Published in final edited form as:

Clin Neurophysiol. 2012 September ; 123(9): 1698–1704. doi:10.1016/j.clinph.2012.05.003.

A checklist for assessing the methodological quality of studies using transcranial magnetic stimulation to study the motor system: An international consensus study

Lucy Chipchase^{a,*}, Siobhan Schabrun^a, Leonardo Cohen^b, Paul Hodges^a, Michael Ridding^c, John Rothwell^d, Janet Taylor^e, and Ulf Ziemann^f

^aThe University of Queensland, NHMRC Centre of Clinical Research Excellence in Spinal Pain, Injury and Health, Australia

^bHuman Cortical Physiology and Neurorehabilitation Section, National Institute of Neurological Disorders and Stroke, NIH, Bethesda, USA

^cNeuromotor Plasticity and Development Group, The Robinson Institute, School of Paediatrics and Reproductive Health, The University of Adelaide, Australia

^dUCL Institute of Neurology, London, UK

^eNeuroscience Research Australia and the University of New South Wales, Australia

^fDepartment of Neurology, Hertie Institute for Clinical Brain Research, University of Tübingen, Germany

Abstract

In the last decade transcranial magnetic stimulation (TMS) has been the subject of more than 20,000 original research articles. Despite this popularity, TMS responses are known to be highly variable and this variability can impact on interpretation of research findings. There are no guidelines regarding the factors that should be reported and/or controlled in TMS studies. This study aimed to develop a checklist to be recommended to evaluate the methodology and reporting of studies that use single or paired pulse TMS to study the motor system. A two round international web-based Delphi study was conducted. Panellists rated the importance of a number of subject, methodological and analytical factors to be reported and/or controlled in studies that use single or paired pulse TMS to study the motor system. Twenty-seven items for single pulse studies and 30 items for paired pulse studies were included in the final checklist. Eight items related to subjects (e.g. age, gender), 21 to methodology (e.g. coil type, stimulus intensity) and two to analysis (e.g. size of the unconditioned motor evoked potential). The checklist is recommended for inclusion when submitting manuscripts for publication to ensure transparency of reporting and could also be used to critically appraise previously published work. It is envisaged that factors could be added and deleted from the checklist on the basis of future research. Use of the TMS methodological checklist should improve the quality of data collection and reporting in TMS studies of the motor system.

*Corresponding author. Address: NHMRC Centre of Clinical Research Excellence in Spinal Pain, Injury and Health, The University of Queensland, Brisbane, Queensland 4072, Australia. Tel.: +61 7 3365 4507; fax: +61 7 3365 6122. ; Email: l.chipchase@uq.edu.au (L. Chipchase)

Keywords

Transcranial magnetic stimulation; Motor system; Consensus; Checklist; Reporting; Expert

1. Introduction

Over the last 20 years, single and paired pulse transcranial magnetic stimulation (TMS) has been used extensively to study the neurophysiological bases of excitability and plasticity in different cortical regions (Rossini and Rossi, 2007; Edwards et al., 2008; Horvath et al., 2011; Groppa et al., 2012). Most experimental work using TMS has been performed on the motor system where stimulation of the motor cortex produces a visible muscle response known as a motor evoked potential (MEP), when recorded as electromyographic activity (Rothwell, 2011).

As a non-invasive brain stimulation tool with few safety risks, applications are rapidly diversifying such that the responses to TMS are being used increasingly as an outcome measure in clinical trials (Ferrerri et al., 2011; Rijntjes et al., 2011) as a predictor of recovery (Manganotti et al., 2012) and as a diagnostic tool (Chen et al., 2008; Edwards et al., 2008). However, variability in MEP responses to TMS has been noted. Wassermann (2002), in a review of a large data set, estimated that 40–50% of the variation in MEP thresholds could be the result of subject variability (between and within) and experimental differences. Such high variability is likely to impact on the reliability and interpretation of research findings.

There is evidence, albeit conflicting, that between subject factors, such as age (Pitcher et al., 2003; Oliviero et al., 2006; Smith et al., 2011), gender (Wassermann, 2002; Inghilleri et al., 2004) and genotype (Cheeran et al., 2008; Voti et al., 2011) and within subject factors, such as caffeine use (Cerqueira et al., 2006) and time of day (Sale et al., 2007) may influence the MEP response to TMS. Similarly, evidence is building for the impact of different experimental techniques on MEP responses including the use of TMS pulses of different waveforms (Sommer et al., 2006), the orientation of the current induced by TMS in the motor cortex (Hill et al., 2000), coil positioning (Conforto et al., 2004), stimulation (trial) frequency, and, in the resting state, sub-threshold activation of corticospinal outputs (Wassermann, 2002). Given the variability in TMS responses and the potential for methodological and physiological differences to influence TMS responses, there is a clear need to report and control as many of these factors as possible.

