital outpatient Physiotherapy Department, Zurich, Switzerland, were collected and analyzed. All patients provided written informed consent for the anonymized use of their medical data for scientific purposes before data collection. The retrospective data analysis was approved by the ethics committee of the Canton of Zurich (BASEC 2016 01120).

The following data from the patient reports were used for analysis: (a) ROM measurements of the unaffected and operated sides recorded 6 weeks after the surgery; (b) follow-up ROM measurements of the operated side at intervals of approximately 2 weeks. Demographic and baseline characteristics of the patients were recorded at the start of the treatment.

Participants

A total of 34 patients referred for treatment after arthroscopic rotator cuff repair, open shoulder stabilization with Latarjet procedure, arthroscopic biceps tenotomy and of a minimum age of 18 years were initially selected; 21 patients met the inclusion criteria and were included in the data analysis. The sample size for this study was determined a priori, based on similar studies in the literature, which typically included 20 to 30 subjects, ensuring sufficient statistical power (15,17,25).

Exclusion criteria were as follows: an unhealthy shoulder on the opposite side, prior shoulder fracture, scoliosis, and documented symptoms of complex regional pain syndrome. For the analysis, only the patients treated by physiotherapists who had experience in treating shoulder conditions for more than 10 years and who had seen more than five patients during the recruiting period were included.

All patients were treated once or twice a week with individual sessions and hydrotherapy in groups, each session lasting 30 minutes. The interventions were (a) instruction and progression of home exercises to increase shoulder ROM and rotator cuff and scapular muscle strength; (b) cognitive behavioral strategies, including goal setting, education, and positive reinforcement; (c) passive GH and scapulothoracic joint mobilization in supine or side position without pain provocation; (d) active joint movement in water of 34°C and swimming as soon as allowed; and (e) soft tissue massage. Treatment procedures after shoulder surgery were based on the patient's condition and followed the standardized guidelines of the surgeon.

ROM measurement

The ROM measurement procedure used in the Physiotherapy Department is a combination of the method originally described by Winkel et al (26) and Cyriax (27), and, to some extent, our own clinical experience. The method was evaluated for its reproducibility. Reliability was excellent across all movement directions (intraclass correlation coefficient [ICC], 0.91-0.99). The standard error of measurement ranged from 2° to 5°, and the smallest detectable change ranged from 5° to 14° (unpublished data). All physiotherapists of the institution participated in a training session to

standardize their shoulder ROM measurements for quality reasons. The shoulder ROM measurements and notations were part of the daily routine and were performed before the regular physiotherapy session.

For shoulder ROM measurements, all movements were performed passively until the end range position. Passive end range position was determined by the tactile perception of a clear resistance to further motion against the stabilizing hand (28). All passive movements were measured either with a standard 205 × 45-mm, double-armed 360° goniometer constructed of clear plastic or by visual estimation. All measurements were conducted with an accuracy of 5°, as this corresponds to the clinical standard.

Global shoulder flexion

For global shoulder flexion, the patients stood with their eyes fixed forward. The examiner moved the patient's arm with one hand in the sagittal plane with the elbow in full extension and the thumb pointing up to the maximal end range position. The other hand rested on the scapula and thorax to secure upright posture. The patient was then asked to hold the elevated arm in position with his other hand while the examiner measured the angle using a goniometer. Anatomical landmarks and measurement device positioning followed the recommendations of Norkin and White (28). The stationary arm of the goniometer was placed parallel to the midline of the thorax, and the moving arm was aligned with the shaft of the humerus and lateral epicondyle (Fig. 1A).

GH motion

For GH motion measurement, the patients were sitting upright on a chair with their feet on the floor. GH flexion was performed in the sagittal plane. The arm was passively moved with one hand, while the other hand immobilized the lower angle of the scapula with the thumb. The angle was measured visually when the scapula began to rotate. The landmarks used for global shoulder flexion were also used here (Fig. 1B).

GH abduction was performed in the plane of the scapula approximately 30° anterior to the frontal plane. One hand was placed on the acromion for stabilization and the other hand moved the arm until the scapula began to move. The landmarks for visual estimation were the sagittal plane and the shaft of the humerus (Fig. 1C).

GH ER was taken by passively placing the patient's arm at 0° of GH abduction with the elbow flexed at 90° with one hand. The medial border of the scapula was stabilized with the fingers of the other hand while the arm was moved in ER. The angle was measured visually from the sagittal plane and the forearm using the olecranon process and ulnar styloid for reference (Fig. 1D).

Statistical analysis

To describe the sample, data are expressed using descriptive statistics. Mean ROM of the healthy side at the start of evaluation was used as a reference for the percentage calculation. Mean ROM value of the operated arm at baseline and

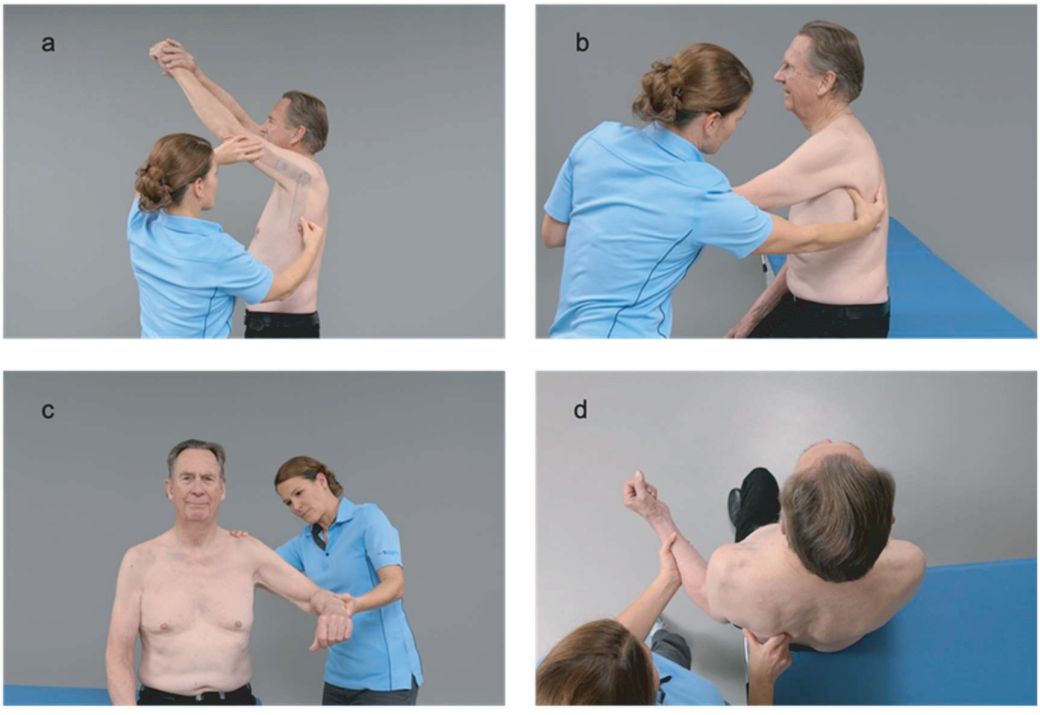


FIGURE 1 - Joint measurement: (a) passive global shoulder flexion measured with goniometer; (b) passive glenohumeral (GH) flexion; (c) passive GH abduction; and (d) passive external rotation. All GH movements were measured visually.

at the end of evaluation were computed. Data are presented separately for the dominant and nondominant sides.

Overhead movement was defined as a global shoulder flexion above 120° and was coded as a dichotomous variable (positive/negative results). To evaluate which GH ROM can be used as a predictor of overhead movement, a receiver operating characteristic (ROC) curve was computed. ROC curves were constructed by plotting sensitivity versus 1-specificity for the absolute data and the percentage data of the opposite side as independent variables. The area under the curve (AUC) was calculated to quantify the accuracy of the predictor. The AUCs can range from 0.50 (no accuracy in distinguishing overheads from nonoverheads) and 1.00 (perfect accuracy). An AUC of 0.75 has been proposed to be clinically useful (29). Significance level was set at $p < 0.05$. The optimal threshold value for each GH movement was determined by selecting the cutoff value closest to 80% specificity. Sensitivity at fixed point of specificity is suitable for determining the validity of a predictor and for comparing two diagnostic tests (30). In addition to the calculation, the measurements were graphically illustrated to exemplify the relationship between GH ROM and overhead mobility.

All statistical analyses were performed under the supervision of an experienced biostatistician using Statistical Package for the Social Sciences (SPSS) version 23.0 (IBM SPSS Statistics Inc., Chicago, IL, USA).

TABLE 1 - Patients' characteristics (n = 21)

Characteristic	Summary
Female/male	5/16
Mean age, years (SD)	49.1 (15.7)
Mean body height, cm (SD)	174.4 (8.3)
Mean body mass, kg (SD)	81.5 (14.2)
Dominant hand: left/right	0/21
Side of surgery: left/right	10/11
Surgery (number)	
Rotator cuff repair	15
Latarjet shoulder stabilization	4
Arthroscopic biceps tenotomy	2
Mean evaluation duration, days	
Rotator cuff repair	92.0
Latarjet shoulder stabilization	69.0
Arthroscopic biceps tenotomy	93.5
Total measurement points	127
Rotator cuff repair	94
Latarjet shoulder stabilization	21
Arthroscopic biceps tenotomy	12

SD = standard deviation.

Results

A total of 34 patients were screened for inclusion. Of these, 21 patients with a total of 127 complete documentation measurements met the inclusion criteria and were analyzed. Reasons for exclusion were as follows: patients treated by therapists with less than five patients during the analyzing period (10), unhealthy shoulder on the opposite side (2), and documented symptoms of complex regional pain syndrome (1). Finally, patients of only two physiotherapists fulfilled the selection criteria. Demographic characteristics of the included patients are summarized in Table 1.

Descriptive data of the measured ROM value at the start and end of evaluation are presented in Table 2. Dominant and nondominant sides presented similar ROM values. Patients exhibited a larger standard deviation (SD) in the ROM on the healthy side for GH AR compared to GH flexion and GH abduction. All cases showed some loss of motion at the start of the evaluation and an improvement in ROM at the end of the evaluation.

The results of the ROC curve analysis are shown in Figure 2. From the 127 measurements, 70 were classified as overhead and 57 as nonoverhead. The absolute and percentage data of all GH movements showed good performance in distinguishing overhead mobility with AUCs ranging from 0.80 to 0.93, which were significant ($p < 0.001$).

The cutoff values of the shoulder ROM are presented in Table 3. The cutoff values closest to 80% specificity, along with their corresponding sensitivity and AUC with 95% confidence interval (CI), are presented separately based on degrees and as a percentage relative to the ROM of the opposite side. The 95% CI of the AUCs exhibit a relatively narrow range of 0.08 to 0.16, confirming the predictor's test strength. GH ER exhibits lower sensitivity compared to GH flexion and GH abduction. This is supported by their respective AUC values.

In addition to the calculation, the measurements were graphically illustrated to exemplify the relationship between GH ROM and the ability to move overhead (Fig. 3). The pattern for GH flexion and GH abduction differs from that of GH ER. The graph illustrates that some patients were able to achieve overhead movement with less than 20° and 20% GH ER, respectively.

Discussion

The goal of this observational study was to evaluate the required GH ROM to achieve overhead mobility in patients after shoulder surgery. Our results showed that overhead mobility can be expected with a GH ROM of 83° for flexion and abduction each and with 53% ER of the contralateral side. In other words, consistent with our hypothesis, overhead mobility needs nearly full ROM for GH flexion and abduction, whereas ER ROM seems less important.

An understanding of normal shoulder ROM is crucial to interpret the results of this investigation. Normative data vary considerably in the literature as many factors can influence ROM. These factors include age, gender, sports activity, and the position of the subject during the examination. Arm dominance is another factor that can influence shoulder ROM (31-33). To minimize the abovementioned variability, the healthy side of the participants was used as a reference. In the present study the mean shoulder ROM for GH flexion was 93° (SD ± 4°), for GH abduction 93° (SD ± 5°), and for GH ER 43° (SD ± 17°) on the dominant side (Tab. 2). Due to the

TABLE 2 - Descriptive data of range of motion value at start and end of evaluation

Shoulder	Global flexion	GH flexion	GH abduction	GH exorotation
At start of evaluation				
Healthy side dominant				
Mean/%	154°/100%	93°/100%	93°/100%	43°/100%
(SD/range)	(5/150-165)	(4/85-100)	(5/90-105)	(17/5-60)
Healthy side nondominant				
Mean/%	155°/100%	96°/100%	93°/100%	47°/100%
(SD/range)	(7/140-165)	(5/90-100)	(5/85-100)	(13/15-75)
Operated side dominant				
Mean/%*	111°/72%	73°/78%	69°/74%	15°/35%
(SD/range)	(23/70-145)	(13/50-90)	(12/50-90)	(15/0-35)
Operated side nondominant				
Mean/%*	109°/70%	65°/68%	62°/67%	–1°/0%
(SD/range)	(24/80-140)	(18/35-90)	(18/35-85)	(7/–20-5)
At end of evaluation				
Operated side dominant				
Mean/%*	137°/89%	90°/97%	87°/94%	38°/88%
(SD/range)	(17/110-135)	(9/70-100)	(7/70-95)	(19/0-60)
Operated side nondominant				
Mean/%*	143°/92%	90°/94%	87°/94%	28°/60%
(SD/range)	(18/100-160)	(8/70-100)	(11/55-95)	(18/0-50)

Operated on the dominant side, n = 11, nondominant side, n = 10; total, n = 21.
GH = glenohumeral; SD = standard deviation.
*Percent of range of motion of the healthy side.

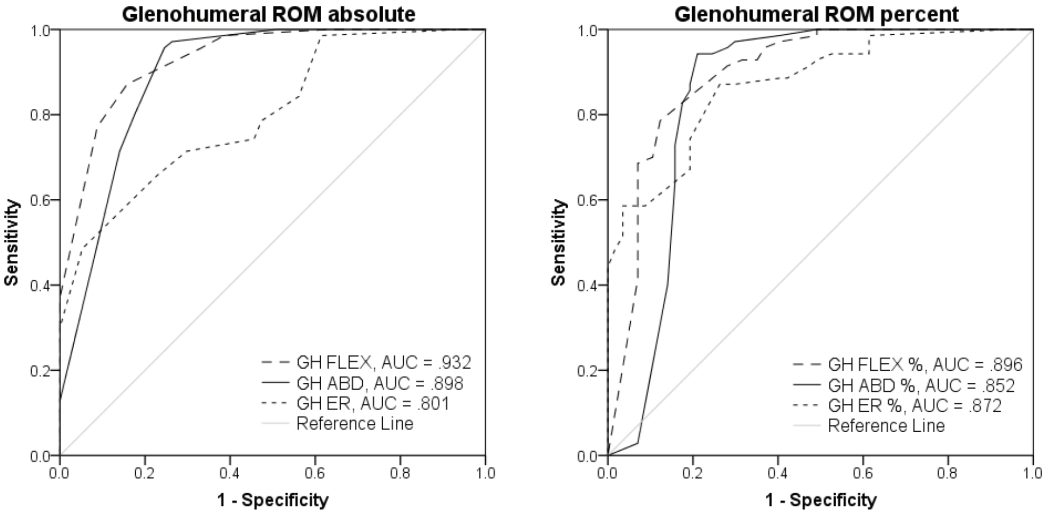
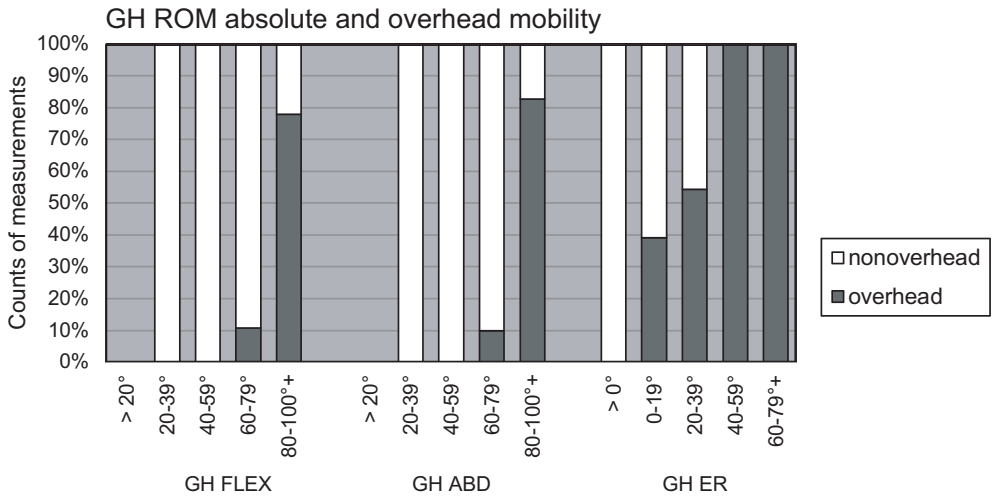


FIGURE 2 - Receiver operating characteristic curves for glenohumeral ROM. It was constructed with the data points of ROM absolute and ROM percentage of the opposite side by plotting sensitivity versus 1-specificity. The greater the area under the curves, the greater was the ability of the predictor to distinguish between overhead mobility (>120° global flexion) and nonoverhead mobility. ABD = abduction; AUC = area under the curve; ER = external rotation; FLEX = flexion; GH = glenohumeral; ROM = range of motion.

wide range of GH ER, the relationship between the affected and healthy sides was used for interpretation.

To the best of our knowledge, this is the first study to investigate the GH cutoff value for overhead shoulder mobility in patients recovering from a shoulder surgery. A comparative analysis of our data with other studies is difficult due to several methodological differences. Previous studies of shoulder kinematics in patients with loss of motion have predominantly described active rather than passive motion (11,12,14,15,34). Both conditions are important to understand the state of the joint, but passive ROM measurements are more useful to obtain the maximal achievable motion.



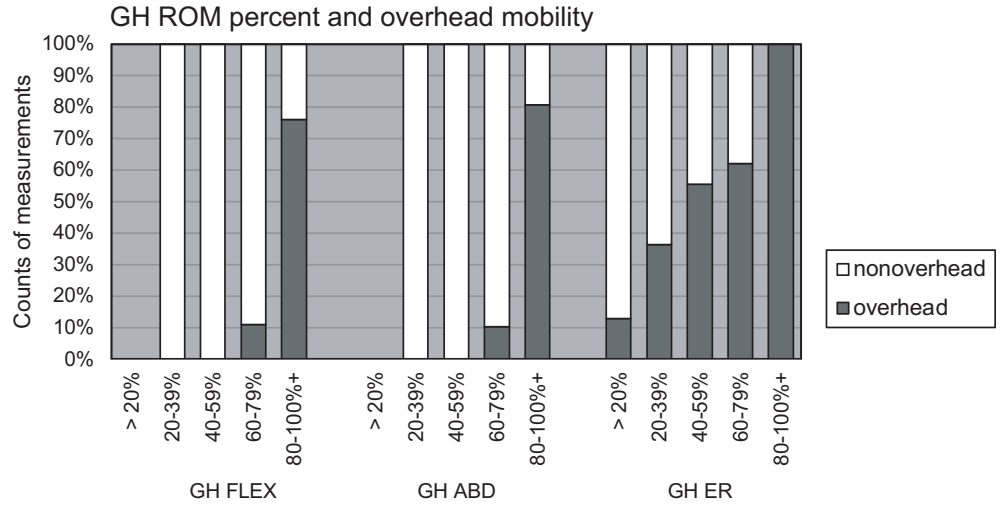


FIGURE 3 - Relationship between the glenohumeral range of motion improvement and the overhead mobility (>120° global flexion). The column chart displays all measurements (n = 127), divided into white (nonoverhead; n = 57) and gray (overhead; n = 70), along with the corresponding glenohumeral range of motion. Glenohumeral movements are demonstrated in absolute values and in percentage of the opposite side values. GH FLEX = glenohumeral flexion; GH ABD = glenohumeral abduction; GH ER = glenohumeral external rotation; GH ROM = glenohumeral range of motion.

TABLE 3 - Glenohumeral ROM cutoff values at 80% specificity that allow overhead mobility

Predictor	80% specificity cutoff value	Sensitivity (%)	AUC	95% CI
GH ROM absolute				
FLEX	83°	87	0.93	0.89-0.97
ABD	83°	80	0.90	0.83-0.95
ER	28°	66	0.80	0.72-0.87
GH ROM percent				
FLEX	87%	90	0.90	0.83-0.95
ABD	85%	87	0.85	0.77-0.93
ER	53%	74	0.87	0.81-0.93

The ROM cutoff values closest to 80% specificity are presented both as absolute values in degrees and as a percentage relative to the opposite side. The corresponding sensitivity at the identified cutoff point is displayed, providing insight into the test's ability to correctly identify true positives at this level of specificity. Along with the AUC and 95% CI, this provides a comprehensive view of the diagnostic performance at the specified specificity level.
ABD = abduction; AUC = area under the curve; CI = confidence interval; ER = external rotation; FLEX = flexion; GH ROM = glenohumeral range of motion.

In addition, passive motion assessments allow clinicians to estimate the amount of isolated GH motion (31). Another factor that affects the results when GH mobility is studied is the type of device used to measure shoulder ROM (35). In clinical practices, shoulder ROM is usually measured with goniometry or visually (36). The accuracy of visual estimation and goniometry varies highly in the current literature (36). However, Warth and Millett (33) have reported that experienced clinicians can measure shoulder ROM with an acceptable precision.

Thus, a direct comparison of our findings with those of previous studies was not possible due to the aforementioned reasons. Nevertheless, consistent with our data, Stenvers (7) reported similar GH ROM values for global shoulder flexion. The author used X-ray cinematography to study the development of global shoulder flexion in subjects with frozen shoulder. Passive shoulder motion in supine position was used for the investigation. The results showed that almost 90° of GH flexion and abduction was necessary before the shoulder could move over a so-called 90° mechanism. This 90°

mechanism was used as an umbrella term for the abnormal motion kinematics of the frozen shoulder and is comparable with the nonoverhead mobility of the present investigation. In a study by Baettig et al (37) patient satisfaction after rotator cuff repair was analyzed. They found active abduction ROM was the only shoulder movement that significantly correlated with higher patient satisfaction in a multivariate analysis. It is known that GH stiffness generates greater impairments in global abduction movements as it does for global flexion movements due to the reduced compensating ability of the scapulothoracic joint in the abduction plane (7). This supports the findings of the present study and indicates that GH mobility plays a key role in shoulder function.

Interestingly, the results demonstrated that some patients could move overhead with considerably restricted GH ER, whereas only a few patients could move overhead with less than 80° of GH flexion and abduction (Fig. 3). A possible explanation was found in a basic study that investigated the effect of selective capsular shortening on passive GH ROM (38). The shortening of the superior part of the capsule resulted in limited GH ER of the adducted arm, whereas GH abduction was not restricted (38). According to Crétual et al (39), GH ER mobility in adduction is least correlated with global shoulder mobility and should therefore be done with the shoulder in abduction. Nevertheless, ER with the arm at the side is commonly used to monitor the development of mobility in patients with restricted shoulder ROM. It has the advantage of being assessed independently of abduction ability, which is often restricted after shoulder surgery (39).

Thus, a question arises about the amount of ER necessary for full global shoulder flexion. However, this topic is controversially discussed in the literature (18,20,21,23). In this context, McClure et al (35) mentioned the importance of scapular upward rotation in full arm elevation for a healthy shoulder. They speculated that scapular upward rotation reduces GH ER requirement. This may be an explanation why some patients were able to move overhead with considerably restricted ER values.

Several limitations should be considered when interpreting the results of our study. First, the data were based on clinical assessments with visual estimation and goniometry, which are not the gold standards for research due to the lack of desired accuracy. Second, when interpreting ROM in clinical practice, the measurement of all ROM directions is important. It provides the information which parts of the capsule are responsible for the specific restrictions (3,38). The present investigation did not evaluate GH internal rotation ROM. The internal rotation in adduction with the hand behind back maneuver is a motion of different joints that requires GH, scapulothoracic, elbow, wrist, and finger movements (40). It is therefore recommended to measure isolated GH internal rotation with the arm abducted. However, internal rotation cannot be measured as recommended in patients with restricted GH abduction. Finally, the present study group consisted of 21 patients who had undergone different shoulder surgery. This nonhomogeneous group may have different impairments. Nevertheless, the mobility of the GH joint is significantly influenced by its biomechanical properties, such as ROM and the surrounding musculoskeletal structures, rather than solely by surgical procedures. The relatively small sample size of 21 patients is consistent with similar studies in the field (15,17,25). Additionally, the AUCs demonstrated a narrow range within the 95% confidence interval (CI), indicating a high level of validity of the results despite the limited sample size.

The main clinical implication of the findings of this study is that the assessment of GH ROM is important for predicting global shoulder flexion mobility. The results imply that nearly normal GH flexion and abduction ROM is required before the shoulder can move above 120° global shoulder flexion. Therefore, a rehabilitation approach that focuses on GH mobility improvement rather than on global shoulder flexion is recommended. The results of this investigation showed a tendency toward greater importance of GH flexion and abduction values than for GH ER, which needs to be confirmed by future research.

Conclusion

This study documents the cutoff values for GH flexion, abduction, and ER ROM that can accurately predict overhead mobility. Results showed that 83° of GH flexion and abduction was required before patients could move their arms above 120° of global shoulder flexion. This means nearly full GH ROM in flexion and abduction is required before overhead mobility is achievable. Consequently, overhead stretches in the presence of GH stiffness should be performed with caution. The cutoff value for GH ER in degrees was inaccurate for interpretation due to the wide range of GH ER of the healthy opposite shoulders. Therefore, it is suggested to use the percent value. About 53% of the ROM of the opposite side for GH ER was required for overhead mobility.

Acknowledgments

The authors thank Marcel Enzler for the technical help and general support and Prof. Dr. Burkhardt Seifert for helping with the statistical analysis. Clinical Trial registration: BASEC 2016 01120, retrospectively registered.

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: LD and JS were responsible for the design of the study. LD conducted the final analysis and wrote the manuscript. LD and SB provided significant contributions to data analysis. JS and HS mentored LD to interpret the data and provided critical feedback for manuscript. All authors read and approved the final manuscript.

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

Ethics approval and consent to participate: This study was approved by the ethics committee of the Canton of Zurich (BASEC 2016 01120). All participants provided written informed consent for the anonymized use of their medical data for scientific purposes.

Consent for publication: The authors affirm that participants provided informed consent to publish the obtained data in this study.

References

1. Magermans DJ, Chadwick EK, Veeger HE, van der Helm FC. Requirements for upper extremity motions during activities of daily living. Clin Biomech (Bristol, Avon). 2005;20(6):591-599. CrossRef PubMed
2. Namdari S, Yagnik G, Ebaugh DD, et al. Defining functional shoulder range of motion for activities of daily living. J Shoulder Elbow Surg. 2012 Sep;21(9):1177-83. CrossRef. Epub 2011 Nov 1. PubMed.
3. Warner JJ, Greis PE. The treatment of stiffness of the shoulder after repair of the rotator cuff. J Bone Joint Surg. 1997; 79(8):1260-1269. CrossRef
4. O'Holleran JD, Kocher MS, Horan MP, Briggs KK, Hawkins RJ. Determinants of patient satisfaction with outcome after rotator cuff surgery. J Bone Joint Surg Am. 2005;87(1):121-126. CrossRef PubMed
5. Vermeulen HM, Obermann WR, Burger BJ, Kok GJ, Rozing PM, van Den Ende CH. End-range mobilization techniques in adhesive capsulitis of the shoulder joint: a multiple-subject case report. Phys Ther. 2000;80(12):1204-1213. CrossRef PubMed
6. Griggs SM, Ahn A, Green A. Idiopathic adhesive capsulitis. A prospective functional outcome study of nonoperative treatment. J Bone Joint Surg Am. 2000;82(10):1398-1407. CrossRef PubMed
7. Stenvers JD. Primary frozen shoulder: a retrospective study on treatment with physiotherapy [Doctoral dissertation]: University of Groningen 1994.
8. Donatelli R, Ruivo RM, Thurner M, Ibrahim MI. New concepts in restoring shoulder elevation in a stiff and painful shoulder patient. Phys Ther Sport. 2014;15(1):3-14. CrossRef PubMed
9. Ludewig PM, Phadke V, Braman JP, Hassett DR, Cieminski CJ, LaPrade RF. Motion of the shoulder complex during multiplanar humeral elevation. J Bone Joint Surg Am. 2009;91(2): 378-389. CrossRef PubMed
10. Fayad F, Hoffmann G, Hanneton S, et al. 3-D scapular kinematics during arm elevation: effect of motion velocity. Clin Biomech (Bristol, Avon). 2006;21(9):932-941. CrossRef PubMed
11. Vermeulen HM, Stokdijk M, Eilers PH, Meskers CG, Rozing PM, Vliet Vlieland TP. Measurement of three dimensional shoulder movement patterns with an electromagnetic tracking device in patients with a frozen shoulder. Ann Rheum Dis. 2002;61(2):115-120. CrossRef PubMed
12. Fayad F, Roby-Brami A, Yazbeck C, et al. Three-dimensional scapular kinematics and scapulohumeral rhythm in patients with glenohumeral osteoarthritis or frozen shoulder. J Biomech. 2008;41(2):326-332. CrossRef PubMed
13. Stenvers JD, van Woerden HH. The primary frozen shoulder [CD-ROM]. Nederland: Onderwijscentrum NSA; 2011.
14. Veeger HE, Magermans DJ, Nagels J, Chadwick EK, van der Helm FC. A kinematical analysis of the shoulder after arthroplasty during a hair combing task. Clin Biomech (Bristol, Avon). 2006;21(suppl 1):S39-S44. CrossRef PubMed
15. Rundquist PJ, Ludewig PM. Correlation of 3-dimensional shoulder kinematics to function in subjects with idiopathic loss of shoulder range of motion. Phys Ther. 2005;85(7):636-647. CrossRef PubMed
16. Magermans DJ, Smits NC, Chadwick EK, Veeger D, van der Helm FC, Rozing PM. Discriminating factors for functional outcome after shoulder arthroplasty. A critical review of the literature. Acta Orthop Belg. 2003;69(2):127-136. PubMed
17. Yang JL, Chang CW, Chen SY, Lin JJ. Shoulder kinematic features using arm elevation and rotation tests for classifying patients with frozen shoulder syndrome who respond to physical therapy. Man Ther. 2008;13(6):544-551. CrossRef PubMed
18. Itoi E, Hsu HC, An KN. Biomechanical investigation of the glenohumeral joint. J Shoulder Elbow Surg. 1996;5(5):407-424. CrossRef
19. Ludewig PM, Reynolds JF. The association of scapular kinematics and glenohumeral joint pathologies. J Orthop Sports Phys Ther. 2009;39(2):90-104. CrossRef PubMed
20. Browne AO, Hoffmeyer P, Tanaka S, An KN, Morrey BF. Glenohumeral elevation studied in three dimensions. J Bone Joint Surg Br. 1990;72(5):843-845. CrossRef PubMed
21. Inui H, Nobuhara K. Glenohumeral relationship in maximum elevation. Surg Radiol Anat. 2014;36(8):755-761. CrossRef PubMed
22. Johnston T. The movements of the shoulder-joint: a plea for the use of the 'plane of the scapula' as the plane of reference for movements occurring at the humero-scapular joint. Br J Surg. 2005;25(98):252-260. CrossRef
23. Matsuki K, Matsuki KO, Yamaguchi S, et al. Dynamic in vivo glenohumeral kinematics during scapular plane abduction in healthy shoulders. J Orthop Sports Phys Ther. 2012;42(2): 96-104. CrossRef PubMed
24. Baettig SJ, Wieser K, Gerber C. Determinants of patient satisfaction following reconstructive shoulder surgery. BMC Musculoskelet Disord. 2017;18(1):458. CrossRef PubMed
25. Ruiz Ibán MA, Alonso Güemes S, Ruiz Díaz R, Diaz Heredia J, de Rus Aznar I, Lorente Moreno R. Assessment of isolated glenohumeral range of motion in patients with adhesive capsulitis can help predict failure of conservative treatment: a pilot study. Knee Surg Sports Traumatol Arthrosc. 2022;30(6):2099-2104. CrossRef PubMed
26. Winkel D, Aufdemkampe G, Meijer O. Nonoperative Orthopedics and Manual Therapy, Part 2/1: Diagnosis of the Extremities. Gustav Fischer Verlag; 1994.
27. Cyriax J. Textbook of orthopaedic medicine. Diagnosis of soft tissue lesion. 8th ed. Bailliere Tindall; 1982:1.
28. Norkin CC, White DJ. Measurement of joint motion: a guide to goniometry. 4th ed. FA Davis Company; 2009.
29. Fan J, Upadhye S, Worster A. Understanding receiver operating characteristic (ROC) curves. CJEM. 2006;8(1):19-20. CrossRef PubMed
30. Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. Indian Pediatr. 2011;48(4):277-287. CrossRef PubMed
31. Clarke GR, Willis LA, Fish WW, Nichols PJ. Preliminary studies in measuring range of motion in normal and painful stiff shoulders. Rheumatol Rehabil. 1975;14(1):39-46. CrossRef PubMed
32. Boone DC, Azen SP. Normal range of motion of joints in male subjects. J Bone Joint Surg Am. 1979;61(5):756-759. CrossRef PubMed
33. Warth RJ, Millett PJ. Physical examination of the shoulder. 1 ed. New York: Springer; 2015: VII, 276 p.
34. Triffitt PD. The relationship between motion of the shoulder and the stated ability to perform activities of daily living. J Bone Joint Surg Am. 1998;80(1):41-46. CrossRef PubMed
35. McClure PW, Michener LA, Sennett BJ, Karduna AR. Direct 3-dimensional measurement of scapular kinematics during dynamic movements in vivo. J Shoulder Elbow Surg. 2001 May-Jun;10(3):269-77. CrossRef. PubMed.

36. Hanney WJ, Kolber MJ, Marshall JS. The reliability of clinical measurements designed to quantify shoulder mobility. Phys Ther Rev. 2011;16(6):413-422. CrossRef
37. Baettig SJ, Wieser K, Gerber C. Determinants of patient satisfaction following reconstructive shoulder surgery. BMC Musculoskelet Disord. 2017 Nov 15;18(1):458. CrossRef. PubMed
38. Gerber C, Werner CM, Macy JC, Jacob HA, Nyffeler RW. Effect of selective capsulorrhaphy on the passive range of motion of the glenohumeral joint. J Bone Joint Surg Am. 2003;85(1): 48-55. CrossRef PubMed
39. Crétual A, Bonan I, Ropars M. Development of a novel index of shoulder's mobility based on the configuration space volume and its link to mono-axial amplitudes. Man Ther. 2015;20(3):433-439. CrossRef PubMed
40. Mallon WJ, Herring CL, Sallay PI, Moorman CT III, Crim JR. Use of vertebral levels to measure presumed internal rotation at the shoulder: a radiographic analysis. J Shoulder Elbow Surg. 1996;5(4):299-306. CrossRef PubMed

Arch Physioter 2024; 14: 56-64
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3056
ORIGINAL RESEARCH ARTICLE

First-contact physiotherapists' perceived competency in a new model of care for low back pain patients: a mixed methods study

Amélie Kechichian[1,2,3], Elsa Viain[2], Thomas Lathière[1,2], François Desmeules[4,5], Nicolas Pinsault[1,2]

[1]THEMAS Team, TIMC-IMAG Laboratory, UMR CNRS-UGA, Grenoble - France
[2]Department of Physiotherapy, University Grenoble-Alpes, Grenoble - France
[3]Interprofessional University Primary Healthcare Center, University Grenoble-Alpes, Saint-Martin d'Hères - France
[4]Maisonneuve-Rosemont Hospital Research Center, University of Montreal Affiliated Research Center, Montreal, Quebec - Canada
[5]School of Rehabilitation, Faculty of Medicine, University of Montreal, Montreal, Quebec - Canada

ABSTRACT

Background: A new advanced practice model of care enables French physiotherapists to perform medical acts for low back pain (LBP) patients as first-contact physiotherapists (FCPs).
Objective: The aim of this study is to determine the self-perceived competency of FCPs and to further explore factors underpinning this feeling.
Methods: A mixed-methods explanatory sequential design was conducted. A survey was used to self-assess the perceived competency of FCPs in performing medical tasks. Semi-structured interviews were then performed to explore determining factors of perceived competency. Inductive thematic analysis was performed.
Results: Nine FCPs answered the survey and were interviewed (mean age 40.1, standard deviation [SD]: ±10.0). FCPs felt very competent with making medical diagnosis (3.44/4, SD: ±0.53), analgesic prescription (3.11, SD: ±0.78) and referring onward to physiotherapy (3.78, SD: ±0.55). They did not feel competent with nonsteroidal anti-inflammatory drug prescription (2.78, SD: ±0.67) and issuing sick leave certificate (2.67, SD: ±1.0). The main identified influencing factors were previous FCPs' experience, training, knowledge, collaboration with family physicians, high responsibility and risk management associated with decision-making.
Conclusion: French FCPs appeared to have the necessary skills to directly manage LBP patients without medical referral. Future training focusing on analgesic prescription and issuing sick leave certificate is however needed.

Keywords: Advanced practice physiotherapy roles, First-contact physiotherapists, Medical acts, Mixed methods, Perceived competency, Training strategies

What is already known about this topic?

- *First-contact physiotherapy is an effective and emerging model of care where advanced practice physiotherapists working in family health teams diagnose and manage patients, including traditional medical acts such as autonomous prescriptions of medications.*

What does the study add?

- *French first-contact physiotherapists in this study reported feeling competent to directly manage patients without medical referral. They, however, needed further training to feel completely competent with medication prescription and issuing sick leave certificate.*

Received: February 23, 2024
Accepted: July 16, 2024
Published online: September 13, 2024

Supplementary material: Interview guide

Corresponding author:
Amélie Kechichian
email: amelie.kechichian@univ-grenoble-alpes.fr

Introduction

Musculoskeletal disorders (MSKDs) affect hundreds of millions of people around the world and can lead to temporary or lifelong disabilities and limitations in participation (1-4). Among MSKD, low back pain (LBP) is the major cause of long-term pain and disability worldwide (3,5-7). The reported lifetime prevalence of LBP is about 40% based on a survey of 54 different countries (8). In France, LBP is the second most

common reason for consulting in family practice (9). While 90% of patients recover within 4-6 weeks following the first pain onset, chronic LBP is the third leading cause of disability and the first cause of occupational disability before the age of 45 in France (9). The early identification and management of patients at risk of poor prognoses represents a major challenge for the healthcare system.

Primary healthcare services in France are reaching a saturation point, and patients are experiencing important delays to access care (10). Considering the aging population and the increasing shortage of physicians, family physicians' burden is expected to continue to increase in the coming decades (3,11). To offer better access and to help reduce physicians' workload, new collaborative care pathways entitled "cooperation protocols" are emerging in French multidisciplinary primary healthcare centers. These models emphasize more autonomous roles for nonmedical healthcare practitioners using task shifting within family healthcare teams (12).

One of these models involves physiotherapists for the management of acute LBP patients. Since the publication of the official legislative text in 2020, the initial LBP consultations can be transferred from family physicians to physiotherapists working within the same multidisciplinary healthcare center (13). Eligible patients aged 20 to 55 years suffering from acute LBP may consult the physiotherapist instead of the family physician. This model expands the usual scope of practice of French physiotherapists, allowing them to work as first-contact practitioners in advanced practice roles (14,15). As described in the United Kingdom, first-contact physiotherapy (FCP) is an emerging advance physiotherapy practice model of care where physiotherapists working in family health teams diagnose and manage patients while that may include traditional medical acts such as autonomous prescriptions of medications (14-16). In the French model, the FCP's role is to diagnose LBP, issue medical sick leave certificates, prescribe low-class analgesic medications (paracetamol or oral anti-inflammatory drugs) and refer patients for additional outpatient physiotherapy in another place if required. This registered healthcare pathway is nonetheless still coordinated by family physicians. The involved FCPs ensure that any necessary information regarding the medical management of patients is accurately conveyed to family physicians (13). Such pathways do match the globally accepted definition of advanced physiotherapy practice models and represents a significant change, as patients in France are traditionally referred by family physicians to the physiotherapist who are not autonomous first-contact providers (17,18).

Our team previously conducted a study regarding physiotherapists' and family physicians' acceptability of this new model prior to its implementation (19). The results highlighted a positive perception of physiotherapists' competencies and skills to adequately manage patients with LBP from the physician's point of view. This study also reported that before the implementation of the FCP model, physiotherapists did not unanimously feel confident in their ability to perform medical tasks, especially regarding the prescription of oral nonsteroidal anti-inflammatory drugs (NSAIDs) or issuing sick leave certificates (19). Family physicians and physiotherapists who finally set up the new FCP model received a 10-hour interprofessional training. The goal of this training was to enable physiotherapists to acquire the competencies for managing patients with LBP as primary contact practitioners, to acquire adequate competencies for red flag identification and patients' referral to physicians, to prescribe appropriate medication and issue sick leave certificate, as well as suitable referral for additional outpatient physiotherapy. Exploring the acquisition of these advanced competencies by physiotherapists working within the new FCP model both helps ensure the quality and safety of this new model and enables a better tailoring of the training provided to physiotherapists.

A successful FCP advanced practice role requires a combination of competencies and skills that can be shaped by perceived self-efficacy (20,21). The self-efficacy theory was developed by Bandura and is defined as an individual's belief in his ability to succeed in a specific task or situation (20). It has been identified as the strongest predictor of clinical performance (22-24). Previous clinical performance experience is one of the principal sources of influence for self-efficacy (20). The French physiotherapists' confidence in performing medical tasks has been evaluated prior to the implementation of the new pathway and we assumed that the said confidence could have changed with working overtime in this new advanced model of care (19,20,25). Given the potential of evolution of the French physiotherapists toward more autonomous advanced practice roles, there is a need to document their acquisition of advanced competencies and skills. The aims of this study are therefore to determine the self-perceived competency of FCPs in their advanced practice role for LBP patients and to further explore factors that influence such perceptions.

Methods

Design

We used a mixed-methods explanatory sequential design to address the research aims. The explanatory sequential design provided the opportunity to collect initially emerging insights from cross-sectional quantitative data and help further explain the results through semi-structured interviews (26,27). This design enabled us to combine both quantitative and exploratory qualitative data so as to provide a deeper insight into how physiotherapists perceive their ability to perform the aforementioned medical tasks (28,29).

Measures

Self-efficacy measure

There is no published instrument to measure healthcare professionals' perceived competency in performing shifted or delegated medical tasks. We therefore designed a tool to measure this construct. This tool took the form of a survey. Its development was guided by Bandura's theory on self-efficacy scale construction guidelines and previous similar studies evaluating self-efficacy and healthcare professionals' perceived competencies using mixed-methods study designs (30-32). We first identified the five medical tasks performed

by physiotherapists in the new pathway that were not part of their usual scope of practice: medical diagnosis, analgesic prescription (acetaminophen), NSAID prescription, sick leave certificate issuance and outpatient physiotherapy referral prescription. The identification of red and yellow flags, respectively signs and symptoms of serious pathologies and psychosocial risk factors for a poor prognosis, was also added to the items. Although red and yellow flags are examined by physiotherapists when receiving patients referred by family physicians, a deeper consideration needs to be given to these tasks in a primary contact role.

The tool was composed of seven items. Each item of the survey assessed one task: medical diagnosis, analgesic prescription (acetaminophen), NSAID prescription, sick leave certificate issuance, outpatient physiotherapy referral prescription, red flag identification, yellow flag identification. The items consisted of a 4-point Likert-type rating scale, ranging from 1 (*not at all competent*) to 4 (*extremely competent*) to self-assess the perceived level of competency of physiotherapists in performing the identified tasks.

Interview guide

Following a review of relevant literature, an initial semi-structured interview guide was developed by one author (E.V.) and completed by a second author (A.K.). Adaptations were made based on the second author's feedback. The interview guide aimed to explore the determining factors of FCPs' perceived competency regarding each task identified in the survey. The interview guide focused on FCPs' experiences and perceptions regarding the activities they carried out, factors that positively or negatively influenced their perceived competency and potential evolutions for the new model of care. Relevant literature and the Consolidated Criteria for Reporting Qualitative Research (COREQ) were considered in the designing of the guide and results' reporting, to ensure the findings' credibility and transparency (33-35).

Participants

To be included, physiotherapists had to work in a multidisciplinary primary healthcare center in France that had set up the FCP advanced model for acute LBP patients' care, having completed the required interprofessional training and having taken care of at least one LBP patient within the FCP pathway. The study was conducted between January and March 2023, one year after the implementation of the model in the primary healthcare centers.

Because of the barriers to the implementation of the model we previously identified in an acceptability study, we anticipated a low deployment of the FCP model in France and thus a relatively small sample size for both qualitative and quantitative steps (19). Efforts were made by the researchers so that all potential participants who met the inclusion criteria in France were contacted. All eligible and voluntary participants were included in the study. All included participants took part in both quantitative and qualitative components of the study using an identical sample strategy for sequential design (29).

Procedures

Participants were identified through the research team's network, by contacting the regional health agencies in each region of France and through the French federations for multidisciplinary primary healthcare centers. Potential participants were contacted by email. The email detailed the aim of the study and mentioned the voluntary participation of physiotherapists. Voluntary participants were asked to complete the informed consent through an electronic standardized form before each interview. An email including a link to complete the online survey was sent to the participants using *LimeSurvey*, a web platform secured by data encryption protocol and hosted by the Grenoble-Alps University server. Individual interviews were conducted virtually (Zoom) by the same research assistant that made initial contact with participants when the online survey was completed in the same day the participant answered the questionnaire. The research assistant used active listening techniques. She did not conduct previous interviews but had a formal university training in qualitative methodology of approximately 10 hours taught by the Physiotherapy Department of Grenoble-Alpes University (34). To profile interviewees, participants' demographic characteristics were collected prior to the interview.

Data analysis

The survey data were anonymized and transferred into a Microsoft Excel spreadsheet. Descriptive analysis was performed for all quantitative data.

Interviews were audio-recorded and transcribed verbatim immediately after the interview. Transcripts were anonymized. Based on the Braun and Clarke process, a thematic analysis of the interviews was performed by the research assistant who conducted the interviews (E.V.) and a physiotherapist researcher (A.K.) (35). QCAmap software was used for this analysis. Both researchers familiarized themselves with the transcripts and independently set up an initial set of codes for the first two interviews using an iterative approach. Discrepancies between the two code sets were reviewed and a final set of codes was defined. The final code set was then applied by one researcher (E.V.) to the seven remaining interviews. Final themes were identified following ongoing critical discussion between researchers (E.V. and A.K.) until a consensus was reached. Throughout the process, data transferability was ensured by documenting the context of the fieldwork so that another reader would be able to decide whether the findings could be applied to another setting (36). It was a major focus that the findings emerged from the data and not from the researcher's perception to ensure their reliability (36).

Results

Participants' description

Nine physiotherapists were included in the study for both quantitative and qualitative data collection (mean age 40.1, standard deviation [SD]: ±10.0). One physiotherapist declined

TABLE 1 - Characteristics of the participants (n = 9)

Physiotherapist	Age	Year of graduation	Experience in multidisciplinary healthcare center (years)	FCP model training duration (hours)	Experience with the FCP model (months)	Number of LBP patients managed in the FCP model
PT1	28	2016	3	6	8	1
PT2	57	1989	10	10	8	5
PT3	34	2010	7	5	10	8
PT4	41	2005	8	10	7	4
PT5	46	1998	2	10	10	3
PT6	36	2008	2	10	10	4
PT7	26	2020	1	10	11	5
PT8	45	2001	9	4	3	5
PT9	48	1998	6	10	2	1

FCP = first-contact physiotherapist; LBP = low back pain; PT = physioterapist.

to participate because she was not available for an interview during the study period. Participants' mean experience duration with the FCP pathway was 7.6 months (SD: ±3.2). FCPs had managed one to eight patients within the model of care prior to this study (mean: 4.0, SD: ±2.2). Characteristics of the participants are presented in Table 1.

Self-perceived competency measure

FCPs felt very competent with making medical diagnosis (3.44/4, SD: ±0.53), analgesic prescription (3.11, SD: ±0.78) and referring onward to another physiotherapist for further rehabilitation (3.78, SD: ±0.55). They did not feel competent with NSAID prescription (2.78, SD: ±0.67) and sick leave certificate issuance (2.67, SD: ±1.0). Results of the questionnaire are presented in Table 2.

TABLE 2 - FCP self-perceived competency in performing tasks in the new model of care (n = 9)

How competent do you feel when performing the following medical tasks?*	Mean (SD)	Min-Max	Median
Red flag identification	3.33 (0.71)	2.0–4.0	3.0
Yellow flag identification	3.22 (0.67)	2.0–4.0	3.0
Making a medical diagnosis	3.44 (0.53)	3.0–4.0	3.0
Analgesic prescription	3.11 (0.78)	2.0–4.0	3.0
NSAID prescription	2.78 (0.67)	2.0–4.0	3.0
Sick leave certificate issuance	2.67 (1.0)	1.0–4.0	3.0
Physiotherapy referral	3.78 (0.55)	3.0–4.0	4.0

FCP = first-contact physiotherapist; NSAID = nonsteroidal anti-inflammatory drug; SD = standard deviation.
*1—not at all competent, 2—not very competent, 3—very competent, 4—extremely competent.

Qualitative interviews analysis

Nine semi-structured interviews were conducted to allow a better understanding of FCPs' perceived competency, influencing factors and readiness to practice in the new FCP model of care. Four main themes were identified: (1) experiences, knowledge, and training are determining factors of FCPs' perceived competency; (2) collaboration with family physicians seems to favor FCPs' perceived competency; (3) higher responsibility and risk management may be associated with lower perceived competency; and (4) formal training and modification of the FCP model could improve FCPs' perceived competency.

Theme 1: Experience, knowledge and training are determining factors of FCPs' perceived competency

Similarities with their usual scope of practice increase FCPs' competency

Previous experience related to the usual scope of practice of FCP was mainly reported as a major influencing factor of perceived competency for the participants. When medical shifted tasks were quite close to the physiotherapist's everyday tasks, their feeling of competency was reported as high as reported for diagnosis: " I'm quite confident, I'm not very worried about diagnostic errors, it's part of my everyday job" (PT4); "I actually feel even more competent than a physician in the diagnosis of low back pain" (PT3).

Regarding red and yellow flag identification, participants also attributed their high perceived competency to their clinical experience: "Given the experience I have with low back pain patients, and within one hour of interviewing and consulting, I feel there are many things I'm capable of identifying" (PT3); "This flag system […], we use that every day" (PT8).

Regarding physiotherapy referrals, participants stated that "My experience and initial assessment allow me to know easily whether or not it is relevant to prescribe further physiotherapy sessions to patients" (PT3); "With the experience I have, I know if the patient needs physiotherapy and a follow-up consultation or not" (PT9).

When evaluating the ability to return to work of patients with LBP, one participant reported being used to "assessing biomechanical factors, psychological factors and deciding whether or not they are compatible with work on a given day" (PT4). Inexperience was mentioned as a factor for a lower self-efficacy associated with NSAID prescription: "we tend, as physiotherapists, to tell patients to take paracetamol to ease the pain, whereas we rarely recommended NSAIDs to our patients, […] it's something we never did before" (PT8).

FCPs' lack of experience with the new model of care

Participants reported that the experience acquired with the FCP pathway contributed to determine their confidence in performing medical tasks. For most of the physiotherapists, the lack of exposure to clinical consultations in the FCP model resulted in a low perceived competency regarding tasks that differed highly from their scope of practice, even if they did not consider the tasks to be complex or challenging "I haven't done it enough [NSAI drugs prescriptions] to feel comfortable with it yet" (PT4); "Regarding drug prescription, it's just a lack of practice in my opinion" (PT5); "Clearly, my experience is growing, ehm, to shift from rather competent to fully competent, that's it" (PT9).

Knowledge and training for medication prescription are insufficient

Participants expressed concern with insufficient knowledge and training regarding the analgesic use and oral NSAID contraindications: "I am not trained with regards to the very developed pharmacopoeia" (PT3); "I don't know the exact nature of the substances I prescribe" (PT9); "Well, there certainly are other more important contraindications to NSAIDs […] that I don't know of" (PT4); "I am clearly not trained enough regarding pharmacological interactions" (PT6). One participant however expressed "Because it was taught during the training, I feel rather competent" (PT7).

Theme 2: Collaboration with family physicians seems to favor FCPs' perceived competency

Interprofessional collaboration fosters FCPs' perceived competency

Collaboration with family physicians was explicitly identified by FCPs as a facilitator impacting positively their feeling in FCP model of care: "I find it quite stressful if the physicians aren't next door" (PT6); "The discussion, the coordination with physicians is very easy. I feel competent because I dare to go ask for information if there is an issue" (PT6).

Some participants expressed the need to be further supervised and to receive additional feedback from family physicians: "I think it could comfort me on whether I made the right choice or not, if the physician tells me I did right, whether there is a sick leave or not" (PT9).

FCPs and family physicians cope with common challenges

Several participants felt reassured knowing that family physicians encounter similar difficulties with decision-making for sick leave certificates and medication prescription: "There is a similar difficulty, that's shared with the physicians, because they go through the same thing"; "they says themselves that they do this approximately, a bit roughly and very much depending on the patients' requests" (PT3); "Physicians are no more competent than we are, in their capacity to know whether or not they should prescribe one or the other, and at which dosage" (PT3); "Even for physicians it is not always clear and they hesitate" (PT5).

Theme 3: Higher responsibility and risk management may be associated with lower perceived competency

Perceived competency is influenced by the level of risk and responsibility

According to most participants, the perceived level of competency with the new medical tasks was reported to be associated with the perceived level of risk when performing the task: "I can never declare myself to be competent because I think we are given an important, a huge responsibility" (PT6); "There are other risks so I'm always a little bit afraid of making a mistake and missing something, of not asking the patients the right question" (PT7).

Low risk associated with inappropriate sick leave certificate issuance seemed to favor a higher level of confidence for FCPs. However, the undesirable effects and potential contraindications associated with NSAIDs use were associated with lower confidence of participants: "Well, I feel that I am not competent enough on the matter, to clearly know if I haven't missed a contraindication" (PT6); "There is an additional apprehension regarding NSAIDs because […] there are more potential consequences" (PT5).

The physiotherapists stated that "additional responsibility" (PT4) associated with "the risk of missing something serious" (PT7) was a barrier to feeling fully competent with their new advanced roles.

Clear guidelines may facilitate clinical decision-making

Participants reported that they would feel more confident in their clinical decision-making process if clear guidelines were available. Regarding the duration of sick leave and analgesic dosage, participants expressed a lack of formal recommendations leaving them with the following questions: "Why do I prescribe a one-day sick leave, why three? Why five?" (PT8); "What is the right dosage for pain killers or NSAIDs?" (PT3).

The FCP model however provided participants clearer recommendations regarding additional physiotherapy referral: "The decision criteria to decide whether or not we prescribe rehabilitation […] Actually they are defined clearly enough so that I can settle on whether or not I prescribe it" (PT3).

Theme 4: Formal training and modification of the FCP model could improve FCPs' perceived competency

Formal educational training is needed

Participants believed that additional extensive educational training about pharmacological prescription and drug safety use is needed to help them become more confident with the prescription of oral analgesic and NSAIDs: "I think we need more formal training on pharmacological matters. I talked about it with family physicians and pharmacists, I took some medical courses, but I did not get exhaustive training on that topic" (PT4). One participant suggested that this training should be associated with regular clinical case presentations so that physiotherapists could update their knowledge and skills.

The FCP model framework should be more flexible

Even if the framework for the FCP model of care was reported to facilitate and help decision-making by most of the participants, one of them felt that the model of care framework definition (eligible patients, allowed new clinical roles and applicability) interfered with his clinical reasoning process: "This model does not require clinical skills […] we only need to answer specific questions and tick boxes, it does not let us think and use our clinical judgment" (PT3). He suggested that this framework should be modified to enable more flexibility, to allow more autonomy for physiotherapists to use their clinical judgment.

Discussion

Main findings

The aim of this study was to determine the FCPs' perceived competency in their first-contact practitioner's role for LBP patients and to further explore factors underpinning these perceptions.

One of the key findings of our study is that physiotherapists felt very, or extremely competent in identifying red and yellow flags and diagnosing acute LBP. Red and yellow flag identification should already be part of the French physiotherapists' practice, thus making this result not all that surprising. However, as physiotherapists usually work based on physician's prescription, they may consider that the identification of red flags has been already done by the physician. It is therefore important to ensure that this skill is mastered in the context of the new FCP. Regarding acute LBP diagnosis, our result is a more significant finding since making a diagnosis is a restricted act that only licensed physicians in France can perform (37). This result shows that physiotherapists, in their advanced practice roles, consider that they have the required skills to adequately determine the condition of LBP patients, and manage them as primary contact practitioners (38-40). Clinical reasoning and differential diagnosis training in the undergraduate training for French physiotherapists is now integrated in several programs (21). This finding is also consistent with other international studies showing that physiotherapists can manage patients with MSKD as primary contact practitioners, or in advanced practice roles, without an increase in adverse events (38-40). The factor that appears to contribute to the physiotherapists' high perceived competency regarding making LBP diagnosis is a previous clinical experience with LBP management. Physiotherapists do routinely look for signs and symptoms of serious pathology in spinal pain patients, even when referred by family physicians. Referring patients to family physicians when they suspect serious pathology is already part of their usual practice, and was therefore not considered as a significant change. This result is also consistent with a previous qualitative study conducted by our team showing that patients in this new model were receptive with being managed autonomously by FCP and were highly confident in the FCPs' ability to perform delegated medical tasks including making a medical diagnosis (41).

All the participants were confident in their ability to adequately refer patients to additional physiotherapy sessions when needed. Participants considered the decision-making about the need for further physiotherapy within their scope of practice. This result is consistent with a previous study conducted in the French context showing that physiotherapists were more likely to confirm their choice of beginning physiotherapy treatments and the physiotherapy approaches they used for evidence-based recommendations for LBP patients' care compared to family physicians' prescribed treatments (42). This study also reported that information required for the referral of patients to physiotherapy by French family physicians was often incomplete (42). Our results strengthen the emerging evidence that French physiotherapists have the adequate skills to independently and directly manage LBP patients including initial diagnosis and decision on further physiotherapy referral.

Another important finding was that participants mostly felt competent with analgesic prescription but expressed being somewhat uncomfortable with oral NSAID prescription. This result is in line with our previous acceptability study that showed a lower level of confidence of physiotherapists and family physicians in the physiotherapists' ability to adequately and safely prescribe oral NSAIDs (19). Other results did not differ between the two studies regarding flags' identification and physiotherapy referral, showing that professionals' perceptions before the implementation of the model were in line with their later feelings (19).

According to the participants, oral NSAID prescription is associated with higher risks and responsibilities because of contraindications and the potential adverse events associated with their use. A lack of knowledge and training regarding medication prescription was suggested as a factor for the participants' low perceived competency. Then, additional training and extensive focus on pharmacological issues should be further considered to strengthen the confidence level of physiotherapists in this advanced practice role. The said training should include clinical practice guidelines on NSAID use, as previous studies have already showed that poor familiarity with these guidelines could explain poor provider adherence (43,44). Another qualitative study conducted in the United Kingdom demonstrated that a clear understanding of responsibility associated with medical tasks is required to further support the deployment of FCP (21). The United

Kingdom developed a national competency framework for FCPs and these roles are developing well (45). The extensive training of French physiotherapists working as FCPs should therefore consider international resources.

Regarding sick leave certificate issuance, the participants' perception varied greatly. For some participants, the assessment of patients' working constraints was already part of their usual practice. For others, the unfamiliar administrative procedure required for issuing sick leaves reduced their perceived level of competencies. According to them, the additional exposure to clinical situations could improve their level of competency. This is consistent with Bandura's theory, which outlined that the repetition of previously successful tasks is more likely to strengthen self-efficacy, whereas lack of exposure or failure may weaken self-efficacy (20). The issuance of sick leave certificate by physiotherapists could be an effective strategy to alleviate medical workload but physiotherapists need to have an extensive training to do so efficiently (19).

Studies about clinical self-efficacy in advanced practice roles have been previously conducted for other healthcare practitioners, such as nurses (46,47). One study showed that peer learning and realistic simulation could result in a positive impact on nursing student's self-efficacy when working in advanced practice roles (47). Future research in advanced practice physiotherapy should focus on the efficacy of learning strategies to maximize skill and competency acquisition regarding medication prescription and sick leave issuance to ensure safe and high-value quality care for patients.

Strengths and limitations

This study is the first to evaluate physiotherapists' self-perceived competency in their first-contact roles in a new LBP advanced practice role in primary care. The mixed methods provided a quantitative perspective to determine FCPs' perceived competency, and the qualitative analysis allowed a deeper exploration of factors that influence such perceptions. Consolidated Criteria for Reporting Qualitative Research (COREQ) were considered in the design of the guide and reporting of the results. Throughout the results, quantitative and qualitative data are consistent. The verbatims clearly reflect a higher feeling of competency for some acts and low for others, in the same way as the quantitative measures do. It reinforces the internal validity of the results.

Some limitations must be considered when interpreting the results. At the time of the study, this new FCP model of care had been deployed in a limited number of primary healthcare centers in France. Only a limited number of physiotherapists working in first-contact roles could therefore be recruited in the study. In the field of implementation research, a multistage strategy for mixed-methods studies should include a purposeful sampling beginning with a quantitative broader view that emphasizes data variation and dispersion, moving then to a narrow qualitative view focusing on similarity or central tendencies (28). Such strategy is recommended to find the optimal balance between internal and external validity of the findings (28). Due to the small number of physiotherapists meeting our inclusion criteria in France, we were unable to recruit a large sample of participants in the first quantitative step of our study that could have provided a broader view of FCPs' perceived competency in France. However, we tried to recruit all voluntary and eligible participants across the country. Findings that were analyzed in our study provided a narrow depth and understanding of FCPs' perceived competency in the French context. They may not be generalizable to all French physiotherapists or to FCPs in other countries. Indeed, the FCP model of care developed by the French authorities slightly differs from the formal international advanced practice physiotherapy models that already exist in several countries worldwide. Our findings may differ from other international contexts, training and practice frameworks.

Conclusion

The overall findings of this study suggest that physiotherapists working as first-contact practitioners in this new model of care in French primary care had a high self-perceived competency when diagnosing LBP and referring patients to additional outpatient physiotherapy care. They however felt less competent with medication prescription and sick leave issuance. The most influential reported factors for FCPs' perceived competency in medical tasks were previous FCPs' experience, training and knowledge, collaboration with family physicians, high responsibility and risk management associated with decision-making.

Our results help the emerging evidence suggesting that French physiotherapists have the necessary skills to directly manage LBP patients without medical referral. Future training focusing on analgesic drug prescription and sick leave certificate issuance is however needed to support physiotherapists' perceived competency in their advanced practice roles. Thus, further research should aim to investigate the most effective training approach to enhance FCPs' perceived competency in performing these medical tasks. Additionally, as the self-efficacy has been identified as the strongest predictor of clinical performance in various healthcare contexts and is therefore linked to quality of patients' care, further research should deeply explore the impact of self-perceived competency on the clinical performance of FCPs in medical acts.

Acknowledgments

This study could not have been completed without the precious collaboration of both URPS-MK ARA and CNO-MK, as well as all the participants who completed this survey.

Disclosures

Conflict of interest: The authors declare no conflicts of interest with respect to the research, authorship and/or publication of this article.

Financial support: This project was funded to A.K. by the "Union Régionale des Professionnels de Santé Masseur-Kinésithérapeutes" of the Auvergne-Rhône-Alpes region in France and the "Conseil National de l'Ordre des Masseurs-Kinésithérapeutes" as part of a doctoral. The funding bodies were not involved in the design of the study, data collection, analysis, interpretation of data and writing the manuscript. No award/grant number.

Ethical approval: The study has been conducted in accordance with the Declaration of Helsinki. Ethical approval was obtained from the French ethics committee: *Comité de Protection des Personnes Ouest*

IV Nantes (*Committee for the protection of subjects West Nantes IV*, number 21.01537.000012).

Consent to participate: All participants offered written informed consent prior to enrolment in the study.

Data availability statement: The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

References

1. World Health Organization. Musculoskeletal health 2021. Online. Accessed February 2024.
2. Cieza A, Causey K, Kamenov K, Hanson SW, Chatterji S, Vos T. Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2021;396(10267):2006-2017. CrossRef PubMed
3. Blyth FM, Briggs AM, Schneider CH, Hoy DG, March LM. The global burden of musculoskeletal pain—where to from here? Am J Public Health. 2019;109(1):35-40. CrossRef PubMed
4. Woolf AD, Erwin J, March L. The need to address the burden of musculoskeletal conditions. Best Pract Res Clin Rheumatol. 2012;26(2):183-224. CrossRef PubMed
5. 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
6. Mattiuzzi C, Lippi G, Bovo C. Current epidemiology of low back pain. J Hosp Manag Health Policy. 2020;4:15. CrossRef
7. Buchbinder R, van Tulder M, Öberg B, et al; Lancet Low Back Pain Series Working Group. Low back pain: a call for action. Lancet. 2018;391(10137):2384-2388. CrossRef PubMed
8. Hoy D, Bain C, Williams G, et al. A systematic review of the global prevalence of low back pain. Arthritis Rheum. 2012;64(6):2028-2037. CrossRef PubMed
9. Haute Autorité de Santé. Management of the patient with common low back pain. Good clinical practice recommendation. Saint-Denis La Plaine 2019. Online. Accessed February 2024.
10. Lapinte A, Legendre B. Renunciation of care: low density is an aggravating factor for poor people. Studies and results. DREES 2021; 1200. Online. Accessed February 2024.
11. Briggs AM, Woolf AD, Dreinhöfer K, et al. Reducing the global burden of musculoskeletal conditions. Bull World Health Organ. 2018;96(5):366-368. CrossRef PubMed
12. Haute Autorité de Santé. Cooperation between health professionals: 2010-2018 report. 2018. Online. Accessed February 2024.
13. Ministère des Solidarités et de la Santé. Decree of the 6th of March 2020 about the authorisation of the cooperation protocol "Management of acute low back pain of less than 4 weeks by the physiotherapist in the context of a multi-professional healthcare center". 2020. Online. Accessed February 2024.
14. Donato EB, DuVall RE, Godges JJ, Zimmerman GJ, Greathouse DG. Practice analysis: defining the clinical practice of primary contact physical therapy. J Orthop Sports Phys Ther. 2004;34(6):284-304. CrossRef PubMed
15. Halls S, Thomas R, Stott H, et al. Provision of first contact physiotherapy in primary care across the UK: a survey of the service. Physiotherapy. 2020;108:2-9. CrossRef PubMed
16. Chartered Society of Physiotherapy. First contact physiotherapy posts in general practice: a guide for implementation in England. Chartered Society of Physiotherapy 2018. Online. Accessed February 2024.
17. WCPT. Advanced physical therapy practice : policy statement. 2019. Online. Accessed February 2024.
18. Dowling M, Beauchesne M, Farrelly F, Murphy K. Advanced practice nursing: a concept analysis. Int J Nurs Pract. 2013;19(2):131-140. CrossRef PubMed
19. Kechichian A, Desmeules F, Girard P, Pinsault N. Acceptability of a task sharing and shifting model between family physicians and physiotherapists in French multidisciplinary primary healthcare centres: a cross-sectional survey. Fam Med Community Health. 2022;10(2):e001644. CrossRef PubMed
20. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev. 1977;84(2):191-215. CrossRef PubMed
21. Langridge N. The skills, knowledge and attributes needed as a first-contact physiotherapist in musculoskeletal healthcare. Musculoskelet Care. 2019;17(2):253-260. CrossRef PubMed
22. Robbins SB, Lauver K, Le H, Davis D, Langley R, Carlstrom A. Do psychosocial and study skill factors predict college outcomes? A meta-analysis. Psychol Bull. 2004;130(2):261-288. CrossRef PubMed
23. Opacic DA. The relationship between self-efficacy and student physician assistant clinical performance. J Allied Health. 2003;32(3):158-166. PubMed
24. Jones A, Sheppard L. Self-efficacy and clinical performance: a physiotherapy example. Adv Physiother. 2011;13(2):79-83. CrossRef
25. Greenhalgh S, Selfe J, Yeowell G. A qualitative study to explore the experiences of first contact physiotherapy practitioners in the NHS and their experiences of their first contact role. Musculoskelet Sci Pract. 2020;50:102267. CrossRef PubMed
26. Creswell JW, Clark VLP. Designing and conducting mixed methods research. Sage Publications 2017.
27. Creswell JW, Hirose M. Mixed methods and survey research in family medicine and community health. Fam Med Community Health. 2019;7(2):e000086. CrossRef PubMed
28. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health. 2015;42(5):533-544. CrossRef PubMed
29. Onwuegbuzie AJ, Collins KM. A typology of mixed methods sampling designs in social science research. Qual Rep. 2007;12(2):281-316.
30. Bandura A. Guide for constructing self-efficacy scales. In: Pajares F, Urdan T, eds. Self-efficacy beliefs of adolescents. Vol 5. Greenwich, CT: Information Age Publishing 2006; 307-337.
31. Jones A, Ingram ME, Forbes R. Physiotherapy new graduate self-efficacy and readiness for interprofessional collaboration: a mixed methods study. J Interprof Care. 2021;35(1):64-73. CrossRef PubMed
32. Sonnenberg LK, Pritchard-Wiart L, Hodgson CS, Yu Y, King S. Assessment of resident physicians' communicator and collaborator competencies by interprofessional clinicians: a mixed-methods study. Teach Learn Med. 2017;29(4):392-401. CrossRef PubMed
33. 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
34. DeJonckheere M, Vaughn LM. Semistructured interviewing in primary care research: a balance of relationship and rigour. Fam Med Community Health. 2019;7(2):e000057. CrossRef PubMed
35. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3(2):77-101. CrossRef

36. Shenton AK. Strategies for ensuring trustworthiness in qualitative research projects. Educ Inf. 2004;22(2):63-75. CrossRef
37. Bunge MA. Medical philosophy: conceptual issues in medicine. World Scientific Publishing Company 2013. CrossRef
38. Demont A, Bourmaud A, Kechichian A, Desmeules F. The impact of direct access physiotherapy compared to primary care physician led usual care for patients with musculoskeletal disorders: a systematic review of the literature. Disabil Rehabil. 2021 Jun;43(12):1637-1648. CrossRef PubMed
39. Ojha HA, Snyder RS, Davenport TE. Direct access compared with referred physical therapy episodes of care: a systematic review. Phys Ther. 2014;94(1):14-30. CrossRef PubMed
40. Babatunde OO, Bishop A, Cottrell E, et al. A systematic review and evidence synthesis of non-medical triage, self-referral and direct access services for patients with musculoskeletal pain. PLoS One. 2020;15(7):e0235364. CrossRef PubMed
41. Kechichian A, Pommier D, Druart L, Lowry V, Pinsault N, Desmeules F. "Cooperation between physicians and physios fosters trust you know": a qualitative study exploring patients' experience with first-contact physiotherapy for low back pain in French primary care. BMC Prim Care. 2024;25(1):69. CrossRef PubMed
42. Demont A, Benaïssa L, Recoque V, Desmeules F, Bourmaud A. Spinal pain patients seeking care in primary care and referred to physiotherapy: a cross-sectional study on patients characteristics, referral information and physiotherapy care offered by general practitioners and physiotherapists in France. PLoS One. 2022;17(9):e0274021. CrossRef PubMed
43. Abraham NS, El-Serag HB, Johnson ML, et al. National adherence to evidence-based guidelines for the prescription of nonsteroidal anti-inflammatory drugs. Gastroenterology. 2005; 129(4):1171-1178. CrossRef PubMed
44. Cavazos JM, Naik AD, Woofter A, Abraham NS. Barriers to physician adherence to nonsteroidal anti-inflammatory drug guidelines: a qualitative study. Aliment Pharmacol Ther. 2008; 28(6):789-798. CrossRef PubMed
45. Chance-Larsen K, Backhouse MR, Collier R, et al. Developing a national musculoskeletal core capabilities framework for first point of contact practitioners. Rheumatol Adv Pract. 2019;3(2):rkz036. CrossRef PubMed
46. Hayes E. Mentoring and self-efficacy for advanced nursing practice: a philosophical approach for nurse practitioner preceptors. J Am Acad Nurse Pract. 1998;10(2):53-57. CrossRef PubMed
47. Kirkpatrick A, Ball S, Connelly S, et al. Intraprofessional simulation's impact on advanced practice and baccalaureate student self-efficacy. Clin Simul Nurs. 2018;16:33-39. CrossRef

Arch Physioter 2024; 14: 65-69
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3128
VIEWPOINT

Treatment fidelity in clinical trials

Chad E. Cook[1,2,3], Bryan O'Halloran[4], Steve Karas[5], Mareli Klopper[6], Jodi L. Young[7]

[1]Department of Orthopaedics, Duke University School of Medicine, Duke University, Durham, North Carolina - USA
[2]Department of Population Health Sciences, Duke University, Durham, North Carolina - USA
[3]Duke Clinical Research Institute, Duke University, Durham, North Carolina - USA
[4]Division of Physical Therapy, Medical University of South Carolina, Charleston, South Carolina - USA
[5]Department of Physical Therapy, Chatham University, Pittsburgh, Pennsylvania - USA
[6]Department of Physical Therapy, Graceland University, Independence, MO - USA
[7]Department of Physical Therapy, Bellin College, Green Bay, Wisconsin - USA

ABSTRACT

In the context of clinical trials, treatment fidelity (TF) has traditionally referred to the extent to which an intervention or treatment is implemented by the clinicians as intended by the researchers who designed the trial. Updated definitions of TF have included an appropriate design of the intervention that was performed in a way that is known to be therapeutically beneficial. This requires careful attention to three key components: (1) protocol and dosage adherence, (2) quality of delivery, and (3) participant adherence. In this viewpoint, we describe several cases in which TF was lacking in clinical trials and give opportunities to improve the deficits encountered in those trials. We feel that along with quality, risk of bias, and certainty of evidence, TF should be considered an essential element of the veracity of clinical trial.

Keywords: Clinical trials, Fidelity, Quality, Treatment

Introduction

How many times have you read a research study and either: (1) had no idea what the treatment intervention consisted of; or (2) realized that the "intervention" that was used in the study was nothing like what you would apply in clinical practice? If you've encountered these two situations while reading literature, you may have been witness to limitations of treatment fidelity (TF).

Despite its importance, TF is often poorly reported in clinical trials (1-3). This is especially the case in behavioral-based studies that require some degree of clinician interpretation of the patient's progress and a modification based on that interpretation (4). It may also be because the definition of TF can vary across studies and contexts. Although TF generally refers to the extent to which an intervention is delivered as intended, ensuring consistency and reliability, terms such as "adherence," "integrity/veracity," or "implementation fidelity" are commonly used, which may not be anchored to the same underlying concept.

In this viewpoint, we focus on perspectives that have a "clinical context" (with a goal of improving clinician interpretation of TF) and provide a modern definition of TF, by describing two key components of TF (adherence and veracity), discuss examples in the literature in which TF was lacking, and provide methods to improve the implementation of interventions in clinical trials. We hope to show that in addition to commonly measured constructs such as quality, risk of bias, and certainty of evidence, TF should be assessed when interpreting the meaningfulness of a clinical trial.

Received: May 13, 2024
Accepted: August 18, 2024
Published online: September 16, 2024

Corresponding author:
Chad E. Cook
email: chad.cook@duke.edu

A modern definition of TF

In the context of clinical trials, TF has historically referred to the extent to which an intervention or treatment adheres to the implementation parameters intended by the researchers who designed the trial (5). Indeed, appropriate implementation is critical as TF is essential in ensuring that the results of the trial accurately reflect the treatment effects of the intended intervention, with no additions or omissions. Adequate TF improves one's interpretation of the outcome data in research studies, improves the likelihood of reproducibility (if studied again), and is essential for clinical translation (5,6). This demands appropriate reporting of treatment structure used in the trial. Perhaps most importantly, TF is one of the few elements in a clinical trial that equally represents components of internal and external validity (5).

Adherence of TF routinely measures protocol and dosage adherence. Adherence can be considered as "did the researchers do as they indicated they would do?" Protocol and dosage adherence reflect the extent to which the intervention was delivered as planned. It involves an assessment of whether the treatment protocol was followed closely, including the dosage, frequency, and duration of the intervention.

Investigators in clinical trials should demonstrate an effort to show that they have optimized dosage capacity by incorporating known parameters of therapeutic effectiveness and an application that is similar to that provided in clinical practice.

Recent consensus-based work has markedly widened the scope of topics that reflect TF (7). In addition to whether the intervention is delivered with a high degree of adherence, TF should include efforts to ensure that the application of the intervention is performed in a way that is known to be therapeutically beneficial (4) (Fig. 1). In other words, was the veracity of the intervention performed and implemented in a manner that should allow someone to improve if performed in a similar clinical situation? To ensure the veracity of TF in a clinical study, one must consider: (a) the quality of delivery and (b) participant adherence.

Quality of delivery assesses both the therapeutic potency of the interventions and the competency of the individuals delivering the intervention. Therapeutic potency reflects whether the clinical parameters such as dosage, time, etc., are performed in a way that allows optimal therapeutic recovery. In a pharmaceutical trial, it would reflect whether the research participant received the appropriate dosages of the medications at appropriate time intervals. Additionally, quality of delivery involves evaluating whether the research administrators have the necessary skills, training, and expertise to deliver the treatment effectively, and ensuring consistency in the provision of interventions between those delivering the treatment.

Participant adherence refers to the extent to which participants engage with and respond to the intervention. It involves monitoring participants' adherence to the intervention protocol, their understanding of the intervention, and their willingness to participate. The selection of appropriately responsive measures that actually assess patient engagement and change in outcomes within the targeted domain is requisite to ensure these measures have meaning.

