Experimental Brain Research
https://doi.org/10.1007/s00221-021-06160-2

RESEARCH ARTICLE

Measurement error and reliability of TMS metrics collected from biceps and triceps in individuals with chronic incomplete tetraplegia

Tarun Arora[1] · Kelsey Potter-Baker[2] · Kyle O'Laughlin[1] · Manshi Li[3] · Xiaofeng Wang[3] · David Cunningham[4,5,6] · Francois Bethoux[7] · Frederick Frost[7] · Ela B. Plow[1,7]

Received: 31 December 2020 / Accepted: 20 June 2021

Abstract

Transcranial magnetic stimulation (TMS) is used to investigate corticomotor neurophysiology associated with functional recovery in individuals with spinal cord injury (SCI). There is insufficient evidence about test–retest measurement properties of TMS in SCI. Therefore, we investigated test–retest agreement and reliability of TMS metrics representing corticomotor excitability, output, gain, map (representation), and inhibition in individuals with cervical SCI. We collected TMS metrics from biceps and triceps muscles because of the relevance of this proximal muscle pair to the cervical SCI population. Twelve individuals with chronic C3–C6 SCI participated in two TMS sessions separated by ≥ 2 weeks. Measurement agreement was evaluated using *t* tests, Bland–Altman limits of agreement and relative standard error of measurement (SEM%), while reliability was investigated using intra-class correlation coefficient (ICC) and concordance correlation coefficient (CCC). We calculated the smallest detectable change for all TMS metrics. All TMS metrics except antero-posterior map coordinates and corticomotor inhibition were in agreement upon repeated measurement though limits of agreement were generally large. Measures of corticomotor excitability, output and medio-lateral map coordinates had superior agreement (SEM% < 10). Metrics representing corticomotor excitability, output, and inhibition had good-to-excellent reliability (ICC/CCC > 0.75). The smallest detectable change for TMS metrics was generally high for a single individual, but this value reduced substantially with increase in sample size. We recommend use of corticomotor excitability and recruitment curve area owing to their superior measurement properties. A modest group size (20 or above) yields more stable measurements, which may favor use of TMS metrics in group level modulation after SCI.

Keywords Transcranial magnetic stimulation · Spinal cord injury · Reliability · Measurement error · Smallest detectable change

Communicated by Winston D Byblow.

✉ Ela B. Plow
plowe2@ccf.org

[1] Plow Lab (ND 20), Department of Biomedical Engineering, Lerner Research Institute, Cleveland Clinic Foundation, 9500 Euclid Ave, ND20, Cleveland, OH 44195, USA

[2] Department of Molecular Science, School of Medicine, The University of Texas Rio Grande Valley, Edinburg, TX, USA

[3] Department of Quantitative Health Sciences, Cleveland Clinic Foundation, Cleveland, OH, USA

[4] MetroHealth Center for Rehabilitation Research, Cleveland, OH, USA

[5] Department of Physical Medicine and Rehabilitaiton, Case Western Reserve University, Cleveland, OH, USA

[6] Cleveland Functional Electrical Stimulation Center, Cleveland, OH, USA

[7] Department of Physical Medicine and Rehabilitation, Cleveland Clinic Foundation, Cleveland, USA

Published online: 10 August 2021

Introduction

Regaining upper extremity movement function is a top priority among individuals with tetraplegia (Anderson 2004). However, recovery of upper extremity function is often slow and incomplete. Understanding the neurophysiologic underpinnings of recovery in individuals with cervical spinal cord injury (SCI) can help develop more effective rehabilitation strategies for functional restoration (Oudega and Perez 2012).

Transcranial magnetic stimulation (TMS) is a non-invasive technique that can help evaluate the residual potential of descending corticomotor pathways. When applied over the motor cortex, TMS produces motor-evoked potentials (MEPs) in muscles, which can be recorded using surface electromyography (EMG). Features of MEPs and ensuing EMG activity are often used to make inferences about corticomotor physiology, including corticomotor excitability (threshold of excitation), output and gain, intracortical mechanisms, and topographical representations (or motor cortical maps) (Rossi et al. 2009).

Given its non-invasive nature, TMS is widely used as a tool to evaluate motor recovery potential in individuals with SCI. Many studies have used TMS to characterize effects of spontaneous functional recovery and intervention-associated improvements after SCI (Hoffman and Field-Fote 2007; Gomes-Osman and Field-Fote 2015a, b; Edwards et al. 2019; Leao et al. 2020). Given the association evidenced between TMS metrics and functional improvements, TMS could serve as an important measurement tool to assess mechanisms of rehabilitation and recovery in SCI. Therefore, it is important to establish the robustness and precision of TMS metrics across test–retest measurements.

Test–retest agreement and test–retest reliability are important properties of a measurement tool (Mokkink et al. 2010). While agreement refers to the capacity of an instrument to produce identical results when administered twice on the same subjects under similar conditions, reliability refers to its capacity to keep ordering among participants consistent (de Vet et al. 2006; Berchtold 2016). Simply put, agreement describes the precision of an instrument while reliability refers to its potential to differentiate among participants. Test–retest agreement is generally described using measurement error. For TMS to be considered useful as an assessment tool in SCI, the measurement error of its metrics needs to be smaller than the change expected to be witnessed across repeated measurements. Test–retest reliability is typically indexed using reliability coefficients like the widely known intra-class correlation coefficient (ICC) (Berchtold 2016). The ICC, which is a reliability parameter evaluates how participants can be differentiated from each other (de Vet et al. 2006; Berchtold 2016). For TMS to be considered reliable across repeated measurements in differentiating among participants with SCI, its metrics need to have higher values of ICC (>0.75) (Koo and Li 2016). The knowledge of these two measurement properties, agreement and reliability, is essential to make prognostic (e.g., changes in an individual/group over time) or diagnostic (e.g., individuals with more severe injuries consistently rated better/worse than less severe injuries) interpretations of the TMS metrics, respectively.

Currently, there is insufficient evidence about test–retest measurement properties of TMS in SCI. Our group has published on the test–retest reliability of TMS metrics related to corticomotor excitability, output and motor cortical maps in a small sample ($n=8$) of individuals with C2–C6 SCI (Potter-Baker et al. 2016). In that study, corticomotor properties were evaluated for pair of muscles, including deltoid ($n=6$), biceps ($n=3$), triceps ($n=1$), extensor digitorum ($n=5$), and first dorsal interossei ($n=1$), which included one muscle innervated above the level of injury and the second muscle innervated below the level of injury. Findings revealed test–retest reliability was good-to-excellent for metrics collected from muscles with higher muscle power but poor for metrics collected from muscles with lower muscle power. Clinically, studying a muscle below and a muscle above the level of injury allows to understand the properties of a less and a more impaired muscle, respectively. However, the differences in corticomotor contributions to proximal and distal muscles (Malcolm et al. 2006; Carson et al. 2013) add to TMS variability between individuals, which affects the test–retest reliability (de Vet et al. 2006). In addition, the previous study did not investigate test–retest agreement properties, which hold important evaluative information for those interested in using TMS for prognosticating changes over time. Therefore, the purpose of the present study was to investigate the test–retest agreement and reliability of TMS metrics across a homogeneous set of muscles in persons with cervical SCI. In the current study, we investigated the biceps and triceps as they represent a functionally relevant muscle pair with biceps involved in actions involving feeding and drinking, while triceps are involved in actions related to weight-shifts during pressure-relief activities and transfers. The biceps muscle is innervated (C5) at or above and the triceps muscle is innervated (C7) below the level of a typical C5–C6 injury. Clinically, evaluation of two muscles one at or above and the other below the level of injury we believed would help determine whether TMS is a useful measurement tool in a more impaired and a less impaired muscle. Neurophysiologically, studying a more homogenous group of muscles will limit the inter-individual differences in TMS outcomes due to differences in the corticomotor properties of these muscles.

