



The test-retest reliability of non-navigated transcranial magnetic stimulation (TMS) measures of corticospinal pathway excitability early after stroke

Kathryn C. Collins, Allan B. Clark, Valerie M. Pomeroy & Niamh C. Kennedy

To cite this article: Kathryn C. Collins, Allan B. Clark, Valerie M. Pomeroy & Niamh C. Kennedy (18 Apr 2024): The test-retest reliability of non-navigated transcranial magnetic stimulation (TMS) measures of corticospinal pathway excitability early after stroke, *Disability and Rehabilitation*, DOI: [10.1080/09638288.2024.2337107](https://doi.org/10.1080/09638288.2024.2337107)

To link to this article: <https://doi.org/10.1080/09638288.2024.2337107>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 18 Apr 2024.



[Submit your article to this journal](#)



Article views: 28



[View related articles](#)



[View Crossmark data](#)

The test-retest reliability of non-navigated transcranial magnetic stimulation (TMS) measures of corticospinal pathway excitability early after stroke

Kathryn C. Collins^a, Allan B. Clark^b, Valerie M. Pomeroy^{c,d} and Niamh C. Kennedy^e

^aFaculty of Health and Social Sciences, Bournemouth University, Bournemouth, UK; ^bNorwich Medical School, University of East Anglia, Norwich, UK; ^cAcquired Brain Injury Rehabilitation Alliance, School of Health Sciences, University of East Anglia, Norwich, UK; ^dNational Institute of Health Research Brain Injury MedTech Cooperative, Cambridge, UK; ^eSchool of Psychology, Ulster University, Coleraine, UK

ABSTRACT

Purpose: Motor evoked potential (MEP) characteristics are potential biomarkers of whether rehabilitation interventions drive motor recovery after stroke. The test-retest reliability of Transcranial Magnetic Stimulation (TMS) measurements in sub-acute stroke remains unclear. This study aims to determine the test-retest reliability of upper limb MEP measures elicited by non-neuronavigated transcranial magnetic stimulation in sub-acute-stroke.

Methods: In two identical data collection sessions, 1–3 days apart, TMS measures assessed: motor threshold (MT), amplitude, latency (MEP-L), silent period (SP), recruitment curve slope in the biceps brachii (BB), extensor carpi radialis (ECR), and abductor pollicis brevis (APB) muscles of paretic and non-paretic upper limbs. Test-retest reliability was calculated using the intra-class correlation coefficient (ICC) and 95% confidence intervals (CI). Acceptable reliability was set at a lower 95% CI of 0.70 or above. The limits of agreement (LOA) and smallest detectable change (SDC) were calculated.

Results: 30 participants with sub-acute stroke were included (av 36 days post stroke) reliability was variable between poor to good for the different MEP characteristics. The SDC values differed across muscles and MEP characteristics in both paretic and less paretic limbs.

Conclusions: The present findings indicate there is limited evidence for acceptable test-retest reliability of non-navigated TMS outcomes when using the appropriate 95% CI for ICC, SDC and LOA values.

Clinical Trial Registration: Current Controlled Trials: ISCRT 19090862, <http://www.controlled-trials.com>

ARTICLE HISTORY

Received 3 August 2021
Revised 20 March 2024
Accepted 24 March 2024

KEYWORDS

Test-retest reliability; transcranial magnetic stimulation (TMS); smallest detectable change; stroke; upper limb

> IMPLICATIONS FOR REHABILITATION

- This study identified that Non-navigated Transcranial Magnetic Stimulation (TMS) demonstrates low reliability of TMS measures in upper limb with variation between muscles and measures in sub-acute stroke
- When using non-navigated TMS to explore corticospinal pathway excitability the individual target muscle and TMS measure should be taken into consideration
- Non-navigated TMS may be more useful in exploring group differences rather than individual differences in corticospinal pathway excitability
- Non-navigated TMS could provide a means of measuring recovery in clinical practice and could inform the development of more effective interventions but this needs further development before it can be used as a clinical recovery biomarker

Introduction


Transcranial Magnetic stimulation (TMS) is a non-invasive method of brain stimulation [1,2] that has been used extensively in healthy participants to probe the Central Nervous System, investigate brain plasticity and to better understand the underlying neurophysiology of motor control [3–5]. This technique has increasingly been applied to clinical populations, in particular stroke for diagnostic or prognostic purposes [6,7] and to evaluate underlying neurological change in clinical trials [8,9].

Rehabilitation of motor function tailored to individual stroke survivors (precision stroke rehabilitation) is promised by the incorporation of biomarkers of recovery, into clinical trials and clinical practice [10]. Biomarkers are expected to advance the science and

practice of precision stroke rehabilitation by distinguishing sub-groups of stroke survivors and ascertaining whether specific rehabilitation interventions drive recovery. An international consensus recommends measuring the functional integrity of the corticospinal tract (CST), using TMS to identify the presence or otherwise of a motor-evoked potential (MEP) [11]. If TMS is going to be incorporated into clinical practice and neurorehabilitation, there must be robust investigation of the psychometric properties of the technique; in particular, the test-retest reliability of TMS outcome measures [12,13].

A systematic review of the reliability of TMS measures in healthy adults and other work [13,14] has highlighted methodological issues in the interpretation of TMS reliability, including

CONTACT Niamh C. Kennedy  n.kennedy@ulster.ac.uk  School of Psychology, Ulster University, Coleraine, UK

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/09638288.2024.2337107>.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

the lack of consideration for the 95% confidence intervals (95% CI) for intraclass correlation coefficient (ICC) values when interpreting the findings, small sample sizes and lack of consideration of measurement error.

The majority of test-retest studies have concentrated on healthy participants [15,16], there has been less investigation of test-retest reliability of TMS measures in chronic stroke individuals [13,17–20]. This lack of evidence is even more pronounced for subacute stroke [13,21].

When TMS is used with neuro-navigation, some TMS measures (such as motor threshold, MEP latency and amplitude) in some distal upper limb muscles have been found to possess good to excellent test-retest reliability in neurologically intact adults [13,16], people within six months of stroke; and people greater than six months after stroke [13]. Consequently, neuro-navigated TMS may be a valuable tool to advance the science of precision stroke rehabilitation. However currently fully equipped TMS laboratories including full neuronavigation systems, are practically exclusively found in Universities or research centres, clinical access to this neuronavigation may be limited. This has a direct impact on this tool's ability to be widely used, if all the equipment including advanced neuronavigation is required for every TMS session to check appropriate stimulation parameters. Consequently, if TMS is used widely in clinical neurorehabilitation, TMS used without neuro-navigation also needs to demonstrate good to excellent test-retest reliability.