Examples and recommendations involving TF in clinical trials

Although critical, it is important to recognize that assessment and implementation of TF procedures in a trial is a challenging process (2). There are numerous studies that

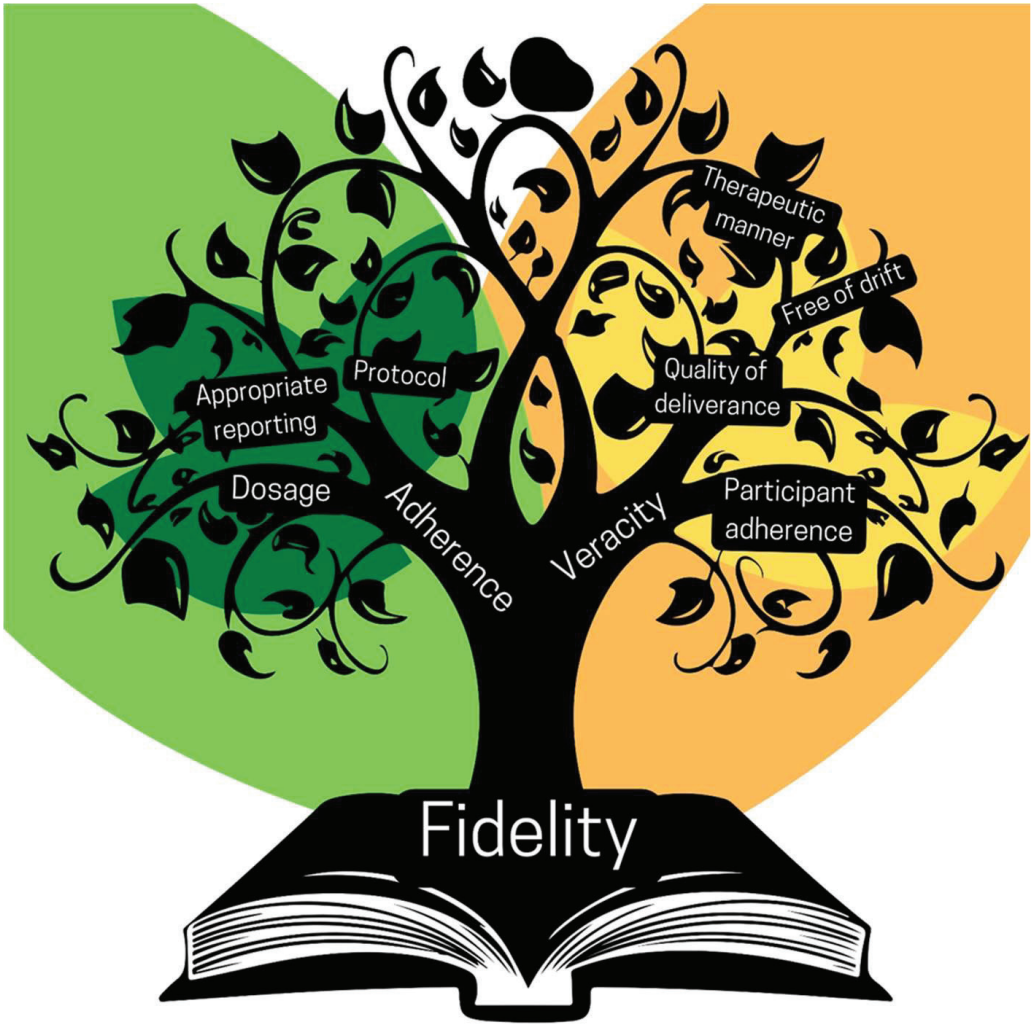


FIGURE 1 - Knowledge tree reflecting the elements of treatment fidelity.

have either experienced or highlighted TF concerns. In this section, we outline examples of TF limitations and provide options for improvement in future studies.

Procedural drift (implementation drift)

Procedural drift is a subcomponent of TF that may influence how a clinician delivers a specific intervention over the course of treatment. It occurs when a clinician chooses the most appropriate intervention based on recommendations at the onset of treatment, and then "drifts" away from using adequate intervention over an episode of care, likely due to their personal beliefs, training, and/or lack of motivation to deviate from their typical model of practice (4). A potential example of procedural drift is the recently published TARGET trial. The TARGET trial (8) reported limited TF in the implementation of a psychologically informed physiotherapy approach, despite initial agreement and formalized training among study clinicians.

Options for improvement

Adding in checklists or manuals that clinicians and researchers can use to improve the quality of specific interventions provided is recommended to limit procedural drift, but adherence to checklists may not always be an easy task due to lack of time, experience, and the belief that the checklists are unnecessary (2). Direct supervision and feedback, videotaping and structured meetings to discuss interventions, along with checklists/manuals, may reinforce the need to limit procedural drift. Early training sessions for clinicians, along with "booster" sessions, to guide the use of appropriate and meaningful interventions may also limit procedural drift in clinical practice. Implementing regular supervised performance reviews with clinicians may assist in determining when adjustments should be made to increase TF (3). Lastly, pretests and the use of specific technologies designed to minimize procedural drift may lend value as well.

Quality and dosage of treatments

A 2021 systematic review (9) was published involving manual therapy interventions vs. sham treatment approaches. In the review, 11 of the 24 reviewed studies (46%) included one visit involving only one technique, applied once. This is not reflective of clinical application nor is it considered to be therapeutic. Further, in many cases, the treatment was applied without interactions with the participants, which did not reflect the contextual aspect of a treatment domain.

Options for improvement

To examine the full treatment effect, including contextual factors and how these are intricately tied to a specific treatment, one must provide the same unique characteristics and components of the intervention, including interpersonal interactions (10). In addition, careful effort should be made to apply the treatment in a manner that is similar to clinical practice and one that reflects clinical practice guideline recommendations.

Vague treatment applications

Recent systematic reviews have found that research reporting and quality of TF remains low across trials investigating exercise therapy and manual therapy for chronic pain, neck pain, and low back pain (11-13). Possible reasons for this deficit include increased time, additional cost, real-world feasibility, and "provider fatigue" from prescriptive and possibly clinician-limiting research designs (14).

Options for improvement

The aforementioned studies exhibited TF limitations, despite the fact that several reporting and fidelity checklists have been developed to monitor the quality of interventions provided in randomized controlled trials (RCTs) for various musculoskeletal conditions. These include the Template for Intervention Description and Replication (TIDieR) and the Consensus on Exercise Reporting Template (CERT), which were both designed to improve the reporting of interventions used in RCTs to assist with methodological transparency and reproducibility of interventions, ultimately leading to improved TF (12,13). In addition to the experimental group, it is imperative that the interventions received by the control group are well described and "controlled." This is commonly an issue in trials and has been identified as a major area of confusion when describing the somewhat innocuous but confusing term of "usual care" (15).

There are also fidelity checklists that have been developed but their effectiveness is questionable. Fidelity checklists are cumbersome, lack succinctness for application, and often include only some of the areas (typically intervention only) that are deemed important to assess (2), frequently failing to address areas such as expertise level of the clinician or procedural drift.

Quality of delivery

In trials that do demonstrate quality reporting of interventions and provide descriptive information on the training and experience of the practitioners' clinical decision-making even while adhering to a strict protocol, TF may still be variable between clinicians. The grade of application in manual therapy, the intensity of resistance in exercise therapy, and the content of the patient instruction including whether to respect or ignore pain are all inherent in physiotherapy interventions. Without consistency of application of these constructs, the same apparent interventions may be applied in a vastly different fashion masking treatment effect.

Options for improvement

One can improve the quality of delivery by training the study providers, and adhering to guiderails of care that are predesigned and incorporated into the training process. This process should be used in both prescriptive and pragmatic clinical trials.

Participant adherence

The recently published PEERC trial (16) is a good example of how participant adherence may have eroded the effect

of one of the treatment arms. In the study, participants with shoulder impingement received a phone-based cognitive behavioral intervention. The authors of the study indicated that there were several instances in which participants took calls: "1) while the patient was driving a car, 2) attending or coaching their youth's sporting events, 3) while at work, 4) while cooking dinner, or 5) during other activities in which they multi-tasked the cognitive behavioral strategies of the PEERC with other daily activities." A cognitive behavioral intervention requires careful attention and active participation to optimize benefits; both of these were absent in many cases in the PEERC trial.

Options for improvement

The necessity of participant adherence should be discussed during the study initiation, and emphasized during the trial. Further, the use of a sensitivity analysis based on those who did and did not adhere to prescribed treatment planning is an option to measure its potential effect.

Unique challenges of TF for physiotherapy and rehabilitation approaches

Measuring TF in physiotherapy can be challenging compared to other areas of healthcare, such as a pharmacological intervention that uses objective laboratory values to determine a treatment regimen, because the nature of physiotherapy is multifaceted, interventions are often clinician-dependent, and interactions between the clinician and patient are uniquely individual (17). Multiple elements impact the delivery, receipt, and enactment of a prescribed physiotherapy treatment intervention and TF may be impacted by the clinician, the patient, or the actual treatment itself (18). The skill of the physiotherapist, the individual needs of the patient, and the distinct interventions required for each individual widely vary across the physiotherapy field, which can lead to significant difficulties in measuring TF.

Multiple covariates associated with the delivery of physiotherapy or other rehabilitation services, such as the time spent with the patient, the setting, and the therapeutic alliance between the patient and provider, can influence TF (18). Because there is so much variation in physiotherapy, a specific checklist may not allow for enough latitude, leading to an unclear interpretation of how high the TF truly is (19). Adaptability within a research protocol, or "flexible fidelity" (20), allows the adjustment of protocol components in response to individual patient differences, such as tailoring exercises based on an individual's pain response or strength. In this context, fidelity can be viewed as adherence to the underlying theory outlined in a treatment protocol, rather than to specific activities or behaviors.

Conclusion

In this viewpoint, we outline the components of TF and provide examples in the literature where TF was lacking. We argue that TF is critical to establishing the evidence base of interventions and determining the circumstances under which an intervention is most effective. Interventions need to be delivered with a high degree of TF, which will allow for greater confidence that the outcomes observed are truly driven by the specific intervention. When TF is not adhered to in clinical research, we may rightly be left to wonder what effect minor modifications of the protocol had on patient outcomes. We suggest that there is a risk that minor modifications could potentially erode the true effect of the treatment and influence clinical outcomes, leading to "evidence" that is erroneously adopted into evolving clinical paradigms.

Disclosures

Conflict of interest: The authors report no conflicts 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 participated in the study concept, writing, and final manuscript preparation.

Data Availability Statement: There are no data associated with this viewpoint.

References

1. Ribeiro LP, Curiel-Montero F, Rodrigues-de-Souza DP, Camargo PR, Alburquerque-Sendín F. Assessment of description and implementation fidelity of clinical trials involving exercise-based treatment in individuals with rotator cuff tears: a scoping review. Braz J Phys Ther. 2024;28(2):101062. CrossRef PubMed
2. Ginsburg LR, Hoben M, Easterbrook A, Anderson RA, Estabrooks CA, Norton PG. Fidelity is not easy! Challenges and guidelines for assessing fidelity in complex interventions. Trials. 2021;22 (1):372. CrossRef PubMed
3. Baker J, Stringer H, McKean C. Ensuring treatment fidelity in intervention studies: developing a checklist and scoring system within a behaviour change paradigm. Int J Lang Commun Disord. 2024;59(1):379-395. An M. CrossRef PubMed
4. Cook CE, Beneciuk JM, George SZ. Procedural drift: an underappreciated element of clinical treatment fidelity. J Orthop Sports Phys Ther. 2022;52(2):63-66. CrossRef PubMed
5. An M, Dusing SC, Harbourne RT, Sheridan SM; START-Play Consortium. START-Play Consortium. What really works in intervention? Using fidelity measures to support optimal outcomes. Phys Ther. 2020;100(5):757-765. CrossRef PubMed
6. Feely M, Seay KD, Lanier P, Auslander W, Kohl PL. Measuring fidelity in research studies: a field guide to developing a comprehensive fidelity measurement system. Child Adolesc Social Work J. 2018;35(2):139-152. CrossRef
7. Sousa Filho LF, Farlie MK, Haines T, et al. Developing an international consensus Reporting guideline for intervention Fidelity in Non-Drug, non-surgical trials: the ReFiND protocol. Contemp Clin Trials. 2024;142:107575. CrossRef PubMed
8. Delitto A, Patterson CG, Stevans JM, et al. Stratified care to prevent chronic low back pain in high-risk patients: the TARGET trial. A multi-site pragmatic cluster randomized trial. EClinicalMedicine. 2021;34:100795. CrossRef PubMed
9. Lavazza C, Galli M, Abenavoli A, Maggiani A. Sham treatment effects in manual therapy trials on back pain patients: a systematic review and pairwise meta-analysis. BMJ Open. 2021;11(5):e045106. CrossRef PubMed
10. Testa M, Rossettini G. Enhance placebo, avoid nocebo: how contextual factors affect physiotherapy outcomes. Man Ther. 2016;24:65-74. CrossRef PubMed

11. Adams SC, McMillan J, Salline K, et al. Comparing the reporting and conduct quality of exercise and pharmacological randomised controlled trials: a systematic review. BMJ Open. 2021;11(8):e048218. CrossRef PubMed
12. McConnell R, Klopper M, Rhon DI, Young JL. The influence of exercise therapy dosing on pain and functional outcomes in patients with subacromial pain syndrome: a systematic review. Shoulder Elbow. 2024;16(1)(suppl):42-58. CrossRef PubMed
13. Kucksdorf JJ, Bartley J, Rhon DI, Young JL. Reproducibility of exercise interventions in randomized controlled trials for the treatment of rotator cuff-related shoulder pain: a systematic review. Arch Phys Med Rehabil. 2024;105(4):770-780. CrossRef PubMed
14. Fuller T, Pearson M, Peters J, Anderson R. What affects authors' and editors' use of reporting guidelines? Findings from an online survey and qualitative interviews. PLoS One. 2015;10(4):e0121585. CrossRef PubMed
15. Pascoe SC, Spoonemore SL Jr, Young JL, Rhon DI. Proposing six criteria to improve reproducibility of "usual care" interventions in back pain trials: a systematic review. J Clin Epidemiol. 2022;149:227-235. CrossRef PubMed
16. Myers H, Keefe FJ, George SZ, et al. Effect of a Patient Engagement, Education, and Restructuring of Cognitions (PEERC) approach on conservative care in rotator cuff related shoulder pain treatment: a randomized control trial. BMC Musculoskelet Disord. 2023;24(1):930. CrossRef PubMed
17. Whyte J, Hart T. It's more than a black box; it's a Russian doll: defining rehabilitation treatments. Am J Phys Med Rehabil. 2003;82(8):639-652. CrossRef PubMed
18. Toomey E, Matthews J, Hurley DA. Using mixed methods to assess fidelity of delivery and its influencing factors in a complex self-management intervention for people with osteoarthritis and low back pain. BMJ Open. 2017;7(8):e015452. CrossRef PubMed
19. Hall AM, Ferreira PH, Maher CG, Latimer J, Ferreira ML. The influence of the therapist-patient relationship on treatment outcome in physical rehabilitation: a systematic review. Phys Ther. 2010;90(8):1099-1110. CrossRef PubMed
20. Palmer JA, Parker VA, Barre LR, et al. Understanding implementation fidelity in a pragmatic randomized clinical trial in the nursing home setting: a mixed-methods examination. Trials. 2019 Nov 28;20(1):656. CrossRef PubMed

Arch Physioter 2024; 14: 70-79
ISSN 2057-0082 | DOI: 10.33393/aop.2024.2733
ORIGINAL RESEARCH ARTICLE

Neurological conditions and community-based physical activity: physical therapists' belief and actions

Louise Declerck[1], Mathilde Gillot[1], Charlotte Goffaux[1], Jean-François Kaux[2,3], Gaëtan Stoquart[1,4]

[1]Neuromusculoskeletal Lab, Institut de Recherche Expérimentale et Clinique, UCLouvain, Brussels - Belgium
[2]ReFORM IOC Research Centre for Prevention of Injury and Protection of Athlete Health, Liège - Belgium
[3]Physical Medicine and Sport Traumatology Department, SportS[2], FIFA Medical Centre of Excellence, FIMS Collaborative Centre of Sports Medicine, University and University Hospital of Liege, Liège - Belgium
[4]Physical and Rehabilitation Medicine Department, Cliniques Universitaire Saint-Luc, Brussels - Belgium

ABSTRACT

Introduction: Physical therapists (PTs) are key actors in physical activity (PA) promotion. However, it remains unclear whether PTs in community settings promote community-based PA such as adapted physical activity (APA) and adaptive sports (AS) to their patients with neurological conditions (NCs). The main purposes were to evaluate the beliefs PTs have of APA and AS, and to explore actions they undertake to promote it to their patients with NCs.
Methods: An online survey was created specifically for the study. PT associations and institutions were contacted and licensed PTs working in community-based settings, treating at least one patient with a NC, were invited to participate. Questionnaires were analyzed only if all mandatory questions had been answered.
Results: A total of 165 questionnaires were analyzed. PTs reported prioritizing active treatment. They viewed APA and AS as beneficial for their patients with NCs; however, its promotion remained largely infrequent due to a number of barriers. The PTs' own level of PA seemed to significantly influence their beliefs of the benefits of APA and AS ($p = 0.001$), while being specialized in neurologic physical therapy enabled the PTs to increase frequency of promotion ($p = 0.003$).
Conclusion: Though community-based PTs are aware of the importance of PA for individuals with NCs, they face difficulties in promoting it to their patients. However, these difficulties are reduced among PTs who are specialized in neurologic physical therapy. Efforts should be made toward educating PTs to neurological pathologies and their specificities when it comes to PA.

Keywords: Health promotion, Neurological rehabilitation, Physical activity, Physical therapists

What's already known about this topic:

- *Adapted physical activity, including adaptive sports, is very beneficial for individuals with disabilities due to neurological conditions. Health care professionals, especially physical therapists, are well placed to vehicle such messages and should promote such activities.*

What does the study add:

- *Physical therapists do not frequently promote adapted physical activity and adaptive sports to their patients with neurological conditions. Lack of knowledge limits their actions. Such barriers are reduced when specializing in neurologic physical therapy.*

Introduction

Though it was long believed that physical activity (PA) was detrimental for people presenting with neurological conditions (NCs) such as stroke, multiple sclerosis (MS), Parkinson's disease (PD), spinal cord injury (SCI), or others (1), sound evidence now clearly demonstrates the beneficial effects of PA on different NCs (2,3), enabling a paradigm shift (4). The literature suggests PA reduces the risk of developing secondary complications (5) and improves autonomy in everyday life (6-10). Finally, in some progressive NCs such as PD or MS, PA may decelerate neurodeterioration (11,12). PA should therefore be a vital part of neurorehabilitation, as recommended by a wide range of condition-specific clinical guidelines (13-16).

Despite this, a majority of individuals with disability due to NCs do not engage in sufficient PA. Indeed, studies suggest that, in the United States, up to 80% of this population are physically inactive (17,18). This finding has been echoed

Received: November 23, 2023
Accepted: September 5, 2024
Published online: September 30, 2024

This article includes supplementary material

Corresponding author:
Louise Declerck
email: louise.declerck@uclouvain.be

in other parts of the world, as authors repeatedly report low level of PA and highly sedentary lifestyles among individuals with NCs (19-22). A drop in PA level is especially great after rehabilitation, when people with NCs return to their communities (23). This lack of compliance to long-term PA may be overcome by making the activity more enjoyable and social. In that sense adapted physical activity (APA) and adaptive sports (AS) allow for PA to be performed in group settings, while under supervision of a trained coach or therapist.

However, individuals with NCs often report lack of knowledge on how, and where, to engage in such PA in the community (24,25). Health care professionals (HCPs) therefore play a vital role in educating their patients toward leading a more active lifestyle (26). In that regard, physical therapists (PTs), defined as exercise experts by the "World Confederation for Physical Therapy," are especially important (27). Moreover, during rehabilitation, individuals with NCs will spend more time with their PT than with any other HCP, making PTs a key reference (28).

While most PTs acknowledge their responsibility in PA promotion among individuals with NCs, implementation in real-life settings remains challenging (27). A qualitative study found that although English PTs believed PA to be important, efforts to promote it to their patients with SCI were lacking (29). However, this study focused specifically on PTs working within SCI-specific rehabilitation centers. Yet, people with NCs do not always have the opportunity to attend highly specialized centers on a long-term basis. Furthermore, PTs who work in community-based settings may encounter ever more difficulty in promoting PA to such patients. It is therefore important to investigate how these PTs use PA, and promote APA and AS among their patients with NCs, within nonspecialized, community settings.

Therefore, the aims of the present study are (i) to explore the perceptions of benefits of APA and AS for individuals with NCs among PTs working within community settings; (ii) to assess if PTs utilize PA in their therapy, and (iii) to explore actions undertaken by the PTs to promote APA and AS as well as barriers to such actions. The secondary objective is to identify PT-related factors influencing PA beliefs and actions. Our hypothesis is that PTs perceive APA and AS as beneficial, but only few utilize PA as a therapeutic tool. Additionally, we except that the majority do not actively promote these activities to their patients with NCs.

Methods

This cross-sectional study was a web-based survey, directed toward French-speaking PTs in Belgium. The study was constructed and written according to Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines, as well as the "Checklist for Reporting Results of Internet E-surveys" guidelines (30). The completed checklist can be found in the Supplementary material I: CHERRIES. Ethical clearance was granted by the Ethics Committee of the Catholic University of Louvain. Participants remained anonymous, and gave their informed consent. Data were treated according to the General Data Protection Regulation.

Convenience sampling was used and participants were invited to respond to the online survey from November 2020 to April 2021. Participation was voluntary.

Eligible participants had to be (i) licensed PTs, (ii) practicing in Brussels or Wallonia (Belgium), (iii) practicing at least partly in a community setting, and (iv) French speakers. Moreover, (v) participants had to be treating at least one patient with a NC when answering the survey. PTs were excluded if their practice setting was solely based in hospitals, clinics, or rehabilitation centers, or if they were retired.

Sample size calculations were performed according to the total number of PTs practicing in Brussels and Wallonia. According to the latest Belgian report, this equaled 12,053 in 2016 (31). As response rates for online surveys approximate 30% (32), and using a margin error of 5% with a confidence level of 90%, the recommended sample size was 225 (33).

An online, adaptive, open questionnaire was created specifically for the study using "Limesurvey." This platform performs IP checks to disable duplicate responses and ensures secure data protection through the Catholic University of Louvain.

Brainstorms among three researchers (one PT and two physicians), with knowledge of the literature available on the topic, were conducted and led to the creation of an initial version of the questionnaire. General guidelines for creating web-based surveys were followed (34): the majority of the questions were mandatory, it was not possible to return to previous questions once answered, questions were mainly closed-ended in order to decrease participation time (35), an adaptive structure was used (i.e., answers to one question determined following questions), and demographic-related questions were placed at the end of the survey (34). A progression bar was added so participants could estimate time to survey completion. Majority of the answers were on a 4-point Likert scale going from 0 (never/not at all) to 3 (always/very).

This first version was critically reviewed by three PTs with experience in neurorehabilitation, and modifications were made. The second version was then tested by another five PTs, who were naïve to the previous version. Their comments allowed final modifications to be made. The questionnaire's final version included 26 questions, with an estimated completion time of 12 minutes. An English version can be found in the Supplementary material II: questionnaire used for the survey (translated from French to English).

Different communication channels were used simultaneously. First, a short message pertaining to our survey's objectives and length, and containing the URL link toward the questionnaire, was published on different Belgian PTs Facebook groups. Second, local and national PT associations were contacted, by mail or phone, in order to diffuse survey link to their members. Third, the published repertoires "kinesithérapie.be" and "abterna.be" were used to contact PTs directly. Only PTs whose contacted details were published were contacted, preferably by phone (if their phone number was published) or by mail. Reminders were sent twice, with a 1-month interval.

Data were exported from Limesurvey into Excel in CSV format. Incomplete questionnaires (where a minimum one

mandatory question was left unanswered) were removed from the analysis. Answers were summarized descriptively, by reporting the absolute and relative frequency.

A score was attributed to the PTs' beliefs of benefits of APA and AS, and another for actions to promote APA and AS. This was done by summing the answers obtained on the Likert scales (i.e., "0: never/not important/not efficient" equaled 0, while "3: very frequently/very important/very efficient" equaled 3). For the total belief score, as this comprised the participants' answers to four questions, maximal score was 12. Higher scores represented more positive beliefs. For the total action score, this related to five questions, with a maximal score of 15. Higher scores represented greater frequency of APA and AS promotion.

Statistical analyses were performed on both total belief and action scores using Statistical Package for the Social Sciences (SPSS; IBM, version 27). First, correlation between beliefs and actions scores was computed through Spearman's test. The correlation coefficient was interpreted as negligible (0-0.10), low (0.11-0.39), moderate (0.40-0.69), strong (0.70-0.89), or very strong (0.90-1) (36). Second, to evaluate the influence of demographics on beliefs and actions, different tests were performed: Spearman's correlations, to explore influence of the number of years PT treated patients with NCs; Kruskal-Wallis tests, first, to evaluate differences according to self-reported level of PA, and second, to explore differences according to percentage of patients with NCs within total patient population; and finally, Mann-Whitney tests, to evaluate differences according to presence of specific training in neurologic physical therapy. When differences were found, they were further analyzed by a chi-square test, to identify which questions led to the significant difference in scores between the groups. For all analyses, a p-value ≤ 0.05 was considered significant.

Results

A total of 255 individuals viewed the questionnaire's introduction page, of which 224 advanced to the next section containing the questions related to eligibility; 33 individuals did not respect the inclusion criteria and were excluded. Of the remaining 191 PTs, 26 did not answer all mandatory questions. Therefore, 165 participants were included for analysis. The number of years practicing physical therapy with patients with NCs ranged from 0.08 (equivalent to 1 month) to 50, with a median of 7 years. While n = 19 participants self-reported low PA levels, the majority reported being moderately (n = 80) and highly (n = 61) physically active. No participant self-reported as not being physically active at all. Only 29% of the sample were specialized in neurologic physical therapy. Demographic parameters of the sample and their patient populations are displayed in Table 1.

TABLE 1 - Demographic variables of the sample

Variable	Categories	n (%) or median (1st; 3rd quartile)
Number of years practicing physical therapy with patients with NCs		7 (3; 20)
Specialized in neurologic physical therapy	-Yes	48 (29%)
	- No	109 (67%)
	- No answer	8 (4%)
Percentage of patients with NCs within overall patient population	- Less than 25%	104 (63%)
	- More than 25% but less than 50%	23 (14%)
	- More than 50% but less than 75%	11 (7%)
	- More than 75% but less than 100%	20 (12%)
	- 100%	7 (4%)
Type of NCs presented by patients*	- Stroke	137 (83%)
	- Parkinson's disease	111 (67%)
	- Multiple sclerosis	74 (45%)
	- Peripheral nerve lesion	61 (37%)
	- Neuromuscular disease	55 (34%)
	- Traumatic brain injury	44 (26%)
	- Spinal cord injury: paraplegia	28 (16%)
	- Spinal cord injury: tetraplegia	20 (12%)
	- Spina bifida	15 (9%)
	- Others	16 (10%)

(*Cont.*)

TABLE 1 - (*Cont.*)

Variable	Categories	n (%) or median (1st; 3rd quartile)
Disability level of patients presenting with NCs	- Majority (over 50%) present with severe disability	17 (9%)
	- Majority (over 50%) present with moderate disability	53 (31%)
	- Majority (over 50%) present with mild disability	65 (38%)
	- Disability level evenly spread among severe, moderate, and mild	30 (22%)
Self-reported PA level	-None	0
	-Low	19 (11%)
	-Moderate	80 (48%)
	-High	61 (37%)
	-No answer	5 (3%)

NCs = neurological conditions; PA = physical activity.
*Note that multiple answers were possible. Therefore, some participants responded positively to a range of categories.

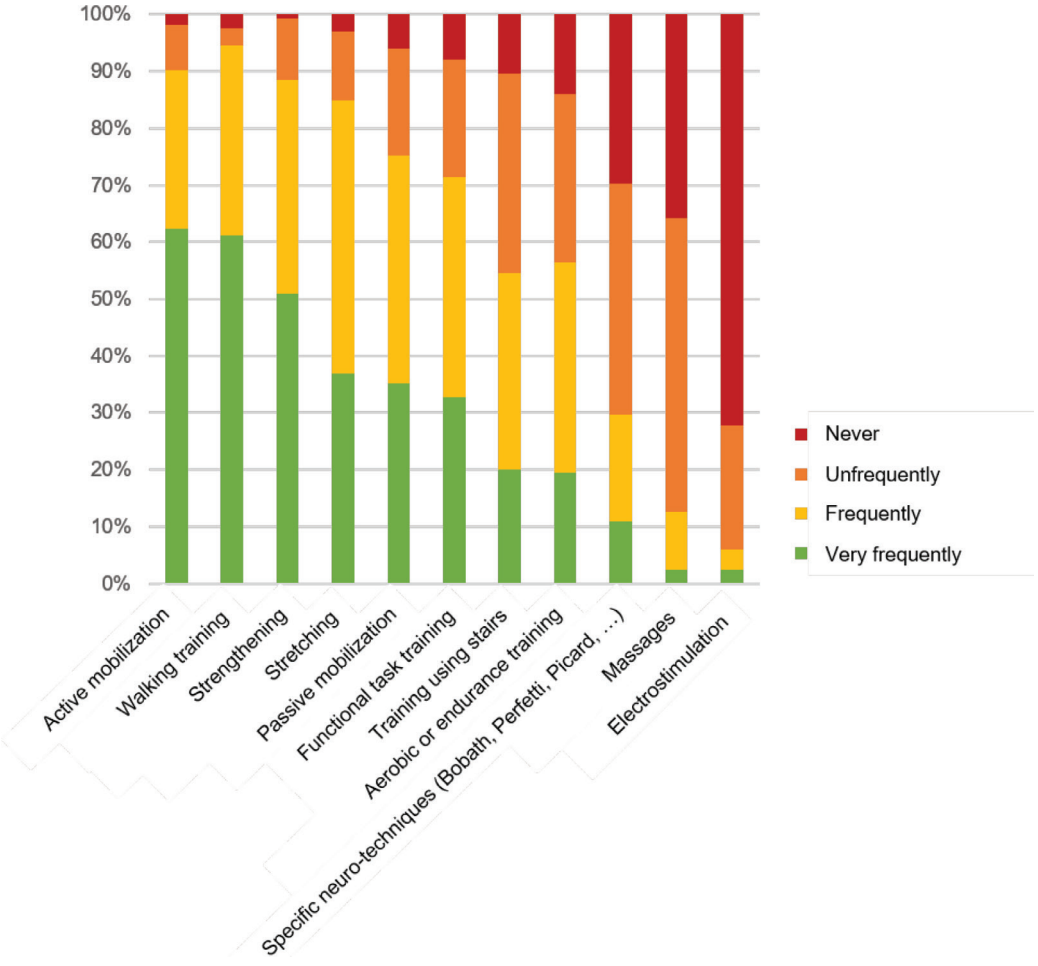


FIGURE 1 - Treatments used by the PTs during therapy. Bar graph demonstrating treatments used by all participants during sessions with patients with neurologic conditions.

Concerning general treatments performed with patients with NCs, the most common were: active mobilization, walking training, resistance training, and stretching. Endurance training was never or infrequently used by 14% and 30% of the sample, respectively. Massages and electrostimulation were the least common treatment options (Fig. 1).

The vast majority of PTs believed APA and AS to be important or very important for the physical and mental health of their patients with NCs. They also believed APA and AS to be effective at improving and maintaining motor function and autonomy (Tab. 2). Altogether, beliefs regarding the benefits of APA and AS were high among PTs, with a median score of 10 (Fig. 2).

TABLE 2 - Beliefs and actions reported by PTs

Items	Not effective/not important/never (0)	Slightly effective/slightly important/rarely (1)	Effective/important/frequently (2)	Very effective/very important/very frequently (3)
Belief 1: Effects of APA or AS on physical health	n = 3 2%	n = 8 5%	n = 60 36%	n = 94 57%
Belief 2: Effects of APA or AS on mental health	n = 2 1%	n = 7 4%	n = 56 34%	n = 100 61%
Belief 3: Effects of APA or AS on motor function	n = 4 2%	n = 5 3%	n = 74 45%	n = 82 50%
Belief 4: Effects of APA or AS on autonomy	n = 4 2%	n = 5 3%	n = 87 53%	n = 69 42%
Action 1: Discuss the subject of APA or AS with patient	n = 31 19%	n = 53 32%	n = 63 38%	n = 18 11%
Action 2: Inquire into patient's habits concerning APA or AS	n = 35 21%	n = 54 33%	n = 61 37%	n = 15 9%
Action 3: Encourage patient to partake in APA or AS outside of physical therapy session	n = 30 18%	n = 38 23%	n = 66 40%	n = 31 19%
Action 4: Guide patient with steps toward participating in APA or AS	n = 76 46%	n = 63 38%	n = 20 12%	n = 6 4%
Action 5: Assess amount of PA undertaken by patient	n = 134 81%	n = 13 8%	n = 15 9%	n = 3 2%

APA = adapted physical activity; AS = adaptive sports; PA = physical therapist; PT = physical activity.

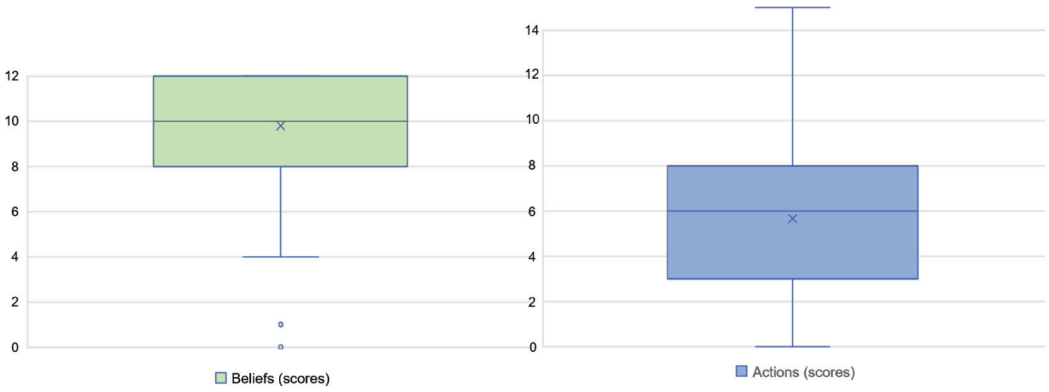


FIGURE 2 - Distribution of belief and action scores of the total sample. Boxplot demonstrating belief and action scores obtained by all participants. Belief scale ranged from 0 to 12, while action scale ranged from 0 to 15.

Concerning actions undertaken to promote APA and AS, half of the sample did not discuss the subject of APA and AS with their patients with NCs, and more than half did not inquire about their patients' habits concerning APA and AS participation. Other actions to promote APA and AS, such as encouraging their patients with NCs to partake in such activities, or helping patients with NCs through the steps toward participating in APA or AS in community settings (including finding accessible sports clubs or centers), remained rare. Finally, 81% of the sample never assessed the amount of PA performed by their patients with NCs (Tab. 2). Accordingly, action scores of the sample were low, with a median of 5 (Fig. 2).

The most common barriers to undertaking actions toward APA or AS promotion are summarized in Table 3. While the most frequent barrier for PTs specialized in neurology was the lack of accessibility regarding information on APA and AS sessions, nonspecialized PTs reported being most limited by the lack of demand for such activities coming from their patients.

Statistical analyses demonstrated significant correlations of moderate intensity (r = 0.48, p = 0.001) between the PTs' belief and action scores. Number of years practicing physical therapy with patients with NCs did not correlate with beliefs (r = 0.06, p=0.460) or actions (r = 0.098, p = 0.217). Likewise, the percentage of patients with NCs within total patient

TABLE 3 - Barriers toward APA and AS promotion

Barriers	Yes, this is a barrier (n, %) of total sample	Yes, this is a barrier (n, %) of sample with specific training in neurology	Yes, this is a barrier (n, %) of sample without training in neurology
Availability of time	90 55%	26 54%	62 56%
Patient demand for such activities	130 79%	34 70%	90 82%
Knowledge on APA and AS	113 69%	27 56%	83 76%
Accessibility to information regarding APA and AS availability	129 78%	36 75%	88 80%

APA = adapted physical activity; AS = adaptive sports.

population did not influence both scores (beliefs p = 0.227, actions p = 0.138).

The presence of specific training within the neurology domain played a significant role on action scores (p = 0.003), whereby PTs with specific training in neurology undertook action to promote APA and AS more frequently than their colleagues (Fig. 3). Specifically, chi-square tests revealed that the actions undertaken significantly more frequently among PTs with training were inquiring into the patients' habits concerning APA and AS (p = 0.040), guiding patients through the steps toward APA and AS sessions (p = 0.033), as well as assessing patients' PA levels (p = 0.001). Training in neurology did not impact belief scores (p = 0.451). Conversely, while self-reported PA levels significantly influenced beliefs (p = 0.001) (Fig. 4), it had no impact on actions (p = 0.148). The highly and moderately active groups had significantly more positive beliefs related to the effects of APA and AS on their patients' physical health (p = 0.010) and motor function (p = 0. 022), in comparison to the group that reported low PA levels.

Discussion

The primary aims of this survey were to explore community-based PTs' beliefs regarding APA and AS, and actions undertaken to promote these activities to individuals with NCs. The findings show that while the PTs believe APA and AS to be very beneficial for their patients with NCs, and commonly use active treatments in their therapy, they rarely undertake actions to promote APA and AS practice. Lack of demand from their patients, as well as lack of information on where APA and AS can be practiced, seem to be the two greatest barriers.

A large majority of the participating PTs had very positive beliefs regarding APA and AS, and favored active treatments to passive ones such as massage. This is in line with

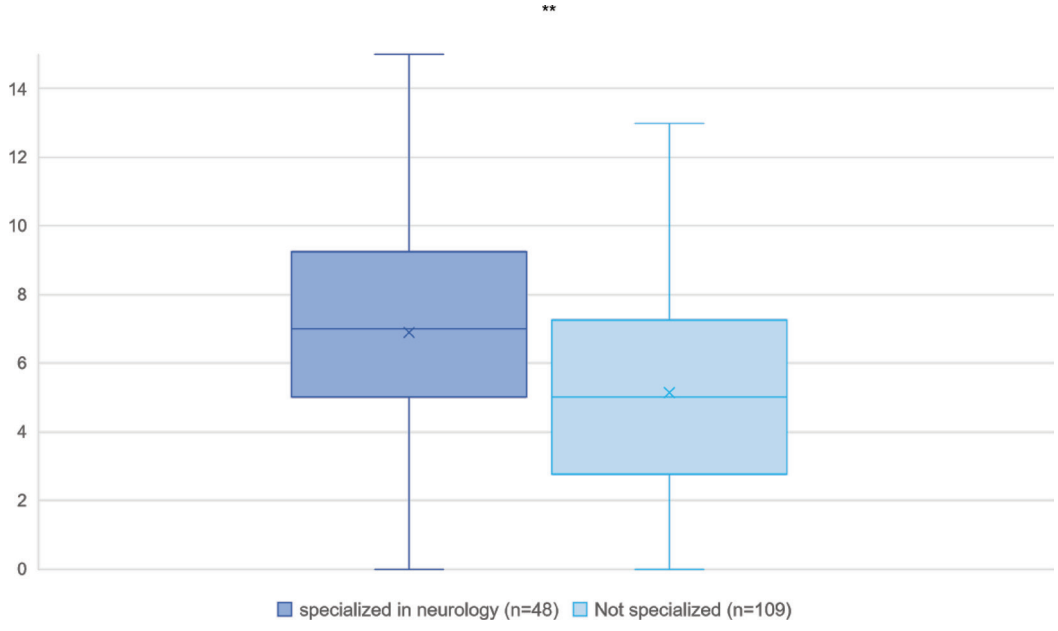


FIGURE 3 - Action scores obtained by physical therapists with and without specific specialization in neurology. Boxplot demonstrating the action scores obtained by the participants with and without specialization in neurologic physical therapy. **Significant difference where p-value is inferior to 0.01.

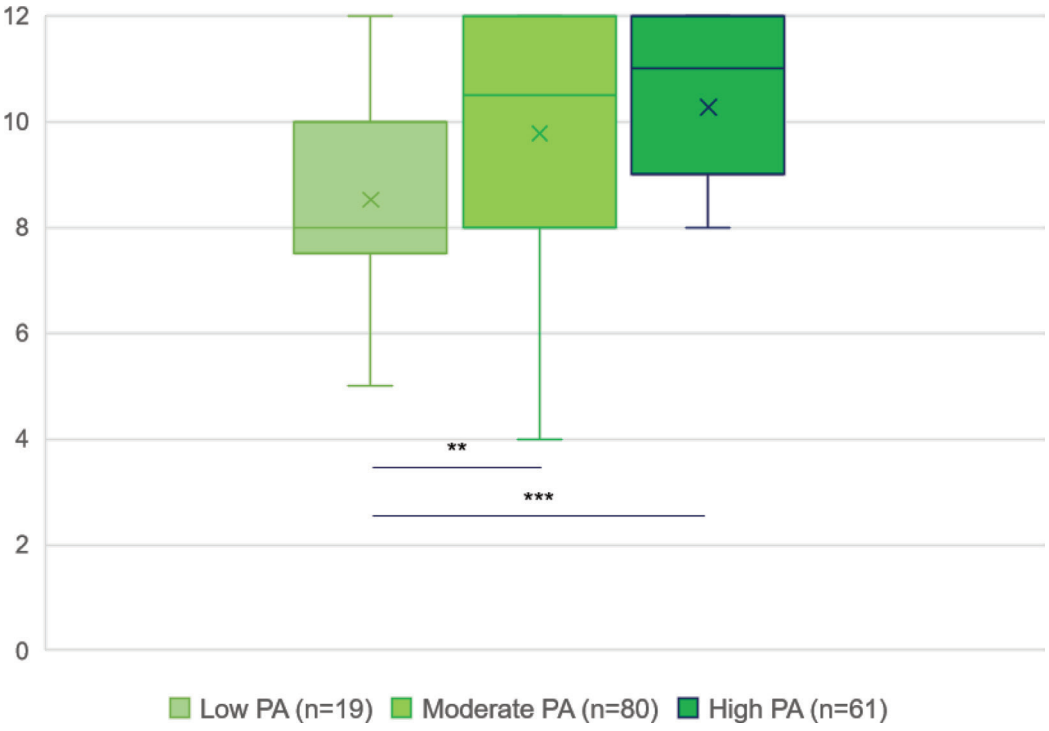


FIGURE 4 - Belief scores obtained by physical therapists reporting low, moderate, or high level of physical activity. Boxplot demonstrating belief scores obtained by the participants engaging in low, moderate, or high self-reported level of physical activity. **Significant difference where p-value is inferior to 0.01. ***Significant difference where p-value is 0.001.

clinical guidelines stating the importance of PA in all stages of neurorehabilitation (37). A range of studies, performed among PTs in other various parts of the world, demonstrated similar positive attitudes toward PA, for all types of patients (29,38-41). Our findings further demonstrate that some demographic factors such as the number of years of practice with patients with NCs, specialization in neurologic physical therapy, as well as percentage of patients with NCs compared to total patient population do not influence beliefs. On the other hand, PTs who self-report as moderately or highly physically active view the effects of APA and AS more positively than PTs with low levels of PA. This seems to be related to APA and AS' effects on physical health and motor capacity specifically. Similarly, Turkish PTs with greater levels of PA were more convinced of the benefits of PA for their patients, than their less active colleagues (42).

However, actions undertaken to promote APA and AS remained infrequent. Only half of the PTs reported discussing APA or AS with their patients or inquiring into their PA habits, and little more than half encouraged their patients to engage in these activities. Moreover, the percentage of PTs who reported promoting APA or AS "very frequently" further dropped to less than 20%. This low percentage is in line with conclusions drawn by two qualitative studies within the field of neurological physical therapy. Indeed, authors of these studies, performed in England and Ireland (29), and New Zealand and Sweden (43), also observed that PA promotion remained predominantly absent from clinical practice. Conversely, Kennedy et al have found that 45% of their sample of 76 American PTs always promoted PA to patients with NCs (44). This difference, noted between Europe and New Zealand, and the United States, could be due to contextual factors such as PT education and reimbursement conditions. International collaborations could be set up in that regard, in order to learn from one another's experience and benefit all parties involved.

The action that was found to be most lacking was guiding patients with NCs through the steps toward enrolling in an APA or AS in the community. Indeed, above 80% of our study's total sample reported never, or only rarely, doing this. Yet, studies show that tailored PA counseling, taking into account the social and environmental conditions unique to each patient, is key in order to increase PA participation (45). To be effective, PA promotion needs to be frequent, repetitive, and include information on how and where to engage in such activities in the community. Educating patients with NCs on where to find this information themselves, as well as who to contact in order to enroll in APA or AS session in the community, is important as it empowers them and creates long-term changes (45).

Interestingly, our data uncovered that the frequency of APA and AS promotion-action was significantly greater following additional training in the field of neurologic physical therapy. Indeed, PTs with additional training in neurology reported undertaking more actions to promote APA or AS to patients with NCs, than their colleagues without training. This related to actions such as inquiring into their patients' APA and AS habits, guiding patients through the steps needed to enroll in an APA or AS program in the community, and finally, monitoring or assessing their patients' PA levels. All these actions allow PA promotion to be tailored according to the individual and his/her needs, and is vital for long-term participation (46). Therefore, efforts should be placed toward training more PTs in neurology, as it allows them to develop essential competencies that seem to be lacking from general PT training. This lack of training was observed by Eisele et

al, who reported that German PTs, working in outpatient settings with all types of patients, believed they required greater competencies in order to promote exercise to patients who do not engage in PA (47).

Moreover, our findings suggest that PTs who are trained in neurology encounter less barriers to APA and AS promotion. Barriers such as accessibility to information on APA and AS availability in the community, as well as knowledge on APA and AS in general, were also less common among this group. Increased knowledge on APA and AS likely contributed to the increased actions undertaken by specialized PTs to promote such activities to their patients. Other authors have also reported on the important role of knowledge on increasing frequency of PA promotion among PTs (44,48). It is therefore of utmost importance that individuals with NCs who are discharged from rehabilitation settings be redirected to PTs in the community who are specialized in neurology. Indeed, these PTs have greater understanding of the special needs of this population, and therefore may provide them with more information on ways to be physically active.

The total sample's greatest barrier to PA promotion seemed to be lack of demand from patients with NCs for such activities. German PTs also reported lack of patient interest for PA as the primary obstacle to exercise promotion for all types of patients (47). However, data suggest that the majority of patients with NCs, such as those with stroke, are interested in PA but lack education on the matter and therefore do not bring up the subject with HCPs (45,49). Moreover, certain tools, including behavior change techniques and education, have shown to be effective for those with low PA motivation (50).

In regard to accessibility to information on APA and AS availability in the community, ranked as the second and first barriers for PTs without and with specialization in neurology, similar results have been observed by Zhu et al. In their sample of 84 Australian PTs working in hospital settings, difficulty locating adequate PA opportunities in the community was cited as one of the most common barriers (51). Indeed, APA and AS still remain poorly developed when compared to sporting activities and opportunities for individuals without a disability (52). A solution could be to develop tools such as websites or applications that display this information in a user-friendly way, and that updates them regularly. Collaborating with patient organizations, which can provide greater insight into the specific needs of their members, should be encouraged when developing this. Such tools then need to be made visible among PTs in order to become engrained in everyday use with patients with NCs. Associations representing PTs at both a national and international level (such as "World Physiotherapy") could be involved in making these tools visible.

Finally, time, or lack of it, seemed to be a barrier for half of the sample. While some authors reported time to be a significant, or even the most significant, barrier (38,53,54), others found only small proportions of the sample to be limited by time (42,55). However, as exercise is now recognized as a vital sign of health (56), it should gain priority in the treatment. This could be facilitated through education and implementation of specific guidelines on PA promotion in the physical therapy practice (29).

Certain study limitations should be considered. First, the sample size of 225 was not reached, though 255 PTs opened the survey. This may be due to our eligibility criteria. Indeed, the sample size calculation was based on the total number of PTs in Brussels and Wallonia, while our study only recruited PTs working in community settings with at least one patient with a NC. Thus, the sample number obtained may be representative of our specific population, though it is impossible to be certain as reports only state total number of PTs. Second, similarly to other self-completed questionnaires, social desirability may have skewed results concerning the frequency of actions undertaken to promote APA and AS. Moreover, participation was voluntary, so recruited PTs may have been highly interested in APA or AS. Yet, as one step of recruitment included contacting PTs one by one, and as the percentages of PTs answering "no" or "rarely" to some questions is high, the influence of these factors likely remained small.

Conclusions

Though PTs practicing in the community view APA and AS as very beneficial for their patients with NCs, and primarily use active therapies within the treatment they provide to these patients, promotion of APA and AS remains infrequent. Certain barriers, including lack of demand for such activities as well as difficulty in obtaining information on the availability of community-based APA and AS, still limit them. However, PTs who are specialized within neurologic physical therapy promote APA and AS more frequently, and report fewer barriers limiting their actions to do so. Effort toward educating more PTs to neurological physical therapy should therefore be made. Moreover, individuals with NCs should be directed to these types of PTs once they return to community settings. International collaborations should be encouraged, in order to inform best practices on PA promotion within individuals with NCs. Finally, tools, which centralize the information on availability of APA and AS sessions, should be created to facilitate visibility of these activities.

Acknowledgments

The work reported in the manuscript has been presented, through a 5-minute oral presentation, at the "XXIe Journées d'Etudes Francophones en Activité Physique Adaptée" in Nantes (France) in March 2023.

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 sector.

Authors' contributions: LD played a role in conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, writing original draft, and writing review and editing. GS played a role in conceptualization, formal analysis, investigation, project administration, supervision, validation, writing original draft, and writing review and editing. JFK played a role in conceptualization, investigation, project administration, validation, and writing review and editing. MG and CG played a role in data curation, formal analysis, methodology, and writing original draft.

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

Clinical Trial Protocol number: Not applicable to this survey.

References

1. Carr J, Shepherd R. The changing face of neurological rehabilitation. Braz J Phys Ther. 2006;10(2). CrossRef
2. Martin Ginis KA, van der Ploeg HP, Foster C, et al. Participation of people living with disabilities in physical activity: a global perspective. Lancet. 2021;398(10298):443-455. CrossRef PubMed
3. Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med. 2020;54(24):1451-1462. CrossRef PubMed
4. Nadeau SE. A paradigm shift in neurorehabilitation. Lancet Neurol. 2002;1(2):126-130. CrossRef PubMed
5. Selph SS, Skelly AC, Wasson N, et al. Physical activity and the health of wheelchair users: a systematic review in multiple sclerosis, cerebral palsy, and spinal cord injury. Arch Phys Med Rehabil. 2021;102(12):2464-2481.e33. CrossRef PubMed
6. Saunders DH, Sanderson M, Hayes S, et al. Physical fitness training for stroke patients. Cochrane Database Syst Rev. 2020;3(3):CD003316. PubMed
7. Hayes S, Galvin R, Kennedy C, et al. Interventions for preventing falls in people with multiple sclerosis. Cochrane Database Syst Rev. 2019;11(11):CD012475; Epub ahead of print. CrossRef PubMed
8. Mehrholz J, Kugler J, Storch A, Pohl M, Hirsch K, Elsner B. Treadmill training for patients with Parkinson's disease. Cochrane Database Syst Rev. 2015;2015(9):CD007830; Epub ahead of print. CrossRef PubMed
9. Ortega-Hombrados L, Molina-Torres G, Galán-Mercant A, Sánchez-Guerrero E, González-Sánchez M, Ruiz-Muñoz M. Systematic review of therapeutic physical exercise in patients with amyotrophic lateral sclerosis over time. Int J Environ Res Public Health. 2021;18(3):1074. CrossRef PubMed
10. Lai B, Lee E, Kim Y, et al. Leisure-time physical activity interventions for children and adults with cerebral palsy: a scoping review. Dev Med Child Neurol. 2021;63(2):162-171. CrossRef PubMed
11. Spielman LJ, Little JP, Klegeris A. Physical activity and exercise attenuate neuroinflammation in neurological diseases. Brain Res Bull. 2016;125:19-29. CrossRef PubMed
12. Motl RW, Pilutti LA. Is physical exercise a multiple sclerosis disease modifying treatment? Expert Rev Neurother. 2016;16(8):951-960. 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. Haselkorn JK, Hughes C, Rae-Grant A, et al. Summary of comprehensive systematic review: rehabilitation in multiple sclerosis: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology. 2015;85(21):1896-1903. CrossRef PubMed
15. Keus S, Munneke M, Graziano M, et al. European Physiotherapy Guideline for Parkinson's Disease. 2014;191. Online. Accessed November 2023.
16. Martin Ginis KA, van der Scheer JW, Latimer-Cheung AE, et al. Evidence-based scientific exercise guidelines for adults with spinal cord injury: an update and a new guideline. Spinal Cord. 2018;56(4):308-321. CrossRef PubMed
17. Carroll DDC-LE, Courtney-Long EA, Stevens AC, et al; Centers for Disease Control and Prevention (CDC). Vital signs: disability and physical activity—United States, 2009-2012. MMWR Morb Mortal Wkly Rep. 2014;63(18):407-413. PubMed
18. Butler EN, Evenson KR. Prevalence of physical activity and sedentary behavior among stroke survivors in the United States. Top Stroke Rehabil. 2014;21(3):246-255. CrossRef PubMed
19. Casey B, Coote S, Galvin R, Donnelly A. Objective physical activity levels in people with multiple sclerosis: meta-analysis. Scand J Med Sci Sports. 2018;28(9):1960-1969. CrossRef PubMed
20. de Hollander EL, Proper KI. Physical activity levels of adults with various physical disabilities. Prev Med Rep. 2018;10:370-376. CrossRef PubMed
21. Joseph C, Conradsson D, Hagströmer M, Lawal I, Rhoda A. Objectively assessed physical activity and associated factors of sedentary behavior among survivors of stroke living in Cape Town, South Africa. Disabil Rehabil. 2018;40(21):2509-2515. CrossRef PubMed
22. Fini NA, Holland AE, Keating J, Simek J, Bernhardt J. How physically active are people following stroke? Systematic review and quantitative synthesis. Phys Ther. 2017;97(7):707-717. CrossRef PubMed
23. van den Berg-Emons RJ, Bussmann JB, Haisma JA, et al. A prospective study on physical activity levels after spinal cord injury during inpatient rehabilitation and the year after discharge. Arch Phys Med Rehabil. 2008;89(11):2094-2101. CrossRef PubMed
24. Moffat F, Paul L. Barriers and solutions to participation in exercise for moderately disabled people with multiple sclerosis not currently exercising: a consensus development study using nominal group technique. Disabil Rehabil. 2019;41(23):2775-2783. CrossRef PubMed
25. Rimmer JH, Riley B, Wang E, Rauworth A, Jurkowski J. Physical activity participation among persons with disabilities: barriers and facilitators. Am J Prev Med. 2004;26(5):419-425. CrossRef PubMed
26. Shirazipour CH, Evans MB, Leo J, Lithopoulos A, Martin Ginis KA, Latimer-Cheung AE. Program conditions that foster quality physical activity participation experiences for people with a physical disability: a systematic review. Disabil Rehabil. 2020;42(2):147-155. CrossRef PubMed
27. Kunstler BE, Cook JL, Freene N, et al. Physiotherapist-led physical activity interventions are efficacious at increasing physical activity levels: a systematic review and meta-analysis. Clin J Sport Med. 2018;28(3):304-315. CrossRef PubMed
28. Keith RA, Cowell KS. Time use of stroke patients in three rehabilitation hospitals. Soc Sci Med. 1987;24(6):529-533. CrossRef PubMed
29. Williams TL, Smith B, Papathomas A. Physical activity promotion for people with spinal cord injury: physiotherapists' beliefs and actions. Disabil Rehabil. 2018;40(1):52-61. CrossRef PubMed
30. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. 2004;6(3):e34. CrossRef PubMed
31. Cellule Planification de l'Offre des Professions des Soins de Santé. Statistiques Annuelles des professionnels des soins de santé en Belgique. Bruxelles: Service public fédéral (SPF) Santé publique, Sécurité de la Chaîne alimentaire et Environnement. Online. Accessed November 2023.
32. Nulty DD. The adequacy of response rates to online and paper surveys: what can be done? Assess Eval High Educ. 2008;33(3):301-314. CrossRef
33. Daniel W. Biostatistics : a foundation for analysis in the health sciences. 7th ed. John Wiley & Sons; 1999.

34. Maisonneuve H, Fournier J-P. Construire une enquête et un questionnaire. E-respect[].2012;1(2):15-21.
35. Sammut R, Griscti O, Norman IJ. Strategies to improve response rates to web surveys: a literature review. Int J Nurs Stud. 2021;123:104058. CrossRef PubMed
36. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. Anesth Analg. 2018;126(5):1763-1768. CrossRef PubMed
37. Harris SR, Winstein CJ. The past, present, and future of neurorehabilitation: from NUSTEP through IV STEP and beyond. Pediatr Phys Ther. 2017;29(suppl 3):S2-S9. CrossRef PubMed
38. Aweto HA, Oligbo CN, Fapojuwo OA, Olawale OA. Knowledge, attitude and practice of physiotherapists towards promotion of physically active lifestyles in patient management. BMC Health Serv Res. 2013;13(1):21. CrossRef PubMed
39. Shirley D, van der Ploeg HP, Bauman AE. Physical activity promotion in the physical therapy setting: perspectives from practitioners and students. Phys Ther. 2010;90(9):1311-1322. CrossRef PubMed
40. Mohan N, Collins E, Cusack T, O'Donoghue G. Physical activity and exercise prescription: senior physiotherapists' knowledge, attitudes and beliefs. Physiother Pract Res. 2012;33(2):71-80. CrossRef
41. Boakye H, Quartey J, Baidoo NAB, Ahenkorah J. Knowledge, attitude and practice of physiotherapists towards health promotion in Ghana. S Afr J Physiother. 2018;74(1):443. CrossRef PubMed
42. Tuna H, Bozan O, Elibol N, Unver B. Are the physical activity habits of Turkish physiotherapists associated with their physical activity promotion and counseling? Physiother Theory Pract. 2022;38(1):189-200. CrossRef PubMed
43. Mulligan H, Fjellman-Wiklund A, Hale L, Thomas D, Häger-Ross C. Promoting physical activity for people with neurological disability: perspectives and experiences of physiotherapists. Physiother Theory Pract. 2011;27(6):399-410. CrossRef PubMed
44. Kennedy W, Curtin C, Bowling A. Access to physical activity promotion for people with neurological conditions: are physical therapists leading the way? Disabil Health J. 2024;17(1):101517. CrossRef PubMed
45. Morris JH, Oliver T, Kroll T, Joice S, Williams B. From physical and functional to continuity with pre-stroke self and participation in valued activities: a qualitative exploration of stroke survivors', carers' and physiotherapists' perceptions of physical activity after stroke. Disabil Rehabil. 2015;37(1):64-77. CrossRef PubMed
46. Zubala A, MacGillivray S, Frost H, et al. Promotion of physical activity interventions for community dwelling older adults: a systematic review of reviews. PLoS One. 2017;12(7):e0180902. CrossRef PubMed
47. Eisele A, Schagg D, Göhner W. Exercise promotion in physiotherapy: a qualitative study providing insights into German physiotherapists' practices and experiences. Musculoskelet Sci Pract. 2020;45:102104. CrossRef PubMed
48. Freene N, Cools S, Bissett B. Are we missing opportunities? Physiotherapy and physical activity promotion: a cross-sectional survey. BMC Sports Sci Med Rehabil. 2017;9(1):19. CrossRef PubMed
49. Prout EC, Mansfield A, McIlroy WE, Brooks D. Patients' perspectives on aerobic exercise early after stroke. Disabil Rehabil. 2017;39(7):684-690. CrossRef PubMed
50. Watson PK, Eitivipart AC, Davis GM, Arora M, Middleton JW, De Oliveira CQ. Effects of behaviour change interventions on physical activity in people with spinal cord injury: a systematic review and meta-analysis. Psychol Sport Exerc. 2023;67:102408. CrossRef PubMed
51. Zhu S, Sherrington C, Jennings M, et al. Current practice of physical activity counselling within physiotherapy usual care and influences on its use: a cross-sectional survey. Int J Environ Res Public Health. 2021;18(9):4762. CrossRef PubMed
52. Declerck L, Stoquart G, Lejeune T, Vanderthommen M, Kaux J-F. Barriers to development and expansion of adaptive physical activity and sports for individuals with a physical disability in sports clubs and centres. Sci Sports. 2021;36(3):202-209. CrossRef
53. Albert FA, Crowe MJ, Malau-Aduli AEO, Malau-Aduli BS. Physical activity promotion: a systematic review of the perceptions of healthcare professionals. Int J Environ Res Public Health. 2020;17(12):4358. CrossRef PubMed
54. Giouridis N, Williams TL, Tomasone JR. Physical activity promotion to persons with spinal cord injury by health and exercise professionals: a scoping review. J Spinal Cord Med. 2023;46(1):126-145. CrossRef PubMed
55. Rethorn ZD, Covington JK, Cook CE, Bezner JR. Physical activity promotion attitudes and practices among outpatient physical therapists: results of a national survey. J Geriatr Phys Ther. 2021;44(1):25-34. CrossRef PubMed
56. Bowen PG, Mankowski RT, Harper SA, Buford TW. Exercise is medicine as a vital sign: challenges and opportunities. Transl J Am Coll Sports Med. 2019;4(1):1-7. CrossRef PubMed

Arch Physioter 2024; 14: 80-82
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3293
EDITORIAL

A decade of growth: preserving the original meaning of research for physiotherapists

Marco Barbero[1], **Andrew Guccione**[2], **Matteo Paci**[3,4]

[1]Rehabilitation Research Laboratory 2rLab, Department of Business Economics, Health and Social Care, University of Applied Sciences and Arts of Southern Switzerland, Manno - Switzerland
[2]George Mason University, Fairfax, Virginia - USA
[3]Department of Experimental and Clinical Medicine, University of Florence, Florence - Italy
[4]Department of Allied Health Professions, Azienda USL Toscana Centro, Florence - Italy

Reflecting on the past does not only mean celebrating milestones but also understanding the path that led us to where we are today. Thirteen years have passed since our journal became a beacon for the Italian physiotherapy community. In 2011, a small visionary group of colleagues from the Società Italiana di Fisioterapia (SIF) recognized the need to foster a solid scientific culture joined to the practical wisdom of clinical experience. From this foresight, the *Italian Journal of Physiotherapy* was born.

During the formative years from 2011 to 2014, the journal entered a significant collaboration with Minerva Medica, a publisher that guided our earliest and most challenging steps in the scientific publishing world. During that period, we struggled with limited resources and low publication numbers similar to other journals, especially in the humanities and social sciences (1). This crucial phase saw the release of four journal volumes, each with four quarterly issues per year. It was a period marked by diligent learning and growth, during which the commitment of our authors, the critical insights of our reviewers, and the leadership of Roberto Gatti as editor-in-chief were pivotal in establishing the journal within the physiotherapy community. Our vision of evidence-based practice, not as a merely academic ideal but as a cornerstone of everyday clinical practice, which was first articulated in our inaugural editorial in 2011, has consistently guided our publications (2). This commitment has always been accompanied by an unwavering focus on methodological rigor and transparency in reporting, principles that are essential to scientific research and shared by the entire editorial board.

In 2015, as we aspired to be an integral part of the borderless international physiotherapy community, the *Italian Journal of Physiotherapy* began expanding from its national audience to an international stage. This expansion reached a turning point when the newly appointed editor-in-chief, Marco Baccini, embarked on a new challenge by initiating a collaboration with BioMed Central, a large open access publisher owned by Springer Nature that produces over 250 scientific journals. The goal was to transform our "national" journal into an "international" one (3). Partnering with BMC brought numerous benefits, including increased visibility and more efficient dissemination. Most importantly, it allowed us to publish and distribute our articles under the terms of the CC BY 4.0 License (Creative Commons Attribution 4.0 International License), fully aligning the journal with the open science initiative (4). This collaboration also marked the relaunch of our journal as the *Archives of Physiotherapy* (AoP).

After four years, Marco Barbero was appointed as the new editor-in-chief. To ensure the highest quality and efficiency in the peer review process, the AoP board was significantly expanded to include more than 60 world-renowned experts. Additionally, the editorial board was restructured into sections reflecting some of the main areas of physiotherapy (Musculoskeletal, Neurology, Geriatrics, Research Methodology and Clinimetrics, Biomechanics, and Movement Analysis). Expert section editors, along with teams of associate editors, were appointed in each of these identified areas to lead the review process. The aim of these changes was twofold. First, we sought to improve the viability of the peer review progress by redistributing the workload more evenly within the editorial board, as the number of annual submissions had exceeded the considerable figure of 100. Secondly, and more importantly, we wanted authors and readers to benefit from the expertise of specialists who could review manuscripts with clinical knowledge and experience in the areas mentioned. This effort was considered crucial to ensure the external validity and clinical utility of published papers, a key aspect that is often overlooked in peer-reviewed publications (5). Two new article types, Viewpoints and Masterclasses, were also introduced. These additions enriched the AoP by providing space for expert opinions and advanced educational content, thus fulfilling its mission to advance the field of physiotherapy. The collaboration with BMC Springer proved to be highly productive. Between 2015 and 2023, we published nine volumes comprising a total of 170 papers with an average rejection rate around 70%, a percentage in line with that of biomedical journals and not far

Received: September 6, 2024
Accepted: September 9, 2024
Published online: October 3, 2024

Corresponding author:
Marco Barbero
email: marco.barbero@supsi.ch

from that of top-tier journals (6). These articles have accumulated over 1,300 citations, highlighting the impact and reach of our authors.

However, a partnership between a society with a single journal and a global publisher managing hundreds of journals has its own challenges, both from a financial and day-to-day editorial management perspective. Therefore, at the beginning of 2024, we ceased our collaboration with BMC Springer and transitioned to AboutScience, a smaller publisher. The adoption of the Open Journal System (OJS), an open-source platform for online journal publishing used by more than 11,500 journals in 2012 (7) and currently exceeding 25,000, was crucial to maintain the financial sustainability of our editorial enterprise. This approach has allowed us to invest more significantly in the diamond open access model, a fundamental consideration for SIF and the Editorial Board as well as for all our funding partners (Federazione Nazionale Ordine Fisioterapisti, Ordine Fisioterapisti Lombardia, Scuola universitaria professionale della Svizzera italiana, Associazione Italiana di Fisioterapia) who have made every effort to ensure that our publication remains freely accessible to all and without any publication fees for authors. Furthermore, working with a smaller publisher allows for a closer, more dynamic partnership. We anticipate that this collaboration will foster innovation and enable us to more effectively address the numerous challenges of modern scientific publishing.

The collaboration is off to a good start. In June, Clarivate announced our journal's first impact factor of 2.1 and placed AoP in the Q1 category for Rehabilitation. In addition, Elsevier's CiteScore has increased significantly, from 2.9 in 2022 to 3.6 in 2023. Both metrics underscore the growing influence and reputation of AoP within the international physiotherapy community and positions the journal among the leading journals in the rehabilitation field.

It has been a long journey, lasting more than 10 years, fostering slow but solid growth, and we are clearly proud of this important achievement but at the same time we look to the future with awe. A speed beyond imagination has been injected into the world of scientific publishing and production has grown at impressive rates. In 2022, approximately 3.3 million scientific articles were published globally and (8) according to a recent study, the global growth rate of scientific production is such that it doubles every 17.3 years (9). The field of physiotherapy is no exception, and we must question its meaning.

Phenomena such as predatory journals, mega-journals, and paper mills are clear examples of the drifts of a market increasingly polluted by financial interests and lucrative publishing models. Predatory journals exploit researchers by charging high fees to publish their articles without providing adequate peer review, thus diluting the quality of published research (10). Meanwhile, mega-journals, which publish a vast number of articles with less rigorous selection criteria, contribute to the proliferation of less impactful research, potentially overwhelming researchers and clinicians with information of variable utility (11). Finally, paper mills produce fraudulent research for profit, often fabricating data, authorship, and entire studies, thereby undermining the integrity of scientific literature (12). The credibility of the scientific publishing world is threatened by phenomena typical of consumer-driven markets, where the relentless pursuit of growth often leads to compromises in quality. The context

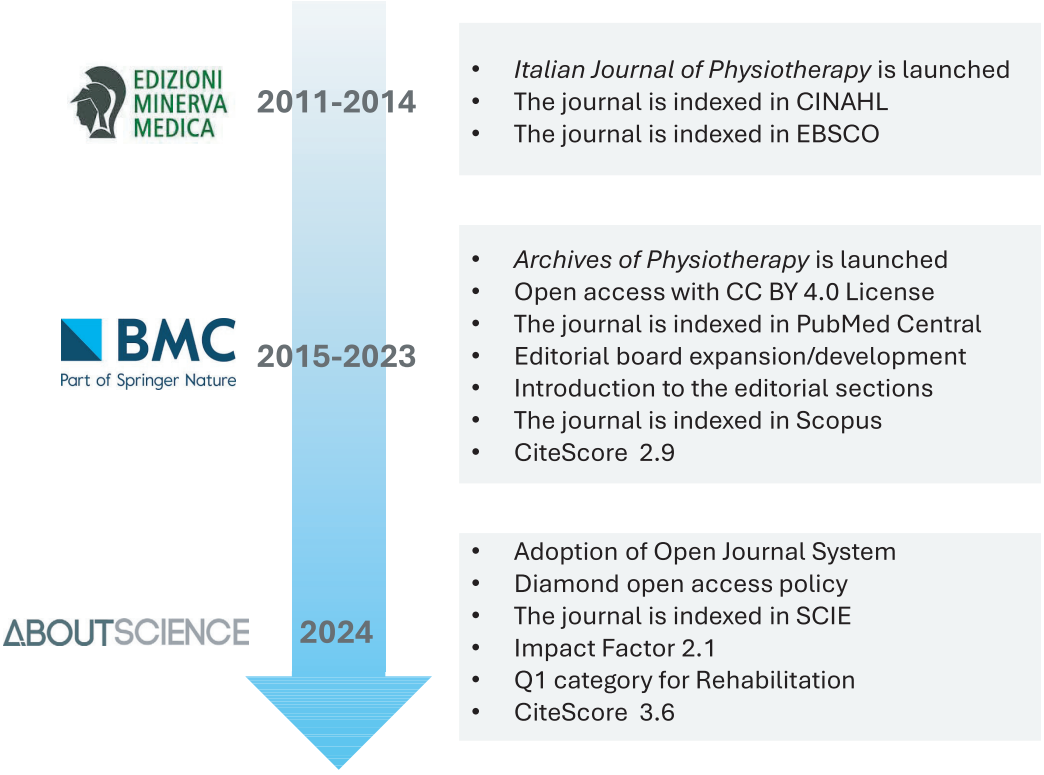


FIGURE 1 - Key milestones and achievements of the journal from 2011 to the present.

is complex and will become even more so with the inevitable adoption of artificial intelligence tools by researchers, which will boost researchers' outputs but not necessarily improve quality (13). But it is important to reflect on the fact that not only is the quality of our scientific publications in danger of being corrupted, but research itself is in danger of losing its original purpose. We view research as a unique opportunity to deeply understand the complexities of physiotherapy practice and ultimately improve our interventions. Having said that, how can we ensure that the AoP continues to uphold the original purpose of research while contributing to the improvement of clinical practice in the physiotherapy community?

It is perhaps that the greatest value we have created lies within our editorial board. Representing the physiotherapy community, our expert and dedicated board members act as gatekeepers against the market forces described earlier as originally highlighted by Zsindely and colleagues (14). Their knowledge, diversity, and expertise not only ensure the quality of the peer review process (15) but also guarantee the integrity and preservation of the original intent of the research published in AoP. This work of oversight and assurance is vitally supported by our reviewers, whose contributions are fundamental to maintaining high standards of quality in our publication. In an era where scientific production is growing at an ever-increasing rate and the risks associated with questionable editorial practices are on the rise, the role of the editorial board becomes even more crucial. Our members are not only called upon to rigorously assess the quality and validity of the research but also to serve as ethical guides, promoting a culture of transparency and responsibility. However, this priority on quality over quantity is made possible, promoted, and shared not only by our board but also by the publisher we have chosen for their commitment to these values, and by the societies that support us. In this sense, the AoP editorial board is not just a guarantor of quality but a flagship for the entire physiotherapy community, committed to upholding the value of research and safeguarding its original purpose. This collective commitment, supported and shared by our editorial partners and funders, will be essential in meeting future challenges and ensuring that the AoP continues to make a meaningful contribution to the improvement of clinical practice within the physiotherapy community. In doing so, we will also preserve the vision of the SIF and of all its original founders.

References

1. Rodríguez-Yunta L, Giménez-Toledo E. Fusión, coedicion o reestructuración de revistas científicas en humanidades y ciencias sociales. Prof Inf. 2013;22(1):36-45. CrossRef
2. Gatti R. The Italian Journal of Physiotherapy: the S.I.F. voice in order to participate at the International Physiotherapists Community. Italian J Physiother. 2011;1(1):1-2. Online. Accessed September 2024.
3. Baccini M, Barbero M, Gatti R. From a national to an international journal: a new opportunity for the physiotherapy community. Arch Physiother. 2015;5(1):1. CrossRef PubMed
4. Ramjoué C. Towards Open Science: the vision of the European Commission. Inf Serv Use. 2015;35(3):167-170. CrossRef
5. Peters J, Langbein J, Roberts G. Policy evaluation, randomized controlled trials, and external validity a systematic review. SSRN Electronic J. 2016. CrossRef
6. Ali MJ. The science and philosophy of manuscript rejection. Indian J Ophthalmol. 2021;69(7):1934-1935. CrossRef PubMed
7. Owen B, Stranack K. The Public Knowledge Project and Open Journal Systems: open source options for small publishers. Learn Publ. 2012;25(2):138-144. CrossRef
8. NSF. Science and engineering indicators. publications output: U.S. trends and international comparisons. 2023. Online. Accessed September 2024.
9. Bornmann L, Haunschild R, Mutz R. Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases. Humanit Soc Sci Commun. 2021;8(1):224. CrossRef
10. Beall J. Predatory publishers are corrupting open access. Nature. 2012;489(7415):179. CrossRef PubMed
11. Ioannidis JPA, Pezzullo AM, Boccia S. The rapid growth of mega-journals: threats and opportunities. JAMA. 2023;329(15):1253-1254. CrossRef PubMed
12. Candal-Pedreira C, Guerra-Tort C, Ruano-Ravina A, et al. Retracted papers originating from paper mills: a cross-sectional analysis of references and citations. J Clin Epidemiol. 2024;172:111397. CrossRef PubMed
13. van Dijk SHB, Brusse-Keizer MGJ, Bucsán CC, van der Palen J, Doggen CJM, Lenferink A. Artificial intelligence in systematic reviews: promising when appropriately used. BMJ Open. 2023;13(7):e072254. CrossRef PubMed
14. Zsindely S, Schubert A, Braun T. Citation patterns of editorial gatekeepers in international chemistry journals. Scientometrics. 1982;4(1):69-76. CrossRef
15. Wu D, Lu X, Li J, Li J. Does the institutional diversity of editorial boards increase journal quality? The case economics field. Scientometrics. 2020;124(2):1579-1597. CrossRef

Arch Physioter 2024; 14: 83-88
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3049
ORIGINAL RESEARCH ARTICLE

Intra- and inter-rater reliability of goniometric finger range of motion using a written protocol

Takuya Nakai[1], Satoru Amano[2,3], Chikako Murao[1], Haruki Taguchi[1], Kayoko Takahashi[2,3]

[1]Department of Occupational Therapy, Kitasato University Hospital, Kanagawa - Japan
[2]School of Allied Health Sciences, Kitasato University, Kanagawa - Japan
[3]Graduate School of Medicine, Kitasato University, Kanagawa - Japan

ABSTRACT

Introduction: Goniometric finger range of motion (ROM) is the most common outcome measure used for functional evaluation of finger joints, but its reliability is not well-evaluated. This study aimed to investigate intra- and inter-rater reliability of goniometric finger ROM using a written protocol for active, passive, and composite movements in healthy adults.
Methods: The design was a single-center, cross-sectional, reliability study. Participants were 20 healthy adults (mean ± standard deviation, 36.4 ± 10.9 years). ROM for active, passive, and composite movements of the fingers was assessed by three occupational therapists with at least 5 years clinical experience in the field of physical disabilities. To standardize the measurement method used, we developed a written protocol, stabilized the wrist position, and trained the evaluators. Intraclass correlation coefficient (ICC) values were used for the reliability analysis. ICC (1,1) was used for intra-rater reliability. ICC (2,1) was used for inter-rater reliability. Hand-shaped heatmaps were used to summarize the reliability data.
Results: Most of the results (88.7%) showed moderate to good intra-rater reliability (ICC ≥ 0.50), while inter-rater reliability showed less (69.0%). Both intra- and inter-rater reliability showed no trends between dominant and non-dominant hands, type of movement, finger, or joint.
Conclusions: Intra-rater reliability was relatively high and using a written protocol was beneficial. Inter-rater reliability tended to be lower, and differences in the physical structure of both raters and participants may have affected inter-rater reliability values.

Keywords: Finger, Range of motion, Reliability, Reproducibility, Standardization

What is already known about this topic:

- *Goniometric finger range of motion (ROM) is the most common outcome measure used for functional evaluation of finger joints.*
- *However, the intra- and inter-rater reliability of finger ROM is not well-evaluated.*

What does the study add:

- *Relative intra-rater reliability was relatively high and using a written protocol was beneficial.*
- *Differences in the physical structure of raters and participants may have affected inter-rater reliability values.*
- *The results of ROM cannot be interpreted in terms of absolute reliability at 2-degree and 5-degree increments.*

Introduction

The fingers are indispensable for performance of tasks. These sophisticated body parts have motor (e.g., grasping and releasing) and sensory (e.g., touching and adjusting) functions.

Received: February 2, 2024
Accepted: September 17, 2024
Published online: October 8, 2024

This article includes supplementary material

Corresponding author:
Kayoko Takahashi
email: kayo.ot@kitasato-u.ac.jp

Range of motion (ROM) is one measure used for functional evaluation of the finger joints (1). When restrictions occur due to disease or disability, ROM is useful for understanding the patient's joint condition, observing changes over time, and evaluating the outcome of an intervention (2). ROM assessment is also frequently used during post-stroke upper limb rehabilitation (3). There is a consensus that ROM should be used for musculoskeletal injuries (4). Santisteban et al's (3) review found that ROM is not only a traditional tool. It remains a first choice for measurement of outcomes associated with the body function categories of the International Classification of Functioning, Disability, and Health. In addition, due to the current emphasis on evidence-based medicine, the need for objective and reliable measures is increasing rapidly.

There are only a few standardized protocols available for finger ROM measurement (e.g., "Methods for Indication and Measurement of Joint Range of Motion" by the Japanese Orthopaedic Association and the Japanese Society of Rehabilitation Medicine (5), *Measurement of Joint Motion: A Guide to Goniometry*, fifth edition by Norkin and White (6)). However, other than definition of the basic and moving axes, some procedures of measurement are not consistent among references. Therefore, repetition of measurements and limb positions can vary across examiners. In clinical settings, examiner bias can be high because therapists commonly use the goniometer manually. Although several previous studies have been reported on the reliability of finger ROM measurement using goniometers, most of them were limited to the certain fingers/joints (5-9) and movement type (10).

Sato et al (11) examined intra- and inter-rater reliability of finger ROM at 2- versus 5-degree intervals. They found that the error was smaller for the 2-degree interval measurement than for the 5-degree interval measurement. This result suggested that smaller angle changes can be captured using a goniometer with smaller measurement intervals. Therefore, it is necessary to verify intra- and inter-rater reliability for all fingers, joints, and types of movement (active, passive, and composite). Thus, the purpose of this study was to investigate the intra- and inter-rater reliability of goniometric finger ROM using a written protocol for active, passive, and composite movements in healthy adults.

Methods

Research design

We used an observational, descriptive study design to examine the intra- and inter-rater reliability of a new protocol for goniometric measurement of finger motions. The risk of bias of the present study was assessed using the COSMIN checklist (Reliability: relative measures) in the supplementary tables. The Kitasato University School of Medicine and Hospital Ethics Committee (2020-027) approved this study.

Participants

The participants were recruited from among the staff members of the hospital where the first author was employed. The exclusion criteria were as follows: (1) history of musculoskeletal condition, such as arthritis, orthopedic conditions involving the upper limbs, (2) neurological, (3) psychiatric conditions, and (4) an unstable general condition due to other complications.

Evaluator

Finger ROM was assessed by three occupational therapists (TN, CM, HT) with ≥5 years of clinical experience in the field of physical disabilities (Rater A/B/C, mean years of experience: 8.3 years).

Procedure

We developed a measurement protocol manual that was based on "Joint Range of Motion Indication and Measurement Methods" by the Japanese Orthopaedic Association and Japanese Society of Rehabilitation Medicine (5) and *Measurement Evaluation for PT/OT: ROM Measurement*, Second Edition (2). To ensure uniformity of the measurement method used, raters received a 15-minute course on the contents of the written protocol and trained for 15 minutes individually using the measurement manual.