Methods

Participants

Individuals with chronic tetraplegia following incomplete (American Spinal Injury Association Impairment Scale (AIS) levels B, C and D) C2–C8 injury were included. Participants were required to be at least 18 years of age and more than 6 months post-injury. We included data from three participants (# 4, 11 and 12) from our previous study where the data was collected using the same methods, by the same examiner, and at least one of the tested muscles was biceps or triceps. Individuals were excluded for contraindications to TMS, such as positive history of seizures, family history of medication-resistant epilepsy in a first-degree relative, and intra-cranial metallic implants or cardiac pacemaker (Rossini et al. 2015). Individuals taking neuro- or psycho-active medications were allowed to participate as long as consistent drug dosing and regimen was maintained (Potter-Baker et al. 2016). The study was approved by the Institutional Review Board of the Cleveland Clinic and the Human Research Protections Office of the U.S. Department of Defense. All participants provided written informed consent.

Muscle power assessments

TMS was used to test corticomotor properties of biceps and triceps muscles on the weaker side of the body. Weaker side was determined based on Upper Extremity Motor Score (UEMS) testing performed by a physical therapist. UEMS is a valid and reliable scale to assess muscle power in individuals with tetraplegia (Graves et al. 2006; Marino et al. 2008). Muscle power was rated on the Medical Research Council (MRC) scale (0–5), where 0 = no movement or contraction; 1 = palpable or visible contraction; 2 = active movement, full range of motion, gravity eliminated; 3 = active movement, full range of motion, against gravity; 4 = active movement, full range of motion, against gravity and some resistance; and 5 = active movement, full range of motion, against gravity and maximal resistance. We did not perform International Standards for Neurological Classification of SCI (ISNCSCI) examination, but obtained the American Spinal Injury Association (ASIA) Impairment Scale (AIS) scores and level of injury from the medical record of the participants (Table 1).

TMS assessments

TMS was performed on two separate occasions, ≥ 2 week apart. Same investigator (KPB) carried out the TMS testing on all the participants and at both time points. Participants were asked to carry out daily activities and typical routine during the 2-week interval, but were asked to refrain from participating in any new upper extremity training or rehabilitation. Every effort was made to keep the time and day of testing the same to remove confounds of diurnal variations and medication effects.

TMS was delivered using a 70 mm figure-of-eight coil connected to a device which delivers monophasic pulses (Magstim 200^2, Magstim, Dyfed, UK). The coil was held tangential to the scalp and oriented postero-lateral to the mid-sagittal axis at a 45^0 angle. Coil targeting was guided based on frameless stereotactic neuro-navigation (Brainsight, Rogue Research, Montreal, QC, Canada). Either T1-weighted magnetic resonance image (MRI) of individual's brain or template brain image was used for reference

Table 1 Demographics and injury characteristic of the participants

Participant ID	Sex	Age (years)	Months post-injury	AIS	Level	Biceps strength MRC	Triceps strength MRC
1	F	28	47	B	C5	2	0
2	F	67	417	B	C4	3	1
3	F	32	79	B	C5	4	2
4	M	48	98	B	C5	3*	2
5	M	62	86	C	C6	5	3
6	M	47	164	D	C4	4	3
7	M	58	368	D	C6	4	2
8	M	36	251	D	C6	5	3
9	M	56	60	D	C5	4	3
10	M	68	135	D	C3	3	4
11	M	52	36	D	C6	3	3*
12	M	56	54	D	C4	3	3*

AIS American Spinal Injury Association Impairment Scale, *MRC* Medical Research Council, *NT* tested

*TMS data for these muscles not available. Note cervical segmental level of injury information is based on medical chart review from the time of the injury

(Potter-Baker et al. 2016). Surface cranial landmarks on participant's head (nasion, left auricular and right auricular) were registered to the corresponding sites on the MRI image. Registration error between surface cranial landmarks and MRI/template landmarks was not allowed to exceed 3 mm.

Participants were tested while seated in a height-adjustable chair or a wheelchair with their forearms resting on a table in front. Shoulders were in slight neutral or abducted and flexed position while elbows were in 90° flexion. Forearms were in supination during testing of the biceps muscle and in neutral rotation during testing of the triceps muscle. TMS-evoked responses were collected from muscles using Ag/AgCl surface EMG electrodes (45 mm diameter). The same investigator palpated the muscles to identify the muscle belly and applied the EMG electrodes in a belly–belly montage during both tests. EMG data were sampled at 4 kHz and filtered using a band-pass filter (10 Hz–2 kHz) (PowerLab 4/25, AD Instruments, Colorado Springs, CO, USA). A ground electrode was place on the test side clavicle.

Prior to TMS data collection, participants underwent evaluation of maximum volitional contraction (MVC) strength of the test muscles. For evaluation of MVC of biceps, participants were asked to flex elbow against investigator's manual resistance force applied to distal forearm (biceps) and for evaluation of MVC of triceps, participants were asked to extend elbow against physical resistance offered by the rigid tabletop underneath. Three trials (3–5 s each) were performed for each muscle with ~ 30-s rest in between. Average EMG of the 3 trials was derived for each muscle and used as MVC.

During TMS data collection, participants were asked to maintain slight active contraction of the test muscle (10–20% MVC). A pre-activated state of the muscle makes it easier to generate descending volleys using TMS (Rossini et al. 2015) {Rossini, 2015 #1}. Patients were given online visual feedback of EMG activity of the test muscle on a computer screen. Frequent breaks were given to avoid fatigue, which included shorter breaks between every 5 TMS pulses and longer breaks when switching between different metrics. Using TMS, first we identified motor hotspots i.e., sites generating criterion-sized MEPs (peak-to-peak amplitude ≥ 200 μV above average baseline contraction) in at least 5/10 consecutive trials at the lowest TMS intensity. The hotspots were identified separately for each muscle during its slight active contraction (10–20% MVC). The lowest intensity of TMS required at the hotspot to elicit criterion