To ensure rigorous research findings and facilitate comparison of data from separate studies, factors that impact on MEP responses should be excluded or controlled as a covariate in multivariate analyses (Wunsch, 2007). Yet, studies using TMS have small sample sizes that preclude covariate analysis. A large body of work investigating a broad range of factors is necessary before all those that impact on TMS responses are identified. In the absence of large datasets, the traditional approach to ensure findings from studies are interpreted and presented correctly is by peer review and editorial decisions, yet the quality of these processes is not guaranteed (Cobo et al., 2011). To assist peer review and critical appraisal of study methodology, reporting checklists and guidelines are being used with increasing frequency. This not only improves the transparency of research reporting and the quality of

data collection but also reduces design and reporting biases. These reporting checklists specify “a minimum set of items required for a clear and transparent account of what was done and what was found in a research study, reflecting in particular, issues that might introduce bias into the research” (EQUATOR Network, 2011).

Checklists are available for specific research methodologies including randomised controlled trials (Moher et al., 2010), observational studies (Vandenbroucke et al., 2009), systematic reviews and meta-analyses (Moher et al., 1999) and on-line surveys (Eysenbach, 2004). The most well known is the CONSORT statement which was developed to provide an evidence-based minimum set of recommendations to prepare reports of findings from randomised controlled trials in order to facilitate complete and transparent reporting while aiding critical appraisal and interpretation (Moher et al., 2010). Given the potential for variability in the parameters of the MEP responses elicited by TMS and the growth in research applications for TMS, guidelines on factors that should be reported and/or controlled in single or paired pulse TMS studies of the motor system are essential to ensure research findings are correctly interpreted.

2. Methods

An observational approach using an on-line Delphi technique was employed. The Delphi technique is a widely used and accepted method for achieving consensus of opinion on a topic area solicited from experts within a field (Keeney et al., 2001; Hsu and Sandford, 2007; Vernon, 2009). Delphi methodology uses a series of sequential questionnaires (commonly referred to as rounds) interspersed by controlled feedback to gain the most reliable consensus of opinion (Thompson, 2009; Vernon, 2009). Controlled feedback is provided by presenting summaries of the data from each round to participants with the process continuing until group consensus is achieved (Hasson et al., 2000).

Compared to group meetings, the Delphi technique has a number of advantages that make it a valuable and valid method for developing consensus. First, participant anonymity is assured which minimizes the possibility that a dominant group member or group pressure for conformity may influence the outcome (Hasson et al., 2000; Hsu and Sandford, 2007). Second, the technique allows involvement of participants worldwide through the use of web-based and e-mail communication (Hsu and Sandford, 2007). Third, the ability to use statistical analysis techniques allows for objective and impartial analysis and summarisation of collected data (Hsu and Sandford, 2007). For these reasons, the Delphi method has been commonly used for the development of consensus-based checklists (Moher et al., 1999; Cook et al., 2010; Mookink et al., 2010). Approval for this study was obtained from the Human Research Ethics Committee of The University of Queensland.

2.1. Steering committee

The role of the steering committee, comprising all eight authors, was to determine the aim of the instrument, generate items and select participants for the Delphi survey. A short email questionnaire was sent to each member of the steering committee. The questionnaire included an open-ended question that asked each member to list the methodological issues that they considered were likely to influence the MEP responses elicited by TMS when used

to study the motor system and that would be important to include in a checklist. A list of items was collated from the responses of the steering committee. Repeated responses were merged. The items were thematically categorised into three major themes: subject factors, methodological factors and analytical factors. The steering committee was asked to nominate experts in the field of TMS to study of the motor system. Names were cross-checked against major research databases (“Web of Knowledge” and “Pub- Med”) to ensure each nominee had published in the field. The final expert panel included 78 potential participants that were located internationally in Oceania, Asia, Europe and North America.

2.2. Questionnaire design

The aim of the questionnaire was to gain consensus among TMS experts regarding the essential criteria for a checklist that would enable critical appraisal and systematic reporting of TMS studies of the motor system. Studies were operationally defined as ‘studies that use single or paired pulse TMS for diagnostic/mechanistic neurophysiological research of the motor system’. Use of TMS in repetitive mode and for therapeutic intervention was not included. Following agreement on the questionnaire format by the steering committee, the web-based questionnaire was pilot tested by five respondents prior to the commencement of the study. These participants were representative of the overall sample. Piloting determined whether the instructions to participants were clear. Following piloting, minor changes to the wording of items were made.