Each participant was seated in a chair facing the table with the assessed side of the arm placed on the table. The forearm position was 0-degree rotation with a 20-degree wrist dorsiflexion. A sheet of paper with a diagram of the basic fixed axis was placed under the arm as a guide (Fig. 1A). The goniometer was placed from the dorsal side of the hand with the long handle (with fixed axis) on the basic axis and the short handle (with meter printed) on the moving axis (Fig. 1B). The thumb was measured first, followed by the index, middle, ring, and little fingers. Measurement of each finger followed the order of metacarpophalangeal (MP), proximal interphalangeal (PIP), and distal interphalangeal (DIP) joints.

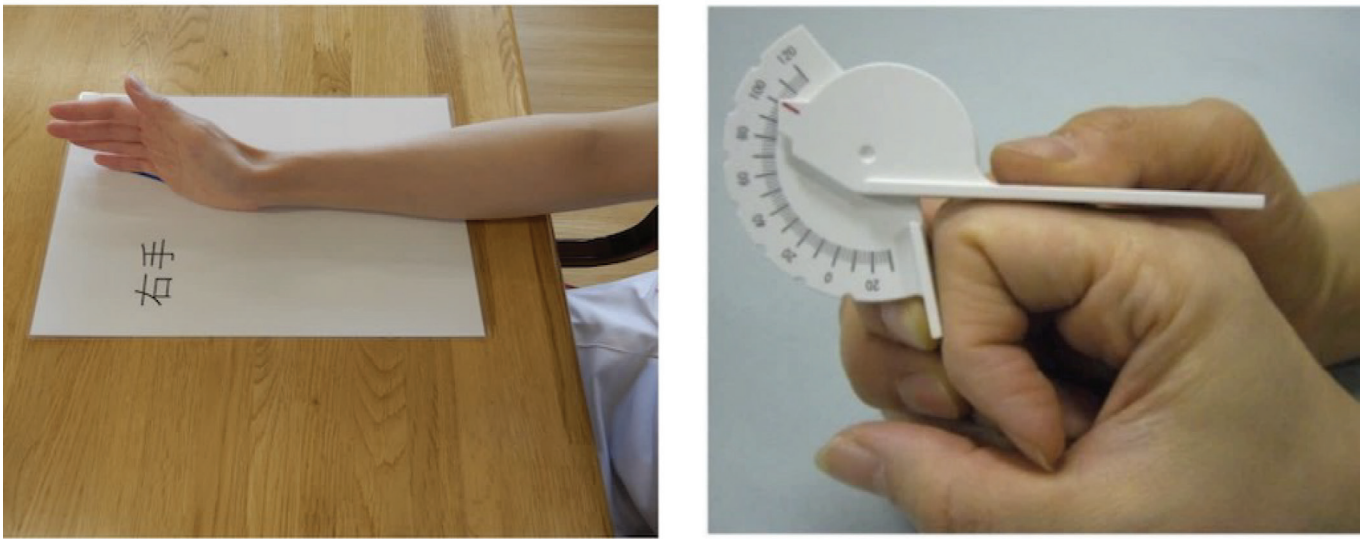
First, active (voluntary) movement was measured with the accompanying verbal instruction, "Please bend XX joint of your XX finger utmost, without moving your wrist." If other fingers were flexed at the same time, the raters instructed the participant to "try to move only your XX (targeted) finger." Second, passive ROM was measured in the same order, with the instructions, "Please relax and let me bend your XX finger's XX joint to the maximum." While measuring the MP joint, extreme flexion of the interphalangeal (IP) joint was avoided, and it kept its natural orientation. The MP and DIP joints were straightened (0-degree flexion/extension) during PIP joint measurement. When the DIP joint was measured, the MP joint was straightened (0-degree flexion/extension) with the PIP joint flexed at 70-90 degrees.

Last, active composite movements of all finger flexion positions were performed following the same orders. The instructions were, "Please bend all fingers utmost without moving your wrist." The thumb was placed closely over the basal phalange of the index finger to avoid interfering with ROM of the other fingers. If the goniometer could not fully contact the joint, we allowed measurement on a line parallel to the basic axis and axis of movement.

All three raters measured all participants twice with at least 24-hour interval to test intra-rater reliability. For inter-rater reliability, the dates of assessment were distributed so a participant was not assessed by more than one rater on the same day. Before each assessment, it was confirmed with the participants that there had been no injury or change in hand function since the last assessment. Assessment was conducted individually in a separate room to ensure the other raters were blinded, and discussion or comparison between rates was strictly prohibited.

Data analysis

Intraclass correlation coefficients (ICCs) were used for the relative reliability analysis (ICC (1,1) for intraclass reliability, ICC (2,1) for inter-rater reliability) (12). R (version 4.0.2) was used for the statistical analysis. We used heatmaps to summarize the reliability data because the study included a large number of values, based on 366 ICC calculations. Heatmaps were



(A) (B)

FIGURE 1 - Goniometric measurement of finger range of motion. A) Alignment of wrist during finger measurement. The axis of movement and basic axis of wrist dorsiflexion are shown on a sheet placed on a desk, so that the 20-degree dorsiflexion fixation is not displaced during measurement. The paper with both the fixed and moving axes was placed under the arm to stabilize the 20-degree dorsiflexion of the measured arm. B) Placement of goniometer on finger. The goniometer was placed from the dorsal side of the hand with the long handle (with fixed axis) on the basic axis and the short handle (with meter printed) on the moving axis. Note: Numerical values are measured to the first digit in 2-degree increments.

also of great value for presentations based on the shape of the hand. However, because heatmaps alone did not include all necessary information, we provide ICC precision data for more in-depth interpretation (Supplementary Table). In addition, minimal detectable change (MDC) was calculated for absolute reliability. The standard error of measurement (SEM) was used to calculate an MDC value with the following formula: $MDC_{95} = 1.96*\sqrt{(2)}*(SEM)$. A SEM value was calculated as $\sqrt{\sigma^2_{error}}$ (square root of the error variance) (13).

Results

Participant demographic characteristics

Twenty healthy adults were included in this study; no participants met the exclusion criteria and no data were missing. The mean ± standard deviation age of the participants was 36.4 ± 10.9 years (33.8 ± 8.3 years) for males and 40.3 ± 13.1 years for females, 40% were female, and 90% were right-handed (Tab. 1).

TABLE 1 - Demographic characteristics of participants (N = 20)

Characteristics	N (%)
Gender, N (%)	
Male	12 (60)
Female	8 (40)
Age, mean (SD)	36.4 (10.9)
Dominant hand, N (%)	
Right	18 (90)
Left	2 (10)

SD = standard deviation.

Relative intra- and inter-rater reliability

Figure 2-5 presents the results for the heatmap of intra-rater reliability of each rater and inter-rater reliability among the three raters. A darker red color indicated a higher ICC value; a lighter color indicated a lower ICC value. Both intra- and inter-rater reliability values showed no trends between dominant and non-dominant hand, type of movement, finger, or joint. Rater C's heatmap tended to be lighter than that of rater A or B. Reliability results varied among the different raters. Compared with intra-rater reliability (Figure 2-4), ICC values for inter-rater reliability were generally low (Figure 5). Detailed ICC information, including precision data, is presented in the supplementary tables.

Absolute intra- and inter-rater reliability

Both intra- and inter-rater reliability values showed no clear trends between dominant and non-dominant hand, type of movement, finger, or joint. Absolute reliability varied depending on the different evaluators, but in many cases MDC fitted between 10 and 15. Compared with intra-rater reliability, MDC values for inter-rater reliability were generally high. Detailed MDC and SEM information, including precision data, is presented in the supplementary tables.

Discussion

This study examined the intra- and inter-rater reliability of goniometric finger ROM measurements with ICC using a written protocol for various type of movements in healthy adults. Koo and Li (12) define moderate reliability (ICC 0.5-0.75), good reliability (ICC 0.75-0.90), and excellent reliability

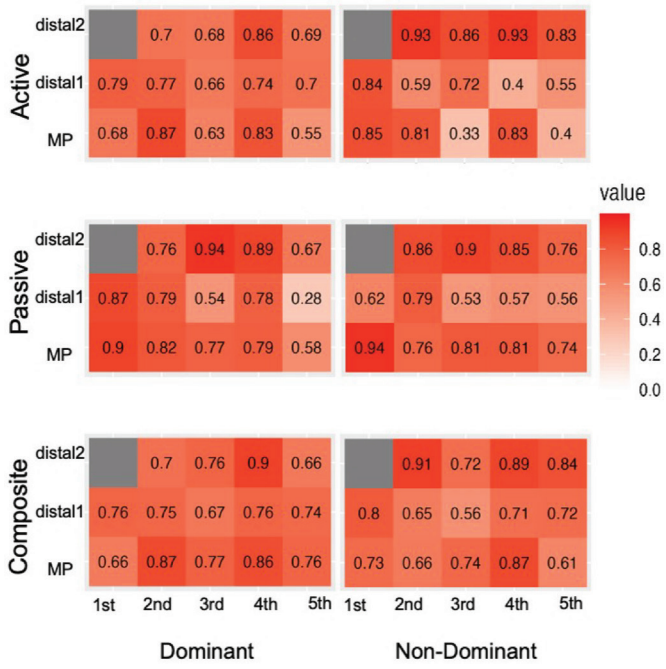


FIGURE 2 - Intraclass correlation coefficient (ICC) values for intra-rater A reliability. Darker red color indicates higher ICC, lighter color indicates lower ICC. The number represents the type of finger. Detailed ICC information and standard error of the measurement (SEM), including precision data, are presented in the supplementary tables.

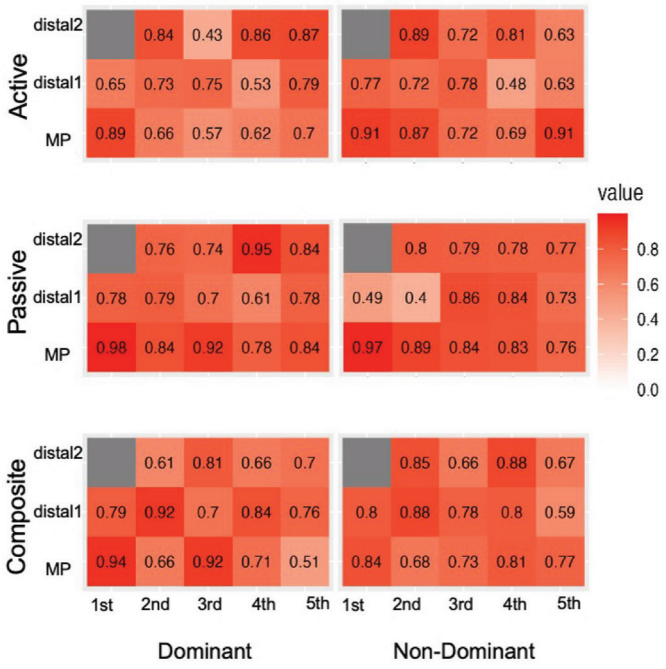


FIGURE 3 - Intraclass correlation coefficient (ICC) values for intra-rater B reliability. Darker red color indicates higher ICC, lighter color indicates lower ICC. The number represents the type of finger. Detailed ICC information and standard error of the measurement (SEM), including precision data, are presented in the supplementary tables.

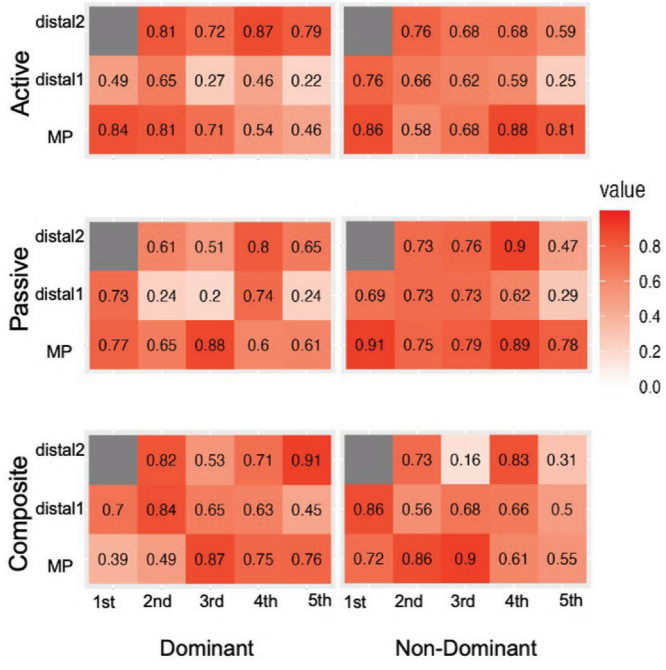


FIGURE 4 - Intraclass correlation coefficient (ICC) values for intra-rater C reliability. Darker red color indicates higher ICC, lighter color indicates lower ICC. The number represents the type of finger. Detailed ICC information and standard error of the measurement (SEM), including precision data, are presented in the supplementary tables.

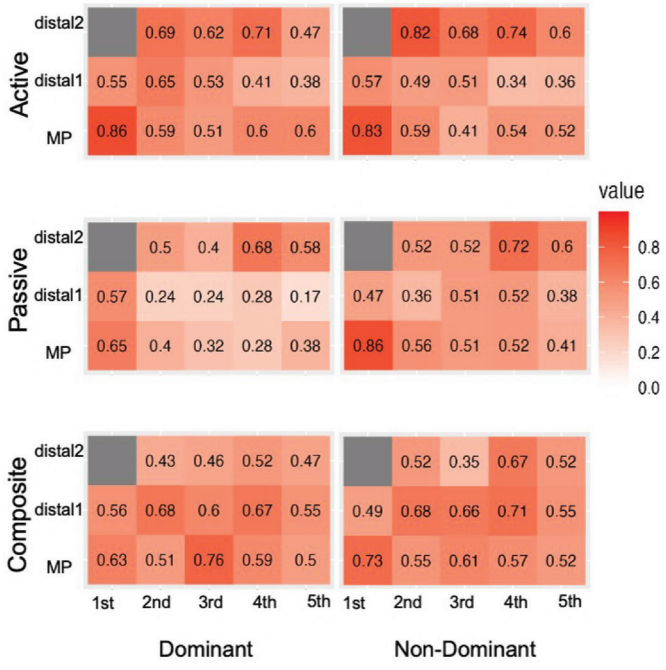


FIGURE 5 - Intraclass correlation coefficient (ICC) values for inter-rater reliability. Note: Darker red color indicates higher ICC, lighter color indicates lower ICC. The number represents the type of finger. Detailed ICC information and standard error of the measurement (SEM), including precision data, are presented in the supplementary tables.

(ICC ≥0.90). This study had a certain degree of reliability in intra-rater reliability. Whereas the ICC tended to have lower inter-rater reliability than intra-rater reliability, the results supported previous studies.

Relative intra-rater reliability

Heatmap analysis revealed a constant dark red color that indicated the presence of a relatively high intra-rater reliability. There were only a few differences in reliability, depending on the type of movement (active, passive/composite), dominant or non-dominant hand, and each finger and each joint. Lewis et al (10) examined intra-rater reliability of the MP, PIP, and DIP joints of the middle finger of the dominant hand in 20 healthy adults. The raters were 10 therapists using Rolyan goniometers to measure both active and passive movement. The ICC values ranged from 0.43 to 0.99. The rater with the highest reliability had ICC values of 0.84-0.99; the rater with the lowest reliability had ICC values of 0.43-0.84. In this study, rater A had the highest reliability (ICC 0.66-0.90 for active composite movement). Thus, the results of this study had acceptable reliability. A certain degree of intra-rater reliability was achieved because we developed a measurement protocol and used raters who were trained to ensure good reproducibility.

Relative inter-rater reliability

For inter-rater reliability, heatmap analysis revealed lighter red color than intra-rater reliability that indicated inter-rater reliability was relatively low compared with the ICC values of intra-rater reliability. Similar results for low inter-rater reliability for finger ROM measurements, compared with intra-rater reliability, have been published (9,10,14). Lewis et al (10) found that inter-rater reliability is lower than intra-rater reliability with ICC values in the range of 0.35-0.85. They also found that errors in ROM angle were due to biarticular muscles and short DIP joints. Ellis et al (14) found that inter-examiner measurements are less reliable than intra-examiner measurements for the comparative reliability of finger ROM measurements using goniometry and wire tracing. They included the amount of force applied to the goniometer, the accuracy of alignment during goniometer application, and identification of anatomical landmarks as reasons for inconsistent measurement outcomes with respect to errors in goniometer measurements. Short et al (15) mentioned that the size of the rater's body (height difference) may affect the interpretation of goniometer readings. In our study, the maximum palm lengths of each rater varied from 19.5, 17.5, and 16.3 cm (average 17.8 cm), and the hand size of each participant also varied. Handling difficulties due to differences in the body structure of both raters and participants may have affected measurement consistency.

Absolute reliability

Measurement error was considered as absolute reliability. Even if the interpretation of relative reliability was acceptable, the results of absolute reliability may not be clinically acceptable. However, rather than clearly judging it to be "clinically unusable," we would like to recommend that medical professionals leave it to the "system" for interpreting ROM. The Mayo Wrist Score (16) is a good example of a practice that takes this approach. In section 3 of the assessment (regarding ROM), the assessment is based on an ordinal scale in increments of approximately 25%, with emphasis on % normal. Even if the ROM is clinically acceptable in terms of relative reliability, medical clinicians should pay attention to the results of this study, which show that the results cannot be interpreted in terms of absolute reliability at 2-degree and 5-degree increments.

Strength of this study

The strength of our study is that we verified the reliability of all active, passive, and composite movements of all joints in all fingers of the participants' dominant and non-dominant hands. In previous studies (7,8), the validation was limited to certain fingers, joints, and types of movement, and this study was the first to compare and validate the results by all joints, fingers, and types of movement. In the clinical setting, ROM should be measured at all affected joints and fingers, and ROM of different types of movement would help define the problem and plan the intervention. Therefore, the results of this study contributed to the field of hand therapy by validating all fingers, all joints, and various types of movements. The results also indicated that a certain degree of intra-rater reliability was obtained.

As with other assessments (17,18), the creation of a manual to reduce variation in measurement methods among raters may have contributed to a certain reliability. In our study, ROM was measured using a written protocol, and multiple trainings were conducted among raters. These components could have helped standardize the measurement methods and improved reliability. ROM angle is significantly affected by the position of the proximal joint. Thus, our manual, with its concrete description of wrist position, could have minimized rater bias and error.

Limitations and direction for future research

One of the limitations of this study is the sample size. According to Borg et al (19), a sample size estimation for a reliability study with three raters requires an ICC planning value of 0.8, a minimum acceptable reliability of 0.6, a power of 80%, and an alpha equal to 0.05, with a necessary sample size of 33 patients. However, the small variability observed in ROM scores in this study may have mitigated the impact of the small sample size on the reliability results. For these reasons, future studies with larger sample sizes are warranted to confirm our findings, particularly in cases involving diseases or pathologies that result in limitations in hand ROM.

The ROM measurement procedure was designed to measure all types of movement of both the dominant and non-dominant hands and to measure all fingers and joints twice; 30 to 40 minutes were required to measure ROM for each participant. This time constraint could have negatively affected rater concentration and the ability to accurately interpret the goniometer scale. Future research should be modified to better reflect actual clinical settings.

This study was a single-center study, and future validation in multicenter studies are recommended. It is also possible that differences in the physical structure of both the raters and participants affected inter-rater reliability. Future validation studies should consider the effects of different body structures of both raters and participants.

Conclusions

This study examined the intra- and inter-rater reliabilities of finger ROM in healthy adults using a finger goniometer. The results indicated that relative intra-rater reliability was relatively acceptable and that inter-rater reliability tended to be lower than intra-rater reliability. In clinical practice, having the same rater is recommended to achieve a certain degree of reliability, regardless of the type of movement or joint, and to capture finger ROM changes over time.

Acknowledgments

This study was supported by the AMED under Grant Number JP21he1302034j0003.

Disclosures

Conflict of interest: There is no conflict of interest for all authors.

Financial support: This study was supported by the AMED under Grant Number JP21he1302034j0003.

Authors' contributions: Conceptualization: NT, SA, KT; Data Curation: NT, CM, HT; Formal Analysis: NT, SA; Funding Acquisition: KT; Investigation: NT, SA, KT; Methodology, NT, SA, KT; Project Administration: NT, CM, HT, SA, KT; Resources: KT, NT; Supervision: KT; Validation: SA; Visualization: SA, KT; Writing – Original Draft: NT, KT; Writing – Review and Editing: NT, CM, HT, SA, KT.

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

References

1. Shiraishi H. What is the evaluation of hand function from various evaluations on the hand. SOBIM Japan. 2010;34(4):291-296. CrossRef
2. Fukuda O. Measurement evaluation for PT/OT: ROM measurement. 2nd ed. MIWA SHOTEN; 2010.
3. Santisteban L, Térémetz M, Bleton J-P, Baron JC, Maier MA, Lindberg PG. Upper limb outcome measures used in stroke rehabilitation studies: a systematic literature review. PLoS One. 2016;11(5):e0154792. CrossRef PubMed
4. Crasto JA, Sayari AJ, Gray RR-L, Askari M. Comparative analysis of photograph-based clinical goniometry to standard techniques. Hand (N Y). 2015;10(2):248-253. CrossRef PubMed
5. Yonemoto K, Isigami S, Kondo T. Joint range of motion and measurement methods. Jpn J Rehabil Med. 1995;32(4):207-217. CrossRef
6. Norkin CC, White DJ. Measurement of joint motion: a guide to goniometry. 5th ed. F A Davis Co; 2016.
7. Hamilton GF, Lachenbruch PA. Reliability of goniometers in assessing finger joint angle. Phys Ther. 1969;49(5):465-469. CrossRef PubMed
8. Burr N, Pratt AL, Stott D. Inter-rater and intra-rater reliability when measuring interphalangeal joints: comparison between three hand-held goniometers. Physiotherapy. 2003;89(11):641-652. CrossRef
9. Ellis B, Bruton A. A study to compare the reliability of composite finger flexion with goniometry for measurement of range of motion in the hand. Clin Rehabil. 2002;16(5):562-570. CrossRef PubMed
10. Lewis E, Fors L, Tharion WJ. Interrater and intrarater reliability of finger goniometric measurements. Am J Occup Ther. 2010;64(4):555-561. CrossRef PubMed
11. Sato A, Oi H, Abe Y. Reliability of hand joint angle measurements using the manual for hand joint range of motion measurements. Jpn Soc Surg Hand. 2015;31(4):512-516.
12. 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
13. Stratford PW, Goldsmith CH. Use of the standard error as a reliability index of interest: an applied example using elbow flexor strength data. Phys Ther. 1997;77(7):745-750. CrossRef PubMed
14. Ellis B, Bruton A, Goddard JR. Joint angle measurement: a comparative study of the reliability of goniometry and wire tracing for the hand. Clin Rehabil. 1997;11(4):314-320. CrossRef PubMed
15. Short N, Almonreoder T, Mays M, et al. Interrater reliability of a novel goniometric technique to measure scapular protraction and retraction. Am J Occup Ther. 2022;76(1):7601205010. CrossRef PubMed
16. Cooney WP, Bussey R, Dobyns JH, Linscheid RL. Difficult wrist fractures. Perilunate fracture-dislocations of the wrist. Clin Orthop Relat Res. 1987;(214):136-147. PubMed
17. Platz T, Pinkowski C, van Wijck F, Kim I-H, di Bella P, Johnson G. Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer Test, Action Research Arm Test and Box and Block Test: a multicentre study. Clin Rehabil. 2005;19(4):404-411. CrossRef PubMed
18. See J, Dodakian L, Chou C, et al. A standardized approach to the Fugl-Meyer assessment and its implications for clinical trials. Neurorehabil Neural Repair. 2013;27(8):732-741. CrossRef PubMed
19. Borg DN, Bach AJE, O'Brien JL, Sainani KL. Calculating sample size for reliability studies. PM R. 2022;14(8):1018-1025. CrossRef PubMed

Arch Physioter 2024; 14: 89
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3326
ERRATUM

Erratum in: Pragmatism in manual therapy trials for knee osteoarthritis: a systematic review

Kyle R. Adams[1,2], Ayodeji O. Famuyide[1,3], Jodi L. Young[1], C. Daniel Maddox[4,5], Daniel I. Rhon[1,6]

[1]Physical Therapy Program, Bellin College, Green Bay, Wisconsin - USA
[2]Department of Physical Therapy, Baylor University, Waco, Texas - USA
[3]Greater Baton Rouge Physical Therapy, Baton Rouge, Louisiana - USA
[4]Department of Physical Therapy, Ivester College of Health Sciences, Brenau University, Gainesville, Georgia - USA
[5]Upstream Rehab Institute, Smyrna, Georgia - USA
[6]Department of Rehabilitation Medicine, School of Medicine, Uniformed Services University of Health Sciences, Bethesda, Maryland - USA

In the article "Pragmatism in manual therapy trials for knee osteoarthritis: a systematic review" it was reported post-publication that two of the studies included were secondary analyses of other studies that the authors had already included. Additionally, the authors modified the reasons for excluding two other studies. The overall conclusions of the systematic review do not change.

We apologize to the readers. The final version of this article, which has been edited to reflect these changes is available online and includes a reference to this correction.

References

1. Adams KR, Famuyide AO, Young JL, Maddox CD, Rhon DI. Pragmatism in manual therapy trials for knee osteoarthritis: a systematic review. Arch Physiother. 2024 Feb 26;14:1-10. CrossRef. PubMed.

Received: October 2, 2024
Accepted: October 2, 2024
Published online: October 21, 2024

Corresponding author:
Kyle R. Adams
email: Kyle_Adams@Baylor.edu

Arch Physioter 2024; 14: 90-95
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3267
ORIGINAL RESEARCH ARTICLE

Indoor and outdoor 10-Meter Walk Test and Timed Up and Go in patients after total hip arthroplasty: a reliability and comparative study

Federico Temporiti[1,2], Chiara Casirati[1], Paola Adamo[1], Davide De Leo[1], Giorgia Marino[1], Guido Grappiolo[3], Roberto Gatti[1,2]

[1]Physiotherapy Unit, IRCCS Humanitas Research Hospital, Rozzano, Milan - Italy
[2]Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, Milan - Italy
[3]Hip and Knee Orthopaedic Surgery Department, IRCCS Humanitas Research Hospital, Rozzano, Milan - Italy

ABSTRACT

Introduction: The 10-Meter Walk Test (10MWT) and Timed Up and Go (TUG) are valid tools for gait performance and mobility assessment after total hip arthroplasty (THA). The study aimed to assess test-retest reliability of 10MWT and TUG in indoor and outdoor environments in patients in acute phase after THA and compare their indoor vs. outdoor performance during these tests.
Methods: Thirty-five inpatients performed 10MWT and TUG in indoor and outdoor settings on the second postoperative day. An additional evaluation session was performed after 1 hour under the supervision of the same operator. Test-retest reliability was assessed using Intraclass Correlation Coefficient (ICC: 2.1) and Minimal Detectable Change (MDC_{95}), while paired t-tests were used to compare indoor vs. outdoor performance.
Results: Indoor (ICC: 0.94, MDC_{95}: 0.13 m/s) and outdoor (ICC: 0.91, MDC_{95}: 0.16 m/s) 10MWT at maximum speed and indoor (ICC: 0.92, MDC_{95}: 2.5 s) and outdoor (ICC: 0.93, MDC_{95}: 2.4 s) TUG revealed excellent reliability. Indoor (ICC: 0.86, MDC_{95}: 0.16 m/s) and outdoor (ICC: 0.89, MDC_{95}: 0.16 m/s) 10MWT at spontaneous speed revealed good reliability. Spontaneous (mean difference [MD]: 0.05 m/s, 95% confidence interval [CI_{95}]: 0.03, 0.07, $p < 0.001$) and maximum (MD: 0.02 m/s, CI_{95}: 0.01, 0.04, $p < 0.001$) 10MWT revealed higher gait speed when performed outdoors compared to indoors.
Conclusions: Indoor and outdoor 10MWT and TUG are reliable tests in acute phase after THA. Higher gait speed during outdoor 10MWT may depend on test score variability, due to MDs being lower than MDC_{95}.

Keywords: Gait performance, Hip arthroplasty, Indoor setting, Mobility, Outdoor setting

What's already known about this topic?

- *The 10MWT and TUG are valid measurements tools, which are widely used for assessing gait performance and mobility of patients in acute phase after THA.*

What does the study add?

- *The 10MWT and TUG in indoor and outdoor settings are reliable tests in acute phase after THA. Higher gait speed was found during 10MWT performed outdoors compared to indoors, but changes are lower than MDC_{95}.*

Introduction

Total hip arthroplasty (THA) represents a successful surgical procedure to reduce pain and improve function and quality of life in patients with end-stage hip osteoarthritis (1). The advancements in surgical techniques (e.g., minimally invasive surgical approaches) and improvements in perioperative care (e.g., prehabilitation and early mobilization protocols) have allowed for length of stay reduction, which decreased from some weeks to a few days in patients undergoing THA (2-4). When considering patients in acute phase after THA, the achievement of clinical stability and functional outcomes represents a milestone to establish the readiness for hospital discharge (3). In fact, functional independence during the execution of basic daily activities and satisfactory levels of walking performance and mobility are required to ensure a safe discharge in these patients (3,5,6).

The 10-Meter Walk Test (10MWT) and Timed Up and Go (TUG) are valid measurement tools for walking performance and mobility assessment in patients after lower limb

Received: August 20, 2024
Accepted: October 14, 2024
Published online: November 4, 2024

Corresponding author:
Federico Temporiti
email: federico.temporiti@humanitas.it

orthopedic surgery (7,8). The 10MWT consists of asking patients to walk along a 10-m walkway at self-paced and maximum speed to detect spontaneous and maximum walking speed by timing the performance (7). During TUG, patients are asked to rise from an armchair, walk at a comfortable pace for 3 m, turn and walk back to the chair and sit down again. The performance is timed to detect the test duration, which is an index of functional mobility (8). However, when considering 10MWT and TUG in patients with THA, the reliability of these tests has only been described in patients with end-stage hip osteoarthritis and in the subacute phase after THA and in a sample of patients suffering from heterogeneous musculoskeletal conditions affecting the lower limb (8-12). In addition, the assessment of walking performance and mobility in patients discharged after THA is usually carried out in a hospital setting (e.g., rehabilitative gyms or ward hallways), which represent an indoor, familiar and supervised setting in which patients have performed a rehabilitative program during postoperative days. However, hospital discharge often induces patients to perform outdoor activities in unfamiliar environments, where the ability to adapt to unexpected perturbations during gait and other functional tasks is required (13). In this context, studies have described motor performance changes between unfamiliar outdoor environment and familiar indoor setting in older adults and patients with gait disorders (14,15). Therefore, it is reasonable to speculate that the execution of motor performance tests such as 10MWT and TUG in indoor and outdoor settings may be more representative of the locomotor performance and mobility in patients discharged in acute phase after THA.

To date, no studies have investigated the reliability of indoor and outdoor 10MWT and TUG in patients discharged in acute phase after THA. Moreover, walking performance and mobility in an indoor vs. outdoor environment have never been compared in these patients. The first study's aim was to assess test-retest reliability of 10MWT and TUG in indoor and outdoor environments in patients in acute phase after THA. The second study's aim was to compare indoor vs. outdoor performance during these tests in patients in acute phase after THA. We hypothesized that indoor and outdoor 10MWT and TUG would result in good to excellent test-retest reliability in patients in the acute phase after THA. Moreover, we expected better 10MWT and TUG scores when these tests were performed indoors compared to outdoors.

Methods

Participants

Thirty-five inpatients with unilateral THA were enrolled on the second postoperative day. Inclusion criteria were age between 40 and 80 years, primary unilateral THA for osteoarthritis (Kellgren-Lawrence grade of at least 3) and readiness for discharge (16). Patients with a Kellgren-Lawrence grade of at least 3 were included in order to select participants who underwent THA for advanced stage of hip osteoarthritis including narrowing of joint space and bone sclerosis. Discharge criteria included the ability to stand up from a standard chair, walk at least 100 m, and perform stairs with crutches. In addition, dry wound, hemoglobin levels higher than 8 g/dL, perceived pain at rest and during walking lower than 4 points on a Visual Analogue Scale (VAS 0-10), and absence of dizziness or nausea were required. Exclusion criteria were revision surgery, perioperative complications, diagnosis of cognitive impairment or psychiatric disorders, and concurrent neurological or musculoskeletal conditions able to influence postoperative functional recovery. All participants were operated under spinal anesthesia by three orthopedic surgeons of the same unit adopting a standardized posterolateral approach with femur-first technique and uncemented implant fixation (17). All patients followed a postoperative in-hospital rehabilitation program under the supervision of a physiotherapist. The rehabilitative protocol consisted of two 30-minute daily sessions including manual therapy techniques to improve hip range of motion, resistance training to enhance strength of lower limb muscles, and task-oriented exercises performed in standing posture for increasing postural stability. In addition, patients were trained on the execution of functional daily tasks, such as getting out of bed, sitting on a chair, walking as tolerated, and stairs performance with crutches (18). The study was carried out at the Physiotherapy Unit of the Humanitas Clinical and Research Center of Milan, Italy. All participants signed a written informed consent as per the Declaration of Helsinki and the ethical committee of our institute approved the study protocol (number: CLF23/04).

Assessment

An experienced physiotherapist enrolled participants on the afternoon of the second postoperative day. Immediately after the enrollment, participants performed the 10MWT at self-paced and maximum speed and TUG in a familiar indoor (rehabilitation gym) and unfamiliar outdoor (straight sidewalk composed of flat tiles in the hospital garden) settings in a randomized order. The indoor and outdoor sessions were interspaced by a 5-minute resting period and the modified Borg scale was used to monitor the participants' fatigue (19). Specifically, participants had to report a fatigue level equal to 0 before each session. Moreover, a wheelchair was used to transfer participants from indoor to outdoor environments in order to avoid fatigue onset. During 10MWT, participants were asked to walk with crutches for 14 m at self-paced speed and as quickly as possible. The initial and final 2 m were used for acceleration and deceleration and the performance was timed using a stopwatch to detect spontaneous and maximum gait speed. Two trials were performed for self-paced and maximum speed conditions, and the mean score was used for data analysis (10). After 10MWT, participants were asked to perform the TUG. In particular, they were asked to rise from an armchair, walk at a comfortable speed for 3 m without crutches, turn and walk back to the chair in order to sit down again. After a familiarization trial, two trials were performed. The performance was timed with a stopwatch to detect test duration, and the best trial was used for data analysis (20). The test execution complied with the most recent guidelines on the use of restrictions and assistive devices in patients in acute phase after THA, which recommended the lack of hip movement restrictions in these patients (21). After 1 hour, the indoor and outdoor sessions were repeated

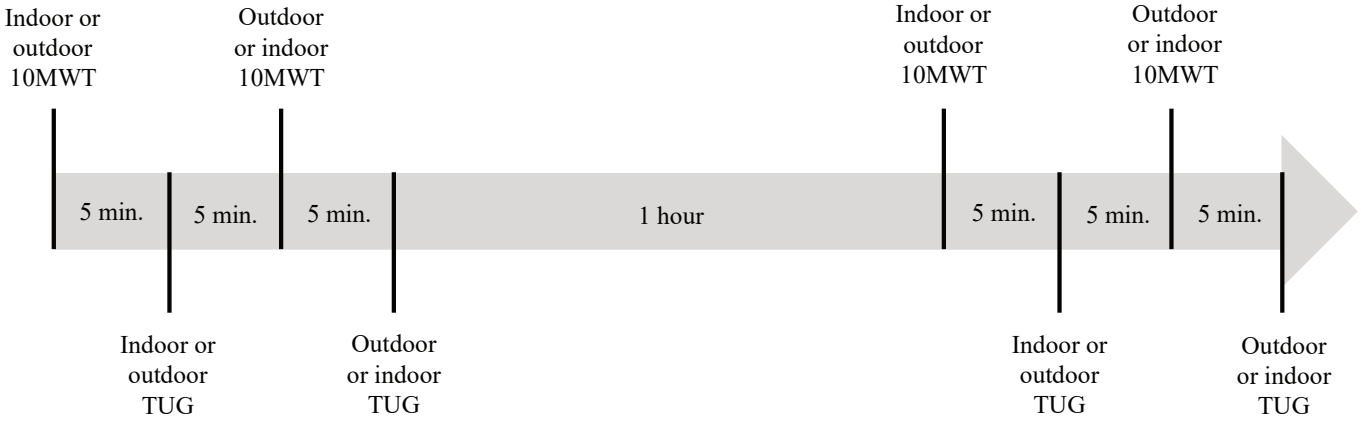


FIGURE 1 - Representation of study design. 10MWT = 10-Meter Walk Test; TUG = Timed Up and Go.

in the same sequence adopted during the first session and under the supervision of the same operator to evaluate the test-retest reliability of 10MWT and TUG in indoor and outdoor settings (Fig. 1).

Perceived pain was assessed by an experienced physiotherapist at the end of each indoor and outdoor session using the Numeric Pain Rating Scale (NPRS), which consists of an 11-point numerical scale with a score ranging from 0 (no pain) to 10 points (maximum pain). Finally, hip function and impact of hip-related signs and symptoms on daily activities were assessed to further characterize study participants. Hip function was assessed through the Harris Hip Score (HHS), which consists of a 10-item questionnaire ranging from 0 (high dysfunction) to 100 (no dysfunction) and exploring pain, hip function, daily activities, hip deformities, and range of motion. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) consists of a 24-item self-administered questionnaire used to assess the impact of hip pain, stiffness, and function on the performance of daily activities.

Data analysis

Sample size was calculated a priori using the methodology proposed by Walter and coworkers (22). Considering two repetitions per subject, alpha error of 0.05, power (1-beta) of 80%, and a minimum acceptable Intraclass Correlation Coefficient (ICC) score of 0.5, 35 participants were required to determine an ICC score of 0.8.

All measurements were checked for normality using the Shapiro-Wilk test, and being normally distributed, were expressed as mean and standard deviation. The ICC 2.1 with a 95% confidence interval (CI_{95}) was adopted to assess the relative reliability and interpreted as excellent (0.9 or greater), good (between 0.75 and 0.9), moderate (between 0.5 and 0.75), and poor (0.5 or lower) (23). In addition, the Standard Error of Measurement (SEM) was adopted to investigate the absolute reliability. It was computed as SEM = SD √1-ICC, where SD represents the standard deviation of the mean of all trials, and expressed in the same measurement unit of the test score (m/s for 10MWT and seconds for TUG) and as a percentage of the mean. Moreover, the minimal detectable change with 95% confidence (MDC_{95}) computed as MDC = 1.96 SEM √2 was adopted to obtain a measure of the change in terms of 10MWT and TUG scores that may be considered as a true change beyond the measurement errors.

Finally, paired t-test was used to compare 10MWT and TUG scores in an indoor vs. outdoor setting. Effect size was also quantified using Cohen's d with 95% CI and interpreted as small (0.2), medium (0.5), or large (0.8 or greater) (24). Data were analyzed using SPSS 28.0 for Windows and the level of significance was set at alpha = 0.05.

Results

All participants completed the evaluation sessions correctly; no dropouts occurred and none of the participants required a longer resting period between indoor and outdoor sessions. Participants had a mean age of 58.5 years (SD: 6.9 years, range: 46-80 years), a mean height of 1.71 m (SD: 0.01 m, range 1.56-1.83 m), a mean weight of 80.6 kg (SD: 16.9 kg, range: 48-102 kg), and a mean body mass index of 26.7 kg/m² (SD: 4.1 kg/m², range: 18.8-33.2 kg/m²). Twenty-two men and 13 women who underwent 22 right-sided and 17 left-sided THA were included. Twenty-four patients had preoperative Kellgren-Lawrence grade 3, while nine patients had preoperative Kellgren-Lawrence grade 4. Finally, participants reported a mean WOMAC score of 51.1 points (SD: 21.1 points) and a mean HHS score of 58.5 points (SD: 10.7 points).

Reliability

Excellent test-retest reliability was found for 10MWT performed indoors (ICC: 0.94, $p < 0.001$, and MDC_{95}: 0.13 m/s) and outdoors (ICC 0.91, $p < 0.001$, and MDC_{95}: 0.16 m/s) at maximum speed, while good test-retest reliability was found for 10MWT performed indoors (ICC: 0.86, $p < 0.001$, and MDC_{95}: 0.16 m/s) and outdoors (ICC: 0.89, $p < 0.001$, and MDC_{95}: 0.16 m/s) at spontaneous speed. Finally, excellent test-retest reliability was found for TUG performed indoors (ICC: 0.92, $p < 0.001$, and MDC_{95}: 2.5 s) and outdoors (ICC: 0.93, $p < 0.001$, and MDC_{95}: 2.4 s) (Tab. 1).

TABLE 1 - Test-retest reliability of 10MWT at self-paced and maximum speed and TUG performed in indoor and outdoor settings

Variables	Test	Retest	ICC [CI 95%]	p-Value	SEM	SEM%
Indoor setting						
10MWT – self-paced speed (m/s)	0.83 ± 0.17	0.92 ± 0.15	0.86 [0.13; 0.96]	<0.001	0.06	6.80
10MWT – maximum speed (m/s)	1.10 ± 0.20	1.16 ± 0.20	0.94 [0.64; 0.98]	<0.001	0.05	4.34
TUG (s)	14.6 ± 3.1	13.5 ± 3.1	0.92 [0.62; 0.97]	<0.001	0.89	6.33
Outdoor setting						
10MWT – self-paced speed (m/s)	0.88 ± 0.17	0.97 ± 0.17	0.89 [0.17; 0.97]	<0.001	0.06	6.06
10MWT – maximum speed (m/s)	1.12 ± 0.20	1.18 ± 0.19	0.91 [0.72; 0.96]	<0.001	0.06	5.22
TUG (s)	14.7 ± 3.4	13.9 ± 3.1	0.93 [0.84; 0.97]	<0.001	0.87	6.05

10MWT = 10-Meter Walk Test; CI = confidence interval; ICC = intraclass correlation coefficient; SEM = standard error of the measurement; TUG = Timed Up and Go.

TABLE 2 - Indoor versus outdoor performance during 10MWT at self-paced and maximum speed and TUG

Variables	Indoors	Outdoors	MD [CI 95%]	p-Value	Cohen's *d*
10MWT – self-paced speed (m/s)	0.88 ± 0.16	0.93 ± 0.17	−0.05 [−0.07; −0.03]	<0.001	0.51
10MWT – maximum speed (m/s)	1.13 ± 0.20	1.15 ± 0.20	−0.02 [−0.05; −0.01]	0.042	0.25
TUG (s)	14.0 ± 3.1	14.3 ± 3.3	−0.3 [−0.59; 0.08]	0.138	–

10MWT = 10-Meter Walk Test; CI = confidence interval; MD = mean difference; SEM = standard error of the measurement; TUG = Timed Up and Go.

Indoor vs. outdoor performance

Participants showed higher gait speed during 10MWT at spontaneous (MD: 0.05 m/s, IC_{95}: 0.03, 0.07 m/s, $p < 0.001$) and maximum speed (MD: 0.02 m/s, IC_{95}: 0.01, 0.04, $p < 0.001$) performed outdoor compared to indoor setting. The effect size was medium for 10MWT at spontaneous speed ($d = 0.51$, IC_{95}: 0.76, 0.26) and small for 10MWT at maximum speed ($d = 0.25$, IC_{95}: 0.01, 0.49). No significant differences were found for TUG performed in indoor or outdoor settings. Finally, no significant differences were found in terms of VAS at the end of the indoor and outdoor sessions (indoor assessment: 1.9 ± 1.2 points, outdoor assessment: 2.3 ± 1.0 points, $p = 0.450$) (Tab. 2).

Discussion

The main finding was that excellent to good reliability was found for 10MWT and TUG performed in indoor and outdoor settings in patients in acute phase after THA. Moreover, higher gait speed was found during 10MWT at self-paced and maximum speed, when this test was performed outdoors compared to indoors.

Literature data have described gait speed during 10MWT as an indicator of functional status in patients after lower limb orthopedic surgery including THA (10,25). When considering available literature, a single study of Unver and coworkers investigated the test-retest reliability of the 10MWT in patients with THA (10). The current study findings agree with the results of Unver and coworkers, which demonstrated excellent test-retest reliability (ICC: 0.96) for 10MWT at maximum speed in patients in the first week after THA (10). However, mean gait speed of patients included in the study of Unver and coworkers was substantially lower than mean gait speed observed in the current study participants (0.22 vs. 1.13 m/s), suggesting differences in terms of patients' characteristics and functional abilities (10). Despite the lack of information, it is reasonable to speculate that hospital discharge criteria were not satisfied in patients enrolled in the study of Unver and coworkers compared to our study participants. In addition, our findings demonstrated good test-retest reliability for 10MWT performed at self-paced speed.

The current study also revealed excellent test-retest reliability for TUG performed indoors and outdoors in patients discharged in acute phase after THA. Our findings revealed ICC values higher than 0.90 both indoors and outdoors and suggested that only changes greater than 2.5 s (MDC_{95}) in terms of TUG score may be interpreted as true changes. When considering existing literature data, studies have investigated the test-retest reliability of TUG performed indoors, showing ICC values ranging from 0.83 to 0.98 in patients after THA (11,20,26). In particular, the clinical features of our study population are similar to the characteristics of patients included in the study of Kirschner and coworkers, which found an ICC value of 0.98 for TUG in patients with THA (26). However, participants included in the aforementioned study had greater body mass index than our study participants and revealed a mean TUG score of approximately 20 s (26). Moreover, Yuksel and coworkers described TUG test-retest reliability of 0.96 and 0.59 in terms of ICC and SEM values in patients at 6 months after THA (20). However, it is reasonable to speculate that patients enrolled by Yuksel and coworkers were extracted from a different population than participants of our study. In fact, Lieberman and co-workers reported that

patients usually achieve a complete restoration of functional abilities at 6 months after THA (27). Furthermore, Doll and coworkers reported an ICC value of 0.83 for TUG in patients at 2 weeks after THA. Lower test-retest reliability in the aforementioned study may depend on the use of different walking aids among participants during TUG, such as one or two crutches or a walker (11). In fact, the use of walking aids might have increased the intrasubject variability between test and retest trials, affecting TUG reliability.

This was the first study that compared indoor versus outdoor performance in patients in acute phase after THA. Conversely to our hypothesis, patients with THA revealed higher spontaneous and maximum gait speed outdoors compared to indoors. This finding was consistent with the results of Schmitt and coworkers, which have described higher gait speed in young and elderly subjects in an outdoor compared to indoor setting, as a result of increased stimuli and multisensory feedback provided by the outdoor environment (15). However, it is worth also highlighting that the magnitude of changes in our study were 0.05 and 0.02 m/s for spontaneous and maximum speed, respectively. These values were lower than the MDC_{95} values described for indoor and outdoor 10MWT in patients in acute phase after THA, suggesting that observed changes may depend on the variability of the test score described in the study population (28). Conversely to Schmitt and coworkers, previous studies have reported no differences between indoor and outdoor performance in older adults and patients with gait disorders for neurological conditions, in agreement with the lack of environment influence on gait abilities and mobility in patients in acute phase after THA (13,29). The results of our study may depend on the fact that the central nervous system tends to redistribute the resources to adequately accomplish the task, when the performance is not maximal (30). In fact, self-paced 10MWT and TUG require submaximal levels of performance and 10MWT at maximum speed was performed using crutches, which might have contributed to limit the task maximality. The lack of task maximality might have hindered potential motor performance changes between indoor and outdoor settings. In addition, the adoption of compensatory mechanisms (e.g., higher reliance on visual inputs) might have played a role in ensuring similar levels of performance between indoor and outdoor settings (13).

The assessment of test-retest reliability of 10MWT and TUG and comparison between indoor and outdoor performance during these tests were carried out in patients in the acute phase after THA. Consistently with time following surgery, patients revealed poor hip function and the presence of hip-related signs and symptoms, as demonstrated by HHS lower than 70 points and WOMAC score of 51.1 points (31,32). In fact, these scores are similar to those reported by previous studies in patients in acute phase after THA (3,31,32).

Some limitations need to be underlined in the current study. First, our findings were extracted from patients in the acute phase after THA showing specific features. In fact, patients had no weight-bearing restrictions on the affected limb and achieved readiness for discharge within the second postoperative day. These factors limit the external validity of our findings and caution is needed to generalize the current results to a broader population undergoing THA. Second, outdoor assessment was carried out using a sidewalk in the hospital garden without ground irregularities or distracting elements, which may be only partially representative of the outdoor setting in which patients usually perform the activities of daily living. Third, mean age of participants was 58.5 years and the age range adopted in the inclusion criteria was slightly different from the age range of the majority of subjects undergoing primary THA in our country (16). Therefore, caution is needed to generalize the current study findings to a broader population of patients with THA. Finally, no instrumental assessment was carried out. In fact, the investigation of the reliability of spatial temporal parameter during 10MWT and TUG might have revealed potential differences between indoor and outdoor performance in acute phase after THA.

Conclusions

Indoor and outdoor 10MWT and TUG were reliable tests to assess walking performance and mobility in patients in acute phase after THA. Moreover, higher gait speed was found during 10MWT at self-paced and maximum speed outdoors compared to indoors, but the relevance of these changes remains questionable.

Acknowledgments

We thank prof. Patricia Taylor for the help on English language revision.

Disclosures

Conflict of interest: The authors have no conflicts of interests.

Funding: The authors received no funding.

Author contributions: FT: conceptualization, methodology, data collection, formal analysis, project administration, and writing. CC: data collection, methodology, visualization, and formal analysis. PA: methodology and data curation. DDL: data curation and visualization. GM: data curation and visualization. GG: conceptualization and supervision. RG: conceptualization, supervision, project administration, and writing.

Data availability statement: 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. Sato EH, Stevenson KL, Blackburn BE, et al. Impact of demographic variables on recovery after total hip arthroplasty. J Arthroplasty. 2024;39(3):721-726. CrossRef PubMed
2. Ohta Y, Sugama R, Minoda Y, et al. Is the anterolateral or posterolateral approach more effective for early postoperative recovery after minimally invasive total hip arthroplasty? J Clin Med. 2022;12(1):139. CrossRef PubMed
3. Temporiti F, Draghici I, Fusi S, et al. Does walking the day of total hip arthroplasty speed up functional independence? A non-randomized controlled study. Arch Physiother. 2020;10(1):8. CrossRef PubMed
4. Papalia R, Zampogna B, Torre G, et al. Preoperative and perioperative predictors of length of hospital stay after primary total hip arthroplasty – our experience on 743 cases. J Clin Med. 2021; 10(21):5053. CrossRef PubMed

5. Yang G, Chen W, Chen W, Tang X, Huang Y, Zhang L. Feasibility and safety of 2-day discharge after fast-track total hip arthroplasty: a Chinese experience. J Arthroplasty. 2016;31(8): 1686-1692.e1. CrossRef PubMed
6. Fraser JF, Danoff JR, Manrique J, Reynolds MJ, Hozack WJ. Identifying reasons for failed same-day discharge following primary total hip arthroplasty. J Arthroplasty. 2018;33(12): 3624-3628. CrossRef PubMed
7. Latham NK, Mehta V, Nguyen AM, et al. Performance-based or self-report measures of physical function: which should be used in clinical trials of hip fracture patients? Arch Phys Med Rehabil. 2008;89(11):2146-2155. CrossRef PubMed
8. Yeung TS, Wessel J, Stratford PW, MacDermid JC. The timed up and go test for use on an inpatient orthopaedic rehabilitation ward. J Orthop Sports Phys Ther. 2008;38(7):410-417. CrossRef PubMed
9. Kennedy DM, Stratford PW, Wessel J, Gollish JD, Penney D. Assessing stability and change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty. BMC Musculoskelet Disord 2005;6:3. CrossRef PubMed
10. Unver B, Baris RH, Yuksel E, Cekmece S, Kalkan S, Karatosun V. Reliability of 4-meter and 10-meter walk tests after lower extremity surgery. Disabil Rehabil. 2017;39(25):2572-2576. CrossRef PubMed
11. Doll H, Gentile B, Bush EN, Ballinger R. Evaluation of the measurement properties of four performance outcome measures in patients with elective hip replacements, elective knee replacements, or hip fractures. Value Health. 2018; 21(9):1104-1114. CrossRef PubMed
12. Halket A, Stratford PW, Kennedy DM, Woodhouse LJ, Spadoni G. Measurement properties of performance-specific pain ratings of patients awaiting total joint arthroplasty as a consequence of osteoarthritis. Physiother Can. 2008;60(3):255-263. CrossRef PubMed
13. Zukowski LA, Tennant JE, Iyigun G, Giuliani CA, Plummer P. Dual-tasking impacts gait, cognitive performance, and gaze behavior during walking in a real-world environment in older adult fallers and non-fallers. Exp Gerontol. 2021;150:111342. CrossRef PubMed
14. Olmos LE, Freixes O, Gatti MA, et al. Comparison of gait performance on different environmental settings for patients with chronic spinal cord injury. Spinal Cord. 2008;46(5):331-334. CrossRef PubMed
15. Schmitt AC, Baudendistel ST, Lipat AL, White TA, Raffegeau TE, Hass CJ. Walking indoors, outdoors, and on a treadmill: gait differences in healthy young and older adults. Gait Posture. 2021;90:468-474. CrossRef PubMed
16. Ciccarelli P, Urakcheeva I, Biondi A, Torre M. Italian Arthroplasty Registry. Annual Report 2021 – Addendum, 1st ed. Roma: Il pensiero scientifico editore, 2022; 29. Online (Accessed August 2024)
17. Loppini M, Longo UG, Caldarella E, Rocca AD, Denaro V, Grappiolo G. Femur first surgical technique: a smart non-computer-based procedure to achieve the combined anteversion in primary total hip arthroplasty. BMC Musculoskelet Disord. 2017;18(1):331. CrossRef PubMed
18. Temporiti F, De Leo D, Adamo P, et al. Impaired modulation of motor and functional performance in patients after total knee arthroplasty: a prospective observational study. BioMed Res Int. 2022;2022:4546836. CrossRef PubMed
19. Grant S, Aitchison T, Henderson E, et al. A comparison of the reproducibility and the sensitivity to change of visual analogue scales, Borg scales, and Likert scales in normal subjects during submaximal exercise. Chest. 1999;116(5):1208-1217. CrossRef PubMed
20. Yuksel E, Unver B, Kalkan S, Karatosun V. 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
21. Ruspi A, De Leo D, Scandelli F, et al. Recommendations on the use of restrictions and assistive devices after total hip arthroplasty: an adolopment of guidelines. Disabil Rehabil. In press.
22. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med. 1998;17(1):101-110. CrossRef PubMed
23. 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
24. Durlak JA. How to select, calculate, and interpret effect sizes. J Pediatr Psychol. 2009;34(9):917-928. CrossRef PubMed
25. van den Akker-Scheek I, Stevens M, Bulstra SK, Groothoff JW, van Horn JR, Zijlstra W. Recovery of gait after short-stay total hip arthroplasty. Arch Phys Med Rehabil. 2007;88(3): 361-367. CrossRef PubMed
26. Kirschner J, Michel S, Becker R, et al. Determination of relationships between symmetry-based, performance-based, and functional outcome measures in patients undergoing total hip arthroplasty. J Pers Med 2023;13(7):1046. CrossRef PubMed
27. Lieberman JR, Dorey F, Shekelle P, et al. Differences between patients' and physicians' evaluations of outcome after total hip arthroplasty. J Bone Joint Surg Am. 1996;78(6):835-838. CrossRef PubMed
28. Dijkers MP, Kropp GC, Esper RM, Yavuzer G, Cullen N, Bakdalieh Y. Reporting on reliability and validity of outcome measures in medical rehabilitation research. Disabil Rehabil. 2002;24(16):819-827. CrossRef PubMed
29. Donovan K, Lord SE, McNaughton HK, Weatherall M. Mobility beyond the clinic: the effect of environment on gait and its measurement in community-ambulant stroke survivors. Clin Rehabil. 2008;22(6):556-563. CrossRef PubMed
30. Clark DJ, Rose DK, Ring SA, Porges EC. Utilization of central nervous system resources for preparation and performance of complex walking tasks in older adults. Front Aging Neurosci. 2014;6:217. CrossRef PubMed
31. Singh JA, Schleck C, Harmsen S, Lewallen D. Clinically important improvement thresholds for Harris Hip Score and its ability to predict revision risk after primary total hip arthroplasty. BMC Musculoskelet Disord. 2016;17(1):256. CrossRef PubMed
32. Gooch K, Marshall DA, Faris PD, et al. Comparative effectiveness of alternative clinical pathways for primary hip and knee joint replacement patients: a pragmatic randomized, controlled trial. Osteoarthritis Cartilage. 2012;20(10): 1086-1094. CrossRef PubMed

Arch Physioter 2024; 14: 96-104
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3092
ORIGINAL RESEARCH ARTICLE

The Italian version of the Postural Assessment Scale for Stroke Patients (PASS): transcultural translation and validation

Elena Lora[1], Noemi Gaudenzi[1], Ada Buriani[1], Antonietta Bacciocchi[1], Lea Godino[2], Mattia Ricco[3], Domenica Gazineo[4]

[1] Department of Continuity of Care and Disability, Physical Medicine and Rehabilitation, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Policlinico di Sant'Orsola, Bologna - Italy
[2] Medical Genetics Unit, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Policlinico di Sant'Orsola, Bologna - Italy
[3] Department of Medical and Surgical Sciences (DIMEC), University of Bologna, Bologna - Italy
[4] Governo Clinico e Qualità, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Policlinico di Sant'Orsola, Bologna - Italy

ABSTRACT

Introduction: The Postural Assessment Scale for Stroke Patients (PASS) is commonly used by health professionals in Italy in several different translations. This study aimed to provide a validated version in Italian. The main focus is on the evaluator, to guarantee a uniform application and interpretation of the statements and scoring for each item in the Italian context.
Methods: A standardized protocol was used for the translation and cross-cultural adaptation. A pilot study conducted using the first draft of the scale led to a revised version, PASS-IT. A principal component analysis (PCA) was performed. The correlation with the Trunk Control Test (TCT) was examined for concurrent validity. In addition, the relationship with the Barthel Index (BI) and the Functional Ambulation Categories (FAC) was tested. Patients with recent stroke were tested for intra-rater (N = 49) and inter-rater agreement (N = 30). Cronbach's alpha, item-to-total correlation, corrected inter-item correlation, the intraclass correlation coefficient (ICC), and measurement error were used to evaluate internal consistency and intra-/inter-rater reliability.
Results: The PCA showed a two-dimensional structure, with high reliability in both subsections ("non-weight-bearing" $\alpha = 0.865$; "weight-bearing" $\alpha = 0.949$). A strong correlation ($\rho > 0.80$) was found with the TCT, the BI, and the FAC. The PASS-IT showed high internal consistency, intra-rater (ICC = 0.942) and inter-rater reliability (ICC = 0.940).
Conclusions: The PASS-IT is a recommended scale, suitable for clinical practice and research in the acute and subacute stage. The introduction of operating instructions resulted in the uniform application. A different order of the items allows faster administration, reducing changes of posture.

Keywords: Cross-cultural adaptation, Outcome assessment, Postural balance, Postural control, Reproducibility of results, Stroke.

What's already known about this topic?

- *The PASS is among the most recommended scales for the assessment of postural control in patients with stroke. Although it is extensively used in Italy in several different translations, a validated version is not available.*

What does the study add?

- *This study aimed to provide an Italian version of the PASS, going through a cross-cultural validation process, adding operating instructions to promote a uniform application and interpretation of the scale among Italian health professionals.*

Introduction

Stroke is currently one of the most common causes of disability and dependence among the older adult population in developed countries (1-3). In the European Union, there are over 9.5 million stroke survivors and a 27% increase is expected in the next three decades, due to population aging and improved survival rates (4).

In addition to compromising the limb mobility in the affected side, a stroke causes an alteration of postural control and balance (5, 6). The ability to maintain balance in the

Received: April 14, 2024
Accepted: November 11, 2024
Published online: December 3, 2024

This article includes supplementary material

Corresponding author:
Elena Lora
email: elena.lora@aosp.bo.it

sitting position, standing, and in postural variations is essential for the recovery of independence, and the close correlation between postural control of the trunk in the acute phase and future functional ability is recognized (6-10).

Being able to predict the degree of recovery at an early stage after the stroke onset allows the medical and rehabilitation team to optimize time, tools, and resources in planning goals and treatment (7, 8, 11). Hence the need to identify valid and adequate assessment tools. The Trunk Control Test (TCT) (12, 13) is probably the most used and feasible in the acute stage. Of the four scales for trunk control compared in 2019 by Fil Balkan et al (14), the TCT was found to be the most time-efficient and with a better predictive value, but showed a floor effect. In clinical practice a ceiling effect is also frequently observed within the first weeks (15). Other scales proposed in the literature, for example, the Fugl-Meyer (16, 17), require a long time and demanding training for their use (18). Others, such as the Trunk Impairment Scale (TIS) (5), evaluate trunk control only in a sitting position.

The Postural Assessment Scale for Stroke Patients (PASS), conceived by Benaim et al and published in 1999 (19), evaluates the ability to maintain stable postures and balance during positional changes. It can be applied to all patients with stroke, even those with minimal postural control, in the first 3 months. The validation studies confirmed the structural validity of the PASS, excellent inter- and intra-operator reliability, high internal consistency (19), and the absence of floor/ceiling (F/C) effects when applied to the target population in the first weeks post-stroke (19, 20). A ceiling effect has been found for patients with high functional ability (21, 22).

Recent studies showed that, compared to the Berg Balance Scale, the PASS is better able to detect balance improvements in patients with severe balance deficits (23), and that it is a valid instrument to assess balance at an early stage (20) but also in the subacute and chronic phase (24, 25). It is an excellent early predictor for autonomy in both basic activities of daily living (ADLs) and instrumental ADLs (IADLs) (7, 19, 26), consistent with the results detected at 3 months with the Functional Independence Measure (FIM) (19) or at 6 months with the Barthel Index (BI) and the Frenchay Activities Index (19, 27). The predictive power is greater than the Fugl-Meyer's (7), even in foreseeing the patient's walking ability after discharge (28). Because of its properties and short evaluation time, the PASS is used worldwide and validated versions have been produced in Portuguese (29), Swedish (30), Spanish (31), Norwegian (32), and Turkish (33).

Moreover, the PASS is among the 14 highly recommended outcome measures selected in 2013 by the American Physical Therapy Association for individuals with stroke (34).

It is therefore advisable to use this tool in the clinical context. Although the PASS is commonly used in Italy, there is no cross-culturally validated version in this language. This study aimed to provide an Italian version of the PASS, going through a cross-cultural validation process, assessing its validity and intra-/inter-rater reliability.

Methods

The study was conducted in two phases (Fig. 1): (1) translation and cross-cultural adaptation of the first version of the Italian PASS (PASS-IT), followed by a pilot study to resolve possible critical issues; (2) a psychometric evaluation, assessing the validity and reliability of the PASS-IT.

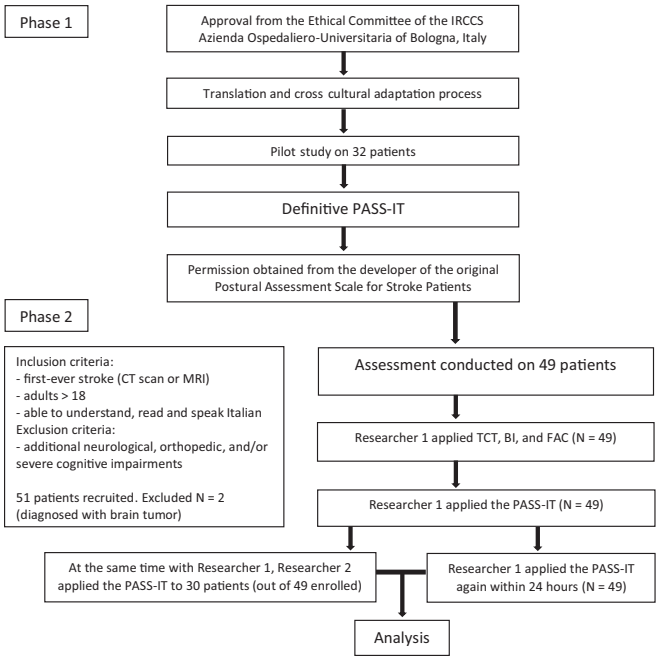


FIGURE 1 - Flowchart of the study.

Translation and cross-cultural adaptation

The translation and cross-cultural adaptation were conducted following the six-step method proposed by Beaton et al (35). Initially, two native Italian physiotherapists produced their own translations separately. Comparing the two translations, a synthesis was produced in agreement. Two back-translations were performed by native English translators, without clinical experience and not familiar with the original scale. The comparison between the original version and the back-translations revealed no substantial differences. Subsequently an expert committee (composed of a methodologist, all the translators involved in the previous phases, a physiotherapist not involved in the translation, a physiatrist, a geriatrician, a stroke unit doctor, a rehabilitation coordinator, and a nursing coordinator, all knowledgeable in English) analyzed the semantic, idiomatic, experiential, and conceptual areas of the scale, choosing the most suitable expression for each item. The changes made at this stage led to the first draft of the PASS-IT.

For a preliminary evaluation of the tool, this version was administered to a sample of 32 patients with recent stroke, admitted to the Stroke Unit and to the Unit of Physical Medicine and Rehabilitation of the IRCCS University Hospital of Bologna. The participants were informed about the study and gave written consent.

Twenty-one physiotherapists were invited to use the PASS-IT for 3 months. A questionnaire was then handed out, investigating clearness of the items, problems encountered,

perception of confusing or missing features, and time needed to administer the scale. The results were reported in a focus group discussion, involving some of the authors, a methodologist doctor, and the physiotherapists, seeking terms which could guarantee the best univocal interpretation. Critical issues were highlighted and resolved by consensus.

Psychometric evaluation

Participants

The sample size of the psychometric evaluation study was determined combining the results of the pilot study (average total score 27.78 ± 8.19 in patients on day 14 ± 3 from stroke onset) and those reported by Koçak et al (average score 17.70 ± 10.08) (33). Since the current study involves patients in a more acute stage, an expected average score of 24 was estimated. Aiming for a statistical power of 80% ($\beta = 0.20$) and a significance level of 0.05 ($\alpha = 0.05$), it was determined that 20 subjects would be necessary to ensure reliable and valid results. A larger sample size was enrolled, in order to offset potential dropouts, provide increased statistical power, and improve the generalizability of the results to a broader population, enhancing the external validity.

The study was conducted on a group of 49 consecutive patients (30 for the inter-rater reliability) admitted to the Stroke Unit of the IRCCS University Hospital of Bologna between February and July 2022. All patients were in the acute or early subacute phase after a stroke onset (mean: 5 ± 2.68 days; range: 1 to 12 days after the event).

The study included patients with a first-ever stroke, confirmed by a cerebral computed tomography (CT) scan or magnetic resonance imaging (MRI); adults ≥ 18; able to understand, read and speak the Italian language.

The study excluded patients with additional neurological, orthopedic, and/or severe cognitive impairments, which could compromise postural control or cooperation.

Data collection

Data were collected by two physiotherapists, both with experience with patients affected by neurological diseases. Researcher 1 was familiar with the PASS-IT, while Researcher 2, who had never used it before, received a short but detailed training session (≈1 hour).

For the intra-rater reliability investigation, the PASS-IT was tested twice by Researcher 1 within 24 hours. During the retest, Researcher 1 had no access to scores collected the first time. For the inter-rater reliability investigation, 30 patients were assessed at the same time by Researchers 1 and 2. No discussion or comparison was allowed between raters. All assessments were conducted bedside in the Stroke Unit, with the bed in a low position and without side rails, using a stopwatch and a PASS-IT form with operating instructions. To assess concurrent validity, Researcher 1 applied to the 49 patients the following test and scales:

- *Postural Assessment Scale for Stroke Patients:* specially designed for individuals with stroke, it evaluates both aspects of postural control: maintaining a posture and changing posture. It has good sensitivity, since it uses 12 items with increasing levels of difficulty in the three fundamental positions (lying, sitting, standing) and in postural variations, with four response options for each item (0 to 3; 3 = best performance), and a total score ranging from 0 to 36. It does not require specific training, nor equipment, except for a stopwatch. It can be safely administered by doctors and physiotherapists; the administration time varies from 1 to 10 minutes (19).
- *Trunk Control Test:* one of the best known and easiest to administer tools to assess trunk control in stroke (12, 13), it evaluates the patient in the lying and sitting position. It consists of four items with three response options (0 = unable to perform movement without assistance; 12 = able to perform movement, but in an abnormal style, e.g., pulls on bed clothes; 25 = able to complete movement normally), and a total score ranging from 0 (minimum) to 100 (maximum, indicating better performance) (12).
- *Barthel Index:* an ordinal scale developed in 1965 for use in rehabilitation patients with stroke and other neuromuscular or musculoskeletal disorders (36), it measures the degree of functional independence or need of assistance of an individual, evaluating 10 common activities of daily living (ADLs) with item scores ranging from 0 to 15, and a total score ranging from 0 (minimum) to 100 (maximum, indicating that no assistance is required to complete the activities). A validated Italian translation was used (37).
- *Functional Ambulation Categories:* developed in 1984, it is a 6-point scale that evaluates how much human support the patient requires when walking, considering different settings. The score ranges from 0 (patient cannot walk) to 5 (independent ambulation on any surface) (38).

Statistical analysis

Data were entered into a dedicated database, arranged by variables and finally analyzed using the Statistical Package for the Social Sciences (SPSS) 28.0 for Windows. Demographic data were analyzed using descriptive statistical tests.

The structural validity of the PASS-IT was evaluated with the explorative factor analysis.

The oblique (Varimax) rotation was used. The appropriateness of the factor analysis was evaluated using the Kaiser-Meyer-Olkin (KMO) test (39) and Bartlett's test. Sampling was considered adequate if KMO was higher than 0.6. Additionally, the result of Bartlett's test of sphericity must be less than 0.05 to indicate validity and suitability of the responses collected for the purpose of the study. The number of factors was determined using the scree plot, the overall variance, and the pattern matrix. Two-tailed p-values less than 0.05 were considered statistically significant.

The internal consistency of the PASS-IT was assessed using the Cronbach's alpha coefficient (α); $\alpha \geq 0.70$ indicates high inter-item correlation and good homogeneity of the scale. Item-to-total correlation, corrected inter-item correlation, and Cronbach's α when the item is deleted were evaluated. Item-to-total correlation represents the correlation between an individual item's score and the total score of all other items in a scale, indicating how well a particular item aligns with the overall construct being measured. A common

cutoff for an acceptable item-to-total correlation is 0.30 or higher, suggesting that the item contributes well to the overall reliability of the scale. Corrected inter-item correlation is the correlation between each item and the sum of the other items, excluding itself, which helps to avoid inflating the correlation value. It indicates how similar an item is to the rest of the items in a scale. Ideally, corrected inter-item correlations should be between 0.20 and 0.50, ensuring that items are related but not redundant. Values below 0.20 suggest the item might not fit well, while values above 0.50 might indicate redundancy. Cronbach's α when the item is deleted represents the internal consistency reliability of a scale after the hypothetical removal of a specific item. Values that increase significantly upon deletion suggest that the item may be negatively contributing to the homogeneity of the scale, whereas minimal changes imply that the item is well-aligned with the overall construct being measured.

For construct validity, the PASS-IT was compared with the TCT. Since previous studies (20, 40) demonstrated the strong positive correlation between the PASS and the level of independence in ADLs and walking, a correlation analysis was also performed to explore the relationship with the BI and the FAC. The correlation was examined using the Spearman's rho (ρ), whose value varies between −1 (perfect negative association) and 1 (perfect positive association), with 0 indicating no association. A correlation of 0.70 or higher, which is considered a strong association, was expected.

F/C effect, occurring when the score does not change from minimum or maximum despite clinical change, is defined as the proportion of participants scoring the lowest (floor) or highest (ceiling) possible score. It is considered to be present if 15% or more achieve the lowest or highest score. It indicates low reliability and limited responsiveness of the scale, since a change of performance in these participants cannot be measured (41).

The intra- and inter-rater reliability was assessed with the intraclass correlation coefficient (ICC) and 95% confidence interval (CI). ICC values of 0.70-0.89 indicate high agreement, 0.90-0.99 very high agreement, 1.00 perfect agreement.

To assess the significance of changes observed in our measurements, the Standard Error of Measurement (SEM) and the Minimal Detectable Change (MDC) were calculated. The SEM was computed to quantify the variability inherent in our measurement process, ensuring an understanding of the precision of our data. The SEM was calculated using the formula: $SEM = SD*\sqrt{(1 - ICC)}$, where SD represents the standard deviation of the baseline measurement.

The MDC was then derived to determine the smallest unit of change that can be detected by the instrument beyond measurement error. The MDC was calculated using the formula: $MDC = SEM \times Z_{1-\alpha/2}\ x$, where $Z_{1-\alpha/2}$ is the z-value corresponding to the desired confidence level (typically 1.96 for a 95% confidence level), and the factor adjusts for the two measurements being compared. This ensures that any observed change equal to or greater than the MDC is unlikely to be due to random measurement error, but rather reflects a true change in the underlying phenomenon being measured.

Since the PASS and the TCT have different ranges (PASS 0-36, TCT 0-100), direct comparison of the SEM and MDC values could lead to misleading interpretations. Larger ranges naturally produce higher absolute values for SEM and MDC, which might not reflect a true difference in the relative precision or variability of the scales. To allow for meaningful comparison, the normalized SEM and normalized MDC were calculated, dividing each value by the respective scale range and expressing the result as a percentage: normalized SEM = (SEM/Range) × 100; normalized MDC = (MDC/Range) × 100. This normalization process enables us to compare the relative error and detectability of changes across both scales, independent of their absolute range, allowing for a more accurate evaluation of the precision and reliability of the two tests.

Results

Translation and cross-cultural adaptation

The focus group with 21 physiotherapists highlighted difficulties related to the interpretation of ambiguous terms. For example, "support" can have several translations in Italian (*Appoggio, Sostegno, Supporto,* or *Assistenza*) with different meanings, confusing the active participation of the patient with the help provided by an external operator. We chose the terms that enjoyed the broadest understanding.

Critical issues were resolved by consensus. To overcome them, a final version was produced (Appendix 1, Supplementary Material), with operating instructions and a different order of the items (Fig. 2). The PASS-IT was submitted to the developers of the original PASS (19), receiving their approval.

Psychometric evaluation

Fifty-one patients had been initially recruited; two were excluded at a later time because they were also diagnosed with a brain tumor. The characteristics of the participants are presented in Table 1.

Further information about participants' sensory disorders, unilateral spatial neglect (USN), upper/lower limb spasticity and function, as well as frequency distributions and percentage of the scores collected for each item are provided in the Supplementary Material (Table S1 and S2, respectively).

PASS-IT	Item	Original PASS
1	**Supine to affected side lateral**	6
2	**Supine to nonaffected side lateral**	7
3	**Supine to sitting up on the edge of the table**	8
4	**Sitting without support**	1
5	**Sitting to standing up**	10
6	**Standing with support**	2
7	**Standing without support**	3
8	**Standing on nonparetic leg**	4
9	**Standing on paretic leg**	5
10	**Standing, picking up a pencil from the floor**	12
11	**Standing up to sitting down**	11
12	**Sitting on the edge of the table to supine**	9

FIGURE 2 - New order of the items.

TABLE 1 - Characteristics of the sample of patients included in the study

		Validity and intra-rater reliability study N = 49	Inter-rater reliability study N = 30
Age (in years)	Mean±SD	71.96±12.46	68.57±13.02
	Median (range)	75 (41-89)	69 (41-89)
Gender	Women	21 (42.86%)	14 (46.67%)
	Men	28 (57.14%)	16 (53.33%)
Diagnosis	Ischemic	40 (81.63%)	24 (80%)
	Hemorrhagic	9 (18.37%)	6 (20%)
Brain injuries	Right	24 (48.98%)	14 (46.67%)
	Left	25 (51.02%)	16 (53.33%)
Days between date of ictus and date of testing	Mean±SD	5±2.68	5±2.59
	Median (range)	4 (1-11)	5 (1-12)

SD=standard deviation.
N = 49 patients included in the validity and intra-rater reliability study;
N = 30 patients included in the inter-rater reliability study.

A principal component analysis was performed. The KMO test (0.880, $p < 0.01$) confirmed the appropriateness of the factor analysis and of the sample size. Bartlett's test of sphericity was < 0.001. The scree plot (Figure S1 in Supplementary Material), the overall variance, and the pattern matrix showed a two-dimensional structure. The two components were studied with a Varimax rotation with Kaiser normalization (Table S3 in Supplementary Material). In factor analysis, items are allocated to factors according to the highest factor loadings, typically using a threshold of 0.3 or 0.4. Two unexpected groups of items were identified: items 1-4 plus 12 (activities performed in lying/sitting position, "non-weight-bearing") and items 5-11 (activities standing, "weight-bearing"). High reliability was found for both subsections: "non-weight-bearing" (ICC = 0.865; 95% CI: 0.795-0.917) and "weight-bearing" (ICC = 0.949; 95% CI: 0.924-0.968). These two subsections, which were not further investigated in this study, do not coincide with those of the original PASS ("maintaining a posture" and "changing posture").

Table 2 shows the scores for the four scales administered by Researcher 1 on the same test occasion: PASS-IT, TCT, BI, and FAC. Spearman's rho (ρ) showed high concurrent validity between the PASS-IT and the TCT ($\rho = 0.845$, $p < 0.001$) and a strong correlation with the BI ($\rho = 0.884$, $p < 0.001$) and the FAC ($\rho = 0.889$, $p < 0.001$).

Table 3 shows the internal consistency results. For each item the median score with Interquartile Range (IQR) is shown, together with the item-to-total correlation, corrected inter-item correlation, and Cronbach's α when the item is deleted. The item-to-total correlation shows a value of 0.390 for the first item, while the others range from 0.663 to 0.939. The corrected inter-item correlation shows a value of 0.363 for the first item, with the others ranging from 0.617 to 0.921. Cronbach's α coefficient, regardless of which item is deleted, is always > 0.90 (range 0.929-0.947).

In the intra-rater reliability study, the mean total score for the PASS-IT is 24.15 ± 10.14 for the first assessment and 24.29 ± 10.16 for the second assessment made by Researcher 1. The mean interval between assessments was 9h43' (SD = 6h54'; range 3h30' to 21h30'). There is high reliability between total scores (ICC = 0.942; 95% CI: 0.914-0.963; $p < 0.001$) and for each item between first and second assessment, with ρ ranging from 0.817 to 0.991.

In the inter-rater reliability study, the mean total score for the PASS-IT is 26.00 ± 9.60 for Researcher 1 and 26.03 ± 9.60 for Researcher 2. The assessments are highly consistent for total scores (ICC = 0.940; 95% CI: 0.903-0.968; $p < 0.001$) and for single items, with ρ ranging from 0.988 to 1.000.

The SEM is 1.72 points for the intra-rater, and 1.63 points for the inter-rater reliability. The $MDC_{95\%}$ is 4.76 based on intra-rater reliability data.

In the PASS-IT, five patients (10.20%) reached the maximum score (mean 24.22/36; range 3 to 36). Nine patients (18.37%) scored between 34 and 36/36. No one scored 0. In the TCT, 29 patients (59.18%) received the highest score (mean 77.63/100; range 0 to 100). The zero score was given to two patients (4.08%).