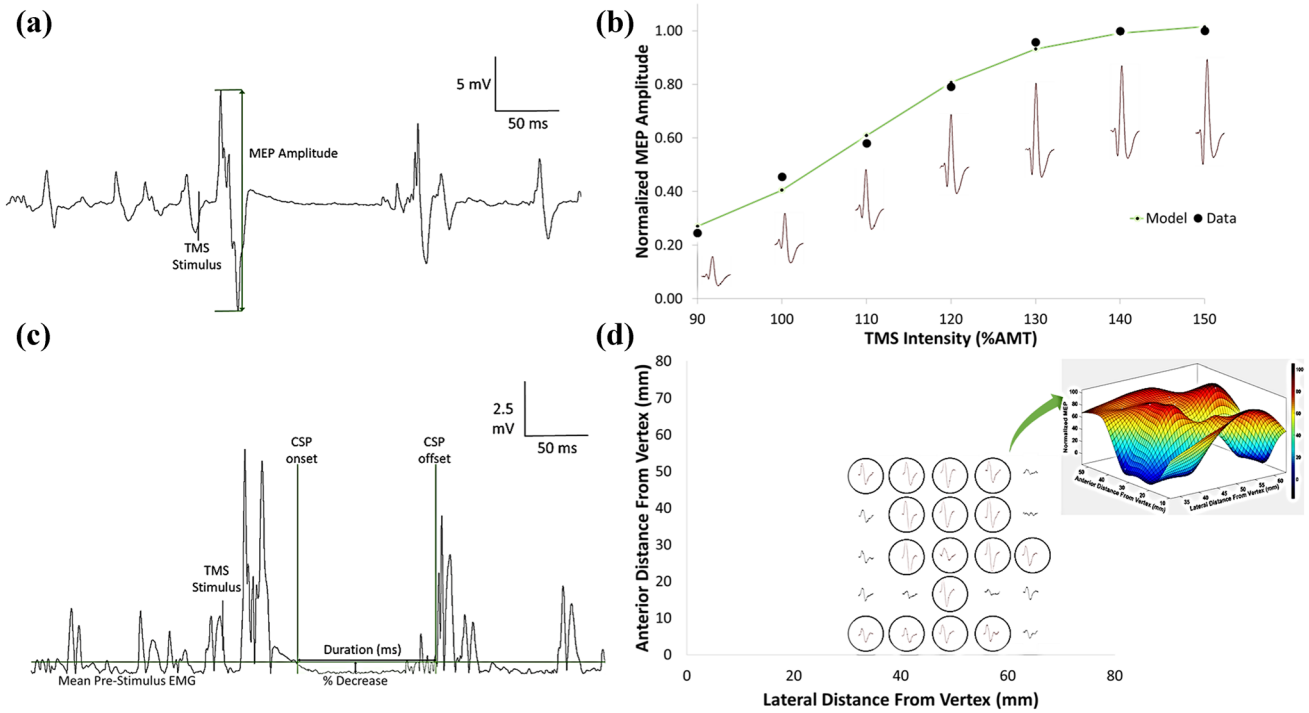

Fig. 1 Schematic representation of TMS metrics from biceps after SCI—**a** motor-evoked potential (MEP) and illustration of how peak-to-peak amplitude is derived, along with representation of the ensuring cortical silent period (CSP) at 120% active motor threshold (AMT), **b** a recruitment curve (RC) with individual averaged (across 15 trials) MEP traces shown for each incremental intensity level tested, **c** rectified EMG trial where CSP or transient interruption of volitional EMG can be seen more clearly after an earlier peak of EMG activity that stands for MEP, and **d** MEP traces collected from the motor map across 5 × 5 sites. Active map sites are marked with circles around them. Map Area in this example was 16 (out of total 25) sites. A colorful contour map is used to show map volume distribution i.e., spread of MEP sizes across the area of the map

MEPs from the test muscle was called active motor threshold (AMT), expressed as a percentage of the maximum stimulator output (%MSO). The AMT was determined online at the time of data collection. During Test 2, to minimize the bias from knowledge of Test 1 AMT values, the evaluators did not review the Test 1 values during the 2-week interval. AMT is a measure of corticomotor excitability and reflects the threshold of activation of descending motor tracts devoted to the contralateral muscle (Rossini et al. 2015). Besides AMT, we collected 15 suprathreshold MEPs from each test muscle using 120% AMT intensity (Fig. 1a). We used average value of peak-to-peak amplitudes of all 15 suprathreshold MEPs as the active motor-evoked potential (AMEP; mV).

Once motor hotspots were identified for each test muscle, we characterized the input–output properties of corticomotor output using what is known as a stimulus–response or recruitment curve (RC) (Fig. 1b). A RC shows the relationship between increments of TMS intensity and corresponding amplitude of MEPs acquired at each tested intensity level. RCs are believed to illustrate the gain and overall output of corticomotor pathways (Devanne et al. 1997). To collect RC, we delivered TMS at the motor hotspot at multiple increments of intensity ranging between 90 and 150% AMT (in increments of 10% MSO) for 15 trials each. Incremental levels of intensity were tested in random order. Average MEP amplitude for each intensity level was calculated and normalized to the maximum MEP amplitude acquired from that hotspot. Normalized MEP values were fit into a Boltzmann sigmoid function (Carson et al. 2013). Steepness or slope of the RC, i.e., the amount of increase in MEP amplitude with unit increment of TMS intensity, was used as a measure of corticomotor gain (RC_Slope) and discarded if the fit with the Boltzmann sigmoid function was poor ($R^2 < 0.70$). The overall corticomotor output was also calculated as area under the fitted RC (RC_Area) (Carson et al. 2013).

From the 15 MEP trials acquired at suprathreshold TMS intensity (120% AMT) during RC data collection, we also acquired the cortical silent period (CSP). CSP is defined as transient interruption of ongoing volitional EMG activity in contracting muscle following contralateral TMS. CSP typically follows a contralateral MEP, though it can appear even at lower intensities where a clear MEP may not be evident. CSP is believed to reflect intracortical and spinal inhibitory mechanisms affecting corticomotor output (Fuhr et al. 1991; Chen et al. 1999). Assessment of CSP holds value among individuals with SCI as weakened corticomotor inhibition has been reported in this population, which is thought of be important for recovery of motor function (Smith et al. 2000). For the measurement of CSP, we rectified MEPs collected at 120%AMT and removed trials where pre-stimulus EMG fell outside the mean + 2 standard deviation of pre-stimulus EMG activity seen across 15 trials. CSP onset was defined as the time point following TMS stimulus artifact where the average EMG signal dropped below the mean pre-stimulus activity and remained there for at least 10 ms. CSP offset was defined as the time point where the EMG activity returned through the pre-stimulus level and remained for at least 50% of 10 ms window. The duration (msec) and the extent (percentage) of decrease in EMG during CSP was calculated using a customized MATLAB script (Fig. 1c). CSP duration was calculated as the difference between CSP offset and onset. The extent of CSP was given as EMG during CSP expressed as a percentage of pre-stimulus EMG (100 ms before TMS stimulus): $\frac{\text{average CSP EMG}}{\text{average prestimulus EMG}} \times 100$.

The final procedure using TMS involved collecting motor cortical maps. Motor cortical maps are believed to reflect corticomotor representation devoted to the test muscle (Rossini et al. 2015). Suprathreshold TMS pulses (110%AMT) were delivered to scalp sites on a 5 × 5 grid (10 mm resolution) centered at the motor hotspot during slight active contraction (10–20% MVC) of the test muscle. Since individuals with SCI can have high AMT values for biceps and triceps, we chose 110% AMT intensity for collecting maps to avoid saturating the machine output. Five trials were collected from each site. All TMS pulses were delivered at ~ 0.2 Hz throughout the session. MEP amplitudes for each trial from every scalp site were recorded (LabChart v. 7.3, ADInstruments, Colorado Springs, CO, USA). Corresponding spatial coordinates of scalp sites were also collected using the neuro-navigation software (Brainsight Inc.). Map Area was defined as number of "active" map sites, i.e., sites that elicited an MEP size of at least 1 standard deviation above the pre-stimulus activity in at least 3 out of 5 trials (Potter-Baker et al. 2016) (Fig. 1d). Map Volume was calculated as the sum of normalized MEPs across all the active sites (Potter-Baker et al. 2016). We calculated center of gravity as the weighted average location MEPs within a map given using *X* and *Y* coordinates in medio-lateral (COG *X*) and antero-posterior (COG *Y*) planes: COG $\sum \frac{\text{MEPi} \times xi}{\text{MEPi}}$ (Potter-Baker et al. 2016). To facilitate analysis, COG *X* values were inverted in participants in whom the right arm was tested.

