Arch Physioter 2025; 15: 250-264
ISSN 2057-0082 | DOI: 10.33393/aop.2025.3428
REVIEW

OPEN ACCESS

Metric properties of the Tonic Stretch Reflex Threshold (TSRT) as a measure of spasticity: a systematic review with meta-analysis

Diego Longo[1], Marco Baccini[2], Giulio Cherubini[1], Aurora Caporaso[3], Mario De Marco[1], Angela Politi[2], Guido Santini[1], Francesca Cecchi[1,2], Maria Angela Bagni[1]

[1]Department of Experimental and Clinical Medicine, University of Florence, Florence - Italy
[2]IRCCS Fondazione Don Gnocchi Onlus, Florence - Italy
[3]Bachelor's degree programme in Physiotherapy, University of Florence, Florence - Italy

ABSTRACT

Introduction: Spasticity is a common symptom after brain injury, often interfering with functional recovery and rehabilitation. The Tonic Stretch Reflex Threshold (TSRT) was proposed as an objective neurophysiological assessment of spasticity that could overcome the limitations of clinical scales. This systematic review aimed to appraise the current evidence on the metric properties of TSRT.

Methods: Electronic databases (MEDLINE, CINAHL, Scopus, Web of Science, and EMBASE) were screened from inception to June 30, 2025, for studies reporting data on reliability, validity, and/or responsiveness of TSRT in adults with stroke. Two reviewers independently selected the studies, assessed the methodological quality, and extracted relevant data. When possible, pooled estimates for each property were computed.

Results: Of the 9804 titles retrieved, 17 were eventually included, to which 2 articles from cross-references were added. We found insufficient values for both intra-rater (two studies, ICC = 0.548, 0.330-0.710) and inter-rater (three studies, ICC = 0.687, 0.511-0.808) reliability, with high measurement error. Data on validity were found in 14 articles, with conflicting results on the association of TSRT with clinical scales of spasticity and motricity, but good ability to discriminate among relevant groups. Only one study investigated responsiveness with an external anchor, finding that TSRT measurements failed to accurately detect improved participants.

Conclusion: Despite the potential of TSRT as a measure of spasticity, its metric properties, particularly reliability, are not fully supported. Future research should prioritize improving its reliability and investigating its validity and responsiveness with neurophysiological measures rather than relying solely on clinical scales.

Keywords: Adult, Muscle spasticity, Reproducibility of results, Stretch reflex, Stroke

What is already known about this topic?

- *Clinical scales for spasticity have limited reliability and validity. The Tonic Stretch Reflex Threshold (TSRT) has been proposed as an objective measure of spasticity for use in both clinical and research settings.*

What the study adds?

- *The reliability of the TSRT measurement is currently not fully supported, and this might affect its validity and responsiveness. Future research should focus on improving reliability through the use of better instrumentation.*

Introduction

Spasticity is one common and disabling symptom after a first motor neuron injury (1), with prevalence ranging from 25% to 38% in stroke survivors (2-4). It interferes with functional recovery and rehabilitation processes, can lead to secondary complications such as muscle retractions, weakness, and pain (5), and impairs activities of daily living and sleep (6).

The pathophysiological mechanisms of post-stroke spasticity are still debated, and several neuromusculoskeletal computational models have been used to study them. Most models focus on reflex gains and thresholds as the main parameters related to its severity, including both neural (alpha and gamma motor neuron firing, afferent input mainly from muscle spindles) and non-neural (viscoelastic elements of the muscular-tendon unit) components (7). The neural component is generally accepted as being caused primarily by supraspinal

Received: December 10, 2024
Accepted: August 21, 2025
Published online: September 19, 2025

This article includes supplementary material

Corresponding author:
Marco Baccini
email: marco.baccini@unifi.it

dysregulation of the spinal reflex loop, to which growing evidence indicates that imbalance of the dorsal and medial reticulo-spinal tracts contributes (8). The uncontrasted excitatory input through the medial reticulo-spinal tract would increase the intrinsic excitability of motoneurons, facilitating persistent inward currents, i.e., depolarizing currents that tend to remain activated and are associated with subthreshold depolarization of spinal motor neurons (8).

There is still no complete agreement on the definition of spasticity. Early in 1980 Lance (9) defined spasticity as "A motor disorder caused by a speed-reflex increase of the stretch tonic reflex with exaggerated tendon response, resulting from a hyperexcitability of the stretch reflex as a component of the syndrome of the first motoneuron". Later, Pandyan et al. (10) extended the concept to all positive symptoms of an upper motor neuron lesion redefining it as "disordered sensorimotor control, presenting as intermittent, or sustained involuntary involvement of muscles". More recently, Dressler et al. (11) suggested a simple, operational definition as "involuntary muscle hyper-activity triggered by rapid passive joint movements," which differentiates it from other symptoms such as rigidity, spasm or dystonia, and from hypertonia, which may also exist at rest. In 2017, 37 experts from 12 European countries participating in two consensus meeting, based on a Delphi approach (12), concluded that the term "hyper-resistance," instead of hypertonia or spasticity, should be preferred to describe the abnormal neuromuscular response during passive stretch, to which both neural and non-neural components contribute, and that the term "spasticity" should be used with care, and only referring, in agreement with Lance (9), to velocity dependent stretch hyperreflexia as part of the neural contributions to hyper-resistance.

The lack of a common definition is at least partly behind the variety of methods for assessing spasticity. In a recent review, 60 different tools were identified (13), including 33 clinical, 18 biomechanical, and eight neurophysiological measures. The Ashworth scale (AS) and its modified version (MAS) are the most used clinical measures of spasticity in the adult population. However, besides showing questionable metric properties (14,15), they assess resistance to passive movement without distinguishing between neural and non-neural components of resistance (13,14), and show low correlation with neurophysiological tests like H-reflex-based measurements (16,17). The poor validity of clinical scales, including the AS, has even been suggested as one of the causes of the failure of many clinical trials in neurology (18).

A direct measurement of spasticity has been proposed in the context of indirect, referent control theory of motor control (19,20). A main mechanism in posture and movement control is the ability to set and reset the spatial threshold of reflexes, named lambda (λ). In the healthy nervous system, the threshold regulatory range is defined by the task-specific ability to activate or relax muscles at each joint position within the biomechanical range of motion (21). Thus, the range of threshold regulation is larger than the biomechanical range, whereas, after neurological injury, the range of possible λ-shifts is limited. Spasticity is the result of a reduction in the central regulatory range of λ thresholds: when the upper limit of λ range (R+) of a muscle is within the biomechanical range of motion, an externally imposed stretching will find an active muscular resistance beginning at R+ and increasing with increasing lengthening. The λ threshold decreases as the rate of muscle stretch increases, and its equivalent parameter at the joint level is called stretch reflex threshold (SRT) (19), which represents the joint angle at which motoneurons of the involved muscles begin to be recruited. This parameter depends on stretch velocity and is commonly referred to as Dynamic Stretch Reflex Threshold (DSRT), whereas the Tonic Stretch Reflex Threshold (TSRT) represents a specific value of DSRT for zero velocity. In particular, when the TSRT is within the biomechanical range and the subject cannot shift it, it discriminates the joint configurations in which the muscles are spastic from those in which they are not (22,23).

Another parameter detected through TSRT measurement is coefficient μ, a time-dimensional parameter that characterizes the velocity sensitivity of TSRT (19). Theoretically, it correlates with the index of sensitivity, changing from dynamic to static muscle spindle afferents, and may be partly controlled by dynamic γ-motor neurons' activity (19). In a velocity/angle plot, μ represents the slope of the dynamic activation area boundary of the muscle and determines a temporal variation in the entry of the muscle into the activation area (19). This parameter is as essential as TSRT, because it is crucial to understand the actual dependence on the velocity of the observed hypertonia.

Surface electromyography (sEMG) combined with a joint angle detection system is used to detect DSRTs (24). The assessment procedure consists of applying several passive stretches to the target muscles at different speeds. For each speed, DSRT is identified as the angle at which reflex EMG activity appears. A linear regression through the DSRTs obtained from repeated stretches is conducted to compute the TSRT (24). In the original procedure, stretches are applied manually, and measures are taken with the Montreal Spasticity Measure instrumentation, which comprises an electrogoniometer, a two-channel surface electromyograph, and dedicated software (24). A complete description on how the TSRT is calculated can be found in a recent critical review on this topic (25), which, however, highlights that the methodologies used by different researchers vary greatly in terms of equipment, number, velocity and mode of application of stretches (manually or with a torque motor), time interval between stretches, criteria for defining the onset of reflex muscle activity, analyses conducted to calculate TSRT and μ. All these features considerably impact the measurement and might explain some inconsistencies in the findings of different research groups.

The aim of this review is to critically appraise, summarize, and meta-analyze the current literature on the metric properties of TSRT as a measure of spasticity in adults with upper motor neuron lesions.

Methods

This review was conducted according to COSMIN (Consensus-Based Standards for the Selection of Health Measurement Instruments) guidelines (26). It was registered prospectively in the PROSPERO database (PROSPERO 2023 CRD42023412289) as a more general systematic review on instrumental methods for assessing spasticity.

Data Sources and Searches

We searched five databases (MEDLINE, EMBASE, CINAHL, SCOPUS, and WOS) for literature published in English on April 20, 2023, with no initial date limit. The search strategy followed the COSMIN guidelines (27) and contained the following words: (1) spasticity, (2) tone, (3) cerebral palsy, (4) stroke, (5) spinal cord injury, (6) upper motor neuron, (7) measure, (8) evaluation, and (9) assessment. A combination of logic Boolean operators was implemented to finalize the existing COSMIN search filters for each database (Supplementary material). The reference list of all relevant articles was also searched for other eligible publications. An updated search was conducted on June 30, 2025, including only the term Tonic Stretch Reflex Threshold or TSRT.