With the new sequence of items, the time of administration decreased (pilot study, mean 12'14", range 5 to 20 minutes; psychometric study, mean: 7'55 ", range 4'35 " to 15'25 ").

Discussion

The PASS-IT, with concise operating instructions, showed a high intra- and inter-rater reliability, reflecting a uniform application and interpretation of the scale. The involvement of a large group of physiotherapists, together with the

TABLE 2 - Scores of administered scales and relationship with the PASS-IT

	Median (IQR)	Minimum	Maximum	Concurrent validity/correlation with the PASS-IT
PASS-IT	27 (18)	3	36	---
TCT	100 (38)	0	100	$\rho = 0.845$, $p < 0.001$
BI	45 (80)	0	100	$\rho = 0.884$, $p < 0.001$
FAC	1 (4)	0	5	$\rho = 0.889$, $p < 0.001$

BI=Barthel index; FAC=Functional Ambulation Categories; IQR=interquartile range; PASS-IT=Postural Assessment Scale for Stroke Patients, Italian; TCT=Trunk Control Test.

TABLE 3 - Internal consistency results

		Intra-rater		
	Median (IQR)	**Item-to-total correlation**	**Corrected inter-item correlation**	**Cronbach's α when item is deleted**
1. Supine to affected side lateral	3 (0)	0.390	0.363	0.947
2. Supine to the nonaffected side lateral	3 (0)	0.663	0.628	0.941
3. Supine to sitting up on the edge of the table	3 (1)	0.800	0.771	0.937
4. Sitting without support	3 (0)	0.734	0.686	0.938
5. Sitting to standing up	3 (3)	0.939	0.921	0.929
6. Standing with support	3 (3)	0.922	0.897	0.930
7. Standing without support	3 (3)	0.933	0.911	0.930
8. Standing on nonparetic leg	0 (1)	0.703	0.643	0.940
9. Standing on paretic leg	0 (1)	0.680	0.617	0.940
10. Standing, picking up a pencil from the floor	2 (3)	0.882	0.843	0.933
11. Standing up to sitting down	3 (3)	0.929	0.907	0.930
12. Sitting on the edge of the table to supine	3 (1)	0.804	0.779	0.938

IQR = interquartile range.

rigorous method followed for the cross-cultural translation, can be considered a strength of this study.

The modified sequence of the items, chosen also for the Swedish (Swe-PASS) (30) and Norwegian versions (Swe-PASS-NV) (32), allowed a quicker and smoother administration, avoiding repeated unnecessary changes of posture. The clear distinction between the original two sections of the scale ("Maintaining a posture" and "Changing posture") is not featured, but the total score of the test does not change.

Although the PASS requires no training for use, the pilot study highlighted the need for operating instructions, consisting of concise and pragmatic indications, shown before each item in the evaluation form, aiming to increase the uniformity both in the administration (setting, instruction given to the patient) and in the interpretation of the results. For example, for items 4, 7, 8, 9 (sitting without support, standing without support, standing on the nonparetic/paretic leg) it is essential to use a stopwatch, as planned by the authors of the PASS (20). Underestimating its use in the pilot study led to imprecise scores.

In items 1-3, 5, 10-12, different scores are expected based on the amount of help received ("much help," "little help," "without help") (19). There are no tools to quantify numerically the help provided, which is subjectively affected by age, technique, experience, build, and physical training of the operator. The Swe-PASS replaced "much help" with "support from two persons" and "little help" with "support from one person" (30). After extensively discussing this aspect, the original definition was preferred for the PASS-IT.

The original PASS describes a setting with a Bobath-type plane 50 cm in height, with the person's feet resting on the floor. For short people, it is difficult to touch the ground from this height, except by moving dangerously forward on the edge of the couch. For taller people, excessive bending of the lower limbs leads to unfavorable leverage. A sitting position with hips and knees bent at 90° is therefore advisable, as recommended by the new instructions for items 4 and 5. When the patient rests on an anti-decubitus mattress, it is essential that the mattress be maximally inflated, to allow stability. Items in standing were tested wearing shoes, for greater safety and hygiene, even if the original PASS doesn't specify whether the patient should be wearing shoes or be barefoot.

While other cross-cultural validation studies evaluated the patients at a chronic stage, or through a video-recorded performance (31, 32), our study was carried out in an acute setting, when a wide variety of factors (functional improvement, caution, fatigue, fear, disorientation) can produce sudden changes. For this reason, like in the Swedish validation study (30), we chose a short interval (< 24 hours) between the intra-rater observations, despite this representing a weakness of the study because of recall bias. A longer interval could result in higher grades at the second assessment, due to the functional recovery of the patient, or because posture changes had been practiced with the physiotherapist or the ward staff.

The value of verbal indications for carrying out the activity was also questioned. To standardize the application of the scale, it is important that no verbal indications be provided that are useful for its performance. Should they become necessary, the verbal indications would be considered as "little help," like in the TCT (12).

The operating instructions emphasize that the PASS, for each of the 12 items, evaluates the ability to perform the activity, without considering its quality. They also draw attention to the fact that only one attempt is allowed for each item.

The PASS does not evaluate the ability to perform activities while seated. When this aspect is essential, it is advisable that more specific tools be used, such as the TIS (5), of

which the validated translation in Italian is available (42), or the Function in Sitting Test (43).

No patient scored zero in item 1 (Supine to affected side lateral), and only one patient scored zero in item 2 (Supine to nonaffected side lateral). This can be partially explained by our exclusion criteria. Breistein et al (32) discussed the possibility that score 0 might be redundant for this item, since even patients with little to no functional independence can be turned on one side with help from one or two persons. We agreed that a zero be scored when medical conditions contraindicate rolling the patient to one side, but also when the patient shows no participation or involvement in the action.

In item 7, the highest score is given if the person "can stand without support for more than one minute and at the same time perform arm movement above the shoulder level." It is important to select specific movements, with a predefined number of repetitions, to obtain comparable results. Persson et al (30) indicate the act of moving "hand/s from the forehead to the neck (like pulling your fingers through your hair)" and bringing the arm back, relaxed along the trunk, without however specifying the number of repetitions. We recommend at least five repetitions of this movement.

As shown in Table S2, at first assessment 59.2% of patients were not able to stand on the non-paretic leg (item 8), 63.3% were not able to stand on the paretic leg (item 9), and 46.9% were not able to pick up a pencil from the floor (item 10). These findings are in line with those of the Swe-PASS (30) and with the results described by Benaim et al (19) for the group of patients evaluated on day 30 after a stroke. In items 8 and 9 it is not clear whether the foot which is not bearing the weight must be lifted off the ground, or whether bending the hip and/or the knee is required. During the pilot study, these items had the worst inter-rater agreement. It was therefore specified that the foot must be lifted off the ground, with the opposite one entirely bearing the weight.

We consider the two-dimensional structure identified by the principal component analysis an unexpected and interesting finding. There is a clear division between items "non-weight-bearing" (1-4 plus 12), where the patients show much better performance, and the "weight-bearing" section (items 5-11), with a floor effect for some tasks. Nothing changes in the way the scale is administered and scored, but we believe that a "non-weight-bearing" and "weight-bearing" subdivision can be relevant to better comprehend the patient's improvement and need for rehabilitation.

The present study confirmed a good correlation with the TCT ($\rho = 0.845$, $p < 0.001$) and the FAC ($\rho = 0.889$, $p < 0.001$), which had not been tested before. Our findings of a good correlation with the BI ($\rho = 0.884$, $p < 0.001$) are consistent with the values reported by Mao et al (20) and Chien et al (40).

The PASS-IT showed very high reliability, in line with Benaim et al (19) and other transcultural validations (30-33).

Considering that our patients were in the acute or early subacute phase after a stroke, and nine of them (18.37%) scored between 34 and 36/36, we agree with Chinsongkram et al (21) and with Wang et al (22) that a ceiling effect is plausible already in the first weeks for patients with high functional ability. The TCT showed a significant ceiling effect, with 29 patients (59.18%) earning the highest score. Considering this, the SEM (PASS = 1.72, normalized 4.77; TCT = 11.30), and the MDC (PASS = 4.76, normalized 13.22; TCT = 31.34), in our clinical practice the PASS-IT is more advisable, being more responsive to slight changes (e.g., going from "with much help" to "with little help"). Moreover, it evaluates the patients also in the standing position and single-leg stance.

The original PASS does not report an MDC value. Hsueh et al (44) indicate an MDC of 1.8 ± 1.7 for acute stroke, but their assessment was conducted at 14 and 30 days from onset. Breistein et al (32) calculated smallest detectable difference (SDD) = 1.9 points (intra-rater) and 2.7 points (inter-rater), but attested that "the measurement error may be considered to be artificially low" due to the use of video recorded assessment, without a real change in the participants' performance.

Since different statistical methods were employed, a direct comparison of the results is often impossible.

Conclusion

The PASS-IT is a valid and reliable tool, suitable for clinical and physiotherapy practice in the acute and subacute stage. The final version of the cross-cultural translation, which includes short operating instructions and a different sequence for the 12 items, overcame critical issues encountered during the pilot study. It serves the purpose of promoting a uniform application and interpretation of the scale among Italian health professionals and researchers. Further study is needed to investigate the potential information provided by considering the scale in its two dimensions, "weight-bearing" and "non-weight-bearing."

Acknowledgments

The authors wish to thank the patients who participated in this study and the physiotherapists of the Unit of Physical Medicine and Rehabilitation of the IRCCS Azienda Ospedaliero-Universitaria di Bologna for their help in thoroughly discussing every critical aspect, and for being such a committed and passionate work team.

A sincere appreciation goes to the reviewers who, with their competence and attention, provided valuable guidance.

Disclosures

Conflict of interest: The authors declare that there is 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.

Ethics approval: The study was approved by the Ethical Committee of the IRCCS University Hospital of Bologna (Reference N. 216/2017/O/Sper).

Authors contribution: EL: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, Writing – Original Draft, Writing – Review & Editing; NG: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Software, Supervision, Validation, Visualization, Writing – Original Draft, Writing – Review & Editing;

AB: Conceptualization, Data Curation, Investigation, Project Administration, Supervision, Validation, Writing – Review & Editing; AB: Conceptualization, Investigation, Project Administration, Resources, Supervision, Validation; LG: Conceptualization, Data Curation, Formal Analysis, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft, Writing – Review & Editing; MR: Conceptualization, Data Curation, Formal Analysis, Resources, Software, Validation, Visualization, Writing – Original Draft, Writing – Review & Editing; DG: Conceptualization, Data Curation, Formal Analysis, Methodology, Resources, Software, Supervision, Validation, Writing – Review & Editing.

Data Availability Statement: The data presented in this study and the printable Italian Version of the PASS are available on request from the corresponding author.

References

1. Adamson J, Beswick A, Ebrahim S. Is stroke the most common cause of disability? J Stroke Cerebrovasc Dis. 2004;13(4):171-177. CrossRef PubMed
2. Hankey GJ. Stroke. Lancet. 2017;389(10069):641-654. CrossRef PubMed
3. Tsao CW, Aday AW, Almarzooq ZI, et al. Heart disease and stroke statistics-2022 update: a report from the American Heart Association. Circulation. 2022;145(8):e153-e639. CrossRef PubMed
4. Wafa HA, Wolfe CDA, Emmett E, Roth GA, Johnson CO, Wang Y. Burden of stroke in Europe: thirty-year projections of incidence, prevalence, deaths, and disability-adjusted life years. Stroke. 2020;51(8):2418-2427. CrossRef PubMed
5. Verheyden G, Nieuwboer A, Mertin J, Preger R, Kiekens C, De Weerdt W. The Trunk Impairment Scale: a new tool to measure motor impairment of the trunk after stroke. Clin Rehabil. 2004;18(3):326-334. CrossRef PubMed
6. Jijimol G, Fayaz RK, Vijesh PV. Correlation of trunk impairment with balance in patients with chronic stroke. NeuroRehabilitation. 2013;32(2):323-325. CrossRef PubMed
7. Hsieh CL, Sheu CF, Hsueh IP, Wang CH. Trunk control as an early predictor of comprehensive activities of daily living function in stroke patients. Stroke. 2002;33(11):2626-2630. CrossRef PubMed
8. Kim TJ, Seo KM, Kim DK, Kang SH. The relationship between initial trunk performances and functional prognosis in patients with stroke. Ann Rehabil Med. 2015;39(1):66-73. CrossRef PubMed
9. Kong KH, Ratha Krishnan R. Truncal impairment after stroke: clinical correlates, outcome and impact on ambulatory and functional outcomes after rehabilitation. Singapore Med J. 2021;62(2):87-91. CrossRef PubMed
10. Martins LG, Molle da Costa RD, Alvarez Sartor LC, et al. Clinical factors associated with trunk control after stroke: A prospective study. Top Stroke Rehabil. 2021;28(3):181-189. CrossRef PubMed
11. Lesser M, Borst J, Dekerlegand J. Use of the Postural Assessment Scale for Stroke Patients in determining acute care discharge recommendations. J Acute Care Phys Ther. 2017;8(3):79-85. CrossRef
12. Collin C, Wade D. Assessing motor impairment after stroke: a pilot reliability study. J Neurol Neurosurg Psychiatry. 1990; 53(7):576-579. CrossRef PubMed
13. Franchignoni FP, Tesio L, Ricupero C, Martino MT. Trunk control test as an early predictor of stroke rehabilitation outcome. Stroke. 1997;28(7):1382-1385. CrossRef PubMed
14. Fil Balkan A, Salcı Y, Keklicek H, Çetin B, Adın RM, Armutlu K. The trunk control: which scale is the best in very acute stroke patients? Top Stroke Rehabil. 2019;26(5):359-365. CrossRef PubMed
15. Verheyden G, Vereeck L, Truijen S, et al. Trunk performance after stroke and the relationship with balance, gait and functional ability. Clin Rehabil. 2006;20(5):451-458. CrossRef PubMed
16. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. Scand J Rehabil Med. 1975;7(1):13-31. CrossRef PubMed
17. Cecchi F, Carrabba C, Bertolucci F, et al. Transcultural translation and validation of Fugl-Meyer assessment to Italian. Disabil Rehabil. 2021;43(25):3717-3722. CrossRef PubMed
18. Sanford J, Moreland J, Swanson LR, Stratford PW, Gowland C. Reliability of the Fugl-Meyer assessment for testing motor performance in patients following stroke. Phys Ther. 1993;73(7):447-454. CrossRef PubMed
19. Benaim C, Pérennou DA, Villy J, Rousseaux M, Pelissier JY. Validation of a standardized assessment of postural control in stroke patients: the Postural Assessment Scale for Stroke Patients (PASS). Stroke. 1999;30(9):1862-1868. CrossRef PubMed
20. Mao HF, Hsueh IP, Tang PF, Sheu CF, Hsieh CL. Analysis and comparison of the psychometric properties of three balance measures for stroke patients. Stroke. 2002;33(4):1022-1027. CrossRef PubMed
21. Chinsongkram B, Chaikeeree N, Saengsirisuwan V, Viriyatharakij N, Horak FB, Boonsinsukh R. Reliability and validity of the Balance Evaluation Systems Test (BESTest) in people with subacute stroke. Phys Ther. 2014;94(11):1632-1643. CrossRef PubMed
22. Wang CH, Hsueh IP, Sheu CF, Hsieh CL. Discriminative, predictive, and evaluative properties of a trunk control measure in patients with stroke. Phys Ther. 2005;85(9):887-894. CrossRef PubMed
23. Huang YJ, Lin GH, Lee SC, Hsieh CL. A comparison of the responsiveness of the Postural Assessment Scale for Stroke and the Berg Balance Scale in patients with severe balance deficits after stroke. J Geriatr Phys Ther. 2020;43(4):194-198. CrossRef PubMed
24. Estrada-Barranco C, Cano-de-la-Cuerda R, Abuín-Porras V, Molina-Rueda F. Postural Assessment Scale for Stroke Patients in acute, subacute and chronic stage: a construct validity study. Diagnostics (Basel). 2021;11(2):365. CrossRef PubMed
25. Liaw LJ, Hsieh CL, Lo SK, Chen HM, Lee S, Lin JH. The relative and absolute reliability of two balance performance measures in chronic stroke patients. Disabil Rehabil. 2008;30(9):656-661. CrossRef PubMed
26. Di Monaco M, Trucco M, Di Monaco R, Tappero R, Cavanna A. The relationship between initial trunk control or postural balance and inpatient rehabilitation outcome after stroke: a prospective comparative study. Clin Rehabil. 2010;24(6):543-554. CrossRef PubMed
27. Schuling J, de Haan R, Limburg M, Groenier KH. The Frenchay Activities Index. Assessment of functional status in stroke patients. Stroke. 1993;24(8):1173-1177. CrossRef PubMed
28. Huang YC, Wang WT, Liou TH, Liao CD, Lin LF, Huang SW. Postural Assessment Scale for Stroke Patients Scores as a predictor of stroke patient ambulation at discharge from the rehabilitation ward. J Rehabil Med. 2016;48(3):259-264. CrossRef PubMed
29. Yoneyama SM, Roiz R de M, Oliveira TM, Oberg TD, Lima NMFV. Validação da versão brasileira da Escala de Avaliação Postural para pacientes após acidente vascular encefálico. Acta Fisiatr. 2008;15(2):96-100. CrossRef
30. Persson CU, Hansson P-O, Danielsson A, Sunnerhagen KS. A validation study using a modified version of Postural Assessment Scale for Stroke Patients: Postural Stroke Study in Gothenburg (POSTGOT). J Neuroeng Rehabil. 2011;8(1):57. CrossRef PubMed

31. Cabanas-Valdés R, Girabent-Farrés M, Cánovas-Vergé D, Caballero-Gómez FM, Germán-Romero A, Bagur-Calafat C. [Spanish translation and validation of the Postural Assessment Scale for Stroke Patients (PASS) to assess balance and postural control in adult post-stroke patients]. Rev Neurol. 2015;60(4):151-158. PubMed
32. Breistein K, Gjelsvik BEB, Jørgensen L. The Postural Assessment Scale for Stroke Patients: translation into Norwegian, cultural adaptation, and examination of reliability. Eur J Physiother. 2017;19(4):207-214. CrossRef
33. Koçak FA, Kurt EE, Koçak Y, Erdem HR, Tuncay F, Benaim C. Validity and interrater/intrarater reliability of the Turkish version of the postural assessment scale for stroke patients (PASS-Turk). Top Stroke Rehabil. 2019;26(5):373-381. CrossRef PubMed
34. Sullivan JE, Crowner BE, Kluding PM, et al. Outcome measures for individuals with stroke: process and recommendations from the American Physical Therapy Association neurology section task force. Phys Ther. 2013;93(10):1383-1396. CrossRef PubMed
35. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000;25(24):3186-3191. CrossRef PubMed
36. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J. 1965;14:61-65. PubMed
37. Castiglia SF, Galeoto G, Lauta A, et al. The culturally adapted Italian version of the Barthel Index (IcaBI): assessment of structural validity, inter-rater reliability and responsiveness to clinically relevant improvements in patients admitted to inpatient rehabilitation centers. Funct Neurol. 2017;22(4):221-228. CrossRef PubMed
38. Holden MK, Gill KM, Magliozzi MR, Nathan J, Piehl-Baker L. Clinical gait assessment in the neurologically impaired. Reliability and meaningfulness. Phys Ther. 1984;64(1):35-40. CrossRef PubMed
39. Yaşar M. Attitudes toward statistics scale: validity and reliability study. Pamukkale University Journal of Education. 2014;02:59-75. Online
40. Chien CW, Lin JH, Wang CH, Hsueh IP, Sheu CF, Hsieh CL. Developing a Short Form of the Postural Assessment Scale for people with stroke. Neurorehabil Neural Repair. 2007;21(1):81-90. CrossRef PubMed
41. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34-42. CrossRef PubMed
42. Lombardi B, Orioli A, Casavola D, Paci M. The Italian version of the Trunk Impairment Scale: development and psychometric properties. Eur J Phys Rehabil Med. 2017;53(4):516-520. CrossRef PubMed
43. Gorman SL, Radtka S, Melnick ME, Abrams GM, Byl NN. Development and validation of the Function In Sitting Test in adults with acute stroke. J Neurol Phys Ther. 2010;34(3):150-160. CrossRef PubMed
44. Hsueh I-P, Chen K-L, Chou Y-T, Wang Y-H, Hsieh C-L. Individual-level responsiveness of the original and short-form postural assessment scale for stroke patients. Phys Ther. 2013;93(10):1377-1382. CrossRef PubMed

Arch Physioter 2024; 14: 105-115
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3245
REVIEW

Red flags for potential serious pathologies in people with neck pain: a systematic review of clinical practice guidelines

Daniel Feller[1,2], Alessandro Chiarotto[2], Bart Koes[2,3], Filippo Maselli[4,5], Firas Mourad[6,7]

[1]Provincial Agency for Health of the Autonomous Province of Trento, Trento - Italy
[2]Department of General Practice, Erasmus MC, University Medical Centre, Rotterdam - The Netherlands
[3]Research Unit of General Practice, Department of Public Health and Center for Muscle and Joint Health, University of Southern Denmark, Odense - Denmark
[4]Department of Human Neurosciences, Sapienza University of Rome, Rome - Italy
[5]Sovrintendenza Sanitaria Regionale Puglia, INAIL, Bari - Italy
[6]Department of Health, LUNEX University of Applied Sciences, Differdange - Luxembourg
[7]Luxembourg Health and Sport Sciences Research Institute A.s.b.l., Differdange - Luxembourg

ABSTRACT

Introduction: We conducted a systematic review of clinical practice guidelines to identify red flags for serious pathologies in neck pain mentioned in clinical practice guidelines, to evaluate agreement in red flag recommendations across guidelines, and to investigate the level of evidence including what study type the recommendations are based on.

Methods: We searched for guidelines focusing on specific and nonspecific neck pain in MEDLINE, EMBASE, and PEDro up to June 9, 2023. Additionally, we searched for guidelines through citation tracking strategies, by consulting experts in the field, and by checking guideline organization databases.

Results: We included 29 guidelines, 12 of which provided a total of 114 red flags for fracture (n = 17), cancer (n = 21), spinal infection (n = 14), myelopathy (n = 15), injury to the spinal cord (n = 1), artery dissection (n = 7), intracranial pathology (n = 3), inflammatory arthritis (n = 2), other systemic disease (n = 6), or unrelated to a specific condition (n = 19). Overall, there is very little agreement (median Fleiss' kappa of 0) between guidelines on the red flags to screen for serious pathologies.

Conclusion: Red flags were mainly supported by expert opinions. We also observed a general lack of consensus among guidelines regarding which red flags to endorse. Considering the current limitations of the evidence, specific recommendations on which red flags to use cannot be provided, except for using the Canadian C-Spine rule for screening posttraumatic fractures.

Keywords: Differential diagnosis, Guidelines, Neck pain, Red flags

What is already known

- *Triaging serious cervical conditions mimicking musculoskeletal neck pain is a mainstay in primary care. Although identifying these pathologies can be challenging for clinicians, their recognition is relevant to determine which patients need to be referred to ensure safe and effective patient care.*

What does this study add

- *Almost all the red flags were only based on mechanism reasoning (Level of evidence 5). Diagnostic accuracy values for red flags were not reported, except for the Canadian C-spine rules. Therefore, clinicians should rely on red flags cautiously, integrating them with sound clinical reasoning.*

Received: August 4, 2024
Accepted: November 11, 2024
Published online: December 4, 2024

This article includes supplementary material

Corresponding author:
Daniel Feller
email: d.feller@erasmusmc.nl

Introduction

Neck pain is a complex biopsychosocial disorder estimated to be the eighth leading cause of Years Lived with Disability globally (1-3). Although benign in the large majority of patients, it is estimated that 1% of neck pain can be caused by underlying serious pathologies, such as malignancy, cervical arterial pathology, myelopathy, congenital craniovertebral anomalies, infection, or fracture (4-7). Screening for

serious pathology masquerading as nonspecific musculoskeletal neck pain is a real challenge for clinicians, particularly in a direct access setting (5,8-10). It has been estimated that a delayed diagnosis of serious cervical pathologies ranges from 5% to 20% of all cases accessing the emergency department with neck pain, with potentially life-threatening consequences in the worst-case scenario (11,12). Therefore, the early recognition of serious cervical pathologies is a mainstay for safe physiotherapy practice and allows clinicians to identify those patients who require referral to another healthcare professional for optimal management and best possible outcomes (13).

As standard practice, red flags have been used to guide physiotherapists in identifying serious cervical pathology (14). Red flags are cues from a patient's medical history and clinical examination potentially associated with a higher risk of serious conditions (15). As practical examples, a past history of cancer is considered a red flag for spinal malignancy, urinary incontinence associated with back pain raises suspicion for a cauda equina syndrome, and pulse changes during palpation of a peripheral artery with neuropathic-like pain in the lower extremities (namely, radicular pain) may suggest the presence of peripheral arterial disease (13,16,17).

The recently released International Federation of Orthopaedic Manipulative Physical Therapists Cervical Framework highlights the need for physiotherapists to use a differential diagnosis tool for informed and safe management of the cervical spine (4,18). Therefore, investigating red flags for neck pain remains a priority for an informed practice and the patient's safety (14). To the best of the authors' knowledge, no systematic review has been published investigating the recommended red flags for neck pain in clinical practice guidelines for their scientific validity. Furthermore, knowledge on the level of evidence red flag recommendations were based on (e.g., systematic reviews of diagnostic test accuracy studies, cross-sectional studies, mechanism-based reasoning) may help clinicians to value the recommendations' strength. Therefore, we aimed to: (1) identify red flags to triage serious pathologies recommended in clinical practice guidelines for neck pain, (2) evaluate the agreement in red flag recommendations across guidelines, and (3) investigate the level of evidence on which the red flag recommendations are based.

Methods

We used the "Preferred Reporting Items for Systematic Reviews and Meta-analyses" (PRISMA) checklist for the reporting of the present manuscript (19). The study protocol was registered on MedRxiv (20).

Eligibility criteria

According to the Classification of Neck Pain and Associated Disorders (NAD) (21), we included guidelines focusing on specific (NAD III) and nonspecific neck pain (NAD I/II). We excluded guidelines for serious neck pain (NAD IV) because we expected them to only address managing these conditions, not identifying them in patients presenting with musculoskeletal neck pain. Also, we excluded guidelines not explicitly focused on neck pain, such as guidelines in which neck pain is only briefly mentioned in the context of other disorders or a more complex topic (e.g., management of chronic pain in general). A document was considered as a clinical practice guideline if it fulfilled the following criteria (adapted from the PEDro criteria for evidence-based clinical practice guidelines (22)): it was produced under the auspices of a health professional association or society, public or private organization, healthcare organization or plan, or government agency; a systematic literature search and review of existing scientific evidence was performed during the guideline development; the guideline was based on published systematic reviews; and the guideline contained systematically developed statements that included recommendations, strategies, or information to guide decisions about appropriate healthcare (22).

We did not apply any restrictions regarding publication date and language. Non-English and non-Italian guidelines were translated using "DeepL Translate" (Online). In addition, we only included the most up-to-date version if multiple versions of the same guideline were present.

Study selection process

Without time restriction, we searched for guidelines in MEDLINE (via PubMed), EMBASE, and PEDro electronic databases on 09/06/2023. Supplementary Material 1 reports the full search strategy for these databases.

Guidelines were also searched through forward and backward citation tracking strategies (Web of Science on 12/07/2023), by consulting experts in the field (top 10 experts on neck pain according to ExpertScape.com on 15/07/2023), and by checking guideline organization databases. The following guideline organization databases were searched: the "Canadian Medical Association Infobase of clinical practice guidelines" (Online), the "Istituto Superiore Sanità – Sistema Nazionale LineeGuida" (Online), the "Guidelines International Network" (Online), the "National Institute for Clinical Excellence – NICE" (Online), the "OPTIMa collaboration" (Online), the "Guideline Central" (Online), the "Scottish Intercollegiate Guidelines Network – SIGN" (Online), and the "Agency for Healthcare Research and Quality" (Online). In addition, we screened the references of two recently published systematic reviews on guidelines for neck pain (23,24).

Duplicates were eliminated using the Deduplicator function of "Systematic Review Accelerator" (25). We used the online electronic systematic review software package (Rayyan QCRI) to organize and track the selection process (26). Two researchers independently performed the study selection process by title/abstract (DF and FMo, or DF and AC) and then by full text (DF and FMo). Any disagreement was resolved by consensus or by the decision of a third author (AC).

Data extraction process

Two reviewers (DF and FMa) performed the data extraction process independently using a standardized Excel form. The data extraction form was piloted on three included guidelines.

Any discrepancies were resolved with a consensus between the two authors and eventually by a third author's decision (AC).

We extracted the following data from each guideline: publication year, language of publication, association(s) or society(ies) which generated the guideline, serious pathologies considered (e.g., malignancy, fracture, infection, congenital craniovertebral anomalies, cervical arteries dysfunctions), reported red flags, if these red flags are presented for individual pathologies or in a more general sense (i.e., not tied to any specific pathology), level of the evidence of each red flag, how red flags were supported (study design, consensus of the guideline committee, or not reported), and, when available, the diagnostic accuracy underpinning each recommendation. We determined the level of evidence for each red flag recommended in the guidelines by extracting the citations provided in each source. The level of evidence was determined using the 2011 Levels of Evidence framework from the Oxford Centre for Evidence-Based Medicine (27). This classification system ranks evidence based on study design, with systematic reviews of cross-sectional studies representing the highest level and mechanism-based reasoning representing the lowest (Tab. 1). Two researchers independently determined the level of the evidence (DF and FMo). Any disagreement was resolved by consensus or by the decision of a third author (AC).

TABLE 1 - Level of evidence for diagnostic questions according to the 2011 framework by the Oxford Centre for Evidence-Based Medicine

Level of evidence	Description
Level 1	Systematic review of cross-sectional studies with consistently applied reference standard and blinding
Level 2	Individual cross-sectional studies with consistently applied reference standard and blinding
Level 3	Non-consecutive studies, or studies without consistently applied reference standards
Level 4	Case-control studies, or poor or non-independent reference standard
Level 5	Mechanism-based reasoning

Data synthesis

We calculated Fleiss' kappa to evaluate the agreement among guidelines recommendations (poor agreement <0.00, slight agreement 0.00–0.20, fair agreement 0.21–0.40, moderate agreement 0.41–0.60, substantial agreement 0.61–0.80, almost perfect agreement 0.81–1.00) (28). Additionally, to summarize the recommendations to triage serious pathologies and the study designs to support recommendations, we computed descriptive statistics (absolute and relative frequencies) and reported the results narratively.

Deviations from the protocol

Deviations from the published protocol were implemented in response to reviewers' requests. Specifically, we determined the level of evidence for each red flag recommended in the guidelines to enhance the rigor of our findings and provide a clearer interpretation of the results in relation to the existing literature.

Equity, diversity, and inclusion statement

The group of authors involved in this study comprises five males from two high-income countries, Italy and the Netherlands. Among these authors, three are physical therapists (AC, FM, and FMo), one is both a physical therapist and a statistician (DF), and the fifth is an epidemiologist (BK). The group maintains a balance in terms of junior, mid-career, and senior researchers. At the time of submission, DF is a first-year PhD student, AC is an assistant professor, while BK is a full professor. FM holds a PhD, and FMo is an assistant professor with clinical and research experience focused on neck pain. Both FM and FMo teach a postgraduate course in screening for referrals for physical therapists in Italy. All the authors have experience in conducting systematic reviews. Additionally, all the authors have attended multiple courses on planning and conducting literature reviews. It is worth noting that our search strategy and data extraction process were not biased toward any specific gender, race, culture, or socioeconomic level.

Results

We retrieved 4,431 records from database investigations, 532 of which were duplicates. Titles and abstracts screening was performed on the remaining 3,899 records; we also retrieved six records from expert consultations, three from guideline organization databases, and seven from citation tracking strategies. In total, 59 reports were selected for full-text analysis. Ultimately, 29 guidelines met the inclusion criteria and were included in the present systematic review (Fig. 1). Supplementary material 2 contains the references to the included guidelines.

Characteristics of the included guidelines

Of the 29 guidelines included in the study, 12 (41%) provided information on red flags for screening serious pathologies. Among the remaining guidelines, 10 (35%) contained recommendations for diagnosing neck pain but did not mention any signs or symptoms to screen for serious pathologies, while 7 (24%) did not provide any diagnostic recommendation. Supplementary material 3 reports the characteristics of the guidelines that do not report red flags.

Of the guidelines reporting red flags, 3 (25%) were developed for patients who suffered from whiplash-associated disorders (29-31), 5 (42%) for patients with NAD grade I to III (32-36), and 4 (33%) for mixed populations (e.g., whiplash and NAD) (1,37-39). Most studies mentioned red flags for specific pathologies (e.g., fracture, cancer, infection), while 3 (25%) described red flags unrelated to a particular disease (e.g., Whalen et al (33) did not specify any particular pathology but identified fever, alongside other signs and symptoms, as a warning sign for serious conditions) (29,33,40). Table 2 reports the complete characteristics of the 12 guidelines reporting on red flags.

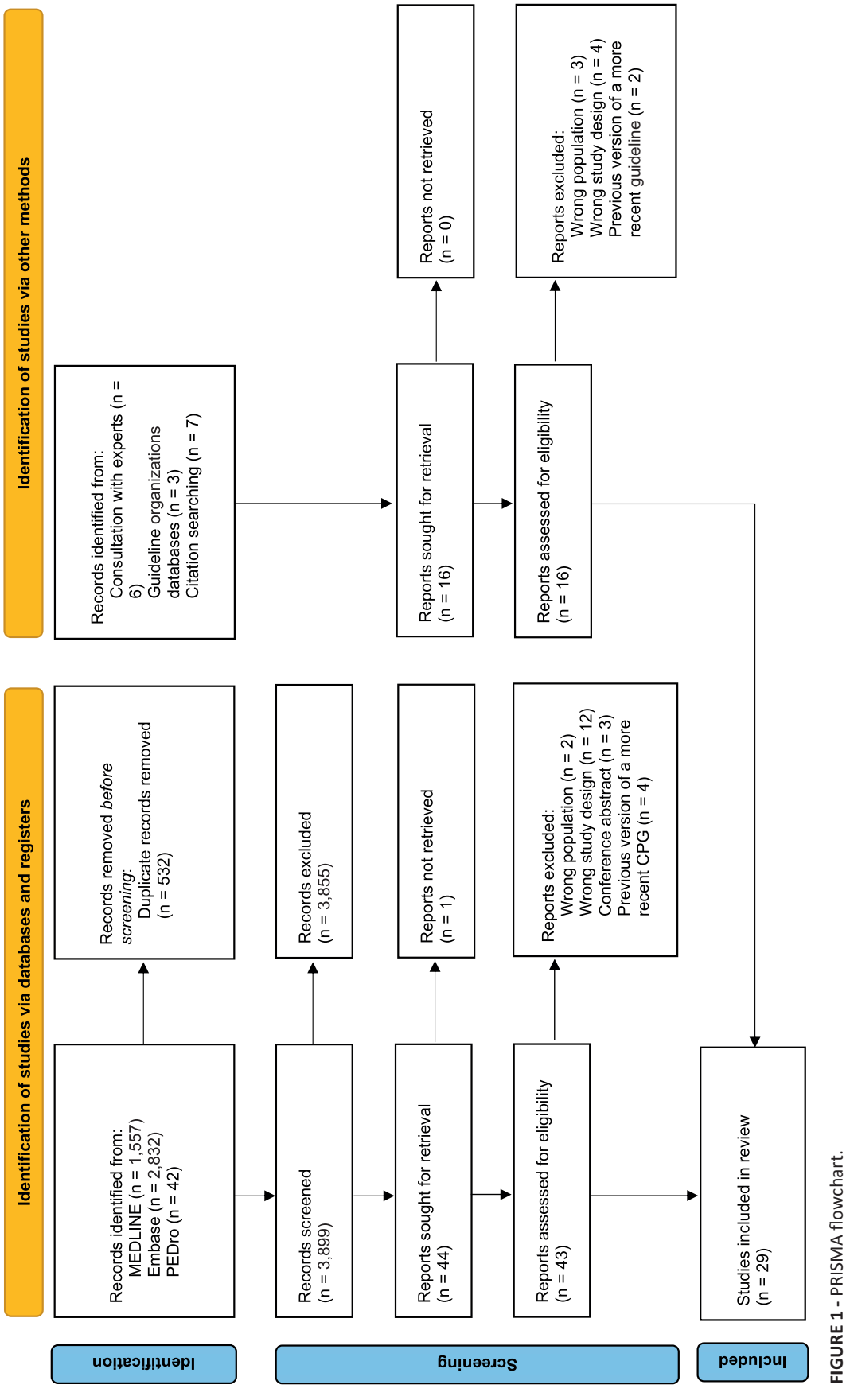


FIGURE 1 - PRISMA flowchart.

TABLE 2 - Characteristics of the guidelines with red flag recommendations

Author and year	Nation	Language	Society or body issuing the guidelines	Population (as reported in the guideline)	Cited pathologies	Strength of recommendation for screening for serious pathologies as reported in the guideline
Papic, 2023	Australia	English	New South Wales State Insurance Regulatory Authority	Acute or chronic WAD (grades I to III)	– Fracture	Strong
Whalen, 2019	USA	English	Scientific Council of the Clinical Compass	Neck pain of any duration	Red flags unrelated to specific disease	Not reported
Bier, 2018	Netherlands	English	Royal Dutch Society for Physical Therapy	NAD grades I to III, irrespective of the duration	– Fracture – Cancer – Vertebral infection – Cervical myelopathy – Spinal cord injury – Vertebral artery dissection – Systemic disease	"Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate"
Blanpied, 2017	USA	English	American Physical Therapy Association	Neck pain, including WAD, headache, and radicular pain	– Fracture – Cancer* – Vertebral infection* – Cervical myelopathy – Arterial insufficiency – Systemic disease* – Intracranial pathology – Systemic disease* – Cardiac involvement* – Unexplained cranial nerve dysfunction*	Strong ("one or more systematic reviews support the recommendation, providing evidence for a strong magnitude of effect")
Karppinen, 2017	Finland	Finnish	Finnish Medical Association Duodecim	Neck pain, including WAD and radiculopathy	– Fracture* – Cancer – Vertebral infection* – Cervical myelopathy – Cervical or carotid dissection – Systemic inflammatory disease*	Not reported
Lemeunier, 2017	France	French	Associations Francaise de Chiropratique and Institut Franco-Europèen de Chiropraxie	Neck pain of any duration	– Fracture – Osteoporotic fracture – Cancer – Vertebral infection – Cervical myelopathy – Carotid/vertebral artery dissection – Intracranial pathology – Inflammatory arthritis	Not reported

(Cont.)

TABLE 2 - Continued

Author and year	Nation	Language	Society or body issuing the guidelines	Population (as reported in the guideline)	Cited pathologies	Strength of recommendation for screening for serious pathologies as reported in the guideline
Bussières, 2016	Canada	English	Canadian Chiropractic Guideline Initiative	NAD grade I to III and WAD grade I to III	– Major structural pathologies* – Other pathologies rather than NAD or WAD*	Not reported
Côté, 2016	Canada	English	Ontario Protocol for Traffic Injury Management Collaboration	NAD grades I-III of less than 6 months duration, and WAD grades I-III of less than 6 months duration	– Fracture/dislocation – Osteoporotic fracture – Cancer – Vertebral infection – Cervical myelopathy – Carotid/vertebral artery dissection – Intracranial pathology – Inflammatory arthritis	Not reported
Scherer, 2016	Germany	German	German Society for General Medicine and Family Medicine	Neck pain of any duration	Red flags unrelated to specific disease	
Monticone, 2013	Italy	English	Italian Society of Physical and Rehabilitation Medicine	Neck pain with or without limb involvement and/or headache	– Fracture – Cancer – Vertebral infection – Cervical myelopathy – Systemic disease	Strongly recommended
Sterling, 2008	Australia	English	South Australian Centre for Trauma and Injury Recovery	WAD grades I to IV (both acute and chronic stages)	– Fracture	"Body of evidence can be trusted to guide practice in most situations"
Moore, 2005	United Kingdom	English	Chartered Society of Physiotherapy	WAD	Red flags unrelated to specific disease	"Evidence obtained from expert committee reports or opinions and/or clinical experience of respected authorities e.g. from the Delphi questionnaire"

NAD = neck pain and associated disorders; WAD = whiplash and associated disorders.

*No red flags provided.

Red flags

Supplementary material 4 summarizes the 114 red flags reported in the guidelines for fracture (number of guidelines = 8, red flags = 17), cancer (number of guidelines = 5, red flags = 21), spinal infection (number of guidelines = 4, red flags = 14), myelopathy (number of guidelines = 5, red flags = 15), injury to the spinal cord (number of guidelines = 1, red flags = 1), cervical artery dissection (number of guidelines = 4, red flags = 7), intracranial pathology (number of guidelines = 3, red flags = 3), inflammatory arthritis (number of guidelines = 2, red flags = 2), other systemic disease (number of guidelines = 2, red flags = 6), and unrelated to a specific condition (number of guidelines = 2, red flags = 19). Additionally, Supplementary material 5 provides the reference to external documents cited by Blanpied et al (37) for the reported red flags. Many red flags (n = 77, 67.5%) were reported only by a minority of the guidelines. As an example, only Bier et al (32) suggested that dysphagia could be a possible red flag for cancer, and only one out of four guidelines that considered spinal infection as a serious pathology mentioned "HIV positivity" as a red flag (36). Furthermore, only a few red flags (n = 7, 6.1%) were recommended by most of the guidelines: five out of seven guidelines (71.4%) mentioning red flags for fractures recommended the Canadian C-spine rule as a screening tool; both the two guidelines reporting red flags for osteoporotic fractures agreed in recommending "history of osteoporosis," "use of corticosteroids," and an "older age" as red flags; all guidelines mentioning cancer as a serious condition recommended a "history of cancer" and an "unexplained weight loss" as flags for this condition; and three out of four guidelines (75%) considering spinal infection reported the presence of fever and a "history of recent infection" as red flags for an infection (Supplementary material 4).

Agreement in red flags recommendations

Overall, there is very little agreement between guidelines on the red flags to screen for serious pathologies (Tab. 3). Notably, for all the pathologies, we found a poor agreement (Fleiss' kappa < 0), except for cancer (slight agreement with a Fleiss' kappa of 0.15) and osteoporotic fractures (perfect agreement with a Fleiss' kappa of 1).

TABLE 3 - Fleiss' kappa values

Pathology	Fleiss' kappa
Fracture	0
Osteoporotic fracture	1
Cancer	0.15
Vertebral infection	−0.14
Cervical myelopathy	−0.17
Arterial dissection	−0.28
Intracranial pathology	−0.50
Inflammatory arthritis	−0.33
Systemic disease	−1

Level of evidence on which the red flag recommendations are based

The Canadian C-spine rules were supported by Level 1 evidence, while the National Emergency X-Radiography Utilization Study (NEXUS) criteria had Level 2 evidence. Ten red flags, such as spasticity for cervical myelopathy and swelling in multiple joints for inflammatory arthritis, did not have any reference to determine their level of evidence. The remaining red flags (102, 89.5%) were based on mechanism-based reasoning, corresponding to Level 5 evidence.

Of all the red flags identified, 36 (31.6%) were supported by systematic reviews in the low back pain field or systematic reviews that did not provide direct information on the diagnostic values of specific signs and symptoms for identifying serious conditions in patients with neck pain. Notably, 10 (8.8%) red flags lacked a reference. A combination of narrative reviews, case series, and guidelines for patients with low back pain supported the remaining red flags (n = 68, 59.6%). Only the Canadian C-spine rules as a screening tool for fractures were supported by systematic reviews and observational studies providing direct information on their diagnostic accuracy (Supplementary material 4). Four guidelines (33%) described the literature used to support the reported red flags. Côté et al (1) reported that the red flags were based on the existing literature on low back pain, Sterling (31) reported that the red flags were supported by one or two primary studies with a low risk of bias, and Lemenunier et al (34) reported that the red flags were supported by studies with an intermediate level of evidence, such as low-powered randomized controlled trials, well-conducted nonrandomized comparative studies, and cohort studies. Lastly, Monticone et al (36) reported that experts' opinions supported their red flags.

Discussion

This review aimed to systematically collect the red flags recommended by the guidelines to screen for serious pathologies masquerading as neck pain. We identified 29 guidelines, 12 of which made recommendations for screening serious pathologies with a total of 114 red flags. Notably, 17 guidelines (59%) did not include screening for serious pathology recommendations, indicating that this topic is overlooked in more than half of the current guidelines. Our analysis showed that only a few red flags were consistently mentioned by the 12 guidelines that reported recommendations for screening serious pathologies, with many red flags (59.6%) reported only by a minority of the guidelines. The agreement between guidelines on the red flags for screening serious pathologies was generally poor, as measured by Fleiss' kappa. Among all the red flags, only the Canadian C-spine rules were well referenced (Level 1 evidence) and had diagnostic value as a screening tool for fractures in patients with neck pain after trauma. All the other red flags were either not referenced or suggested by mechanism-based reasoning (Level 5 evidence).

There are three main reasons for the heterogeneity in the recommended red flags. First, there is a lack of secondary studies, such as systematic reviews, specifically conducted to identify red flags for neck pain. Except for the Canadian C-spine rules, all the red flags were supported by primary

studies or systematic reviews that did not aim to summarize the diagnostic values of red flags for neck pain or were not supported at all. For example, the guideline by Côté et al (1) reported that "as there is a paucity of literature on red flags for neck pain, the list of red flags was informed by the low back pain literature." Most of the included guidelines cited Nordin et al's (41) review as a reference to support the recommended red flags. However, this review does not contain results on the red flags for which it is used as a reference. Notably, there is no strong evidence for most of the red flags for neck pain, and, therefore, the guidelines mainly relied on studies conducted in other fields and expert opinions to make their recommendations, resulting in high variability in the red flags provided in each guideline. Second, guidelines frequently presented the same red flags but offered a different cutoff or definition due to the absence of a universally agreed definition or a different healthcare system. As an example, four guidelines agreed on older age as a red flag for cancer. However, three guidelines reported "age above 60" (1,32,34), while one reported "age above 50" (36). Thus, the heterogeneity in the red flags can also be attributed to a lack of an agreed definition for almost all red flags. Third, the guidelines are customized to align with the specific health policies of the countries where they are created. For instance, the way patients can see a physiotherapist varies between countries, with some allowing direct access and others requiring a physician's referral. These disparities may have led to heterogeneity in the suggested red flags. Our results also highlight that certain serious medical conditions have received less attention in the guidelines. As an example, only three guidelines reported red flags for intracranial pathologies, and only two reported red flags for inflammatory arthritis. This lack of knowledge of clinical predictors may reflect the diagnostic delay in certain pathologies, such as axial spondyloarthritis (42).

Our review also aimed to gather data on the diagnostic accuracy of the red flags. Several guidelines have presented the diagnostic accuracy of the Canadian C-spine rule, revealing its accuracy as a screening tool for fractures with a sensitivity of almost 100%. Papic (30) also highlighted that a positive Canadian C-spine rule reduces unnecessary imaging by 44% by mentioning preliminary results of a Cochrane review (43). The Canadian C-spine rule is a decision tool that combines several red flags with a high sensitivity. Accordingly, the combination of red flags of serious lower back pathologies was found to increase their diagnostic accuracy positively (44). Notably, in our review, the diagnostic accuracy for all other red flags was not reported. Therefore, it is unclear how these signs and symptoms may affect the likelihood of a serious condition. In addition, their combination could not be investigated. This indicates that the clinical influence of these red flags remains, at best, uncertain.

Implication for practice

Clinicians are responsible for screening for underlying serious conditions when managing patients with neck pain. Of the 29 included clinical practice guidelines, only 12 recommended screening for serious non-musculoskeletal disorders. This recommendation consistently received a "strong" indication in favor whenever the strength of the recommendation was provided. However, there seems to be a lack of consensus on which red flags to use, almost all red flags are merely based on mechanism-based reasoning (Level 5 of evidence), and a report or reference to their diagnostic accuracy is often lacking. For these reasons, specific recommendations on which red flags to use cannot be provided, except for using the Canadian C-Spine rule for screening posttraumatic fractures. In fact, this rule is recommended by multiple guidelines based on systematic reviews of the literature (Level 1 evidence). Additionally, we have access to diagnostic accuracy values that support the Canadian C-Spine rules as an excellent screening tool, with sensitivity approaching 100%.

It is important to consider that the absence of clear red flags does not rule out the presence of a serious underlying condition. In addition, due to the rarity of many serious pathologies, one of the difficulties in differential diagnosis and in investigating the diagnostic accuracy of red flags is that some of these conditions may be present but clinically unmanifested (6). Although red flag testing remains the best tool to screen for serious cervical pathology, red flags when used in isolation are often uninformative (45,46). However, when combined within a broad clinical reasoning framework to determine the level of suspicion about serious pathology, they may help clinicians make the best judgment on the appropriate clinical action (e.g., further investigation or referral) in a continuous monitoring process (46,47). Within this reasoning pathway, the evidence to support red flags should be considered in the context of the patient's health profile (e.g., risk factors, medications, comorbidities, age, and gender) (47).

It is also important to consider that not all red flags masquerade severe medical conditions and that not all conditions and their stage require an emergency referral. Based on the level of concern, the decision might be: to begin a trial of therapy keeping an alert to clinical features that change unexpectedly in patients with no concerning features; begin a trial of therapy with watchful waiting in patients with few concerning features; urgent referral in patients with some concerning features – such as suspected myelopathy with long-lasting symptoms; or emergency referral in patients with some concerning features that might benefit from early specialized intervention – such as suspected myelopathy with new-onset neurological signs or symptoms. After evaluating the presence of red flags and considering the patient's clinical profile, clinicians must use their clinical reasoning to thoughtfully weigh the risks and benefits when deciding whether to refer the patient or not. For a deeper discussion on integrating red flags in clinical reasoning, we invite readers to refer to Finucane et al (13), Rushton et al (14), de Best et al (48), and Kranenburg et al (47).

Implication for future research

Future research should focus on conducting secondary studies like scoping and systematic reviews to map and/or summarize all the evidence regarding using red flags in people with neck pain. Primary studies should also be conducted to determine red flags' diagnostic accuracy and identify

additional signs and symptoms that could indicate less considered pathologies in the current guidelines, such as intracranial pathologies and inflammatory arthritis. Since serious pathologies are rare in patients with neck pain, conducting cross-sectional and prospective cohort studies is challenging. Hence, it would be better to rely on retrospective studies like case-control observational studies, even though they might have a higher risk of bias (49). Additionally, it would be helpful to study the diagnostic value in terms of discrimination and calibration of clusters of red flags, such as diagnostic predictive models (50). Finally, it would be beneficial to establish a clear and agreed definition for the most frequently reported red flags in the literature to prevent any future research wastage. As an example, the literature could define the duration and dosage of corticosteroid usage or establish a standard age threshold for identifying a person at risk of cancer. Such standardizations would ensure that the red flags are consistently and accurately reported across various studies, leading to more reliable and comparable research outcomes.

Comparison with the low back pain field

In line with our findings, it has been observed that there is high heterogeneity in the red flags presented in the guidelines for individuals with low back pain. Verhagen et al (51,52) found no agreement between guidelines on which red flags should be recommended, paucity of diagnostic accuracy, and insufficient empirical support for most red flags. However, in contrast to neck pain, a significant amount of research has recently been conducted regarding red flags for the low back pain field. For instance, in 2020, the IFOMPT released a framework to clarify the role of red flags in identifying serious pathology (47). Additionally, the Cochrane Collaboration published two systematic reviews of red flags to screen for cancer and fractures in patients with low back pain (45,46).

Strengths and limitations of the present systematic review

This study followed a rigorous methodology. Notably, we published a protocol with the study's objectives, the search strategy was comprehensive, including the consultation with experts in the neck pain field, and all the phases were performed independently by two authors. Nonetheless, this study has some limitations. First, we translated non-English and non-Italian guidelines using "DeepL Translate." DeepL is a software based on artificial intelligence that is highly precise in translating scientific papers (53). However, the translation would have probably been more accurate with the help of a human native speaker. Second, determining whether a paper should be classified as a guideline can be challenging. To decide if a document had to be considered a guideline, we employed the PEDro criteria for evidence-based clinical practice guidelines. However, even with these criteria and even though we consulted the top experts in the neck pain field asking them for additional guidelines we did not retrieve with our initial search, there is still a possibility that a guideline may have been misjudged as not being a guideline. Third, we determined the level of evidence on which the red flags recommendations are based using the 2011 Oxford Centre for Evidence-Based Medicine Levels of Evidence (27). This classification system primarily focuses on study design rather than the quality or applicability of the evidence to clinical practice. As a result, we may have overlooked important nuances, particularly in the case of red flags based on lower-level or mechanism-based reasoning. Fourth, we assessed the strength of recommendations for screening for serious pathologies by referring directly to the descriptions in the guidelines (see Tab. 2). In some cases, such as the Bier et al. guidelines (32), the description of the strength of recommendation (e.g., "Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate") was unclear, as the guideline did not provide an effect estimate (i.e., diagnostic accuracy values). This shows that some guidelines have imprecise reporting, and the strength of the recommendation is often based on a general statement.

Conclusions

Our review observed significant heterogeneity in the red flags recommended in guidelines for neck pain, with a general lack of consensus between guidelines for which red flags to endorse. Most red flags were not supported by a reference or were supported only by mechanism-based reasoning. Also, evidence for the accuracy of recommended red flags was lacking, except for the Canadian C-spine rule for fractures. Addressing the gaps in the current literature is a mainstay for future research. This includes conducting secondary studies to systematically summarize the available red flags and primary studies to determine the diagnostic accuracy of signs and symptoms that may suggest a serious medical condition. According to the current limitations of the evidence, specific recommendations on which red flags to use cannot be provided, except for using the Canadian C-Spine rule for screening posttraumatic fractures. Therefore, clinicians should use the red flags mentioned in the guidelines cautiously and integrate them into a sound clinical reasoning process.

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 contribution: DF: Conceptualization, Study selection, Data extraction, Data analysis, Writing; AC: Conceptualization, Study selection, Writing; BK: Conceptualization, Writing; FMa: Conceptualization, Data extraction, Writing; FMo: Conceptualization, Study selection, Writing.

Data availability statement: Not applicable. The data presented in this study are available as supplementary material to this article.

References

1. Côté P, Wong JJ, Sutton D, et al. Management of neck pain and associated disorders: a clinical practice guideline from the Ontario Protocol for Traffic Injury Management (OPTIMa) Collaboration. Eur Spine J. 2016;25(7):2000-2022. CrossRef PubMed
2. Safiri S, Kolahi AA, Hoy D, et al. Global, regional, and national burden of neck pain in the general population, 1990-2017:

systematic analysis of the Global Burden of Disease Study 2017. BMJ. 2020;368:m791. CrossRef PubMed
3. Cieza A, Causey K, Kamenov K, Hanson SW, Chatterji S, Vos T. Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2021;396(10267):2006-2017. CrossRef PubMed
4. Rushton AB, Verra ML, Emms A, et al. Development and validation of two clinical prediction models to inform clinical decision-making for lumbar spinal fusion surgery for degenerative disorders and rehabilitation following surgery: protocol for a prospective observational study. BMJ Open. 05 22;8(5):e021078. CrossRef
5. Mourad F, Giovannico G, Maselli F, Bonetti F, Fernández de las Peñas C, Dunning J. Basilar impression presenting as intermittent mechanical neck pain: a rare case report. BMC Musculoskelet Disord. 2016;17(1):7. CrossRef PubMed
6. Mourad F, Giudice A, Maritati G, et al. A guide to identify cervical autonomic dysfunctions (and associated conditions) in patients with musculoskeletal disorders in physical therapy practice. Braz J Phys Ther. 2023;27(2):100495. CrossRef PubMed
7. Faletra A, Bellin G, Dunning J, et al. Assessing cardiovascular parameters and risk factors in physical therapy practice: findings from a cross-sectional national survey and implication for clinical practice. BMC Musculoskelet Disord. 2022;23(1):749. CrossRef PubMed
8. Maselli F, Piano L, Cecchetto S, Storari L, Rossettini G, Mourad F. Direct access to physical therapy: should Italy move forward? Int J Environ Res Public Health. 2022;19(1):555. CrossRef PubMed
9. Maselli F, Storari L, Mourad F, Barbari V, Signorini M, Signorelli F. Headache, loss of smell, and visual disturbances: symptoms of SARS-CoV-2 infection? A case report. Phys Ther. 2023;103(4):pzad017. CrossRef PubMed
10. Mourad F, Cataldi F, Patuzzo A, et al. Craniopharyngioma in a young woman with symptoms presenting as mechanical neck pain associated with cervicogenic headache: a case report. Physiother Theory Pract. 2021;37(4):549-558. CrossRef PubMed
11. Platzer P, Hauswirth N, Jaindl M, Chatwani S, Vecsei V, Gaebler C. Delayed or missed diagnosis of cervical spine injuries. J Trauma. 2006;61(1):150-155. CrossRef PubMed
12. Sizer PS Jr, Brismée JM, Cook C. Medical screening for red flags in the diagnosis and management of musculoskeletal spine pain. Pain Pract. 2007;7(1):53-71. CrossRef PubMed
13. Finucane LM, Downie A, Mercer C, et al. International framework for red flags for potential serious spinal pathologies. J Orthop Sports Phys Ther. 2020;50(7):350-372. CrossRef PubMed
14. Rushton A, Carlesso LC, Flynn T, et al. International framework for examination of the cervical region for potential of vascular pathologies of the neck prior to musculoskeletal intervention: International IFOMPT Cervical Framework. J Orthop Sports Phys Ther. 2023;53(1):7-22. CrossRef PubMed
15. Heick J, Lazaro RT, eds. Goodman and Snyder's differential diagnosis for physical therapists: screening for referral. 7th ed. Elsevier; 2022.
16. Feller D, Giudice A, Faletra A, et al. Identifying peripheral arterial diseases or flow limitations of the lower limb: important aspects for cardiovascular screening for referral in physiotherapy. Musculoskelet Sci Pract. 2022;61:102611. CrossRef PubMed
17. Feller D, Giudice A, Maritati G, et al. Physiotherapy screening for referral of a patient with peripheral arterial disease masquerading as sciatica: a case report. Healthcare (Basel). 2023;11(11):1527. CrossRef PubMed
18. Mourad F, Lopez G, Cataldi F, et al. Assessing cranial nerves in physical therapy practice: findings from a cross-sectional survey and implication for clinical practice. Healthcare (Basel). 2021;9(10):1262. CrossRef PubMed
19. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. The BMJ. Online. (Accessed August 2024)
20. Feller D, Chiarotto A, Koes B, Maselli F, Mourad F. Red flags for potential serious pathologies masquerading as musculoskeletal neck pain: a protocol for a systematic review of clinical practice guidelines. 2023.06.20.23291691. CrossRef
21. Guzman J, Hurwitz EL, Carroll LJ, et al; Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. A new conceptual model of neck pain: linking onset, course, and care. Spine. 2008;33(4)(suppl):S14-S23. CrossRef PubMed
22. PEDro, Physiotherapy evidence database. Indexing criteria and codes. Online. (Accessed August 2024)
23. Parikh P, Santaguida P, Macdermid J, Gross A, Eshtiaghi A. Comparison of CPG's for the diagnosis, prognosis and management of non-specific neck pain: a systematic review. BMC Musculoskeletal Disorders. 2019;20(1):N.PAG-N.PAG. CrossRef
24. Corp N, Mansell G, Stynes S, et al. Evidence-based treatment recommendations for neck and low back pain across Europe: a systematic review of guidelines. Eur J Pain. 2021;25(2):275-295. CrossRef PubMed
25. 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
26. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan – a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210. CrossRef PubMed
27. Howick J, Chalmers I, Glasziou P, et al. Explanation of the 2011 Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence (Background Document). Online. (Accessed August 2024)
28. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull. 1971;76(5):378-382. CrossRef
29. Moore A, Jackson A, Jordan J, et al. Clinical guidelines for the physiotherapy management of whiplash associated disorder. Chartered Society of Physiotherapy; 2005.
30. Papic. Australian Clinical Guidelines for Health Professionals Managing People with Whiplash-Associated Disorders, 4th ed. Online (Accessed August 2024)
31. Sterling M, TRACsa: Trauma and Injury Recovery. Clinical guidelines for best practice management of acute and chronic whiplash-associated disorders – UQ eSpace. 2008. Online (Accessed August 2024)
32. Bier JD, Scholten-Peeters WGM, Staal JB, et al. Clinical practice guideline for physical therapy assessment and treatment in patients with nonspecific neck pain. Phys Ther. 2018;98(3):162-171. CrossRef PubMed
33. Whalen W, Farabaugh RJ, Hawk C, et al. Best-practice recommendations for chiropractic management of patients with neck pain. J Manipulative Physiol Ther. 2019;42(9):635-650. CrossRef PubMed
34. Haute Autorité de Santé. Label de la HAS - Évaluation du patient atteint de cervicalgie et prise de décision thérapeutique en chiropraxie. Online. (Accessed August 2024)
35. Scherer M, Chenot JF. Pragmatic and effective treatment of painful neck. Deutsche Gesellschaft für Allgemeinmedizin. MMW Fortschr Med und Familienmedizin. 2006. CrossRef
36. Monticone M, Iovine R, de Sena G, et al; Italian Society of Physical and Rehabilitation Medicine (SIMFER). The Italian Society of Physical and Rehabilitation Medicine (SIMFER) recommendations for neck pain. G Ital Med Lav Ergon. 2013;35(1):36-50. PubMed

37. Blanpied PR, Gross AR, Elliott JM, et al. Neck pain: revision 2017. J Orthop Sports Phys Ther. 2017;47(7):A1-A83. CrossRef PubMed
38. Bussières AE, Stewart G, Al-Zoubi F, et al. The treatment of neck pain-associated disorders and whiplash-associated disorders: a clinical practice guideline. J Manipulative Physiol Ther. 2016;39(8):523-564.e27. CrossRef PubMed
39. Karpinnen et al. Liikunta (ylläpito lopetettu). Duodecim. 2017. Online (Accessed August 2024)
40. Nordin M, Carragee EJ, Hogg-Johnson S, et al. Assessment of neck pain and its associated disorders: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. Spine (Phila Pa 1976). 2008;33(4 Suppl):S101-122. CrossRef PubMed
41. Nordin M, Carragee EJ, Hogg-Johnson S, et al; Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. Assessment of neck pain and its associated disorders: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. Spine. 2008;33(4)(suppl):S101-S122. CrossRef PubMed
42. Zhao SS, Pittam B, Harrison NL, Ahmed AE, Goodson NJ, Hughes DM. Diagnostic delay in axial spondyloarthritis: a systematic review and meta-analysis. Rheumatology (Oxford). 2021;60(4):1620-1628. CrossRef PubMed
43. Saragiotto BT, Maher CG, Lin CC, Verhagen AP, Goergen S, Michaleff ZA. Canadian C-spine rule and the National Emergency X-Radiography Utilization Study (NEXUS) for detecting clinically important cervical spine injury following blunt trauma. Cochrane Libr. 2018;2018(4):CD012989. CrossRef
44. Maselli F, Rossettini G, Viceconti A, Testa M. Importance of screening in physical therapy: vertebral fracture of thoracolumbar junction in a recreational runner. BMJ Case Rep. 2019;12(8):e229987. CrossRef PubMed
45. Han CS, Hancock MJ, Downie A, et al. Red flags to screen for vertebral fracture in people presenting with low back pain. Cochrane Database Syst Rev. 2023;8(8):CD014461. CrossRef PubMed
46. Henschke N, Maher CG, Ostelo RW, de Vet HC, Macaskill P, Irwig L. Red flags to screen for malignancy in patients with low-back pain. Cochrane Database Syst Rev. 2013;2013(2):CD008686. CrossRef PubMed
47. Kranenburg HA, Kerry R, Taylor A, Mourad F, Puentedura E, Hutting N. Correspondence re: de Best et al. J Physiother. 2024;70(1):78. CrossRef PubMed
48. de Best RF, Coppieters MW, van Trijffel E, et al. Risk assessment of vascular complications following manual therapy and exercise for the cervical region: diagnostic accuracy of the International Federation of Orthopaedic Manipulative Physical Therapists framework (The Go4Safe project). J Physiother. 2023;69(4):260-266. CrossRef PubMed
49. Rutjes AW, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PM. Case-control and two-gate designs in diagnostic accuracy studies. Clin Chem. 2005;51(8):1335-1341. CrossRef PubMed
50. van Smeden M, Reitsma JB, Riley RD, Collins GS, Moons KG. Clinical prediction models: diagnosis versus prognosis. J Clin Epidemiol. 2021;132:142-145. CrossRef PubMed
51. Verhagen AP, Downie A, Maher CG, Koes BW. Most red flags for malignancy in low back pain guidelines lack empirical support: a systematic review. Pain. 2017;158(10):1860-1868. CrossRef PubMed
52. Verhagen AP, Downie A, Popal N, Maher C, Koes BW. Red flags presented in current low back pain guidelines: a review. Eur Spine J. 2016;25(9):2788-2802. CrossRef PubMed
53. Takakusagi Y, Oike T, Shirai K, et al. Validation of the reliability of machine translation for a medical article from Japanese to English using DeepL translator. Cureus. 2021;13(9):e17778. CrossRef

Arch Physioter 2024; 14: 116-130
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3244
ORIGINAL RESEARCH ARTICLE

Perspectives, perceptions, and expectations of subjects with frozen shoulder: a web-based Italian survey

Fabrizio Brindisino[1], Fabiola Garzonio[1], Andrea Turolla[2,3], Arianna Andriesse[4], Fabrizio Pulina[5], Davide Cucchi[6,7], Filip Struyf[8], Davide Venturin[1]

[1]Department of Medicine and Health Science "Vincenzo Tiberio", University of Molise, Campobasso - Italy
[2]Department of Biomedical and Neuromotor Sciences, Alma Mater Studiorum, University of Bologna, Bologna - Italy
[3]Unit of Occupational Medicine, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna - Italy
[4]Medical Translation Private Practice c/o Andriesse Medical Translator, Lecce - Italy
[5]Physiotherapy Private Practice c/o Pulina Physiotherapist, Genova - Italy
[6]Department of Orthopaedics and Trauma Surgery, University of Bonn, Bonn - Germany
[7]Universitätsklinikum Bonn, Bonn - Germany
[8]Department of Rehabilitation Sciences and Physiotherapy, Faculty of Medicine and Health Sciences, University of Antwerp, Wilrijk - Belgium

ABSTRACT

Introduction: Frozen shoulder (FS) is a musculoskeletal disorder affecting the glenohumeral joint. This condition leads to disability and a worsening in quality of life. Despite its considerable impact on patients and its economic burden, research on the psychological and social implications of FS—as well as patients' perspectives and needs—is limited. This study aims to explore the perspectives, perceptions, and expectations of individuals suffering from FS, providing a comprehensive understanding of their experiences and needs.

Methods: A cross-sectional observational study was conducted following STROBE guidelines. A 59-question survey was administered to Italian individuals diagnosed with FS from April 1 to July 1, 2023.

Results: All 110 participants completed the survey. Most preferred an experienced and empathetic physiotherapist (73.64%) and relied primarily on physiotherapy (49.09%) for FS management. Additionally, 45.45% were open to a multidisciplinary approach. Subjects reported reducing night pain (71.82%) and achieving full range of motion (ROM) recovery (70.91%) as their top priorities. Participants reported a notable shift in their mood from "pre" to "post" FS, with many experiencing fear and catastrophizing thoughts and perceiving a lack of social support. Furthermore, 27.27% were open to cortisone use, while 25.45% considered electrophysical agents beneficial for managing the painful phase of FS.

Conclusion: These results underscore a strong preference for empathetic physiotherapists and the value of a multidisciplinary approach. Addressing night pain and restoring ROM are crucial priorities—emphasizing the need for tailored and shared decision-making. Additionally, these findings highlight the importance of addressing psychological well-being alongside physical symptoms.

Keywords: Adhesive capsulitis, Frozen shoulder, Irritable mood, Psychological, Rehabilitation, Stress

What is already known about this topic

- *Frozen shoulder primarily affects working-age individuals and is characterized by severe pain, restrictions in multidirectional shoulder movement, and a significant economic burden. These issues have a stressful impact on physical, personal, and social aspects of individuals' lives.*

What does the study add

- *Individuals with frozen shoulder prefer skilled and empathetic physiotherapists, indicating a potential shift in treatment paradigms. The prevalence of catastrophizing tendencies and perceived lack of social support further underscore the need to address psychological well-being as part of patient care.*

Received: August 3, 2024
Accepted: November 13, 2024
Published online: December 9, 2024

This article includes supplementary material

Corresponding author:
Fabrizio Brindisino
email: fabrizio.brindisino@unimol.it

Introduction

Frozen shoulder (FS) is a condition affecting the glenohumeral joint (1,2), with a prevalence in the general population estimated to be between 2% and 5%, and with a higher incidence in women and subjects aged 40–60 years. The exact etiology of FS remains unclear, despite extensive research into its etiopathogenesis, biological characteristics,

progression, fibrotic processes evolution, and joint changes (3,4), although several risk and predisposing factors have been identified—for example, diabetes mellitus, hypothyroidism, cardiovascular diseases, hyperlipidemia, and endotoxemia (5-8). Bilateral presentation, diabetes mellitus, thyroid disorders, and autonomic symptoms are recognized as biological factors associated with a poorer prognosis (5,6,9). In addition, psychological factors—for example, pain-related fear, depression, anxiety, catastrophizing (10), and self-perceived mental and physical health—significantly influence both subjective and objective clinical outcomes (11).

Despite its significant impact on subjects' lives (12-14), limited research has explored the psychological and social implications of FS. FS primarily affects subjects of working age, and it is characterized by severe pain, multiplanar shoulder movement restrictions, and a potentially significant economic burden. These factors alter the physical, personal, and social dimensions of those affected (12,13). FS symptoms impact various areas of life, including the work environment, and often lead to introversion and isolation (12,13). Moreover, family members are often called upon to support individuals with FS, sometimes leading to feelings of guilt over their dependency (13). Thus, the burden of FS extends beyond physical symptoms (15), affecting daily life through intense pain, disrupted sleep, perceived limitations, loss of independence, altered self-perception, and uncertainty about the condition (12,14). This may trigger emotional-cognitive alterations, influencing subjects' perception of pain and disability (16-18).

Several qualitative research studies have explored the psychological dimensions of subjects with FS (13,14,19), highlighting their subjective perspectives on rehabilitation. However, these studies often lack conclusive results on other specific issues—limited to understanding of subjects' experience and thereby hindering clinicians' ability to tailor effective management strategies and treatments.

Given the generally modest improvements seen in FS patients—particularly in terms of pain reduction and range of motion (ROM) recovery (20,21)—it is crucial to gain a deeper understanding of the psychological factors associated with FS. This includes examining patients' emotional states, the challenges they face during their condition and treatment, and their focus on achieving personal goals and returning to normalcy (13).

Therefore, the aim of this study is to explore the perspectives, perceptions, and expectations of subjects affected by FS through a cross-sectional survey.

Materials and methods

Study design

This study was designed as a cross-sectional observational study and conducted as an online-based survey. Results were reported following the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) checklist (22) and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) (23) reporting guidelines. The study protocol has been submitted and approved by the Technical Scientific Committee of the University of Molise (Italy)—Prot. n. 10/2023. All the study-related procedures were performed according to the principles of the Declaration of Helsinki (24).

Setting, sampling, and recruiting

This study enrolled Italian and Italian-speaking subjects diagnosed with FS according to Kelley's guideline (25) criteria. Specifically, participants had painful shoulder with stable or worsening reduced external rotation with the arm by the side (<50% compared with the contralateral limb) over the past month, along with at least a 25% loss of active and passive ROM in two other planes, and negative x-ray (25). All subjects presented to the authors' private practice for their first physiotherapy consultation for FS, with no prior treatments.

Recruitment was voluntary and conducted over a 3-month period (from April 1 to July 1, 2023)—similarly to previous studies (26,27) and other international surveys (28-30). The timeframe was deemed adequate based on prior surveys on similar topics. Participants received no incentives, and duplicate responses were prevented using a single-user authentication. Additionally, no modifications were allowed after the survey completion. All potential participants were invited to participate via a link generated by Google Form.

Informed consent

All potential participants received a link to an information letter containing details about the investigators' identity, aim of the survey, inclusion criteria, data protection and dissemination of results, estimated time required for survey completion, and a clear informed consent statement ("If you voluntarily agree to participate in the survey, please scan this QR code or follow the link below; if not, you can close this document"). Access to the survey was granted only upon approving this consent. This method has been used in other surveys (29,30). The information letter is detailed in Appendix 1—Information Letter.

Survey development and pre-testing

The questionnaire was designed to investigate the perspectives, perceptions, expectations, needs, beliefs, and behaviors of subjects suffering from FS. Additionally, questions regarding other important priorities for FS subjects were included, for example, pain characteristics, awareness, treatment, disability, frustration due to prolonged and debilitating shoulder pain, impact on social relationships, skepticism from others, loss of independence, altered self-perception, experiences and expectations regarding healthcare providers, struggle for normalcy, and cognitive and emotional sense of uncertainty (12-14,19), as suggested by previous qualitative studies on this topic. The goal was to gather comprehensive insights that could inform better clinical management and improve outcomes for FS patients.

A draft of this cross-sectional survey was developed by six researchers—three physiotherapists and three orthopedic surgeons—specializing in shoulder diseases. Additionally, a psychotherapist was consulted to ensure the survey ability to assess psychological themes accurately. The final version

of the survey, comprising 59 questions, was approved by the project lead and all team members.

Content validity was evaluated through a two-round pre-testing process. The initial round involved testing the questionnaire with four individuals who had previously experienced FS, followed by a second round with 10 subjects currently suffering from FS—in order to spot possible overlooked themes and clarify any confusing questions. Participants currently suffering from FS highlighted the need to address themes such as "unbearable pain," "long-lasting complaints," and "moments of discouragement." In response, the researchers and psychologist developed specific questions (Q47, Q51, Q52) to address these concerns, which were validated by the participants and incorporated into the final survey. Consensus on the survey structure was achieved through an online meeting with all involved parties.

Final version of the survey

The final version of the survey included 2 introduction questions (email address and consent to participate), 7 demographic questions, 3 questions on current levels of day and night pain and stiffness, and 47 topic-specific questions—as detailed in Appendix 2. All questions allowed for one response only. The demographic section comprised seven multiple-choice questions on sex, geographical origin, age, education, profession, time since FS onset, and number of clinicians consulted before diagnosis (Q3 to Q9). Additionally, three questions further assessed perceived day pain, night pain, and stiffness, utilizing a Numeric Rating Scale (NRS) (Q10, Q11, and Q12).

The questionnaire used a hybrid structure, combining multiple-choice (i.e., Q13-17, 23, 52, 58, and 59) and 5-point Likert scale questions (i.e., Q18-22, 24-51, 53-57). This approach is consistent with other previously published surveys (31,32) and aimed to gather detailed data on the importance that each subject attributes to various aspects of FS—with a particular focus on exploring fear and catastrophizing tendencies, in order to assess the psychological burden associated with FS.

Specifically, the technical questions covered: eventual prior diagnosis and imaging assessment (two questions, Q13 and Q14); expectation regarding health professionals and care process (four questions, Q15 to Q18); information about FS and its effects (four questions, Q19 to Q22); beliefs on treatment (four questions, Q23 to Q26); subjects' priorities (seven questions, Q27 to Q33); past (five questions, Q34 to Q38) and current (four questions, Q39 to Q42) mood; fear about their condition and the future (three questions, Q43 to Q45); expectation and catastrophizing thoughts related to pain, sense of self, struggle for normality (six questions, Q46 to Q51); social support, relationships, frustration, feeling of not being understood, loss of independence, skepticism from others (six questions; Q52 to Q57), and subjects' preferences regarding treatment (two questions, Q58 and Q59).

Data analysis

Data extraction and processing were performed using Excel—with all data stored in an encrypted, password-protected file. After survey completion, the anonymized data were forwarded for blind statistical analysis to a statistician (AT). Data analysis was performed using STATA 18 SE (33), with results reported as absolute and relative (percentage) frequencies of responses.

Results

Demographic

One-hundred and ten subjects were invited to complete the survey, and all provided their consent (100% completion rate), with no missing answers. On average, participants spent 11.22 minutes to complete the survey, as highlighted by the software.

Most participants were female (n = 72; 65.5%), aged 40 to 50 years (n = 47; 42.7%), and from northern Italy (n = 51; 46.4%). Most held a high school degree (n = 56; 50.9%) and were employed in non-physical jobs (n = 67; 60.9%).

Regarding the duration of FS, most participants had been experiencing symptoms for 5 months or longer (n = 69; 62.7%). Prior to diagnosis, most participants had consulted with one (n = 44; 40%) or two physicians (n = 30; 27.3%). Detailed demographic information is provided in Table 1 (Q3 to Q9).

Current level of day and night pain and stiffness

Participants reported a range of different day and night pain and stiffness levels. Most reported NRS pain scores between 5 and 8 during the day (n = 73; 66.4%) and between 7 and 10 at night (n = 68; 61.8%). Additionally, most rated their stiffness with an NRS score between 7 and 10 (n = 77; 70%). Detailed ratings of pain and stiffness are provided in Table 1 (Q10 to Q12).

Technical questions

Results showed that a significant number of FS patients had not undergone imaging investigations (n = 26; 23.6%). Among those who did, magnetic resonance imaging (MRI) was the most commonly prescribed, either alone (n = 24; 21.8%) or combined with x-ray (n = 17; 15.5%) (Q13). Interestingly, FS was frequently misdiagnosed as rotator cuff pathology (n = 48; 43.6%), with only 31.82% (n = 35) of cases receiving an initial correct diagnosis of FS (Q14).

Regarding interactions with physiotherapists, most participants (n = 81; 73.6%) preferred an experienced, empathetic, and caring physiotherapist (Q15). Moreover, the majority (n = 57; 51.8%) believed that physiotherapists should consider both anatomical and psychological aspects (e.g., fear, worry, anxiety, anger, lack of confidence) of FS. However, 30% of respondents (n = 33) indicated that functional outcomes should be the primary focus for physiotherapists (Q16).

Participants received several explanations about the natural history of FS from their clinicians. Some described three phases (freezing, frozen, and thawing) (n = 32; 29.1%), while others referred to two phases (pain predominant and stiffness predominant) (n = 17; 15.5%), and some did not specify any phases (n = 23; 20.9%) (Q17).