Figure 1 shows the schematic representation of TMS metrics acquired in this study. The calculations associated with each metric are reported in Online appendix A.

Statistical and data analysis

All statistical analyses were carried out using R studio software and SAS 9.4 for Linux (SAS, Cary, North Carolina). For each muscle, differences between repeated baseline scores (Test 1–Test 2, called Change Score) and the average of repeat baseline scores (mean of Test 1 and Test 2, called

Mean Score) were computed. Bland–Altman plots were developed to estimate limits of agreement between the two measurements (Online Appendices B and C). We plotted reference lines parallel to the *x*-axis at the mean Change Score ± 1 × SD and 1.96 × SD of Change Score. Normality for differences in test–retest measurements was investigated by plotting histograms and using the Shapiro–Wilk test. Overall agreement between repeated measurements was evaluated using paired *t* tests where non-significant results meant the measurements were in agreement. The extent of the agreement was assessed by estimating measurement error upon repeated measurements such that a higher error would indicate lesser agreement. Measurement error was estimated using standard error of measurement (SEM) (de Vet et al. 2006), which was calculated as the square root of variance within sessions: $\mathrm{SEM} = \sqrt{\sigma 2\ \text{within sessions}}$. The measurement error was also expressed relative to the mean value (SEM%) of the measurement: $\mathrm{SEM\%} = \frac{\mathrm{SEMeas}}{\mathrm{Mean}} \times 100$. SEM shows the magnitude of absolute measurement error between Test 1 and Test 2, while SEM% shows the error relative to the mean. An SEM% value of less than 10% is indicative of high measurement stability (Schambra et al. 2015), which aids in understanding the noisiness of the measurement. Using the SEM values, we also identified the smallest detectable change for an individual, which is the smallest change in an individual that is above the inherent measurement error and can be reliably detected; $\mathrm{SDC_{indiv}} = \mathrm{SEMeas} \times \sqrt{2} \times 1.96$. These smallest detectable values are smaller for a group of individuals ($\mathrm{SDC_{group}}$) and reduce further with an increase in sample size. We calculated $\mathrm{SDC_{group}}$ as: $\mathrm{SDC_{group}} = \frac{\mathrm{SDC_{indiv}}}{\sqrt{\text{group size}}}$ and plotted against hypothetical sample sizes to visualize how the values reduced with an increased sample size and reported them for $N = 1, 5, 10, 20$ and 30.

We calculated reliability using the intra-class correlation coefficients (ICC) and concordance correlation coefficients (CCC). ICC was calculated using the following:

$\mathrm{ICC} = \frac{\sigma 2\text{between}}{\sigma 2\text{within} + \sigma 2\text{between}}$, where σ2between was the variance between participants and σ2within was the variance within sessions. ICC values can range between 0 and 1, where values greater than 0.75 indicate good-to-excellent reliability, 0.50 to 0.75 moderate reliability, and values below 0.50 indicate poor reliability (Koo and Li 2016). CCC was calculated using the following: $\mathrm{CCC} = \frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2 + (\mu_x + \mu_y)^2}$, where μ_x and μ_y are the means of the two variables and σ_x^2 and σ_y^2 are the corresponding variance. ρ is the correlation coefficient between the two variables. CCC is considered the non-parametric alternative for assessing reliability and was included in the analysis due to our smaller sample size. The values range between − 1 (perfect disagreement) and + 1 (complete agreement). No clear demarcations have been defined for CCC, therefore, for this study we used CCC > 0.75 to represent good-to-excellent reliability and lower values as poor-to-moderate reliability. To study the association between impairment levels and measurement error, we calculated the correlation between change in values of each variable from Test 1 to Test 2, and the MRC scores separately for both muscles.

Results

Twelve individuals with tetraplegia [mean ± sd (Range) age = 51 ± 13 years (28–68 years)] with C3–C6 injuries (4 with AIS B, 1 with AIS C, and 7 with AIS D) participated (Table 1). Participants did not report any lifestyle changes between the two TMS sessions. We were able to obtain test–retest data from biceps muscle in 11 participants, and triceps muscle in 10 participants. All participants, except one, had a lower muscle power for triceps [average MRC (range) = 2.3 (0–4)] than biceps [average MRC (range) = 3.6 (2–5)]. We could acquire MEPs in triceps muscle even in persons with muscle power or MRC grade of 0. We could not obtain RCs from the triceps muscle in 5 participants (#1,3,5,6,9) and the biceps muscle in 3 participants (#2,3,12) because either high values of AMT precluded testing at TMS intensities above 100% MSO or there was a poor fit of the Boltzmann model ($R^2 < 0.70$). CSPs were found in the biceps of all but one participant (#9) and in the triceps of all but three participants (#2,6,7). The differences in test–retest measures were found to be normally distributed upon Shapiro–Wilk test. There were no significant correlations between the change in scores on repeated measurements, and MRC grades for biceps and triceps muscles meaning that persons with lower muscle power did not necessarily have greater variation in TMS measures from one time point to another.

Measurement error properties

Table 2 reveals findings of *t* tests used to ascertain agreement across repeated measurements. Based on *t* test results, all TMS metrics were in agreement between Test 1 and Test 2 except COG Y ($p = 0.03$) and CSP % ($p = 0.04$) collected in biceps. COG Y for biceps had a change 5.8 (8.1) mm while CSP% for biceps had a change of 6.46 (7.58) %. Using Bland–Altman plots, acceptable limits of agreement are typically compared to a priori values derived from clinical consideration, but such clinical values for TMS metrics are not known. Limits of agreement found in our study were larger than changes typically seen with interventions. For example, limits of agreement for AMT for biceps and triceps was − 13.41 to + 10.32% MSO and − 19.93% to 15.73% MSO, respectively (Table 3). Across different metrics, agreement was superior for metrics characterizing corticomotor excitability and output from motor hotspot. SEM% values were

Table 2 Mean and standard deviation for TMS metrics on day 1 and 2, and mean change scores and one-sample *t* test scores for the change scores from biceps and triceps