The results of the test-retest reliability using non-navigated TMS in stroke survivors six months after stroke, are variable and range from poor to good in the upper and lower limbs. The MEP amplitude demonstrated the most variable findings, ICC values ranging from ICC = 0.205 in the vastus lateralis [22] to an ICC = 0.98 (lower level of the 90% confidence interval 0.94) in the (first dorsal interosseous) FDI muscle [18]. However, participants in these studies were not clinically representative, particularly because people with severe paresis were excluded. In addition, these findings may not be transferable to the sub-acute phase after stroke. This is important because most people receive the majority of rehabilitation during the first few weeks after stroke when neural plasticity is more rapid and therefore considered the optimal time for targeted interventions [23,24]. During this time, it is essential to know whether a change in any TMS-derived measure is attributable to neural factors or measurement 'error' [25]. In sub-acute stroke participants it has been demonstrated for a range of TMS measures including MEPs, total conduction time in APB muscle [21] and FDI [13] that there was more variability resulting in lower ICC, higher measurement error than in healthy groups and chronic stroke groups. There is also a lack of examination of the test-retest reliability of TMS measures in a range of muscles, with the majority of studies targeting the hand muscles [13,18,20,21]. This missing information is crucial, given that TMS values can vary between M1 representations of different muscles along a proximal-distal gradient [26] and do not target the clinically key forearm muscles.

To address this, in the current study the muscles of investigation were the biceps brachii, extensor carpi radialis, and abductor pollicis brevis of the paretic and less paretic limbs. These muscles were selected because they are essential for the completion of activities of daily living such as dressing and grooming and are commonly affected following stroke. To ensure that the most commonly used TMS measures were included in this study a range of measures were assessed including motor threshold (MT), amplitude, latency (MEP-L), silent period (SP), and recruitment curve slope.

This study aims to determine the test-retest reliability of MEP measures, elicited using non-neuro-navigated TMS in people between two and 60d of stroke onset.

Materials and methods

Design and ethics

A prospective correlational study embedded within a randomised clinical trial (RCT) [27].

Ethical approval was obtained for this study and the study was conducted in accordance with the Declaration of Helsinki. Participant testing was conducted at the University Exercise Laboratory and a UK University Hospital Stroke Rehabilitation Centre, participants were tested at the same location on both occasions. Identical hardware, software, equipment, and consumables were used in both locations.

Test-retest reliability was assessed over two identical TMS sessions separated by one to three days. The short inter-session time was due to the neural recovery and cortical reorganisation being at its most rapid in the first three months since ictus [23] and similar to a previous sub-acute reliability study [13].

Participants

Thirty adults in the sub-acute phase (early) after stroke were recruited from an ongoing RCT but had not started the intervention period and met the inclusion criteria: aged 18+ years, after anterior circulation ischemic or hemorrhagic stroke within the previous 2–60d; scoring >11/33 on Motricity Index pinch subsection with more paretic upper limb; unable to complete Nine-Hole Peg Test in 50s or less with paretic upper limb (to determine participants had enough residual upper limb motor control/movement to complete the RCT intervention); no obvious spatial neglect and no contraindications to TMS [27]. As the study was investigating the test-retest reliability of TMS measures, only participants who exhibited an MEP (in at least one of the target muscles in either hemisphere) at the baseline time point of the trial were invited to participate in this study. A full medical history, details of stroke, and clinical scores were obtained for all stroke participants (Table 1).

Table 1. Participant characteristics.

Age	73.3 ± 10.6years (range: 46–91)
Gender	Male <i>n</i> = 16 Female <i>n</i> = 14
Stroke type ^a	Ischemic = 26 Haemorrhagic = 4
Stroke side ^b	Right <i>n</i> = 23 Left <i>n</i> = 7
Side of brain lesion ^c	Right <i>n</i> = 9 Left <i>n</i> = 21
Time since stroke (days) ^d	36 ± 16 (range: 11–70)
ARAT total score Baseline ^e	29 ± 17.4 (range: 0–53)
WMF total score baseline ^f	42 ± 16.28 (range: 4–66)
Grip force	7.08 ± 6.35 kg (range 0–20)
Pinch Force	3.18 ± 2.73 kg (range 0–7.5)
Medication	<i>N</i> = 29 including medication for Hypertension (<i>n</i> = 8), Cholesterol (<i>n</i> = 9), Blood thinner (<i>n</i> = 14), Atrial Fibrillation (<i>n</i> = 5), Diabetes (<i>n</i> = 4), Osteoarthritis (<i>n</i> = 4); SSRI (<i>n</i> = 2), Thyroid hormone treatment (<i>n</i> = 1)

^aType of stroke as determined by MRI.

^bStroke side refers to more paretic side of the body, more stroke affected side.

^cSide of brain lesion as determined by MRI.

^dTime since stroke in days at randomisation.

^eThe Action Research Arm Test (ARAT) total score at Baseline (max score 57-Scores of less than 10 points, between 10 and 56 points, and 57 points correlate with poor, moderate, and good recovery respectively).

^fThe Wolf Motor Function Test (WMF) total score baseline (max score 75 Lower scores are indicative of lower functioning levels).

Procedure

Procedures were identical in sessions one and two. The experimenter had extensive experience in the use of TMS. The experimenter was the same for both sessions. Participants were seated comfortably in an office chair with their forearms relaxed on pillow placed on their lap. Arms were consistently positioned across sessions.

Surface EMG electrodes were placed in parallel along the muscle fibers of the bilateral biceps brachii (BB) and extensor carpi radialis longus (ECR) (ConMed Cleartrace ECG surface electrodes 20 mm), cup electrodes (Nicolette) placed over the abductor pollicis brevis (APB), ground electrode at the olecranon process. The EMG signal was sampled at 1000 Hz, amplified 1000x, band-pass filtered at 15–450 Hz, and saved for offline analysis. This process was standardised across all sessions.

Single pulse monophasic TMS was delivered using a Magstim 200² (Magstim Company Ltd) stimulator with a figure-of-8 coil (90 mm in diameter) [28]. The coil was placed so that the axis of intersection between the two loops was orientated at approximately 45° to the sagittal plane, to induce a posterior to anterior current flow across the motor strip in the primary motor cortex (M1). Once the hotspot for each muscle was established, the coil position was marked on the head with a permanent pen and used throughout the experiment. The hot spot was determined for each muscle individually at each session [29]. Resting motor threshold (rMT), was established as the lowest stimulation intensity at which MEPs with peak-to-peak amplitude of approximately 50 µV were evoked in at least 5 out of 10 MEPs, at that hotspot location. Active Motor threshold (AMT) was established when half of the successive TMS pulses (in at least 5 out of 10 trials while the participants maintained a constant background contraction) produced an MEP > 200 µV [30]. There is evidence that after stroke there is variability in force production when maintaining a specific percentage of the maximal voluntary contraction in both upper and lower limbs [31,32]. Furthermore, there is greater variation with lower percentages of the MVC, which within this experiment would be between 10–20% MVC [32]. This evidence along with the participant burden and fatigue of completing multiple upper limb assessments prior to TMS data collection participants were asked to maintain a slight muscle contraction during TMS data collection with the muscle of interest during active trials and this was monitored by palpation, vision, and visual EMG by the Experimenter (a qualified physiotherapist).