2.3. Procedure

The web-based questionnaire was distributed to the expert panellists who were asked to participate in a maximum of three iterations of the web-based questionnaire. For the first round, participants were sent a letter of invitation along with a link to the online questionnaire. A reminder was emailed if a response was not received after a further 2 weeks. The same approach and timeline were used for the subsequent rounds.

Round 1 of the questionnaire included a section on demographics of the panel members. A second section requested that participants indicate their opinion on the importance of controlling a range of subjective (13 items), methodological (24 items) and analytical (2 items) factors when conducting TMS (single and paired pulse) studies of the motor system. A five-point Likert scale was provided with responses ‘unsure’, ‘not important’, ‘somewhat important’, ‘important’ and ‘very important’ in ascending order. Control, in this instance, meant that the factor was controlled by methodological consistency (e.g. using the same equipment), research design (e.g. randomisation or matching) or by statistical analysis (e.g. as a covariate factor). In addition, participants were asked their opinion regarding whether the factor should be reported in journal publications. Reporting meant that the item was adequately and clearly described. Again, a five point Likert scale was used with the following options: ‘unsure’, ‘never’, ‘some of the time’, ‘most of the time’ and ‘always’, again in ascending order. For each group of items, an ‘other, please specify’ question was included in the event that participants considered factors not present in the provided list should be included.

Two to three rounds are generally sufficient to collect the information and to reach consensus (Hasson et al., 2000). Once consensus is reached, further rounds are not needed. It was envisaged that three rounds would be sufficient to reach consensus but if consensus was reached after two rounds, the process could be terminated. Participants were to be provided feedback in the form of group responses for each factor from the previous round. In light of this information, participants were to be requested to re-score each item on the same Likert scale used in round 1. New items added to the questionnaire based on responses to the 'other, please specify' question from the first round, were highlighted.

2.4. Data analysis

All data were entered into Microsoft Excel and Statistical Package for the Social Sciences (SPSS, Version 17). The responses for each Delphi round were reported as the mean of the five-point Likert scale. The steering group discussed, via email, participants' qualitative and quantitative answers after each round. Based on these discussions, items with an insufficient consensus rate were excluded and items proposed by participants were added, modified or expanded.

After the first round, factors were only retained if 60% of the sample indicated that it was 'important' or 'very important' to control the factor or that the factor should be reported 'most of the time' or 'always'. This decision was made by the steering committee after viewing the results from the first round. Following this step, fountain graphs from responses to the first and second round were generated. A fountain graph plots the mean of the Likert scale of each item against the standard deviation to demonstrate the extent of the group's opinion and the level of agreement for all of the items in the Delphi study (Greatorex and Dexter, 2000). Comparisons of fountain graphs for each round provided a statistical story of how the panel's opinion changed and the amount of change between rounds.

Interquartile deviation (IQD) scores for responses to each factor in the second round were calculated. The interquartile range is the absolute value of the difference between the 75th and 25th percentiles, with smaller values indicating higher degrees of consensus (Rayens and Hahn, 2000). An IQD of 1.00 or less has been identified as an indicator of consensus (Raskin, 1994). In addition, the percentage agreement in excess of 60% of generally positive respondents was used. Consensus to include a factor on the checklist was defined by an IQD 1.0 and 60% positive responses of the panellists.

Additionally, Cronbach's coefficient alpha was used as an indicator of consistency of the panellists' responses in the second round of the questionnaire. Where the responses of the panellists are highly correlated, they are considered to be internally consistent or homogeneous (Graham et al., 2003). Previous work has used an alpha of 0.80 as an indicator that consensus has been reached (Graham et al., 2003; Palter et al., 2011).

3. Results

From the 78 experts invited to participate in the first round, 42 (53.8%) responses were received. The characteristics of the expert panel are presented in Table 1. The expert panel had a mean (SD) of 13.3 (6.5) years experience using TMS (single or paired pulse) to study

the motor system. In total, the panel had 520 years of experience. Seventy-four per cent of participants held PhDs with over 60% working in universities in various research roles. From the 42 respondents from round 1, 39 responses (92.9%) were received following round 2.

In the first round, there were 13 subject factors, 24 methodological factors and two analysis factors as well as an 'other, please specify' question for each of these areas. Using the criterion of a 60% cut-off for panellists' positive responses, seven subject factors, three methodological factors and one analysis factor were excluded (Table 2).