Study Selection

We included studies that investigated any metric property (28) of the TSRT measurement in adults with spasticity, with no restrictions regarding spasticity etiology. We included only studies where the TSRT was measured following the procedure described by Calota et al. (24), so excluding the articles where this term was used with a different meaning, e.g., to simply indicate the joint angle of onset of reflex muscle contraction triggered by a passive stretch, even if instrumentally measured. Subjects with cerebral palsy outcomes were included if they were assessed when older than 18 years. Studies reporting data on metric properties were included even if this evaluation was not the primary aim. Two authors independently performed the study selection. After removing duplicates, titles, abstracts, and, when needed, full texts were analyzed for eligibility. Disagreements were solved by a third researcher.

Data Extraction

Two authors independently extracted the following data from the articles: authors, year of publication, population, comparator tool, when applicable, muscles evaluated, metric properties evaluated, statistics conducted, and values found for each property, as detailed below. In case of disagreement, consensus was reached with the intervention of a third author.

Reliability

Intraclass correlation coefficients (ICC) were considered the most appropriate index, and the statistics used for their computation were also recorded, when provided; other correlation coefficients were also considered appropriate if the study assessed and excluded significant differences between the two measurements. Standard Error of the Measurement (SEM), Minimal Detectable Change (MDC), Coefficient of Variation, and Bland-Altman's Limits of Agreement were considered appropriate indices of measurement error.

Criterion validity

If the study used a reference measure for spasticity, the correlation with the TSRT was noted. If the study dichotomized the threshold measurement result (e.g., spasticity present/absent), diagnostic accuracy values (sensitivity, specificity, area under the receiver operating characteristic curve, AUC) were collected.

Construct validity

All data about the a priori hypotheses tested were collected, including convergent (e.g., correlation between TSRT and motor impairment severity) and discriminative validity, i.e., the ability of the TSRT to discriminate among relevant subgroups.

Responsiveness

Any index of responsiveness, both internal (e.g., effect size) or external (e.g., correlation with an anchor measure of change), was collected. When provided, the Minimal Clinically Important Difference (MCID) computed in the study was noted.

Quality Assessment

Two authors independently assessed the risk of bias of the included articles, and disagreements were resolved by consensus, with the intervention of a third author. For reliability studies, the Quality Appraisal for Reliability (QAREL) checklist was used (29), which consists of 11 items that can be answered "yes" if the item is satisfied, "no" if the item is not satisfied, "unclear" if the information is insufficient, or "not applicable." Although the COSMIN checklists on risk of bias (30) also cover reliability, the QAREL checklist was preferred for reliability because it lists all possible sources of error in more detail. The COSMIN checklists (30) were used instead for validity (boxes 8 and 9) and responsiveness (box 10). Each COSMIN item may be rated as "very good," "adequate," "doubtful," or "inadequate," when applicable, and the lowest score assigned to any item of a box is taken to establish the risk of bias.

Data Synthesis and Analysis

Pooled estimations were obtained after Fisher's z transformation of correlation coefficients to normalize their distribution (31), including the estimation of the prediction interval (PI), as recommended (32). All meta-analyses were based on random effects models to account for the heterogeneity of estimates (33). Heterogeneity and inconsistency were assessed using the Q statistic and the I^2 statistic, respectively (34).

When missing, SEM and MDC_{95} were computed provided that authors reported the necessary data, by applying the formulas $SEM = DS \times \sqrt{(1-ICC)}$ and $MDC = SEM \times 1{,}96 \times \sqrt{2}$.

Convergent validity was assessed through the correlation of TSRT and μ with clinical scales of spasticity and motricity. When relevant indices were not reported in the article, they were computed provided that the authors reported the necessary data. When data were provided in graphical form, values were extracted from graphs using Microsoft PowerPoint 2021. If this was not possible, the authors were contacted and asked to provide the requested values or raw data for computation.

Results

Out of 9804 titles retrieved from the databases, 17 (35-51) were ultimately included (see Fig. 1 for details on the selection process). Cross-referencing led to the inclusion of two other studies (19,52). Table 1 shows the characteristics of the included studies.

All studies enrolled individuals with stroke outcomes, more often in the chronic phase (at least six months after the event) (19,35,37,42,43,45,49,51,52). Four studies enrolled participants in the subacute phase (less than 6 months) (44,46, 48, 50), and five studies had a mixed population (36,38-41), whereas in one study (47), this information is not reported. In one study (44), the sample also included two patients with spinal cord injury and one with traumatic brain injury. Sample size ranged from 4 to 55 participants. TSRT was measured at the elbow flexors (19,35-37,39-50,52), elbow extensors (19,36,39), and ankle plantar flexors. (38,51)

Four articles (35,37,38,46) evaluated TSRT reliability, whereas in 14 articles (19,35,36,39,41-48,51,52) data about TSRT validity were reported or retrievable. One article (46) studied both internal and external responsiveness of TSRT, using the upper limb motor section of the Fugl-Meyer Assessment Scale (53) (FMA-UL) as an anchor for the latter.

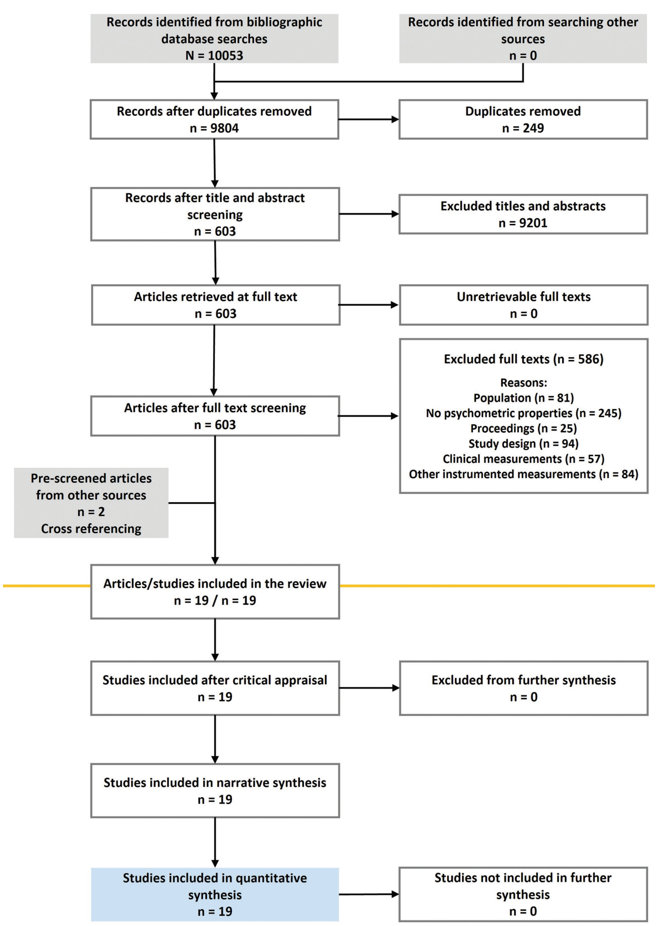

FIGURE 1 - Flow-chart of the process of study selection.

However, we were able to extract data about internal responsiveness from three other studies (40,49,50).

Reliability and measurement error

Data on intra-rater and inter-rater reliability were reported by two (35,37) and three (35,38,46) articles, respectively. Ferreira et al. (37) did not compute any reliability coefficient, so we analyzed their data to calculate the $ICC_{2,1}$, including only participants with stroke who completed all the assessments. Only one of these studies (46) provided the measurement error, but we computed it from the raw data reported in the other three articles (35,37,38). All studies obtained a QAREL score <50% (Fig. 2), indicating a possible high risk of bias.

Intra-rater reliability

The intra-rater reliability of biceps brachii TSRT measurement in a sample of subjects with stroke outcome was assessed by Calota et al. (35) (15-18 participants assessed twice, with a 2-7 days delay, by three different raters) and by Ferreira et al. (37) (14 participants assessed three times, with 2-5 days delay, by the same rater). We meta-analyzed the data of these studies (Fig. 3, top), finding a pooled ICC = 0.548 (95% CI: 0.330, 0.710; 95% PI: 0.330, 0.710; sample size = 64), with no heterogeneity and high consistency (Q = 0,778, p = 0.855; I^2 = 0,00%). The pooled MDC_{95} was 41°.

Inter-rater reliability

TSRT inter-rater reliability was evaluated for biceps brachii (35,46) and triceps surae (38), involving two (46), three (35), or nine (38) raters. Since Calota et al. (35) reported data measured by three raters at two different occasions, data from four comparisons were meta-analyzed. The pooled ICC was 0.687 (95% CI: 0.511, 0,808; 95% PI: 0.346, 0.868; I^2 = 50,41%; Q = 6,049, p = 0.109; sample size = 123) (Fig. 3, bottom). Including only data on biceps brachii resulted in a lower value of pooled ICC (0.618; 95% CI: 0.471, 0,732; 95% PI: 0.471, 0,732; sample size = 95), with no heterogeneity (Q = 0,454, p = 0,796) and inconsistency (I^2 = 0.00%). The MDC_{95} was 14.55° for triceps surae and 35.16° for biceps brachii.

We classified the overall rating of inter-rater and intra-rater reliability and measurement error of TSRT as insufficient for using the measure both at the group and individual levels. We rated the level of evidence as low (risk of bias and imprecision).