Most participants felt adequately informed ("disagree" = 37.3%; n = 41) (Q18) and supported ("disagree" = 36.4%; n = 40)

TABLE 1 - Demographic characteristics of respondents and answers for technical questions

	Question	Answers	Frequency (N = 110)	Percentage (%)
Q1	Email	Anonymized	110	100
Q2	Consent form agreement. Do you want to complete the survey?	Yes	110	100
		No	0	0
Q3	Gender	Female	72	65.45
		Male	38	34.55
Q4	Italian Region of provenience	Northern Italy	51	46.36
		Central Italy	20	18.18
		Southern Italy	39	35.45
Q5	Age	≤39 years old	2	1.82
		40-50	47	42.73
		51-60	36	32.73
		61-65	18	16.36
		≥ 66	7	6.36
Q6	Educational level	Elementary school	3	2.73
		Middle school	9	8.18
		High school	56	50.91
		University degree	42	38.18
Q7	Work type	Mainly inactive (most of the time spent in the same position)	67	60.91
		Mainly dynamic (most of the time spent performing different activities/often changing position)	43	39.09
Q8	For how long have you been experiencing frozen shoulder?	More than 5 months	69	62.73
		3 months or less than 5 months	24	21.82
		More than a month and less than 3 months	17	15.45
		Less than a month or a month	0	0
Q9	How many doctors examined you before you were diagnosed with frozen shoulder?	1	44	40.00
		2	30	27.27
		3	24	21.82
		>3	12	10.91
Q10	On a scale from 0 to 10, where 0 means no pain and 10 means the worst pain you have ever felt, how would you rate your daytime pain?	0 no pain	5	4.55
		1	6	5.45
		2	2	1.82
		3	7	6.36
		4	7	6.36
		5	14	12.73
		6	17	15.45
		7	25	22.73
		8	17	15.45
		9	7	6.36
		10 worst pain ever	3	2.73

(*Continued*)

TABLE 1 - (*Continued*)

	Question	Answers	Frequency (N = 110)	Percentage (%)
Q11	On a scale from 0 to 10, where 0 means no pain and 10 means the worst pain you have ever felt, how would you rate your nighttime pain?	0 no pain	7	6.36
		1	7	6.36
		2	2	1.82
		3	3	2.73
		4	10	9.09
		5	9	8.18
		6	4	3.64
		7	15	13.64
		8	22	20.00
		9	14	12.73
		10 worst pain ever	17	15.45
Q12	On a scale from 0 to 10, where 0 means no stiffness and 10 means the worst stiffness imaginable, how would you rate your stiffness?	0 no stiffness	0	0
		1	1	0.91
		2	2	1.82
		3	4	3.64
		4	5	4.55
		5	8	7.27
		6	13	11.82
		7	11	10.00
		8	34	30.91
		9	18	16.36
		10 worst stiffness ever	14	12.73
Q13	Which imaging tests have doctors recommended since your frozen shoulder diagnosis?	None	26	23.64
		MRI	24	21.82
		X-ray	17	15.45
		X-ray and MRI	17	15.45
		X-ray and MRI	10	9.09
		Ultrasound	9	8.18
		X-ray and ultrasound	7	6.36
		Arthro MRI	0	0
Q14	Before being diagnosed with frozen shoulder, did you receive a different diagnosis? If yes, please specify.	Yes, rotator cuff pathology (impingement, rotator cuff injury, tendinopathies)	48	43.64
		No, frozen shoulder is the first diagnosis I have received	35	31.82
		Yes, but I don't remember what	13	11.82
		Yes, periarthritis	10	9.09
		Yes, arthrosis	3	2.73
		Yes, rheumatologic issue	1	0.91
Q15	When considering physiotherapy treatment, what qualities or attributes do you prefer in a physiotherapist?	The physiotherapist should be expert, empathetic, and caring about my shoulder condition.	81	73.64
		I prefer a physiotherapist with specific expertise in managing shoulder pathology.	23	20.91
		I would like a physiotherapist who acts as a supportive partner and builds a relationship of trust.	5	4.55
		I prefer a straightforward approach where the physiotherapist focuses solely on assessing and treating the frozen shoulder.	1	0.91

	Question	Answers	Frequency (N = 110)	Percentage (%)
Q16	In your opinion, what is the most important factor for your physiotherapist to consider?	Both anatomical and psychological (fear, worry, anxiety, anger, no confidence) aspects of frozen shoulder	57	51.82
		Functional outcomes (range of movement, pain, stiffness) about frozen shoulder	33	30.00
		More anatomical aspect than psychological one	17	15.45
		More psychological aspect than anatomical one	3	2.73
Q17	How did clinicians explain the development of your frozen shoulder?	They provided a detailed explanation, including the three phases of frozen shoulder, timing, and therapies.	32	29.09
		I received a satisfactory explanation about my condition, but no mention of phases.	23	20.91
		They provided a detailed explanation, including the two phases of frozen shoulder, timing, and therapies.	17	15.45
		I did not receive a clear explanation about my condition.	13	11.82
		They gave a brief explanation, including the three phases of frozen shoulder, timing, and therapies.	10	9.09
		Different clinicians provided varying explanations.	9	8.18
		They gave a brief explanation, including the two phases of frozen shoulder, timing, and therapies.	6	5.45
Q18	How much do you agree with the following sentences:	I totally agree	8	7.27
	I was not informed about my pathology.	I agree	21	19.09
		Neither agree nor disagree	15	13.64
		I disagree	41	37.27
		I totally disagree	25	22.73
Q19	I received unhelpful explanations that did not improve my ability to manage my condition.	I totally agree	5	4.55
		I agree	22	20.00
		Neither agree nor disagree	19	17.27
		I disagree	40	36.36
		I totally disagree	24	21.82
Q20	I received explanations that increased my anxiety and worried me about the potential for recovery failure.	I totally agree	8	7.27
		I agree	18	16.36
		Neither agree nor disagree	22	20.00
		I disagree	32	29.09
		I totally disagree	30	27.27
Q21	I received explanations that helped me cope with discouragement, reassured me, encouraged me, and allowed me to manage pessimistic thoughts about my condition.	I totally agree	20	18.18
		I agree	39	35.45
		Neither agree nor disagree	26	23.64
		I disagree	23	20.91
		I totally disagree	2	1.82

(*Continued*)

TABLE 1 - (*Continued*)

	Question	Answers	Frequency (N = 110)	Percentage (%)
Q22	I received encouraging explanations that reduced my fear of movement as much as possible.	I totally agree	22	20.00
		I agree	48	43.64
		Neither agree nor disagree	16	14.55
		I disagree	22	20.00
		I totally disagree	2	1.82
Q23	Who do you believe is best equipped to manage your frozen shoulder?	Physiotherapist	54	49.09
		All aforementioned professionals when their expertise is needed	50	45.45
		Medical doctor (orthopedic, general practitioner, etc.)	3	2.73
		Medical doctor expert in pain management (algologist)	3	2.73
		Psychologist	0	0
Q24	To what extent do you agree with the following statement: If I put all my efforts into physiotherapy, I am confident I will fully recover from frozen shoulder	I totally agree	31	28.44
		I agree	44	40.37
		Neither agree nor disagree	20	18.35
		I disagree	14	12.84
		I totally disagree	0	0
Q25	These treatments are unhelpful, and I don't believe I will return to my previous condition	I totally agree	0	0
		I agree	21	19.09
		Neither agree nor disagree	20	18.18
		I disagree	42	38.18
		I totally disagree	27	24.55
Q26	If I put all my efforts into physiotherapy treatment, I will improve my situation, even if I don't achieve a complete recovery	I totally agree	4	3.64
		I agree	28	25.45
		Neither agree nor disagree	29	26.36
		I disagree	40	36.36
		I totally disagree	9	8.18
Q27	How much is important for you to achieve these results?	Not important at all	1	0.91
	Manage day-time pain.	Unimportant	3	2.73
		Neutral	10	9.09
		Important	43	39.09
		Very important	53	48.18
Q28	Manage night pain.	Not important at all	1	0.91
		Unimportant	3	2.73
		Neutral	5	4.55
		Important	22	20.00
		Very important	79	71.82

	Question	Answers	Frequency (N = 110)	Percentage (%)
Q29	Restore the full range of movement	Not important at all	0	0
		Unimportant	2	1.82
		Neutral	5	4.55
		Important	25	22.73
		Very important	78	70.91
Q30	Improve sleep quality	Not important at all	4	3.64
		Unimportant	0	0
		Neutral	5	4.55
		Important	35	31.82
		Very important	66	60.00
Q31	Improve autonomy in activities of daily living (showering, getting dressed, driving, etc.)	Not important at all	0	0
		Unimportant	1	0.91
		Neutral	9	8.18
		Important	35	31.82
		Very important	65	59.09
Q32	Improve occupational, leisure, and social activities	Not important at all	0	0
		Unimportant	2	1.82
		Neutral	18	16.36
		Important	29	26.36
		Very important	61	55.45
Q33	How much is important for you to be reassured by the physiotherapist about your clinical condition?	Not important at all	0	0
		Unimportant	1	0.91
		Neutral	18	16.36
		Important	49	44.55
		Very important	42	38.18
Q34	Which of these following sentences better describes your mood about frozen shoulder/adhesive capsulitis?	Not at all	19	17.27
	I'm feeling angry.	A little	16	14.55
		Moderately	32	29.09
		A lot	33	30.00
		Very much	10	9.09
Q35	I'm feeling sad/overcome.	Not at all	22	20.00
		A little	11	10.00
		Moderately	21	19.09
		A lot	33	30.00
		Very much	23	20.91
Q36	I'm feeling blue/low mood.	Not at all	17	15.45
		A little	15	13.64
		Moderately	24	21.82
		A lot	34	30.91
		Very much	20	18.18
Q37	I'm feeling powerless.	Not at all	23	20.91
		A little	19	17.27
		Moderately	29	26.36
		A lot	29	26.36
		Very much	10	9.09

(*Continued*)

TABLE 1 - (*Continued*)

	Question	Answers	Frequency (N = 110)	Percentage (%)
Q38	I feel like I can react.	Not at all	1	0.91
		A little	21	19.09
		Moderately	25	22.73
		A lot	45	40.91
		Very much	18	16.36
Q39	How many times, BEFORE the onset of frozen shoulder/adhesive capsulitis, did you feel: Angry	Never	6	5.45
		Rarely	43	39.09
		Sometimes	51	46.36
		Often	10	9.09
		Always	0	0
Q40	Sad/overcome	Never	12	10.91
		Rarely	36	32.73
		Sometimes	52	47.27
		Often	8	7.27
		Always	2	1.82
Q41	Blue/low mood	Never	18	16.36
		Rarely	33	30.00
		Sometimes	48	43.64
		Often	10	9.09
		Always	1	0.91
Q42	Powerless	Never	26	23.64
		Rarely	51	46.36
		Sometimes	27	24.55
		Often	5	4.55
		Always	1	0.91
Q43	How much do you agree with the following sentences? I'm afraid that moving my shoulder will make my condition worse.	I totally agree	0	0
		I agree	25	22.73
		Neither agree nor disagree	21	19.09
		I disagree	47	42.73
		I totally disagree	17	15.45
Q44	I fear that frozen shoulder will cause irreversible damage to my shoulder.	I totally agree	2	1.82
		I agree	31	28.18
		Neither agree nor disagree	31	28.18
		I disagree	33	30.00
		I totally disagree	13	11.82
Q45	I fear I will never be able to return to my previous activities.	I totally agree	9	8.18
		I agree	41	37.27
		Neither agree nor disagree	22	20.00
		I disagree	25	22.73
		I totally disagree	13	11.82
Q46	How often have you had these thoughts? I will never raise my arm as I used to do before	Never	19	17.27
		Rarely	10	9.09
		Sometimes	35	31.82
		Often	43	39.09
		Always	3	2.73

	Question	Answers	Frequency (N = 110)	Percentage (%)
Q47	Pain is terrible and it will never end.	Never	23	20.91
		Rarely	15	13.64
		Sometimes	30	27.27
		Often	40	36.36
		Always	2	1.82
Q48	All I do to heal is useless.	Never	29	26.36
		Rarely	21	19.09
		Sometimes	38	34.55
		Often	20	18.18
		Always	2	1.82
Q49	My life is ruined.	Never	49	44.55
		Rarely	25	22.73
		Sometimes	27	24.55
		Often	8	7.27
		Always	1	0.91
Q50	I'm feeling overwhelmed by this condition.	Never	25	22.73
		Rarely	20	18.18
		Sometimes	38	34.55
		Often	26	23.64
		Always	1	0.91
Q51	I'm worried because I know this is a long-term pathology.	Never	9	8.18
		Rarely	18	16.36
		Sometimes	28	25.45
		Often	41	37.27
		Always	14	12.73
Q52	If you have experienced moments of demoralization or discouragement about your situation, how did you manage them?	I asked for advice to a clinician. He/she listened to me.	42	38.18
		I let off steam with a loved one.	22	20.00
		I have never had moments of demoralization/ discouragement.	12	10.91
		I asked for advice to a clinician. He/she did not listen to me.	12	10.91
		I didn't share my discomfort with anyone.	11	10.00
		I felt abandoned and unable to manage those moments.	8	7.27
		I have taken the initiative to call a psychologist.	3	2.73
Q53	How well do you think people around you understand the seriousness of your situation? Are they supporting you in managing your pathology?	Not at all	9	8.18
		A little	46	41.82
		Moderately	29	26.36
		A lot	24	21.82
		Very much	2	1.82
Q54	How much do you agree with the following sentences?	I totally agree	6	5.45
	Other people fully understand my condition and they support me.	I agree	30	27.27
		Neither agree nor disagree	35	31.82
		I disagree	29	26.36
		I totally disagree	10	9.09

(*Continued*)

TABLE 1 - (*Continued*)

	Question	Answers	Frequency (N = 110)	Percentage (%)
Q55	Other people fully understand my condition, but they don't support me as I wish.	I totally agree	1	0.91
		I agree	39	35.45
		Neither agree nor disagree	27	24.55
		I disagree	37	33.64
		I totally disagree	6	5.45
Q56	Nobody really understands my situation	I totally agree	12	10.91
		I agree	34	30.91
		Neither agree nor disagree	22	20.00
		I disagree	32	29.09
		I totally disagree	10	9.09
Q57	I don't feel supported at all	I totally agree	8	7.27
		I agree	23	20.91
		Neither agree nor disagree	24	21.82
		I disagree	39	35.45
		I totally disagree	16	14.55
Q58	If clinicians provide you with home exercises during the rehabilitation process, which method would you prefer to remember how to perform them?	Video with a phone and text messages	52	47.27
		Booklet	42	38.18
		No one preferred	11	10.00
		Draw made by your physiotherapis	5	4.55
Q59	Which additional therapy would you prefer to combine with physiotherapy to better manage your painful phase?	Cortisone (oral or injection)	30	27.27
		Therapeutic modalities (laser, diathermy, transcutaneous electrical nerve stimulation, shockwave therapy)	28	25.45
		Massage	18	16.36
		No one preferred	18	16.36
		Non-steroidal anti-inflammatory drugs	16	14.55

Data are reported as absolute and relative frequencies.

MRI = magnetic resonance imaging; N = number; Q = questions.

in managing FS (Q19). They also reported reduced anxiety and concerns about recovery failure due to the information provided ("disagree" = 29.1%; n = 32) (Q20).

Respondents agreed that clinicians provided helpful and reassuring information to manage discouragement ("agree" = 35.5%; n = 39) (Q21)—which contributed to increased encouragement and reduced kinesiophobia ("agree" = 43.6%; n = 48) (Q22).

While most participants primarily relied on physiotherapists for FS management (n = 54; 49.1%), they were also open to collaborative approach involving physicians, algologists, and psychologists when necessary (n = 50; 45.5%) (Q23).

Most participants believed that their efforts in physiotherapy would lead to complete recovery ("agree" = 40.4%; n = 44) (Q24), rather than just partial improvement (n = 28; 25.5%) (Q26) and found treatments to be beneficial (n = 42; 38.2%) (Q25).

Most participants identified several goals as "very important" (Q27-Q33): specifically, night pain (n = 79; 71.9%), full ROM restoration (n = 78; 71%), improvement of sleep quality (n = 66; 60%), autonomy in activities of daily living (n = 65; 59.1%), participation in social and leisure activities (n = 61; 55.5%), and daytime pain (n = 53; 48.2%). Lastly, reassurance from the physiotherapist (n = 49; 44.6%) was also considered "important."

Regarding the emotional impact of FS (Q34-Q38), many respondents stated they felt "a lot" angry (n = 33; 30%), sad or overwhelmed (n = 33; 30%), experiencing a blue or low mood (n = 34; n = 30.1%), and feeling powerless (n = 29;

26.4%). However, most respondents considered themselves to be reactive (n = 45; 40.9%).

Before FS onset (Q39-Q42), most respondents reported that they "rarely" felt angry (n = 43; 39.1%), sad or overwhelmed (n = 36; 32.7%), blue or experienced low mood (n = 33; 30%), or felt powerless (n = 51; 46.4%).

The survey also investigated respondents' fear (Q43-Q45). Most subjects with FS did not fear worsening their condition through shoulder movement ("disagree" = 42.7%; n = 47), or believe that FS will irreversibly damage their shoulder ("disagree" = 30%; n = 33). However, many were concerned they might never return to their previous activities ("agree" = 37.3%; n = 41).

Regarding catastrophizing (Q46-Q51), most respondents "often" thought that they would never regain full arm elevation (n = 43; 39.1%), that the pain was unbearable, that they felt trapped in a never-ending situation (n = 40; 36.4%), and that they were worried about the prolonged duration of their condition (n = 41; 37.3%).

Most respondents "sometimes" felt that all their efforts for healing were useless (n = 38; 34.6%), that they were overwhelmed by the situation (n = 38; 34.6%); however, they "never" thought that FS had ruined their lives (n = 49; 44.6%).

Six questions investigated social support, with most respondents reporting that they sought advice from a clinician who was ready to listen to them (n = 42; 38.2%) when they felt discouraged due to FS (Q52). A total of 41.8% (n = 46) felt that those around them had "little" understanding of the seriousness of their condition and provided inadequate support (Q53). In particular, 35 (31.8%) respondents were unsure whether people fully supported and understood their condition (Q54); however, most patients felt somewhat supported (n = 39; 35.5%) (Q57), although not as much as they would have hoped (n = 39; 35.5%) (Q55). Consistent with previous questions, 30.9% (n = 34) felt that others did not truly understand their situation (Q56).

To improve therapy adherence, participants preferred being filmed with a phone and receiving text messages for home exercises (47.3%, n = 52) (Q58). Most were also open to cortisone treatment (oral or injection) (27.3%, n = 30) or electrophysical agents (25.5%, n = 28) for managing the painful phase of FS (Q59).

Discussion

This study's main contribution provides a comprehensive insight into the psychological and social dimensions of FS. A key finding is the participants' strong preference for informed, empathetic guidance from healthcare professionals, particularly clinicians and physiotherapists. The survey also identified treatment priorities, emphasizing the need to alleviate night pain and restore ROM. Additionally, the study revealed significant levels of fear and catastrophizing among FS patients, which can affect treatment outcomes. These findings highlight the necessity for a multidisciplinary approach to FS management that addresses both psychological and physical aspects.

This study included 110 participants, predominantly females aged 40 to 50 years, consistent with FS demographics reported in the literature (34); most were from northern Italy and held non-physical jobs. This demographic information provides a typical profile of FS patients.

A notable issue was the high rate of misdiagnosis, highlighting the challenge of diagnosing FS, which is often only confirmed once stiffness is well-established (35). More than 20% of patients stated that imaging investigations were not prescribed, raising concerns about adherence to diagnostic guidelines and the potential underuse of tools that could identify other conditions mimicking FS and beyond physiotherapists' expertise (36). However, the utilization rate of MRI with or without x-ray appears relatively high compared to rates in other surveys (37-39).

Our sample reported a wide range of physical impairments due to FS—including both day and night pain and stiffness—revealing a considerable heterogeneity among respondents. Many participants reported moderate to severe pain and stiffness levels, emphasizing the significant impact of FS on daily life. Interestingly, participants viewed FS as affecting both biological and psychological aspects, with many believing that physiotherapists should address both in their treatment. This supports the need for a multidimensional approach to FS management, as emphasized in previous research (10,18).

Some participants noted inconsistencies in how clinicians explained the progression of FS, aligning with prior research (40) and trends in primary studies (34). Such inconsistencies may cause confusion, undermine trust, and affect treatment adherence (12,14). Despite this, most respondents were satisfied with the information provided, finding it helpful and supportive in managing their FS.

The respondents' perspectives on their condition revealed a mix of positive and challenging aspects. Many believed in the effectiveness of physiotherapy and anticipated a full recovery. However, they also reported persistent fear and concerns about long-term impact of FS on their daily activities—along with catastrophizing thoughts about pain and their future. These findings align with other qualitative studies (12,19), highlighting the ongoing struggle for normalcy experienced by those living with FS (12).

To the best of the authors' knowledge, this study was the first to ask participants to rate the importance of different priorities in subjects suffering from FS. Night pain, ROM restoration, and psychological reassurance emerged as key priorities for the participants, providing new evidence on this topic and suggesting treatments that align with patients' expectations. While a previous study identified pain relief as a main priority (14), our findings partially agree with this result but highlight additional concerns. Given that priorities may vary among individuals, clinicians should routinely investigate these preferences to enhance shared decision-making and patient engagement (41).

No consensus was found in the literature regarding whether psychological aspects could trigger FS or vice versa (42-45). This survey aimed to clarify this by examining the emotional experiences of FS patients. Participants reported a shift in their mood, with increased anger, sadness, and powerlessness after developing FS. These findings suggest that psychological distress is more a consequence than a cause of

FS, supporting previous research (12,13). The insidious onset, sleep deprivation, and significant pain and disability associated with FS—particularly in middle-aged individuals—may lead to the development or to the amplification of psychological symptoms. The prolonged recovery and limitations in using the affected arm may significantly impact daily life, work, and hobbies, contributing to psychological distress. FS significantly affects mental health, leading to feelings of anger, overwhelming, and powerlessness compared to before the onset of this condition—although some participants reported to be "reactive." Additionally, concerns about the underlying cause of pain could exacerbate catastrophizing and pain-related beliefs, further diminishing arm function and increasing disability (17,18).

Jones et al (14) reported that subjects often experience delays in receiving a definitive diagnosis of FS, a finding consistent with the experiences reported by participants in this survey, who consulted with multiple clinicians before receiving a diagnosis.

Delays or misdiagnoses, particularly during the initial phase, when pain and disability are most severe and quality of life is compromised, can worsen anxiety and depression. Such delays contribute to altered pain beliefs, unanswered questions, and uncertainty—potentially fostering distrust and leaving patients in a state of ongoing psychological fragility.

Similarly, social support emerged as a critical aspect in our sample. Many valued the understanding provided by clinicians but reported dissatisfaction with support from their social circles. This aligns with previous research, which describes FS as a hidden disability, leading to frustration over others' inability to recognize its seriousness (19). Additionally, family members often bear the burden of providing support, leading to feelings of guilt in the patients.

FS also contributes to disrupted routines, causing a sense of isolation and uncanniness, described as a form of anxiety and fear stemming from the realization of one's solitary existence (19).

The preference for technology-based support—like videos and text messages for home exercises—suggests that such tools could enhance treatment adherence.

Limitation of this study

This survey represents one of the most extensive studies providing valuable insights into the psychological and social dimensions experienced by subjects suffering from FS—highlighting the need for a comprehensive, patient-centered approach, as recommended in prior studies (18).

However, there are limitations. Self-reported data may be affected by participants' current emotional states, potentially leading to inaccuracies. While efforts were made to ensure content validity through literature review, expert consultation, and pilot testing, this survey's psychometric properties were not extensively validated. Additionally, participants' perspectives and needs may evolve over time. Social desirability bias may influence responses, especially regarding interactions with healthcare professionals. Additionally, the sample may not fully represent the Italian population, as participants were recruited from specific areas of Italy and from a single private physiotherapy practice, limiting the generalizability of the findings.

Implications for clinical practice

This survey highlights areas for improvement in physiotherapy practice. In terms of diagnosis, clinicians should carefully consider clinical presentation and disease progression, along with appropriate use of imaging, to reduce misdiagnosis of FS. Additionally, physiotherapists should enhance their therapeutic skills, as well as their abilities in communication, empathy, and patient care, as patients expect clinicians to be engaged and empathetic. Moreover, establishing a strong therapeutic relationship that aligns with patients' preferences is a key element of patient-centered care and has been positively linked to better clinical outcomes in physiotherapy (46). Notably, patients experience significant mood changes before and after FS—including increased feelings of anger, sadness, and low mood. Patients also emphasize the importance of feeling heard and reassured when expressing their fears. In light of these emotional changes and specific needs, adopting a biopsychosocial approach to patient care is essential. Additionally, catastrophizing thoughts and a lack of social support were noted—aligning with findings from previous studies (12-14,18,19). Physiotherapists should therefore be prepared to address these factors, as psychological interventions led by physiotherapists have shown promise in improving health outcomes (47). However, this approach may require additional training or collaboration within multidisciplinary teams to ensure the most effective and comprehensive care.

Clinicians should incorporate a holistic assessment of all patient domains from the initial evaluation and monitor these aspects consistently throughout rehabilitation, moving beyond the traditional biomechanical focus. From the patients' perspective, treatment priorities emphasize the need for physiotherapists to focus on relieving night pain and improving ROM, to better align with patient goals and increase satisfaction (41).

Future research

Given that this study included only Italian-speaking participants, future research should consider administering the survey in multiple languages to capture cultural nuances that might affect responses. While this study offers a snapshot of the participants' experiences, a longitudinal design would provide insights into how challenges and perceptions evolve over time. Such surveys could also help tailor rehabilitation approaches at the beginning and throughout therapy.

Incorporating more robust and validated measures could further enhance the reliability of the findings. Addressing these considerations in future research will deepen our understanding of FS and improve care and outcomes for affected individuals.

Conclusion

This survey highlights the complex challenges faced by individuals with FS, underlining the need for a comprehensive

rehabilitation approach that addresses both physical and psychological aspects. Participants showed a clear preference for informed and empathetic physiotherapists and recognized the benefits of a multidisciplinary approach, suggesting a potential shift in treatment paradigms. Night pain and ROM recovery emerged as critical priorities, emphasizing the need for personalized interventions. The high levels of fear, catastrophizing tendencies, and perceived lack of social support highlight the need to address psychological well-being alongside physical symptoms—especially given the significant mood changes observed from "pre" to "post" FS. This study encourages future research on integrated, patient-centered approaches to FS management.

Acknowledgments

The authors extend their gratitude to all subjects who participated in the survey.

Disclosures

Conflict of interest: The authors report there are no competing interests to declare.

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

Authors' contributions: Concept/idea: FB, FG, DV; Research design: FB, FG, DV; Writing: FB, FG, AT, FS, DC, AA; Data collection: AT, FB, FP; Data analysis: AT, FB; Project management: FB, AT, DV, FP; Consultation (including review of manuscript before submitting): FB, FG, AT, FS, DC, FP, DV, AA; Final approval of the Manuscript: FB, FG, AT, FS, DC, FP, DV, AA

Ethical approval and consent to participate: Ethical approval was obtained from the Ethics Committee of the University of Molise (Italy) with the registration number Prot. n. 10/2023. All procedures were conducted as part of this.

Data availability statement: All data generated or analyzed during this study are included in this published article (and its supplementary information files).

References

1. Lewis J. Frozen shoulder contracture syndrome – Aetiology, diagnosis and management. Man Ther. 2015;20(1):2-9. CrossRef PubMed
2. Cucchi D, Di Giacomo G, Compagnoni R, et al. A high level of scientific evidence is available to guide treatment of primary shoulder stiffness: the SIAGASCOT consensus. Knee Surg Sports Traumatol Arthrosc. 2024;32(1):37-46. CrossRef PubMed
3. Kraal T, Lübbers J, van den Bekerom MPJ, et al. The puzzling pathophysiology of frozen shoulders – a scoping review. J Exp Orthop. 2020;7(1):91. CrossRef PubMed
4. Fields BKK, Skalski MR, Patel DB, et al. Adhesive capsulitis: review of imaging findings, pathophysiology, clinical presentation, and treatment options. Skeletal Radiol. 2019;48(8): 1171-1184. CrossRef PubMed
5. Ando A, Sugaya H, Hagiwara Y, et al. Identification of prognostic factors for the nonoperative treatment of stiff shoulder. Int Orthop. 2013;37(5):859-864. CrossRef PubMed
6. Eljabu W, Klinger HM, von Knoch M. Prognostic factors and therapeutic options for treatment of frozen shoulder: a systematic review. Arch Orthop Trauma Surg. 2016;136(1):1-7. CrossRef PubMed
7. Chuang SH, Chen YP, Huang SW, Kuo YJ. Association between adhesive capsulitis and thyroid disease: a meta-analysis. J Shoulder Elbow Surg. 2023;32(6):1314-1322. CrossRef PubMed
8. André P, Laugerette F, Féart C. Metabolic endotoxemia: a potential underlying mechanism of the relationship between dietary fat intake and risk for cognitive impairments in humans? Nutrients. 2019;11(8):11. CrossRef PubMed
9. Mertens MG, Meeus M, Noten S, et al. Understanding the clinical profile of patients with frozen shoulder: a longitudinal multicentre observational study. BMJ Open. 2022;12(11):e056563. CrossRef PubMed
10. Brindisino F, Minnucci S, Sergi G, Lorusso M, Struyf F, Innocenti T. Does the psychological profile of a patient with frozen shoulder predict future outcome? A systematic review. Physiother Res Int. 2024;29:e2056. CrossRef PubMed
11. De Cristofaro L, Brindisino F, Venturin D, Andriesse A, Pellicciari L, Poser A. Prognostic factors of nonsurgical intervention outcomes for patients with frozen shoulder: a retrospective study. Disabil Rehabil. 2024;1-8. CrossRef PubMed
12. King WV, Hebron C. Frozen shoulder: living with uncertainty and being in "no-man's land". Physiother Theory Pract. 2023; 39(5):979-993. CrossRef PubMed
13. Lyne SA, Goldblatt FM, Shanahan EM. Living with a frozen shoulder – a phenomenological inquiry. BMC Musculoskelet Disord. 2022;23(1):318. CrossRef PubMed
14. Jones S, Hanchard N, Hamilton S, Rangan A. A qualitative study of patients' perceptions and priorities when living with primary frozen shoulder. BMJ Open. 2013;3(9):e003452. CrossRef PubMed
15. Page MJ, O'Connor DA, Malek M, et al. Patients' experience of shoulder disorders: a systematic review of qualitative studies for the OMERACT Shoulder Core Domain Set. Rheumatology (Oxford, England). 2019. CrossRef PubMed
16. De Baets L, Matheve T, Dierickx C, Bijnens E, Jans D, Timmermans A. Are clinical outcomes of frozen shoulder linked to pain, structural factors or pain-related cognitions? An explorative cohort study. Musculoskelet Sci Pract. 2020;50:102270. CrossRef PubMed
17. De Baets L, Matheve T, Traxler J, Vlaeyen J, Timmermans A. Pain-related beliefs are associated with arm function in persons with frozen shoulder. Shoulder Elbow. 2020;12(6):432-440. CrossRef PubMed
18. Brindisino F, Silvestri E, Gallo C, et al. Depression and anxiety are associated with worse subjective and functional baseline scores in patients with frozen shoulder contracture syndrome: a systematic review. Arthrosc Sports Med Rehabil. 2022;4(3):e1219-e1234. CrossRef PubMed
19. Bilsborough Smith C, Nadesan K, Cairns M, Chester R, Lewis J. Living with frozen shoulder. 'Here are the risks. I want the injection'. An interpretative phenomenological analysis. Musculoskelet Sci Pract. 2023;65:102755. CrossRef PubMed
20. Maund E, Craig D, Suekarran S, et al. Management of frozen shoulder: a systematic review and cost-effectiveness analysis. Health Technol Assess. 2012;16(11):1-264. CrossRef PubMed
21. Shaffer B, Tibone JE, Kerlan RK. Frozen shoulder. A long-term follow-up. J Bone Joint Surg Am. 1992;74(5):738-746. CrossRef PubMed
22. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. 2004;6(3):e34. CrossRef PubMed
23. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Int J Surg (London, England). 2014;12:1495-1499. CrossRef PubMed

24. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013;310(20):2191-2194. CrossRef PubMed
25. Kelley MJ, Shaffer MA, Kuhn JE, et al. Shoulder pain and mobility deficits: adhesive capsulitis. J Orthop Sports Phys Ther. 2013;43(5):A1-A31. CrossRef PubMed
26. Brindisino F, De Santis A, Rossettini G, et al. Post-surgery rehabilitation following rotator cuff repair. A survey of current (2020) Italian clinical practice. Disabil Rehabil. 2022;44(17):4689-4699. CrossRef PubMed
27. Brindisino F, Ristori D, Lorusso M, et al. Subacromial impingement syndrome: a survey of Italian physiotherapists and orthopaedics on diagnostic strategies and management modalities. Arch Physiother. 2020;10(1):16. CrossRef PubMed
28. Bury J, Littlewood C. Rotator cuff disorders: a survey of current (2016) UK physiotherapy practice. Shoulder Elbow. 2018; 10(1):52-61. CrossRef PubMed
29. Pieters L, Voogt L, Bury J, et al. Rotator CUFF disorders: a survey of current physiotherapy practice in Belgium and the Netherlands. Musculoskelet Sci Pract. 2019;43:45-51. CrossRef PubMed
30. Littlewood C, Mazuquin B, Moffatt M, Bateman M. Rehabilitation following rotator cuff repair: a survey of current practice (2020). Musculoskelet Care. 2021;19(2):165-171. CrossRef PubMed
31. Faletra A, Bellin G, Dunning J, et al. Assessing cardiovascular parameters and risk factors in physical therapy practice: findings from a cross-sectional national survey and implication for clinical practice. BMC Musculoskelet Disord. 2022;23(1):749. CrossRef PubMed
32. Mourad F, Yousif MS, Maselli F, et al. Knowledge, beliefs, and attitudes of spinal manipulation: a cross-sectional survey of Italian physiotherapists. Chiropr Man Therap. 2022;30(1):38. CrossRef PubMed
33. StataCorp. Stata Statistical Software: Release 18. StataCorp LLC; 2023. Online (Accessed August 2024)
34. Mertens MG, Meeus M, Verborgt O, et al. Exploration of the clinical course of frozen shoulder: a longitudinal multicenter prospective study of functional impairments. Braz J Phys Ther. 2023;27(4):100539. CrossRef PubMed
35. Millar NL, Meakins A, Struyf F, et al. Frozen shoulder. Nat Rev Dis Primers. 2022;8(1):59. CrossRef PubMed
36. Ewald A. Adhesive capsulitis: a review. Am Fam Physician. 2011;83(4):417-422. PubMed
37. Pandey V, Chidambaram R, Modi A, et al. Trends in practice among shoulder specialists in the management of frozen shoulder: a consensus survey. Orthop J Sports Med. 2022;10(10):23259671221118834. CrossRef PubMed
38. Cho CH, Lee YH, Kim DH, Lim YJ, Baek CS, Kim DH. Definition, diagnosis, treatment, and prognosis of frozen shoulder: a consensus survey of shoulder specialists. Clin Orthop Surg. 2020;12(1):60-67. CrossRef PubMed
39. Kraal T, Visser C, Sierevelt IN, Beimers L. How to treat a frozen shoulder? A survey among shoulder specialists in the Netherlands and Belgium. Acta Orthop Belg. 2016;82:78-84. PubMed
40. Abrassart S, Kolo F, Piotton S, et al. 'Frozen shoulder' is ill-defined. How can it be described better? EFORT Open Rev. 2020;5(5):273-279. CrossRef PubMed
41. Hoffmann TC, Lewis J, Maher CG. Shared decision making should be an integral part of physiotherapy practice. Physiotherapy. 2020;107:43-49. CrossRef PubMed
42. Chiaramonte R, Bonfiglio M, Chisari S. A significant relationship between personality traits and adhesive capsulitis. Rev Assoc Med Bras. 2020;66(2):166-173. CrossRef PubMed
43. Fleming A, Dodman S, Beer TC, Crown S. Personality in frozen shoulder. Ann Rheum Dis. 1975;35(5):456-457. CrossRef PubMed
44. Wright V, Haq AM. Periarthritis of the shoulder. I. Aetiological considerations with particular reference to personality factors. Ann Rheum Dis. 1976;35(3):213-219. CrossRef PubMed
45. Debeer P, Franssens F, Roosen I, Dankaerts W, Claes L. Frozen shoulder and the Big Five personality traits. J Shoulder Elbow Surg. 2014;23(2):221-226. CrossRef PubMed
46. Miciak M, Mayan M, Brown C, Joyce AS, Gross DP. The necessary conditions of engagement for the therapeutic relationship in physiotherapy: an interpretive description study. Arch Physiother. 2018;8(1):3. CrossRef PubMed
47. Silva Guerrero AV, Maujean A, Campbell L, Sterling M. A systematic review and meta-analysis of the effectiveness of psychological interventions delivered by physiotherapists on pain, disability and psychological outcomes in musculoskeletal pain conditions. Clin J Pain. 2018;34(9):838-857. CrossRef PubMed

Arch Physioter 2024; 14: 131-137
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3191
ORIGINAL RESEARCH ARTICLE

Does the modified shuttle test exhibit a ceiling effect in healthy and cystic fibrosis children and adolescents?

Marta Amor Barbosa[1], Fernanda Salazar Pérez[1], Fernanda Maria Vendrusculo[2], Maria Amélia Bagatini[2], Evanirso da Silva Aquino[3], Márcio Vinícius Fagundes Donadio[1,2]

[1]Department of Physiotherapy, Faculty of Medicine and Health Sciences, Universitat Internacional de Catalunya (UIC), Barcelona - Spain
[2]Laboratory of Pediatric Physical Activity, Infant Center, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre (RS) - Brazil
[3]Pontifícia Universidade Católica de Minas Gerais (PUCMG), Campus Betim and Hospital Infantil João Paulo II, FHEMIG - Brazil

ABSTRACT

Introduction: The modified shuttle test-15 (MST-15) is a valid alternative for assessing exercise capacity when a cardiopulmonary exercise testing is not feasible. This study aims to describe the percentage of healthy and cystic fibrosis (CF) children and adolescents reaching the MST-15 ceiling. Additionally, it examines associations between MST-15 distance and demographic, anthropometric, and lung function data.
Methods: This retrospective cross-sectional study involved 286 healthy volunteers (11.5 ± 3.3 years) and 70 CF patients (11.9 ± 4.4 years). Data on age, gender, weight, height, body mass index, lung function, and MST-15 were collected. The ceiling effect was determined by the absolute and relative number of participants reaching the 15th level. Univariate linear regression and correlation analyses were conducted to explore associations with MST-15 distance.
Results: A ceiling effect for the MST-15 was found in 19 healthy participants (6.6%) and 1 CF patient (1.4%). The ceiling effect was correlated with age (r = 0.777 for healthy; r = 0.538 for CF), with no cases under 10 years and reaching 25% in healthy participants aged 17-19. Regression analysis showed significant associations between age and MST-15 distance in healthy participants (β = 53.6) and CF patients (β = 32.1). Additionally, sex was significantly associated with MST-15 distance in healthy participants (β = 107.0), and FEV_1 with MST-15 distance in CF patients (β = 31.0).
Conclusions: The ceiling effect on the MST-15 is age-dependent, with no occurrences observed in children under 10 years and a gradual increase in incidence as participants age.

Keywords: Adolescent, Child, Cystic fibrosis, Exercise test, Exercise tolerance

What is already known?

- *The modified shuttle test-15 (MST-15) is a valid alternative for assessing exercise capacity when a cardiopulmonary exercise test is not feasible or recommended, but it may be submaximal for some children and adolescents.*

What does the study add?

- *The ceiling effect on the MST-15 is age-dependent, with no occurrences in children under 10 and gradually increasing with age. The MST-15 effectively evaluates functional exercise capacity in most children and adolescents with CF.*

Introduction

Cystic fibrosis (CF) is a hereditary, autosomal recessive disease caused by a mutation in the gene responsible for encoding the cystic fibrosis transmembrane conductance regulator protein (CFTR). The absence or dysfunction of this protein results in a multisystemic disease, leading to obstruction in secretory glands (1). As the disease progresses, exercise capacity declines due to a multifactorial etiology (2), including chronic infection (3), lung function impairment (4), peripheral muscle dysfunction (5), and ventilation impairment (6). This reduced exercise capacity is associated with a higher risk of hospitalization for pulmonary exacerbations (7) and a poorer prognosis leading to increased mortality (8).

International guidelines strongly recommend including exercise capacity assessments as a standard component in regular evaluations for individuals with CF (9). The gold standard for assessing exercise capacity is the cardiopulmonary exercise test (CPET), despite its logistical limitations such as testing time, space, cost, and the need for specialized

Received: July 3, 2024
Accepted: November 12, 2024
Published online: December 18, 2024

Corresponding author:
Marta Amor Barbosa
email: mamor@uic.es

expertise (10). Cost-effective alternatives encompass field tests like the six-minute walking test (6MWT) and the shuttle tests (9). While the 6MWT has been thoroughly studied, the physiological responses vary when compared to a CPET, mainly because the former involves submaximal effort, especially in patients with low-severity disease. In contrast, the shuttle tests stand out as validated incremental protocols designed to assess maximal exercise capacity (11).

The original shuttle test consisted of 12 levels (ST-12) of progressively increasing speed, covering a 10-m distance until the subject could no longer maintain the pace or experienced fatigue, dyspnea, or signs of alertness (12). The ST-12 was developed for adults with chronic obstructive pulmonary disease and later validated for application in children with CF (13). Given that the ST-12 did not elicit a maximal response in patients with minimal disability, it was modified by adding three levels and allowing patients to run (14–16). Despite this expansion, the 15-level modified shuttle test (MST-15) may still be a submaximal test for some patients, limiting its applicability for those with high exercise capacity. Subsequently, a novel version, the 25-level modified shuttle test (MST-25), has been developed (17,18).

To date, only two reports have identified a ceiling effect for the MST-15. The first, a conference abstract, reported a ceiling effect in 6% of adult CF patients and 31% of healthy peers (17). The second found that 40% of children and adolescents with CF reached the 15th level of the MST-15 (18). However, these findings are based on either the authors' clinical experience (17) or small sample sizes (18), which may introduce bias into the results. Our hypothesis is that the impact of the ceiling effect of the MST-15 in children and adolescents is relatively low, even within a cohort encompassing both CF and healthy individuals. Therefore, this study aims to provide a description of the percentage of healthy and CF children and adolescents reaching the ceiling of the MST-15. New evidence on the topic may help to guide healthcare professionals in choosing adequate tests to evaluate exercise capacity. Secondarily, we examine the associations between MST-15 distance and demographic, anthropometric, and lung function data.

Methods

A retrospective cross-sectional study was conducted on previously collected data from the Pediatric Physical Activity Laboratory. The sample of healthy participants from our database comprised volunteer children and adolescents of both genders from public and private schools in Southern Brazil, who met the following inclusion criteria: (i) age between 6 and 19 years, (ii) body mass index (BMI) between the 5th and 85th percentile (19), (iii) absence of chronic or acute neurological, orthopedic, respiratory, cardiac, or endocrine diseases contraindicating participation in school physical education, and (iv) forced expiratory volume in 1 second/forced vital capacity (FEV_1/FVC) above the lower limit of the reference values (20). Patients of both sexes diagnosed with CF were recruited from databases of two specialized CF centers, as a part of clinical assessments or annual reviews, based on the following inclusion criteria: (i) confirmed CF diagnosis through genetic testing, (ii) age between 6 and 19 years, and (iii) regular follow-up at two specialized CF centers. Patients were excluded if they exhibited signs of hemodynamic instability (altered blood pressure or heart rate (HR) responses), exacerbation of respiratory symptoms within the last 30 days (increased cough and expectorated sputum, changes in secretion, a decline in lung function by more than 10%), and/or osteoarticular or musculoskeletal changes that could interfere with the performance of the test. None of the participants were under use of CFTR modulator therapy.

Demographic (age and gender) and anthropometric (weight, height, BMI) data were collected from all participants. Weight measurement was conducted in an upright position using a calibrated digital scale (G-tech, Glass 1 FW, Rio de Janeiro, Brazil, or 110 F, Welmy, São Paulo, Brazil) with a precision of 100 g. Height was obtained with a portable stadiometer (AlturaExata, TBW, São Paulo, Brazil), accurate to 1 mm, and participants were barefoot. BMI was calculated and expressed in kg/m². For participants with CF, additional clinical and genetic data were presented. *Pseudomonas aeruginosa* chronic colonization was defined as the persistent presence of the bacterium in the oropharyngeal swab or sputum culture for at least 6 months or in three consecutive collections (21).

Lung function assessments were performed through spirometry using a KOKO spirometer (Louisville, CO, USA), following the criteria established by the American Thoracic Society-European Respiratory Society (ATS/ERS) (22). The parameters under evaluation encompassed FVC, FEV_1, the FEV_1/FVC ratio, and forced expiratory flow between 25% and 75% of FVC ($FEF_{25–75\%}$). Data were presented in absolute values, z-scores, and as a percentage of predicted values derived from an international reference equation (23).

The MST-15 was conducted following the guidelines outlined by the ATS/ERS (24). Participants were advised to abstain from vigorous physical activity, avoid consuming caffeine within the 24 hours leading up to the test, ensure a minimum of 8 hours of sleep the night before, and have a light meal. A 10-m circuit was demarcated by two cones positioned at a 9-m distance, leaving half a meter on each side to account for the change of direction. Participants navigated the circuit at a pace signaled by an acoustic cue. The initial speed was set at 0.5 m/s and increased by 0.17 m/s at each level until the completion of the test. Initial instructions and encouragement during the test adhered to standardized protocols. The test was finished under the following circumstances: the participant failed to reach the cone on two consecutive occasions, the participant did not sustain the required speed, the participant reached the maximum distance of 1,500 m, or the participant exhibited peripheral oxygen saturation (SpO_2) below 85%. Prior to the commencement and upon completion of the test, measurements were taken for HR, SpO_2 (Nonin®, Minneapolis, USA), blood pressure (BIC sphygmomanometer, Itupeva, Brazil), and the modified Borg scale score for dyspnea (rated from 1 to 10). HR and SpO_2 were continuously monitored throughout the test. The distance covered in the test was calculated by counting the total number of laps and expressed in meters and as a percentage of predicted values (25).

The normality of the data was assessed using the Kolmogorov-Smirnov test, and the results were reported by presenting the mean and standard deviation or median and interquartile range, according to distribution. Categorical variables were expressed as absolute and relative frequencies. Comparisons between groups were performed using the Student's t-test or the Mann-Whitney U-test. A linear regression analysis was conducted to assess the association of sex, age, BMI (z-score), and FEV_1 (z-score) with the distance covered in the MST-15. For each variable, β-coefficients were presented along with their 95% confidence intervals and the associated p-values. Additionally, a correlation analysis was performed to examine the relationship between age and distance covered in the MST-15 for both the CF group and the healthy group, with p-values reported. Pearson correlation coefficients were categorized as low (r = 0.0 to 0.3), moderate (r = 0.3 to 0.7), or high (r = 0.7 to 1.0). Scatter plots illustrating the linear relationship between age and distance covered in the MST-15 were also presented for both groups. In all cases, significance was considered when $p \leq 0.05$.

Results

The healthy participant group included 286 volunteers, comprising 140 males with a mean age of 11.5 ± 3.3 years. Spirometry parameters indicated normal lung function, as expected (FEV_1 103.0 ± 16.8% of predicted and FVC 98.9 ± 16.6% of predicted). The CF sample comprised 70 individuals, with a mean age of 11.9 ± 4.4 years, consisting of 47 males. Genotyping revealed that over half of the participants had a heterozygote F508del mutation (57.1%), and 12 were chronically colonized by *P. aeruginosa*. Spirometry assessments indicated a mild decline in lung function, with a mean FEV_1 of 77.2 ± 23.9% of the predicted, while FVC was 83.8 ± 20.9% of the predicted. The characteristics of the study sample are shown in Table I.

In the MST-15, healthy participants achieved a mean level of 12.7 ± 1.5, covering a distance of 1065.9 ± 232.9 m (110.2 ± 23.5% of the predicted), while participants with CF attained an average level of 10.5 ± 2.1 and covered a distance of 760.5 ± 261.3 m (71.2 ± 21.2% of the predicted). All parameters related to the evaluation of exercise capacity using the MST-15 are presented in Table II.

A ceiling effect for the MST-15 was observed in 19 (6.6%) participants in the healthy group and in only 1 (1.4 %) patient in the CF group. For healthy participants, we conducted a subgroup analysis by age, categorizing individuals into the following age ranges: 6–10, 11–13, 14–16, and 17–19 years. Our findings showed that the ceiling effect was age-related, with higher occurrences in older participants, while no cases were observed in those under 10 years. Figure 1 illustrates the percentage of healthy participants exhibiting the ceiling effect on the MST-15 across each age group. This analysis was not performed in the CF group, as only one participant, an 18-year-old male, reached the 15th level. In addition, the correlation analysis between age and distance covered in the MST-15 (Fig. 2) revealed a high correlation for the healthy group (r = 0.777; p < 0.001) and a moderate correlation for the CF group (r = 0.538; p < 0.001).

TABLE 1 - Characteristics of the study sample

Variables	Healthy (n = 286)	CF (n = 70)	p-Value
Demographics			
Age (years)	11.5 ± 3.3	11.9 ± 4.4	0.446
Male, n (%)	140 (49.0)	47 (67.1)	**0.006**
Anthropometrics			
Weight (kg)	43.3 ± 16.5	39.3 ± 15.3	0.061
Height (cm)	146.3 ± 17.8	145.6 ± 17.4	0.755
BMI (kg/m^2)	19.5 ± 4.2	17.0 ± 3.4	**0.002**
BMI (z-score)	0.3 ± 1.1	−0.5 ± 1.2	**<0.001**
Genotyping			
F508del homozygous, n (%)	–	19 (27.1)	–
F508del heterozygous, n (%)	–	40 (57.1)	–
Other mutations, n (%)	–	11 (15.7)	–
PA chronic airway colonization, n (%)	–	12 (17.1)	–
Lung function			
FEV_1 (L)	2.5 ± 0.9	1.9 ± 0.9	**<0.001**
FEV_1 (% predicted)	103.0 ± 16.8	77.2 ± 23.9	**<0.001**
FEV_1 (z-score)	0.3 ± 1.4	−1.9 ± 2.0	**<0.001**
FVC (L)	2.8 ± 1.0	2.4 ± 1.1	**0.002**
FVC (% predicted)	98.9 ± 16.6	83.8 ± 20.9	**<0.001**
FVC (z-score)	−0.1 ± 1.4	−1.42 ± 1.8	**<0.001**
FEV_1/FVC (absolute)	0.9 ± 0.1	0.8 ± 0.1	**<0.001**
FEV_1/FVC (% predicted)	103.5 ± 8.4	90.6 ± 12.8	**<0.001**
FEV_1/FVC (z-score)	0.7 ± 1.5	−1.0 ± 1.5	**<0.001**
$FEF_{25-75\%}$ (L/min)	3.1 ± 1.0	1.9 ± 1.2	**<0.001**
$FEF_{25-75\%}$ (% predicted)	105.6 ± 22.8	65.6 ± 37.8	**<0.001**
$FEF_{25-75\%}$ (z-score)	0.9 ± 11.5	−1.9 ± 2.1	**0.044**

BMI = body mass index; CF = cystic fibrosis; $FEF_{25-75\%}$ = forced expiratory flow between 25% and 75% of vital capacity; FEV_1 = forced expiratory volume in 1 second; FVC = forced vital capacity; PA = ***Pseudomonas aeruginosa***. Significant values ($p \leq 0.05$) are highlighted in bold.

Table III displays the association of the MST-15 distance with demographic, anthropometric, and lung function variables, using a univariate linear regression analysis. The analyses revealed significant associations between age and MST-15 distance for both healthy (β-coefficient = 53.6, p < 0.001) and CF groups (β-coefficient = 32.1, p < 0.001). Additionally, we identified significant associations between sex and MST-15 distance for the healthy group (β-coefficient = 107.0, p < 0.001) and between FEV_1 and MST-15 distance for the CF group (β-coefficient = 31.0, p = 0.05).

TABLE 2 - Evaluation of the exercise capacity using the modified shuttle test

Variables	Healthy (n = 286)	CF (n = 70)	p-Value
Rest			
HR (bpm)	94 ± 11	93 ± 15	0.686
SpO_2 (%)	98 ± 1	97 ± 2	**<0.001**
SBP (mmHg)	110 ± 12	98 ± 14	**<0.001**
DBP (mmHg)	74 ± 8	62 ± 9	**<0.001**
Borg for dyspnea	0 (0-0)	0 (0-0)	**<0.001**
Peak exercise			
HR (bpm)	202 ± 8	178 ± 17	**<0.001**
SpO_2 (%)	98 ± 2	93 ± 4	**<0.001**
SBP (mmHg)	137 ± 22	118 ± 23	**<0.001**
DBP (mmHg)	82 ± 12	68 ± 9	**<0.001**
Borg for dyspnea	6 (4-10)	5 (5-8)	0.123
MST-15 level	13 ± 2	11 ± 2	**<0.001**
MST-15 distance (m)	1069 ± 226	761 ± 261	**<0.001**
MST-15 (% of predicted)	110.5 ± 22.8	71.2 ± 21.2	**<0.001**

CF = cystic fibrosis; DBP = diastolic blood pressure; HR = heart rate; MST-15 = 15-level modified shuttle test; SBP = systolic blood pressure; SpO_2 = peripheral oxygen saturation.
Significant values (p ≤ 0.05) are highlighted in bold.

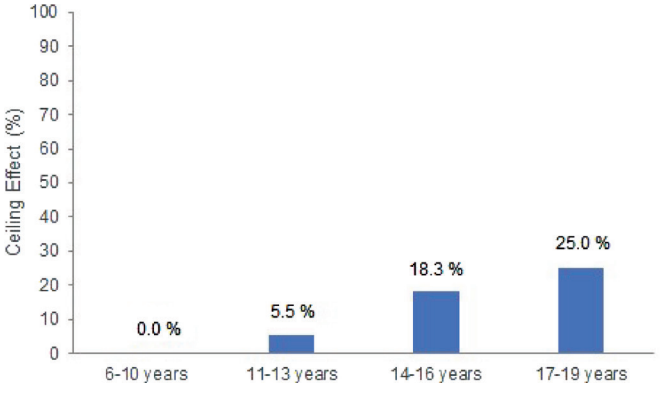


FIGURE 1 - Ceiling effect in healthy participants on the 15-level modified shuttle test (MST-15) by age group.

Discussion

The results obtained in the present study reveal that 1.4% of CF participants and 6.6% of healthy children and adolescents reached the end of the 15th level of the MST-15. Our findings further indicate that the ceiling effect is age-dependent, with no evidence of a ceiling effect observed in participants younger than 10 years and a maximum of 25% in healthy participants aged 17 to 19.

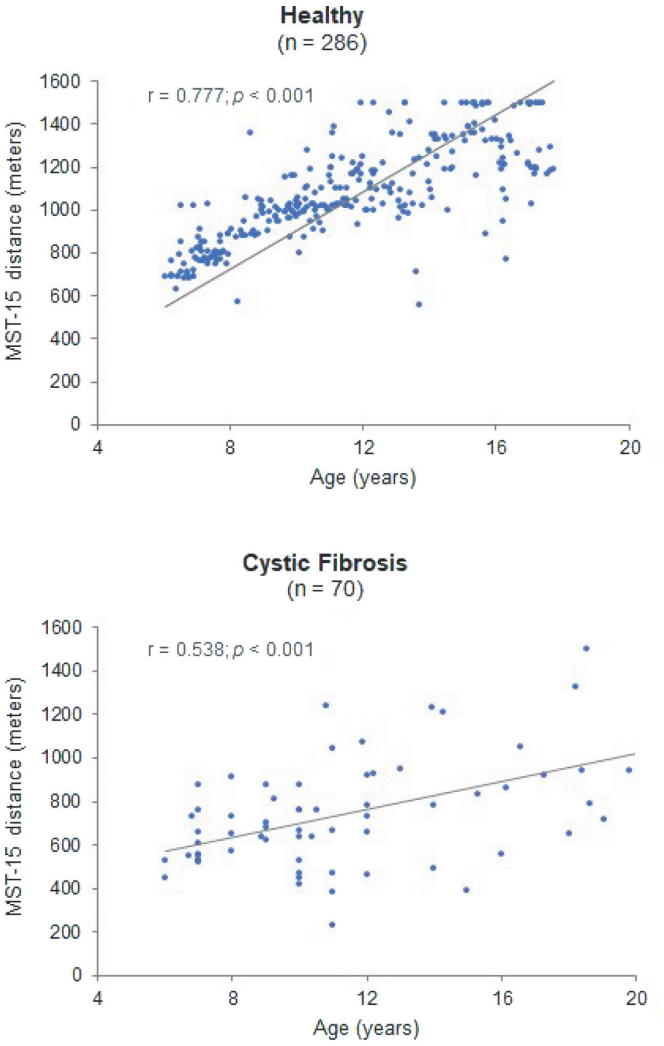


FIGURE 2 - Scatter plot of age and distance covered in the 15-level modified shuttle test (MST-15) for healthy and cystic fibrosis groups.

The ERS recently updated its guidance and standard operating procedures for functional exercise testing in CF. The ERS now recommends using the MST-15 for patients with moderate to severe lung disease, and the MST-25 for those with mild-to-moderate lung disease, considering the possibility that some individuals may complete the MST-15 without reaching their exercise capacity limits (9). Our findings suggest that, even among healthy participants with preserved lung function, the 15 levels of the MST-15 are sufficient to assess exercise capacity in children under 10 years. However, consideration should be given to extending the MST-15 as children transition into adolescence, as a ceiling effect (up to 25%) may occur. It is important to note that our cohort did not include healthy or CF adults. Therefore, the incidence of the ceiling effect in this population warrants further investigation to enhance understanding of its impact on the MST-15 within a broader demographic.

Our results contrast with a recent study reporting that 40% of children and adolescents with CF reached the

TABLE 3 - Association of the modified shuttle test distance with anthropometric and lung function variables

	Healthy (n = 286)			CF (n = 70)		
Variables	**β-coeff**	**95% CI**	**p-Value**	**β-coeff**	**95% CI**	**p-Value**
Sex (male)	100.7	49.4 to 152.0	<0.001*	93.9	−37.8 to 225.6	0.16
Age (years)	53.4	48.4 to 58.5	<0.001*	32.1	20.0 to 44.3	<0.001*
BMI (z-score)	9.3	−15.0 to 33.6	0.45	13.8	−37.6 to 65.1	0.59
FEV_1 (z-score)	−18.5	−36.6 to −0.3	0.55	31.0	−0.4 to 62.4	0.05*

β-coeff = β-coefficient; 95% CI = 95% confidence interval; BMI = body mass index; CF = cystic fibrosis; FEV_1 = forced expiratory volume in 1 second.
*Significant association ($p \leq 0.05$) using a linear regression model.

maximum level in the MST-15, indicating a high ceiling effect (18). While no clear or definitive factors explain these discrepancies, some considerations may shed light on the observed variations in the proportion of children and adolescents demonstrating a ceiling effect across our study and that previous report (18). Differences in sample characteristics could play a role, as a smaller sample size (n = 20 vs. n = 70) might not capture the broader clinical variations seen in typical practice. Additionally, it is possible that a higher exercise capacity (mean MST-25 distance of 1408 ± 298 m) and the introduction of new CFTR modulator therapy in 55% of the sample could have contributed to the observed high exercise performance (18).

Various demographic and clinical variables demonstrated associations with the distance covered in the MST-15. In both healthy children and those with CF, age was significantly associated with the distance covered during the test, with a greater effect observed in the healthy group compared to the CF group. As expected, the simultaneous growth in height and muscle mass with age leads to a corresponding increase in exercise capacity. These findings align with previous reports indicating that age is a variable contributing to the prediction of exercise capacity in pediatrics (26-28). Our findings also demonstrate that, in healthy subjects, age is a key determinant of exercise performance, exhibiting a more consistent linear relationship with the distance covered. In contrast, the weaker association between age and distance in the CF participants is likely due to disease-related factors that may overshadow the linear relationship between age and exercise capacity. In the healthy group, sex was also associated with the MST-15 distance, which can be explained by boys generally having greater muscle mass compared to girls (29). However, sex was not associated with MST-15 performance in the CF group, contrary to a previous study that indicated sex plays a role in predicting exercise capacity in patients with CF (28). The fact that sex was not associated with the distance achieved in the CF group, along with the lower associations between age and MST-15 distance compared to the healthy group, could be attributed to the greater impact of lung function on exercise capacity in children and adolescents with CF. This aligns with previous findings that report a moderate to strong correlation between FEV_1 and MST-15 (14, 15, 30), while other studies show a moderate correlation only among patients with FEV_1 < 67% of predicted (27, 28). It has been suggested that exercise intolerance in patients with CF may depend more on FEV_1 in those with severe lung disease, while in those with mild-to-moderate lung disease, the limitation may be more related to the magnitude of ventilatory responses to exercise (30). Surprisingly, the β-coefficient for BMI (z-score) did not demonstrate significant association with the distance covered in the MST-15 for either the healthy or the CF group. This result contradicts a study of prediction equations for MST-15 distance in children and adolescents, where sex, age, and BMI accounted for 48% of the variability in MST-15 performance (25).

In this rapidly changing landscape for CF, characterized by the increasing implementation of CFTR modulator therapy, it may be needed to reevaluate the applicability of field tests, including the MST-15. The use of CFTR modulator therapy could potentially lead to a long-term improvement in functional capacity and could significantly increase exercise capacity beyond the expected (31, 32). It has been observed that many patients with CF, as part of their disease management, engage in a significant amount of physical activity, often surpassing even their healthy peers (33). High levels of physical activity contribute to preserving exercise capacity in children and adolescents with CF (34), despite exhibiting lower ventilatory efficiency and reduced respiratory reserve during exertion (6). In individuals with mild-to-moderate lung disease who followed exercise recommendations before starting CFTR modulator therapy, lung function and ventilatory responses to exercise are likely to normalize. This could lead to a significant enhancement in exercise capacity, prompting the utilization of MST-25 whenever CPET is not available. Therefore, the MST-25 could be considered an alternative to the MST-15 for managing potential unpredictable increases in exercise capacity, especially in adolescents. Nonetheless, it is also important to highlight that CFTR modulator therapy is not available worldwide, as well as there are patients ineligible or considered as non-responders for the use of the therapy (35).

Field tests will continue to have an important value in assessing functional exercise performance, especially when CPET is not available or recommended. MST-15 has demonstrated excellent reliability (14-16); a good correlation between distance covered and peak oxygen consumption (VO_2peak), HR, and breathlessness assessed in CPET (36); a minimal clinically important difference (MCID) of 97 m (37);

responsiveness to antibiotic therapy (15); and predictability of risk for hospitalization (38). Although MST-25 has also shown good reliability and correlation with VO_2peak in children with CF (18), further research is required to define its psychometric properties, responsiveness, and MCID before standardizing its usage, as well as larger sample size studies allowing results to be generalizable to the CF population. While other 20-m shuttle tests are available (13, 39), they are typically regarded as impractical in most clinical settings.

Limitations

A limitation of this study is the lack of measurements regarding the physical activity levels of the participants, hampering our ability to assess the influence of this variable on MST-15 performance. Additionally, none of the participants with CF had initiated CFTR modulator therapy, preventing us from examining the treatment's influence on the ceiling effect of the MST-15.

Conclusion

The results obtained in the present study indicate that the ceiling effect on the MST-15 is age-dependent, with no occurrences observed in children under 10 years and a gradual increase as participants age. Furthermore, our findings indicate that MST-15 efficiently evaluates functional exercise capacity in most children and adolescents with CF. Until further evidence to support the use of alternative field tests are available, the MST-15 remains a valid option for assessing functional exercise capacity, particularly in younger children, when the gold standard is unavailable or not recommended.

Disclosures

Conflict of interest: The authors state that there are no conflicts of interest to declare.

Financial support: This work was supported by Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior – Brazil (CAPES) – Finance Code 001, and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Authors' contributions: MAB: Conceptualization, Methodology, Formal Analysis, Writing – original draft; FSP: Methodology, Writing – original draft; FMV: Investigation, Writing – review and editing; MAB: Investigation, Writing – review and editing; EdSA: Investigation, Resources, Writing – review and editing; MVFD: Formal Analysis, Resources, Supervision, Funding Acquisition, Writing – review and editing.

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 confidentiality reasons.

Ethics approval statement: The performance of tests was approved by the Research Ethics Committees of all involved institutions, with the assigned protocol numbers 52583416.5.0000.5336, 54142716.8.3001.5119, and 29351620.6.2006.5336.

Patient consent statement: All legal guardians and participants aged at least 18 signed an informed consent form, while children and adolescents under the age of 18 signed an assent form.

References

1. Cutting GR. Cystic fibrosis genetics: from molecular understanding to clinical application. Nat Rev Genet. 2015;16(1):45-56. CrossRef PubMed
2. Almajed A, Lands LC. The evolution of exercise capacity and its limiting factors in cystic fibrosis. Paediatr Respir Rev. 2012;13(4):195-199. CrossRef PubMed
3. van de Weert-van Leeuwen PB, Slieker MG, Hulzebos HJ, Kruitwagen CLJJ, van der Ent CK, Arets HGM. Chronic infection and inflammation affect exercise capacity in cystic fibrosis. Eur Respir J. 2012;39(4):893-898. CrossRef PubMed
4. Pianosi P, LeBlanc J, Almudevar A. Relationship between FEV_1 and peak oxygen uptake in children with cystic fibrosis. Pediatr Pulmonol. 2005;40(4):324-329. CrossRef PubMed
5. Gruet M, Troosters T, Verges S. Peripheral muscle abnormalities in cystic fibrosis: Etiology, clinical implications and response to therapeutic interventions. J Cyst Fibros. 2017;16(5):538-552. CrossRef PubMed
6. Donadio MVF, Barbosa MA, Vendrusculo FM, et al. Mechanisms of ventilatory limitation to maximum exercise in children and adolescents with chronic airway diseases. Pediatr Pulmonol. 2023;58(11):3293-3302. CrossRef PubMed
7. Pérez M, Groeneveld IF, Santana-Sosa E, et al. Aerobic fitness is associated with lower risk of hospitalization in children with cystic fibrosis. Pediatr Pulmonol. 2014;49(7):641-649. CrossRef PubMed
8. Vendrusculo FM, Heinzmann-Filho JP, da Silva JS, Perez Ruiz M, Donadio MVF. Peak oxygen uptake and mortality in cystic fibrosis: systematic review and meta-analysis. Respir Care. 2019;64(1):91-98. CrossRef PubMed
9. Saynor ZL, Gruet M, McNarry MA, et al; European Cystic Fibrosis Society Exercise Working Group. Guidance and standard operating procedures for functional exercise testing in cystic fibrosis. Eur Respir Rev. 2023;32(169):230029. CrossRef PubMed
10. Tomlinson OW, Trott J, Williams CA, et al. Challenges in implementing routine cardiopulmonary exercise testing in cystic fibrosis clinical practice: a single-centre review. SN Compr Clin Med 2020 23. 2020;2(3):327-331. CrossRef
11. Saglam M, Vardar-Yagli N, Savci S, et al. Six minute walk test versus incremental shuttle walk test in cystic fibrosis. Pediatr Int. 2016;58(9):887-893. CrossRef PubMed
12. Singh SJ, Morgan MDL, Scott S, Walters D, Hardman AE. Development of a shuttle walking test of disability in patients with chronic airways obstruction. Thorax. 1992;47(12):1019-1024. CrossRef PubMed
13. Selvadurai HC, Cooper PJ, Meyers N, et al. Validation of shuttle tests in children with cystic fibrosis. Pediatr Pulmonol. 2003;35(2):133-138. CrossRef PubMed
14. Bradley J, Howard J, Wallace E, Elborn S. Validity of a modified shuttle test in adult cystic fibrosis. Thorax. 1999;54(5):437-439. CrossRef PubMed
15. Bradley J, Howard J, Wallace E, Elborn S. Reliability, repeatability, and sensitivity of the modified shuttle test in adult cystic fibrosis. Chest. 2000;117(6):1666-1671. CrossRef PubMed
16. Rogers D, Smith P, John N, Oliver W, Doull I. Validity of a modified shuttle walk test as a measure of exercise tolerance in paediatric CF patients. J Cyst Fibros. 2002;1(suppl 1):22. Online
17. Elkins M, Dentice R, Bye P. Validation of the MST-25: an extension of the modified shuttle test (MST). J Cyst Fibros. 2009;8:S70. CrossRef

18. Corda J, Holland A, Berry CD, Westrupp N, Cox NS. Validation of the 25 level modified shuttle test in children with cystic fibrosis. Pediatr Pulmonol. 2023;58(8):2240-2248. CrossRef PubMed
19. ABRAN – Associação Brasileira de Nutrologia. Calculadoras, IMC Infantil. Online. (Accessed July 2024).
20. Quanjer PH, Stanojevic S, Cole TJ, et al; ERS Global Lung Function Initiative. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. Eur Respir J. 2012;40(6):1324-1343. CrossRef PubMed
21. Lee TWR, Brownlee KG, Conway SP, Denton M, Littlewood JM. Evaluation of a new definition for chronic *Pseudomonas aeruginosa* infection in cystic fibrosis patients. J Cyst Fibros. 2003;2(1):29-34. CrossRef PubMed
22. Graham BL, Steenbruggen I, Miller MR, et al. Standardization of Spirometry 2019 Update. An official American Thoracic Society and European Respiratory Society Technical Statement. Am J Respir Crit Care Med. 2019;200(8):e70-e88. CrossRef PubMed
23. Quanjer PH, Hall GL, Stanojevic S, Cole TJ, Stocks J; Global Lungs Initiative. Age- and height-based prediction bias in spirometry reference equations. Eur Respir J. 2012;40(1):190-197. CrossRef PubMed
24. Holland AE, Spruit MA, Troosters T, et al. An official European Respiratory Society/American Thoracic Society technical standard: field walking tests in chronic respiratory disease. Eur Respir J. 2014;44(6):1428-1446. CrossRef PubMed
25. Lanza FC, Zagatto EP, Silva JC, et al. Reference equation for the incremental shuttle walk test in children and adolescents. J Pediatr. 2015;167(5):1057-1061. CrossRef PubMed
26. Amedro P, Matecki S, Pereira dos Santos T, et al. Reference values of cardiopulmonary exercise test parameters in the contemporary paediatric population. Sport Med – Open. 2023;9(1):68. CrossRef
27. Vendrusculo FM, Aquino ES, Campos NE, et al. Determinants of exercise capacity assessed with the modified shuttle test in individuals with cystic fibrosis. Respir Care. 2020;65(5):643-649. CrossRef PubMed
28. Doeleman WR, Takken T, Bronsveld I, Hulzebos EHJ. Relationship between lung function and Modified Shuttle Test performance in adult patients with cystic fibrosis: a cross-sectional, retrospective study. Physiotherapy. 2016;102(2):184-188. CrossRef PubMed
29. Rogol AD. Sex steroids, growth hormone, leptin and the pubertal growth spurt. Endocr Dev. 2010;17:77-85. CrossRef PubMed
30. Pastré J, Prévotat A, Tardif C, Langlois C, Duhamel A, Wallaert B. Determinants of exercise capacity in cystic fibrosis patients with mild-to-moderate lung disease. BMC Pulm Med. 2014;14(1):74. CrossRef PubMed
31. Rysgaard UK, Pedersen CL, Jensen JH, et al. Change in exercise capacity measured by Cardio-pulmonary Exercise Testing (CPET) in Danish people with cystic fibrosis after initiation of treatment with Lumacaftor/Ivacaftor and Tezacaftor/Ivacaftor. J Cyst Fibros. 2022;21(5):844-849. CrossRef PubMed
32. Wark PAB, Cookson K, Thiruchelvam T, Brannan J, Dorahy DJ. Lumacaftor/Ivacaftor improves exercise tolerance in patients with cystic fibrosis and severe airflow obstruction. BMC Pulm Med. 2019;19(1):106. CrossRef PubMed
33. Valencia-Peris A, Lizandra J, Moya-Mata I, Gómez-Gonzalvo F, Castillo-Corullón S, Escribano A. Comparison of physical activity and sedentary behaviour between schoolchildren with cystic fibrosis and healthy controls: a gender analysis. Int J Environ Res Public Health. 2021;18(10):5375. CrossRef PubMed
34. Radtke T, Smith S, Nevitt SJ, Hebestreit H, Kriemler S. Physical activity and exercise training in cystic fibrosis. Cochrane Database Syst Rev. 2022;8(8):CD002768. PubMed
35. Ramananda Y, Naren AP, Arora K. Functional consequences of CFTR interactions in cystic fibrosis. Int J Mol Sci. 2024;25(6):3384. CrossRef PubMed
36. Vendrusculo FM, Heinzmann-Filho JP, Campos NE, Gheller MF, de Almeida IS, Donadio MVF. Prediction of peak oxygen uptake using the modified shuttle test in children and adolescents with cystic fibrosis. Pediatr Pulmonol. 2019;54(4):386-392. CrossRef PubMed
37. Del Corral T, Gómez Sánchez Á, López-de-Uralde-Villanueva I. Test-retest reliability, minimal detectable change and minimal clinically important differences in modified shuttle walk test in children and adolescents with cystic fibrosis. J Cyst Fibros. 2020;19(3):442-448. CrossRef PubMed
38. Donadio MVF, Vendrusculo FM, Campos NE, et al. The modified shuttle test as a predictor of risk for hospitalization in youths with cystic fibrosis: a two-year follow-up study. J Cyst Fibros. 2021;20(4):648-654. CrossRef PubMed
39. Léger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 metre shuttle run test for aerobic fitness. J Sports Sci. 1988;6(2):93-101. CrossRef PubMed

Arch Physioter 2024; 14: 138-154
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3282
REVIEW

Adverse events related to physiotherapy practice: a scoping review

Yiran Wang[1], Narelle Dalwood[1], Melanie Farlie[1,2], Annemarie L. Lee[1,3]

[1]Department of Physiotherapy, School of Primary and Allied Health Care, Monash University, Frankston, Victoria - Australia
[2]Monash Centre for Scholarship in Health Professions Education (MCSHE), Faculty of Medicine, Nursing and Health Sciences, Monash University, Clayton, Victoria - Australia
[3]Institute for Breathing and Sleep, Austin Health, Heidelberg, Victoria - Australia

ABSTRACT

Introduction: While adverse events related to physiotherapy are possible, the type of adverse event and the area of physiotherapy practice in which they occur are not well understood. The purpose of this scoping review was to establish adverse events related to physiotherapy practice and understand the nature of these events and the circumstances in which they occurred.
Methods: Relevant literature from January 2014 to February 2024 was gathered from five electronic databases. Studies reporting adverse events within any physiotherapy practice (intervention or assessment) were eligible. Two reviewers independently assessed title and abstract, and full texts. Findings were synthesised by clinical streams.
Results: A total of 58 studies met the inclusion criteria. Common adverse events described in musculoskeletal physiotherapy involving manual therapy, exercise and electrotherapy were increased pain and stiffness. Cardiorespiratory physiotherapy interventions involving early mobilisation, exercise and airway clearance therapy reported desaturation and haemodynamic instability. Neurological physiotherapy studies reported falls and fatigue during gait and balance training and exercise. Oncology and aged care interventions involving exercise, balance training and lymphoedema management reported increased pain and muscle strain while studies including pelvic floor muscle training reported the adverse event of vaginal discomfort.
Conclusion: This review identified adverse events occurring during physiotherapy interventions or assessment procedures. Increased monitoring and proactive safety measures may be necessary to ensure patient safety during these treatments.

Keywords: Adverse events, Patient safety, Physiotherapy, Scoping review

What is already known about this topic:

- *Adverse events within clinical trials and observational studies across physiotherapy practice have been documented. However, the adverse events and the nature of physiotherapy practice during which these events have occurred are not well understood.*

What this study adds:

- *This review summarises adverse events attributable to physiotherapy across a range of clinical practice areas. The awareness of these events highlights the importance of clinicians adapting and monitoring their practice to maximise patient safety.*

Introduction

Patient safety is important in all healthcare settings. However, preventable adverse events do occur and are a significant challenge globally. Recent data generated by the World Health Organization indicate that within hospital settings, unsafe healthcare practices contribute to 134 million adverse events annually (1,2). An adverse event is defined as an incident in which harm resulted to a person receiving healthcare (3). A serious adverse event is defined as any undesirable experience occurring during intervention which requires further medical attention or extended hospital stays (4). The healthcare treatment may involve a procedure, medication or a specific intervention, and the type of adverse events can have a wide range of severity, including injury, specific signs or symptoms, psychological harm or trauma (5). Adverse events may be unintended or a side effect of treatment, with the potential for either no harm, rapid recovery, the possibility of an extended hospital stay or significant clinical deterioration requiring additional medical attention (6).

Received: August 26, 2024
Accepted: November 26, 2024
Published online: December 23, 2024

This article includes supplementary material

Corresponding author:
Annemarie Lee
email: Annemarie.Lee@monash.edu

While patient safety is the foundation of healthcare practice, procedures or interventions associated with unintended harm can arise as a result of medical, nursing and allied health management, including physiotherapy (7). Although physiotherapy procedures and treatments are commonly acknowledged for their safety, particularly when implemented by qualified professionals (8), adverse events do occur (4). Musculoskeletal physiotherapy has been associated with a range of risks, including those related to manual therapy (e.g. increases in pain beyond baseline following treatment) and electrotherapy (4,9). In the field of cardiac surgery, 20% of physiotherapy interventions within intensive care were associated with adverse events, with 10% of these linked to negative outcomes (10). Similarly, early mobilisation in critically ill patients has been linked to haemodynamic and respiratory changes which have raised potential safety concerns (11). In physiotherapy management of patients with Parkinson's disease, common adverse events reported include falls, pain or discomfort and hypotension (12).

Early awareness and recognition of potential risks are vital for the safety of physiotherapy interventions and are key strategies to reduce the occurrence of adverse events (13). The benefit of this practice extends to those of graduate-entry physiotherapy students to facilitate the reduction in risk of harm in clinical situations (14). Instruction for students regarding potential adverse events across a variety of clinical fields of physiotherapy may be instrumental in developing risk management skills and contribute to enhanced patient safety, a core professional expectation of clinicians (15). Given the diverse field of physiotherapy practice, it is of clinical value to identify adverse events directly attributed to physiotherapy interventions and the nature of those adverse events. The collation of this information can be used to improve the awareness of clinicians and physiotherapy students of potential adverse events related to clinical practice. This may further promote the implementation of mitigating strategies to minimise or eliminate their occurrence (16). Furthermore, the problems with adequate systems to capture adverse events and the poor quality of the data that are collected is well documented (17). Learning from the adverse events that are reported can assist us identify priorities for investing in improved systems or supplementary data collection for this process. This scoping review is a step towards achieving this.

A scoping review was chosen to enable a broad inclusion of studies regardless of study design or quality (18). The objective of this study was to: (i) establish the adverse events related to physiotherapy practice; and (ii) describe the nature of these events.

Methods

The scoping review methodology involved documentation of a structured protocol including: eligibility criteria, information sources, selection of sources of evidence, data charting process and synthesis of results. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) was used to guide the reporting (Appendix 1) (19).

Eligibility criteria

The review included studies that met the following criteria: (i) peer-reviewed literature; (ii) studies published from January 2014 to February 2024; (iii) studies conducted in physiotherapy settings; (iv) reporting of adverse events or serious adverse events (as defined within each study) during or after the physiotherapy intervention or assessment procedure and was deemed by the study to be attributable to the physiotherapist-prescribed intervention or assessment procedure; and (v) studies published in English. Exclusion criteria were: studies involving adverse events in physiotherapy students rather than patients.

Information sources

The process of identifying potentially relevant studies included searching the following bibliographic databases from January 2014 to February 2024: Scopus; Physiotherapy Evidence Database (PEDro); Excerpta Medica Database (Embase); Medical Literature Analysis and Retrieval System Online (MEDLINE); Psychological Information Database (PsycINFO); and Cumulative Index to Nursing and Allied Health Literature (CINAHL). The time frame of 2014 to 2024 was selected in order to focus on studies published in the last 10 years due to their relevance to recent or current physiotherapy practice. As physiotherapy practice continues to evolve, it is likely that some practice procedures and technology from more than 10 years ago are not consistently equivalent to current practice. For some practices, continuous quality improvement in healthcare would enable a proportion of adverse events to be minimised by controls in place. Adverse events which occurred more than 10 years during physiotherapy practice, if persistent, are likely to be captured in a search limited to the last 10 years. The search strategies were developed and further refined through team discussion. The search strategy applied in each of the databases is outlined in Appendix 2.