New material was refined and added to the questionnaire following suggestions made by respondents to the 'other, please specify' question. Additional subject factors were 'medical conditions that may alter responses to TMS e.g. thyroid dysfunction, diabetes, limb injury' and 'history of specific repetitive motor usage e.g. proficiency playing a musical instrument, intensive typing'. The additional analysis factor was 'size of unconditioned MEP'. Thus, in the second round, there were eight subject factors, 21 methodological factors and two analysis factors.

The mean and standard deviation for each of the 39 factors in round one and each of the 31 factors in round two are displayed in Fig. 1A and B, respectively. These fountain graphs allow the overall picture and convergence of opinion to be appreciated (Greatest and Dexter, 2000). They depict the convergence of 'important' and 'very important' or 'most of the time' and 'always' responses between the two rounds. The results in round two show convergence of opinion for the factors included in this questionnaire with smaller SDs and a greater number of means between 4 and 5.

In order to measure consensus among respondents following the second round questionnaire, the inter-quartile deviation (IQD) and percentage of positive responses were calculated for each factor (Table 3). All factors had an IQD of one or less indicating consensus was reached. Two factors were identified as not important according to the 60% cut off for positive responses. These were controlling 'gender' (46% positive responses) and reporting 'relaxation of other muscles' (51% positive responses). Therefore, these were excluded from the final checklist.

For the analysis of Cronbach's alpha, respondents with missing responses for one or more factors were excluded ($n = 7$). The analysis of the 31 factors in round two resulted in a Cronbach's alpha of 0.88. This indicates that consensus was reached. The final list of items included in the checklist is presented in Table 4. There are 27 items for single pulse studies and 30 items for paired pulse studies.

4. Discussion

The goal of this study was to design a checklist to facilitate evaluation of methodology and reporting of studies that use single or paired pulse TMS to evaluate the motor system. The final TMS methodological checklist was developed through an on-line survey using the Delphi technique and consists of 27 items for single pulse studies and 30 items for paired pulse studies. Eight items relate to subjects (e.g. age, gender), 21 to methodology (e.g. coil

type, stimulus intensity) and two to analysis (e.g. unconditioned MEP size). The final checklist has two primary uses. First, researchers could use the checklist to ensure that the necessary subject, methodological and analytical information is considered in their research design. It is recommended that this checklist be included when submitting manuscripts for publication to ensure transparency of reporting. Inclusion of this data would assist editors, peer reviewers and readers in the evaluation and interpretation of the study findings. This is particularly important given the potential for variability in TMS responses to impact on the interpretation of research findings. Second, the checklist is likely to aid critical appraisal of studies particularly when reviewing multiple papers that address a similar research question. This may be of value when conducting critical reviews or meta-analyses. Depending on the study design, items from the checklist that should be reported and/or controlled can be used to determine whether the results of a study were potentially confounded by variables that may have impacted TMS measures. It is important to note that the specific combination of factors that should be reported or controlled for a specific study may depend on the hypotheses and goals. Thus, issues that are crucial for one investigation may be less important for others. Thus the hypotheses and goals of a study need to be considered when using the TMS methodological checklist to critically appraise its methodology.

The consensus findings reported by the expert panel generally concur with current research evidence. For example, most data indicate no differences in MEP response characteristics between genders (Wassermann, 2002; Pitcher et al., 2003; Livingston et al., 2010). As noted by the expert panel, gender is important to report but not to control. Similarly, published studies indicate that age influences cortical excitability and the age of subjects was considered both important to report and control, depending on the study methodology (Pitcher et al., 2003; Smith et al., 2011). There is also good evidence that methodological factors included in the final checklist such as the TMS pulse waveform (Sommer et al., 2006), the orientation of the current induced by TMS (Hill et al., 2000), coil positioning (Conforto et al., 2004), stability of coil location (Ahdab et al., 2010) and EMG electrode placement (Rossini et al., 1999) are important contributors to TMS variability.

A number of factors were excluded from the first round (Table 1) despite the existence of some evidence that they may impact on MEP responses to TMS. For example, caffeine intake, genetics (siblings) and hormonal/menstrual cycle have been demonstrated to affect MEP responses (Wassermann, 2002; Inghilleri et al., 2004; Cerqueira et al., 2006). This suggests uncertainty about the significance of the findings in these areas and there may be a need for further research. Alternatively, the response may reflect the personal experimental evidence of the panel members, or perhaps the logistics of being able to control these factors. The conflict between the opinions of the expert panel and available research findings highlights the dynamic nature of research in this field and the importance of updating the tool as new data emerge. Until then, the most considered approach has been reach consensus opinion until hard evidence is available to support or refute these views.