An active recruitment curve was obtained from 100% AMT to 130% AMT increasing in 10% increments (bilateral limbs). Five TMS pulses were delivered at each intensity. This process was repeated for each muscle (BB, ECR, APB) in both upper limbs. When collecting active recruitment curves the non-paretic BB hotspot (which corresponds to stimulation of non-stroke lesioned hemisphere) was always probed first, followed by non-paretic ECR, non-paretic APB, paretic (stimulation of lesioned hemisphere) BB, paretic ECR and paretic APB. The order of the TMS measures was purposefully fixed so that any order effect, if present, would have a consistent influence on measurement variability. All participants were asked to attend to their upper limb during data collection, instructed to relax all other muscles and the level of arousal and fatigue was monitored by an experimental team [33].

Data analysis

Active Recruitment curve was collected for each muscle in both hemispheres. From these the following measurements were

extracted MEP amplitude, silent period (measured at 130% active motor threshold), MEP latency (at 120% active motor threshold), and recruitment curve slope.

Following completion of data collection, all MEPs were visually assessed (to determine if there was any electrical noise or a lack of an MEP and trials without an MEP or with electrical noise were not analysed). Then using a custom-made script, values were extracted for: peak-to-peak MEP amplitude; visual analysis of the MEP was used to determine MEP latency [18], and the Cortical Silent Period [34]. The cortical silent period was assessed at 130% AMT as there is evidence that the cortical silent period has the smallest coefficient of variation at this threshold (vs 110% and 120% AMT [30]). The raw MEP amplitudes were fitted using both a Boltzmann sigmoid function [29,35] and a linear function [36] to estimate the slope of the recruitment curve (RC). Data were fitted to both sigmoidal and linear function as there has been issues fitting data, especially from more proximal upper limb muscles [35,37].

The data were tested for normality and found to exhibit a normal distribution. The test-retest reliability of TMS measures was assessed using the intra-class correlation coefficient (ICC) model 2,1 with associated 95% confidence intervals [95%CI] and the limits of agreement (LOA). The ICC model was used to determine agreement between the two sessions [12] and interpreted with the associated 95% CI such that an ICC > 0.70 is acceptable reliability, determined by the lower 95% CI. The LOA provided further estimates of measurement agreement and evaluated if there was a biased pattern of error (systematic or random).

The SDC provides a value for the minimum change that needs to be observed, in order to be confident that the observed change is real and not, potentially, a product of measurement error in the instrument [38]. The smallest detectable change (SDC) was calculated using the formula: $1.96 \times \sqrt{2} \times SEM$ [21].

Results

Sixty-seven stroke survivor participants in one of the three trial centres were screened for inclusion. Of these, 43 people (63%) were suitable to have TMS, 36 (84%) of those provided informed consent of whom 30 (83%) were included in data analysis. The 6 people who were consented but not included in the analysis had only completed one of the two required sessions. Participant characteristics are detailed in Table 1. Thirty participants were included with a mean age of 73.3 (range 46–91) years and a mean of 36 (SD 16 d, range 11–70) days after stroke. Participants were included irrespective of mild, moderate or severe upper limb dysfunction as demonstrated by a mean Action Research Arm Test score of 29 of a possible total of 57 (range 0–53) and a mean Wolf Motor Function Test score of 42 of a maximum possible total of 75 (range 4–66) [27]. These scores demonstrate that testing was conducted on a clinically representative sample.

No adverse TMS events occurred. There were 10% of trials across the participants not included in the analysis due to either a MEP not being present, or electrical noise inhibiting detection and analysis of the MEP.

Test-retest reliability of MEP characteristics

The test-retest reliability of TMS measures (including AMT, silent period, MEP latency and MEP amplitude) in individuals with stroke was variable within and between MEP characteristics and muscles (Tables 2–4).

Table 2. Test-retest reliability and smallest detectable difference (SDC) of: motor threshold (% stimulator output), motor evoked potential latency (MEP-L, ms) and silent period (ms).

	Non-Paretic UL	Paretic UL	Non-Paretic UL		Paretic UL	
	ICC [95% CI]	ICC [95% CI]	Mean difference (95% LOA)	SDC	Mean difference (95% LOA)	SDC
RMT						
BB	0.47 [0.14,0.71] ^a	0.67 [0.38,0.84] ^b	-0.40 (-19.94; 19.87)	22.20	-3.67 (-27.27;19.93)	25.76
ECR	0.81 [0.64,0.90] ^a	0.73 [0.49,0.80] ^c	-0.29 (-14.49; 13.92)	12.68	-6.04 (-31.77;19.69)	18.72
APB	0.66 [0.39,0.83] ^d	0.77 [0.54,0.89] ^e	-1.60 (-11.81;8.61)	21.01	-5.17 (-25.37;15.04)	12.63
AMT						
BB	0.65 [0.38,0.82] ^a	0.69 [0.45,0.84] ^d	-0.20 (-13.87;12.76)	12.30	-2.18 (-18.90;14.55)	14.66
ECR	0.76 [0.56,0.88] ^a	0.84 [0.68,0.93] ^d	-0.29 (-8.37;7.04)	7.64	0.09 (-16.34;10.88)	13.85
APB	0.64 [0.37,0.80] ^d	0.81 [0.62,0.91] ^f	-0.26 (-12.33;9.59)	10.81	-1.96 (-19.86;15.95)	17.52
MEP-L						
BB	0.71 [0.46,0.85] ^g	0.65 [0.32,0.84] ^g	0.20 (-3.18;3.59)	3.31	0.27(-3.84;4.38)	14.51
ECR	0.35 [0.00,0.64] ^g	0.30 [0.00,0.64] ^g	-0.34 (-4.13;3.45)	7.15	0.23 (-9.79;10.24)	8.72
APB	0.67 [0.40,0.83] ^c	0.77 [0.49,0.91] ^b	0.55 (-2.99;4.09)	3.66	-0.05 (-2.67;2.56)	2.63
SP						
BB	0.53 [0.20,0.75] ^f	0.45 [0.04,0.72] ^b	-7.90 (-88.15;72.34)	85.50	-3.43 (-74.35;67.49)	72.79
ECR	0.19 [0.00,0.51] ^f	0.75 [0.50,0.89] ^b	-13.89 (-107.43;79.65)	97.22	-10.66 (-68.06;46.75)	56.53
APB	0.66 [0.38,0.83] ^g	0.84 [0.62,0.94] ⁱ	-11.67 (-78.38;55.65)	71.86	-0.68 (-48.72;47.35)	51.56

^an=30; ^bn=24; ^cn=27; ^dn=29; ^en=25; ^fn=26; ^gn=28; ^hn=22; ⁱn=18.