Validity

Convergent and discriminative validity of TSRT was evaluated by analyzing the correlation with measures of similar constructs (clinical scales of spasticity) or of a construct that might theoretically be associated (motor function), and comparing values measured in different groups (healthy, subacute or chronic stroke, Parkinson's disease), respectively. The quality of ratings of included studies is shown in Table 2. Usually, the authors did not state their a-priori hypotheses about the expected correlations.

TABLE 1 - Characteristics of included studies

Study	Participants, number, and condition (M/F)	Time since event, months (mean (SD))	Severity of paresis/ spasticity[§]	Age, years (mean (SD))	Muscle evaluated	Property evaluated[#]	Devices	Number of passive stretches	Algorithm for muscle activity onset	TSRT value[§]
Levin et al., 2000 (19)	11 CS (7/4)	47.3 (22.4)	FMA-UL: 37.3 (19.5); CSI: 10.6 (3.9)	41.2 (12.1)	EF-EE[ǂ]	Convergent validity (CSI, FMA-UL)	sEMG, manipulandum coupled to a torque motor and a digital resolver	56	>2 SD over baseline for 50ms at least	118.0 (29.2)
Calota et al., 2008 (35)	20 CS (16/4)	77.9 (111.7)	MAS: 1.7 (1.3)	62.9 (12.9)	EF	Reliability Convergent validity (MAS)	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	135.4 (24.1)
Kim et al., 2011 (36)	9 SS, 6CS (7/8)	6.7 (9.1)	MAS: 1.4 (0.4)	63.5 (15.6)	EF	Convergent validity (MAS)	sEMG, electrogoniometer	About 50	>2 SD over baseline	123.6 (4.6)
Ferreira et al., 2013 (37)	14 CS (11/3)	53.9 (38.4)	NR	60.21 (14.40)	EF	Reliability	sEMG, electrogoniometer	30	Modified double threshold protocol	188.7 (37.6)
Blanchette et al., 2016 (38)	2 SS, 26 CS (21/7)	47.9 (43.8)	NR	57.4 (9.7)	PF	Reliability	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	66.0 (13.1),
Turpin et al., 2017 (39)	2 SS, 11 CS (12/1)	62.3 (69.2)	FMA-UL: 36.6 (13.9); CSI: 9.1 (2.6)	56.1 (9.1)	EF-EE	Convergent validity (CSI, FMA-UL)	sEMG, manipulandum coupled to a torque motor and optical encoder	15	>3SD over baseline	105.3 (20.5)
Vieira et al., 2017 (40)	4 SS, 12 CS (12/4)	7.6 (2.5)	NR	63.9 (8.5)	EF	Internal responsiveness*	sEMG, electrogoniometer	30	>2SD over baseline for 200ms at least	69.7 (19.3)
Levin et al., 2018 (41)	13 SS, 20 CS (26/7)	20.4 (23.6)	FMA-UL: 22.2 (15.8); MAS: 1,4 (0.4)	52.2 (8.7)	EF	Convergent validity (MAS)	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	85.1 (13.5)
Subramanian et al., 2018 (42)	12 CS (10/2)	54.7 (71.4)	FMA-UL: 39.2 (12.0); CSI: 9.5 (2.8)	58.5 (9.5)	EF	Convergent validity (CSI, FMA-UL)*	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	134.8 (15.1)
Marques et al., 2019 (43)	10 CS (8/2)	35.4 (29.2)	MAS: 1.4 (0.4)	68.2 (11.6)	EF	Convergent validity (MAS)	sEMG, electrogoniometer	30	>2SD over baseline for 200ms at least	49.8 (57.8)
Zhang et al., 2019 (44)	14 SS, 1 TBI, 1 SCI (14/2)	1,7 (0.3)	MAS: 1.7 (0.6)	54 (10)	EF	Convergent validity (MAS)	sEMG, IMU	15-20	>3 SD over baseline	49.9 (21.7)
Alves et al., 2021 (45)	22 CS (16/6)	71.4 (139.3)	MAS: 1.4 (0.6)	56.1 (11.3)	EF	Convergent validity (MAS)	sEMG, electrogoniometer	30	>2 SD over baseline for 200ms at least	41.4 (33.7)

(*Continued*)

TABLE 1 - (*Continued*)

Study	Participants, number, and condition (M/F)	Time since event, months (mean (SD))	Severity of paresis/ spasticity§	Age, years (mean (SD))	Muscle evaluated	Property evaluated#	Devices	Number of passive stretches	Algorithm for muscle activity onset	TSRT value§
Frenkel-Toledo et al., 2021 (46)	55 SS (37/18)	2.2 (1.4)	FMA-UL: 29.6 (15.9); MAS: 1.4 (0.4)	54.3 (10.7)	EF	Reliability Convergent validity (MAS) Responsiveness	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	100.6 (20.3)
Wang et al., 2021 (47)	16 Stroke (11/5)	NR	NR	51.6 (13.6)	EF	Convergent validity (MAS)*	Exoskeletal device coupled to sEMG, torque sensor, and angular position sensor	9	NR	66.69 (19.29)
Frenkel-Toledo et al., 2022 (48)	45 SS (25/16)	2,2 (1,3)	FMA-UL: 32.3 (12.2)	53.66 (11.27	EF	Convergent validity (FMA-UL, MAS)	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	107.4 (25.7)
Marques et al., 2022 (49)	10 CS (8/2)	36 (20.8)	NR	59.7 (9.07)	EF	Internal responsiveness*	sEMG, electrogoniometer	6 at least	>2SD over baseline for 200ms at least	35.2 (28.3)
Levin et al., 2023 (50)	35 SS (25/10)	2.2 (1.4)	FMA-UL: 32.6 (13.1); MAS: 1.4 (0.4)	53.7 (11.3)	EF	Internal responsiveness*	MSM (sEMG, electrogoniometer)	20 at least	>2 SD over baseline for 25 ms at least	101.0 (20.2)
Longo et al., 2023 (51)	10 CS (6/4)	134.4 (124.3)	MAS: 2.6 (0.7)	52.1 (15.0)	PF	Convergent validity (MAS)*	MSM (sEMG, electrogoniometer)	20 at least	>3SD over baseline for 25 ms at least	88.2 (13.8)
Mullick et al., 2013 (52)	10 CS (6/4)	58.3 (19.5)	CSI 8.5 (2.2)	65.3 (12.2)	EF-EE	Convergent validity (CSI)	sEMG, manipulandum coupled to a torque motor and precision gauges	30	>2SD over baseline for 50 ms at least	133.7 (16.2)

§When multiple assessments were conducted (test-retest, pre-post intervention), data from the first assessment are reported, pooling data of different groups or different raters when needed.
#The asterisk indicates that the estimation of metric properties was not an aim of the study.
‡Elbow extensors assessed in 4 participants.
M/W = men/women (ratio); SS = subacute stroke; CS = chronic stroke; TBI = traumatic brain injury; SCI = spinal cord injury; FMA-UL = Fugl-Meyer assessment scale, upper limb motor section; CSI = Composite Spasticity Index; MAS = Modified Ashworth Scale; EF = elbow flexors; EE = elbow extensors; PF = plantar flexors; sEMG = surface Electromiography; IMU = inertial measurement units; MSM = Montreal Spasticity Measure; NR = not reported; SD = Standard Deviation; TSRT=Tonic Stretch Reflex Threshold.

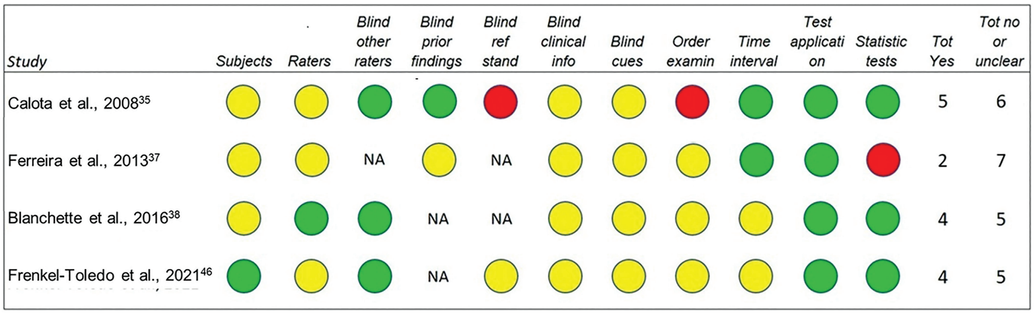

FIGURE 2 - Quality ratings of studies on TSRT reliability.