Selection of sources of evidence

Covidence (Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia, 2024) and Endnote 20 were used for data screening and extraction. The retrieved references were imported into Endnote 20, where duplicates were identified and removed. These references were subsequently imported into Covidence for the screening process. Two reviewers (YW and ALL) independently conducted the initial title and abstract screening. These reviewers evaluated the eligibility criteria for each study; any disagreement was resolved through discussion. Following the title and abstract screening, the two reviewers independently evaluated the full text of the selected studies to make final decisions.

Data charting process

Data extraction was performed using Google Sheets and Microsoft Excel. Extracted data included study design, patient condition, number of participants, demographics (age and sex), nature of the physiotherapy interventions, the location of the physiotherapy assessment or intervention (e.g. hospital setting – inpatients or outpatients; primary care – private

practice, community, home) and related adverse or serious adverse event(s) reported during physiotherapy interventions. One reviewer (YW) extracted the relevant data from the selected studies and the team evaluated the data systematically. Any disagreements arising during this process were resolved via team discussion.

Synthesis of results

Findings were synthesised in tables, grouped by physiotherapy clinical stream: musculoskeletal; cardiorespiratory; neurological; oncology, aged care and pelvic health (20).

Results

Selection of sources of evidence

After 261 duplicates were removed, 1,104 studies were identified from searches of electronic databases and review article references. Based on the title and abstract, 1,015 studies were excluded, with 89 full-text articles to be retrieved and assessed for eligibility. Of these, 31 were excluded for reasons outlined in Figure 1. The remaining 58 studies were considered eligible for this scoping review.

Characteristics of sources of evidence

The included studies were published from 2014 to 2024. Study designs included 36 randomised controlled trials; one randomised cross-over study; 20 non-randomised interventional, cross-sectional, cohort or feasibility studies; and one case study. Sample sizes ranged from one to 1,208 participants. Across the studies, the age of participants ranged from a median of eight months to a mean of 80 years. Regarding areas of physiotherapy practice, 22 studies reported on adverse events on musculoskeletal physiotherapy (21-42), 20 in cardiorespiratory (43-62), eight in neurological physiotherapy (63-70), five in oncology (71-75), one in aged care (76) and two in pelvic health (77,78).

Synthesis of findings

In musculoskeletal physiotherapy, all adverse events were reported by patient participants. The majority of included studies involved pain management for chronic conditions such as osteoarthritis (21,25-28), impingement syndrome (22), meniscal injury (29), buttock pain (31), tendinopathy or foot fractures (32,33), post-orthopaedic procedure rehabilitation following total hip or total knee arthroplasty or hip surgery (23,24,30), neck or back pain (34-36), shoulder conditions (37-40) or non-specific regions (41,42) (Tab. 1 and Supplementary Table 1). For management of a range of lower limb conditions (including following surgery) or back or neck pain, interventions provided included manual therapy, heat therapy, strength exercises, functional training, gait retraining and education. Findings indicated that commonly reported adverse events during or after these interventions were increased pain, stiffness, swelling, headaches and worsening of symptoms (Tab. 1 and Supplementary Table 1) with other serious or non-serious adverse events consisting of musculoskeletal tissue disorders and falls. For those with shoulder conditions receiving exercise, pain, muscle soreness and tendon complications were apparent.

In cardiorespiratory physiotherapy, the patient conditions, physiotherapy interventions and types of adverse events are outlined in Tab. 2 and Supplementary Table 2. The majority of adverse events were collated from patient

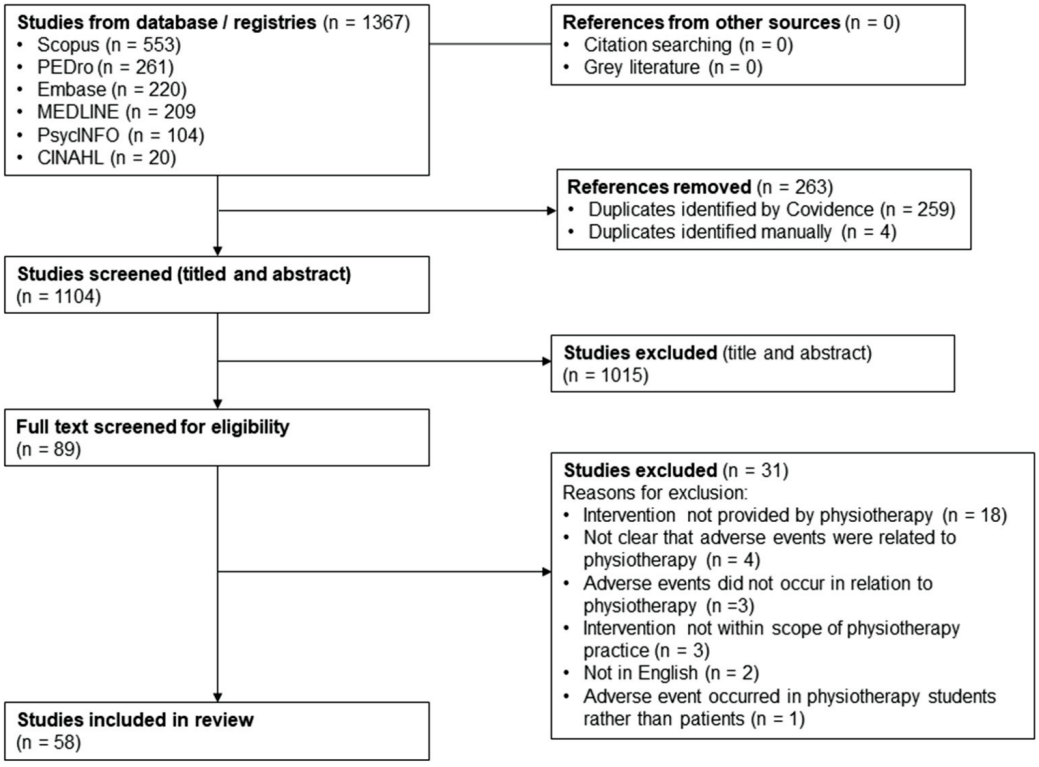


FIGURE 1 - Flow chart of included studies.

TABLE 1 - Musculoskeletal physiotherapy interventions for lower limb conditions and associated adverse events

Condition	Study design	No., age (mean±SD) years	Physiotherapy interventions			Adverse events and method of reporting			Setting and supervision provision
			ROM exercises	Strength exercises	Other interventions	Pain/muscle soreness	Stiffness	Other non-serious or serious adverse events	
Lower limb joint conditions									
Hip osteoarthritis (21)	RCT	102, 65±9	✓	✓	Manual therapy, functional balance, and gait drills	✓, ↑ post-Rx	✓	–	PP (S) and home (US)
						Collated from logbook during Rx and via questionnaires during follow-up			
Femoroacetabular impingement syndrome (22)	RCT	177, 35±9	–	–	Personalised hip therapy: physiotherapist-led rehabilitation	✓, change from baseline NR	–	–	Hospital – outpatients (S)
						Collated from questionnaires			
Hip fracture with surgical repair (23)	RCT	210, 80±8	✓	✓	Endurance, balance, function training Sensory-level TENS	–	–	Falls, femur/hip fracture, dehydration, dyspnoea	Home (visits by therapist) (S)
						Collated from telephone interviews			
Hip arthroplasty for osteoarthritis (24)	RCT	34, 65±8	–	✓	–	✓, ↑ post-Rx	–	Wound oozing, hypotension	Home (visits by therapist) (S)
						Collated from scoring on VAS and observations			
Knee osteoarthritis (25)	RCT	32, 63±8	–	–	Extracorporeal shockwave therapy	–	–	Reddening of skin, burning sensation, swelling	Hospital – outpatient (S)
						Collated from VAS and questionnaires			
Knee osteoarthritis (26)	RCT	102, 70±8	–	–	Neuromuscular exercises	–	–	Musculoskeletal disorders, connective tissue disorders	Outpatient clinic (S)
						Collated from questionnaire pain subscale			
Knee pain (27)	RCT	128, ≥50	–	✓	–	✓, ↑ post-Rx	–	Swelling	PP (S) and home (US)
						Collated from questionnaire			
Antero- or retro-patella pain (28)	RCT	16, 32±10	–	–	Gait retraining	✓, ↑ post-Rx	–	–	PP (S) and home (US)
						Collated from physiotherapy Rx notes			
Knee pain and meniscal tear (29)	RCT	161, 57±7	–	✓	Cardiovascular, coordination/balance exercise	–	–	Repeat knee surgery, acute myocardial infarction, sudden death	Outpatient (S) and home (US)
						Collated from follow-up Ax and reports			

(*Continued*)

TABLE 1 - (*Continued*)

Condition	Study design	No., age (mean±SD) years	Physiotherapy interventions			Adverse events and method of reporting			Setting and supervision provision
			ROM exercises	Strength exercises	Other interventions	Pain/muscle soreness	Stiffness	Other non-serious or serious adverse events	
Post-knee arthroplasty (30)	RCT	621, 70±8	–	–	Personalised home-based rehabilitation and standard post-operative physiotherapy	–	–	Musculoskeletal disorders, connective tissue disorders	Outpatient (S) and home (US)
						Collated from patient reports			
Pain between the buttock band and the rib arch (31)	RCT	113, 31±12	–	–	Hydrotherapy, TENS, and infrared thermal therapy	✓, ↑ from baseline	–	–	Outpatient (S)
						Collated from self-report			
Achilles tendinopathy (32)	Feasibility study	15, 38±9		✓	Double leg jump progression Single leg hops and running	–	–	Muscle tears	PP (S)
						Collated from daily diary			
Calcaneal or talar fracture (33)	RCT	50, 18-70	✓	–	Manual therapy	–	–	Wound complications, deep vein thrombosis	Outpatient (S)
						Collation from self-reported questionnaires, chart review			

✓ = applied; – = not applied; No. = number; NR = not reported; PP = private practice; RCT = randomised controlled trial; ROM = range of motion; Rx = treatment; S = supervised; SD = standard deviation; TENS = transcutaneous electrical nerve stimulation; US = unsupervised; VAS = visual analogue scale.

TABLE 2 - Cardiorespiratory physiotherapy interventions for the critically unwell and associated adverse events

Condition	Study design	No., age (mean±SD) years	Physiotherapy interventions			Adverse events			Setting and supervision provision
			Early mobilisation	Airway clearance therapy	Breathing exercises/ oxygen therapy	Cardiac issues	Oxygen-related issues	Other non-serious or serious adverse events	
Critically unwell patients undergoing MV									
Adults requiring MV (43)	RCT	371, 61±15	✓	–	–	Altered blood pressure, cardiac arrhythmia	Oxygen desaturation, tachypnoea	Pain or agitation, removal of invasive line	Hospital – inpatient (S)
						AEs reported by clinicians and patients Collated from PRO and case analyses			
Adults requiring MV (44)	RCT	115, 65±15	Endurance and resistance training	–	Respiratory therapy	Haemodynamic instability	Desaturation	–	Hospital – inpatient (S)
						AEs reported by researchers (clinicians and nurses) Collated from chart review (standard monitoring)			
Adults requiring MV >48 hours (45)	Process evaluation	36, 56±18	Bilateral lower limb in-bed cycling	–	–	–	Desaturation, increased RR	–	Hospital – inpatient (S)
						AEs reported by principal investigators Collated from pre-defined safety criteria			
Adults requiring MV <48 hours (46)	RCT	200, 65 (46-74)#	Early, goal-directed mobilisation	–	–	Hypotension	Desaturation	Dislodgement of arterial line Dislodgement of nasogastric tube	Hospital – inpatient (S)
						AEs summarised by treating clinicians Collated from chart review, patient record, bedside exam			
Children requiring MV (47)	RXT	34, 1 (0-15)		Postural changes, endotracheal instillation of saline or mucolytics, endotracheal suction, manual techniques	Manual or ventilator lung inflation	Haemodynamic instability	Transient alterations in oxygen saturation	Acute haemodynamic instability, pneumothorax, cardiac arrest 30 minutes post-PT	Hospital – inpatient (S)
						AEs reported by physiotherapists Collated from case analyses			
Adults requiring MV for 120 hours (48)	RCT	99, 58 (42-67)#	Progressive mobilisation as tolerated	–	–	Tachycardia, hypotension	Tachypnoea, desaturation	Arterial catheter removal, rectal tube removal	Hospital – inpatient (S)
						AEs reported by treating PT and OT Collated from chart review of therapy notes			

(*Continued*)

TABLE 2 - (*Continued*)

Condition	Study design	No., age (mean±SD) years	Physiotherapy interventions			Adverse events			Setting and supervision provision
			Early mobilisation	Airway clearance therapy	Breathing exercises/ oxygen therapy	Cardiac issues	Oxygen-related issues	Other non-serious or serious adverse events	
Critically unwell patients									
Patients considered functionally independent before hospitalisation (49)	Prospective observational study	142, 51 (43-64)[#]	Five times Sit-to-Stand Test	–	–	–	Dyspnoea	Muscle pain/ fatigue, chest pain	Hospital – inpatient (S)
						AEs reported by physiotherapists			
						Collated from checklist and physiological monitoring			
Uncomplicated AMI (50)	Cross-sectional study	152, 56±10	6-Minute Walk Test	–	–	Tachycardia, angina, hypotension	Dyspnoea, desaturation	Dizziness, palpitation	Hospital – inpatient (S)
						AEs reported by physiotherapists and medical staff			
						Collated from physiological responses			
Acute respiratory infection (51)	Case study	1, 59	–	Mechanical insufflation-exsufflation, expiratory vibrations, manual assisted cough, suction with saline		–	Desaturation	–	Hospital – inpatient (S)
						AEs reported by physiotherapists			
						Collated from observation during Rx			
Post-extubation (52)	Observational study	258, ≥18	Ambulation	–	NIV	Alteration in blood pressure	–	Vertigo	Hospital – inpatient (S)
						AEs reported by physiotherapists			
						Collated from observation of treatment			
Admitted to the medical ICU (53)	Retrospective analysis	99, 65 (52-72)[#]	Electrical muscle stimulation, passive or active range of motion, and mobility	–	–	Bradycardia	Respiratory distress, desaturation	Intolerance, tracheostomy removal	Hospital – inpatient (S)
						AEs reported by clinicians			
						Collated from observation during mobility Rx			
Covid-19 (54)	Feasibility	93; N/R	Prone positioning and supine turning	–	–	–	–	Endotracheal tube leak, one airway obstruction secondary to the body habitus	Hospital – inpatient (S)
						AEs reported by clinicians			
						Collated from observations and chart review			
Covid-19 (55)	Observational	84; 56±11	Active mobilisation in bed and activities of daily living training	✓	Assist invasive mechanical ventilation, lung expansion techniques	Tachycardia, hypotension, hypertension	Desaturation	Falls	Hospital – inpatient (S)
						AEs reported by physiotherapists			
						Collated from observation of Rx and chart review			

[#]Data are median (interquartile range).
✓ = applied; – = not applied; AE = adverse events; AMI = acute myocardial infarction; Covid-19 = coronavirus-19; MV = mechanical ventilation; NIV = non-invasive ventilation; No. = number; N/R = not reported; OT = occupational therapy; PRO = patient-reported outcomes; PT = physiotherapy; RCT = randomised controlled trial; RR = respiratory rate; Rx = treatment; RXT = randomised cross-over trial; S = supervised; SD = standard deviation.

reports or physiotherapists, via a mix of monitoring and chart review. Of those individuals who were critically ill who may or may not have required mechanical ventilation (43-55), physiotherapy interventions provided in the intensive care setting included early mobilisation, endurance and resistance training, strength and functional exercise testing, electrical muscle stimulation and respiratory therapy. Findings from 10 studies indicated that adverse events during these treatments were haemodynamic instability, episodes of angina, oxygen desaturation, elevated respiratory rate, vertigo and falls, line or tube dislodgement and airway obstruction during prone positioning (Tab. 2). For individuals following cardiac or abdominal surgery (56-59), interventions included exercises for breathing and upper and lower limbs, passive mobilisation, oxygen therapy, non-invasive ventilation and suction. Similar adverse events with haemodynamic instability, desaturation, dyspnoea and pain were reported (Supplementary Table 2). For those deconditioned due to COVID-19 following an acute hospital stay or related to stay-at-home orders, resistance training was linked to falls (60). Physiotherapy for managing acute respiratory infections or asthma consisted of breathing exercises and airway clearance therapy (61,62). Asthma exacerbations or episodes of desaturation were reported with these therapies (Supplementary Table 1).

In neurological physiotherapy, the patient population, interventions and adverse events are outlined in Tab. 3. Most adverse events were reported by patients or clinicians, from self-reports, observations or chart review. For patients diagnosed with stroke or an acute brain haemorrhage (63-65), the interventions consisted of treadmill or gait training. Common adverse events reported during or after these interventions were increased pain, falls and symptoms of intolerance of the activity. For those with Parkinson's disease or multiple sclerosis, physiotherapy management consisted of gait and balance training and exercise prescription (66-68). Adverse events reported in relation to these interventions were pain, falls and haemodynamic intolerance. For patients with sport-related concussion, interventions included submaximal aerobic training, sport-specific exercises and imagery techniques and were linked to headaches, dizziness and exacerbation of symptoms during exercise (69). For patients with peripheral nervous system disorders who underwent supervised aerobic exercises, pain and fatigue were the most commonly reported adverse events (70).

In oncological physiotherapy, key interventions for those with breast or other types of cancer (71-74) were resistance and aerobic exercise training, balance training and whole-body vibration (Tab. 4). These treatments were linked to increased pain, falls, haemodynamic instability, muscle strain and fatigue as adverse events. In lymphoedema physiotherapy, manual lymphatic drainage led to discomfort, lymphangitis attacks and oedema displacement (75), as reported from both patients and clinicians via self-report or monitoring. In aged care, active mobilisation exercises, lower limb strengthening, walking and balance for those with dementia were linked to increased pain from baseline measures (76). Pelvic floor muscle training resulted in vaginal discomfort, spotting and greater pain (77,78).

Discussion

This scoping review identifies adverse events related to a range of physiotherapy interventions across a mix of clinical fields. In musculoskeletal physiotherapy, increased muscle pain or soreness and to a lesser extent joint stiffness were the most commonly reported adverse events. Within cardiorespiratory physiotherapy interventions, the adverse events most commonly reported were haemodynamic or respiratory instability, while in neurological, oncological physiotherapy and aged care management, increased pain, fatigue, falls and cardiovascular intolerance were the most commonly reported adverse events. In pelvic health physiotherapy, the predominant adverse event during pelvic floor muscle training was discomfort.

For musculoskeletal physiotherapy, experiencing a certain level of pain or muscle soreness during exercise or manual therapy is not unusual, as specific exercise training and manual therapy techniques including joint mobilisation can lead to temporary muscle soreness due to the mechanical stress applied to the muscles and connective tissues (79). However, the level of pain is expected to remain within a tolerable range and be temporary in nature. Excessive or prolonged pain is considered an adverse response (80); this is the type of pain which has been reported in the included studies. The identification of these adverse events suggests that symptom monitoring during these interventions would be important to regulate the adjustment of treatment intensity to minimise pain or soreness (81). Range-of-motion exercises performed too aggressively or with excessive force have the potential to cause temporary stiffness in the area being treated (82). To minimise this effect, gradual progressions and individualised approaches may be necessary to improve the safety of this type of intervention (82). Experienced clinicians may be more likely to notice subtle signs and consistently tailor interventions, but physiotherapy students may benefit from targeted education about potential adverse events that may occur during physiotherapy treatment. 'Clinical noticing' is arguably a skill to be emphasised as practical skills are developed and refined (83). This may be key to minimising the occurrence of adverse events during these interventions when delivered by students in clinical care (84).

For cardiorespiratory physiotherapy, the adverse events described related to haemodynamic and respiratory intolerance during selected interventions. This is not an unexpected outcome given the nature of patients being critically unwell (85). The occurrence of these adverse events reinforces the importance of regular monitoring of these responses in patients undergoing treatments including exercise, early mobilisation and airway clearance techniques, in order to detect possible intolerance and enable adjustment to interventions to accommodate these clinical responses (86). Furthermore, changes in heart rate, blood pressure and patient reports of dyspnoea or dizziness during interventions highlight the necessity of monitoring clinical signs and subjective symptoms on an individual patient basis (87). This knowledge is critical for physiotherapy students to be aware of, as they gain clinical experience in the management of

TABLE 3 - Neurological physiotherapy interventions for acute and chronic conditions and associated adverse events

Condition	Study design	No., age (mean±SD) years	Physiotherapy interventions			Adverse events			Setting and supervision provision
			Strength and endurance training	Balance and coordination exercises	Other intervention types	Pain	Falls	Other non-serious and serious adverse events	
Stroke									
Subacute phase of ischaemic or haemorrhagic stroke (63)	RCT	200, 69±12	Treadmill-based, bodyweight supported training	–	–	–	✓	Fatigue, dizziness	Hospital – inpatient (S)
						AEs reported by patients Collated from self-reports			
Stroke (64)	RCT	38, 71±13	Treadmill training	–	Gait training	–	✓	Chest pain, syncope, nausea and SOB	Hospital – inpatient (S)
						AEs reported by physiotherapists and patients Collated from patient report and observations			
Acute condition									
Subarachnoid haemorrhage or subdural hematoma (65)	Observational pilot study	50, 61±14	Mobilisation as tolerated	–	–	–	–	Symptoms of light headedness, hypotension, feeling unwell	Hospital – inpatient (S)
						AEs reported by physiotherapists and patients Collated from self-report and physiotherapy monitoring during Rx			
Chronic conditions									
Parkinson's disease (66)	Exploratory clinical study	100, 75±9	✓	✓	Movement strategy training, gait training, and progressive resistance training	✓, ↑ from baseline	✓	Hypotension or hypertension, cardiorespiratory arrest, hip fracture attributed to a fall	Hospital – inpatient (S) and home (US)
						AEs reported by physiotherapists and patients Collated from self-report or witnessed by physiotherapists			
Mild to moderate Parkinson's disease (67)	Interventional study	5; 72 (69-80)	–	Balance training regime emphasising specific and highly challenging exercises	–	–	✓	Dizziness related to low blood pressure	Hospital – outpatient (S)
						AEs reported by physiotherapists and patients Collated from patient experience and physiotherapy monitoring			
Multiple sclerosis (68)	Feasibility study	17, 54 (18-70)	✓	✓	Stretching exercise	–	–	Fatigue	Home (S) – virtual group exercise
						AEs reported by patients Collated from self-report			

Condition	Study design	No., age (mean±SD) years	Physiotherapy interventions			Adverse events			Setting and supervision provision
			Strength and endurance training	Balance and coordination exercises	Other intervention types	Pain	Falls	Other non-serious and serious adverse events	
Traumatic injury									
Sport-related concussion (69)	RCT	10, 16±2	Submaximal aerobic training	Light coordination and sport-specific exercises	Visualisation and imagery techniques, home exercise programme	–	–	Worsening of symptoms, headache, dizziness	Hospital – outpatient (S) and home (US)
						AEs reported by physiotherapists and patients Collated from PRO, measurement and observation by treating physiotherapist			
Peripheral nervous system disorders									
Neuromuscular diseases (70)	Prospective pilot study	31, 58 (20-76)[#]	Supervised aerobic exercises	–	–	✓, ↑ from baseline	–	Fatigue	Home (S)
						AEs reported by physiotherapists and patients Collated from self-report, log books and patient records			

[#]Data are median (interquartile range).
✓ = applied; – = not applied; AEs = adverse events; No = number; PRO = patient-reported outcomes; RCT = randomised controlled trial; Rx = treatment; S = supervised; SD = standard deviation; SOB = shortness of breath; US = unsupervised.

TABLE 4 - Oncology, aged care and pelvic health interventions and associated adverse events

Condition	Study design	No., age (mean±SD) years	Physiotherapy interventions			Adverse events			Setting and supervision provision
			Musculoskeletal	Neurological	Other intervention types	Pain	Falls	Other non-serious or serious adverse events	
Oncology (adults and paediatrics)									
Curable breast, prostate or colorectal cancer (71)	RCT	577, 59±12	Supervised, group-based and home-based resistance training	–	–	✓, ↑ from baseline	–	Muscle strains Dizziness Injured finger	Public gym (S) and home (US)
						AEs reported by physiotherapists and patients Collated from self-report and monitoring			
Non-metastatic or metastatic breast cancer (72)	Experimental study	20, 61±10	Resistance Training	–	–	✓, ↑ from baseline	–	–	Hospital – outpatient (S)
						AEs reported by patients Collated from self-report			
Non-central nervous system cancer (73)	Retrospective chart review	147, 9±4	Strengthening and endurance exercise, jumping	Balance training	Stretching	–	✓	Tachycardia Headaches Fatigue	Hospital – inpatient (S) and outpatient (US)
						AEs reported by patients, physiotherapists and oncology doctors Collated from chart review and patient report			
Paediatric cancer (74)	Exploratory feasibility study	11, 12 (7-17)	–	–	Whole-body vibration	–	–	Bleeding observed in patients with bleeding tendencies and low platelets	Hospital – inpatient (S)
						AEs reported by physiotherapists and patients Collated from self-report and observation			
Lymphoedema									
Primary or secondary lymphedema (75)	RCT	194, 58±12	–	–	Manual lymphatic drainage	–	–	Discomfort Lymphangitis attacks Oedema displacement	Hospital – outpatient (S)
						AEs reported by physiotherapists and patients Collated from patient experience and monitoring			

Condition	Study design	No., age (mean±SD) years	Physiotherapy interventions			Adverse events			Setting and supervision provision
			Musculoskeletal	Neurological	Other intervention types	Pain	Falls	Other non-serious or serious adverse events	
Aged care									
Community-dwelling individuals with dementia (76)	RCT	54, 80±6	Active limb mobilisation exercises, lower limb strengthening exercises, and walking	Balance exercises	Whole-body vibration	✓, ↑ from baseline	–	–	Community – outpatient (S)
						AEs reported by patients			
						Collated from self-report			
Pelvic health									
Symptoms of stress or mixed urinary incontinence (77)	RCT	362, 68±6	–	–	Pelvic floor muscle training	–	–	Vaginal discomfort during intravaginal biofeedback	PP (S) and home (US)
								Vaginal spotting	
						AEs reported by participants			
						Collated from self-report			
Vaginal prolapse (78)	RCT	414, 46±5	–	–	Individual home pelvic floor muscle training programme	✓, ↑ from baseline	–	Shortness of breath and chest pain during pelvic floor muscle training	PP (S) and home (US)
						AEs reported by patients			
						Collated from self-report			

AE = adverse events; No. = number; RCT = randomised controlled trial; S = supervised; SD = standard deviation; US = unsupervised; ✓ = applied; – = not applied.

acutely unwell patients. While less frequent, dislodgement of tubes or lines remains an area requiring careful management, which likely involves multiple healthcare team members to ensure patient safety (88). The reported occurrence of falls in those who were deconditioned illustrates the need for a heightened level of awareness of this potential adverse event in this patient population to enable risk mitigation (89).

Increased pain and risk of falls during gait retraining after stroke are likely to be attributed to impairments in balance or postural instability (90), while their occurrence as part of Parkinson's disease management is associated with freezing of gait or difficulty dual tasking (91). It is not unforeseen that patients with neurological pathology such as stroke are at greater risk of falls during treatment, given impairments such as weakness, sensory deficits and poor balance. In addition, these patient populations may have a heightened fear of falling, which can lead to hesitation and guarding, all factors which increase the risk of falls (92,93). To minimise the risk of falls, optimal therapist body mechanics and safe patient handling techniques are required (94), and an awareness of the appropriate level of assistance and decluttered environments is also important in ensuring safety for patients engaged in gait retraining and walking practice (95). For those with Parkinson's disease, clinician-directed education regarding safe mobility and cueing strategies, together with healthcare team collaboration to ensure optimal symptom management prior to physiotherapy interventions are potential approaches to minimise the risk of these adverse events (96). This may be further supported by interdisciplinary communication and sharing knowledge regarding an individual patient's clinical status to assist in reducing risk. Fatigue is a common side effect in multiple sclerosis and conditions affecting the peripheral nervous system (97). Education on pacing and energy conservation techniques including breaking down activities and incorporating rest periods (98) can reduce the occurrence of this type of adverse event. Adjusting the intensity and intervention duration according to a patient's tolerance is also crucial to reduce the risk of exacerbating fatigue (99).

Within oncology and aged care physiotherapy, the commonly reported adverse event of increased pain during exercises and walking may be caused by muscle wasting or reduced bone density secondary to specific cancer treatments or age-related deconditioning (100,101). The nature of these diagnoses indicates these undesirable outcomes are not unexpected. Collaborating with other healthcare professionals, such as dietitians and pharmacists, to improve the nutrition intake and provide medication for slowing muscle atrophy is important for these patient populations. A multidisciplinary approach, together with proactive education for patients and caregivers, may enhance muscle function and exercise performance and minimise discomfort during physiotherapy interventions (102). Within paediatric oncology, early recognition of changes in clinical signs and symptoms secondary to cancer and its related treatment (103) may be supported by the adoption of a proactive approach. This enables children's caregivers in monitoring individual responses and providing education to enable adjustments to be made in a timely manner (104).

In pelvic health physiotherapy, discomfort during pelvic floor muscle training can be attributed to muscle overwork or irritation (105). Conducting an individualised training programme may be necessary to provide instructions on correct techniques and minimise the discomfort from improper contractions (106).

Identifying adverse events across a range of clinical areas in physiotherapy provides valuable information for optimising patient-centred care. This information has direct applications for the practice of physiotherapy clinicians and physiotherapy students. Awareness of adverse events associated with physiotherapy in different practice contexts can inform adjustments to interventions based on early awareness, close monitoring of patient responses, and education of patients and caregivers of the potential risks to enhance safety. In addition, this information may be used to develop physiotherapy students' knowledge and understanding of situational awareness in a mix of clinical areas of practice, having the potential to inform and implement timely mitigating strategies to reduce the risk of adverse events (107).

The mean age range included across the studies is vast. For musculoskeletal physiotherapy, the mean age of participants was 18 to 80 years; for cardiorespiratory physiotherapy, the mean age was 1 to 65 years; for neurological physiotherapy, age ranged from 16 to 73 years. A wide range was also apparent for oncology physiotherapy and pelvic health. Although the common comorbidities in each study have not been reported in this review, for some participants, the presence of co-existing conditions may have influenced the occurrence of adverse events during physiotherapy assessment or intervention practices and therefore contributed to an undesirable outcome.

Many challenges related to identifying and tracking adverse events in hospital and healthcare settings have been reported (108), and it is recommended that more than one method be used to identify adverse events comprehensively (109). There is potential for using routinely collected data in electronic health records implemented in hospitals and primary care to develop automated adverse event recording systems (110), but confidence in data quality is needed. One factor affecting data quality is variability in the terminology used when reporting adverse events (111,112). In this scoping review, we sought to identify what is known about adverse events in physiotherapy and have identified how adverse events have been reported in a range of physiotherapy practice areas across hospital and community-based settings, and the terminology used to describe adverse events in the physiotherapy context.

There are limitations to this scoping review. This scoping review only included the adverse events or serious adverse events that occurred during physiotherapy interventions and assessments and that were defined by the study authors as adverse events attributable to interactions with physiotherapy. It is likely that there is a variation between studies regarding the definition and threshold for adverse events and criteria for attributing these events to physiotherapy. The search strategy focused on selected terms (adverse events); this terminology may not be consistently used within studies to describe events which are considered undesirable

outcomes. For this reason, it is possible that some relevant studies which identified the occurrence of adverse events were not included. We did not choose to include systematic reviews of all types of physiotherapy interventions, as not all reviews are guaranteed to incorporate mention of adverse events as part of their data extraction. Nor did we include systematic reviews of adverse events related to interventions which are within the scope of physiotherapy practice, as studies greater than 10 years contributed to the collated data and this may not be reflective of current practice. It is possible that some adverse events related to physiotherapy interventions occurred after the intervention had been delivered. Therefore, there is a small possibility that this review has potentially missed important adverse events associated with physiotherapy.

This review highlighted adverse events related to physiotherapy interventions across various clinical settings. Increased awareness of adverse events reported in studies of physiotherapy interventions provides an opportunity to focus on clinical awareness when tailoring interventions to individuals, implementation of preventive strategies and designing curriculum related to patient safety in physiotherapy education programmes.

Disclosures

Conflict of interest: The authors have no conflicts of interest.

Financial support: The authors have no funding or financial support to disclose.

Author contributions: YW contributed to the conception of the work, undertook the search strategy and data extraction, prepared figures and tables, authored and reviewed drafts of the manuscript, and approved the final draft. ND contributed to the conception of the work, reviewed drafts of the manuscript and approved the final draft. MF contributed to the conception of the work, reviewed drafts of the manuscript and approved the final draft. ALL contributed to the conception of the work, undertook the search strategy, authored and reviewed drafts of the manuscript, and approved the final draft.

Data availability statement: Data sharing not applicable. No new data were created or analysed in this study and more data are available as supplementary material to this article.

References

1. World Health Organization. Global patient safety action plan 2021-2030. 2021. Online. (Accessed August 2024)
2. Voskanyan YV. [Safety of patients and adverse events related thereto in medicine]. Angiol Sosud Khir. 2018;24(4):11-17. PubMed
3. Australian Institute of Health and Welfare. Australia's Health 2018. Canberra 2018. Online (Accessed August 2024)
4. Carlesso LC, Macdermid JC, Santaguida LP. Standardization of adverse event terminology and reporting in orthopaedic physical therapy: application to the cervical spine. J Orthop Sports Phys Ther. 2010;40(8):455-463. CrossRef PubMed
5. Rajendran D, Mullinger B, Fossum C, Collins P, Froud R. Monitoring self-reported adverse events: A prospective, pilot study in a UK osteopathic teaching clinic. Int J Osteopath Med. 2009;12(2):49-55. CrossRef
6. He Y, Sarntivijai S, Lin Y, et al. OAE: the ontology of adverse events. J Biomed Semantics. 2014;5(1):29. CrossRef PubMed
7. Gonzalez-Caminal G, Camps Gomez A, Gomar-Sancho C. Patient safety in physiotherapy: are errors that cause or could cause harm preventable? In: Salen PN, Stawicki SP, eds. Contemporary topics in patient safety. Vol 2. IntechOpen; 2023. CrossRef
8. Sricharoenchai T, Parker AM, Zanni JM, Nelliot A, Dinglas VD, Needham DM. Safety of physical therapy interventions in critically ill patients: a single-center prospective evaluation of 1110 intensive care unit admissions. J Crit Care. 2014;29(3):395-400. CrossRef PubMed
9. Partridge C, Kitchen C. Adverse effects of electrotherapy used by physiotherapists. Physiotherapy. 1999;85(6):298-303. CrossRef
10. Cassina T, Putzu A, Santambrogio L, Villa M, Licker MJ. Hemodynamic challenge to early mobilization after cardiac surgery: a pilot study. Ann Card Anaesth 2016;19:425-432. CrossRef PubMed
11. Eggmann S, Nydahl P, Gosselink R, Bissett B. We need to talk about adverse events during physical rehabilitation in critical care trials. EClinicalMedicine. 2024;68:102439. CrossRef PubMed
12. Osborne JA, Botkin R, Colon-Semenza C, et al. Physical therapy management of Parkinson Disease: clinical practice guideline from the American Physical Therapy Association. Phys Ther 2022;102(4):1-36. CrossRef PubMed
13. King J, Anderson CM. Patient safety and physiotherapy: what does it mean for your clinical practice? Physiother Can. 2010;62(3):172-179. CrossRef PubMed
14. Phillips AC, Mackintosh SF, Bell A, Johnston KN. Developing physiotherapy student safety skills in readiness for clinical placement using standardised patients compared with peer-role play: a pilot non-randomised controlled trial. BMC Med Educ. 2017;17(1):133. CrossRef PubMed
15. Australian Health Practitioner Regulation Agency. Regulating Australia's health practitioners. 2024. Online. (Accessed August 2024)
16. Walshe N, Ryng S, Drennan J, et al. Situation awareness and the mitigation of risk associated with patient deterioration: A meta-narrative review of theories and models and their relevance to nursing practice. Int J Nurs Stud. 2021;124:104086. CrossRef PubMed
17. Macrae C. The problem with incident reporting. BMJ Qual Saf. 2016;25(2):71-75. CrossRef PubMed
18. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol. 2018;18(1):143. CrossRef PubMed
19. 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
20. Australian Physiotherapy Association. The APA valuing skills series: physiotherapy range of clinical practice, Position paper. 2023. Online
21. Bennell KL, Egerton T, Martin J, et al. Effect of physical therapy on pain and function in patients with hip osteoarthritis: a randomized clinical trial. JAMA. 2014;311(19):1987-1997. CrossRef PubMed
22. 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
23. Magaziner J, Mangione KK, Orwig D, et al. Effect of a multicomponent home-based physical therapy intervention on ambulation after hip fracture in older adults: the CAP randomized clinical trial. JAMA. 2019;322(10):946-956. CrossRef PubMed

24. Mikkelsen LR, Petersen AK, Mechlenburg I, Mikkelsen S, Søballe K, Bandholm T. Description of load progression and pain response during progressive resistance training early after total hip arthroplasty: secondary analyses from a randomized controlled trial. Clin Rehabil. 2017;31(1):11-22. CrossRef PubMed
25. Zhong Z, Liu B, Liu G, et al. A randomized controlled trial on the effects of low-dose extracorporeal shockwave therapy in patients with knee osteoarthritis. Arch Phys Med Rehabil. 2019;100(9):1695-1702. CrossRef PubMed
26. Bandak E, Christensen R, Overgaard A, et al. Exercise and education versus saline injections for knee osteoarthritis: a randomised controlled equivalence trial. Ann Rheum Dis. 2022;81(4):537-543. CrossRef PubMed
27. Bennell KL, Nelligan RK, Kimp AJ, et al. What type of exercise is most effective for people with knee osteoarthritis and co-morbid obesity? the TARGET randomized controlled trial. Osteoarthritis Cartilage. 2020;28(6):755-765. CrossRef PubMed
28. Bonacci J, Hall M, Saunders N, Vicenzino B. Gait retraining versus foot orthoses for patellofemoral pain: a pilot randomised clinical trial. J Sci Med Sport. 2018;21(5):457-461. CrossRef PubMed
29. van de Graaf VA, Noorduyn JCA, Willigenburg NW, et al; ESCAPE Research Group. Effect of early surgery vs physical therapy on knee function among patients with nonobstructive meniscal tears: the ESCAPE randomised clinical trial. JAMA. 2018;320(13):1328-1337. CrossRef PubMed
30. Barker KL, Room J, Knight R, et al; CORKA Trial group. Home-based rehabilitation programme compared with traditional physiotherapy for patients at risk of poor outcome after knee arthroplasty: the CORKA randomised controlled trial. BMJ Open. 2021;11(8):e052598. CrossRef PubMed
31. Peng MS, Wang R, Wang YZ, et al. Efficacy of therapeutic aquatic exercise vs physical therapy modalities for patients with chronic low back pain: a randomized clinical trial. JAMA Netw Open. 2022;5(1):e2142069. CrossRef PubMed
32. Sancho I, Morrissey D, Willy RW, Barton C, Malliaras P. Education and exercise supplemented by a pain-guided hopping intervention for male recreational runners with midportion Achilles tendinopathy: A single cohort feasibility study. Phys Ther Sport. 2019;40:107-116. CrossRef PubMed
33. Albin SR, Koppenhaver SL, Van Boerum DH, McPoil TG, Morgan J, Fritz JM. Timing of initiating manual therapy and therapeutic exercises in the management of patients after hindfoot fractures: a randomized controlled trial. J Man Manip Ther. 2018;26(3):147-156. CrossRef PubMed
34. Kim JI, Han CH, Jeon JH, et al. Effectiveness and safety of polydioxanone thread embedding acupuncture compared to physical therapy in the treatment of patients with non-specific chronic neck pain: an assessor-blinded, randomised, controlled, clinical trial. J Pain Res. 2021;14:201-211. CrossRef PubMed
35. Saper RB, Lemaster C, Delitto A, et al. Yoga, physical therapy, or education for chronic low back pain: a randomized noninferiority trial. Ann Intern Med. 2017;167(2):85-94. CrossRef PubMed
36. Sharma S, Jensen MP, Moseley GL, Abbott JH. Results of a feasibility randomised clinical trial on pain education for low back pain in Nepal: the Pain Education in Nepal-Low Back Pain (PEN-LBP) feasibility trial. BMJ Open. 2019;9(3):e026874. CrossRef PubMed
37. Liaghat B, Skou ST, Søndergaard J, Boyle E, Søgaard K, Juul-Kristensen B. Short-term effectiveness of high-load compared with low-load strengthening exercise on self-reported function in patients with hypermobile shoulders: a randomised controlled trial. Br J Sports Med. 2022;56(22):1269-1276. CrossRef PubMed
38. Rangan A, Brealey SD, Keding A, et al; UK FROST Study Group. Management of adults with primary frozen shoulder in secondary care (UK FROST): a multicentre, pragmatic, three-arm, superiority randomised clinical trial. Lancet. 2020;396(10256):977-989. CrossRef PubMed
39. Brealey S, Northgraves M, Kottam L, et al. Surgical treatments compared with early structured physiotherapy in secondary care for adults with primary frozen shoulder: the UK FROST three-arm RCT. Health Technol Assess. 2020;24(71):1-162. CrossRef PubMed
40. Kjær BH, Magnusson SP, Henriksen M, et al. Effects of 12 weeks of progressive early active exercise therapy after surgical rotator cuff repair: 12 weeks and 1-year results from the CUT-N-MOVE randomised controlled trial. Am J Sports Med. 2021;49(2):321-331. CrossRef PubMed
41. Stuart S, Armstrong M, Sewell J, Dixon C, Morris R. Acupuncture for whiplash-associated disorder following road traffic collision: a physiotherapy service evaluation. Acupunct Med. 2020;38(4):272-278. CrossRef PubMed
42. Funabashi M, Carlesso LC. Symptoms patients receiving manual therapy experienced and perceived as adverse: a secondary analysis of a survey of patients' perceptions of what constitutes an adverse response. J Man Manip Ther. 2021;29(1):51-58. CrossRef PubMed
43. Hodgson CL, Bailey M, Bellomo R, et al; TEAM Study Investigators and the ANZICS Clinical Trials Group. Early active mobilization during mechanical ventilation in the ICU. N Engl J Med. 2022;387(19):1747-1758. CrossRef PubMed
44. Eggmann S, Verra ML, Luder G, Takala J, Jakob SM. Effects of early, combined endurance and resistance training in mechanically ventilated, critically ill patients: a randomised controlled trial. PLoS One. 2018;13(11):e0207428. CrossRef PubMed
45. Nickels MR, Aitken LM, Barnett AG, Walsham J, McPhail SM. Acceptability, safety, and feasibility of in-bed cycling with critically ill patients. Aust Crit Care. 2020;33(3):236-243. CrossRef PubMed
46. Schaller SJ, Anstey M, Blobner M, et al; International Early SOMS-guided Mobilization Research Initiative. Early, goal-directed mobilisation in the surgical intensive care unit: a randomised controlled trial. Lancet. 2016;388(10052):1377-1388. CrossRef PubMed
47. Shannon H, Stocks J, Gregson RK, Dunne C, Peters MJ, Main E. Clinical effects of specialist and on-call respiratory physiotherapy treatments in mechanically ventilated children: a randomised crossover trial. Physiotherapy. 2015;101(4):349-356. CrossRef PubMed
48. Patel BK, Wolfe KS, Patel SB, et al. Effect of early mobilisation on long-term cognitive impairment in critical illness in the USA: a randomised controlled trial. Lancet Respir Med. 2023;11(6):563-572. CrossRef PubMed
49. de Melo TA, Silva Guimarães F, Lapa E, Silva JR. The five times sit-to-stand test: safety, validity and reliability with critical care survivors at ICU discharge. Arch Physiother. 2022;13(1):2. CrossRef PubMed
50. Diniz LS, Neves VR, Starke AC, Barbosa MPT, Britto RR, Ribeiro ALP. Safety of early performance of the six-minute walk test following acute myocardial infarction: a cross-sectional study. Braz J Phys Ther. 2017;21(3):167-174. CrossRef PubMed
51. Apps C, Morris K, Allum L, et al. Use of mechanical insufflation exsufflation and manual techniques in an intubated adult with COVID-19 positioned in prone – a case study. Physiother Res Int. 2022;27(4):e1961. CrossRef PubMed
52. Takei M, Nozawa E, Sousa ML, et al. Physiotherapy intervention is safety in early postoperative patients after cardiac surgery? Eur Respir J. 2017;50:PA2560. Online (Accessed August 2024)

53. Lee H, Ko YJ, Suh GY, et al. Safety profile and feasibility of early physical therapy and mobility for critically ill patients in the medical intensive care unit: beginning experiences in Korea. J Crit Care. 2015;30(4):673-677. CrossRef PubMed
54. Rollinson TC, Rose J, McDonald LA, et al. The PhLIP team: feasibility of a physiotherapy-led intensive prone positioning team initiative during the COVID-19 pandemic. Aust Crit Care. 2023;36(6):974-979. CrossRef PubMed
55. Rossi V, Santambrogio M, Del Monaco C, et al. Safety and feasibility of physiotherapy in ICU-admitted severe COVID-19 patients: an observational study. Monaldi Arch Chest Dis. 2022;92(4). CrossRef PubMed
56. Sousa MLA, Coimbra VRM, Takei MT, Melo CCA, Feltrim MIZ, Nozawa E. Physiological abnormalities and adverse events during physical therapy in the intensive care unit after cardiac surgery: A prospective observational study. Braz J Phys Ther. 2021;25(5):623-631. CrossRef PubMed
57. Lockstone J, Parry SM, Denehy L, Robertson IK, Story D, Boden I. Non-Invasive Positive airway Pressure thErapy to Reduce Postoperative Lung complications following Upper abdominal Surgery (NIPPER PLUS): a pilot randomised control trial. Physiotherapy. 2022;117:25-34. CrossRef PubMed
58. Maffei P, Wiramus S, Bensoussan L, et al. Intensive early rehabilitation in the intensive care unit for liver transplant recipients: a randomised controlled trial. Arch Phys Med Rehabil. 2017;98(8):1518-1525. CrossRef PubMed
59. Boden I, Sullivan K, Hackett C, et al. Intensive physical therapy after emergency laparotomy: pilot phase of the Incidence of Complications following Emergency Abdominal surgery Get Exercising randomized controlled trial. J Trauma Acute Care Surg. 2022;92(6):1020-1030. CrossRef PubMed
60. Stevens-Lapsley JE, Derlein D, Churchill L, et al. High-intensity home health physical therapy among older adult veterans: a randomized controlled trial. J Am Geriatr Soc. 2023;71(9):2855-2864. CrossRef PubMed
61. Bruton A, Lee A, Yardley L, et al. Physiotherapy breathing retraining for asthma: a randomised controlled trial. Lancet Respir Med. 2018;6(1):19-28. CrossRef PubMed
62. Corten L, Morrow BM. Use of airway clearance therapy in children hospitalised with acute lower respiratory tract infections in a South African paediatric hospital. S Afr J Physiother. 2020;76(1):1367. CrossRef PubMed
63. Nave AH, Rackoll T, Grittner U, et al. Physical Fitness Training in Patients with Subacute Stroke (PHYS-STROKE): multicentre, randomised controlled, endpoint blinded trial. BMJ. 2019;366:l5101. CrossRef PubMed
64. Baer GD, Salisbury LG, Smith MT, Pitman J, Dennis M. Treadmill training to improve mobility for people with sub-acute stroke: a phase II feasibility randomized controlled trial. Clin Rehabil. 2018;32(2):201-212. CrossRef PubMed
65. O'Shea K, Stiller K. Early progressive mobilisation following acute subarachnoid or subdural haemorrhage: an observational pilot study. Int J Ther Rehabil. 2016;23(7):339-346. CrossRef
66. Caniça V, Bouça-Machado R, Rosa MM, Ferreira JJ; CNS Physiotherapy Study Group. Adverse events in physiotherapy interventions in Parkinsonian patients. Mov Disord Clin Pract. 2022;9(6):744-750. CrossRef PubMed
67. Conradsson D, Löfgren N, Ståhle A, Franzén E. Is highly challenging and progressive balance training feasible in older adults with Parkinson's disease? Arch Phys Med Rehabil. 2014;95(5):1000-1003. CrossRef PubMed
68. Garg H, Rutherfurd C, Labrum J, Hawley B, Gard E, Davis J. Feasibility, outcomes, and perceptions of a virtual group exercise program in multiple sclerosis. J Neurol Phys Ther. 2024;48(1):54-63. CrossRef PubMed
69. Chan C, Iverson GL, Purtzki J, et al. Safety of active rehabilitation for persistent symptoms after paediatric sport-related concussion: a randomised controlled trial. Arch Phys Med Rehabil. 2018;99(2):242-249. CrossRef PubMed
70. Voorn EL, Koopman FS, Nollet F, Brehm M-A. Individualised aerobic exercise in neuromuscular diseases: A pilot study on the feasibility and preliminary effectiveness to improve physical fitness. Phys Ther. 2021;101(3):pzaa213. CrossRef PubMed
71. Demmelmaier I, Brooke HL, Henriksson A, et al. Does exercise intensity matter for fatigue during (neo-)adjuvant cancer treatment? The Phys-Can randomized clinical trial. Scand J Med Sci Sports. 2021;31(5):1144-1159. CrossRef PubMed
72. Schlüter K, Schneider J, Rosenberger F, Wiskemann J. Feasibility of high-intensity resistance training sessions in cancer survivors. J Strength Cond Res. 2022;36(9):2643-2652. CrossRef PubMed
73. Gilchrist L, Tanner LR. Safety of symptom-based modification of physical therapy interventions in paediatric oncology patients with and without low blood counts. Rehabil Oncol. 2017;35(1):3-8. CrossRef
74. Oschwald V, Prokop A, Maas V, et al. Whole-body vibration training for inpatient children and adolescents receiving chemotherapy for first cancer diagnosis: an exploratory feasibility study. Ger J Exerc Sport Res. 2022;53(1):30-36. CrossRef
75. Forner-Cordero I, Muñoz-Langa J, DeMiguel-Jimeno JM, Rel-Monzó P. Physical therapies in the decongestive treatment of lymphedema: A randomized, non-inferiority controlled study. Clin Rehabil. 2021;35(12):1743-1756. CrossRef PubMed
76. Lam FMH, Liao LR, Kwok TCY, Pang MYC. Effects of adding whole-body vibration to routine day activity program on physical functioning in elderly with mild or moderate dementia: a randomized controlled trial. Int J Geriatr Psychiatry. 2018;33(1):21-30. CrossRef PubMed
77. Dumoulin C, Morin M, Danieli C, et al; Urinary Incontinence and Aging Study Group. Group-based vs individual pelvic floor muscle training to treat urinary incontinence in older women: a randomized clinical trial. JAMA Intern Med. 2020;180(10):1284-1293. CrossRef PubMed
78. Hagen S, Glazener C, McClurg D, et al. Pelvic floor muscle training for secondary prevention of pelvic organ prolapse (PREVPROL): a multicentre randomised controlled trial. Lancet. 2017;389(10067):393-402. CrossRef PubMed
79. Semmons J. The role of specialist physiotherapy in a pain management clinic – traditional and novel approaches. Anaesth Intensive Care Med. 2022;23(7):405-408. CrossRef
80. Ryall C, Coggon D, Peveler R, Reading I, Palmer KT. Pain tolerance in patients presenting to primary care and physiotherapy services with upper limb disorders. Occup Environ Med. 2007;64(5):349-351. CrossRef PubMed
81. Malfliet A, Ickmans K, Huysmans E, et al. Best evidence rehabilitation for chronic pain part 3: low back pain. J Clin Med. 2019;8(7):1063. CrossRef PubMed
82. Hirata K, Yamadera R, Akagi R. Associations between range of motion and tissue stiffness in young and older people. Med Sci Sports Exerc. 2020;52(10):2179-2188. CrossRef PubMed
83. Clement T, Bolton J, Griffiths L, Cracknell C, Molloy E. 'Noticing' in health professions education: time to pay attention? Med Educ. 2023;57(4):305-314. CrossRef PubMed
84. Graj E, Sheen J, Dudley A, Sutherland-Smith W. Adverse health events associated with clinical placement: A systematic review. Nurse Educ Today. 2019;76:178-190. CrossRef PubMed
85. Rocha A, Arbex FF, Sperandio PA, et al. Heart or lungs? Uncovering the causes of exercise intolerance in a patient with chronic cardiopulmonary disease. Ann Am Thorac Soc. 2018;15(9):1096-1104. CrossRef PubMed

86. Nazir A, Hasri IP. Pathophysiology and rehabilitation management of exercise intolerance in COVID-19 patients. Ann Thorac Med. 2022;17(2):87-93. CrossRef PubMed
87. Stiller K, Phillips A, Lambert P. The safety of mobilisation and its effect on haemodynamic and respiratory status of intensive care patients. Physiother Theory Pract. 2004;20(3):175-185. CrossRef
88. Munoz-Mozas G. Solving the problem of IV dislodgement. Br J Nurs. 2022;31(2):S4-S7. CrossRef PubMed
89. Shafizadeh M, Parvinpour S, Ali K. Effect of home-based exercise on falls in community-dwelling older adults: an umbrella review. Sport Sci Health. 2022;19:743-756. CrossRef PubMed
90. Roelofs JMB, Zandvliet SB, Schut IM, et al. Mild stroke, serious problems: limitations in balance and gait capacity and the impact on fall rate, and physical activity. Neurorehabil Neural Repair. 2023;37(11-12):786-798. CrossRef PubMed
91. Gilat M, Ginis P, Zoetewei D, et al. A systematic review on exercise and training-based interventions for freezing of gait in Parkinson's disease. NPJ Parkinsons Dis. 2021;7(1):81. CrossRef PubMed
92. Landers MR, Jacobson KM, Matsunami NE, McCarl HE, Regis MT, Longhurst JK. A vicious cycle of fear of falling avoidance behavior in Parkinson's disease: a path analysis. Clin Park Relat Disord. 2021;4:100089. CrossRef PubMed
93. Xie Q, Pei J, Gou L, et al. Risk factors for fear of falling in stroke patients: a systematic review and meta-analysis. BMJ Open. 2022;12(6):e056340. CrossRef PubMed
94. Eng JJ. Fitness and mobility exercise (FAME) program for stroke. Top Geriatr Rehabil. 2010;26(4):310-323. CrossRef PubMed
95. Moore SA, Boyne P, Fulk G, Verheyden G, Fini NA. Walk the Talk: current evidence for walking recovery after stroke, future pathways and a mission for research and clinical practice. Stroke. 2022;53(11):3494-3505. CrossRef PubMed
96. Alatawi SF. A scoping review of the nature of physiotherapists' role to avoid fall in people with Parkinsonism. Neurol Sci. 2021;42(9):3733-3748. CrossRef PubMed
97. Chaudhuri A, Behan PO. Fatigue in neurological disorders. Lancet. 2004;363(9413):978-988. CrossRef PubMed
98. Vatwani A, Margonis R. Energy conservation techniques to decrease fatigue. Arch Phys Med Rehabil. 2019;100(6):1193-1196. CrossRef PubMed
99. Khan F, Amatya B. Management of fatigue in neurological disorders: implications for rehabilitation. J Int Soc Phys Rehabil Med. 2018;1(2):9-36. CrossRef
100. Griffith K, Wenzel J, Shang J, Thompson C, Stewart K, Mock V. Impact of a walking intervention on cardiorespiratory fitness, self-reported physical function, and pain in patients undergoing treatment for solid tumors. Cancer. 2009;115(20):4874-4884. CrossRef PubMed
101. Niederstrasser NG, Attridge N. Associations between pain and physical activity among older adults. PLoS One. 2022;17(1):e0263356. CrossRef PubMed
102. Wilson CM, Arena SK, Boright LE. State of the art physiotherapist-led approaches to safe aging in place. Arch Physiother. 2022;12(1):17. CrossRef PubMed
103. van Deuren S, Boonstra A, van Dulmen-den Broeder E, Blijlevens N, Knoop H, Loonen J. Severe fatigue after treatment for childhood cancer. Cochrane Database Syst Rev. 2020;3(3):CD012681. CrossRef PubMed
104. Toro C, Felmingham B, Jessop S, et al. Cardio-oncology recommendations for pediatric oncology patients: an Australian and New Zealand Delphi consensus. JACC Adv. 2022;1(5):100155. CrossRef PubMed
105. National Institute of Clinical Excellence. Pelvic floor dysfunction: prevention and non-surgical management. NICE guideline, London, 2021. Online. (Accessed August 2024)
106. Hagen S, Stark D, Glazener C, et al; POPPY Trial Collaborators. Individualised pelvic floor muscle training in women with pelvic organ prolapse (POPPY): a multicentre randomised controlled trial. Lancet. 2014;383(9919):796-806. CrossRef PubMed
107. Karanfilian B, Saks N. Measuring the effects of an observation training program for first-year medical students. Med Sci Educ. 2018;28(4):649-653. CrossRef
108. Harrison R, Walton M, Manias E, et al. The missing evidence: a systematic review of patients' experiences of adverse events in health care. Int J Qual Health Care. 2015;27(6):424-442. CrossRef PubMed
109. Olsen S, Neale G, Schwab K, et al. Hospital staff should use more than one method to detect adverse events and potential adverse events: incident reporting, pharmacist surveillance and local real-time record review may all have a place. Qual Saf Health Care. 2007;16(1):40-44. CrossRef PubMed
110. Bates DW, Levine D, Syrowatka A, et al. The potential of artificial intelligence to improve patient safety: a scoping review. NPJ Digit Med. 2021;4(1):54. CrossRef PubMed
111. Bruce J, Russell EM, Mollison J, Krukowski ZH. The measurement and monitoring of surgical adverse events. Health Technol Assess. 2001;5(22):1-194. CrossRef PubMed
112. Rafter N, Hickey A, Condell S, et al. Adverse events in healthcare: learning from mistakes. QJM. 2015;108(4):273-277. CrossRef PubMed

Arch Physioter 2024; 14: 155-169
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3217
ORIGINAL RESEARCH ARTICLE

Factors contributing to non-compliance with active physiotherapy guidelines among chronic low back pain patients in India

G. Shankar Ganesh[1,2], **Abdur Raheem Khan**[2], **Ashfaque Khan**[2]

[1]Department of Physiotherapy, Composite Regional Centre for Skill Development, Rehabilitation, and Empowerment of Persons with Disabilities, Lucknow - India
[2]Department of Physiotherapy, Integral University, Lucknow - India

ABSTRACT

Introduction: Physiotherapists exhibit different degrees of adherence to clinical guidelines for low back pain (LBP). The preferences and expectations of their patients significantly influence physiotherapists' adherence to these guidelines. Therefore, it is crucial to have a comprehensive analysis of the patients' perspectives, which can identify the factors that prevent the implementation of an active approach.

Methods: We conducted semi-structured interviews with patients suffering from non-specific chronic LBP (CLBP). We transcribed the semi-structured interviews verbatim and conducted an inductive thematic analysis to uncover themes related to the participants' expectations and experiences of consultations with physiotherapists for CLBP.

Results: In total, we interviewed thirty-three individuals, with 14 women and 19 men (mean age 53 + 12 years). Our thematic analysis discovered six overarching themes that are relevant to patients' expectations and experiences. We identified several sub-themes under the "physiotherapist-related factors" and "patient-related factors" themes. Additional themes recognized were guideline-related factors**,** institution-related factors**,** healthcare-related factors, and health information. A significant number of participants expressed dissatisfaction with the short timeframe allocated by the physiotherapist.

Conclusions: Multiple participants expressed dissatisfaction with their experience, particularly about the quality of explanations and the nature of the exercises provided. This emphasizes the importance of patient education, and physiotherapists should consider suggesting active interventions that the family, society, and culture can more easily accept. Accordingly, the formulation of future guidelines for nations like India should take into account these patient expectations and perspectives.

Keywords: Exercise, Implementation science, Low back pain, Motivation, Patients, Qualitative research

What's already known about this topic?

- *Clinical practice guidelines (CPGs) inform healthcare providers to follow pre-set recommendations to guide health intervention decisions.*
- *The low adoption of LBP guidelines in physiotherapy is well known.*
- *Patient preferences and expectations can impact healthcare providers' CPG adherence.*

What does the study add?

- *Multiple elements, both internal (non-consideration of local conditions and target audience, resource constraints) and external to the domain of CPGs, such as societal (perception of physical activity within the community), familial (gender roles, family expectations), and cultural influences (preference for traditional exercises and outdoor exercise limitations for women), contribute to patients' acceptance of CPG recommendations.*
- *Therefore, it is essential for physiotherapists to prescribe culturally relevant interventions while improving communication to assist patients in achieving their functional goals.*

Received: July 20, 2024
Accepted: December 16, 2024
Published online: December 29 2024

Corresponding author:
G. Shankar Ganesh
email: shankarpt@rediffmail.com; gsganesh@student.iul.ac.in

Introduction

Low back pain (LBP) is the most common musculoskeletal disorder globally, and the financial burdens associated with its treatment and impact on society are substantial. LBP is a prominent contributor to disability in low- and middle-income countries (LMICs) (1). The prevalence of disability

caused by LBP has increased by over 50% in these regions since 1990 (2). For individuals with chronic diseases, effective management of their disorders is essential to mitigate their effects, enhance health outcomes, avert additional disability, and decrease healthcare expenses (3). A recent study examining the main characteristics of LBP treatment in LMICs found that the care provided did not consistently adhere to the most up-to-date and reliable evidence (4). This presents a significant challenge for contemporary healthcare systems, necessitating the provision of more effective care.

Clinical practice guidelines (CPGs) have been published to assist healthcare professionals in adopting the pre-established recommendations designed to influence decisions on health interventions. Physical treatments for chronic LBP (CLBP) include graded activity or exercise programs that specifically target function gains (5). The majority of the CPGs recommend exercise therapy as the initial therapeutic option for regular use (6-7). When compared to no treatment, usual care, or placebo, exercise therapy is linked to significant improvements in functional outcomes (7). Improved adherence to guidelines is projected to enhance treatment outcomes and result in cost savings (8). Nevertheless, the anticipated enhancements in patient outcomes and a decrease in healthcare-associated expenses have not materialized (9). The existing understanding regarding the suitability of guideline recommendations, mostly originating from high-income nations, for LMICs remains uncertain (5).

Adherence to therapy, defined as the degree to which patients comply with the prescribed advice from their healthcare provider, is a crucial element of ongoing health management (10). DiMatteo observed that 24.8% of the patients had a typical prevalence of not following healthcare recommendations (11). Due to the prevalent issue of non-adherence, a significant portion of patients fail to achieve the full potential benefits of therapy, leading to unfavorable health outcomes, diminished quality of life, and heightened healthcare expenses (12). Several factors have been proposed as the causes for inadequate compliance with treatment suggestions. These factors encompass the patient's socioeconomic position, lack of agreement between providers and patients, misconceptions regarding the role of interventions, reduced motivation due to a perceived lack of treatment success, limited understanding of health information, resistance to the health belief paradigm, and social stigma (13-14).

Research conducted in several countries has shown that physiotherapists have varying levels of adherence to clinical guidelines for LBP (15-17). In our prior research, we observed that Indian physiotherapists generally adhered to CPGs while treating patients with LBP. Further, the use of certain procedures, such as the use of electrical modalities and ordering X-rays for patients, was not supported by current evidence (18). One of the components that greatly affects healthcare professionals' adherence to CPGs is the preferences and expectations of their patients (19). Expectations can be defined as the prevailing notion that a clinical outcome will materialize (20). Our previous study focused on investigating the patient's expectations and factors that impact adherence to physiotherapists' treatment recommendations for CLBP (21). The findings indicated that in the Indian context, patients' expectations regarding diagnosis, inclination towards passive therapies and medical care, and their behavior in seeking information are reliable indicators. One significant drawback of the aforementioned research was its quantitative approach, which made it challenging to ascertain the underlying reasons or motivations behind participants' responses. The process of developing recommendations for treatment should involve both the patient and the physiotherapist in a collaborative manner (22). Hence, the patients' abilities, experiences, anticipations, and inclinations hold significant significance in the process of treatment decision-making, alongside the clinical competence of physiotherapists.

The mechanisms by which physiotherapy interventions modify musculoskeletal pain are likely highly intricate and contingent upon various aspects associated with the physiotherapist, the patient, and the environment (23). Given that efficacy trials often overlook the aspects that influence patients' underlying beliefs and expectations (24), it is crucial to comprehend the factors that contribute to patients' adherence to physiotherapy treatment recommendations. Factors pertaining to patient expectations are correlated with clinical results, treatment satisfaction, and behavioral influence (25). Gaining a comprehensive understanding of patient expectations for the management of LBP through physiotherapy in India is crucial for devising tactics that present the most significant obstacles to the adoption of CPGs. Given the variability of patient preferences and expectations towards treatment across different cultures, our objective was to examine the expectations of Indian patients regarding physiotherapy recommendations for CLBP.

Methods

Study design and setting

This study employed a qualitative approach to elucidate the underlying meanings of quantitative data from earlier research (21). This study was conducted in the Uttar Pradesh state of India from January 2023 to December 2023. Medical practitioners are the primary initial contact clinicians, and physiotherapists do not typically operate as first-contact practitioners independently. Patients with LBP have the option to consult a physiotherapist with or without a referral from a medical practitioner. The choice of treatment center is dependent on the patient's financial situation. Individuals from lower and middle socioeconomic strata prefer free health services provided in government settings, whereas those from the upper-middle and upper classes prefer private hospitals and clinics. Physiotherapy services are currently not covered by any insurance plan, and the majority of patients personally bear the cost of these services. Insurance payment may be provided to patients who receive therapy when they are admitted as inpatients, depending on the healthcare policy. Currently, there are no existing nationwide guidelines for physiotherapy for LBP. For CLBP, non-steroidal anti-inflammatory medications, electrophysiological modalities, and exercise are the most often reported interventions (4). Occasionally, physicians admit patients experiencing more intense pain and radicular pain as inpatients; these individuals may undergo multimodal treatment, benefit from enhanced

medical supervision, and receive treatment more frequently. The treatment period for CLBP varies based on the intervention approach, often lasting approximately 12 weeks.

The Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist was utilized to facilitate the design and reporting of the qualitative research (26). Ethics approval for this study was obtained from the Integral University Ethics Committee (IIAHSR/DO/PT/2022/23). The survey's reference was established by using the guidelines for the treatment recommendations of CLBP (27) and the findings of the Ganesh et al. (21) study.

Participants

The authors employed a purposive sample technique (28) to select participants who were actively seeking care for CLBP. The participants were recruited from the authors' professional networks, as well as from their workplaces, neighboring hospitals, and physiotherapy clinics. Participants were deemed eligible if they were at least 18 years old and had experienced non-specific LBP (with no clear etiology) for a minimum of 12 weeks. In order to encompass all demographic categories, such as sex (male/female), nature of work (employed/non-employed), residence (rural/urban), socioeconomic status (upper middle/lower class), education (formal/informal), marital status (single/married), and age, consecutive registrations were included until a satisfactory number of participants in each parameter category were reached. Subsequently, the remaining categories were filled by registrations that followed one after another. This approach was employed to guarantee a sample of patients that accurately represents the population. Every participant was provided with information regarding the objective and methodologies of the study. Participants who agreed to take part in the study were required to provide both verbal and written consensus for the interview, recording, and release of anonymized data.

Interviews

The interviews were conducted from March 2023 to December 2023. In order for the participants to be able to freely express their expectations, the interviews were conducted by experienced interviewers belonging to other disciplines such as community medicine and public health. Three interviewers conducted all the interviews. While one of the interviewers conducted the interviews, the other two experts observed and supervised the process to maintain uniformity in the interview procedure by overseeing the active interviewer's approach. There was no prior relationship among the interviewers and participants.

The semi-structured interviews were conducted using a predetermined list of open-ended questions, which can be located in Table 1. This approach was chosen for its capacity to enable participants to articulate their viewpoints and expand on their personal narratives in a systematic and complete manner while still maintaining the interview's emphasis on the intended course of action (29). Furthermore, additional questions arose during the interviewers' and respondents' conversations. The interviewers queried the participants about their perspectives on the CPG's recommendations for patients with CLBP, their expectations, and the factors influencing their adherence or non-adherence to the guideline.

The interview guide and protocol subsequently underwent a pilot testing phase, with two test interviews conducted before the actual interviews. Following the initial two rounds, the interviewers exchanged input with one another in order to improve the interviewing procedure's efficiency. The presentation of the questions during one of the interviews was conversational rather than in a systematic order, which led to the exclusion of the responses from the final analysis to prevent potential bias in the results. We determined the number of interviews based on the point of saturation, a stage where we could no longer discern any new information from the interviews (30). We conducted the interviews in person, with an average duration of 1 hour. We created field notes alongside the interviews. We recorded the interviews in audio format and subsequently transcribed them verbatim. We provided the participants with the transcripts of the interviews to rectify errors and provide further remarks (31).

TABLE 1 - Interview Questions

Interview Questions	Sub-questions
What is your opinion on the use of imaging techniques, such as X-rays, CT, MRI, etc., in the evaluation and treatment of chronic low back pain (CLBP)?	• Do you think that spending money on imaging can enhance your recovery? • What are your anticipated outcomes from investigations? • Do you think the results of the investigation have influenced the outcomes of your management?
Has your healthcare provider introduced you to the CPG guidelines for LBP?	• What interventions have been recommended or administered previously and currently? • By whom were those recommendations made? • Which of these choices is most suitable for you? • Where did you provide an explanation for why one should adhere to these recommendations? • Have you been assigned exercises as part of the recommendations? • Have you been invited to actively participate in the care of your lower back pain? • Will you follow these recommendations? If not, what are the reasons for not adhering to it?

(*Continued*)

TABLE 1 - (*Continued*)

Interview Questions	Sub-questions
What is your opinion on incorporating exercises into your treatment recommendations?	• What is your opinion on incorporating exercises into your treatment recommendations? • What information have you received on the inclusion of exercises in your rehabilitation? • Which specific exercises have been recommended to you? • Are you willing to go along with these recommendations? • What factors do you believe determine the appropriateness or inappropriateness of these recommendations for you?
What is your opinion on incorporating physiotherapy into your treatment plan in accordance with guideline recommendations?	• What are your thoughts on receiving physiotherapy to treat your pain? • What are your expectations about the physiotherapy recommendations? • What makes your therapy expectations ideal for you? • What is your rationale for believing that certain recommendations you receive are inappropriate for you? • What additional elements do you believe influence your decision about accepting or not accepting therapeutic recommendations?

Data Analysis

The study utilized a four-step grounded theory approach to analyze the free-text responses (verbatims) with the recorded interviews as the unit of analysis (32). Initially, two analysts (SG and ARK, along with the interview moderator) manually and independently carried out the inductive coding process on the free-text answers provided by five randomly selected participants to establish a coding framework. The code framework underwent further refinement. Following that, the two analysts examined five random transcripts and improved the framework through further discussions. The analysts then used a combination of inductive and deductive methodologies to examine the remaining transcripts.

Next, two analysts (SG and ARK) used axial coding to generate a comprehensive list of codes (or sub-themes) by engaging in iterative discussions and reviewing the free-text responses. Furthermore, the process of selective coding was employed to establish themes by categorizing sub-themes that shared similarities (32). The analysts then systematically arranged the themes according to the study's objective. If quotes regarding diagnosis and management contradicted the guidelines, the researchers deemed them to be non-adherent to the implementation of an active approach. The researchers deemed the inclusion of additional content in the guideline suggestions compliant. This study identifies and discusses the factors that contribute to non-adherence. Table 2 displays the characteristics of the participants, and Table 3 provides a depiction of the coding tree. The initial step involved reading the transcripts and identifying phrases, sentences, or paragraphs that were pertinent to adherence to the guidelines. These identified sections were then tagged and categorized. The interview transcripts were coded using the qualitative data analysis software program Atlas.ti, version 8.4.20. Moreover, the analysts consolidated codes related to the same type of consideration into a separate category. We carefully examined the categories to identify recurring trends and establish comprehensive themes (33). SG and ARK engaged in a thorough discussion, carefully considering each step, until they reached a mutual agreement. The final coding framework was subsequently deliberated with an additional author (AK) and two outside experts possessing specialized knowledge in the content area. Finally, the analysts carried out a member validation process, where all participants reviewed and confirmed the accuracy of the identified themes, codes, code descriptions, and quotations. The authors gave each participant the opportunity to review, provide input, and provide their approval for the final draft of the findings.

Results

Study population

Fifty-five patients diagnosed with CLBP expressed their willingness to participate in this study, and a total of 37 interviews were done. Four interviewees were unable to complete their interviews, primarily due to issues with interview scheduling and other unforeseen circumstances that prevented them from fully participating. Thus, we conducted a total of 33 semi-structured individual interviews. Based on interviews 29–33, researchers determined that saturation had been achieved as only one more theme emerged.

Table 2 - Sociodemographic and clinical characteristics of participants (n = 33)

	n (%)
Age (years)*	53 (24-59)
Sex	14 women, 19 men
Years of CLBP (years)*	7.8 (1-17)
Employment	(employed-16, non-employed-6, house wife11)
Education	(college and above-14; High school and below-19)
Marital Status	(married-26; single-7)
Residence	(rural- 15; urban-18)
Household	(joint family-9; nuclear family-24)
Socio-economic status	(upper-9, middle-18, lower-6)
Nature of physiotherapy care provider	(private-12, public- 21)

*mean (range)

The qualitative analysis identified a total of six themes from the participant's perspective in identifying the factors that prevent the implementation of an active approach for LBP. The themes "physiotherapist-related factors" and "patient-related factors" revealed the highest numbers of sub-themes. Additional themes identified were guideline-related factors, institution-related factors, healthcare-related factors, and health information.

Factors influencing non-adherence

Overall, the participants interviewed expressed that their preferences, beliefs, and expectations do not align with the treatment recommendations offered by the physiotherapists. The participants discussed various factors that led to their non-adherence to the physiotherapy recommendations as outlined in the CPGs. Nevertheless, while comparing their concerns with the recommendations outlined in the guidelines, it becomes apparent that elements outside the realm of CPGs and societal/familial/cultural influence play a role in influencing their acceptability.

The six main themes from the analysis of free-text responses related to the study objective are displayed in Table 3 and explored in detail below:

1. Guideline-related factors

a. Culturally inappropriate recommendations

The majority of participants cited the CPG recommendation, including its development process, as one of the reasons they did not support the presented advice (table 3). A significant number of participants, particularly females from rural origins, expressed concerns about the widespread endorsement of exercises that may not align with their cultural norms, including their dressing attributes and the acceptance of these exercises within their community.

"These exercises prescribed are not for me, who drapes a saree" (traditional attire).- (R5, 56F, school educated, housewife, rural location)

b. Onus is on the care seeker

Another significant issue expressed regarding the guideline recommendations is that the responsibility for recovery has been entirely placed on the patients rather than the healthcare practitioners.

"The entire recommendation is made up to give healthcare providers as many reasons as possible to point us" (R8, 45M, college educated, employed, urban location)

c. Focus on biopsychosocial perspective

The guidelines propose the 'biopsychosocial approach' as a possible framework, but a minority of participants reject it. The concept has been deeply ingrained that any pathology results in pain and impairment, and extrinsic factors connected to social or geographical environments do not have a role to play.

"The doctors shift their inability to find a reason and identify a cure to reasons that are hypothetical" (R12, 29M, college educated, employed, urban location)

d. Involvement of patients with back pain in guidelines development?

Although participants do not doubt the advice of their healthcare professionals, they do not accept the notion that these recommendations are being made by professionals outside their home country without consulting or involving either the clinicians or patients from their place of residence.

"It's hard to accept that someone else has chosen what kind of care I should get." (R10, 25M, PhD student, urban location)

2. Institution related factors

a. Focus on modern equipment purchase and development of physical infrastructure

Participants attribute their non-acceptance of the recommendations provided by the CPGs to several variables associated with institutions. The main argument given is the institutional preference for investing in equipment and physical infrastructure rather than human resources.

"There is so much investment in new machines and equipment rather than the care that is provided to us" (R6, 39M, college educated, employed, urban location)

b. Focus on electrotherapy and an overcrowded department that lacks privacy

Another element of importance is the physiotherapy department's preference for administering electrotherapy versus exercise therapy prescriptions in patients. The majority of participants expressed that government facilities that are overcrowded and lack privacy are not suitable for exercise.

"Billing for exercises is cheaper compared to electrotherapy. If exercise is effective, shouldn't it cost more?" *(R3, 31F, college educated, job seeker, rural location)*

3. Patient-related factors

a. Insufficient patient engagement in goal setting and disregard for patient expectations

One of the primary issues is the failure to involve patients in setting priorities and not appealing to their treatment expectations. The participants indicated a reduced level of engagement with the physiotherapist, and they perceived themselves as passive recipients of therapy.

"Inside the department, I feel like a circus animal, and my only task is to listen to my ringmaster (physiotherapist)." (R6, 39M, College educated, employed, urban location)

b. No variability in treatment despite shifts in symptoms and disregard for lived experience

Another element is the patients' observation of the physiotherapist's reluctance to modify recommendations for treatment despite complaints of fluctuating symptoms.