ICC: intraclass correlation coefficient; CI: confidence interval; LOA: limits of agreement; UL: upper limb; RMT: resting motor threshold; BB: biceps brachii; ECR: extensor carpi radialis; APB: abductor pollicis brevis; AMT: active motor threshold; MEP-L: motor evoked potential latency.

Table 3. Test-re-test reliability and smallest detectable difference (SDC) of motor evoked potential (MEP) amplitude (μ v) at different stimulation intensities.

	Non-Paretic UL	Paretic UL	Non-Paretic UL		Paretic UL	
	ICC [95% CI]	ICC [95% CI]	Mean difference (95% LOA)	SDC	Mean difference (95% LOA)	SDC
BB						
100% AMT	0.46 [0.14, 0.70] ^a	0.35 [0.00, 0.63] ^b	0.17 (-0.60;0.94)	0.77	0.03 (-0.71;0.77)	2.94
110% AMT	0.57 [0.28, 0.77] ^a	0.24 [0.00, 0.55] ^b	0.20 (-0.91;1.30)	0.89	0.30 (-2.03;2.63)	1.95
120% AMT	0.55 [0.25, 0.76] ^a	0.23 [0.00, 0.54] ^b	0.17 (-1.50;1.85)	1.19	0.26 (-1.63;2.16)	1.70
130% AMT	0.51 [0.20, 0.73] ^a	0.11 [0.00, 0.47] ^c	0.24 (-1.85;2.33)	1.57	0.55 (-2.73;3.82)	2.75
ECR						
100% AMT	0.46 [0.14, 0.70] ^a	0.39 [0.02, 0.67] ^c	-0.32 (-2.79;2.15)	2.42	-0.07 (-1.19;1.05)	1.23
110% AMT	0.43 [0.10, 0.67] ^a	0.34 [0.00, 0.63] ^d	-0.57 (-3.25;2.11)	2.62	-0.18 (-1.87;1.50)	1.49
120% AMT	0.35 [0.02, 0.62] ^a	0.31 [0.00, 0.62] ^c	-0.59 (-3.65;2.47)	2.99	-0.16 (-1.82;1.49)	1.54
130% AMT	0.41 [0.08, 0.66] ^a	0.33 [0.00, 0.65] ^e	-0.54 (-3.37;2.28)	2.72	-0.12 (-1.80;1.56)	1.59
APB						
100% AMT	0.77 [0.56, 0.88] ^b	0.34 [0.00, 0.65] ^f	0.05 (-1.71;1.61)	1.55	-0.21 (-3.42;2.99)	2.90
110% AMT	0.69 [0.44, 0.84] ^b	0.74 [0.48, 0.87] ^f	-0.29 (-2.47;1.89)	1.96	-0.04 (-2.67;2.53)	1.94
120% AMT	0.67 [0.40,0.83] ^b	0.88 [0.75, 0.95] ^g	-0.04 (-2.85;2.76)	2.46	-0.06 (-2.08;1.96)	1.88
130% AMT	0.66 [0.38, 0.82] ^b	0.77 [0.49, 0.90] ^h	-0.06 (-3.06;2.94)	2.65	0.08 (-2.99;3.15)	3.00

^an=30; ^bn=29; ^cn=25; ^dn=27; ^en=24; ^fn=26; ^gn=23; ^hn=21.

ICC: intraclass correlation coefficient; CI: confidence interval; LOA: limits of agreement; UL: upper limb; BB: biceps brachii; AMT: active motor threshold; ECR: extensor carpi radialis; APB: abductor pollicis brevis.

For RMT, AMT, SP and MEP latency the majority of estimated ICC values were good or moderate, the exceptions were non-paretic RMT BB, ECR (both limbs) latency and ECR silent period which had poor ICC values (Table 2). MEP amplitude test-retest reliability in RC stimulation intensities across all muscles had estimated ICC values of poor or moderate with the exceptions of good ICC values in APB 100% AMT (both paretic and non-paretic limbs) and APB 120/130% AMT paretic limb (Table 3). The RC slope generally had poor test-retest reliability across all muscles in both arms regardless of whether the fit was sigmoidal or linear (Table 4). However, across all MEP measures (AMT, RMT, latency, SP, MEP amplitude, RC slope) the ICC confidence intervals were wide, with the lower end falling below acceptable test-retest reliability (Tables 2–4).

The LOA analysis demonstrates small mean differences between session 1 and session 2, suggesting small group differences between the two sessions. However, the upper and lower 95% LOA were wide, suggesting variation in individuals in turn indicating measurement error (Tables 2–4). For example, the mean difference (LOA) for: MEP latency was 0.23 (-9.79, 10.24) ms and AMT ECR was 0.09 (-16.34, 10.88) percentage of stimulator output

(Table 2). There were differences in the mean difference and LOA between paretic and non-paretic limbs, generally the non-paretic limb demonstrated smaller mean differences and narrower 95% LOA (exceptions SP ECR and APB, and latency APB). For example, the motor threshold in non-paretic ECR mean difference and 95% LOA were -0.29 (-14.49; 13.92) whereas the paretic ECR demonstrated -6.04 (-31.77; 19.69) measured as a percentage of stimulator output.

Smallest detectable change (SDC) in MEP characteristics

Example group SDC's for motor threshold ranged from 7.64% of stimulator output for AMT of non-paretic ECR to 25.76% for RMT of paretic BB, (Supplementary Table 1 and 2) indicating that a change of greater than 7.64 or 25.76% of stimulator output would indicate a change in measurement above that of overall measurement error in the respective muscles. Generally, the non-paretic limb demonstrated smaller SDCs than the paretic limb in 13 of the MEP characteristics investigated. The few exceptions to this were the SP for all muscles and the RMT and MEP-L in APB.

Table 4. Test-retest reliability of the recruitment curve slope (RCS) for all participants.