	Biceps					Triceps				
	N	Test 1 Mean (SD)	Test 2 Mean (SD)	Change score Mean (SD)	*t* (*p*)	N	Test 1 Mean (SD)	Test 2 Mean (SD)	Change score Mean (SD)	*t* (*p*)
AMT, %MSO	11	50.6 (14.4)	52.2 (13.1)	1.55 (6.06)	0.85 (0.42)	10	62.2 (13.5)	64.3 (14.7)	2.1 (9.1)	0.73 (0.48)
AMEP, mV	11	2.1 (2.7)	1.6 (2.3)	−0.43 (0.94)	−1.51 (0.16)	10	0.44 (0.24)	0.37 (0.20)	−0.07 (0.20)	−1.13 (0.29)
RC_Area	8	37.7 (9.3)	35.7 (8.1)	−2.04 (4.54)	−1.27 (0.25)	5	38.3 (7.0)	40.7 (5.4)	2.41 (6.07)	0.89 (0.42)
RC_Slope	8	0.20 (0.06)	0.22 (0.07)	0.018 (0.053)	0.94 (0.38)	5	0.20 (0.06)	0.18 (0.05)	−0.022 (0.079)	−0.62 (0.57)
Map volume	11	1018.4 (394.4)	976.9 (299.2)	−41.5 (359.4)	−0.38 (0.71)	10	967.7 (342.4)	864.1 (299.1)	−103.6 (283.8)	−1.16 (0.28)
Map Area, Count/25	11	14.7 (5.8)	15.1 (3.9)	0.36 (5.26)	0.23 (0.82)	10	13.8 (4.8)	13.6 (5.7)	−0.2 (4.7)	−0.13 (0.90)
COGX, mm	11	42.8 (6.6)	41.0 (3.8)	−1.8 (3.8)	−1.57 (0.15)	10	38.7 (5.6)	36.8 (5.4)	−2.0 (7.6)	−0.82 (0.43)
COGY, mm	11	25.9 (9.8)	31.7 (10.0)	5.8 (8.1)	2.38 (0.04)*	10	27.4 (13.0)	32.6 (8.8)	5.2 (8.0)	2.03 (0.07)
CSP_Durn, ms	10	76.1 (32.1)	82.7 (28.6)	6.60 (11.49)	1.09 (0.31)	7	91.3 (46.0)	81.0 (35.8)	−10.28 (26.87)	1.01 (0.35)
CSP %	10	42.7 (8.8)	36.2 (5.2)	−6.46 (7.58)	−2.70 (0.03)*	7	36.4 (6.9)	39.1 (4.8)	2.69 (5.10)	1.40 (0.21)

*Significant results with $p < 0.05$ suggesting lack of agreement between repeated measurements

less than 10% for AMT, RC_Area and COGX of biceps and AMT of triceps (though SEM% for RC_Area of triceps was also close to 10% at 10.62%).

In Table 4, we report SDC values of TMS metrics collected across biceps and triceps muscles for different hypothetical sample sizes based on measurement error findings described above. SDC_{indiv} values ($N = 01$) collected for biceps and triceps are large indicating substantial changes in TMS metrics are needed to be reliably detected in an individual with tetraplegia (Table 4). When the study sample size is increased, SDC values reduce considerably (Fig. 2). For example, SDC_{indiv} for AMT from biceps is 12% MSO, suggesting only a change larger than this value can be reliably detected in an individual. In contrast, SDC_{group} collected across 20 individuals ($N = 20$) is only 3% MSO, indicating even a small change can be reliably detected if the sample size is increased. Exceeding a small change is more realistic with interventions and stresses the importance of larger sample sizes in studies of persons with SCI.

Reliability

Table 5 shows reliability—ICC and CCC—values for all TMS metrics collected from biceps and triceps muscles. Reliability was good-to-excellent (ICC and CCC > 0.75) for several metrics collected from biceps including AMT, AMEP, RC_Area, and CSP duration. For triceps, only AMT and CSP duration were found to have good-to-excellent reliability.

Discussion

The objective of the present study was to investigate test–retest measurement properties of TMS metrics collected from biceps and triceps muscles in the weaker upper limb of persons with tetraplegia after incomplete cervical SCI. Overall, TMS metrics characterizing corticomotor excitability, output, gain and corticomotor map size agree upon repeated measurements, i.e., these metrices did not significantly differ between repeated measurements. However, metrics characterizing antero-posterior coordinates of maps and corticomotor inhibition for biceps muscle had poor agreement across repeated measurements. We also observed that of all TMS metrics, measures of corticomotor excitability were most stable with the smallest relative measurement error and good-to-excellent reliability. Measures of corticomotor gain and maps had higher measurement error and poor-to-moderate reliability. Below we explain how these findings carry value for making TMS assessments in the cervical SCI population.

Table 3 Measurement error for the TMS metrics from biceps and triceps

	Biceps				Triceps			
	N	SEM	SEM%	Limits of agreement	*N*	SEM	SEM%	Limits of agreement
Excitability and output								
AMT, %MSO	11	4.23	8.22*	− 13.41 to 10.32	10	6.28	9.93*	− 19.93 to 15.73
AMEP, mV	11	0.7	37.71	− 1.41 to 2.26	10	0.14	35.31	− 0.32 to 0.47
RC_Area	8	3.33	9.07*	− 6.87 to 10.94	5	4.20	10.62	− 14.3 to 9.48
Gain								
RC_Slope	8	0.04	17.79	− 0.12 to 0.09	5	0.05	27.94	− 0.13 to 0.18
Maps								
Map volume	11	244.1	24.5	− 663.0 to 746.0	10	204.0	22.3	− 452.5 to 659.8
Map area, Count/25	11	3.6	23.8	− 10.7 to 9.9	10	3.2	23.0	− 9.0 to 9.4
COGX, mm	11	2.8	6.8*	− 5.6 to 9.2	10	5.3	14.0	− 12.9 to 16.8
COGY, mm	11	6.8	23.6	− 21.6 to 10.0	10	6.5	21.7	− 20.9 to 10.6
Inhibition								
CSP_Durn, ms	10	13.70	17.26	− 44.24 to 31.04	7	19.03	22.10	− 42.39 to 62.94
CSP %	10	6.83	17.33	− 8.39 to 21.31	7	3.84	10.19	− 12.69 to 7.30

*Measures with low relative measurement error (SEM% < 10%)

Table 4 Smallest detectable change (SDC) values for the TMS metrics from biceps and triceps for an individual (*N* = 01) and a group of individuals (*N* = 5, 10, 20, 30)

	Biceps SDC					Triceps SDC				
	N = 01	*N* = 05	*N* = 10	*N* = 20	*N* = 30	*N* = 01	*N* = 05	*N* = 10	N = 20	*N* = 30
AMT, %MSO	12	5	4	3	2	17	8	6	4	3
AMEP, mV	1.94	0.87	0.61	0.43	0.35	0.40	0.18	0.13	0.09	0.07
RC_Area	9.23	4.13	2.92	2.06	1.69	11.64	5.21	3.68	2.60	2.13
RC_Slope	0.10	0.04	0.03	0.02	0.02	0.14	0.06	0.04	0.03	0.03
Map volume	676.6	302.6	214.0	151.3	123.5	565.4	252.9	178.8	126.4	103.2
Map Area, Count/25	10	4	3	2	2	9	4	3	2	2
COGX, mm	7.8	3.5	2.5	1.7	1.4	14.6	6.5	4.6	3.3	2.7
COGY, mm	18.9	8.5	6.0	4.2	3.5	18.0	8.1	5.7	4.0	3.3
CSP_Durn, ms	37.98	16.99	12.01	8.49	6.93	52.75	23.59	16.68	11.80	9.63
CSP %	18.94	8.47	5.99	4.24	3.46	10.65	4.76	3.37	2.38	1.94

AMT and map area values are rounded off to the nearest whole number

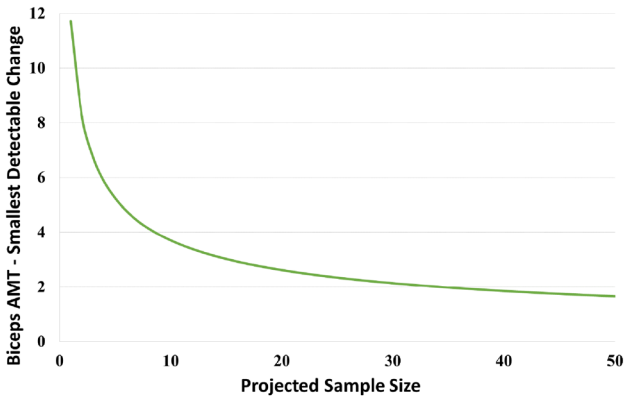

Fig. 2 Using example of biceps AMT to show that the smallest detectable change reduces considerably with increase in sample size allowing reliable assessment of smaller changes in TMS metrics