"There is hardly any variation in the exercise provided, regardless of whether bending forward or backward is painful." (R20, 42F, college educated, housewife, rural location)

Table 3 - Themes and sub-themes mapped to study objective

Themes	Sub-themes	Description of link: How does the theme contribute to non-adherence to guideline recommendations?	Exemplar quotes
Guideline Related Factors	Culturally inappropriate recommendations.	Participants believe that a significant number of the recommendations should be culturally appropriate for Indian society and accepted by the community in which they live in order to be accepted.	"These exercises prescribed are not for me, who drapes a saree" (traditional attire). (R1, 55F, no formal education, housewife, rural location) "I am the head of my family and need to show a strong character. Performing recommended exercises, such as the Superman and Fire Hydrant exercises, reduces my value." (R28,58M, school education, self-employment, rural location) "The recommended exercises should be appropriate for the community in which I live. I accept walking; I reject anything else." (R2,52M, college education, employed, rural location) "At their best, these exercises are acceptable inside a hospital or department, not in the community." (R21, 47M, school educated, unskilled labor, rural location) "I've seen patients with other illnesses perform walking as an exercise. I am okay with it; I will not perform other strange exercises." (R23, 57F, no formal education, housewife, rural location)
	Onus is on the care seeker	Participants perceive that exercise and self-management have shifted the entire burden of recovery onto them, while the physiotherapist assumes a mentoring role. According to participants, this does not align with the typical duties of a healthcare professional.	"I went to seek a cure. The physiotherapist laid all responsibilities on me for recovery from lifestyle changes to self-exercise." (R8, 45M, employed, college educated, urban location) "It's easy for the healthcare provider. If I recover, his recommendations win, and if I don't, I haven't adhered to them." (R12, 29M, Employed, college-educated, urban location)
	Focus on biopsychosocial perspective	Participants question the appropriateness of treatment recommendations aimed at psychological and social factors for addressing physical pain.	"I have endured pain for many years and know from my experience that my work aggravates pain. Assigning psychological variables or changes to the brain for this behavior questions my mental strength" (R24,44M, college educated, employed, urban location)
	Involvement of patients with back pain in guidelines development?	Participants assert that a representative from one of their members should be involved in the guideline formulation process to ensure their expectations are considered, hence facilitating the appropriateness and acceptance of the recommendations.	"If at all there is a patient representative during this recommendation development process, I am confident that they must be someone who represents the privileged class rather than someone from the middle class, such as us." (R14, 44F, college educated, housewife, urban location)
Institution related factor	Focus on modern equipment purchase and development of physical infrastructure	Participants believe that acquiring the most advanced equipment improves any department's infrastructure. Clinics bill patients based on the number of electrotherapy devices they prescribe but charge less for exercise sessions. Consequently, participants consider electrotherapy to be a more appropriate recommendation than exercises.	"If exercise is all-powerful, why are there advanced machines and expensive electrotherapy sessions?' (R18, 54M, college educated, self-employment, urban location) "I see physiotherapists giving different people the same exercises to do. I see that more effort is being put into improving equipment and facilities than into giving people more advanced training. Because of this, I think that electrical therapies are better than exercises." (R11, 56M, Employed, college-educated, urban location)

Themes	Sub-themes	Description of link: How does the theme contribute to non-adherence to guideline recommendations?	Exemplar quotes
	Overcrowded department that lacks privacy	Participants perceive that numerous physiotherapy departments, particularly those at government facilities, are overcrowded and lack sufficient privacy. The participants believe such locations are unsuitable for exercising, so they avoid the recommendations.	"The government-run physiotherapy centers are too overcrowded. It is very difficult to perform exercises when someone watches you." (R13, 54F, school-educated, housewife, rural location) "There will be 3-5 people in the same room where one is receiving therapy. It's sometimes embarrassing to receive physiotherapy services." (R20, 42F, college educated, housewife, rural location)
Patient-related factors	Disregard for patient expectations	Participants felt that most clinics do not provide enough information about their diagnosis or assessment results, nor do they engage in discussions about their treatment expectations and course of care. Consequently, they tend not to accept imposed recommendations that do not align with their expectations.	"After listening to my complaint, the physiotherapist does some tests on me, the results of which are not explained." (R6, 39M, College educated, employed, urban location)
	Insufficient patient engagement in goal-setting		"The physiotherapist dictates my plan of care with junior staff. I hardly get time to discuss my progress or problems with the senior staff. My treating physio says that the senior physio is told every day about my progress" (R4, 31M, college educated, employed, urban location) "They (physiotherapists) say exercises are most important, but there has been no review of my exercise schedule since the first day.' (R22,53F, college educated, employed, urban location)
			"The physiotherapist says that self-management is the best way to go. After that, he provides a treatment plan. If sticking to it is what makes it work, should I not participate in the care planning process?"- (R2,52M, college education, employed, rural location).
	Financial factors	The participants suspect that healthcare organizations may raise the costs associated with LBP treatment under the guise of a bundled package, as specified in the recommendations. Consequently, the participants desire the advice to align with their preferences and expectations rather than favoring a multidisciplinary approach to care.	"I am confident that this presents an opportunity to increase the already high costs of physiotherapy" (R6, 39M, college educated, self-employment, urban location) "Why does my physiotherapist charge so much when he insists on self-management?" (R9,46M, no formal education, self-employment, urban location)
	No variability in treatment despite shifts in symptoms and disregard for lived experience	Despite differing complaints and symptoms, participants believe the suggested exercises are uniform and do not cater to individual needs. Consequently, they deem exercises irrelevant to their needs.	"Both me and my friend were prescribed walking as a home exercise despite having different complaints." (R24,44M, college educated, employed, urban location) "I have been continuing the same set of exercises for 8 weeks now, irrespective of my improvement or deterioration of symptoms." (R4, 31M, college educated, employed, urban location)
Physiotherapist related factors	Divergence in treatment recommendations among settings and physiotherapists	Participants think that there are significant differences in LBP treatment recommendations between healthcare settings and physiotherapists, raising the possibility that these recommendations are invalid and, therefore, unacceptable.	"If recommendations are universal, why is there so much variability in treatment prescriptions across different set-ups and among physiotherapists working in the same unit?" (R24,44M, college educated, employed, urban location)
	Absence of close supervision	Participants contend that physiotherapists have diminished their 'hands-on' care for patients under the pretext of self-management. The participants desire a human element in the treatment and dispute the self-management approach.	"The physiotherapist says I am in charge of my back and instructs me to self-exercise, which is not to my liking..... Where is the human touch in terms of treatment?" (R23, 57F, no formal education, housewife, rural location)

(*Continued*)

TABLE 3 - (*Continued*)

Themes	Sub-themes	Description of link: How does the theme contribute to non-adherence to guideline recommendations?	Exemplar quotes
	Lack of experience in managing LBP	Participants hold the belief that not all physiotherapists demonstrate equal interest and competence in treating LBP and that even if recommendations are appropriate, their implementation is not effective.	"I want physiotherapists to demonstrate the exercises so that I can learn. I have seen therapists who periodically treat children assigned to me, and they are unable to demonstrate or answer queries to satisfaction." (R18, 54M, college educated, self-employment, urban location) "The physiotherapist says my technique is wrong. But, he never demonstrates all the exercise progressions." (R17, 57F, housewife, no formal education, rural location) "I think a physiotherapist is unable to target the diseased structures with his management. That's why he pass the ball on to me, saying no structures are identified as sources of pain or refer to complex factors as sources of pain." (R12, 29M, employed, college educated, urban location)
	Ineffective communication in conveying the rationale behind recommendations	Participants suggest that physiotherapists lack the communication skills necessary to persuade them regarding the foundations for the recommendations given. .	"There is hardly an explanation by my physiotherapist why I should stretch or strengthen my muscles." (R24,44M, college educated, employed, urban location) "I don't think physiotherapists are able to explain the concepts behind the exercises well. For example, he says that if I am not interested in participating in trunk muscle training, he recommends walking or cycling as an alternative. How will exercises that target the lower limb be a substitute for trunk muscles?" (R6, 39M, college educated, employed, urban location) "My physiotherapist puts things so unclearly. He says everything is in my head, and I need to get rid of my pain from the head. My family thinks I have some psychological disorder." (R17, 57F, housewife, no formal education, rural location)
	Non-compliance with CPGs	Participants claim that physiotherapists do not comply with the stipulations set forth in the guidelines, hence making patient acceptance unlikely.	"When he is called away for other work, my physiotherapist instructs the assistant to apply electrotherapy modalities for me. Given the importance of exercises, shouldn't I be required to participate in exercises?" (R22,53F, college educated, employed, urban location)
	Recommendations are inconsistent with expectations.	Participants state that the recommended interventions are incongruent with the personal and cultural expectations of their comprehension of pain. This hinders their acceptance of active treatment strategies.	"My physiotherapist says pain while exercising is acceptable. How can pain be an answer to my pre-existing pain?" (R1, 55F, no formal education, housewife, rural location) "I want pain relief. My physiotherapist says, 'if you improve your function, the pain will come down.' Isn't my function reduced because of pain?" (R5, 56F, housewife, school educated, rural location) My daily activities involve a significant amount of physical activity. Why should I engage in other exercises?" (R17, 57F, housewife, no formal education, rural location) "There are days when low back pain severely limits my function, and I do not wish to work. If exercise is management, this is not for me." (R15,33F, college education, rural location) "Statements like "low back pain has no source" and "exercises are the best pain management" would have led my entire family to believe I had been lying for all these years." (R13, 54F, school educated, housewife, rural location) "Going to physiotherapy clinics for electrotherapy is a relaxing experience. I get to know and interact with other people. Exercises do not provide this experience." (R15,33F, college education, rural location)

Themes	Sub-themes	Description of link: How does the theme contribute to non-adherence to guideline recommendations?	Exemplar quotes
Healthcare related factors	Not on par with the medical practitioners	Participants perceive that physiotherapists do not occupy a central role within the healthcare system or the hierarchy of back pain management. Given their limited capacity to order investigations and establish diagnoses, the participants regard the physiotherapist's position as supportive, leading them to believe that their recommendations may not be strictly adhered to.	"The physiotherapists take orders from medical practitioners, and I will listen to their (medical doctors') recommendations, which are easy to comply." (R8, 45M, employed, college educated, urban location)
	Lack of consistency in recommendations among healthcare providers	The recommendations provided by physiotherapists contradict those of other healthcare providers, and this casts doubt on the validity of physiotherapists' recommendations.	"My physiotherapist says it's okay for me to bend forward; my doctor says no to bending forward. How can two different recommendations be appropriate?" (R18, 54M, college educated, self-employment, urban location) "Other health care providers do not appear to endorse exercises as strongly as a physiotherapist" (R14, 44F, college educated, housewife, urban location)
	No flexibility in treatment recommendations despite investigation outcomes.	Participants are of the opinion physiotherapists hesitate to alter treatment recommendations, leading to skepticism regarding the validity of the recommendations based on the investigation's results.	" My physiotherapists shows not much enthusiasm in interpreting these outcomes." (R14, 44F, college educated, housewife, urban location) "My treatment remains the same irrespective of the MRI findings." (R22, 53F, college educated, employed, urban location)
	Availability of alternative healthcare options	Participants believe that there are many alternative treatment methods available and that following physiotherapists' recommendations is not necessary.	"When we do not get better with the treatments and therapies provided in hospitals and clinics, we can rely on local healers and traditional therapies. There is no pressure on us to engage in exercises, which we do not prefer". (R32,58M, school education, employed, rural location)
	Negative expectations concerning recovery	Participants anticipate minimal positive effects from physiotherapists' recommendations due to the prolonged duration of their symptoms and their past experiences with various treatments.	"When they can't even confirm the possible reasons for pain, how do they propose a management plan? " (R32, 58M, school education, employed, rural location) "After such a long treatment, I am convinced there is no cure for my pain. All I expect is a non-aggravation of pain and some treatment that can ease my pain" (R22, 53F, college educated, employed, urban location)
Health information	Impact of media on healthcare options	Participants indicate that several media sources and healthcare providers disseminate supplementary information on exercise and other physiotherapeutic methods, making it challenging to determine appropriateness.	"None of the techniques my physiotherapist demonstrates resemble what I see on social media" (R10, 25M, PhD student, urban location) "I wonder if my physiotherapist possesses the same knowledge as the person who provides information on social media." (R10, 25M, PhD student, urban location) "There is so much conflicting information provided by healthcare practitioners that I do not know whom to trust" (R8, 45M, employed, college educated, urban location)

c. Financial factors

The financial aspect is a significant factor that causes the participants to disagree with the recommendations. When prescriptions are based on international guidelines, the participants fear that treatment costs will increase.

"Due to the lack of standardization in healthcare costs in India, we worry that private healthcare organizations may escalate costs for therapy under the cover of international prescriptions." *(R18, 54M, college educated, self-employment, urban location)*

4. Physiotherapist related factors

a. Divergence in treatment recommendations among settings and physiotherapists

Most participants identified multiple issues associated with physiotherapists' care as potential causes of non-adherence. The participants mentioned the inconsistency among physiotherapists when it comes to treatment prescriptions. The participants observed that variations exist not only among physiotherapists employed in different healthcare settings but also among therapists working within the same institutions. Given the wide range of differences, the participants are skeptical about the validity of broad recommendations.

"*When it comes to physiotherapists, everything varies, from assessment to treatment prescriptions to home education*." *(R4, 31M, college educated, employed, urban location)*

b. Absence of close supervision

Another frequent critique is that physiotherapists do not offer regular supervision for exercise sessions, even though patients are expected to self-manage their symptoms.

"*My physiotherapist once demonstrated all the exercises; nobody supervised me after that." (R17, 54M, college educated, self-employment, urban location)*

c. Lack of experience in managing low back pain

The participants highlighted the insufficient availability of a sufficient number of physiotherapists, particularly those with competence in managing LBP. Another reason is the participants' gauge that physiotherapists who provide treatment do not refer them to other physiotherapists for their opinion or advice.

"*My physiotherapist dedicates the majority of her time to patients with paralysis. I hardly get any attention.*" *(R14, 44F, college educated, housewife, urban location)*

"*My physiotherapist does not always conduct treatment sessions directly. Students or attendants can take on that role depending on the physio's availability.*" *(R17, 54M, college educated, self-employment, urban location)*

d. Ineffective communication in conveying the rationale behind recommendations

In addition, they observe that physiotherapists are unable to adequately explain the rationale behind the treatment recommendations they make, even when those suggestions are listed in CPGs.

"*When I say exercises hurt, my physiotherapist asks me to continue them, saying painful exercises are better than painless exercises. I don't understand how*?" *(R19, 54F, school educated, self-employment, rural location)*

Participants believe that physiotherapists fail to effectively communicate the results of assessments or provide treatment recommendations as outlined in the CPG in a way that patients may easily understand.

"My physiotherapist says I will get confused if he explains the causes of the symptoms." (R20, 42F, college educated, housewife, rural location)

e. Expensive recommendations with little benefit

Participants also remarked on the link between the recommendations made by the CPGs and the financial advantages for the treatment clinics or physiotherapists.

"Anything endorsed overseas costs more in India." (R18, 54M, college educated, self-employment, urban location)

f. Recommendations are inconsistent with expectations.

Furthermore, the inability to meet expectations is the primary driver of non-adherence. There is a lack of consistency between patients and physiotherapists in meeting common goals.

"When the physician referred me to physiotherapy, he said I would be provided with electromassage. That's what I expect." (R13, 54F, school educated, housewife, rural location)

"Physiotherapy should be relaxing for aching muscles and joints. Not aggravate it." (R23, 57F, no formal education, housewife, rural location)

Participants also express the view that therapy recommendations, such as exercises, put them in a difficult situation because their complaints of pain and discomfort seem unreal.

"My family believes that if I can exercise, I can do all of my work at home. Nobody will now believe I have real pain." (R23, 57F, no formal education, housewife, rural location)

5. Healthcare related factors

a. Not on par with the medical practitioners

Participants believe that physiotherapists in India are not positioned at the highest level of the clinical hierarchy and express reservations regarding their proficiency in the field of pain management.

" Physiotherapists take orders from doctors like any other allied health worker." (R4, 31M, college educated, employed, urban location)

a. Lack of consistency in recommendations among healthcare providers

Other participants contend the recommendations provided by physiotherapists and other healthcare practitioners are contradictory.

"My doctor says to wear a belt (lumbar corset), and my physiotherapist says there is no advantage to belts." (R17, 54M, college educated, self-employment, urban location)

b. No considerations for laboratory and radiographic investigations

During the interview, participants expressed concern that physiotherapists often disregard radiological findings and insist on making the same recommendations.

"I am not sure why I have to spend so much on investigations if it's not going to change my physiotherapy treatment." (R14, 44F, college educated, housewife, urban location)

c. Availability of alternative healthcare options

Participants indicate that there are numerous alternative treatment options available in India, which may render it unnecessary to adhere to uncomfortable recommendations.

"*There is no pressure on us to engage in exercises, which we do not prefer." (R1, 55F, no formal education, housewife, rural location)*

6. Health Information

a. Impact of media on healthcare options

The existence of healthcare professionals on diverse social media platforms, where many showcase their approaches and offer guidance and treatment suggestions, leads to significant confusion when it comes to choosing treatment choices for participants.

"The patients show instant results on the videos I see on social media. Such techniques are not listed as recommendations" (R10, 25M, PhD student, urban location)

"From nutrition to surgery, there are so many treatment choices. I am uncertain about whom to contact and which treatment options are good for me." (R6, 39M, College educated, employed, urban location)

Discussion

We conducted a qualitative analysis involving 33 participants to examine their perspectives on the physiotherapy recommendations for CLBP, as recommended by the CPGs. To enhance compliance with recommendations, it is essential to comprehend the underlying causes of non-adherence. The main findings of this research reveal that the determinants of non-adherence to guideline recommendations are multifaceted, involving various stakeholders such as the patient, physiotherapist, and institution, as well as the disregard of cultural, societal, and familial influences during the development of guidelines, among other healthcare-related factors.

Although CPGs offer us evidence, the relevance of this knowledge to particular patients in low-resource settings remains questionable. This ambiguity stems from the necessity to customize treatments based on the distinct attributes and requirements of individual patients. The skepticism shown by participants regarding the guideline recommendations may be ascribed to the distinctive impact of gender norms, familial support, religious practices, and communal influence in India. Indian cultures are frequently regarded as stringent, religious, familial, and abundant in traditions (34). Conventional gender roles can restrict women's participation in exercise, as societal expectations frequently emphasize home responsibilities and the maintenance of cultural traditions above exercise (35). Likewise, individuals, particularly females from disadvantaged socioeconomic backgrounds, bear greater family duties, which may considerably influence their exercise options. It has been found that individuals' attitudes and behaviors are influenced by societal norms related to the acceptability of specific exercises (36). Individuals are increasingly inclined to integrate yoga or meditation into their routines owing to its cultural acceptance and spiritual advantages (34).

Patients have the belief that if they are able to visually perceive a specific cause for their pain, their healthcare practitioners will possess a greater understanding of how to effectively address the issue. Participants anticipated receiving a definitive diagnosis, and, as a result, most of them expected their results to be meticulously analyzed and appropriate management to be prepared appropriately. This illustrates the prevailing notion that pain is solely a biological phenomenon and highlights the lack of widespread understanding regarding the assessment and treatment of LBP. This concurs with previous research indicating that participants' yearning for a diagnosis may stem from a need for reassurance regarding the source of their symptoms (37). Recent research indicates that pathoanatomical diagnoses exhibit a weak correlation with symptoms and outcomes in non-specific LBP (38). Patients appear oblivious to the weak association between imaging findings and symptomatology in lower back pain. There is a need to provide additional education regarding the indications for imaging tests in order to effectively manage expectations and influence patients' beliefs.

It is remarkable that individuals did not regard regular evaluation and physical examination as a crucial component of assessing LBP. This viewpoint is accorded with patients' belief that physical medicine and rehabilitation physicians are the most qualified providers for diagnosing and treating LBP, possibly attributable to their capacity to demand imaging studies (39). Further, this opinion may be associated with the view that participants hold the belief that their physiotherapists possess limited abilities and expertise in evaluating and treating LBP. The participants' perspectives align with findings from another study indicating that patients doubt physiotherapists' ability to diagnose back pain, highlighting the necessity for physiotherapists to enhance confidence in their knowledge regarding the therapy and diagnosis of LBP (40).

Patients with CLBP may have viewed their bodies as 'broken machines' due to the biomedical interpretation of their pain (41), leading them to believe that they should avoid exercises and functional movements to prevent further structural damage (42). The patients' preference for passive therapy can be attributed to the perceived advantages associated with the utilization of advanced equipment, the immediate and temporary alleviation that these treatments can provide, past experiences with previous care, or their existing beliefs and assumptions about CLBP, as well as the passive nature of decision-making. While various passive treatment approaches are suggested in CPGs (27), it is necessary to investigate if they can effectively contribute to the development of a therapeutic alliance and the establishment

of trust between patients and physiotherapists. It is crucial to acknowledge that exercises may not be suitable for all patients with LBP at every stage of their treatment, regardless of the advice provided in CPGs.

The ambiguity surrounding the most effective treatment for CLBP might result in people attending clinic appointments that fail to match their expectations, ultimately leading to inadequate adherence to the prescribed treatment. Our study also revealed that participants expected to receive appropriate care from a competent professional in order to receive the appropriate treatment, notwithstanding their lack of knowledge regarding which expert could assist them or who the most competent care provider for their condition was. This finding aligns with a recent qualitative study that revealed individuals suffering from LBP experience ambiguity over the appropriate course of treatment they should pursue (43). Participants expressed skepticism over whether consulting a physiotherapist was the best course of action to begin their treatment. The participants had a lack of clarity on the role of the physiotherapy, as they expressed a preference for receiving passive treatments, while failing to grasp the significance of exercise. Participants also express cynicism regarding the efficacy of self-management methods despite multiple research highlighting the significance of these interventions for managing LBP (44). Patients' ability to navigate the healthcare system could impede their access to care and results (40). Thus, further information is required to address patients' perceptions and expectations regarding the role of physiotherapy and what constitutes effective evidence-based practice for the treatment of LBP.

The relationship between the clinician and the patient is a crucial aspect of delivering healthcare and ensuring the impact of therapy. A prior study has demonstrated a significant correlation between the role and relationships of healthcare providers with patients and the likelihood of patients adhering to their treatment plan (45). For example, elderly individuals appear to prioritize engagement and rapport with the physiotherapist over alterations in symptoms when considering therapy suggestions (46). The effectiveness of provider-patient communication and patient-centered practice, which includes shared decision-making and hearing input from patients on treatment goals, has been identified as crucial factors in enhancing patients' overall adherence to treatment and quality of care (47). Therefore, it is crucial for healthcare providers to actively involve patients in treatment and diagnosis instructions (48). Continuity of care is a crucial element in healthcare. To enhance care continuity, it is vital for healthcare practitioners to offer clear guidelines and directions regarding a patient's treatment plan (49) and to respect patient appointments. Enhancing and cultivating enduring provider-patient relationships would be advantageous for patients. Therefore, the act of providing patient education is highly valuable in encouraging compliance with healthcare advice (50).

The participants' fear of increased expenditures linked to guideline recommendations is similar to the copayments incurred by patients during physiotherapist specialist consultations in the United States, potentially leading to a substantial rise in their out-of-pocket expenses (40). Multiple studies on private providers have revealed that patient overcharging is rampant in India (51), primarily due to insufficient regulation (52) and the higher investment in infrastructure compared to public facilities (53). Modifications are also required on the part of healthcare infrastructure and/or policymaking in order to be cost-effective.

The impact of the information disseminated by media in influencing perspectives has the capacity to create unjustified expectations (54) and adverse emotional reactions. The acquisition of new information or knowledge from different media outlets can either increase patients' faith in their healthcare provider and improve their understanding of their health situation (55) or have the opposite effect. Therefore, we suggest that it is crucial to meet the information needs of patients and take steps to be vigilant in monitoring magazines and websites that distribute health-related information, making sure that the material is supported by scientific research. It is crucial for healthcare practitioners, experts, consumers, and researchers to have an open conversation in order to close the gap between the treatment-related health information that patients obtain from the internet and other sources and the evidence that is now available. Given that participants see participating in exercises as not being part of the treatment culture, it may be deduced that greater levels of behavioral and lifestyle modifications will be linked to better levels of overall adherence. It is essential for healthcare practitioners to possess strong interpersonal and communication skills in order to effectively engage in joint decision-making. Given the limited infrastructure and human resources for rehabilitation services, policymakers should prioritize investment in exercise promotion activities (56) and effective pain education strategies (57). It is important for policymakers to ensure that healthcare professionals receive evidence-based healthcare that is relevant to the complex healthcare system, which includes a busy practice and limited clinical time with patients. Furthermore, establishing a robust therapeutic alliance and allocating additional time for patient interaction will enable physiotherapists to develop a treatment strategy that aligns with evidence-based guidelines.

Patients reported receiving conflicting messages from various healthcare providers regarding the management of LBP, resulting in uncertainty regarding whose advice to adhere to (58). Given the inconsistencies in CPG treatment recommendations, we suggest all first-contact healthcare professionals advocate for exercise as an intervention in order to minimize variation in treatment recommendations. Medical professionals should act as gatekeepers, deciding on the precise type and quantity (number of treatment sessions and frequency per week) of physiotherapy services to prescribe (59). There is also an immediate need to address the communication gap between patients, physicians, and physiotherapists.

Another notable finding in the study is the patient's acknowledgment of heightened treatment expenses linked to electrotherapy prescriptions, notwithstanding their lack of preference for exercises and self-management initiatives. Similarly, despite not favoring self-management strategies, participants expressed a desire for consultation on goal setting and treatment preferences. Patients' unmet

requirements for healthcare services, their expectation of rapid cures through more pain-centered passive treatments, their perception of a lack of empathy from healthcare personnel, and the inadequacy of the offered services may all contribute to this situation (60). Future studies should look at the divergent viewpoints of participants regarding these topics.

Interestingly, the findings of this study conducted in India exhibited numerous similarities to a study conducted in Belgium (61) about patient-reported impediments to the adoption of guidelines for an active physiotherapeutic approach to LBP in clinical practice. Patients at both locations preferred passive treatments, faced difficulties in understanding the precise objectives of the therapy, and reported issues in assuming responsibility and adhering to exercises. The stigmatization of psychological issues is prevalent, and individuals reported receiving contradictory guidance from various healthcare practitioners. Barriers to the adoption of CPGs in both developed and LMICs must be taken into account during the development of these guidelines.

The strengths of the study are as follows. The interview guide provided a comprehensive understanding of the anticipated outcomes and first-hand encounters of those living with CLBP. Due to the prolonged length of pain, we were able to gain a more comprehensive understanding of the diverse range of experiences among participants. The methodology of analyzing interviews was intended to ensure that the findings were derived only from the data and not influenced by the researcher's perception. Not including a physiotherapist in the interview team may have created an environment where participants felt comfortable and were able to provide their comments without feeling pressured or coerced. A noteworthy advantage of our study is the inclusion of people receiving treatment from both government and private establishments.

It is important to take into account the limitations of this study. The patients' interviews were contingent upon the accuracy of their recollection of events that occurred during their physiotherapy sessions. Their perceptions may have been influenced by the positive or negative progression of their symptoms over time. Given the extensive duration of treatment, it is conceivable that participants may not have recorded all facets of the patient experiences. Furthermore, we did not gather data regarding patients' assessment of the consultations they got from other healthcare providers. The generalizability of the research findings from India to other nations or circumstances may be limited due to potential changes in cultural backgrounds or contextual factors.

Conclusion

A patient's perception and expectation of a physiotherapy intervention recommendation might be influenced by their comprehension of CLBP, as well as the therapist's instruction and implementation of the treatment. The perception of exercise can be significantly influenced by the social environment and culture. Indian physiotherapists should consider suggesting active interventions that are culturally appropriate, and developing communication skills could enhance their ability to manage patient expectations that contradict guideline suggestions, hence potentially improving adherence to guidelines.

Acknowledgments

This work is acknowledged under Integral University manuscript number IU/R&D/2024-MCN0003047.

Disclosures

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

Financial support: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contributions: Conceptualization: SG; Methodology: SG, ARK, AK; Data Collection Supervision: SG, ARK, AK; Formal analysis and investigation: SG, ARK, AK; Writing: SG; Writing - review and editing: All authors read and approved the final manuscript

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

References

1. Ferreira ML, de Luca K, Haile LM, et al. GBD 2021 Low Back Pain Collaborators. Global, regional, and national burden of low back pain, 1990-2020, its attributable risk factors, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. Lancet Rheumatol. 2023;5(6):e316-e329. CrossRef PubMed
2. Clark S, Horton R. Low back pain: a major global challenge. Lancet. 2018;391(10137):2302. CrossRef PubMed
3. Lorig KR, Ritter P, Stewart AL, et al. Chronic disease self-management program: 2-year health status and health care utilization outcomes. Med Care. 2001;39(11):1217-1223. CrossRef PubMed
4. Sharma S, Pathak A, Parker R, et al. How low back pain is managed-a mixed methods study in 32 countries. Part 2 of Low Back Pain in Low- and Middle-Income Countries Series. J Orthop Sports Phys Ther. 2024;54(8):560-572. CrossRef PubMed
5. Foster NE, Anema JR, Cherkin D, et al. Lancet Low Back Pain Series Working Group. Prevention and treatment of low back pain: evidence, challenges, and promising directions. Lancet. 2018;391(10137):2368-2383. CrossRef PubMed
6. Kamper SJ, Yamato TP, Williams CM. The prevalence, risk factors, prognosis and treatment for back pain in children and adolescents: an overview of systematic reviews. Best Pract Res Clin Rheumatol. 2016;30(6):1021-1036. CrossRef PubMed
7. Hayden JA, Ellis J, Ogilvie R, et al. Exercise therapy for chronic low back pain. Cochrane Database Syst Rev. 2021;9(9):CD009790. PubMed
8. Hall AM, Scurrey SR, Pike AE, et al. Physician-reported barriers to using evidence-based recommendations for low back pain in clinical practice: a systematic review and synthesis of qualitative studies using the Theoretical Domains Framework. Implement Sci. 2019;14(1):49. CrossRef PubMed
9. Beyera GK, O'Brien J, Campbell S. Healthcare utilization for low back pain: a systematic review and meta-analysis of population-based observational studies. Rheumatol Int. 2019;39(10):1663-1679. CrossRef PubMed

10. Miller TA. Health literacy and adherence to medical treatment in chronic and acute illness: a meta-analysis. Patient Educ Couns. 2016;99(7):1079-1086. CrossRef PubMed
11. DiMatteo MR. Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. Med Care. 2004;42(3):200-209. CrossRef PubMed
12. Burke LE, Ockene IS. Compliance in Healthcare and Research. Furtura; 2001. Online
13. Iuga AO, McGuire MJ. Adherence and health care costs. Risk Manag Healthc Policy. 2014;7:35-44. CrossRef
14. Buszko K, Obońska K, Michalski P, et al. The adherence scale in chronic diseases (ASCD). The power of knowledge: the key to successful patient—health care provider cooperation. Med Res J. 2016;1(1):37-42. CrossRef
15. Husted M, Rossen CB, Jensen TS, et al. Adherence to key domains in low back pain guidelines: a cross-sectional study of Danish physiotherapists. Physiother Res Int. 2020;25(4):e1858. CrossRef PubMed
16. de Souza FS, Ladeira CE, Costa LOP. Adherence to back pain clinical practice guidelines by brazilian physical therapists: a cross-sectional study. Spine. 2017;42(21):E1251-E1258. CrossRef PubMed
17. Bahns C, Happe L, Thiel C, et al. Physical therapy for patients with low back pain in Germany: a survey of current practice. BMC Musculoskelet Disord. 2021;22(1):563. CrossRef
18. Ganesh GS, Khan AR, Khan A. A survey of Indian physiotherapists' clinical practice patterns and adherence to clinical guidelines in the management of patients with acute low back pain. Musculoskeletal Care. 2023;21(2):478-490. CrossRef PubMed
19. Schers H, Braspenning J, Drijver R, et al. Low back pain in general practice: reported management and reasons for not adhering to the guidelines in the Netherlands. Br J Gen Pract. 2000;50(457):640-644. PubMed
20. Wiles R, Cott C, Gibson BE. Hope, expectations and recovery from illness: a narrative synthesis of qualitative research. J Adv Nurs. 2008;64(6):564-573. CrossRef PubMed
21. Ganesh GS, Khan AR, Khan A, et al. Factors predicting non-adherence to treatment recommendations for patients with chronic low back pain in India: a cross-sectional survey. J Manipulative Physiol Ther. 2023;46(5-9):239-253. CrossRef PubMed
22. World Health Organization. WHO Handbook for Guideline Development. 2nd ed. 2014 Online (Accessed July 2024)
23. Whyte J, Hart T. It's more than a black box; it's a Russian doll: defining rehabilitation treatments. Am J Phys Med Rehabil. 2003;82(8):639-652. CrossRef PubMed
24. Hidalgo B, Detrembleur C, Hall T, et al. The efficacy of manual therapy and exercise for different stages of non-specific low back pain: an update of systematic reviews. J Man Manip Ther. 2014;22(2):59-74. CrossRef PubMed
25. Bialosky JE, Bishop MD, Cleland JA. Individual expectation: an overlooked, but pertinent, factor in the treatment of individuals experiencing musculoskeletal pain. Phys Ther. 2010;90(9):1345-1355. CrossRef PubMed
26. 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
27. George SZ, Fritz JM, Silfies SP, et al. Interventions for the management of acute and chronic low back pain: revision 2021. J Orthop Sports Phys Ther. 2021;51(11):CPG1-CPG60. CrossRef PubMed
28. Bryman A. Social Research Methods. Oxford Univ Pr Canada; 2013. Online (Accessed July 2024)
29. Dicicco-Bloom B, Crabtree BF. The qualitative research interview. Med Educ. 2006;40(4):314-321. CrossRef PubMed
30. Dworkin SL. Sample size policy for qualitative studies using in-depth interviews. Arch Sex Behav. 2012;41(6):1319-1320. CrossRef PubMed
31. Shah A. Using data for improvement. BMJ. 2019;364:l189. CrossRef PubMed
32. Ayre J, McCaffery KJ. Research note: thematic analysis in qualitative research. J Physiother. 2022;68(1):76-79. CrossRef PubMed
33. Sutton J, Austin Z. Qualitative research: data collection, analysis, and management. Can J Hosp Pharm. 2015;68(3):226-231. CrossRef PubMed
34. Ramanathan S, Crocker PR. The influence of family and culture on physical activity among female adolescents from the Indian diaspora. Qual Health Res. 2009;19(4):492-503. CrossRef PubMed
35. Farver JAM, Bhadha BR, Narang SK. Acculturation and psychological functioning in Asian Indian adolescents. Soc Dev. 2002;11(1):11-29. CrossRef
36. Rio CJ, Saligan LN. Understanding physical activity from a cultural-contextual lens. Front Public Health. 2023;11:1223919. CrossRef
37. Holt N, Pincus T, Vogel S. Reassurance during low back pain consultations with GPs: a qualitative study. Br J Gen Pract. 2015;65(639):e692-e701. CrossRef PubMed
38. George SZ, Lentz TA, Beneciuk JM, et al. Framework for improving outcome prediction for acute to chronic low back pain transitions. Pain Rep. 2020;5(2):e809-e809. CrossRef PubMed
39. Sharpe JA, Martin BI, Fritz JM, et al. Identifying patients who access musculoskeletal physical therapy: a retrospective cohort analysis. Fam Pract. 2021. CrossRef PubMed
40. Sharpe JA, Thackeray A, Fritz JM, et al. Patients' use of physical therapy for lower back pain: a qualitative study. Musculoskelet Sci Pract. 2021;56:102468. CrossRef PubMed
41. Setchell J, Costa N, Ferreira M, et al. Individuals' explanations for their persistent or recurrent low back pain: a cross-sectional survey. BMC Musculoskelet Disord. 2017;18(1):466. CrossRef PubMed
42. Bunzli S, Smith A, Watkins R, et al. What do people who score highly on the tampa scale of kinesiophobia really believe? A mixed methods investigation in people with chronic non-specific low back pain. Clin J Pain. 2015;31(7):621-632. CrossRef PubMed
43. Costa N, Mescouto K, Dillon M, et al. The ubiquity of uncertainty in low back pain care. Soc Sci Med. 2022;313:115422. CrossRef PubMed
44. Longtin C, Décary S, Cook CE, et al. What does it take to facilitate the integration of clinical practice guidelines for the management of low back pain into practice? Part 1: A synthesis of recommendation. Pain Pract. 2021;21(8):943-954. CrossRef PubMed
45. Bussell JK, Cha E, Grant YE, et al. Ways health care providers can promote better medication adherence. Clin Diabetes. 2017;35(3):171-177. CrossRef PubMed
46. Maiers M, Vihstadt C, Hanson L, et al. Perceived value of spinal manipulative therapy and exercise among seniors with chronic neck pain: a mixed methods study. J Rehabil Med. 2014;46(10):1022-1028. CrossRef PubMed
47. Tofan G, Bodolica V, Spraggon M. Governance mechanisms in the physician-patient relationship: a literature review and conceptual framework. Health Expect. 2013;16(1):14-31. CrossRef PubMed
48. Kamimura A, Weaver S, Armenta B, et al. Patient centeredness: the perspectives of uninsured primary care patients in the United States. Int J Care Coord. 2019;22(1):19-26. CrossRef

49. Kamimura A, Panahi S, Ahmmad Z, et al. Continuity of care: perspectives of uninsured free clinic patients. J Patient Exp. 2019;6(4):305-310. CrossRef PubMed
50. Jin J, Sklar GE, Min Sen Oh V, et al. Factors affecting therapeutic compliance: a review from the patient's perspective. Ther Clin Risk Manag. 2008;4(1):269-286. CrossRef PubMed
51. Hooda SK. Growth of formal and informal private healthcare providers in India: structural changes and implications. J Health Care Finance. 2017;44(2):FALL Online
52. Prinja S, Bahuguna P, Pinto AD, et al. The cost of universal health care in India: a model based estimate. PLoS One. 2012; 7(1):e30362. CrossRef PubMed
53. Basu S, Andrews J, Kishore S, et al. Comparative performance of private and public healthcare systems in low- and middle-income countries: a systematic review. PLoS Med. 2012;9(6): e1001244. CrossRef PubMed
54. Hummer M, Fleischhacker WW. Ways of improving compliance. In: Lader M, Naber D, eds. Difficult Clinical Problems in Psychiatry. Martin Dunitz; 1999:229-238.
55. Anderson JG, Rainey MR, Eysenbach G. The impact of CyberHealthcare on the physician-patient relationship. J Med Syst. 2003;27(1):67-84. CrossRef PubMed
56. Slade SC, Patel S, Underwood M, et al. What are patient beliefs and perceptions about exercise for non-specific chronic low back pain? A systematic review of qualitative studies. Clin J Pain. 2014;30(11):995-1005. CrossRef PubMed
57. Boutevillain L, Dupeyron A, Rouch C, et al. Facilitators and barriers to physical activity in people with chronic low back pain: a qualitative study. PLoS One. 2017;12(7):e0179826. CrossRef PubMed
58. Hoffmann TC, Del Mar CB, Strong J, et al. Patients' expectations of acute low back pain management: implications for evidence uptake. BMC Fam Pract. 2013;14:7. CrossRef
59. Brechtel T. Heil- und Hilfsmittel. In: Thielscher C, ed. Handbuch Medizinökonomie | Grundlagen Und System Der Medizinischen Verorgung. Vol 1. Springer Gabler; 2021:365-373. CrossRef
60. Chou L, Ranger TA, Peiris W, et al. Patients' perceived needs for medical services for non-specific low back pain: a systematic scoping review. PLoS One. 2018;13(11):e0204885. CrossRef
61. Bogaert L, Brumagne S, Léonard C, et al. Physiotherapist- and patient-reported barriers to guideline implementation of active physiotherapeutic management of low back pain: a theory-informed qualitative study. Musculoskelet Sci Pract. 2024;73:103129. CrossRef PubMed

Arch Physioter 2024; 14: 170-181
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3228
ORIGINAL RESEARCH ARTICLE

Effect of Pilates exercises on symptoms of irritable bowel syndrome in women: a randomized controlled trial

Doaa Mohammed Mahmoud Allam[1], Heba Ali Abdel Ghaffar[2], Amira Mostafa Elshamy[3], Mohamed Abdelmotaal Safa[4], Ahmad Mahdi Ahmad[2]

[1]Department of Physical Therapy for Internal Medicine, Faculty of Physical Therapy, Kafr El-Sheikh University, Kafr el-Sheikh - Egypt
[2]Department of Physical Therapy for Cardiovascular and Respiratory Disorders, Faculty of Physical Therapy, Cairo University, Giza - Egypt
[3]Department of Medical Biochemistry, Faculty of Medicine, Tanta University, Tanta - Egypt
[4]Department of Internal Medicine, Faculty of Medicine, Tanta University, Tanta - Egypt

ABSTRACT

Objective: The treatment of irritable bowel syndrome (IBS) is challenging, calling for therapeutic strategies other than pharmacological treatment. Therefore, this study aimed to investigate the effects of Pilates exercises on IBS symptoms and severity, frequency of complete spontaneous bowel movements, fatigue, anxiety, depression, and body weight in women with IBS.

Methods: Sixty women with IBS, aged 20-45, completed this study. They were randomly assigned to two equal groups: a study group (n = 30) and a control group (n = 30). The study group received an 8-week Pilates exercise program (2 sessions per week) in addition to dietary advice, while the control group received dietary advice only. Inclusion criteria were women, IBS diagnosed based on Rome IV Diagnostic criteria, constipation-predominant IBS, and moderate to severe IBS. The outcome measures were the IBS severity scoring system (IBS-SSS), the frequency of complete spontaneous bowel movements, the modified fatigue impact scale (MFIS), hospital anxiety and depression (HADS) scale, and body weight (BW).

Results: The study group showed more significant improvements than the control group in total IBS-SSS score (Cohen d = 0.73, $p < 0.001$), frequency of complete spontaneous bowel movements (Cohen d = 0.50, $p < 0.001$), total MFIS score (Cohen d = 0.74, $p < 0.001$), anxiety (Cohen d = 0.56, $p < 0.001$), and depression (Cohen d = 0.64, $p < 0.001$). The study group also showed a significant reduction in body weight compared to baseline ($p < 0.05$). The control group showed significant improvements in all outcomes, except body weight, compared to baseline ($p < 0.05$).

Conclusion: Pilates exercises, used in addition to dietary advice, may significantly improve IBS symptoms and severity, frequency of complete spontaneous bowel movements, and alleviate fatigue, anxiety, and depression moderately more than dietary advice alone in women with constipation-predominant IBS. Nevertheless, dietary advice alone may also significantly improve these outcomes in this cohort.

Keywords: Anxiety/depression, Fatigue, Irritable bowel syndrome, Pilates exercises, Women

What is already known about this topic

- *Pilates exercises can relieve symptoms of fatigue, anxiety, and depression in several populations.*
- *There is a gap in the literature concerning the potential effect of Pilates exercises on the gastrointestinal and extra-gastrointestinal symptoms of irritable bowel syndrome (IBS).*

What this study adds

- *This is the first study to reveal the effectiveness of adding Pilates exercises to dietary advice as a lifestyle therapy in women with IBS.*
- *Pilate exercises added to dietary advice may offer more relief for IBS symptoms and clinically meaningful changes in IBS severity than dietary advice alone in women.*
- *Pilate exercises with dietary advice may increase the frequency of complete spontaneous bowel movements and reduce fatigue, anxiety, and depression to a greater extent than dietary advice alone in women with IBS.*

Received: July 26, 2024
Accepted: December 5, 2024
Published online: December 31, 2024

Trial registration
ClinicalTrials.gov, NCT05832801

Corresponding author:
Ahmad Mahdi
email: ahmed.mahdy@pt.cu.edu.eg

Introduction

The incidence of irritable bowel syndrome (IBS) is 4.1% worldwide (1), with women being more affected than men (2). IBS with constipation is the most prevalent subtype of IBS and most commonly affects women with overweight and obesity (3). Frequent abdominal pain and annoying bloating associated with irregular bowel habits are the most common symptoms accompanying IBS (4). The frequency and

extent of pain vary according to changes in bowel habits (4). Non-gastrointestinal symptoms such as fatigue, anxiety, and depression can also be present (5). The underlying cause of IBS is not yet clear but is known to be multifactorial, involving gut-brain axis dysfunction (6). Dysregulation of the gut-brain axis, a bidirectional pathway connecting the brain to the gut via the vagus nerve, may contribute to the perception of abdominal pain, discomfort, anxiety, and depression by the brain based on sensory input from the gut (3,7). Research has shown that managing IBS symptoms can be challenging (8). Since no established medical treatment can change IBS's natural course (9), management plans have recently focused on pathophysiology and symptom relief (10). Dietary intervention is the first target therapy to alleviate IBS symptoms, as consuming particular foods can exacerbate symptoms (11). Additionally, Pilates exercises can have a potential role in relieving extra-gastrointestinal symptoms such as fatigue, anxiety, and depression (12).

To our knowledge, no prior research has assessed Pilates exercises' effectiveness in managing patients with IBS. Thus, this study aimed to determine the effects of Pilates exercises in combination with dietary advice on IBS symptoms and severity, the frequency of complete spontaneous bowel movements, fatigue, anxiety and depression, and body weight compared to dietary advice alone in women with IBS. We hypothesized that Pilates exercises might influence the gastrointestinal symptoms of IBS based on the fact that they comprise a variety of postures combined with deep breathing, which can stimulate the vagal tone (13), thus optimizing the gastrointestinal tract (GIT) function and relieving constipation and associated symptoms. We also hypothesized that Pilates exercises might have positive effects on non-gastrointestinal symptoms of fatigue, anxiety, and depression associated with IBS based on two randomized trials in patients with multiple sclerosis (MS) (12) and post-menopausal women (14) and also based on a recent meta-analysis investigating the effect of Pilates exercises on depression in women with various medical conditions (e.g., MS, schizophrenia, chronic low back pain, type 2 diabetes, and breast cancer) (15). The results of this study may aid efforts targeting lifestyle approaches to improve the symptoms of IBS.

Methods

This study follows the CONSORT 2010 Statement Guidelines for reporting randomized controlled trials (16).

Study design and settings

This is a single-centered, parallel-group, randomized, controlled study. This study recruited patients by referral and continued from July 2023 to February 2024. The Ethics Committee of the senior author's institution approved the study's protocol (NO: P.T.REC/012/004218). This study has adhered to the Helsinki Declaration's guidelines. Patients provided informed consent before the beginning of the study. The study was prospectively registered at ClinicalTrials.gov (registration No.: NCT05832801), and no changes were made to the protocol after the study commenced.

Randomization and concealed allocation

A simple randomization was employed with a 1:1 allocation ratio using a randomization table created by a computer software program. The allocation sequence was hidden using opaque, sealed envelopes with sequential numbers. The participants and the allocator were unaware of the upcoming allocation.

Implementation and blinding

The randomization sequence was generated by a researcher who was not involved in study interventions. Enrollment and assignment of subjects were performed by the physiotherapist involved in study interventions. For practical considerations, the assessor of the outcomes was not blinded to the allocation of subjects. Also, both the subjects and therapist were not blinded either to the dietary advice or the Pilates exercises due to the nature of the interventions.

Subjects

Sixty women with irritable bowel syndrome were recruited for this study by referral from a physician. Eligibility criteria were as follows: women aged 20-45 (i.e., the most common age for developing IBS) (17), a body mass index of 25-34.9 kg/m^2 (i.e., IBS is mostly prevalent in overweight and obese subjects) (5,18), a diagnosis of IBS based on Rome IV Diagnostic criteria (4,19) constipation-predominant IBS established by a physician according to the Bristol stool form scale (20), and moderate-to-severe IBS (IBS severity score>174) (21). Exclusion criteria were organic gastrointestinal disorders, thyroid dysfunctions, concurrent cardiovascular, respiratory, renal, hepatic disorders, pregnancy, hematological disease, neurological/musculoskeletal problems, psychiatric disease, fibromyalgia, and previous history of stomach or intestinal excision. Eligible subjects were randomly assigned to a study group (Pilates and Dietary Advice) (n=30) and a control group (Dietary Advice) (n=30). Both groups received dietary advice without pharmacological treatment (e.g., laxatives).

Power analysis

A priori power analysis could not be performed due to the lack of similar studies in this research area. A post-hoc power analysis was conducted based on the total IBS-SSS score data from the present study using the G* Power software program (3.1.9.4). The post hoc power analysis revealed that 60 patients achieved 99% power at alpha = 0.05.

Evaluation

History taking and clinical evaluation

Thorough medical history-taking and clinical evaluation were performed for patient selection. The demographic, anthropometric, and clinical features of eligible patients were recorded at baseline. BMI was calculated at baseline

by dividing body weight in kilograms by height in meters squared (22).

Outcome measures

The primary outcome measure was the IBS symptom severity scoring system (IBS-SSS). The secondary outcome measures were the number of complete spontaneous bowel movements, the modified fatigue impact scale (MFIS), the hospital anxiety and depression (HADS) scale, and body weight (BW).

Irritable Bowel Syndrome Severity Scoring System (IBS-SSS)

The IBS-SSS is a valid and reliable patient-based measure that evaluates the severity of IBS symptoms through five clinically significant items over ten days (21), as follows: (1) frequency and (2) severity of abdominal pain; (3) degree of abdominal distention or tightness; (4) dissatisfaction with bowel habits; and (5) affection of IBS on quality of life. A greater score denotes worse conditions. Each item is rated on a visual analog scale (VAS) from 0 to 100, resulting in an overall score ranging from 0 to 500 (21). Based on the data collected, the IBS-SSS scores are divided into three categories: mild symptoms (from 75 to 174), moderate symptoms (from 175 to 299), and severe symptoms (from 300 to 500) (21). Additionally, a 95-point reduction in total IBS-SSS scores is clinically meaningful, indicating an improvement in symptoms (23).

Frequency of complete spontaneous bowel movements

Complete spontaneous bowel movement is defined as a sense of complete evacuation without laxatives, enemas, or suppositories on the day of the bowel movement or the preceding day [24]. A participant with a weekly complete spontaneous bowel movement frequency rate of three or more and an increase of one or more from baseline is considered a responder [24]. The patients reported the number of their complete spontaneous bowel movements per week at baseline and post-intervention. The minimal clinically important change in the number of complete spontaneous bowel movements is 1.3 times per week for subjects suffering from functional constipation treated with acupuncture (25).

Modified Fatigue Impact Scale (MFIS)

This questionnaire measures the impact of exhaustion on life in subjects experiencing fatigue-like symptoms (26). The Arabic version of the questionnaire was used (27). The 21 items in the MFIS are divided into three categories: physical (nine items), cognitive (10 items), and psychosocial (two items). For all items, participants rate their agreement using a 5-point Likert scale, where 0 means “never” and 4 means “almost always.” The total score (0–84) comprises sub-scores for physical (0–36), cognitive (0–40), and psychosocial (0–8) functioning (26). A greater score is worse. The Arabic version of MFIS exhibited high reliability and concurrent validity in MS (27).

The Hospital Anxiety and Depression Scale (HADS)

To evaluate anxiety and depression, the Arabic version of the Hospital Anxiety and Depression Scale (HADS) was used (28). The Arabic version of HADS is a valid and reliable instrument (28). HADS has two subscales (i.e., anxiety and depression), and each subscale has 7 items. Each item is rated on a 4-point Likert scale (0-3) (29). Each subscale has a normal range of 0-7, a borderline range of 8-10, and a range denoting depression or anxiety of 11-21 (29). The minimal clinically important change is 1.17-2.13 for anxiety symptoms and 1.48-2.54 for depression symptoms (30).

Body weight

Body weight was measured at baseline and after 8 weeks.

Interventions

The interventions are reported following the TIDieR checklist (31).

Dietary/lifestyle advice

Dietary advice is an essential component in the management of IBS (32). Patients in both groups were instructed to follow dietary advice as per the guidelines from the National Institute for Health and Care Excellence (NICE) and the British Dietetic Association (BDA) (33-35) for 8 weeks. The dietary pieces of advice are outlined in Table 1. A physician provided dietary advice through one face-to-face interview in a private clinic at the beginning of the study, and patients were instructed to report their diet using diaries. Then, compliance with the dietary advice was assessed regularly by a physiotherapist who checked patients’ diaries via the Watts-up application. All patients adhered to the dietary advice given. No modifications were made to the dietary advice throughout the study, as patients adhered well to the intervention.

TABLE 1 - Dietary advice

Dietary advice as per the guidelines from the National Institute for Health and Care Excellence (NICE) and the British Dietetic Association (BDA) (33-35).
Integrating a healthier eating habit, having food at the same time every day with regular intervals
Never eat too little or too much.
Staying properly hydrated
Preventing processed, fatty, and spicy food
Limiting caffeine, carbonated, and alcoholic drinks
Limiting fiber intake to soluble fibers starts with a low dose and builds up gradually.
Avoiding insoluble fibers, gas-producing foods like beans, and sweeteners
Awareness of dietary intolerance
An additional advice for increasing physical activity.

Exercise Intervention

Pilates exercises may stimulate the vagal tone, thus enhancing the GIT function (13), and have a role in relieving fatigue, anxiety, and depression (12). Patients in the study group received Pilates exercises and dietary advice for eight weeks. The Pilates exercise program, modified from the protocol by Silva et al. (36), is described in Table 2. A mat and a gymnastic Swiss ball for adults were used. Initially, patients received instructions on the exercises and YouTube's educational videos, as shown in Table 2. Then, an experienced physiotherapist guided the subjects in Pilates exercises through face-to-face sessions at a physiotherapy clinic. The sessions were scheduled two times per week for a total of 16 sessions. The sessions started with a 5-minute warm-up composed of repeated sit-ups and back extensions on a Swiss ball and 1 set of 30 seconds of hamstring stretch, knee to chest, and global stretch for trunk and back. The session ended with a 5-minute cool-down period with the same activities as the warm-up. The total session lasted for 45-50 minutes. From the 1st to the 3rd week, patients performed each exercise for one set of 8 repetitions that reached 10 repetitions by the end of the 4th week. From the 5th to the 8th week, two sets of 10 repetitions were performed. At this time, exercises were personalized so that the patients who couldn't perform the two sets continually were allowed to rest for 3–5 minutes between sets. No adjustments were required, as every patient could complete the activities as directed. All patients in this group adhered to the scheduled sessions.

TABLE 2 - Pilates exercises prescription

Pilates exercises [modified from Silva et al. (36)]	Description
1-Swan	• The subject lies prone with hands resting in the direction of the shoulders, • Extends the elbows, maintaining the head in line with the spine and extending the trunk. • Returns to the starting position. • https://www.youtube.com/watch?v=hFoF-9UhJJc
2-One leg up -down	• The subject lies flat, arms outstretched alongside the body. • Elevates the leg in extension with plantar flexion. • Returns to the starting position • https://www.youtube.com/watch?v=x06RH8Dig8g
3- Leg circles	• The subject lies flat, arms outstretched alongside the body and supported on the ground. • Elevates the leg in extension with plantar flexion. • Makes circles with the leg • https://www.youtube.com/watch?v=zeUW8LFUEPo
4-Single leg stretch	• The subject lies flat and flexes the right leg as much as she can towards the chest by putting the left hand on the right knee and the right hand on the right ankle. • Extends the left leg at a 30° angle. • Alternates the leg slowly • https://www.youtube.com/watch?v=dJWsTv3IhOo
5-Saw	• The subject sits erect, with legs apart at hip width, and arms extended apart at shoulder height. • Twists the spine to the left slowly from the waist. • Moves the right arm towards the left foot and the left arm back at shoulder height. • Returns to the original position and switches sides. • https://www.youtube.com/watch?v=1XcU-WsTcaU
6-Side kicks: front and back	• The subject assumes side lying, elbow flexed, and hand resting under the head. • Keeps the upper leg aligned with the hips and slowly brings the extended leg forward. • Returns to the starting position. • https://www.youtube.com/watch?v=ZAhTD9Mitck
7-The hundred	• The subject lies flat, elbow extended, with the shoulder, hips, and knees at 90°. • Extends the knee at approximately 45°, with a slight bending of the trunk (removing the shoulder blades from the mat) and chin towards the chest. • Returns to the starting position • https://www.youtube.com/watch?v=HRsDeUrW1BA
8-Pelvic lift on the ball	• The subject lies flat, legs straight, with heels on the ball. Squeezes the glutei and lifts the hips from the mat as high as possible. • Returns to the starting position • https://www.youtube.com/watch?v=Pxu5Mp6_ezc

Pilates exercises [modified from Silva et al. (36)]	Description
9-Sit-ups with the ball.	• The subject lies flat while holding the ball over the head and legs at 45. • Brings the ball toward the legs and hold it. • Returns to the starting position • https://www.youtube.com/watch?v=L0rZ6V_SB-U
10-Stretching on the ball for muscle relaxation.	• The subject lies supine and prone on the ball in each position for 30 seconds. This was a part of cooling down. • https://www.youtube.com/watch?v=7Lp5X9aazkg

Statistical analysis

Descriptive statistics were used to describe baseline patients' characteristics (age, weight, height, BMI, IBS duration) as means ±SD. A two-way analysis of variance (ANOVA) was employed to compare outcome measures within and between groups. The chi-squared test was utilized to compare IBS-SSS severity categories between the groups after interventions. All statistical tests were set at a significance level of $p < 0.05$. Mean differences (MD) and 95% confidence interval (CI) were computed. The effect size was assessed by calculating the absolute mean changes from baseline (change scores) and the standardized mean difference between the two groups (Cohen d). Cohen d was interpreted as per Cohen [37], as <0.2= trivial effect; 0.2-0.49 = small effect; 0.5-0.79 = moderate effect; and ≥ 0.8= large effect. The change scores were calculated by subtracting the pre-intervention mean score from the post-intervention mean score. Mean percent changes were also calculated. Statistical Package for Social Sciences (SPSS) software version 24 was used for all analyses.

Results

Participants flow

Of 79 women with IBS, 15 did not fulfill the enrolment criteria, and 4 refused to participate. Sixty women with IBS were randomized to two groups: a study group (Pilates and Dietary Advice) (n = 30) and a control group (Dietary Advice) (n = 30). No losses occurred after randomization, and all participants in each group received the intended interventions and were included in the final analysis. The flow of participants can be seen in Figure 1. Eligible participants were recruited from July 2023 to February 2024.

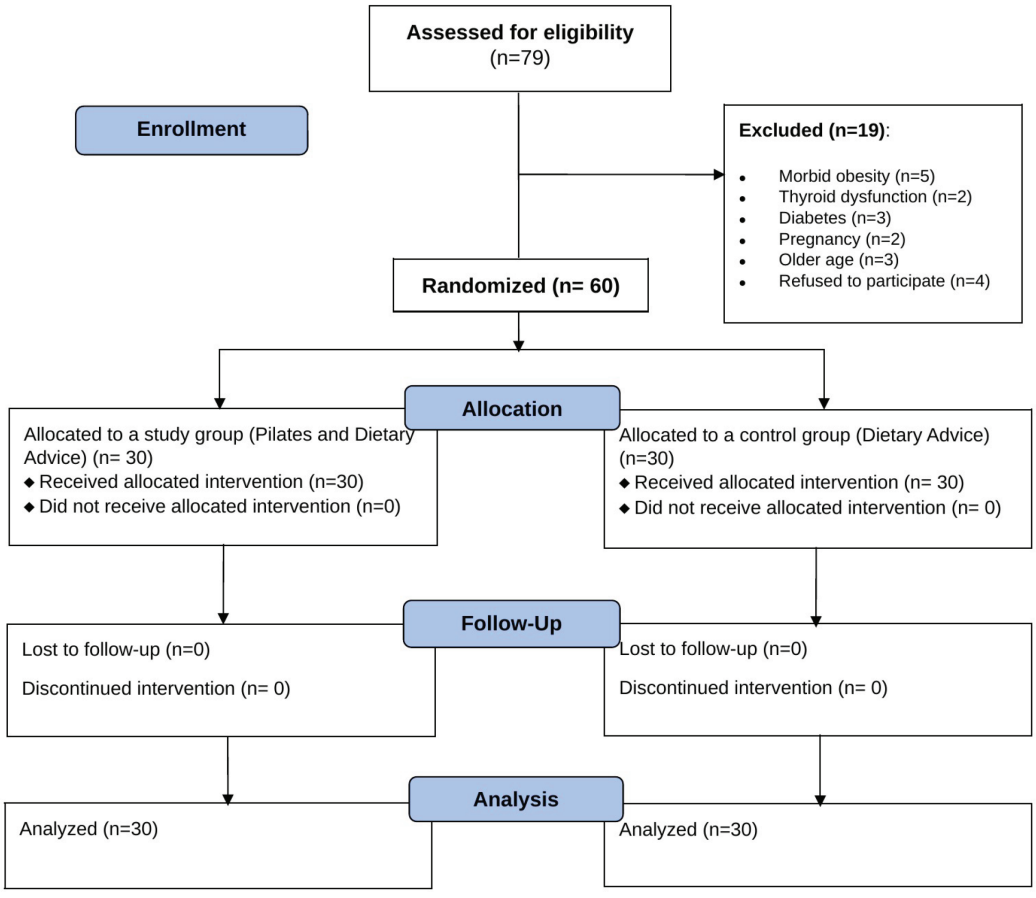


FIGURE 1 - Flowchart of the study.

Baseline data

The age and anthropometric and clinical characteristics of patients in the two groups are listed in Table 3.

Results of irritable bowel syndrome severity scoring system (IBS-SSS)

The total IBS-SSS displayed significant reductions in the study and control groups ($p < 0.001$) compared to the baseline (Table 4). However, the study group showed a more significant reduction in the total IBS-SSS score than the control group ($p < 0.001$), with a moderate effect size (Cohen d = 0.73) (Table 4). The percentage changes from baseline in total IBS-SSS were ↓66.75% versus ↓38.28% in the study and control groups, respectively (Table 4). Additionally, the severity of IBS was significantly lower in the study group than in the control group post-intervention ($p<0.001$) (i.e., 30 mild cases and 0 moderate cases in the study group versus 16 mild cases and 14 moderate cases in the control group) (Table 4).

Frequency of complete spontaneous bowel movements

The mean values of the frequency of complete spontaneous bowel movements showed a significant increase in the study and control groups ($p < 0.001$) compared to the baseline (Table 5). However, the study group presented a more significant increase in the frequency of complete spontaneous bowel movements than the control group ($p < 0.001$) with a moderate effect size (Cohen d = 0.5) (Table 5). The percentage changes in the frequency of complete spontaneous bowel movements were ↑97.17 % versus ↑67.69% in the study and control groups, respectively (Table 5).

Results of Modified Fatigue Impact Scale (MFIS)

The subscores of cognitive, physical, and psychosocial subscales of MFIS and the total score of MFIS reduced significantly in the study and control groups compared to baseline ($p < 0.001$) (Table 5). However, the study group showed significantly greater improvements in the MFIS total score ($p < 0.001$) than the control group with a moderate effect size (Cohen d = 0.74) (Table 5). The percentage changes in MFIS were ↓55.52% versus ↓28.83% in the study and control groups, respectively (Table 5).

Results of Hospital Anxiety and Depression Scale (HADS)

The mean values of anxiety and depression scores reduced significantly in the study and control groups compared to baseline ($p < 0.001$); however, the study group showed significantly more reductions in the mean scores of anxiety and depression with moderate effect sizes (d = 0.56, d = 0.64, respectively) than the control group ($p < 0.001$) (Table 5). The percent changes in anxiety were↓53.45% versus ↓27.64% in the study and control groups, respectively; and in depression, they were↓59.1% versus ↓26.33% in the study and control groups, respectively (Table 5).

Body weight (BW)

After interventions, BW displayed a significant reduction of 4.91% in the study group only ($p <0.001$), with no significant difference between the two groups ($p=0.17$) and trivial effect size (Cohen d=0.19) (Table 5).

TABLE 3 - Baseline data for both groups

Variables		Study group (Pilates and Dietary Advice) ($n_1 = 30$)	Control group (Dietary Advice) ($n_2 = 30$)
Age (Years)		29.4 ± 7.66	30.33 ± 8.63
Body weight (kg)		77.15 ± 11.16	77.63 ± 9.85
Height (cm)		160.38 ± 5.57	161.32 ± 4.58
BMI (kg/m2)		29.9 ± 3.23	29.73 ± 3.24
IBS Duration (Years)		7.38 ± 5.95	7.22 ± 5.98
Total IBS-SSS score		312.57 ± 80.1	283.7 ± 62.71
Severity	Moderate	11 (36.66%)	17(56.66%)
	Severe	19(63.33%)	13(43.33%)
Frequency of bowel movements		2.47 ± 0.94	2.37 ± 0.89
Modified fatigue impact scale		59.87 ± 9.72	58.27 ± 12.1
Anxiety		13.47 ± 3.7	14 ± 3.3
Depression		10.83 ± 2.59	11.13 ± 3.63

Data are expressed as Means ± SD and frequencies (percent distribution). BMI: Body Mass Index; IBS: Irritable Bowel Syndrome; IBS-SSS: Irritable Bowel Syndrome Severity Scoring System

TABLE 4 - Results of IBS-SSS within and between groups

IBS-SSS			Study group (Pilates and Dietary Advice) (n_1 = 30)	Control group (Dietary Advice) (n_2 = 30)	MD	95% CI	Effect size Cohen d	p-value
Pain severity	Pre		59.63 ± 19.75	59 ± 16.01	0.63	(−8.66,9.93)	–	0.89
	Post		21.47±10.37	39.37±16.07	−17.9	(−24.91, −10.89)	0.67	<0.001**
	p-value		<0.001*	<0.001*				
	Change score		−38.16	−19.79				
	% Mean change		↓64%	↓33.27%				
Pain duration	Pre		66.33 ± 24.7	52.33±24.31	14	(1.33,26.67)	–	0.03**
	Post		15 ± 6.3	28.33±19.67	−13.3	(−20.99, −5.68)	0.46	0.001**
	p-value		<0.001*	<0.001*				
	Change score		−51.33	−24				
	% Mean change		↓77.39%	↓45.86%				
Abdominal distention	Pre		61.7 ± 24.03	42.23 ± 23.23	19.47	(7.25,31.68)	–	0.002**
	Post		18.87 ± 15.49	30.67 ± 21.54	−11.8	(−21.49, −2.1)	0.32	0.018**
	p-value		<0.001*	<0.001*				
	Change score		−42.83	-11.56				
	% Mean change		↓69.42%	↓27.37%				
Defecation satisfaction	Pre		64.6 ± 25.91	62.9 ± 22.09	1.7	(−10.74,14.14)	–	0.79
	Post		24.53±15.34	36.1±16.32	−11.5	(−19.75, −3.38)	0.36	0.006**
	p-value		<0.001*	<0.001*				
	Change score		−40.07	−26.8				
	% Mean change		↓62.03%	↓42.61%				
General QOL	Pre		71.2 ± 18.51	67.23 ± 19.99	3.97	(−5.99,13.92)	–	0.43
	Post		24.4 ± 9.07	40.63 ±17.62	−16.2	(−23.53, -8.94)	0.58	<0.001**
	p-value		<0.001*	<0.001*				
	Change score		−46.8	−26.6				
	% mean change		↓65.73%	↓39.57%				
Total score	Pre		312.57 ± 80.1	283.7 ± 62.71	28.87	(−8.31,66.04)	–	0.13
	Post		103.93 ± 27.95	175.10 ± 62.80	−71.1	(−96.53, −45.8)	0.73	<0.001**
	p-value		<0.001*	<0.001*				
	Change score		−208.64	−108.6				
	% mean change		↓66.75%	↓38.28%				
Severity (based on IBS-SSS total score)	Post	Mild	30 (100%)	16(53.33%)				<0.001¶
		Moderate	0 (0%)	14(46.66%)				

Data are expressed as Means ± SD, change scores and percent mean changes from baseline, and frequencies and percent distributions. IBS-SSS: Irritable Bowel Syndrome Severity Scoring System; QOL: Quality of life; MD: mean difference; CI: confidence interval. *Significant p-value ($p < 0.05$) based on the two-way ANOVA within-group comparison; **significant p-value ($p < 0.05$) based on the two-way ANOVA between-groups comparison; ¶ significant p-value ($p < 0.05$) based on Chi-square test.

TABLE 5 - Results of the frequency of bowel movements, MFIS, HADS, and body weight within and between groups

Outcomes			Study group (Pilates and Dietary Advice) (n_1 = 30)	Control group (Dietary Advice) (n_2 = 30)	MD	95% CI	Effect size Cohen d	p-value
Frequency of bowel movements		Pre	2.47 ± 0.94	2.37 ± 0.89	0.1	(−0.37,0.57)	–	0.67
		Post	4.87 ± 1.04	3.77 ± 1.44	1.1	(0.54,1.66)	0.5	<0.001**
		p-value	<0.001*	<0.001*				
		Change score	2.4	1.4				
		% Mean change	↑97.17%	↑67.69%				
Modified fatigue impact scale (MFIS)	Cognitive	Pre	26.83 ± 5.87	26.37 ± 5.37	0.47	(−2.6,3.53)	–	0.76
		Post	12.83 ± 5.02	18.13 ± 6.12	−5.3	(−8.19, −2.41)	0.47	0.001**
		p-value	<0.001*	<0.001*				
		Change score	−14	−8.24				
		% Mean change	↓52.18%	↓31.25%				
	Physical	Pre	26.77 ± 5.19	26.53 ± 6	0.23	(−2.67,3.13)	–	0.87
		Post	11.83 ± 2.1	18.9 ± 4.44	−7.0	(−9.54, −4.6)	0.74	0.001**
		p-value	<0.001*	<0.001*				
		Change score	−14.94	−7.63				
		% Mean change	↓55.81%	↓28.76%				
	Psychosocial	Pre	5.83 ± 1.34	5.37 ± 1.54	0.47	(−0.28,1.2)	–	0.22
		Post	1.97 ± 1.38	4.3 ± 1.66	−2.3	(−3.12, −1.54)	0.78	<0.001**
		p-value	<0.001*	<0.001*				
		Change score	−3.86	−1.07				
		% Mean change	↓69.3%	↓19.93%				
	Total	Pre	59.87 ± 9.72	58.27 ± 12.1	1.6	(−4.07,7.27)	–	0.57
		Post	26.63 ± 9.48	41.47 ± 10.43	−14.8	(−19.98, −9.68	0.74	<0.001**
		p-value	<0.001*	<0.001*				
		Change score	−33.23	−16.8				
		% Mean change	↓55.52%	↓28.83%				
Hospital Anxiety and Depression Scale (HADS)	Anxiety	Pre	13.47 ± 3.7	14 ± 3.3	−0.5	(−2.35,1.28)	–	0.56
		Post	6.27 ± 2.7	10.13 ± 3.94	−3.8	(−5.66, −2.08)	0.56	<0.001**
		p-value	<0.001*	<0.001*				
		Change score	−7.2	−3.87				
		% Mean change	↓53.45%	↓27.64%				
	Depression	Pre	10.83 ± 2.59	11.13 ± 3.63	−0.3	(−1.93,1.33)	–	0.71
		Post	4.43 ± 2.6	8.2 ± 3.27	−3.7	(−5.29, −2.24)	0.64	<0.001**
		p-value	<0.001*	<0.001*				
		Change score	−6.4	−2.93				
		% Mean change	↓59.1%	↓26.33%				
Body weight (kg)		Pre	77.15 ± 11.16	77.63 ± 9.85	−0.48	(−5.92, 4.96)	–	0.86
		Post	73.36 ± 10.27	77.08 ± 10.24	−3.72	(−9.02, 1.57	0.19	0.17
		p-value	<0.001*	0.15				
		Change score	−3.79	−0.55				
		% Mean change	↓4.91%	↓0.71%				

Data are expressed as Mean ± SD and percent mean changes from baseline. MD: mean difference, CI: confidence interval; *Significant p-value ($p < 0.05$) based on the two-way ANOVA within-group comparison; **significant p-value ($p < 0.05$) based on the two-way ANOVA between-groups comparison

Discussion

The purpose of the study was to investigate the effect of Pilates exercises on IBS severity scoring system assessed by IBS-SSS, the number of complete spontaneous bowel movements, fatigue assessed by MFIS, anxiety, and depression assessed by HADS, and body weight (BW) in women with constipation-predominant IBS. To our knowledge, this study is the first randomized control trial that investigates the effectiveness of Pilates exercises in IBS. The main findings of this study are: (i) women with constipation-predominant IBS who received Pilates exercises in addition to dietary advice for IBS showed significantly greater improvements in IBS symptoms and severity as assessed by the IBS-SSS, with a moderate effect size in its total score, than their counterparts who received dietary advice only; (ii) Pilates exercises in combination with dietary advice led to a significantly higher frequency of complete spontaneous bowel movements, less fatigue as assessed by MFIS, and lower anxiety and depression levels as assessed by HADS with moderate effect sizes than dietary advice alone in women with constipation-predominant IBS, (iii) body weight was reduced significantly only following Pilates exercises and dietary advice compared to the baseline value.

This study showed that pilates exercises, in addition to dietary advice, showed more significant improvement in the IBS symptoms/severity assessed by IBS-SSS (i.e., pain duration, pain severity, abdominal distension, defecation satisfaction, and general QoL) than dietary advice alone. In a similar context, Fani et al. (38) revealed that six weeks of aerobic exercises significantly improved the severity of IBS symptoms. Interestingly, the change scores in total IBS-SSS scores in the study and control groups were 208.64 and 108.6 points, respectively, representing clinically important changes, as a 95-point reduction in total IBS-SSS scores is clinically meaningful (23). However, there was a difference in clinical significance in the improvements of IBS-SSS between the two groups in favor of the study group. This is because all patients in the study group had mild symptoms after the interventions, in contrast to 16 patients with mild symptoms and 14 patients with moderate symptoms in the control group post-intervention. Moreover, there was a significant difference with a moderate effect size in the improvement of IBS-SSS total score between the groups in favor of the study group (i.e., Cohen d = 0.73).

Another finding in this study was that the self-reported frequency of complete spontaneous bowel movements per week increased by a significantly greater degree in patients who received Pilates exercises and dietary advice than in patients who received dietary advice alone. This finding complements the previous findings in this study (i.e., the improvement in IBS symptoms and severity). Also, this study reveals the effectiveness of Pilates exercises in treating symptoms of constipation in our patients who had constipation-predominant IBS. In a similar context, Daley et al. (39) showed that exercise can significantly improve constipation symptoms compared to usual care in patients with IBS. Additionally, Gao et al. (40), in their systematic review, concluded that exercise can have a major role in alleviating constipation symptoms. The mechanism underlying the constipation-relieving effect of Pilates exercises is that Pilates has the advantage of combining deep breathing with body movements, which can activate the vagal tone (13). Within this, Liu et al. (41) reported that 6 weeks of slow deep breathing exercises improved the number of bowel movements compared to the control group, which can be attributed to improvement in the parasympathetic activity in patients with constipation-predominant IBS. Worth noting is that Ai et al. (25) reported that a mean increase of ≥1.3 times/week in complete spontaneous bowel movement in patients with severe functional constipation treated with acupuncture suggests clinical significance. Given that the mean increases (change scores) in the frequency of complete spontaneous bowel movements were 2.4 and 1.4 times per week, the changes in this outcome may be of clinical importance in both groups. However, there was a significant difference with a moderate effect size in the improved frequency of complete spontaneous bowel movements between the groups in favor of the study group (Cohen d = 0.50).

The current study also showed that Pilates exercises and dietary advice improved overall fatigue symptoms (i.e., cognitive, physical, and psychosocial) compared to the baseline and dietary advice alone. In the present study, the mean changes (i.e., change scores) in total MFIS scores were 33.23 and 16.8 points in the study and control group, respectively. Alawami and Abdulla (42) reported that a mean change of 14.68 or more points in total MFIS score may indicate a minimal detectable change (MDC) in fatigue of clinical importance in patients with MS. Although this MDC in fatigue, assessed by MFIS, was investigated in MS, we think this threshold value may help interpret the changes in fatigue assessed by MFIS in response to our interventions. Nevertheless, there was a significant difference with a moderate effect size in the improvement of MFIS total score between groups in favor of the study group (Cohen d = 0.74).

Pilates exercises played a role in the enhancement of fatigue symptoms in patients with MS (12,43,44), healthy young female participants (45,46), post-menopausal women (14,47), and women with breast cancer (48). A recent meta-analysis reported that moderate aerobic exercises or combination training approaches for 2-10 weeks positively affected fatigue in subjects suffering from chronic conditions (49). In contrast, Johannesson et al. (50) demonstrated that moderate physical activity did not enhance fatigue symptoms in patients with IBS.

This study also showed that more significant improvements in anxiety and depression were found in the study group than in the control group. This finding can be supported by a recent systematic review by Ju et al. (15), which showed that Pilates exercises can be considered an additional treatment method for alleviating depression and anxiety symptoms in female patients. In addition, Pilates was effective in improving anxiety and depression symptoms in women with type 2 diabetes (51), middle-aged women with obesity (52), and post-menopausal women (53). On the other hand, some studies reported that Pilates exercises did not lead to any significant improvement in anxiety or depression in female patients with fibromyalgia (54) or dysmenorrhea (55). It is worth noting that changes in HADS anxiety symptoms of 1.17-2.13 points and in HADS depression symptoms

of 1.48-2.54 points were considered clinically important in patients with chronic pain (30). The change scores in HADS anxiety and depression symptoms in both groups in the present study were greater than those cut-off values, suggesting clinically meaningful improvements. However, there were significant differences with moderate effect sizes in the improvement of HADS anxiety symptoms (Cohen d = 0.56) and HADS depression symptoms (Cohen d = 0.64) between groups in favor of the study group.

The positive effects of Pilates on IBS symptoms, fatigue, anxiety, and depression may be explained based on the fact that Pilates can activate the parasympathetic nervous system and optimize the function of the hypothalamic-pituitary-adrenal axis, which regulates several body processes, including digestion, mood, and emotions, energy storage, and production (52,56,57). It should also be noted that the control group, which received dietary advice, only showed significant improvement in IBS symptoms, frequency of complete spontaneous bowel movements, anxiety, and depression compared to baseline. This may be because the diet they received could have modulated gut flora composition and function, thus optimizing the gut-brain axis pathway and reducing the brain's perception of abdominal pain, discomfort, and anxiety [58]. As such, it may be unsurprising that the patients who received both Pilates exercises and dietary advice experienced greater improvements in their IBS symptoms than those who received dietary advice alone, owing to the combined effects of both interventions.

The last finding in this study was that the study group showed a significant reduction in body weight compared to the baseline value. Pilates effectively reduced body weight in subjects with overweight or obesity (59,60), MS (44), young women (46), women with type 2 diabetes (51), and postmenopausal women (53). On the contrary, a study by Park et al. (52) reported that Pilates exercises did not cause any change in body weight in female participants with obesity. Also, the meta-analysis by Cavina et al. (61) reported that mat Pilates had no advantage over the control condition for reducing body weight in the general population. Interestingly, the mean percent change in body weight from baseline in the study group was approximately 4.91%, and this reduction could be of clinical relevance. It was shown that a weight reduction of 5% from baseline is "clinically meaningful" (62) and is associated with a decline of 13% in intra-hepatic triglycerides (63). Elevated triglyceride levels were found to exaggerate IBS symptoms (64).

Finally, this study has limitations, and its findings should be interpreted in that context. There was a lack of control over potential hormonal factors (e.g., effects of menstrual cycles or use of contraceptive pills). Also, the current study only included women with overweight and grade I obesity suffering from constipation-predominant IBS, which may limit the generalization of the results. Moreover, this study did not include abdominal circumference or other metrics of obesity. Nevertheless, the current study has several strengths. This study is the first to investigate the effectiveness of Pilates exercises in women with IBS. Additionally, a variety of patient-reported outcomes were used in this study to produce an extensive assessment of the patients at baseline and in response to the interventions. Moreover, the results of this study may have practical applications for physiotherapists and healthcare providers interested in the complementary treatment of IBS.

Conclusion

When used in tandem with dietary advice, Pilates exercises may be an effective therapeutic intervention that could significantly reduce IBS symptoms and severity, increase the frequency of complete spontaneous bowel movements, reduce fatigue, and relieve anxiety and depression to a more moderate extent than dietary advice alone in women with IBS. However, dietary advice alone may also improve these outcomes in these patients.

Disclosures

Financial support: No funding was received for this research.