Non-paretic UL active	Paretic UL active			
RCS-sigmoidal				
ICC (95% CI)				
BB	0.21	[0.00, 0.77]	0.16	[0.00, 0.78]
ECR	0.22	[0.00, 0.63]	0.78	[0.00, 0.98]
APB	0.12	[0.00, 0.68]	0.00	[0.00, 0.91]
M-diff (95% LOA)				
BB	0.13	(0.00,0.65)	0.00	[0.00,0.23]
ECR	0.00	(0.00,0.35)	0.21	(0.00,0.54)
APB	0.26	(0.00,0.57)	0.31	(0.00,0.60)
RCS-linear				
ICC (95% CI)				
BB	0.48	[0.00; 0.76]	0.03	[0.14, 0.20]
ECR	0.01	[0.22, 0.20]	0.02	[0.04, 0.08]
APB	0.04	[0.34, 0.25]	0.04	[0.55, 0.48]
M-diff (95% CICI LOA)				
BB	0.01	(-0.07;0.08)	0.02	(-0.07;0.10)
ECR	-0.01	(-0.08;0.07)	0.00	(-0.04;0.04)
APB	0.01	(-0.08;0.11)	0.02	(-0.07;0.12)

UL: upper limb; ICC: intra-class correlation coefficient; 95% CI = 95% confidence intervals; BB: biceps brachii; ECR: extensor carpi radialis; APB: Abductor Pollicis Longus; M-diff: mean difference; LOA: limits of agreement.

Discussion

This present study characterised the test-retest reliability of frequently used TMS measures in sub-acute stroke participants in a range of upper limb muscles across two separate sessions. To the best of our knowledge, this study is the first to investigate a variety of upper limb muscles across the distal-proximal gradient, including the clinically meaningful forearm (ECR) and Biceps muscles, in a sub-acute stroke population. Using the different measures of reliability and agreement, the authors conclude that TMS has low reliability between sessions at an individual level early after stroke.

If TMS is to be used in clinical practice and research to ascertain whether interventions promote recovery or compensation, then the SDC must be considered. The smallest detectable change (SDCs) of eight MEP characteristics in three upper limb muscles bilaterally (48 values) is included. The SDC results for a range of MEP characteristics demonstrate high values, with the paretic limb generally having higher values than the non-paretic limb and variation across the three target muscles. The smallest values were for MEP latency and the largest for the silent period. The sub-acute stroke participants had substantially higher values, for example, in AMT, 10.8% compared to 4% for healthy participants [17]. This demonstrates the impact that post-stroke neurological changes can have on TMS measures. It, therefore, warrants more work to identify these stroke-related changes to ensure measurement error is not larger than any potential observed change. This is especially pertinent if TMS measures are to be used as a biomarker for recovery [11].

Limits of agreement (LOA) is used to determine measurement error which is otherwise known as agreement [13]. If there is agreement between measures, the measurement error is small, and the more reliable and less variable the measure is. The LOA were wide across all TMS measures and muscles, indicating poor absolute reliability, high levels of measurement error or a lack of agreement between individuals. Measurement error is subject to change depending on muscle, population and TMS measure in focus. Previous work has suggested that more proximal muscles such as Biceps may have higher measurement error because it receives fewer cortico-motor neuronal projections than distal muscles [39].

ICC was used to measure the degree to which stable individuals in a sample maintain their position relative to each other [40]. This study found that most TMS measures were not reliable due to the lower end of 95% CI of the ICC being below 0.7. Despite many of the estimated ICC values being within moderate or good reliability, they must be interpreted in conjunction with 95% CI values. To date, reliability results have been compared across studies including in different muscles and different samples. This has fundamental limitations due to the effect age has on TMS measures [37] and that TMS measures do not respond uniformly across all muscle representations [35]. Comparing ICC values is also statistically problematic, as they are dependent on sample variation [37]. However, much of the previous literature concluded that the ICC values were reliable but did not provide a 95% CI value; they may not have concluded acceptable reliability if adherence to the requirement of a lower 95% CI of at least 0.70 was used.

Previous research has demonstrated that distal muscles have a greater response to brain stimulation compared to proximal muscles [26]; the greater response (of distal muscles) to stimulation may be a contributing factor to their reliability. This may be in part due to the neurophysiological control mechanisms of upper limb muscles. Distal muscles such as FDI may have more reliable responses to TMS due to the higher density of cortico-spinal neurons projecting to the muscle and the larger motor map accessible to TMS in comparison to a proximal muscle such as Biceps [41]. The more proximal muscles also appear to have weaker cortical inhibitory control than distal muscles [26,35]. There has been relatively little research investigating the reliability of TMS measures from more proximal muscles such as BB, with the majority in healthy volunteers [39,41,42]. Even in healthy participants, it has been demonstrated that reliability is lower than in the hand muscles [41,42] and dependent on the TMS measure. The present study has a lower ICC for RMT (0.67) than Sankarasubramanian and colleagues (ICC 0.74)[39]. However, the latter did not provide 95% CI values. Therefore, although the authors concluded the ICC suggests that RMT intensity in BB was reliable, we concluded using the lower end of the 95% CI RMT was not. Kamen and colleagues [43] displayed higher ICC values in RMT than the current study (0.95/0.99) as did [42] (0.882) but these were both in healthy individuals and it is established that reliability tends to be lower in stroke patients regardless of the muscle [13]. Given the small sample of studies with BB as a target muscle for TMS reliability, it is suggested that further investigation is needed [13,39,42].

As outlined in the introduction, very few studies have investigated reliability in a sub-acute stroke population [13,20]. The ICC [95% CIs] of the TMS measures from sub-acute stroke population, in the study reported here were lower than those found for neuro-navigated TMS in sub-acute stroke survivors. Neuro-navigated TMS improves spatial accuracy of coil placement on the hotspot and aims to decrease variation in MEP amplitude [44–46]. This study purposely aimed to investigate the reliability of TMS measures whilst not using neuronavigation. This was to examine the transferability of this tool to a clinical setting, where neuronavigation is less widely used in comparison to neurophysiological lab settings. The authors consider this an important development in translational clinical neuroscience as many (but not all) of the stroke clinical trials involving TMS have used neuronavigation. However, if this technique is to be applied in a clinical setting the test-retest reliability without neuro-navigation needs to be considered.

Studies using neuronavigated TMS have produced some measures that demonstrate high ICC or low measurement error in

healthy and stroke participants. However, some MEP characteristics using neuro-navigated TMS are not reliable, namely: recruitment curve slope, short-interval intracortical inhibition, and intracortical facilitation [13]. So, it is possible that even with the use of neuro-navigation, some MEP characteristics do not have acceptable reliability in subacute stroke [13,47]. If neuro-navigated TMS does produce reliable measures of MEP characteristics in muscles of interest in a clinically relevant population, its feasibility in clinical practice will require further investigation.