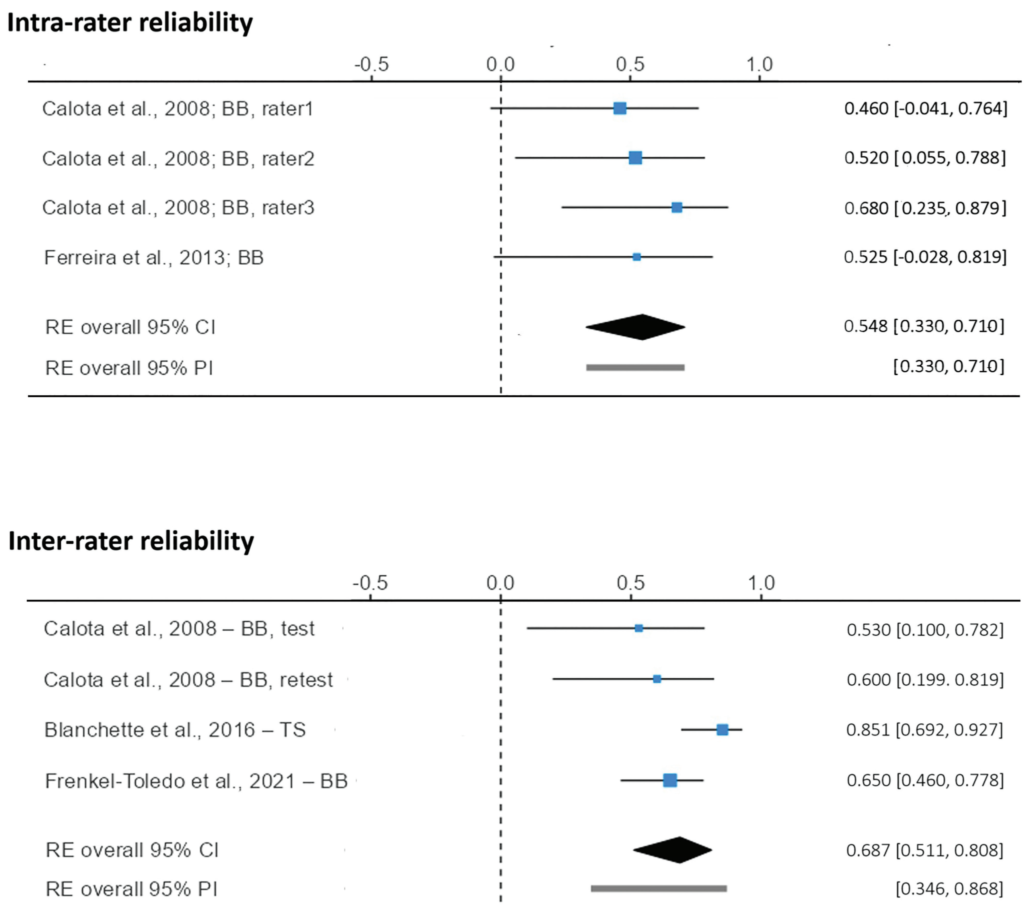

FIGURE 3 - Meta-analysis of studies on intra-rater reliability (top) and inter-rater reliability (bottom) of TSRT. *(BB = biceps brachii).*

We could retrieve data on the association of TSRT with spasticity clinical scales from 14 articles that used the Modified Ashworth Scale (MAS) (35,36,41,43-48,51) or the Composite Spasticity Index (CSI) (19,39,42,52). CSI (54) sums up three features, namely resistance to passive movement, clonus, and exaggerated tendon reflexes. Only few studies reported the Spearman correlation coefficient (19,35,43,45,46,48), so we computed it from the raw data of the other articles (36,39,41,42,44,47,51,52), in two cases by extracting values from graphs (44,47), and in one case from data provided on request by the authors (39). Figure 4 shows the meta-analysis conducted aggregating data from all studies, resulting in a pooled coefficient of −0.305 (95% CI: −0.500, −0.081; 95% PI: −0.822, 0.487; sample size = 283) with high heterogeneity and inconsistency (Q = 66.862, $p < 0.001$; I^2 = 74.57%). Pooling the data separately for each comparator, heterogeneity persists with MAS (rho = −0.327; 95% CI: −0.557, −0.050; PI: −0.853, 0.528; Q = 60.060, $p < 0.001$; I^2 = 81.68%; sample size = 238), not with CSI (rho = −0.240; 95% CI: −0.537, 0.011; PI: −0.670, 0.311; Q = 6.717, p = 0.243; I^2 = 25.56%; sample size = 45).

We estimated the association of biceps brachii TSRT with motor impairment as measured by the FMA-UL from data presented in five studies (19,39,41,42,48). Only one

TABLE 2 - Quality appraisal of studies on TSRT construct validity and responsiveness

	Convergent validity (COSMIN box 9a)					Discriminative validity (COSMIN box 9b)			
Study	**1**	**2**	**3**	**4**	**Risk-of-bias score**	**5**	**6**	**7**	**Risk-of-bias score**
Levin et al., 2000 (19) – CSI	VG	D	VG	D*	doubtful				
Levin et al., 2000 (19) – FMA-UL	VG	A	VG	D*	doubtful				
Calota et al., 2008 (35) – MAS	VG	D	VG	D*	doubtful				
Kim et al., 2011 (36) – MAS	VG	D	VG	D*	doubtful				
Turpin et al., 2017 (39) – CSI	VG	A	VG	D*	doubtful				
Turpin et al., 2017 (39) – FMA-UL	VG	A	VG	D*	doubtful				
Levin et al., 2018 (41) – MAS	VG	D	VG	D*	doubtful				
Levin et al., 2018 (41) – FMA-UL	VG	A	VG	D*	doubtful				
Levin et al., 2018 (41) – chronic vs subacute patients						VG	VG	D*	doubtful
Subramanian et al., 2018 (42) – CSI	VG	A	NA^	D*	doubtful				
Subramanian et al., 2018 (42) – FMA-UL	VG	A	NA^	D*	doubtful				
Mar+ques et al., 2019 (43) – MAS	VG	D	VG	D*	doubtful				
Zhang et al., 2019 (44) – MAS	VG	D	A	D*	doubtful				
Zhang et al., 2019 (44) – healthy vs spastic participants						VG	VG	D*	doubtful
Alves et al., 2021 (45) – MAS	VG	D	VG	D*	doubtful				
Frenkel-Toledo et al., 2021 (46) – MAS	VG	D	VG	D*	doubtful				
Frenkel-Toledo et al., 2021 (46) –FMA-UL	VG	A	VG	D*	doubtful				
Wang et al., 2021 (47) – MAS	VG	D	NA^	D*	doubtful				
Frenkel-Toledo et al., 2022 (49) – MAS	VG	D	VG	D*	doubtful				
Frenkel-Toledo et al., 2022 (49) –FMA-UL	VG	A	VG	D*	doubtful				
Longo et al., 2023 (51)	VG	D	NA^	D*	doubtful				
Mullick et al., 2013 (52) – CSI	VG	A	VG	D*	doubtful				
Mullick et al., 2013 (52) – UPDRS	VG	VG	VG	D*	doubtful				
Mullick et al., 2013 (52) – healthy, stroke vs Parkinson						VG	VG	D*	doubtful
	Responsiveness (COSMIN box 10a and 10d)								
Study	**10a_1**	**10a_2**	**10a_3**		**10d_11**	**10d_12**	**10d_13**		**Risk-of-bias score**
Frenkel-Toledo et al., 2021 (46)	VG	NA	D**		VG	VG	VG		doubtful

1: Is it clear what the comparator instrument(s) measure(s)? 2: Were the measurement properties of the comparator instrument(s) sufficient? 3: Was the statistical method appropriate for the hypotheses to be tested? 4: Were there any other important flaws in the design or statistical methods of the study? 5: Was an adequate description provided of important characteristics of the subgroups? 6: Was the statistical method appropriate for the hypotheses to be tested? 7, 10a_3, 10d_13: Were there any other important flaws in the design or statistical methods of the study? 10a_1: For continuous scores: were correlations between change scores, or the AUC calculated? 10a_2: For dichotomous scales: were sensitivity and specificity determined? 10d_11: Was an adequate description provided of the intervention given? 10d_12: Were statistical methods appropriate for the before-after comparison being made? VG = very good; AD = adequate; D = doubtful; IN = inadequate; N/A = not applicable.
* Blinding to the comparator or to groups.
** Suitability of the anchor.
^Not applicable because the study did not aim to evaluate the TSRT metric properties.

study (19) also measured triceps brachii TSRT and reported a significant positive correlation with motor impairment for both muscles, whereas one study reported a non-significant negative correlation (48); two studies (39,41) just reported the absence of significant correlation, and one study (42) did not investigate such an association. We computed the Spearman correlation from the raw data presented in the articles (41,42) or provided on request by the authors (39). The meta-analysis (Fig. 5, top) showed a pooled correlation coefficient = 0.117 (95% CI: −0.140, 0.360; PI: −0.293, 0.492; Q = 6.970; p = 0.223; I^2 = 28.26%; sample size = 109).

Only two studies also investigated the association between FMA-UL and µ of biceps brachii (19,41), with conflicting results, i.e., non-significant negative correlation (41), and significant positive correlation (19). A pooled rho = 0.119 (95% CI: −0.391, 0.574; PI: −0.575, 0.714; Q = 3.439; p = 0.064; I^2 = 50.9%; sample size = 44) resulted from the meta-analysis (Fig. 5, bottom).

We classified the overall rating of TSRT and µ convergent validity with clinical scales of motor impairment or spasticity as indeterminate, with a very low level of evidence (imprecision, risk of bias, and very serious inconsistency).