Type of muscles: proximal vs distal and more vs less impaired

In general, agreement and reliability values for TMS metrics seen in patients with SCI in our current study are lower than those previously reported in healthy individuals (Beaulieu et al. 2017) and individuals with stroke (Schambra et al. 2015). Difference in nature of muscles tested may have contributed to the differences in findings. We tested agreement and reliability of TMS metrics across proximal upper limb muscles, whereas majority of previous work in healthy individuals and individuals with stroke has involved test of distal upper limb muscles. TMS metrics obtained from proximal muscles are shown to have low agreement and reliability values even when collected across healthy individuals

Table 5 Reliability measurement properties for the TMS metrics from biceps and triceps

	Biceps			Triceps		
	N	ICC	CCC	*N*	ICC	CCC
Excitability and output						
AMT	11	0.91*	0.9*	10	0.8*	0.78*
AMEP	11	0.92*	0.92*	10	0.58	0.55
RC_Area	8	0.86*	0.84*	5	0.55	0.49
Gain						
RC_Slope	8	0.70	0.67	5	0.09	0.03
Maps						
Map volume	11	0.50	0.47	10	0.60	0.58
Map area	11	0.47	0.43	10	0.64	0.6
COGX	11	0.73	0.71	10	0.07	0.05
COGY	11	0.59	0.56	10	0.68	0.66
Inhibition						
CSP_Durn	10	0.80*	0.78*	7	0.79*	0.76*
CSP %	10	0.34	0.31	7	0.60	0.57

ICC intra-class correlation coefficient, *CCC* concordance correlation coefficient

*Measures with good-to-excellent reliability (ICC and CCC > 0.75)

(Brasil-Neto et al. 1992; Carson et al. 2013; Sankarasubramanian et al. 2015). Therefore, measurement properties of TMS metrics are dependent upon types of muscles tested. Reduced agreement and reliability of TMS values collected across proximal muscles can be ascribed to physiologic factors, such as fewer corticomotor projections distributed over a relatively wider area than distal muscles (Brasil-Neto et al. 1992).

Following SCI, there are anatomic and physiologic changes in corticomotor projections (Oudega and Perez 2012), which very likely added to the variability of neurophysiologic measurements performed using TMS. For example, there is an increase in motor threshold of corticospinal pathways, and an impaired modulation of intracortical inhibitory pathways following SCI (Oudega and Perez 2012). Due to corticomotor reorganization, there is also a possibility of more corticomotor projections to muscles rostral to injury (Levy Jr et al. 1990; Streletz et al. 1995), which may have contributed to better agreement values for biceps as compared to triceps. While it is not possible to tease out how much variability emerged from the injury itself, we tried to explore whether the level of impairment (characterized using muscle power or MRC grades) had any association with degree of variability witnessed. Lack of correlation between MRC grades and agreement values indicates that variability of TMS metrics is not related to the severity of

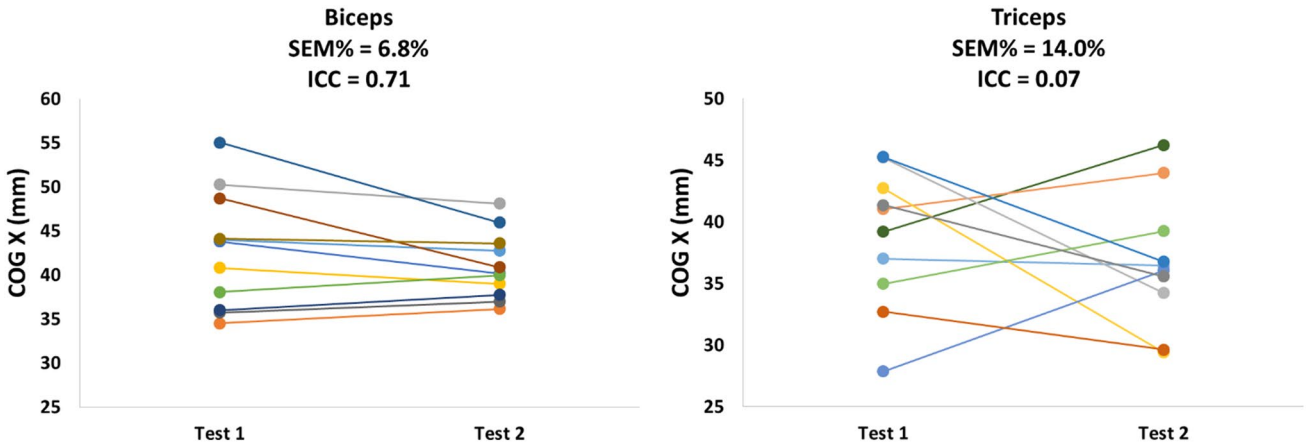

Fig. 3 Changes in COG *X* from Test 1 to Test 2 for each participant for biceps and triceps. Note a greater change in the magnitude of COGX and ranking order for each participant from Test 1 to Test 2 for triceps as compared to biceps, which corresponds with a greater measurement error and poor reliability for the triceps. SEM%—Standard Error of Measurement Relative to the Mean, and ICC—Intraclass Correlation Coefficient

motor impairment. Although, these results must be interpreted with caution as there was limited variability in the distribution of impairment scores—all but one participant had biceps MRC ≥ 3 and triceps MRC ≤ 3.

In general, we observed test–retest measurement properties for biceps were better than those for triceps. For example, Fig. 3 shows comparison between the two muscles upon the COG *X* metric. Note extent of change in COG *X* for different individuals was smaller for biceps than triceps, also reflected in smaller SEM% (6.8% vs 14%, respectively). Similarly, participants' rank-order across repeated testing was also more consistent for biceps than triceps, with ICC values being 0.7 vs 0.07, respectively. For biceps, participants who had relatively higher COG X values at Test 1 also had higher COG X values at Test 2, which was not the case for triceps. Differences in corticomotor contributions to the two muscles may explain differences in the measurement error as biceps generally has lower AMTs than triceps (a sign of relatively stronger corticomotor contributions) (Palmer and Ashby 1992; Brouwer and Hopkins-Rosseel 1997). Also, typically biceps have a better muscle power than triceps in persons with cervical SCI (since innervation for triceps at C7 lies below a typical C5–C6 injury), which was also seen in our study. For some of the metrics (e.g., AMEP and RC_Area), triceps had a smaller between-subject variance as evidenced in the Bland–Altmann plots (see Appendix C), which may have contributed to low ICC scores.