Other factors could influence our results, including a heterogeneous sample: this current study included participants with a wide range of functional impairment, including those with very low clinical scores indicating a high level of impairment. Previous work investigating reliability in stroke was conducted with participants with less motor impairment [13]. However, in the current study, the authors included a spectrum of stroke survivors from mild to severe motor impairment to have a sample closer to the clinical population. The target muscle could also be a factor as this study aimed to investigate the clinically functionally relevant muscles, both proximal and distal. The differing muscle response may be explained by the proximal to distal gradient [48] and the functional role of each muscle [49]. There are other potential influences on the reliability of TMS-derived measures of MEP characteristics. Inter-individual variability in response to TMS has been reported due to factors such as gene expression, e.g., brain-derived neurotrophic factor, BDNF [50]. Also, the TMS data collection method may contribute to differences in findings between studies [40]. Standardisation of data collection and processing [33] would facilitate improved comparisons between studies and may contribute to improved measurement reliability.

The key strengths of this study are that the stroke participants were representative of the sub-acute stroke population with no contraindication to TMS early (mean 39 d). MEP characteristics of clinical interest for discerning recovery from compensation were investigated for both proximal and distal upper limb muscles. The test-retest reliability in stroke survivors has been focused on chronic stroke populations (greater than six months after stroke) ranging from 9 to 23 with an average of 19 participants. There are only two studies that focus on sub-acute TMS reliability with 18 and 20 participants. To the best of the author's knowledge, the achieved sample size of 30 is the largest stroke survivor sample in a TMS reliability study. The sample of stroke participants included a wide range of impairment level, including individuals more severely affected, a population that is often less represented in stroke research.

This study addresses statistical and procedural limitations of previous studies. In particular, the lack of use of the lower 95% CI of an ICC value and not including a means of assessing measurement error (for example SDC) [13,17]. In response to the recent systematic review [35], the critical appraisal tool proposed by IFCN [33] was used to check TMS procedures and improve the quality of reporting. This demonstrated that 90% of the factors were reported, except for the following: "History of specific repetitive motor activity," "Prior motor activity of the muscle to be tested," and "time between MEP trials" [33].

There were some limitations to this study. The main one is that five TMS pulses were delivered at each stimulation intensity, which is below the number advised [51,52]. It remains possible that more pulses would have changed the reliability values. It is also possible that increasing the number of stimulation intensities could influence the reliability findings. It has been suggested that increasing the number of TMS trials can reduce SDC and measurement error [53]; future studies should consider this. However, five stimuli have been used previously (as well as recommended in previous guidelines) and

reported to have high reliability [30,54,55]. The use of non-navigated TMS can be seen as a limitation, but this was part of a conscious decision by the research team to reflect the current clinical environment where neuro-navigation is rarely available. Several pragmatic decisions to reduce the burden on stroke participants very early after stroke could be viewed as limitations. These included not conducting an RMT recruitment curve, not having sub-acute stroke participants perform an MVC, and not having stroke participants maintain a specific percentage of MVC contraction during AMT which may have impacted the reliability. While the experimenter provided consistent instructions and used observation to evaluate the muscle contraction during TMS this method could be more objective and standardised in future research and in clinical settings. For example, using a dynamometer for MVC and maintaining a percentage of the MVC as well as using EMG feedback in a similar way to this study but having feedback to the participant to maintain a specific contraction during TMS data collection [56].

Conclusions

In summary, this study found unacceptable test-retest reliability for a variety of non-navigated TMS measures (motor threshold, MEP latency, cortical silent period, recruitment curve slope) in the Biceps, ECR or APB muscles in sub-acute stroke participants. The different conclusions from earlier studies may result from this study addressing many of the limitations of previous studies (including the use of 95% CI). The present findings indicate there is still limited evidence for acceptable reliability of TMS outcomes in sub-acute stroke. The next steps could be to undertake a direct comparison between the reliability of neuro-navigated and non-neuro-navigated TMS focused on the same muscles in the sub-acute stroke population. The results of a definitive comparative study would provide the knowledge required to inform the future use of biomarkers to advance knowledge of the mechanisms of recovery and thereby the science of precision stroke rehabilitation.

Professor Pomeroy is grateful for the support of the National Institute for Health Research (NIHR) Brain Injury MedTech Co-operative based at Cambridge University Hospitals NHS Foundation Trust and University of Cambridge. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Acknowledgements

We acknowledge the contribution of the FAST INdiCATE Investigators for contributions to the conduction of the FAST INdiCATE Trial namely: Susan M Hunter, Heidi Johansen-Berg, Nick Ward, Elizabeth Chandler, Christopher John Weir, John Rothwell, Alan M Wing, Michael J Grey, Garry Barton, Nick Malachy Levy, Claire Havis, Roger N Lemon, Jane Burridge, and Amy Drummond. In addition, we are grateful to all participants in the study reported here who gave their time for research.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