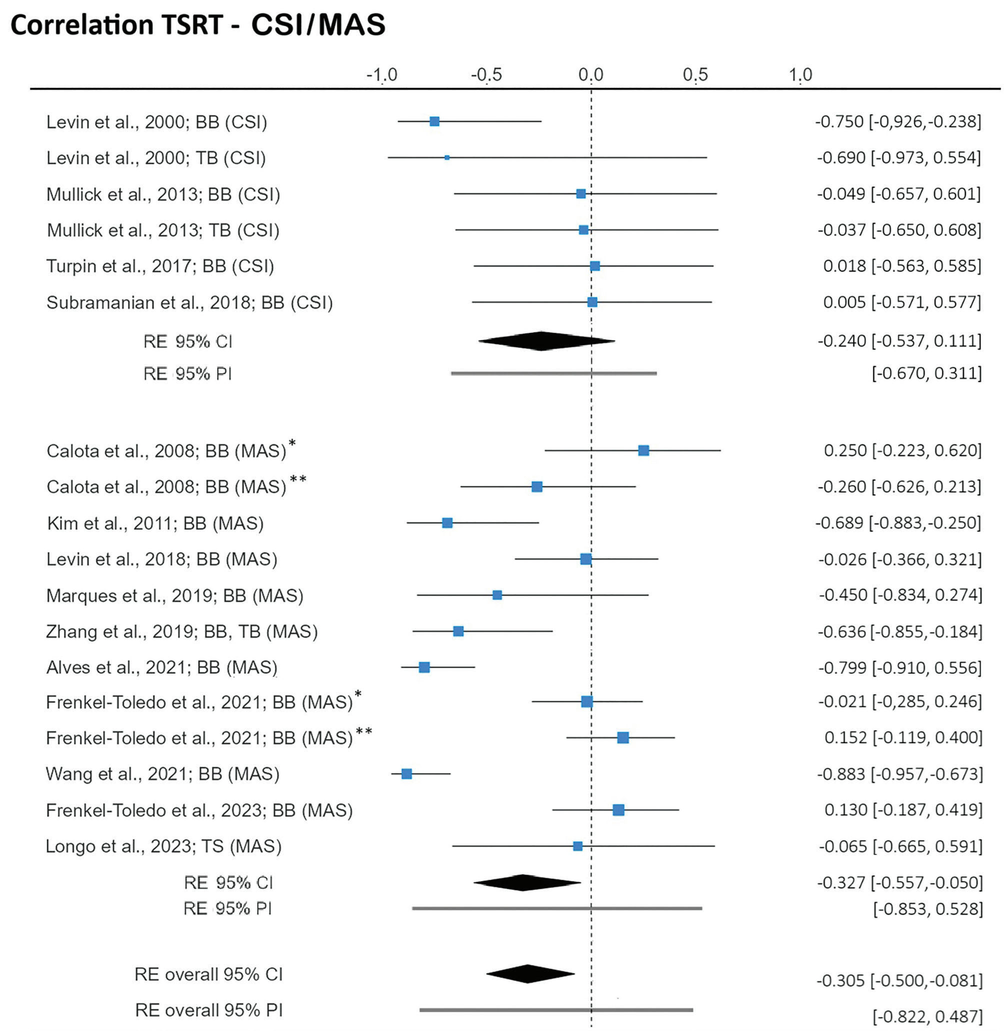

FIGURE 4 - Meta-analysis of studies on TSRT convergent validity: correlation with clinical scales of spasticity. (*BB = biceps brachii; TB = triceps brachii; MAS = Modified Ashworth Scale; CSI = Composite Spasticity Index; *measured at first assessment (test); **measured at second assessment (retest)*).

As for discriminative validity, significant differences were found between TSRT values measured in healthy and spastic subjects (44,52), and between TSRT and µ values measured in participants with stroke and Parkinson's disease (52). Conversely, no differences were found between persons with subacute and chronic stroke (41). We classified the overall rating of TSRT discriminative validity as positive, with a low level of evidence (imprecision, risk of bias).

Responsiveness

We found only one study (46) that comprehensively evaluated this property in 44 participants who received 10 sessions of upper limb intervention combining Virtual Reality and real/sham transcranial direct current stimulation. Contrary to MAS, TSRT changed significantly in the whole group, detecting a medium treatment effect (mean change: 8.79 ± 24.96; Effect Size = 0.40; Standardized Response Mean = 0.35). External responsiveness was examined using a receiver-operating characteristic curve approach and classifying participants as improved or unimproved according to a change in FMA-UL above or below, respectively, the MDC of this scale. A cut-off value of 6.8° TSRT change was found with near significance (Area Under Curve, AUC = 0.671, $p = 0.056$). The risk of bias of this study was rated as doubtful due to the doubtful suitability of the chosen anchor (Table 2).

We retrieved data on internal responsiveness (Effect Size, ES) in three other studies not specifically designed to evaluate this psychometric property. We found large ESs in a study (40) enrolling a small number of persons with stroke outcomes who were treated with biofeedback (N = 6, ES = 1.17) or conventional physical therapy (N = 6, ES = 0.86), and in another study (49) where 10 participants with chronic stroke practiced 15 sessions of training based on serious game and virtual reality (ES = 0.92). A third study (50) reported an ES = 0.570 in a group of 46 participants with chronic stroke who were engaged in 10 sessions of personalized reaching training. Considering the very different rehabilitation interventions, which probably have different effectiveness, we decided not to aggregate the data from these studies.

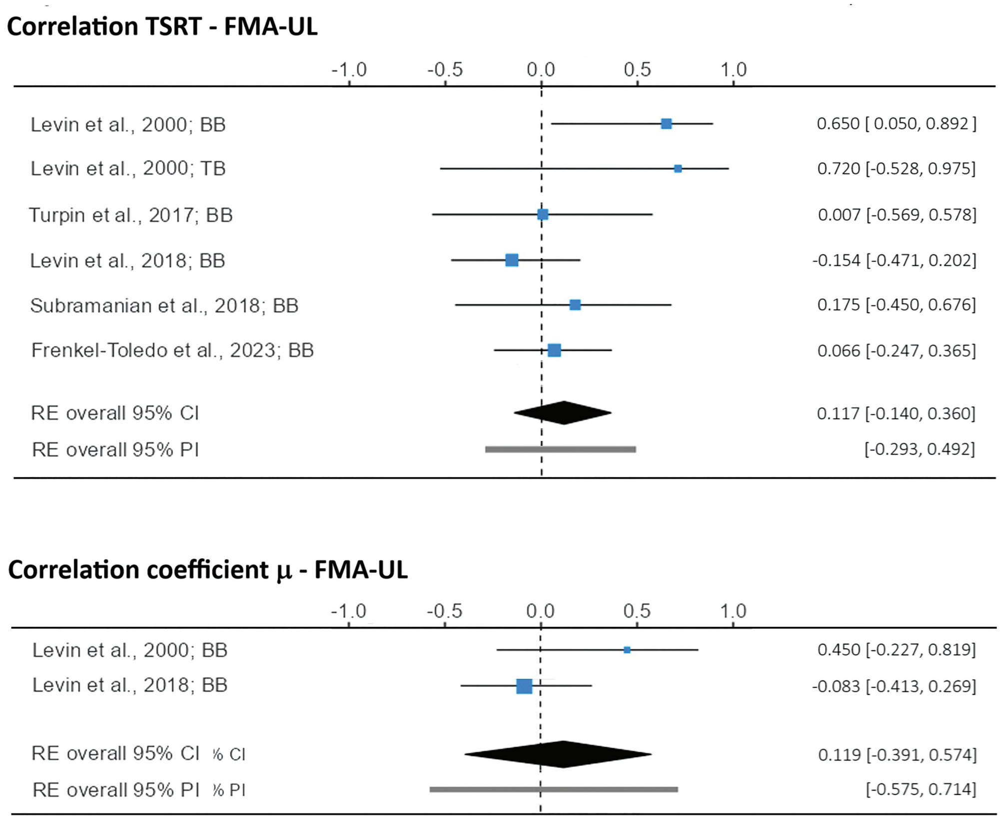

FIGURE 5 - Meta-analysis of studies on convergent validity of TSRT (top) and coefficient μ (bottom): correlation with Fugl-Meyer Assessment Scale – motricity section. *(BB = biceps brachii; TB = triceps brachii).*

We classified the overall rating of TSRT responsiveness as insufficient, with a very low level of evidence (imprecision, risk of bias).

Discussion

This systematic review investigated the psychometric properties of TSRT as a potential objective measure of spasticity, finding only four articles on reliability and up to 14 articles reporting data on its convergent validity, even if many of them do not explicitly refer to TSRT validity. This is surprising since reliability is a prerequisite for exploring criterion and construct validity of any assessment tool (28), but it is rather common in this field. A narrative review on new tools for assessing spasticity (55), i.e., tools that were developed from 2016 to 2022, identified four new clinical scales, eight medical imaging methods, and three spasticity assessment devices. From the data presented, it appears that none of them have been subjected to a thorough evaluation of metric properties, and in particular, reliability. In a recent systematic review (56) aimed at evaluating the quality of clinical assessment tools for spasticity, the Tardieu scale was the only clinician-reported instrument recommended, since it is consistent with the definition of spasticity as a speed-dependent increase of stretch reflexes. The authors, however, did not meta-analyze data on Tardieu scale metric properties and pointed out that its reliability has not been fully established.

We will first discuss the results concerning TSRT reliability because, in our opinion, they may partly explain the conflicting results regarding validity.

Reliability

The data presented show that TSRT reliability, either intra- or inter-rater, is not sufficiently supported. Although different criteria are used for ICC interpretation, the values found in the present review should be considered low when referring either to the COSMIN criteria (30) (ICC = 0.70 at least), or to the more conservative criteria proposed by Fitzpatrick et al. (57) (ICC = 0.70 or = 0.90 at least for use in assessing groups or individual performance, respectively). The insufficient reliability is confirmed by the high measurement error for both biceps brachii [over 35° out of a total range of about 150° (58)] and triceps surae [14.55° out of a total passive range of approximately 74° (58)]. For biceps brachii, MDC_{95} is far higher than the MCID of TSRT reported by Frenkel-Toledo et al. (46), so it should be rated as negative also according to COSMIN criteria (30).