Choice of TMS metrics: corticomotor excitability, output, gain, maps or inhibition

Regardless of the choice of muscles, certain TMS metrics appear to have more favorable measurement properties than others. Table 6 provides recommended TMS metrics on the basis of good-to-excellent reliability (ICC and CCC > 0.75) and/or low measurement error (SEM < 10%). For example, AMT was found to have low measurement error, and

Table 6 Recommendations for use of TMS metrics based on absence or presence of good-to-excellent reliability (ICC and CCC > 0.75) and/or low measurement error (SEM < 10%)

	Metric	Biceps	Triceps
Excitability and Output	AMT	High Reliability and Low ME*	High Reliability and Low ME*
	AMEP	High Reliability# High ME	Low Reliability High ME
	RC_Area	High Reliability* Low ME	Low Reliability High ME
Gain	RC_Slope	Low Reliability High ME	Low Reliability High ME
Maps	Map Volume	Low Reliability High ME	Low Reliability High ME
	Map Area	Low Reliability High ME	Low Reliability High ME
	COGX	Low Reliability# Low ME	Low Reliability High ME
	COGY	Low Reliability High ME	Low Reliability High ME
Inhibition	CSP_Durn	High Reliability# High ME	High Reliability# High ME
	CSP %	Low Reliability High ME	Low Reliability High ME

Recommended*: Good-to-excellent reliability and low measurement error
Selectively Recommended#: Either good-to-excellent reliability or low measurement error
Use with caution: Neither good-to-excellent reliability nor low measurement error
ME: Measurement error

good-to-excellent reliability for both muscles. Previous studies have also reported measures of motor threshold have lower measurement error and higher reliability compared to other TMS metrics (Malcolm et al. 2006; Sankarasubramanian et al. 2015; Potter-Baker et al. 2016). Thus, our findings provide further support that threshold for exciting corticomotor output at the motor hotspot is a relatively stable TMS metric to use for repeated measurements. For measures of corticomotor output, measurement error is lower for metrics collected at range of intensities from a single site (the motor hotspot alone). For example, RC_Area had a smaller measurement error compared to Map Volume or Area. This is in line with previous research that shows that MEP variability is higher as the distance from the hotspot is increased, and when fewer stimuli are delivered (5 for maps versus 15 for RC measurements) (Brasil-Neto et al. 1992).

Recruitment curve slope had higher measurement error and poor-to-moderate reliability for both biceps and triceps. A previous study by Carson et al. has similarly reported poor reliability of recruitment curve slope from proximal muscles including biceps and triceps (Carson et al. 2013). In the same study, Carson et al. reported that goodness of fit to sigmoid function was inferior in proximal muscles than distal muscles (Carson et al. 2013). In our study also, the goodness of fit as measured using R^2 values was poor (< 0.70) for biceps muscle in 2 participants. In addition, we could not collect complete RC in participants with high AMT due to inability to test at intensities above 100% MSO which may have further compromised measurement of slope. A second factor contributing to poor test–retest properties of recruitment curve slope could be the level of background contraction. We had asked participants with tetraplegia to maintain slight (10–20% MVC) contraction to minimize fatigue. But a previous study in healthy individuals has shown that RC curve for biceps plateaus ~ 70% MSO when background contraction of 50% MVC is maintained (van Kuijk et al. 2009) and in the same study, it was reported that measurement error of recruitment curve slope was not different for biceps versus abductor digiti minimi muscles. Therefore, differences in background contraction maintained in our study versus previous studies in healthy and in general the poor goodness of fit of RC collected from proximal muscles may have contributed to high measurement error and poor-to-moderate reliability.

Motor maps were found to have low measurement error when represented using weighted average location of corticomotor excitability in medio-lateral plane (COG *X*). But the same was not true for weighted average location of corticomotor excitability in the antero-posterior plane (COG *Y*). Previous findings from our group and from others have also noted that COG *X* tends to be more reliable than COG *Y* upon repeated measurements (Potter-Baker et al. 2016) (Malcolm et al. 2006). Posterior shift in cortical representation of forearm muscles is common after incomplete chronic tetraplegia (Green et al. 1999; Freund et al. 2011). Similar posterior shifts in corticomotor excitability have been reported for maps of hand muscles after fatiguing exercise in healthy adults (Cunningham et al. 2016). Therefore, injury-induced reorganization of corticomotor representation and the susceptibility of these maps to fatigue in individuals with SCI may contribute to larger day-to-day variability in the *Y*-direction.

Among the measures for corticomotor inhibition, CSP duration was found to have good-to-excellent reliability for both muscles, which means the rank-order among participants was maintained over repeated measurements. However, measurement error associated with CSP duration was higher than for many other TMS metrics. This is not uncommon, while there was consistency in inter-individual differences, absolute values of in CSP duration measurements were not similar (or identical) from one testing time point to the next. High variability of CSP duration across participants could have contributed to these findings. The early part of the silent period is mainly due to spinal mechanisms, which may have contributed to inter-individual differences in CSP duration (Chen et al. 1999). These results suggest CSP duration might be useful for diagnostic purposes such as for consistently differentiating among injured participants but may not be stable across time to allow drawing inferences of longitudinal change.

Smallest detectable change: single individual vs group

To the best of our knowledge, this is the first report documenting SDC values of TMS metrics from biceps and triceps muscles in the SCI population. When using TMS for evaluation of an individual with SCI, only those changes that exceed SDC_{indiv} values can be attributed to real change; anything smaller could be considered a confound of random variation in measurement. From our results, SDC_{indiv} values for TMS metrics in SCI appear to be high and it is less likely that interventions could produce changes in an individual that exceed SDC_{indiv}. However, our findings also reveal that when using TMS for evaluation of a group of individuals with SCI, SDC_{group} values that need to be exceeded become smaller with increases in group size. Therefore, when studying TMS metrics in moderate-to-large groups (20–50) of individuals with SCI, it becomes more likely that the interventions could produce changes that exceed SDC_{group}, and thereby contribute to evidence of use of TMS metrics as biomarkers of recovery at a group level, but not at the level of single individual. A previous study in individuals with stroke has also suggested that TMS metrics can be reliably used as a biomarker to detect longitudinal changes when used in moderate-sized groups (Schambra et al. 2015). Overall, the

results of our study can be extrapolated to other studies only if the population and muscles being studied are the same and the samples have similar test–retest variance and between-subject variance.

Limitations

This study had a small sample size, especially for recruitment curve parameters, which limits the generalizability of results. Although we asked our participants to maintain the same lifestyle between the two sessions, the longer interval of ≥ 2 weeks could have led to variability due to uncontrolled factors. However, the values obtained using longer interval can be applied more meaningfully to longitudinal studies, which typically occur over longer periods. We used 5/10 criteria for AMT determination to limit fatigue in our participants, which may also have contributed to higher variability in the data as newer recommendations suggest more stringent criteria of 10/20 positive MEPs (Rossi et al. 2009). In addition, adaptive methods based on threshold-tracking algorithms may increase the accuracy of motor threshold determination and further lower their day-to-day variability (Groppa et al. 2012). The reliability of the output measures, such as AMEP may be limited in this study as we did not normalize the MEP amplitudes to their M-max values for the test muscle. However, we normalized other outcome measures, including recruitment curve area and Map Volume to the peak MEP values. One can argue that using the beginning of MEP as the CSP onset may remove the ambiguity that may be introduced by the end of the MEP (i.e., potential rebound after the MEP), while retaining the variations in MEP latency (Van Kuijk et al. 2014). We chose to exclude the MEP from CSP to limit the variability introduced by the MEP itself, which can be altered in the SCI population. Lastly, our results cannot be used to make interpretations about the clinical meaningfulness of TMS-associated changes as it was beyond the scope of this study to establish the minimal clinically important differences.

Conclusions

All TMS metrics except antero-posterior coordinates of map center of gravity and corticomotor inhibition were in agreement; however, the limits of agreement were generally large. Metrics representing corticomotor excitability, output, and inhibition had good-to-excellent reliability. The smallest detectable change appears to be higher in a single individual with incomplete tetraplegia but reduces substantially in groups of modest sample size. Therefore, TMS measures could not be reliably used as a biomarker to detect individual change, but may be able to detect group level modulation after SCI. Metrics of corticomotor excitability (active motor threshold) and output (recruitment curve area for biceps) had superior test–retest agreement and reliability properties, which makes them more favorable for performing TMS assessments of biceps and triceps function in SCI.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00221-021-06160-2.