References

- [1] Barker AT, Jalinous R, Freeston IL. Non-invasive magnetic stimulation of human motor cortex. *Lancet*. 1985;1(8437):1106–1107. doi:10.1016/s0140-6736(85)92413-4.
- [2] Rossini PM, Barker AT, Berardelli A, et al. Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: basic principles and procedures for routine clinical application. Report of an IFCN committee. *Electroencephalogr Clin Neurophysiol*. 1994;91(2):79–92. doi:10.1016/0013-4694(94)90029-9.
- [3] Rothwell JC. Techniques and mechanisms of action of transcranial stimulation of the human motor cortex. *J Neurosci Methods*. 1997;74(2):113–122. doi:10.1016/s0165-0270(97)02242-5.
- [4] Hallett M. Transcranial magnetic stimulation: a primer. *Neuron*. 2007;55(2):187–199. doi:10.1016/j.neuron.2007.06.026.
- [5] Reis J, Swayne OB, Vandermeeren Y, et al. Contribution of transcranial magnetic stimulation to the understanding of cortical mechanisms involved in motor control. *J Physiol*. 2008;586(2):325–351. doi:10.1113/jphysiol.2007.144824.
- [6] Talelli P, Greenwood RJ, Rothwell JC. Arm function after stroke: neurophysiological correlates and recovery mechanisms assessed by transcranial magnetic stimulation. *Clin Neurophysiol*. 2006;117(8):1641–1659. doi:10.1016/j.clinph.2006.01.016.
- [7] Stinear CM, Barber PA, Petoe M, et al. The PREP algorithm predicts potential for upper limb recovery after stroke. *Brain*. 2012;135(Pt 8):2527–2535. doi:10.1093/brain/aww146.
- [8] Marconi B, Filippi GM, Koch G, et al. Long-term effects on cortical excitability and motor recovery induced by repeated muscle vibration in chronic stroke patients. *Neurorehabil Neural Repair*. 2011;25(1):48–60. doi:10.1177/1545968310376757.
- [9] Khedr EM, Shawky OA, El-Hammady DH, et al. Effect of anodal versus cathodal transcranial direct current stimulation on stroke rehabilitation: a pilot randomized controlled trial. *Neurorehabil Neural Repair*. 2013;27(7):592–601. doi:10.1177/1545968313484808.
- [10] Pomeroy VM, Tallis RC. Need to focus research in stroke rehabilitation. *Lancet*. 2000;355(9206):836–837. doi:10.1016/s0140-6736(99)08143-x.
- [11] Boyd LA, Hayward KS, Ward NS, et al. Biomarkers of stroke recovery: consensus-based core recommendations from the stroke recovery and rehabilitation roundtable. *Int J Stroke*. 2017;12(5):480–493. doi:10.1177/1747493017714176.
- [12] Portney LG, Watkins MP. *Foundations of clinical research: applications to practice*. Upper Saddle River (NJ): Pearson/Prentice Hall; 2009.
- [13] Schambra HM, Ogden RT, Martínez-Hernández I, et al. The reliability of repeated TMS measures in older adults and in patients with subacute and chronic stroke. *Front Cell Neurosci*. 2015;9:335. doi:10.3389/fncel.2015.00335.
- [14] Moksink 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. doi:10.1016/j.jclinepi.2010.02.006.
- [15] Cacchio A, Cimini N, Alosi P, et al. Reliability of transcranial magnetic stimulation-related measurements of tibialis anterior muscle in healthy subjects. *Clin Neurophysiol*. 2009;120(2):414–419. doi:10.1016/j.clinph.2008.11.019.
- [16] Tedesco Triccas L, Hughes AM, Burridge JH, et al. Measurement of motor-evoked potential resting threshold and amplitude of proximal and distal arm muscles in healthy adults. A reliability study. *J Rehabil Assist Technol Eng*. 2018;5:2055668318765406. Mar doi:10.1177/2055668318765406.
- [17] Beaulieu LD, Massé-Alarie H, Ribot-Ciscar E, et al. Reliability of lower limb transcranial magnetic stimulation outcomes in the ipsi- and contralesional hemispheres of adults with chronic stroke. *Clin Neurophysiol*. 2017;128(7):1290–1298. doi:10.1016/j.clinph.2017.04.021.
- [18] Koski L, Lin JC, Wu AD, et al. Reliability of intracortical and corticomotor excitability estimates obtained from the upper extremities in chronic stroke. *Neurosci Res*. 2007;58(1):19–31. doi:10.1016/j.neures.2007.01.007.
- [19] Wheaton LA, Villagra F, Hanley DF, et al. Reliability of TMS motor evoked potentials in quadriceps of subjects with chronic hemiparesis after stroke. *J Neurol Sci*. 2009;276(1-2):115–117. doi:10.1016/j.jns.2008.09.012.
- [20] Liu H, Au-Yeung SS. Reliability of transcranial magnetic stimulation induced corticomotor excitability measurements for a hand muscle in healthy and chronic stroke subjects. *J Neurol Sci*. 2014;341(1-2):105–109. doi:10.1016/j.jns.2014.04.012.
- [21] Hoonhorst MH, Kollen BJ, Van Den Berg PS, et al. How reproducible are transcranial magnetic stimulation-induced MEPs in subacute stroke? *J Clin Neurophysiol*. 2014;31(6):556–562. doi:10.1097/WNP.0000000000000114.
- [22] Cacchio A, Paoloni M, Cimini N, et al. Reliability of TMS-related measures of tibialis anterior muscle in patients with chronic stroke and healthy subjects. *J Neurol Sci*. 2011;303(1-2):90–94. doi:10.1016/j.jns.2011.01.004.
- [23] Cramer SC. Repairing the human brain after stroke: I. Mechanisms of spontaneous recovery. *Ann Neurol*. 2008;63(3):272–287. doi:10.1002/ana.21393.
- [24] Wahl A, Schwab ME. Finding an optimal rehabilitation paradigm after stroke : enhancing fiber growth and training of the brain at the right moment. *Front Hum Neurosci*. 2014;8:381. doi:10.3389/fnhum.2014.00381.
- [25] de Vet HC, Terwee CB, Knol DL, et al. When to use agreement versus reliability measures. *J Clin Epidemiol*. 2006;59(10):1033–1039. doi:10.1016/j.jclinepi.2005.10.015.
- [26] Martin PG, Gandevia SC, Taylor JL. Theta burst stimulation does not reliably depress all regions of the human motor cortex. *Clin Neurophysiol*. 2006;117(12):2684–2690. doi:10.1016/j.clinph.2006.08.008.
- [27] Hunter SM, Johansen-Berg H, Ward N, et al. Functional strength training and movement performance therapy for upper limb recovery early poststroke—efficacy, neural correlates, predictive markers, and cost-effectiveness: FAST-INdiCATE trial. *Front Neurol*. 2017;8:733. doi:10.3389/fneur.2017.00733.
- [28] Wassermann EM. Variation in the response to transcranial magnetic brain stimulation in the general population. *Clin Neurophysiol*. 2002;113(7):1165–1171. doi:10.1016/s1388-2457(02)00144-x.
- [29] Carroll TJ, Riek S, Carson RG. Reliability of the input-output properties of the cortico-spinal pathway obtained from transcranial magnetic and electrical stimulation. *J Neurosci Methods*. 2001;112(2):193–202. doi:10.1016/s0165-0270(01)00468-x.
- [30] Groppa S, Oliviero A, Eisen A, et al. A practical guide to diagnostic transcranial magnetic stimulation: report of an IFCN committee. *Clin Neurophysiol*. 2012;123(5):858–882. doi:10.1016/j.clinph.2012.01.010.
- [31] Lindberg PG, Roche N, Robertson J, et al. Affected and unaffected quantitative aspects of grip force control in hemiparetic patients after stroke. *Brain Res*. 2012;1452:96–107. doi:10.1016/j.brainres.2012.03.007.