One source of error might be the data collection procedure and, in particular, the assessment of joint angles through electrogoniometry. Very few studies evaluated the reliability of electrogoniometry, with results that raise serious doubts about this property. Two studies assessed the reliability of range of motion measurement at the ankle (59,60) in small samples (<20) of healthy subjects. The first study (59) only stated that no significant differences were found between the repeated measurements, whereas the other (60) reported high test-retest reliability (ICC = 0.987; SEM = 3,63°), but computing the average measure ICC, which overestimates reliability. Only one non-recent study (61) evaluated the reliability of electrogoniometry at the elbow in 23 female healthy volunteers, finding very low ICCs (ranging 0.00-0.43).

The main issue of this kind of assessment is the difficulty of avoiding small shifts of the goniometers' arms and axis during limb mobilization that might generate artefacts in data recording. In order to reduce this source of error, it would likely be better to replace the electrogoniometer with more reliable motion analysis instruments. Marker-based motion capture systems, for example, are considered a gold standard in human biomechanics assessment (62). Such systems are hardly feasible as a routine evaluation procedure in clinical practice, but a recent study reports that markerless systems could be quite reliable, with significant advantages in terms of usability and rapidity of procedures (63). Indeed, some authors (19,39,52) measured the TSRT with different devices (a horizontal manipulandum coupled to a torque motor, with an angular position sensor placed on the axis of the manipulandum to measure joint angles) that could reduce the random variability of the measurement. Unfortunately, none of them investigated the reliability of this alternative procedure.

sEMG has been more widely investigated in terms of reliability at both elbow (64-67) and ankle muscles (68-70), and showed high test-retest reproducibility in different conditions. However, a key point for TSRT measurement is the detection of the onset of muscle activity from EMG signals, which depends on the criteria adopted. Silva et al. (71) compared the performance of four different criteria – two of which were employed in two studies (36,38) included in the present review – and found that the TSRTs measured by each method differed significantly from one another. However, this should affect the validity of the measurements, rather than their reliability.

Validity

The lack of consistent correlations between TSRT and clinical measures of spasticity is somewhat surprising. However, some authors underline the difficulty in distinguishing between spasticity and spastic dystonia, the latter considered due to the secondary soft tissue changes that occur in upper motor neuron syndrome (72-74) because both features may co-occur in these patients (72). A clinical evaluation of the resistance to passive stretch could hardly discriminate between these intrinsic components of spasticity that are strictly related to the phase of recovery after stroke. For this reason, we could expect different results when enrolling chronic or subacute subjects, but such a difference was not evidenced by the results of our review. We found conflicting results between studies that enrolled only chronic patients (19,42,52) and between studies that enrolled only subacute patients (44,46). In both cases, we find studies reporting strong negative correlation (19,44) or no correlation (42,46, 52). Therefore, the phase after stroke does not seem to be a relevant factor for this association. Indeed, recent findings show that architectural changes in muscles are driven by both reduced mobility and muscle overactivity, which increase passive resistance and promote progression to fibrosis, and may also increase stretch sensitivity and reflex mechanisms (75).

In a very recent article, Piscitelli et al. (76) reported data about the correlation between elbow flexors TSRT measured in subjects with acute, subacute, or chronic stroke, and clinical scales of motor impairment (FMA-UL) and spasticity (MAS). The authors found a significant, negative correlation between TSRT and MAS in the total sample (N = 247) and in the subgroup of acute/early subacute participants (N = 158), and a significant, positive correlation between TSRT and FMA-UL in the subgroup of chronic participants (N = 33). We decided not to include these data in our review because in this article, the authors pooled data from eight studies conducted in their laboratory, five of which are unpublished studies and therefore cannot be assessed for risk of bias. Moreover, we already included two (35,39) of the three published studies reported in this article in our meta-analysis on the correlation of TSRT with clinical scales, both enrolling only, or overwhelmingly, participants with chronic stroke. The first one (35) found conflicting results on correlation with MAS at two different assessments, and the other (39) found no correlation with CSI nor with FMA-UL. We could not include in the meta-analysis the third study that enrolled participants with subacute stroke because the published article (50) reports no data on correlation with clinical scales and only provides data on changes in TSRT after a period of training, which was included in our analysis of TSRT internal responsiveness. Nevertheless, the pooled correlation coefficients found in the present meta-analysis are about the same as those reported by Piscitelli et al. (76) for the whole sample, both for MAS (−0.327 vs −0.32) and FMA-UL (0.117 vs 0.08), although in our analyses the high heterogeneity resulted in wide prediction intervals.

We believe that insufficient TSRT reliability may be a factor behind the high inter-study variability, but the lack of reliability of clinical scales (MAS and CSI) might also contribute. Indeed, poor to moderate reliability was found for both the original and the modified AS (14,77-79). Not sufficient data are available for CSI reliability; instead, it was evaluated only in two studies (80,81) in very limited samples of patients who had a stroke (N = 10), with different results. The authors reported ICC = 0.87 (80) and ICC = 0.97 (81), but did not compute the measurement error. From the data presented in the two articles, we estimated an MDC_{95} of 3.4 and 0.67, respectively, i.e., quite different values.

Conflicting results might also depend on the method used to measure TSRT, which varied among studies. However, this does not seem to be the case, as conflicting data were also reported in studies that adopted the same method. For example, two studies on participants with chronic stroke (19,52) adopted same device, same algorithm to detect EMG activity onset, and same clinical scale as a comparator (CSI). However, results ranged from a strong, negative correlation (19) to no correlation (52).

Similar considerations can be made for the conflicting results on the association between TSRT and motor scales; in this case, problems mainly concern the reliability of TSRT, since the reliability of the FMA scale is fairly well established (82,83). However, the lack of correlation with motor scales is not necessarily an unexpected finding; in fact, TSRT and FMA assess quite different aspects of motor control, and we may find people with severe paresis who have either high spasticity or no spasticity. Thus, the conflicting findings might also be related to the very limited sample sizes, ranging from 4 to 33 participants, who are likely not fully representative of the population with stroke outcomes. It is worth noting that one study (39) found that motor impairments as measured by FMA-UL were negatively associated with the ability to modulate stretch reflex thresholds between active and passive stretch rather than to passive-state values. This is an interesting finding, because during active stretch, inhibitory circuits responsible for reciprocal inhibition may be involved (84), as well as central inhibition of motor neurons (85), and both may be related to impairment of active movement (72). In our opinion, this finding could support the validity of TSRT measurements more than the correlation between motor scales and TSRT during passive stretch, but it needs to be confirmed in studies with adequate power.

Given the lack of reliable and valid clinical scales for spasticity, and the different constructs assessed by motor function scales, not necessarily correlated with spasticity as defined by Lance (9), we would have expected to find studies comparing TSRT to other neurophysiological measures, such as H-reflex-related measures. The H-reflex is a neurophysiological technique for evaluating spasticity that offers insights into spinal cord excitability and sensorimotor integration and can be employed to investigate different spinal mechanisms involved in motor control, such as presynaptic and reciprocal inhibition (86). Indeed, we found two articles (87,88) that assessed the association between some parameters detected by H-reflex test and a parameter somewhat related to TSRT, named Reflex Threshold Angle (RTA), measured in plantar flexors of patients with post-stroke spasticity. Both studies found that dysfunctions in reciprocal and presynaptic inhibition, as measured by H-reflex, correlated very weakly with RTA. However, RTA is defined as the angle at which a simultaneous increase in torque and EMG activity occurs during a passive stretch (87). It seems fully comparable with DSRT rather than with TSRT, so these studies were excluded from the present review. Future research should investigate whether TSRT, which specifies where the spastic zone begins along the muscle length, shows a stronger association with H-reflex parameters.

Conversely, the findings on between-groups differences, although obtained on still very limited samples, strongly support the discriminative validity of the measure, in particular with regards to differences between spasticity (participants who had a stroke) and rigidity (participants with Parkinson's disease). It should be noted that, in group comparisons, less precision of measurements is required, so insufficient reliability has less impact on the results. This indirectly supports the hypothesis that reliability problems affect the results of correlations with clinical scales.

Responsiveness

Responsiveness of TSRT has not been largely investigated, but the single study that comprehensively evaluated this property (46) reported rather negative results. Although the estimation of responsiveness is still a matter of debate, and several different methods are employed (89), only the use of an external anchor makes it possible to evaluate whether an instrument is able to detect important changes (90). In this respect, Frenkel-Toledo et al. (46) found that TSRT measurements failed in detecting participants who had changed, since the AUC was not significantly different from 0.50. However, two factors should be considered when interpreting their results. First, the sample size (N = 44) is too low for evaluating responsiveness (30). Most important, the authors used changes in FMA-UL as an external anchor, a questionable choice considering the low correlation between the two measures, as highlighted above. Future research on this property should enrol an adequate number of subjects and choose different anchors. We believe that subjective anchors such as the Global Rating of Perceived Change questionnaire are hardly applicable to this symptom, and instrumental measures of spasticity should be preferred. Conceivably, H-reflex-related measures are better suited to detect true changes in spasticity and could therefore be a more appropriate anchor in responsiveness studies.

Limitations

Some limitations impose caution in interpretating the results, primarily the low quality of many studies, the small sample size, and the high heterogeneity of many meta-analytical estimates. A more direct limitation might be the failure to retrieve all relevant titles, since about 10% of the articles were not found through database searches. However, this can largely be explained by the fact that in these studies, the evaluation of the TSRT metric properties was not an objective or was not explicitly stated. Indeed, 1 out of 2 articles included by cross-referencing (52) did not conduct any analyses of metric properties, and we re-analyzed the data reported in this article to compute correlations between TSRT and clinical scales, whereas the other article (19) is a study on motor control where the authors just mention results on correlation between TSRT and clinical scales in the main text. Nevertheless, we can wonder whether useful data, particularly on the association with clinical scales, might be found in other published studies.