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

References

1. Sperber AD, Bangdiwala SI, Drossman DA, et al. Worldwide prevalence and burden of functional gastrointestinal disorders, results of Rome Foundation Global Study. Gastroenterology. 2021;160(1):99-114.e3. CrossRef PubMed
2. Kim YS, Kim N. Sex-gender differences in irritable bowel syndrome. J Neurogastroenterol Motil. 2018;24(4):544-558. CrossRef PubMed
3. Galica AN, Galica R, Dumitrascu DL. Epidemiology of irritable bowel syndrome in Albania. J Gastrointestin Liver Dis. 2021; 30(3):334-338. CrossRef PubMed
4. Mearin F, Lacy BE, Chang L, et al. Bowel disorders. Gastroenterology. 2016;150(6):1393-1407. CrossRef PubMed
5. Frändemark Å, Jakobsson Ung E, Törnblom H, et al. Fatigue: a distressing symptom for patients with irritable bowel syndrome. Neurogastroenterol Motil. 2017;29(1):e12898. CrossRef PubMed
6. Tang HY, Jiang AJ, Wang XY, et al. Uncovering the pathophysiology of irritable bowel syndrome by exploring the gut-brain axis: a narrative review. Ann Transl Med. 2021;9(14):1187. CrossRef PubMed
7. Mayer EA, Ryu HJ, Bhatt RR. The neurobiology of irritable bowel syndrome. Mol Psychiatry. 2023;28(4):1451-1465. CrossRef PubMed
8. Ford AC, Lacy BE, Talley NJ. Irritable bowel syndrome. N Engl J Med. 2017;376(26):2566-2578. CrossRef PubMed
9. Ford AC, Sperber AD, Corsetti M, et al. Irritable bowel syndrome. Lancet. 2020;396(10263):1675-1688. CrossRef PubMed
10. D'Silva A, MacQueen G, Nasser Y, et al. Yoga as a therapy for irritable bowel syndrome. Dig Dis Sci. 2020;65(9):2503-2514. CrossRef PubMed
11. Portincasa P, Bonfrate L, de Bari O, et al. Irritable bowel syndrome and diet. Gastroenterol Rep (Oxf). 2017;5(1):11-19. CrossRef PubMed
12. Fleming KM, Coote SB, Herring MP. Home-based Pilates for symptoms of anxiety, depression and fatigue among persons with multiple sclerosis: an 8-week randomized controlled trial. Mult Scler. 2021;27(14):2267-2279. CrossRef PubMed
13. Yoo KT. Effects of mat Pilates on the autonomic nervous system in the elderly women. J Korean Soc Phys Med. 2022;17(4): 27-35. CrossRef

14. Aibar-Almazán A, Hita-Contreras F, Cruz-Díaz D, et al. Effects of Pilates training on sleep quality, anxiety, depression and fatigue in post-menopausal women: a randomized controlled trial. Maturitas. 2019;124:62-67. CrossRef PubMed
15. Ju M, Zhang Z, Tao X, et al. The impact of Pilates exercise for depression symptoms in female patients: a systematic review and meta-analysis. Medicine (United States). 2023;102(41): E35419. CrossRef
16. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. J Pharmacol Pharmacother. 2010;1(2):100-107. CrossRef PubMed
17. Hadjivasilis A, Tsioutis C, Michalinos A, et al. New insights into irritable bowel syndrome: from pathophysiology to treatment. Ann Gastroenterol. 2019;32(6):554-564. CrossRef PubMed
18. Pickett-Blakely O. Obesity and irritable bowel syndrome: a comprehensive review. Gastroenterol Hepatol (N Y). 2014;10(7): 411-416. PubMed
19. Drossman DA. Functional gastrointestinal disorders: history, pathophysiology, clinical features, and Rome IV. Gastroenterology. 2016;150(6):1262-1279. CrossRef
20. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. Scand J Gastroenterol. 1997;32(9):920-924. CrossRef PubMed
21. Francis CY, Morris J, Whorwell PJ. The irritable bowel severity scoring system: a simple method of monitoring irritable bowel syndrome and its progress. Aliment Pharmacol Ther. 1997;11(2):395-402. CrossRef PubMed
22. Nuttall FQ. Body mass index: obesity, BMI, and health: a critical review. Nutr Today. 2015;50(3):117-128. CrossRef PubMed
23. Spiegel B, Bolus R, Harris LA, et al. Measuring irritable bowel syndrome patient-reported outcomes with an abdominal pain numeric rating scale. Aliment Pharmacol Ther. 2009;30(11-12):1159-1170. CrossRef PubMed
24. Cao H, Wang Y, Su J, Zhao H, Han M, Liu JP. Acupuncture for chronic constipation. Cochrane Database Syst Rev. 2023; 2023(4):CD004117. PubMed
25. Ai YK, Liu ZS, Liu BY, et al. Minimal clinically important difference of the frequency of bowel movement for patients with chronic severe functional constipation treated with acupuncture. Zhongguo Zhenjiu. 2022;42(11):1247-1250. Chinese. PubMed
26. Fatigue guidelines development panel of the Multiple Sclerosis Council for clinical practice guidelines. Fatigue and multiple sclerosis. Evidence-based management strategies for fatigue in multiple sclerosis. Paralyzed Veterans of America; 1998. Online
27. Khalil H, Al-Shorman A, Alghwiri AA, et al. Cross cultural adaptation and psychometric evaluation of an Arabic version of the modified fatigue impact scale in people with multiple sclerosis. Mult Scler Relat Disord. 2020;39:101878. CrossRef PubMed
28. Terkawi AS, Tsang S, AlKahtani GJ, et al. Development and validation of Arabic version of the Hospital Anxiety and Depression Scale. Saudi J Anaesth. 2017;11(5)(suppl 1):S11-S18. CrossRef PubMed
29. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand. 1983Jun;67(6):361-70. CrossRef PubMed
30. Grönkvist R, Vixner L, Äng B, et al. Measurement error, minimal detectable change, and minimal clinically important difference of the Short Form-36 Health Survey, Hospital Anxiety and Depression Scale, and Pain Numeric Rating Scale in patients with chronic pain. J Pain. 2024;25(9):104559. CrossRef PubMed
31. 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
32. Jayasinghe M, Karunanayake V, Mohtashim A, et al. The role of diet in the management of irritable bowel syndrome: a comprehensive review. Cureus. 2024;16(2):e54244. CrossRef PubMed
33. National Institute for Health and Clinical Excellence (NICE). Irritable bowel syndrome in adults. Diagnosis and management of irritable bowel syndrome in primary care. Clinical Guideline 61 Update 2017. Online (Accessed July 2024)
34. McKenzie YA, Bowyer RK, Leach H, et al. IBS Dietetic Guideline Review Group on behalf of Gastroenterology Specialist Group of the British Dietetic Association. British Dietetic Association systematic review and evidence-based practice guidelines for the dietary management of irritable bowel syndrome in adults (2016 update). J Hum Nutr Diet. 2016;29(5):549-575. CrossRef PubMed
35. Vasant DH, Paine PA, Black CJ, et al. British Society of Gastroenterology guidelines on the management of irritable bowel syndrome. Gut. 2021;70(7):1214-1240. CrossRef PubMed
36. Silva HJA, Lins CAA, Nobre TTX, et al. Mat Pilates and aquatic aerobic exercises for women with fibromyalgia: a protocol for a randomized controlled blind study. BMJ Open. 2019;9(2):e022306. CrossRef PubMed
37. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Routledge; 1988. CrossRef
38. Fani M, Mostamand J, Fani M, et al. The effect of aerobic exercises among women with mild and moderate irritable bowel syndrome: a pilot study. J Bodyw Mov Ther. 2019;23(1):161-165. CrossRef PubMed
39. Daley A, Grimmett C, Roberts L, et al. The effects of exercise upon symptoms and quality of life in patients diagnosed with irritable bowel syndrome: a randomized controlled trial. J Sci Med Sport. 2010;12(2):E49. CrossRef
40. Gao R, Tao Y, Zhou C, et al. Exercise therapy in patients with constipation: a systematic review and meta-analysis of randomized controlled trials. Scand J Gastroenterol. 2019;54(2):169-177. CrossRef PubMed
41. Liu J, Lv C, Wang W, et al. Slow, deep breathing intervention improved symptoms and altered rectal sensitivity in patients with constipation-predominant irritable bowel syndrome. Front Neurosci. 2022;16:1034547. CrossRef PubMed
42. Alawami AS, Abdulla FA. Psychometric properties of an Arabic translation of the modified fatigue impact scale in patients with multiple sclerosis. Disabil Rehabil. 2021;43(22):3251-3259. CrossRef PubMed
43. Miri H, Mehrabian H, Peyvandi MG, et al. Effect of pilates training on balance, muscular endurance, fatigue, and quality of life among women with multiple sclerosis. Health, Sport. Rehabilitation (Stuttg). 2023;9(4):73-84. Online
44. Eftekhari E, Etemadifar M. Impact of clinical mat Pilates on body composition and functional indices in female patients with multiple sclerosis. View project. Crescent J Med Biol Sci. 2018;5(4):297-305. Online
45. Amzajerdi A, Keshavarz M, Ezati M, et al. The effect of Pilates exercises on sleep quality and fatigue among female students dormitory residents. BMC Sports Sci Med Rehabil. 2023;15(1):67. CrossRef PubMed
46. Akbaş E, Ünver B. A six-week pilates exercise protocol for improving physical and mental health-related parameters. Malaysian J Movement. Healthc Exec. 2018;7(2):65-79.
47. Mohamed DMA, Hussien AMN, Kamel HEDH, et al. Effect of Pilates exercises on fatigue in post-menopausal women. Bull Fac Phys Ther. 2023;28(1):4-9. CrossRef

48. Eyigor S, Karapolat H, Yesil H, et al. Effects of pilates exercises on functional capacity, flexibility, fatigue, depression and quality of life in female breast cancer patients: a randomized controlled study. Eur J Phys Rehabil Med. 2010;46(4):481-487. PubMed
49. Barakou I, Sakalidis KE, Abonie US, et al. Effectiveness of physical activity interventions on reducing perceived fatigue among adults with chronic conditions: a systematic review and meta-analysis of randomized controlled trials. Sci Rep. 2023;13(1):14582. CrossRef PubMed
50. Johannesson E, Simrén M, Strid H, et al. Physical activity improves symptoms in irritable bowel syndrome: a randomized controlled trial. Am J Gastroenterol. 2011;106(5):915-922. CrossRef PubMed
51. Alizadeh R, Zolfaghari Z, Nosratzehi S, et al. Effect of Pilates on anxiety, depression, and anthropometric indices of women with type 2 diabetes. Medical-Surgical Nursing Journal. 2023;12(2):e143695. CrossRef
52. Park HY, Jung K, Jung WS, et al. Effects of online Pilates and face-to-face Pilates intervention on body composition, muscle mechanical properties, cardiometabolic parameters, mental health, and physical fitness in middle-aged women with obesity. Healthcare (Basel). 2023;11(20):2768. CrossRef PubMed
53. Farzaneh A, Koushkie Jahromi M, Salesi M, Rostami R. The ameliorating impact of pilates training on anxiety and body mass index in menopausal symptoms. Womens Health Bull. 2020;7(4):46-53. Online
54. Komatsu M, Avila MA, Colombo MM, et al. Pilates training improves pain and quality of life of women with fibromyalgia syndrome. Rev Dor. 2016;17(4):274-278. CrossRef
55. Song BH, Kim J. Effects of Pilates on pain, physical function, sleep quality, and psychological factors in young women with dysmenorrhea: a preliminary randomized controlled study. Healthcare (Basel). 2023;11(14):2076. CrossRef PubMed
56. Najafi P, Hadizadeh M, Cheong JPG, et al. Effects of Tele-Pilates and Tele-Yoga on biochemicals, physical, and psychological parameters of females with multiple sclerosis. J Clin Med. 2023;12(4):1585. CrossRef PubMed
57. Robayo A. Hypotalamus-Pituitary-Adrenal (HPA) Axes and Their Relationship with Stress, Mood, Personality, and Neurocognitive Functioning. In: Lopez-Soto T, Garcia-Lopez A, Salguero-Lamillar FJ, eds. The Theory of Mind Under Scrutiny: Psychopathology, Neuroscience, Philosophy of Mind and Artificial Intelligence. Springer Nature Switzerland; 2024:341-365.
58. Ribeiro G, Ferri A, Clarke G, et al. Diet and the microbiota - gut - brain-axis: a primer for clinical nutrition. Curr Opin Clin Nutr Metab Care. 2022;25(6):443-450. CrossRef PubMed
59. Hushmandi K, Jamali J, Saghari S, et al. The effect of eight weeks of pilates exercises on anthropometric indices and subjective well-being in obese middle-aged women. Open Public Health J. 2023;16(1):1-8. CrossRef
60. Wang Y, Chen Z, Wu Z, et al. Pilates for overweight or obesity: a meta-analysis. Front Physiol. 2021;12:643455. CrossRef PubMed
61. Cavina APS, Pizzo Junior E, Machado AF, et al. Effects of the mat Pilates method on body composition: systematic review with meta-analysis. J Phys Act Health. 2020;17(6):673-681. CrossRef PubMed
62. Williamson DA, Bray GA, Ryan DH. Is 5% weight loss a satisfactory criterion to define clinically significant weight loss? Obesity (Silver Spring). 2015;23(12):2319-2320. CrossRef PubMed
63. Magkos F, Fraterrigo G, Yoshino J, et al. Effects of moderate and subsequent progressive weight loss on metabolic function and adipose tissue biology in humans with obesity. Cell Metab. 2016 Apr 12;23(4):591-601. CrossRef PubMed
64. Xu M, Liu D, Tan Y, et al. A Mendelian randomization study on the effects of plasma lipids on irritable bowel syndrome and functional dyspepsia. Sci Rep. 2024;14(1):78. CrossRef PubMed

Arch Physioter 2024; 14: 182-188
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3155
ORIGINAL RESEARCH ARTICLE

Evolving trends of systematic reviews on virtual reality for stroke rehabilitation

Silvia Bargeri[1], Mariachiara Baggio[2], Stefania Guida[1], Greta Castellini[1], Silvia Gianola[1]

[1]Unità di Epidemiologia Clinica - IRCCS Istituto Ortopedico Galeazzi, Milan - Italy
[2]University of Verona, Verona - Italy

ABSTRACT

Objective: Virtual reality (VR) therapies have increasingly been adopted across medical fields, including neurorehabilitation for stroke recovery. Evidence from several systematic reviews (SRs) was explored, covering different aspects. We aim to explore overlaps, gaps, and trends of SRs focusing on VR stroke rehabilitation providing a foundation for improving the field and addressing current limitations.

Materials and methods: We moved from a recent overview of reviews, searching multiple databases for all published SRs and the international database of prospectively registered SRs (PROSPERO) for ongoing SRs. Data extraction of study characteristics and methodological quality of SRs using AMSTAR 2 were obtained from a recent overview of reviews. Two independent reviewers conducted data analysis and visualization by the trend over time of published SRs with their included primary studies and ongoing SRs, methodological quality and other SR characteristics.

Results: The data set consisted of 58 SRs, including 345 primary studies and 45 ongoing SRs, published between 2007 and 2022. The number of published and ongoing SRs significantly increased over time (R^2 = 0.8654; R^2 = 0.747, respectively). In the last three years, Asia accounts for the majority of publications (31%). Overall, the main outcome assessed over time was upper extremity function and activity in 67.2% of SRs. Most of the published SRs were judged "critically low" (77.6%). The number of included studies increased over time reaching a median of 17 studies with a median of 493 participants.

Conclusions: In stroke rehabilitation, the published and ongoing SRs on VR have risen over time in terms of the number of publications, with some concerns about methodological quality and representation of countries around the world.

Keywords: Randomized controlled trials as topic, Rehabilitation, Stroke, Systematic review, Virtual reality

What is already known about this topic:

- *In recent years, VR technologies rapidly spread across medical specialties, including neurorehabilitation. Recent research trends highlight various VR therapies, but systematic reviews (SR) on VR for stroke rehabilitation, crucial for clinicians and policymakers, remain unexplored.*

What does the study add:

- *The number of SRs on VR has increased, including those ongoing, with larger sample sizes and diverse outcomes. However, concerns about methodological quality and global representation exist. Authors should check protocol registries and plan innovative synthesis methods.*

INTRODUCTION

Stroke is one of the major causes of disability and death worldwide, with the highest incidence in the elderly population (1). In 2019, ischemic heart disease and stroke were the top-ranked causes of Disability-Adjusted Life-Years (DALYs) both in the 50–74-year-old group and in the 75-year-old and elder group (2,3), impacting motor functions, activities of daily living (ADL), social participation and quality of life (4).

The use of technology in rehabilitation after stroke has been proposed worldwide in the past three decades, with an increasing interest in virtual reality (VR) (5). VR technology has the advantages of creating more realistic environments to imitate the real world, providing repetitive training for specific tasks, increasing the sense of participation, and stimulating near-life experiences that patients cannot otherwise achieve (6). In fact, VR rehabilitation utilizes virtual environments and objects to deliver visual and auditory feedback to the user. This feedback can be experienced through various platforms, such as head-mounted displays, projection systems, or flat screens, with equipment ranging from basic tools

Received: June 4, 2024
Accepted: December 5, 2024
Published online: December 31, 2024

This article includes supplementary materials

Corresponding author:
Stefania Guida
email: stefania.guida@grupposandonato.it

like a joystick to technologies such as sensors or cameras. The implementation of VR in rehabilitation shows helpful results in motor function recovery, especially in the upper extremities and lower extremities for balance, gait, and posture. The way that VR can be used in multiple different conditions suggests the efficacy and versatility of the application (6).

In recent years, especially during the pandemic, there has been a rapid spread of VR technologies (7,8) across all medical specialties (9), including the neurorehabilitation field (10). The most recent research trends cover more defined types of VR therapy, embracing different study designs offering information regarding the current hotspots in the field (10). However, to our knowledge, the characteristics, and extent of the highest study design for synthesizing the evidence (i.e., SRs) and informing clinicians, patients, and policymakers focused on stroke VR rehabilitation was not investigated yet. Therefore, we aimed to evaluate overlaps, gaps, and trends of published and ongoing SRs on VR for stroke rehabilitation over the years, along with their general characteristics and methodological quality.

METHODS

Study design

We conducted a cross-sectional study based on an overview of reviews (11). We adapted items from the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) (12), assuming as units of analysis the included SRs (see **Supplementary 1**).

Search strategy and data collection

We moved from a recent overview of reviews (11) investigating VR on stroke (CRD42022329263), which included 58 SRs and 345 unique primary studies.

The search was launched in multiple databases (the Cochrane Database of Systematic Reviews, EMBASE, MEDLINE, Scopus, ISI Web of Science, CINAHL, PsycINFO, PEDro, Otseeker, Healthevidence.org, Epistemonikos), including PROSPERO for ongoing SRs from inception up to January 17, 2023. We selected SRs published in English including adults with any diagnosis of stroke. The treatment investigated was any kind of immersive, semi-immersive, or non-immersive VR intervention, either with or without conventional therapy (e.g., usual care, exercises). Details of eligibility criteria are reported in our previous publication (11).

Data collection

We obtain the dataset of the related overview (11) to collect information about the general characteristics and methodological quality of SRs. Particularly, we used the following general characteristics: years of publication, countries of the corresponding author, description of outcomes, references and year of included primary studies, number of included primary studies, sample size, journal of publication, and journal impact factor (JIF), methodological quality appraised by A Measurement Tool to Assess SR (AMSTAR) 2 and categorized into critically low, low, moderate, and high methodological quality (13).

Data synthesis

We used descriptive statistics for general characteristics and methodological quality expressed as median and interquartile range (IQR) or absolute value and frequency.

The data chart and Pearson's correlation coefficient were used to assume linearity assumption. Data were transformed into a logarithmic scale. Prediction of the percentage of published and ongoing SR data based on year was computed through linear regression models plotted with confidence and prediction intervals at 95%. All analyses were performed using the R Core Team (2023) (14).

We then visually described the trend of the following variables: median JIF, number of included primary studies and participants, outcomes, and methodological quality using Microsoft Excel 2019 and RAWGraphs 2.0 (15).

RESULTS

Publication trend

Overall, the data set consisted of 58 SRs published between 2007 and 2022, including 345 primary studies published between 1999 and 2021, and 45 ongoing SRs from 2014 to 2022. We found a significant increase in the number of published and ongoing SRs over the years (R^2 = 0.8654; R^2 = 0.747) respectively) (**Figure 1**). In **Supplementary 2**, we reported the trend of SRs and primary studies publication for each year (**Figure S1-S2**, respectively).

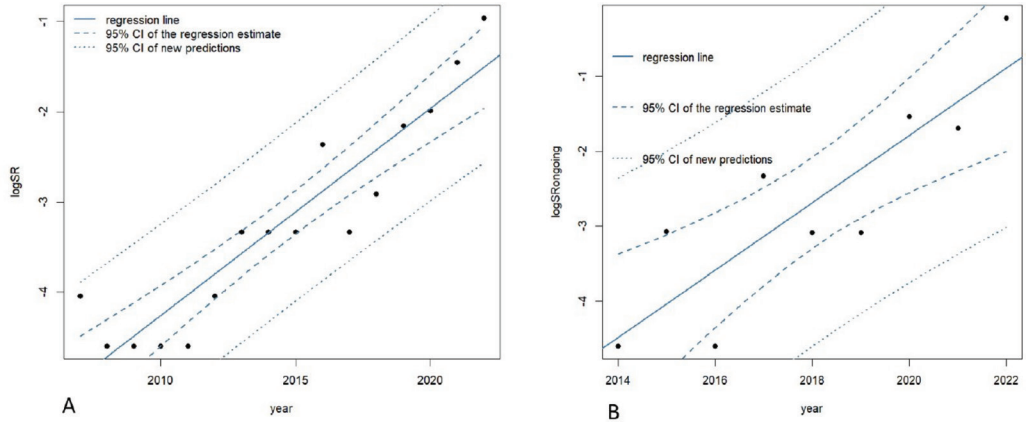


FIGURE 1 - A) Trend of published SR and B) Trend of ongoing SR over the years. Legend: SRs, systematic reviews.

General characteristics of published systematic reviews

In **Table 1** we showed overall general characteristics of published SRs. Most SRs include RCTs only (81%). Overall, 42 SRs (72.4%) conducted meta-analyses. Many SRs (69%) included mixed onset of stroke. The most common journal of publication was the Journal of Stroke and Cerebrovascular Diseases (8.6%) (**Supplementary 2, Table S1**). The median JIF was 3.25 (2.18-4.61). The distribution of the median JIF over the years is reported in **Supplementary 2, Figure S3.**

TABLE 1 - General characteristics of published SRs

General Characteristics	Overall (N = 58)
Population, Median (IQR)	492.5 (224.8-1082.8)
N. of included primary studies, median (IQR)	17 (8-30.3)
N of outcomes, median (IQR)	3 (2-4)
Presence of conflicts of interest, N (%)	6 (10.3)
Non-industry funding, **N** (%)	28 (48.3)
Presence of meta-analyses, **N** (%)	42 (72.4)
SR including RCTs only, **N** (%)	47 (81.0)
Country, n (%)	
Africa	1 (1.7)
America	13 (22.4)
Asia	24 (41.4)
Europe	16 (27.6)
Oceania	4 (6.9)
JIF, Median (IQR)	3.25 (2.18-4.61)
Outcomes, **N** (%)	Total (n = 58)
Upper limb function and activity	39 (67.2)
Gait and balance	36 (62.1)
ADL	37 (63.8)
Participation	28 (48.3)
Cognitive and mental function	20 (34.5)
Adverse events	18 (31.0)

Primary studies and participants of included SRs

The number of primary studies included in an SR ranged from a minimum of three to a maximum of 87, with a median of 17 per SR. More than 85% of primary studies were RCTs. The number of participants included in the SRs ranged from a minimum of 60 to a maximum of 3540, with a median of 493. **Supplementary 2, Figure S4-S5** shows the distribution of primary studies and participants over the years.

Countries

As regards the Asian states, China is the first country for the number of reviews, with 15 publications (25.9% of all SRs). In Europe, 16 SRs were published in eight different countries; Spain holds 5 publications (8.6% of all SRs), whereas Italy and Belgium hold 3 reviews each one (5.2%). America is the third continent for publications of SRs, with Brazil involved in 6 studies (10.3%), Canada in 4 (6.9%), and the USA in 3 (5.2%). In Oceania, 4 SRs were published. Only one review was conducted in Africa. The distribution of the continents where SRs have been conducted both generally and over the years is illustrated in **Figure 2A-B.** In the last three years we found an absolute increase in publications in Asia (n = 18, 31% of the overall sample).

Outcomes

Overall, the main outcome assessed was Upper Extremity Function and Activity in 67.2% of SRs (n = 39), followed by Activities of Daily Living (n = 37) and Gait and Balance (n = 36). Less investigated outcomes were Participation (n = 28), Cognitive and Mental function (n = 20), and Adverse Events (n = 18) (Table 1). **Figure 3** shows the outcome distribution over the years. In the last three years, 28 SRs (48.3% of the overall sample) assessed Upper Limb Function and Activity.

Methodological quality

According to the AMSTAR 2 tool, 45 SRs (77.6%) were judged critically low, 12 low (20.7%), and one high (1.7%). Overall, 69% of SRs have no recorded protocol, and 89.7% do not describe motivations for the excluded primary studies. The 96.6% of SRs accomplished exhaustive bibliographic research. As regards the quality of the primary studies included in the 58 SRs, all authors used an adequate tool to measure the risk of bias (100%), and 74.1% included the assessment of primary studies in their SRs' results (**Supplementary File 2, Figure S6**). **Figure 4** shows the methodological quality distribution over the years.

DISCUSSION

Main findings

We analyzed the frequency and characteristics of 58 published SRs and 45 ongoing SRs covering the scientific diffusion of VR technologies for stroke from 2007 to 2022 and from 2014 to 2022, respectively, with an increasing trend over the years.

VR research is becoming more influential around the world, with over one-fifth of all countries involved in the scientific progress in this field (16). Asia represents the most influential country regarding VR rehabilitation in stroke adult people, with a particular increase in SR's publications in the last three years (31%). In particular, China has been the leader in SR publications since it started in 2007 to cover not only the VR scientific field but also all medical fields, as it is in the top publishing countries, as reported by recent publications (17,18). With its aging population, China faces increasing challenges for stroke care and prevention, with a prevalence of stroke survivors of 58.1 million, four times higher than other countries (19). It has been argued that this rapid increase in the number of SRs could be due to multiple reasons, including the easiest widely accessible tools for doing SRs and meta-analyses (20), the pressures of academia (21), or industry of contracting companies "operating in the domain of evidence synthesis" to produce these publications,

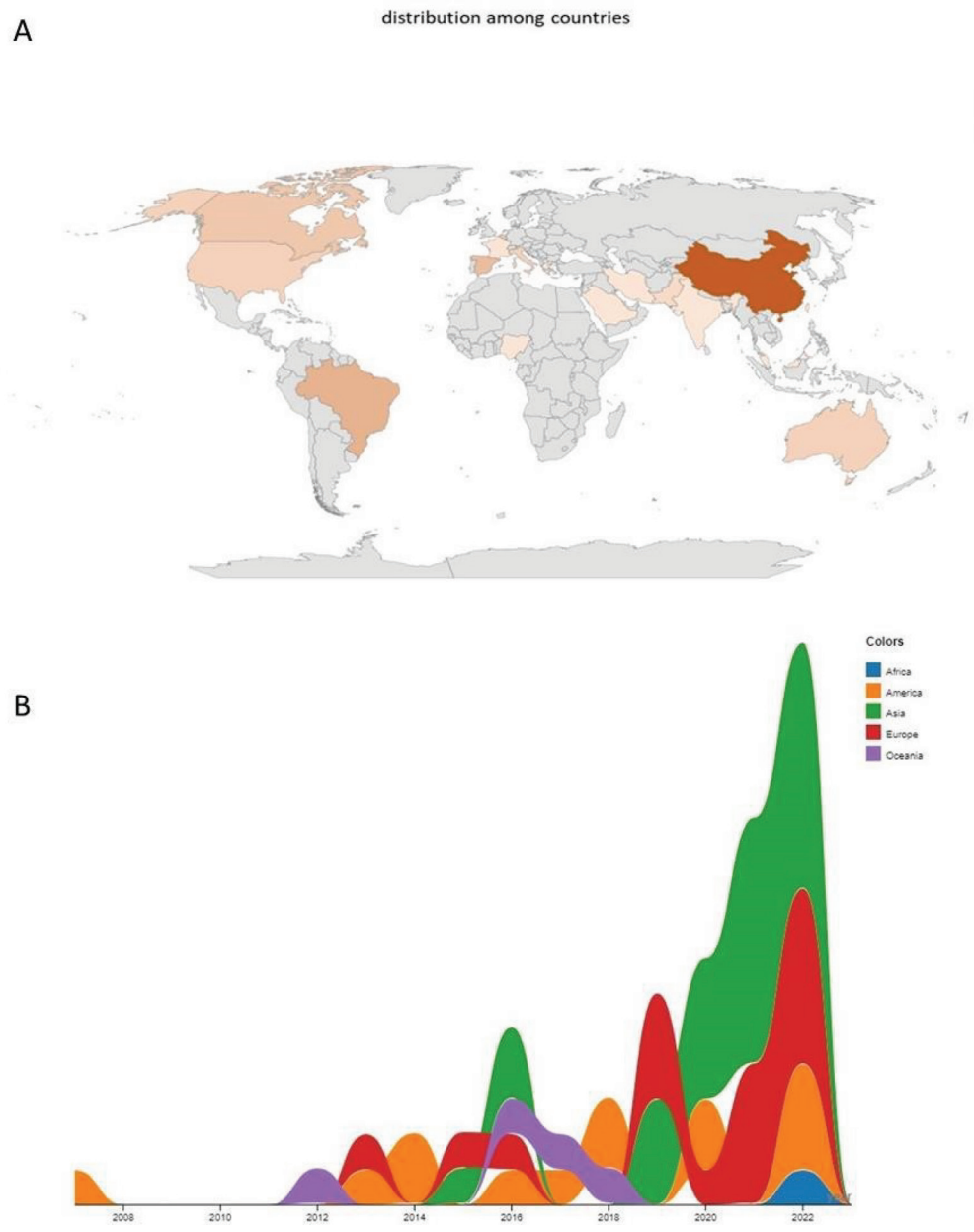


FIGURE 2 - Countries of corresponding authors of SRs A) overall and B) over the years. Legend: the size is proportional to the number of SRs for each year by country

many of which probably remain unpublished (22). However, we cannot exclude that this increase is due to the growth of collaborative research with China and other countries, as already reported (23)

As the number of SRs has increased over time, the number of participants included in SRs and the number of included primary studies increased. This phenomenon might be conditioned by the fact that the latest published SRs can also include primary studies as well as participants of the oldest SRs. However, the number of primary studies retrieved from SRs may be underestimated considering that a time span exists from running the search strategy and publication of SRs (24); therefore, we cannot exclude that some published primary studies were not included in the SRs as well as new primary studies might be actually still ongoing. The year 2019-2020 was represented by a global pandemic emergency with limited possibilities to undergo primary studies. This might explain the substantial growth of secondary literature studies rather than primary studies in the last recent years.

Looking at the health outcomes, upper limb function and activity, participation, and cognitive and mental function are becoming more assessed in SRs as a sign of implementation of a core outcome set that can optimize the quality of post-stroke rehabilitation (24).

We also found that the methodological quality of SRs remained critically low over time, even if SRs were published in journals with a median JIF of 3.25. It has been found that JIF may have little to no association with study results or methodological quality (25). Low methodological quality can be mainly due to some critical flaws in the protocol and in

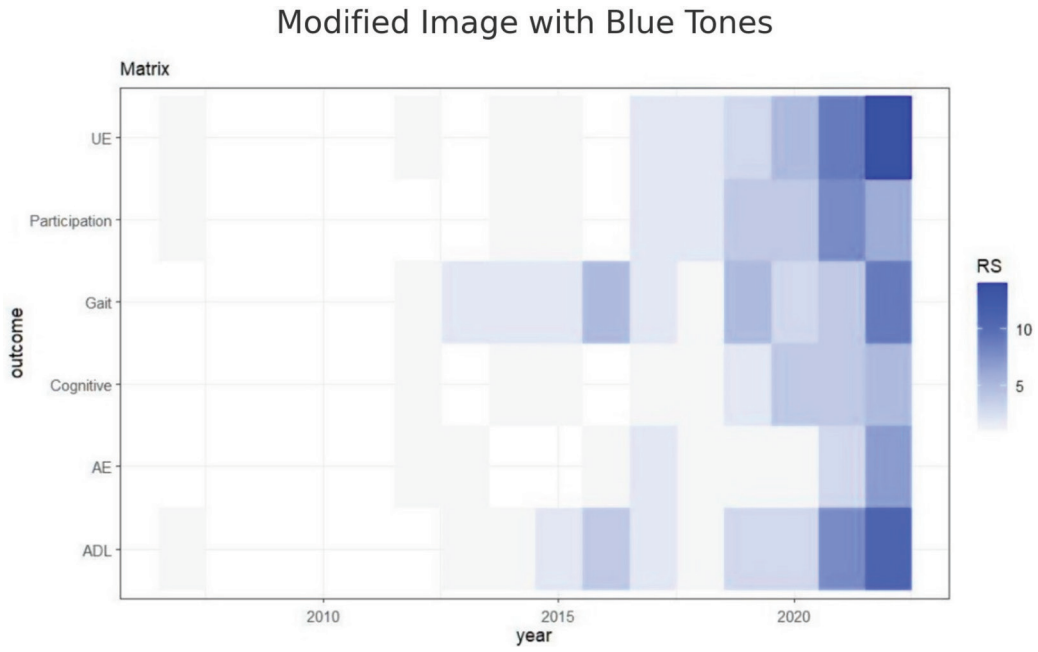


FIGURE 3 - Outcome assessed in SRs over years. Legend: AE, adverse events, UE, upper limb

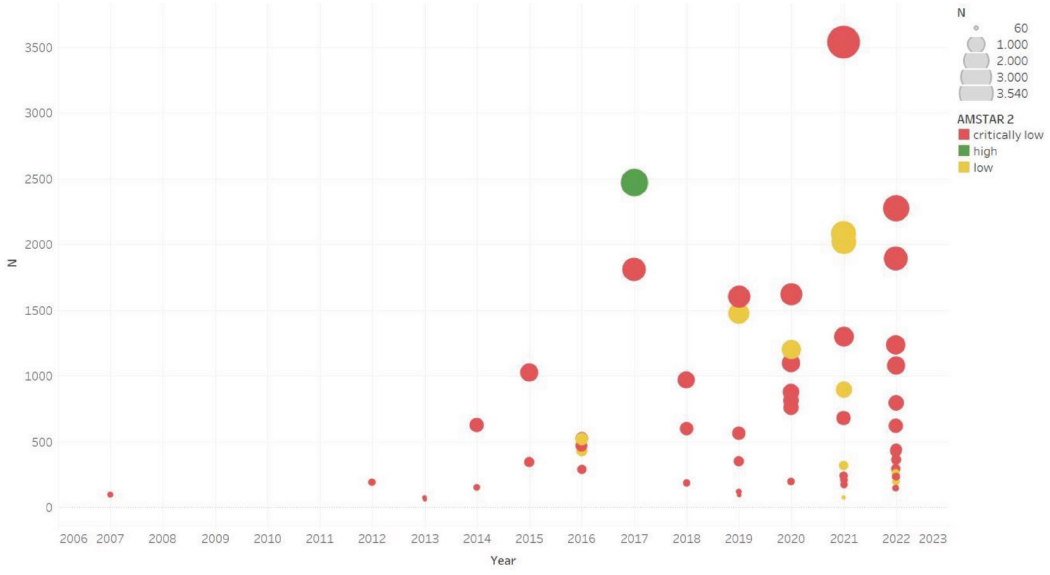


FIGURE 4 - Methodological quality of SRs over the years. Legend: the bubble size is proportional to the number of included participants in each SR

the excluded studied justification items. In fact, according to AMSTAR 2, 69% of our included SRs have no explicit statement that the study methods were established prior to the conduct of the review. On one hand, these findings are discouraging, as more than half of the authors (55.8%) register their protocol prior to publishing their systematic review/meta-analysis (26). Pre-registration does not guarantee that the protocol is complete but allows readers to be aware of methods for conducting and reporting in a transparent way (22). On the other hand, in the last years, we collected many ongoing SRs pre-registered in PROSPERO as a starting point of improvement for this flaw. However, authors should use registers such as PROSPERO to check overlapping questions covered by already existing SRs to avoid redundant meta-analyses with inconsistency and discordant findings. As well, journal editors should keep in mind that much has already been published. Multiple overlapping SRs can facilitate the origin of disputes. This is well known also in other fields such as the case of thrombolytic therapy for pulmonary embolism (27). As well, the large increase in the number of published and ongoing SRs over the years raises concerns about research waste. It should be ethical and reasonable to systematically review what is already known before deciding to perform any new study (28).

The other important flaw in SRs is related to the non-reporting of studies excluded from SRs. We found the 89.7%

don't provide a list of the excluded studies and the motivation. This can lead to selective inclusion for outcome non-reporting bias impacting meta-analytic effects (29-31).

Recently, a meta-epidemiological study found that 58,8% of the included SRs (n = 131) excluded studies due to "no relevant outcome data" (32) despite it not being recommended by the scientific community (33) since this may be a consequence of selective outcome reporting and therefore compromise the systematic review reliability.

Strength and limitation

This review aimed to summarize all the publications and trends about the application of VR in the neurorehabilitation field in adult people with stroke. To analyze this specific sample, we included a total of 103 papers (58 published and 45 ongoing SRs) without limits in a publication year. An exhaustive research was conducted in many different databases. This trend study is linked to an extensive overview of reviews in accordance with the Cochrane Guidelines; moreover, the review protocol was registered in the PROSPERO database.

Limitations of this work need to be acknowledged. The research was conducted for only English-language publications, and the children population was excluded from the sample. We did not extract data from primary studies focused on VR rehabilitation, but we reported the trend of publications of those included in the SRs. Ongoing SRs were searched on PROSPERO and not on other registries. Nevertheless, PROSPERO seems to be the most common database used for protocol registration [71.3%, n = 270 (26)]. Thus, we cannot be sure that all SRs included the whole body of evidence.

Conclusion

The highest synthesis of evidence as published SRs, including those ongoing on VR, has risen over time in terms of the number of publications, sample size, and eligible outcomes, with some concerns about methodological quality and representation of countries around the world. To avoid waste of research, authors should check protocol registries before embarking on a new systematic review. As well they should consider planning innovative research methods for synthetizing the amount of literature available.

Disclosures

Conflict of interest: The authors declare no conflict of interest

Financial support: This study was supported and funded by the Italian Ministry of Health. The APC was funded by the Italian Ministry of Health – Ricerca Corrente. The funder had no role in study design, data collection and analysis, decision to publish, or manuscript preparation.

Author's contributor role: SB: conceptualization, data curation, data analysis, writing the original draft, reviewing and editing; MB: data curation, data analysis, writing the original draft, reviewing and editing; SG: data analysis, writing the original draft, reviewing & editing; GC: methodology, writing the original draft, reviewing & editing; SG: conceptualization, methodology, reviewing and editing;

Data Availability Statement: Data publicly available: The data presented in this study are openly available in the OSF online repository at https://osf.io/v6k75/

References

1. Sacco RL, Kasner SE, Broderick JP, et al. American Heart Association Stroke Council, Council on Cardiovascular Surgery and Anesthesia; Council on Cardiovascular Radiology and Intervention; Council on Cardiovascular and Stroke Nursing; Council on Epidemiology and Prevention; Council on Peripheral Vascular Disease; Council on Nutrition, Physical Activity and Metabolism. An updated definition of stroke for the 21st century: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*. 2013;44(7):2064-2089. CrossRef PubMed
2. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020;396(10258):1204-1222. CrossRef PubMed.
3. Katan M, Luft A. Global Burden of Stroke. *Semin Neurol*. 2018;38(2):208-211. CrossRef PubMed
4. Hatem SM, Saussez G, Della Faille M, et al. Rehabilitation of motor function after stroke: A multiple systematic review focused on techniques to stimulate upper extremity recovery. Front Hum Neurosci. 2016;10:442. CrossRef PubMed
5. Laffont I, Van Dokkum L, Froger J, et al. Stroke patients: emerging rehabilitation techniques. Ann Phys Rehabil Med. 2012;55: e145-e146. CrossRef
6. Amirthalingam J, Paidi G, Alshowaikh K, et al. Virtual reality intervention to help improve motor function in patients undergoing rehabilitation for cerebral palsy, Parkinson's disease, or stroke: A systematic review of randomized controlled trials. Cureus. 2021;13(7):e16763. CrossRef PubMed
7. Ahmadi Marzaleh M, Peyravi M, Azhdari N, et al. Virtual reality applications for rehabilitation of COVID-19 patients: A systematic review. Health Sci Rep. 2022;5(6):e853. CrossRef PubMed
8. Pallavicini F, Pepe A, Clerici M, et al. Virtual reality applications in medicine during the COVID-19 pandemic: Systematic review. JMIR Serious Games. 2022;10(4):e35000. CrossRef PubMed
9. Kacmaz KS, Kaçmaz C. Bibliometric analysis of research in pediatrics related to virtual and augmented reality: A systematic review. Curr Pediatr Rev. 2024;20(2):178-187. CrossRef PubMed
10. Guo QF, He L, Su W, et al. Virtual reality for neurorehabilitation: A bibliometric analysis of knowledge structure and theme trends. Front Public Health. 2022;10:1042618. CrossRef. PubMed
11. Bargeri S, Scalea S, Agosta F, et al. Effectiveness and safety of virtual reality rehabilitation after stroke: an overview of systematic reviews. EClinicalMedicine. 2023;64:102220. CrossRef PubMed.
12. 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. Ann Intern Med. 2007;147(8):573-577. CrossRef PubMed
13. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j4008. CrossRef PubMed
14. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Online (Accessed June 2024)
15. Mauri M, Elli T, Caviglia G, et al. RAWGraphs: A visualisation platform to create open outputs. In Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter, 2017; 28: 1-5. CrossRef
16. Pawassar CM, Tiberius V. Virtual reality in health care: Bibliometric analysis. JMIR Serious Games. 2021;9(4):e32721. CrossRef PubMed

17. Fontelo P, Liu F. A review of recent publication trends from top publishing countries. Syst Rev. 2018;7(1):147. CrossRef PubMed
18. Hoffmann F, Allers K, Rombey T, et al. Nearly 80 systematic reviews were published each day: Observational study on trends in epidemiology and reporting over the years 2000-2019. J Clin Epidemiol. 2021;138:1-11. CrossRef PubMed
19. World stroke organization (WSO). Global Stroke Fact Sheet 2022. Extracted from Global Burden of Disease Stroke Statistics Worldwide for the year 2019. Online (Accessed June 2024)
20. Ye XF, Yu DH, He J. The rise in meta-analyses from China. Epidemiology. 2013;24(2):335-336. CrossRef PubMed
21. Yang ZP, Ye XF, Fan DM. Meta-analysis is victim to Chinese academic and educational systems. J Formos Med Assoc. 2013 May;112(5):235-6. CrossRef PubMed
22. Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q. 2016;94(3):485-514. CrossRef PubMed
23. Wagner CS, Park HW, Leydesdorff L. The continuing growth of global cooperation networks in research: A conundrum for national governments. PLoS One. 2015;10(7):e0131816. CrossRef PubMed
24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-174. CrossRef PubMed
25. Saginur M, Fergusson D, Zhang T, et al. Journal impact factor, trial effect size, and methodological quality appear scantly related: a systematic review and meta-analysis. Syst Rev. 2020;9(1):53. CrossRef PubMed
26. Tawfik GM, Giang HTN, Ghozy S, et al. Protocol registration issues of systematic review and meta-analysis studies: a survey of global researchers. BMC Med Res Methodol. 2020;20(1):213. CrossRef PubMed
27. Riva N, Puljak L, Moja L, et al. Multiple overlapping systematic reviews facilitate the origin of disputes: the case of thrombolytic therapy for pulmonary embolism. J Clin Epidemiol. 2018;97:1-13. CrossRef PubMed
28. Clarke M, Alderson P, Chalmers I. Discussion sections in reports of controlled trials published in general medical journals. JAMA. 2002;287(21):2799-2801. CrossRef PubMed
29. Page MJ, Forbes A, Chau M, et al. Investigation of bias in meta-analyses due to selective inclusion of trial effect estimates: empirical study. BMJ Open. 2016;6(4):e011863. CrossRef PubMed
30. Kirkham JJ, Altman DG, Williamson PR. Bias due to changes in specified outcomes during the systematic review process. PLoS One. 2010;5(3):e9810. CrossRef PubMed
31. Page MJ, McKenzie JE, Kirkham J, et al. Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions. Cochrane Database Syst Rev. 2014;2014(10):MR000035. CrossRef PubMed
32. Lazzarini SG, Stella Yousif M, Bargeri S, et al. Reasons for missing evidence in rehabilitation meta-analyses: a cross-sectional meta-research study. BMC Med Res Methodol. 2023;23(1):245. CrossRef PubMed
33. Higgins JPT, Lasserson T, Chandler J, et al. Methodological Expectations of Cochrane Intervention Reviews. Cochrane: London, Version February 2022. Online (Accessed June 2024)

Arch Physioter 2024; 14: 189-195
ISSN 2057-0082 | DOI: 10.33393/aop.2024.3370
MASTERCLASS

Integrating spirituality into physical therapy: exploring its emerging role as a recognized determinant of health

Alessandra N Garcia[1], TR Goins[1], Bridget Eubanks[1], Valerie E Joseph[2], Adam English[3]

[1]Department of Doctor of Physical Therapy, Campbell University, Lillington - USA
[2]Department of Psychology, Campbell University, Lillington - USA
[3]Department of Christian Studies, Campbell University, Lillington - USA

ABSTRACT

This masterclass explores the increasing recognition of spirituality as a vital aspect of patient care, alongside other Social Determinants of Health (SDH) such as economic stability and education. The distinction between spirituality and religion is clarified, with spirituality described as a broader, more personal experience that can exist both within and outside of religious contexts. Research demonstrates that spirituality influences health in mostly positive ways, particularly in areas like mental health, resilience, and coping, making it a critical component of holistic, patient-centered care. In physical therapy, incorporating a patient's spirituality into their plan of care can enhance cultural competence and foster a more holistic care approach. However, many Physical Therapists (PTs) express uncertainty in addressing spiritual concerns, often due to limited training or unclear role expectations. The authors suggest that integrating tools like the Inclusive Spiritual Connection Scale (ISCS), Functional Assessment of Chronic Illness Therapy-Spiritual Well-Being Scale (FACIT-Sp), Spiritual Well-Being Questionnaire (SWBQ), the Spiritual Health and Life-Orientation Measure (SHALOM), or the Spiritual Transcendence Scale (STS), along with enhanced education, could help therapists incorporate spirituality into practice more seamlessly. Integration of spirituality enables PTs to deliver more complete, personalized care that addresses the whole person. Ultimately, the authors advocate for recognizing spirituality as a key determinant of health and an important component of healthcare to ensure more inclusive treatment.

Keywords: Cultural competency, Physical therapy, Spirituality

Background

"A closed mouth can't get fed… They have to know a little bit about your circumstances to be able to help you", was a response from a participant in a qualitative study when asked about their perception of how social factors may be relevant to their healthcare (1). Their comment highlights the importance of recognizing and addressing individual social factors in patient care, reflecting the growing emphasis in healthcare on integrating SDH into clinical practice. In recent years, there has been growing recognition of the critical role that SDH, such as education, economic stability, and neighborhood environment, play in shaping well-being, health outcomes, and health-related behaviors (2).

Similarly, spirituality, like social and economic factors, is being recognized as vital in shaping patient well-being (3,4). Although not yet regarded on the same level, there is growing recognition of the effort to fully integrate spirituality into healthcare and consider it an important SDH (5,6). As this awareness grows, healthcare providers are more frequently encouraged to consider spirituality as part of comprehensive care, especially in disciplines like physical therapy, where patient-centered approaches are crucial (5,7). In this masterclass, we explore the concept of spirituality and its domains, as well as spirituality as a potential determinant of health. We also review evidence of spirituality's impact on health outcomes, explore its implications for physical therapy practice and education, and present practical examples of incorporating spirituality into clinical practice with an emphasis on assessment and measurement tools.

Spirituality

Although often used interchangeably, spirituality and religion are multidimensional concepts that represent distinct ideas. The definition of spirituality varies across academic disciplines, and the dimensions assessed in different studies are often inconsistent (6,8). Spirituality can be defined as *"a dynamic and intrinsic aspect of humanity through which persons seek ultimate meaning, purpose, and transcendence and experience relationship to self, family, others, community,*

Received: November 5, 2024
Accepted: December 30, 2024
Published online: December 31, 2024

Corresponding author:
Alessandra N. Garcia
email: alessandra.garcia.pt@gmail.com

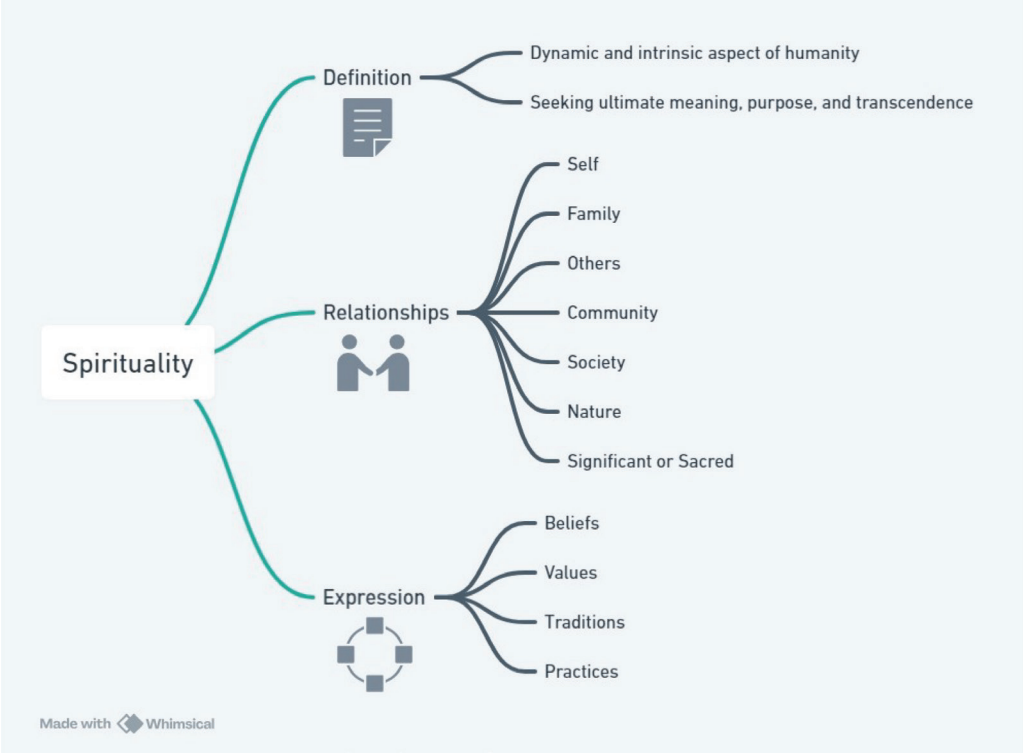


FIGURE 1 - Concept of spirituality

society, nature, and the significant or sacred. Spirituality is expressed through beliefs, values, traditions, and practices." (5,9) (Fig. 1) Religion, by contrast, names *"the search for significance that occurs within the context of established institutions that are designed to facilitate spirituality."* (5,6,9,10) In this sense, spirituality can be viewed as a broader concept than religion, encompassing an individual's search for and connection with what they perceive as transcendent or sacred (11–14). Religion, on the other hand, serves as one potential pathway for this search and connection. While spirituality often describes personal experiences and beliefs within a religious framework, it can also extend beyond specific traditions or faith communities, as reflected in the popular expression "spiritual but not religious." (15)

Both religion and spirituality encompass significant social dimensions (16). Religion often fosters community through shared practices, traditions, and institutions, providing a sense of belonging and collective identity. Spirituality, while more individualized, can also create social bonds by encouraging connections with others through shared values, mutual support, and collective experiences of meaning and purpose. These social aspects may contribute to the overall well-being of individuals by offering support networks and a sense of community (17). The past fifty years of research have tended to emphasize four domains of spirituality: personal, communal, environmental, and transcendental (18). Examples of practices that nurture personal domains, exemplify community and the search for spiritual relatedness, access the environmental domain, and manifest the transcendental domain are outlined in Table 1.

TABLE 1 - Examples of practices that nurture the four domains of spirituality

Personal Domain	Communal	Environmental	Transcendental
Contemplation, prayer, meditation	Church or community gathering attendance	Spending time in nature (walking, hiking, camping, gardening, boating)	Breathing exercises and meditation
Scripture/spiritual reading	Shared meals	Volunteering for roadside clean-up	Prayer and worship of the Divine
Chanting, recitation	Communal singing, rituals & sacraments	Contributing to recycling efforts	The study of holy texts
Activities that affirm one's sense of identity and help bring about self-awareness (e.g., sound/ energy healing)	Acts of community service	Tai chi, qigong	Similar practices evoke a sense of peace and oneness with God and the Divine

Spirituality: A Domain of Health and Wellness and a Potential Determinant of Health

Spirituality is a vital domain of health and wellness for person-centered care (7,19,20). Person-centered care, as opposed to patient-centered care, takes a holistic approach

to healing body, mind, and spirit that supports PTs in developing more individualized treatment plans to enhance overall health and well-being (3,4). In this context, 'spiritual health' is defined as *"a state of being in which an individual effectively manages life's challenges, leading to the realization of one's full potential, meaning, and purpose, and fulfillment from within."* (4) Additionally, 'spiritual wellness' may be understood as *"a positive sense of meaning and purpose in life"* (3) or "*the development of an appreciation for the depth and expanse of life and natural forces that exist in the universe.*" (21)

In other words, spiritual wellness can be considered an integral component of holistic health, with spirituality serving as a primary driver. Health comprises physical, psychological, emotional, social, and spiritual dimensions, all of which interact and influence one another (4). Despite the importance of each element, spirituality remains underrepresented in research and is less extensively documented (6,9). At the very least, evidence suggests spirituality can have a mediating role, linking other health determinants to outcomes. But it is also becoming clear that spirituality has a broader influence on well-being (6,9). More recently, Delphi panels, consisting of clinicians, public health experts, researchers, health system leaders, and medical ethicists, have recommended recognizing spirituality as a 'determinant of health', alongside other social factors, due to its demonstrated influence on health outcomes, which will be discussed further (9).

Impact of Spirituality on Health Outcomes

The distinction between religion and spirituality intersects with health outcomes in complex ways, and existing research provides far more insight into the relationship between religion and health than spirituality and health. This discrepancy might stem from the fact that while there are standardized methods for measuring religious beliefs and practices, fewer such means exist for assessing spirituality, in part because spirituality is so often highly individualized. While numerous operational definitions of spirituality exist across disciplines, studies that isolate spirituality as a distinct construct, separate from religion or psychological well-being, are comparatively limited.

A systematic review based on high-quality evidence and expert analysis (9) identified various secular and religious dimensions of spirituality, such as community involvement and prayer, that may be linked to health outcomes. For example, the frequency of attendance at religious services was associated with a reduced mortality risk. Additionally, those who attended services more frequently exhibited lower rates of smoking, alcohol use, marijuana use, and illicit drug use compared to individuals with less frequent or no attendance. The frequency of attendance at religious services was also connected to better quality of life, including higher life satisfaction, improved mental health, fewer depressive symptoms, and reduced suicidal behaviors. Among adolescents, frequent attendance was associated with lower levels of unsafe sexual behavior, smoking, and substance use, including alcohol, marijuana, and illicit drugs. While most existing literature on spirituality and health situates spirituality within a religious framework, (22) we continue to see similar patterns of connection between secular spiritual expressions and positive health outcomes. For example, spiritual well-being and secular reverence have been associated with reduced levels of cardiovascular risk markers and shorter hospital stays after open-heart surgery, respectively (23,24).

Spiritual practices and beliefs can serve as coping mechanisms for stress and anxiety, but they can also help reorient people's perspectives and help them develop attitudes of resiliency and positivity. Gathering together a broad survey of studies, Mueller, Plevak, and Rummans found that spiritual practice and spiritual well-being are associated with more positive outlooks in persons with "cancer, HIV disease, heart disease, limb amputation, and spinal cord injury." (25) Religious practice serves as a helpful coping mechanism for people with asthma, anxiety, and feelings of stress. Spirituality does not, of course, guarantee better health outcomes. In a review of 3,300 empirical studies on religiosity and spirituality, approximately 12% of the studies reported negative associations between spirituality/religiosity (e.g., spiritual struggle or distress) and various health outcomes (e.g., general well-being, depression, anxiety, cancer) (26).

It is important to note that the benefits attributed to the religious dimensions of spirituality, as previously discussed, may also stem from the inherently social nature of religious institutions (e.g., churches, mosques, synagogues, and temples). These institutions often provide social support, opportunities for interaction, and community activities (27). Nearly nine out of ten U.S. adults (89%) believe religious institutions foster community connection and unity (27). Similarly, 87% recognize their significant role in assisting the poor and vulnerable, while three-quarters credit them with upholding and promoting societal morality (27). 90% or more of Christians, along with 88% of Muslims, Jews, and Hindus, regard religious institutions as unifying forces in society (27). Non-religious individuals also recognize this role, including 85% of agnostics, 81% of those without a specific religious identity, and 75% of atheists (27). Similarly, most Christians (90%), adherents of non-Christian faiths (82%), and religiously unaffiliated individuals (78%) agree that religious institutions play a crucial role in supporting the poor and needy (27).

Spiritual and religious values can, at times, amplify fear, refusal of treatment, and distrust of medical institutions and practitioners. For example, religious affiliation may be linked to vaccine hesitancy rates in certain religious communities (28,29) and negative mental health effects because of shame and guilt associated with uncured illnesses (30). In addition, while many Americans acknowledge the positive societal contributions of religious institutions, roughly half also voice concerns about their behavior. These concerns include being overly focused on money and power, excessively rule-driven, and too involved in political matters (27).

Implications of Spirituality for Physical Therapy Practice, and Education

Incorporating spirituality into physical therapy practice holds significant implications for practice and education. The integration of spirituality can enhance patient-centered care, promote holistic healing, and improve overall treatment adherence. These implications will be explored in greater detail throughout the discussion.

Physical Therapy Practice Implications

Spirituality, as a source of personal meaning and values, may play a critical role in delivering person-centered care and promoting cultural competence in physical therapy (7,9,31,32). A culturally competent system can deliver more holistic and effective care by adhering to key values and principles that focus on designing and implementing services tailored to the unique needs of individuals, children, and families (33). A culturally competent system requires understanding a person or family's cultural identity, as well as their levels of assimilation, in order to effectively apply the principle of "starting where the individual or family is." (33) Additionally, cultural competence involves collaborating with natural and informal support networks within diverse communities, such as neighborhood organizations, civic and advocacy groups, ethnic and social organizations, religious institutions, and, where appropriate, spiritual healers (33). Recognizing the role of spirituality within these networks is crucial, as it often forms a vital component of cultural identity and well-being.

It is worth noting that cultural competence is a fundamental aspect of the American Physical Therapy Association's (APTA) vision to *'transform society by optimizing movement to improve the human experience.'* (33) Several of APTA's guiding principles for achieving this vision are directly linked to cultural competence (33). For example, '*consumer-centricity'* emphasizes the need for PTs to prioritize patient values, goals, and individual needs in care, which includes recognizing cultural (including spiritual) factors (33). *'Access/equity'* focuses on addressing health disparities and SDHs through innovative and inclusive care models, which may involve considering patients' spiritual beliefs (33). *'Advocacy'* highlights the role of PTs in promoting patient-centered care by driving change in healthcare systems (33). Integrating spirituality into these principles ensures holistic, culturally competent care that respects and responds to patients' diverse backgrounds (33).

In addition, research indicates that addressing the spiritual needs of patients in healthcare can enhance patients' psychological well-being and satisfaction with care (34). This aligns with the biopsychosocial model, which emphasizes the importance of addressing psychological and social factors alongside physical health (35). Recognizing the spiritual dimensions of health allows PTs to create more comprehensive treatment plans that resonate with patients' values and beliefs. Moreover, spiritual considerations can foster better therapeutic relationships. When PTs engage with patients on a spiritual level, it cultivates trust and openness, facilitating more effective communication (36). This can lead to increased motivation and adherence to treatment plans, as patients feel more understood and supported (37).

In clinical practice, it is not uncommon for patients and therapists to hold differing beliefs, including those related to spirituality, its meaning, and its implications. When such differences arise, it is crucial for PTs to cultivate self-awareness, which means acknowledging one's own spirituality in a way that does not presume anything about the patient's spirituality or impose anything on the patient. In addition, it is crucial for PTs to broaden their appreciation for diverse expressions of faith and spirituality in their patients. This can be done through didactic training and hands-on experience (33). While PTs do not provide direct spiritual care, they should be prepared to respectfully acknowledge and support a patient's spiritual framework as it intersects with their physical therapy journey (33). For specific spiritual needs, similar to other SDH that fall outside the scope of physical therapy, PTs can refer patients to qualified professionals such as chaplains, spiritual care providers, religious leaders (e.g., priests, pastors, imams, or rabbis), spiritual counselors, or mental health professionals with expertise in spirituality (38). These referrals ensure patients receive appropriate support tailored to their beliefs and needs (38).

Educational Implications

The integration of spirituality in physical therapy education is crucial for preparing future practitioners to address the diverse needs of their patients. Current curricula often lack comprehensive training on spiritual care, which can lead to discomfort or inadequacy in handling these topics in practice (39). However, evidence has demonstrated the role of spirituality in patients' ability to reshape and interpret life events, as a coping mechanism, as a tool for pain management, and as a part of wellness and cultural competence (40). Educators are encouraged to incorporate spirituality into existing courses to ensure that learning experiences are intentional and address their role in patient care, ethical considerations, and therapeutic alliance. Furthermore, training programs that include reflective practices on spirituality can enhance students' self-awareness and empathy (41). This not only prepares students to engage more effectively with patients but also promotes their own well-being, helping to mitigate burnout and compassion fatigue often experienced in the healthcare field (42).

Relatedly, part of the curriculum should address religious trauma, including how to recognize it in oneself as a clinician and in one's patients, and what resources are available to navigate it. A large study published in 2023 gave evidence that as many as 1 in 5, or 20%, of U.S. adults, have suffered or are suffering from hurtful and harmful experiences with religion. At the very least, PTs need to be trained to approach the subject of spirituality with sensitivity and empathy, allowing the patient to lead and articulate the terms of engagement (43).

Examples of how to Incorporate Spirituality into Clinical Practice: Focus on Assessment and Measures

To our knowledge, there is no universally accepted "best" method for incorporating spirituality into clinical practice, as effective integration depends on the specific context, individual patient needs, and provider expertise. Nevertheless, best practices generally encompass a combination of comprehensive assessment, clear communication, and personalized interventions that respect and align with patients' beliefs and values. This section highlights key gaps in integrating spirituality into clinical practice, offers examples of how PTs can assess and measure patients' spirituality, and describes a systematic review with consensus-suggested implications

for how to address spirituality in serious illness and health outcomes (9).

Despite the guidance of the APTA and the plethora of literature supporting the need to incorporate a patient's spirituality into their session, PTs may not be incorporating these principles into their plan of care (44). Research shows (44) that while 96% of PTs believe spiritual well-being is a vital component of health, only 30% feel that addressing spiritual concerns falls within their responsibilities. That leads to the question of whose role it is to address spirituality in healthcare. Secondly, what barriers do PTs encounter if and when they address spirituality? Research findings also revealed a lack of education in taking spiritual history and navigating spiritual beliefs. This raises another question: What is the PT's obligation in terms of moral responsibility, code of ethics, and policy to address the patient's spirituality in the plan of care?

Similar to physical therapy, other healthcare disciplines recognize a gap in their integration of spiritual assessment in patient care (e.g., The Joint Commission and the American College of Physicians) (7,45). The Joint Commission on Accreditation of Healthcare Organizations (JCAHO) requires the administration of a spiritual assessment as of 2001. JCAHO's requirements address three areas (1) important spiritual practices, (2) denomination or faith tradition, and (3) significant spiritual beliefs with a list of questions to help guide the clinician in the assessment. Allowing the patient to answer neutrally is the goal of the JCAHO to assist with respecting the individual's spirituality. With this knowledge, JCAHO mandates healthcare organizations, including those with physical therapy, to take into consideration the spiritual needs of the patient and the barriers to creating a policy that is consistent across the field.

Balboni et al. highlighted that the limited provision of spiritual care for patients with serious illnesses is partly due to care team members not adequately addressing patient spirituality (9). Screening for spiritual needs is frequently overlooked, possibly due to time constraints, a belief that addressing spirituality falls outside the clinician's responsibilities, or discomfort in discussing such topics with patients (9), as well as due to concerns about provider competence and/or patient vulnerability in the medical environment (7). In clinical practice, clinicians may incorporate brief assessments to perform initial screenings that address spiritual care (9). This may include asking straightforward spiritual history questions such as, "Is spirituality or faith important to you in relation to your health and illness?" or "Do you have, or would you like, someone to talk to about spiritual or faith matters?" (9) This approach ensures that spiritual needs are integrated into comprehensive patient care and reflects respect for the patient's spiritual values (9).

Other examples of spirituality measures include multidimensional scales, such as the Bio-psycho-socio-spiritual Inventory (BioPSSI), Perceived Wellness Survey (PWS), Positive Mental Health Measurement Scale (PMH), and the WHOQOL-100 (46). Examples of questionnaires that more directly and comprehensively assess spirituality include the ISCS, (22,47) Functional Assessment of Chronic Illness FACIT-Sp, (12,48) SWBQ, (49,50) the SHALOM (51), and the STS (52).

The ISCS assesses levels of spiritual connection within spiritually diverse populations and is designed for use with clients in healthcare settings (22,47). It is a unidimensional scale consisting of 13 items tapping into the level and centrality of spiritual connection (22,47). It also contains 1 frame of reference item in which clients indicate the source of their spiritual connection (e.g., nature, God/Allah, multiple gods, and the universe) (22,47). The frame of reference items allows clinicians to quickly and directly understand if spirituality is important to their client and if so, the source of their clients' spiritual connection (22,47). This item has the capacity to reduce provider discomfort and increase client agency within the exchange (22,47). ISCS is grounded in theory, operationalizes spirituality from an expansive framework, utilizes inclusive language, and demonstrates strong psychometric properties (22,47). With a growing portion of US adults engaging in diverse spiritual expressions, the ISCS allows for a more inclusive assessment approach (53).

The FACIT-Sp is a cross-culturally validated measure designed to assess a patient's state of spiritual well-being, utilizing a multidimensional framework (48). It consists of 12 items with a Likert-type response scale and contains a subscale (Meaning and Peace) designed to assess spirituality from a more expansive framework (48). FACIT-Sp is a useful measure if the assessment of current spiritual well-being state is desired (in contrast to measurement of spiritual well-being as a trait), (48) as it frames questions based on the respondents' last 7 days (22).

Both the SWBQ and SHALOM instruments are available in multiple languages and have validated measurement properties (49–51). The SWBQ is a 20-item tool that assesses four key dimensions of spirituality: personal (reflecting one's internal sense of meaning, purpose, and values), communal (focused on the quality of interpersonal relationships, including love, justice, and hope), environmental (concerned with the individual's connection to nature and the environment), and transcendental (addressing beliefs in and relationships with a higher power, such as God, and the associated faith, adoration, and worship) (49,50).

The SHALOM (51) consists of two sets of 20 items, identical to those in the SWBQ. In the first set, respondents are asked to indicate what they believe the ideal levels for each descriptor should be (the "ideal" component) (51). In the second set, they are asked to rate how the descriptors reflect their own experiences over the past six months (the "lived experience" component) (51). In contrast to the SWBQ, which assesses spirituality by evaluating current states or traits without comparing them to ideal levels, the SHALOM incorporates both ideal and lived experiences. This allows for a comparison between an individual's spiritual aspirations and their actual experiences, helping to identify gaps in spiritual well-being (49–51).

Lastly, the STS is a 24-item measure designed to assess spirituality as an expansive and central element of the human experience (52). The measure utilizes a 5-point Likert-type response scale across three subscales (connectedness, universality, and prayer fulfillment) (52). Evidence of adequate internal consistency and convergent validity has been demonstrated (52). A unique aspect of the STS is the peer

evaluation form that can be completed concurrently. The STS peer evaluation form provides an avenue for engaging core social support figures in a client's holistic health journey.

In 2022, a systematic review and multidisciplinary Delphi panel (9) recommended six implications, three for addressing spirituality in serious illness and three for addressing spirituality in health outcomes. The primary implications, supported by strong evidence, for how to address spirituality in serious illness were: "(1) incorporate spiritual care into care for patients with serious illness; (2) incorporate spiritual care education into training of interdisciplinary teams caring for persons with serious illness; and (3) include specialty practitioners of spiritual care in care of patients with serious illness." The primary implications, supported by strong evidence, for how to address spirituality in health outcomes were: "(1) incorporate patient-centered and evidence-based approaches regarding associations of spiritual community with improved patient and population health outcomes; (2) increase awareness among health professionals of evidence for protective health associations of spiritual community; and (3) recognize spirituality as a social factor associated with health in research, community assessments, and program implementation." This systematic review was based on high-quality evidence and expert appraisal, and the referenced recommendations underscore the importance of integrating spirituality into healthcare practice, education, and research to enhance patient care and health outcomes.

Conclusion

Incorporating spirituality into physical therapy practice has significant implications for enhancing patient care and addressing diverse health outcomes. A holistic approach that includes the spiritual dimension allows clinicians to address the complex needs of their patients more effectively, leading to improved health outcomes and greater patient satisfaction. Recognizing spirituality as a key determinant of health supports the development of treatment plans that align with patients' values, fostering stronger therapeutic relationships. As the field of physical therapy evolves, integrating spirituality into both clinical practice and educational curricula will be essential for providing compassionate, culturally competent care. This holistic perspective may not only enhance patient care but can also enable PTs to connect with their patients on a deeper level, potentially fostering resilience. A variety of spirituality assessment tools, such as ISCS, FACIT-Sp, SWBQ, SHALOM, and STS, provide healthcare providers with comprehensive frameworks for evaluating spiritual well-being, ensuring a more inclusive and personalized approach to meeting patients' diverse spiritual needs.

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' Contributor Role: AG and AE were responsible for the conceptualization of the study. All authors participated in writing the manuscript and critically revised it for important intellectual content. The final manuscript was approved by all authors, who have read and agreed to the published version.

Data Availability Statement: Data sharing not applicable.

References

1. Kiles TM, Cernasev A, Leibold C, et al. Patient perspectives of discussing social determinants of health with community pharmacists. J Am Pharm Assoc (2003). 2022;62(3):826-833. CrossRef
2. Rethorn ZD, Cook C, Reneker JC. Social Determinants of Health: If You Aren't Measuring Them, You Aren't Seeing the Big Picture. J Orthop Sports Phys Ther. 2019;49(12):872-874. CrossRef PubMed
3. Bezner JR. Promoting Health and Wellness: Implications for Physical Therapist Practice. Phys Ther. 2015;95(10):1433-1444. CrossRef PubMed
4. Dhar N, Chaturvedi SK, Nandan D. Spiritual health, the fourth dimension: a public health perspective. WHO South-East Asia J Public Health. 2013;2(1):3-5. CrossRef PubMed
5. Puchalski CM, Vitillo R, Hull SK, et al. Improving the spiritual dimension of whole person care: reaching national and international consensus. J Palliat Med. 2014;17(6):642-656. CrossRef PubMed
6. Long KNG, Symons X, VanderWeele TJ, et al. Spirituality As A Determinant Of Health: Emerging Policies, Practices, And Systems. Health Aff (Millwood). 2024;43(6):783-790. CrossRef PubMed
7. Pearce MJ. Addressing religion and spirituality in health care systems. In: An Applied Psychology of Religion and Spirituality. APA handbooks in psychology. American Psychological Association; 2013:527-541. APA Handbook of Psychology, Religion, and Spirituality; vol 2., CrossRef.
8. Idler E, Idler EL, eds. Religion: The Invisible Social Determinant. In: Religion as a Social Determinant of Public Health. Oxford University Press; 2014:0. CrossRef
9. Balboni TA, VanderWeele TJ, Doan-Soares SD, et al. Spirituality in Serious Illness and Health. JAMA. 2022;328(2):184-197. CrossRef PubMed
10. Oman D. Defining religion and spirituality. In: Handbook of the Psychology of Religion and Spirituality. 2nd ed. The Guilford Press; 2013:23-47. Online (Accessed November 2024)
11. Paul Victor CG, Treschuk JV. Critical Literature Review on the Definition Clarity of the Concept of Faith, Religion, and Spirituality. J Holist Nurs. 2020;38(1):107-113. CrossRef PubMed
12. Monod S, Brennan M, Rochat E, et al. Instruments measuring spirituality in clinical research: a systematic review. J Gen Intern Med. 2011;26(11):1345-1357. CrossRef PubMed
13. Pargament KI, ed. Searching for the sacred: Toward a nonreductionistic theory of spirituality. In: APA Handbook of Psychology, Religion, and Spirituality (Vol 1): Context, Theory, and Research. APA handbooks in psychology®. American Psychological Association; 2013:257-273. CrossRef
14. Sessanna L, Finnell DS, Underhill M, et al. Measures assessing spirituality as more than religiosity: a methodological review of nursing and health-related literature. J Adv Nurs. 2011;67(8):1677-1694. CrossRef PubMed
15. Kallo BAA Michael Rotolo, Patricia Tevington, Justin Nortey, et al. Who are 'spiritual but not religious' Americans? Pew Research Center. December 7, 2023. Online (Accessed November 2024)
16. Cox D. Religion, the Spiritual Dimension and Social Development. In: Midgley J, Pawar M, eds. Future Directions in Social Development. 2017:187-204, CrossRef.

17. Villani D, Sorgente A, Iannello P, et al. The Role of Spirituality and Religiosity in Subjective Well-Being of Individuals With Different Religious Status. Front Psychol. 2019;10:1525. CrossRef PubMed
18. Fisher J. Development and Application of a Spiritual Well-Being Questionnaire Called SHALOM. Religions (Basel). 2010; 1(1):105-121. CrossRef
19. Gang JA, Gang GR, Gharibvand L. Integration of Spirituality and Whole Person Care into Doctor of Physical Therapy Curricula: A Mixed Methods Study. J Allied Health. 2022;51(2):130-135. PubMed
20. U.S. Department of Veteran Affairs. Spirit & Soul – Whole Health. Online (Accessed November 2024)
21. National Wellness – National Wellness Institute. March 18, 2020. Online (Accessed November 2024)
22. Hoots VM. Conceptualization and Measurement of Spirituality: Towards the Development of a Nontheistic Spirituality Measure for Use in Health-Related Fields. Electronic Theses and Dissertations. Published online December 1, 2017. Online (Accessed November 2024)
23. Ai AL, Wink P, Shearer M. Secular reverence predicts shorter hospital length of stay among middle-aged and older patients following open-heart surgery. J Behav Med. 2011;34(6):532-541. CrossRef PubMed
24. Holt-Lunstad J, Steffen PR, Sandberg J, et al. Understanding the connection between spiritual well-being and physical health: an examination of ambulatory blood pressure, inflammation, blood lipids and fasting glucose. J Behav Med. 2011;34(6):477-488. CrossRef PubMed
25. Mueller PS, Plevak DJ, Rummans TA. Religious involvement, spirituality, and medicine: implications for clinical practice. Mayo Clin Proc. 2001;76(12):1225-1235. CrossRef PubMed
26. Koenig HG. Religion, spirituality, and health: a review and update. Adv Mind Body Med. 2015;29(3):19-26. PubMed
27. Wormald B. Chapter 3: Views of Religious Institutions. Pew Research Center. November 3, 2015. Online (Accessed November 2024)
28. Tiwana MH, Smith J. Faith and vaccination: a scoping review of the relationships between religious beliefs and vaccine hesitancy. BMC Public Health. 2024;24(1):1806. CrossRef PubMed
29. Alsuwaidi AR, Hammad HAA, Elbarazi I, et al. Vaccine hesitancy within the Muslim community: islamic faith and public health perspectives. Hum Vaccin Immunother. 2023;19(1):2190716. CrossRef PubMed
30. Bowler K. A History of the American Prosperity Gospel. Oxford University Press; 2013. CrossRef
31. Institute of Medicine (US) Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. National Academies Press (US); 2001. Online (Accessed November 2024)
32. APTA. Cultural Competence in Physical Therapy. Online (Accessed November 2024)
33. APTA. Achieving Cultural Competence. January 15, 2020. Accessed September 16, 2024. Online (Accessed November 2024)
34. Hodge DR. Administering a two-stage spiritual assessment in healthcare settings: a necessary component of ethical and effective care. J Nurs Manag. 2015;23(1):27-38. CrossRef PubMed
35. Engel GL. The need for a new medical model: a challenge for biomedicine. Science. 1977;196(4286):129-136. CrossRef PubMed
36. Giske T, Cone P. Comparing Nurses' and Patients' Comfort Level with Spiritual Assessment. Religions (Basel). 2020;11(12):671. CrossRef
37. de Brito Sena MA, Damiano RF, Lucchetti G, et al. Defining Spirituality in Healthcare: A Systematic Review and Conceptual Framework. Front Psychol. 2021;12:756080. CrossRef PubMed
38. APTA. APTA Guide for Professional Conduct. Online (Accessed November 2024)
39. Puchalski C, Ferrell B, Virani R, et al. Improving the quality of spiritual care as a dimension of palliative care: the report of the Consensus Conference. J Palliat Med. 2009;12(10):885-904. CrossRef PubMed
40. Sargeant DM. Teaching Spirituality in the Physical Therapy Classroom and Clinic. J Phys Ther Educ. 2009;23(1):29-35. CrossRef
41. Gonçalves JPB, Lucchetti G, Menezes PR, et al. Religious and spiritual interventions in mental health care: a systematic review and meta-analysis of randomized controlled clinical trials. Psychol Med. 2015;45(14):2937-2949. CrossRef PubMed
42. Kabat-Zinn J. Mindfulness-based interventions in context: Past, present, and future. Clin Psychol (New York). 2003;10(2): 144-156. CrossRef
43. Slade D, Smell A, Wilson E, et al. Percentage of U.S. Adults Suffering from Religious Trauma. SHERM Journal. 2023. Online (Accessed November 2024)
44. Oakley ET, Sauer K, Dent B, et al. Physical Therapists' Perception of Spirituality and Patient Care: Beliefs, Practices, and Perceived Barriers. J Phys Ther Educ. 2010;24(2):45-52. CrossRef
45. Hodge DR. A template for spiritual assessment: a review of the JCAHO requirements and guidelines for implementation. Soc Work. 2006;51(4):317-326. CrossRef PubMed
46. Lindert J, Bain PA, Kubzansky LD, et al. Well-being measurement and the WHO health policy Health 2010: systematic review of measurement scales. Eur J Public Health. 2015;25(4):731-740. CrossRef PubMed
47. Hoots V. Measurement of Nontheistic and Theistic Spirituality: Initial Psychometric Qualities of the Inclusive Spiritual Connection Scale. Electronic Theses and Dissertations. Published online December 1, 2020. Online (Accessed November 2024)
48. Peterman AH, Fitchett G, Brady MJ, et al. Measuring spiritual well-being in people with cancer: the functional assessment of chronic illness therapy--Spiritual Well-being Scale (FACIT-Sp). Ann Behav Med. 2002;24(1):49-58. CrossRef PubMed
49. Gomez R, Watson S. A Reevaluation of the Factor Structure, Reliability, and Validity of the Spiritual Well-Being Questionnaire (SWBQ). J Relig Health. 2023;62(3):2112-2130. CrossRef PubMed
50. Gomez R, Fisher JW. Domains of spiritual well-being and development and validation of the Spiritual Well-Being Questionnaire. Pers Individ Dif. 2003;35(8):1975-1991. CrossRef
51. Fisher JW. Validation and Utilisation of the Spiritual Well-Being Questionnaire: SHALOM. J Relig Health. 2021;60(5):3694-3715. CrossRef PubMed
52. Piedmont RL. Does Spirituality Represent the Sixth Factor of Personality? Spiritual Transcendence and the Five-Factor Model. J Pers. 1999;67(6):985-1013. CrossRef
53. Kallo BAA Michael Rotolo, Patricia Tevington, Justin Nortey, et al. Spirituality Among Americans. Pew Research Center. December 7, 2023. Online (Accessed November 2024)