- [32] Chow JW, Stokic DS. Force control of quadriceps muscle is bilaterally impaired in subacute stroke. *J Appl Physiol.* 2011;111(5):1290–1295. doi:10.1152/jappphysiol.00462.2011.
- [33] Chipchase L, Schabrun S, Cohen L, et al. A checklist for assessing the methodological quality of studies using transcranial magnetic stimulation to study the motor system: an international consensus study. *Clin Neurophysiol.* 2012;123(9):1698–1704. doi:10.1016/j.clinph.2012.05.003.
- [34] Damron LA, Dearth DJ, Hoffman RL, et al. Quantification of the corticospinal silent period evoked via transcranial magnetic stimulation. *J Neurosci Methods.* 2008;173(1):121–128. doi:10.1016/j.jneumeth.2008.06.001.
- [35] Carson RG, Nelson BD, Buick AR, et al. Characterizing changes in the excitability of corticospinal projections to proximal muscles of the upper limb. *Brain Stimul.* 2013;6(5):760–768. doi:10.1016/j.brs.2013.01.016.
- [36] Massie CL, Malcolm MP. Considerations for stimulus–response curves in stroke: an investigation comparing collection and analysis methods. *Int J Neurosci.* 2013;123(3):175–183. doi:10.3109/00207454.2012.738734.
- [37] Houde F, Laroche S, Thivierge V, et al. Transcranial magnetic stimulation measures in the elderly: reliability, smallest detectable change and the potential influence of lifestyle habits. *Front Aging Neurosci.* 2018;10:379.
- [38] Geerinck A, Alekna V, Beaudart C, et al. Standard error of measurement and smallest detectable change of the sarcopenia quality of life (SarQoL) questionnaire: an analysis of subjects from 9 validation studies. *PLoS One.* 2019;14(4):e0216065. doi:10.1371/journal.pone.0216065.
- [39] Sankarasubramanian V, Roelle SM, Bonnett CE, et al. Reproducibility of transcranial magnetic stimulation metrics in the study of proximal upper limb muscles. *J Electromyogr Kinesiol.* 2015;25(5):754–764. doi:10.1016/j.jelekin.2015.05.006.
- [40] Beaulieu LD, Flamand VH, Massé-Alarie H, et al. Reliability and minimal detectable change of transcranial magnetic stimulation outcomes in healthy adults: a systematic review. *Brain Stimul.* 2017;10(2):196–213. doi:10.1016/j.brs.2016.12.008.
- [41] Brasil-Neto JP, McShane LM, Fuhr P, et al. Topographic mapping of the human motor cortex with magnetic stimulation: factors affecting accuracy and reproducibility. *Electroencephalogr Clin Neurophysiol.* 1992;85(1):9–16. doi:10.1016/0168-5597(92)90095-s.
- [42] Corp DT, Rogers MA, Di Lazzaro V, et al. Intrasession reliability of single and paired pulse TMS evoked from the biceps brachii representation of the human motor cortex. *Brain Stimul.* 2015;8(3):660–661. doi:10.1016/j.brs.2015.01.402.
- [43] Kamen G. Reliability of motor-evoked potentials during resting and active contraction conditions. *Med Sci Sports Exerc.* 2004;36(9):1574–1579. doi:10.1249/01.mss.0000139804.02576.6a.
- [44] Julkunen P, Säisänen L, Danner N, et al. Comparison of navigated and non-navigated transcranial magnetic stimulation for motor cortex mapping, motor threshold and motor evoked potentials. *Neuroimage.* 2009;44(3):790–795. doi:10.1016/j.neuroimage.2008.09.040.
- [45] Sollmann N, Hauck T, Obermüller T, et al. Inter- and intraobserver variability in motor mapping of the hotspot for the abductor pollicis brevis muscle. *BMC Neurosci.* 2013;14(1):94. doi:10.1186/1471-2202-14-94.
- [46] Forster MT, Limbart M, Seifert V, et al. Test-retest reliability of navigated transcranial magnetic stimulation of the motor cortex. *Oper. Neurosurg.* 2014;10(Suppl 1):51–56. doi:10.1227/NEU.0000000000000075.
- [47] Ngomo S, Leonard G, Moffet H, et al. Comparison of transcranial magnetic stimulation measures obtained at rest and under active conditions and their reliability. *J Neurosci Methods.* 2012;205(1):65–71. doi:10.1016/j.jneumeth.2011.12.012.
- [48] Chen R. Studies of human motor physiology with transcranial magnetic stimulation. *Muscle Nerve.* 2000;23(S9):S26–S32. doi:10.1002/1097-4598(2000)999:9<::AID-MUS6>3.0.CO;2-I.
- [49] Waller SM, Yang CL, Magder L, et al. Impaired motor preparation and execution during standing reach in people with chronic stroke. *Neurosci Lett.* 2016;630:38–44. doi:10.1016/j.neulet.2016.07.010.
- [50] Kleim JA, Chan S, Pringle E, et al. BDNF val66met polymorphism is associated with modified experience-dependent plasticity in human motor cortex. *Nat Neurosci.* 2006;9(6):735–737. doi:10.1038/nn1699.
- [51] Rossini PM, Burke D, Chen R, et al. Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: basic principles and procedures for routine clinical and research application. An updated report from an IFCN committee. *Clin Neurophysiol.* 2015;126(6):1071–1107. doi:10.1016/j.clinph.2015.02.001.
- [52] Biabani M, Farrell M, Zoghi M, et al. The minimal number of TMS trials required for the reliable assessment of corticospinal excitability, short interval intracortical inhibition, and intracortical facilitation. *Neurosci Lett.* 2018;674:94–100. doi:10.1016/j.neulet.2018.03.026.
- [53] Bastani A, Jaberzadeh S. Correction: a higher number of TMS-elicited MEP from a combined hotspot improves intra- and inter-session reliability of the upper limb muscles in healthy individuals. *PLOS One.* 2013;8(9):33. doi:10.1371/annotation/ca8c5e69-bdd5-4ac2-9c42-993792c37b33.
- [54] Christie A, Fling B, Crews RT, et al. Reliability of motor-evoked potentials in the ADM muscle of older adults. *J Neurosci Methods.* 2007;164(2):320–324. doi:10.1016/j.jneumeth.2007.05.011.
- [55] Cavaleri R, Schabrun SM, Chipchase LS. The number of stimuli required to reliably assess corticomotor excitability and primary motor cortical representations using transcranial magnetic stimulation (TMS): a systematic review and meta-analysis. *Syst Rev.* 2017;6(1):48. doi:10.1186/s13643-017-0440-8.
- [56] Lin LF, Huang YZ, Hu CJ, et al. Using surface electromyography to guide the activation during motor-evoked potential measurement: an activation control method for follow-up studies. *Brain Inj.* 2015;29(13–14):1661–1666. doi:10.3109/02699052.2015.1075150.