Conclusions

Despite the potential of TSRT as an objective measure of spasticity, its metric properties are currently not fully supported, particularly with regard to reliability. Likely, the high measurement error contributes at least in part to the inconsistent results regarding convergent validity with clinical measures of spasticity. In order to reduce the random error, the use of alternative instruments to the electrogoniometer is recommended, as already implemented by some authors. Future research should prioritize improving the reliability of

TSRT by using alternative instrumentation. Once reliability is established, the validity and responsiveness of the measure should be evaluated using neurophysiological measures as a reference standard rather than relying solely on clinical spasticity scales, because the reliability and validity of the latter are also questionable.

Disclosures

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

Financial support: The funders played no role in the design, conduct, or reporting of this study.

Author's contributor role: DL: study design, studies selection process, quality appraisal, writing – draft; MB: study design, quality appraisal, data analysis, writing – editing; GC: studies selection process, writing – editing; AC: studies selection process, writing – editing; MDM: quality appraisal, writing – draft; AP: data extraction, writing – draft; GS: data extraction, writing – editing; FC: study design, writing – editing; MAB: study design, writing – editing.

Data Availability Statement: The data presented in this study are not publicly available but are available on request from the corresponding author.

References

1. Gallichio JE. Pharmacologic management of spasticity following stroke. Phys Ther. 2004;84(10):973-981. CrossRef
2. Schinwelski M, Sławek J. Prevalence of spasticity following stroke and its impact on quality of life with emphasis on disability in activities of daily living. Systematic review. Neurol Neurochir Pol. 2010;44(4):404-411. CrossRef
3. Zeng H, Chen J, Guo Y, et al. Prevalence and risk factors for spasticity after stroke: a systematic review and meta-analysis. Front Neurol. 2021;11:616097. CrossRef
4. Cheng H, Fang X, Liao L, et al. Prevalence and factors influencing the occurrence of spasticity in stroke patients: a retrospective study. Neurol Res. 2023;45(2):166-172. CrossRef
5. Kheder A, Nair KP. Spasticity: pathophysiology, evaluation and management. Pract Neurol. 2012;12(5):289-298. CrossRef
6. Bethoux F, Marrie RA. A cross-sectional study of the impact of spasticity on daily activities in multiple sclerosis. Patient. 2016;9(6):537-546. CrossRef
7. da Silva VA, da Silva RL, Withers JW, et al. Neuromusculoskeletal modeling of spasticity: A scoping review. PLoS One. 2025;20(5):e0320153. CrossRef
8. Li S, Chen YT, Francisco GE, Zhou P, Rymer WZ. A Unifying Pathophysiological Account for Post-stroke Spasticity and Disordered Motor Control. Front Neurol. 2019;10:468. CrossRef
9. Lance JW. The control of muscle tone, reflexes, and movement: Robert Wartenberg Lecture. Neurology. 1980;30(12):1303-1313. CrossRef
10. Pandyan AD, Gregoric M, Barnes MP, et al. Spasticity: clinical perceptions, neurological realities and meaningful measurement. Disabil Rehabil. 2005;27(1-2):2-6. CrossRef
11. Dressler D, Bhidayasiri R, Bohlega S, et al. Defining spasticity: a new approach considering current movement disorders terminology and botulinum toxin therapy. J Neurol. 2018;265(4):856-862. CrossRef
12. van den Noort JC, Bar-On L, Aertbeliën E, et al. European consensus on the concepts and measurement of the pathophysiological neuromuscular responses to passive muscle stretch. Eur J Neurol. 2017;24(7):981-e38. CrossRef
13. Aloraini SM, Alyosuf EY, Aloraini LI, Aldaihan MM. Assessment of spasticity: an overview of systematic reviews. Phys Ther Rev. 2022;27(4):290-303. CrossRef
14. Fleuren JF, Voerman GE, Erren-Wolters CV, et al. Stop using the Ashworth Scale for the assessment of spasticity. J Neurol Neurosurg Psychiatry. 2010;81(1):46-52. CrossRef
15. Pandyan AD, Johnson GR, Price CI, et al. A review of the properties and limitations of the Ashworth and modified Ashworth Scales as measures of spasticity. Clin Rehabil. 1999;13(5):373-383. CrossRef
16. Stokic DS, Yablon SA, Hayes A. Comparison of clinical and neurophysiologic responses to intrathecal baclofen bolus administration in moderate-to-severe spasticity after acquired brain injury. Arch Phys Med Rehabil. 2005;86(9):1801-1806. CrossRef
17. Martens G, Deltombe T, Foidart-Dessalle M, et al. Clinical and electrophysiological investigation of spastic muscle overactivity in patients with disorders of consciousness following severe brain injury. Clin Neurophysiol. 2019;130(2):207-213. CrossRef
18. Hobart JC, Cano SJ, Zajicek JP, et al. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations [published correction appears in Lancet Neurol 2008 7: 25]. Lancet Neurol. 2007;6(12):1094-1105. CrossRef
19. Levin MF, Selles RW, Verheul MH, et al. Deficits in the coordination of agonist and antagonist muscles in stroke patients: implications for normal motor control. Brain Res. 2000;853(2):352-369. CrossRef
20. Feldman AG. Indirect, referent control of motor actions underlies directional tuning of neurons. J Neurophysiol. 2019;121(3):823-841. CrossRef
21. Feldman AG, Levin MF. Spatial control of reflexes, posture and movement in normal conditions and after neurological lesions. J Hum Kinet. 2016;52(1):21-34. CrossRef
22. Latash ML, Levin MF, Scholz JP, et al. Motor control theories and their applications. Medicina (Kaunas). 2010;46(6):382-392. CrossRef
23. Feldman AG. The nature of voluntary control of motor actions. In: Latash ML, Lestienne F, eds. Motor Control and Learning. Springer; 2006. CrossRef
24. Calota A, Levin MF. Tonic stretch reflex threshold as a measure of spasticity: implications for clinical practice. Top Stroke Rehabil. 2009;16(3):177-188. CrossRef
25. Levin MF, Piscitelli D, Khayat J. Tonic stretch reflex threshold as a measure of disordered motor control and spasticity - a critical review. Clin Neurophysiol. 2024;165:138-150. CrossRef
26. Elsman EBM, Mokkink LB, Terwee CB, et al. Guideline for reporting systematic reviews of outcome measurement instruments (OMIs): PRISMA-COSMIN for OMIs 2024. Qual Life Res. 2024;33(8):2029-2046. CrossRef
27. Terwee CB, Jansma EP, Riphagen II, et al. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res. 2009;18(8):1115-1123. CrossRef
28. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737-745. CrossRef
29. Lucas NP, Macaskill P, Irwig L, et al . The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). J Clin Epidemiol. 2010;63(8):854-861. CrossRef
30. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Qual Life Res. 2018;27(5):1171-1179. CrossRef