Acknowledgements We thank Yin-Liang Lin for assistance with data collection, Daniel Janini for the initial analysis, and Patrick Chabra, Corin Bonnett, and Sarah Roelle for the recruitment efforts. We also thank all research participants for their valuable time and efforts.

Funding The U.S. Army Medical Research Acquisition Activity, 820 Chandler Street, Fort Detrick MD 21702-5014 is the awarding and administering acquisition office. This work was supported by the USAMRAA through the Spinal Cord Injury Research Programs under Award Nos. W81XWH1810530 and W81XWH1110707 to EP. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the Department of Defense or U.S. Army). This work was also supported by Conquer Paralysis Now grant (CPN1512) to EP and Cleveland Clinic Research Programs Committee award (2016-195) to KPB.

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

Declarations

Conflict of interest On behalf of all the authors, the corresponding author states that there is no conflict of interest.

Ethical approval The study was approved by the Institutional Review Board of the Cleveland Clinic and the Human Research Protections Office of the U.S. Department of Defense.

Consent to participate All the participants provided written informed consent.

References

Anderson KD (2004) Targeting recovery: priorities of the spinal cord-injured population. J Neurotrauma 21:1371–1383. https://doi.org/10.1089/neu.2004.21.1371

Beaulieu L-D, Flamand VH, Masse-Alarie H, Schneider C (2017) Reliability and minimal detectable change of transcranial magnetic stimulation outcomes in healthy adults: a systematic review. Brain Stimul 10:196–213

Berchtold A (2016) Test–retest: agreement or reliability? Methodol Innov 9:2059799116672875

Brasil-Neto JP, McShane LM, Fuhr P et al (1992) Topographic mapping of the human motor cortex with magnetic stimulation: factors affecting accuracy and reproducibility. Electroencephalogr Clin Neurophysiol Potentials Sect 85:9–16

Brouwer B, Hopkins-Rosseel DH (1997) Motor cortical mapping of proximal upper extremity muscles following spinal cord injury. Spinal Cord 35:205–212

Carson RG, Nelson BD, Buick AR et al (2013) Characterizing changes in the excitability of corticospinal projections to proximal muscles of the upper limb. Brain Stimul 6:760–768

Chen R, Lozano AM, Ashby P (1999) Mechanism of the silent period following transcranial magnetic stimulation evidence from epidural recordings. Exp Brain Res 128:539–542

Cunningham DA, Janini D, Wyant A et al (2016) Post-exercise depression following submaximal and maximal isometric voluntary contraction. Neuroscience 326:95–104

de Vet HCW, Terwee CB, Knol DL, Bouter LM (2006) When to use agreement versus reliability measures. J Clin Epidemiol 59:1033–1039

Devanne H, Lavoie BA, Capaday C (1997) Input-output properties and gain changes in the human corticospinal pathway. Exp Brain Res 114:329–338

Edwards DJ, Cortes M, Rykman-Peltz A et al (2019) Clinical improvement with intensive robot-assisted arm training in chronic stroke is unchanged by supplementary tDCS. Restor Neurol 37:167–180

Freund P, Rothwell J, Craggs M et al (2011) Corticomotor representation to a human forearm muscle changes following cervical spinal cord injury. Eur J Neurosci 34:1839–1846

Fuhr P, Agostino R, Hallett M (1991) Spinal motor neuron excitability during the silent period after cortical stimulation. Electroencephalogr Clin Neurophysiol 81:257–262

Gomes-Osman J, Field-Fote EC (2015a) Improvements in hand function in adults with chronic tetraplegia following a multi-day 10Hz rTMS intervention combined with repetitive task practice. J Neurol Phys Ther 39:23

Gomes-Osman J, Field-Fote EC (2015b) Cortical vs. afferent stimulation as an adjunct to functional task practice training: a randomized, comparative pilot study in people with cervical spinal cord injury. Clin Rehabil 29:771–782

Graves DE, Frankiewicz RG, Donovan WH (2006) Construct validity and dimensional structure of the ASIA motor scale. J Spinal Cord Med 29:39–45

Green JB, Sora E, Bialy Y et al (1999) Cortical motor reorganization after paraplegia: an EEG study. J Neurol 53:736

Groppa S, Oliviero A, Eisen A et al (2012) A practical guide to diagnostic transcranial magnetic stimulation: report of an IFCN committee. Clin Neurophysiol 123:858–882

Hoffman LR, Field-Fote EC (2007) Cortical reorganization following bimanual training and somatosensory stimulation in cervical spinal cord injury: a case report. Phys Ther 87:208–223

Koo TK, Li MY (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 15:155–163

Leao MTD, Wiesinger L, Ziemann U et al (2020) Rapid motor cortical reorganization following subacute spinal cord dysfunction. Brain Stimul 13:783–785

Levy WJ Jr, Amassian VE, Traad M, Cadwell J (1990) Focal magnetic coil stimulation reveals motor cortical system reorganized in humans after traumatic quadriplegia. Brain Res 510:130–134

Malcolm MP, Triggs WJ, Light KE et al (2006) Reliability of motor cortex transcranial magnetic stimulation in four muscle representations. Clin Neurophysiol 117:1037–1046

Marino RJ, Jones L, Kirshblum S et al (2008) Reliability and repeatability of the motor and sensory examination of the international standards for neurological classification of spinal cord injury. J Spinal Cord Med 31:166–170

Mokkink LB, Terwee CB, Patrick DL et al (2010) The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 63:737–745

Oudega M, Perez MA (2012) Corticospinal reorganization after spinal cord injury. J Physiol 590:3647–3663

Palmer E, Ashby P (1992) Corticospinal projections to upper limb motoneurones in humans. J Physiol 448:397–412

Potter-Baker KA, Janini DP, Frost FS et al (2016) Reliability of TMS metrics in patients with chronic incomplete spinal cord injury. Spinal Cord 54:980

Rossi S, Hallett M, Rossini PM et al (2009) Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. Clin Neurophysiol 120:2008–2039

Rossini PM, Burke D, Chen R et al (2015) 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 126:1071–1107

Sankarasubramanian V, Roelle SM, Bonnett CE et al (2015) Reproducibility of transcranial magnetic stimulation metrics in the study of proximal upper limb muscles. J Electromyogr 25:754–764

Schambra HM, Ogden RT, Martínez-Hernández I et al (2015) The reliability of repeated TMS measures in older adults and in patients with subacute and chronic stroke. Front Cell Neurosci 9:335

Smith HC, Savic G, Frankel HL et al (2000) Corticospinal function studied over time following incomplete spinal cord injury. Spinal Cord 38:292–300

Streletz LJ, Belevich JKS, Jones SM et al (1995) Transcranial magnetic stimulation: cortical motor maps in acute spinal cord injury. Brain Topogr 7:245–250

van Kuijk AA, Pasman JW, Hendricks HT et al (2009) Predicting hand motor recovery in severe stroke: the role of motor evoked potentials in relation to early clinical assessment. NeuroRehabilitation 23:45–51

Van Kuijk AAA, Bakker CD, Hendriks JCM et al (2014) Definition dependent properties of the cortical silent period in upper-extremity muscles, a methodological study. J Neuroeng Rehabil 11:1–9