31. Fisher RA. 014: On the "probable error" of a coefficient of correlation deduced from a small sample. Metron. 1921;1:3-32.
32. IntHout J, Ioannidis JP, Rovers MM, et al. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open. 2016;6(7):e010247. CrossRef
33. Der Simonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-188. CrossRef
34. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539-1558. CrossRef
35. Calota A, Feldman AG, Levin MF. Spasticity measurement based on tonic stretch reflex threshold in stroke using a portable device. Clin Neurophysiol. 2008;119(10):2329-2337. CrossRef
36. Kim KS, Seo JH, Song CG. Portable measurement system for the objective evaluation of the spasticity of hemiplegic patients based on the tonic stretch reflex threshold. Med Eng Phys. 2011;33(1):62-69. CrossRef
37. Ferreira JF, Moreira V, Lima CS, et al. Development, test and validation of a mechatronic device for spasticity quantification. IJARS. 2013;10(5). CrossRef
38. Blanchette AK, Mullick AA, Moïn-Darbari K, et al. Tonic stretch reflex threshold as a measure of ankle plantar-flexor spasticity after stroke. Phys Ther. 2016;96(5):687-695. CrossRef
39. Turpin NA, Feldman AG, Levin MF. Stretch-reflex threshold modulation during active elbow movements in post-stroke survivors with spasticity. Clin Neurophysiol. 2017;128(10): 1891-1897. CrossRef
40. Vieira D, Silva MB, Melo MC, et al. Effect of myofeedback on the threshold of the stretch reflex response of post-stroke spastic patients. Disabil Rehabil. 2017;39(5):458-467. CrossRef
41. Levin MF, Solomon JM, Shah A, et al. Activation of elbow extensors during passive stretch of flexors in patients with post-stroke spasticity. Clin Neurophysiol. 2018;129(10):2065-2074. CrossRef
42. Subramanian SK, Feldman AG, Levin MF. Spasticity may obscure motor learning ability after stroke. J Neurophysiol. 2018;119(1):5-20. CrossRef
43. Marques IA, Silva MB, Silva AN, et al. Measurement of post-stroke spasticity based on tonic stretch reflex threshold: implications of stretch velocity for clinical practice. Disabil Rehabil. 2019;41(2):219-225. CrossRef
44. Zhang X, Tang X, Zhu X, et al. A regression-based framework for quantitative assessment of muscle spasticity using combined emg and inertial data from wearable sensors. Front Neurosci. 2019;13:398. CrossRef
45. Alves CM, Rezende AR, Marques IA, et al. SpES: a new portable device for objective assessment of hypertonia in clinical practice. Comput Biol Med. 2021;134:104486. CrossRef
46. Frenkel-Toledo S, Solomon JM, Shah A, et al. Tonic stretch reflex threshold as a measure of spasticity after stroke: Reliability, minimal detectable change and responsiveness. Clin Neurophysiol. 2021;132(6):1226-1233. CrossRef
47. Wang C, Peng L, Hou ZG, et al. The assessment of upper-limb spasticity based on a multi-layer process using a portable measurement system. IEEE Trans Neural Syst Rehabil Eng. 2021;29:2242-2251. CrossRef
48. Frenkel-Toledo S, Levin MF, Berman S, et al. Shared and distinct voxel-based lesion-symptom mappings for spasticity and impaired movement in the hemiparetic upper limb. Sci Rep. 2022;12(1):10169. CrossRef
49. Marques IA, Alves CM, Rezende AR, et al. Virtual reality and serious game therapy for post-stroke individuals: a preliminary study with humanized rehabilitation approach protocol. Complement Ther Clin Pract. 2022;49:101681. CrossRef
50. Levin MF, Berman S, Weiss N, et al. ENHANCE proof-of-concept three-arm randomized trial: effects of reaching training of the hemiparetic upper limb restricted to the spasticity-free elbow range. Sci Rep. 2023;13(1):22934. CrossRef
51. Longo D, Santini G, Cherubini G, et al. The muscle shortening maneuver in individuals with stroke: a consideration-of-concept randomized pilot trial. Top Stroke Rehabil. 2023;30(8): 807-819. CrossRef
52. Mullick AA, Musampa NK, Feldman AG, et al. Stretch reflex spatial threshold measure discriminates between spasticity and rigidity. Clin Neurophysiol. 2013;124(4):740-751. CrossRef
53. Fugl-Meyer AR, Jääskö L, Leyman I, et al. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. J Rehabil Med. 1975;7(1):13-31. CrossRef
54. Levin MF, Hui-Chan CW. Relief of hemiparetic spasticity by TENS is associated with improvement in reflex and voluntary motor functions. Evok Pot. 1992;85(2):131-142. CrossRef
55. He J, Luo A, Yu J, et al. Quantitative assessment of spasticity: a narrative review of novel approaches and technologies. Front Neurol. 2023;14:1121323. CrossRef
56. Gal O, Baude M, Deltombe T, et al. Clinical outcome assessments for spasticity: review, critique, and recommendations. Mov Disord. 2025;40(1):22-43. CrossRef
57. Fitzpatrick R, Davey C, Buxton MJ, et al. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess. 1998;2(14):i-74. CrossRef
58. Neumann DA. Kinesiology of the Musculoskeletal System. 3rd ed. Mosby; 2016.
59. Rome K, Cowieson F. A reliability study of the universal goniometer, fluid goniometer, and electrogoniometer for the measurement of ankle dorsiflexion. Foot Ankle Int. 1996;17(1):28-32. CrossRef
60. Bronner S, Agraharasamakulam S, Ojofeitimi S. Reliability and validity of electrogoniometry measurement of lower extremity movement. J Med Eng Technol. 2010;34(3):232-242. CrossRef
61. Goodwin J, Clark C, Deakes J, et al. Clinical methods of goniometry: a comparative study. Disabil Rehabil. 1992;14(1):10-15. CrossRef
62. Robertson DGE, Caldwell GE, Hamil J. Research Methods in Biomechanics. 2nd ed. Human Kinetics; 2013.
63. Das K, de Paula Oliveira T, Newell J. Comparison of markerless and marker-based motion capture systems using 95% functional limits of agreement in a linear mixed-effects modelling framework. Sci Rep. 2023;13(1):22880. CrossRef
64. Gaudet G, Raison M, Maso FD, et al. Intra- and intersession reliability of surface electromyography on muscles actuating the forearm during maximum voluntary contractions. J Appl Biomech. 2016;32(6):558-570. CrossRef
65. Lee J, Jung MY, Kim SH. Reliability of spike and turn variables of surface EMG during isometric voluntary contractions of the biceps brachii muscle. J Electromyogr Kinesiol. 2011;21(1):119-127. CrossRef
66. Carius D, Kugler P, Kuhwald HM, et al. Absolute and relative intrasession reliability of surface EMG variables for voluntary precise forearm movements. J Electromyogr Kinesiol. 2015;25(6):860-869. CrossRef
67. Howatson G, Glaister M, Brouner J, et al. The reliability of electromechanical delay and torque during isometric and concentric isokinetic contractions. J Electromyogr Kinesiol. 2009;19(5):975-979. CrossRef
68. Gupta A, Mudie KL, Clothier PJ. The reliability of determining the onset of medial gastrocnemius muscle activity during a stretch-shorten-cycle action. J Electromyogr Kinesiol. 2014;24(5):588-592. CrossRef
69. Gallina A, Pollock CL, Vieira TM, et al. Between-day reliability of triceps surae responses to standing perturbations in people

post-stroke and healthy controls: a high-density surface EMG investigation. Gait Posture. 2016;44:103-109. CrossRef
70. Cè E, Rampichini S, Limonta E, et al. Fatigue effects on the electromechanical delay components during the relaxation phase after isometric contraction. Acta Physiol (Oxf). 2014;211(1):82-96. CrossRef
71. Silva MB, Silva AN, Naves EL, et al. An improved approach for measuring the tonic stretch reflex response of spastic muscles. Comput Biol Med. 2017;80:166-174. CrossRef
72. Sheean G, McGuire JR. Spastic hypertonia and movement disorders: pathophysiology, clinical presentation, and quantification. PM R. 2009;1(9):827-833. CrossRef
73. Marinelli L, Currà A, Trompetto C, et al. Spasticity and spastic dystonia: the two faces of velocity-dependent hypertonia. J Electromyogr Kinesiol. 2017;37:84-89. CrossRef
74. Trompetto C, Marinelli L, Mori L, et al. Pathophysiology of spasticity: implications for neurorehabilitation. BioMed Res Int. 2014;354906.
75. Raghavan P. Muscle physiology in spasticity and muscle stiffness. Toxicon. 2025;259:108350. CrossRef
76. Piscitelli D, Khayat J, Feldman AG, et al. Clinical relevance of the tonic stretch reflex threshold and μ as measures of upper limb spasticity and motor impairment after stroke. Neurorehabil Neural Repair. 2025;39(5):386-399. CrossRef
77. Ansari NN, Naghdi S, Arab TK, et al. The interrater and intrarater reliability of the Modified Ashworth Scale in the assessment of muscle spasticity: limb and muscle group effect. NeuroRehabilitation. 2008;23(3):231-237. CrossRef
78. Blackburn M, van Vliet P, Mockett SP. Reliability of measurements obtained with the modified Ashworth scale in the lower extremities of people with stroke. Phys Ther. 2002;82(1):25-34. CrossRef
79. Haas BM, Bergström E, Jamous A, et al. The inter rater reliability of the original and of the modified Ashworth scale for the assessment of spasticity in patients with spinal cord injury. Spinal Cord. 1996;34(9):560-564. CrossRef
80. Levin MF, Hui-Chan C. Are H and stretch reflexes in hemiparesis reproducible and correlated with spasticity? J Neurol. 1993;240(2):63-71. CrossRef
81. Ng SS, Hui-Chan CW. The timed up & go test: its reliability and association with lower-limb impairments and locomotor capacities in people with chronic stroke. Arch Phys Med Rehabil. 2005;86(8):1641-1647. CrossRef
82. Gladstone DJ, Danells CJ, Black SE. The fugl-meyer assessment of motor recovery after stroke: a critical review of its measurement properties. Neurorehabil Neural Repair. 2002;16(3):232-240. CrossRef
83. Hochleitner I, Pellicciari L, Castagnoli C, et al. Intra- and interrater reliability of the Italian Fugl-Meyer assessment of upper and lower extremity. Disabil Rehabil. 2023;45(18):2989-2999. CrossRef
84. Morita H, Crone C, Christenhuis D, et al. Modulation of presynaptic inhibition and disynaptic reciprocal Ia inhibition during voluntary movement in spasticity. Brain. 2001;124(4):826-837. CrossRef
85. Nielsen JB, Crone C, Hultborn H. The spinal pathophysiology of spasticity--from a basic science point of view. Acta Physiol (Oxf). 2007;189(2):171-180. CrossRef
86. Knikou M. The H-reflex as a probe: pathways and pitfalls. J Neurosci Methods. 2008;171(1):1-12. CrossRef
87. Allison SC, Abraham LD, Allison SC. Correlation of quantitative measures with the modified Ashworth scale in the assessment of plantar flexor spasticity in patients with traumatic brain injury. J Neurol. 1995;242(10):699-706. CrossRef
88. Annaswamy T, Mallempati S, Allison SC, et al. Measurement of plantarflexor spasticity in traumatic brain injury: correlational study of resistance torque compared with the modified Ashworth scale. Am J Phys Med Rehabil. 2007;86(5):404-411. CrossRef
89. Husted JA, Cook RJ, Farewell VT, et al. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol. 2000;53(5):459-468. CrossRef
90. Beaton DE, Bombardier C, Katz JN, et al. Looking for important change/differences in studies of responsiveness. OMERACT MCID Working Group. Outcome Measures in Rheumatology. Minimal Clinically Important Difference. J Rheumatol. 2001;28:400-405.