The expert panel consisted of international researchers and academics that each had, on average, 13.3 years experience using TMS. The final TMS methodological checklist represents the consensus opinion of 39 experts in TMS throughout Asia, Oceania, Europe and North America. In addition, we used several analytical methods (inter-quartile deviation,

percentage agreement scores and fountain graphs) to determine convergence of agreement between rounds and when consensus was reached. Cronbach's alpha, which was 0.88 for the second round, should be considered substantial and is consistent with reliability scores obtained for validated scales in clinical use (Graham et al., 2003). Our methodical and analytical approach, as well as the solicitation of international expert opinion, ensures the developed checklist has content validity.

Recommendations for the size of Delphi expert panels range from 10 to 50. Our sample size of 42 for the first round (response rate 53.8%) and 39 in round two (92.9%) represents an appropriate sample size with minimal attrition between rounds. The high response rate likely reflects the considerable interest in the topic and the loss of only three participants between rounds minimizes attrition bias. The use of initial steering committee for the Delphi study ensured that a preliminary list of methodological issues was developed along with an international sample of potential expert panel members. While responses to each Delphi round were anonymous, there is the slight possibility that the views of the steering committee may have propagated through the Delphi rounds. This is unlikely to have affected responses to closed questions but may have affected item generation from the categorisation of responses to open questions.

In summary, the Delphi methodology allowed the development of a consensus-based checklist by international experts in TMS. The checklist could be used to evaluate the methodology and reporting of studies that use single or paired pulse TMS to study the motor system. We recommend that the checklist be used at the time of designing a single or paired pulse TMS study of the motor system. In addition, information on compliance with the checklist could be provided in the methods section of scientific manuscripts of single and paired pulse TMS studies to provide transparent control and reporting of subject, methodological and analytical factors that may impact on responses to TMS. The items developed in this study represent a core set of subject, methodological and analytical factors that should be reported in every TMS study (single and paired pulse) of the motor system. Thus, when submitting a manuscript, authors may explicitly highlight, within the methodology, which items have been reported, and where necessary, which items have been controlled. Compliance with the TMS methodology checklist could be reported by indicating that items on the checklist that pertain, for example to single pulse TMS, have been reported or controlled. It is however important to stress that additional factors (not mentioned on the checklist) might have to be considered as essential information dependent on the specific aims and settings of a given study.

Further refinements will need to be made to the checklist over time as new evidence emerges about factors that influence responses to TMS. For example, the temporal order of different single pulse (e.g. stimulus intensity curve) and paired pulse conditions (e.g. interstimulus interval curve) was not raised by the expert panel nor was the use of the triple pulse technique. Furthermore, specificity on some methodological factors is required. For example, there are different methods for measuring the motor threshold such as the relative frequency method, adaptive methods, the two-threshold methods and the supervised parametric estimation (Groppa et al., 2012). As our understanding of whether these factors

impact on the variability of TMS responses improves, factors such as these may be added to the checklist.

Acknowledgments

We are grateful to all panel members who participated in both rounds of this Delphi consensus study: Abbruzzese G., Byblow W., Cantello R., Celnik P., Chen R., Chipchase L., Classen J., Di Lazzaro V., Dimyan M., Epstein C., Floel A., Galea M., Golaszewski S., Hamada M., Khedr E., Lefaucheu J., Macefield V., Mall V., McDonnell M., Orth M., Paulus W., Perez M., Pitcher J., Quartarone A., Rossi S., Rothkegel H., Rothwell J., Sale M., Sandrini M., Schabrun S., Schambra H., Semmler J., Taylor J., Thickbroom G., Tsao H., Ugawa Y., Wittenberg G., Ziemann U., Zijdwind I.

References

- Ahdab R, Ayache SS, Brugieres P, Goujon C, Lefaucheur JP. Comparison of “standard” and “navigated” procedures of TMS coil positioning over motor, premotor and prefrontal targets in patients with chronic pain and depression. *Clin Neurophysiol.* 2010; 40:27–36.
- Cerqueira V, de Mendonça A, Minez A, Dias AR, de Carvalho M. Does caffeine modify corticomotor excitability? *Clin Neurophysiol.* 2006; 36:219–26.
- Cheeran B, Talelli P, Mori F, Koch G, Suppa A, Edwards M, et al. A common polymorphism in the brain-derived neurotrophic factor gene (BDNF) modulates human cortical plasticity and the response to rTMS. *J Physiol.* 2008; 586:5717–25. [PubMed: 18845611]
- Chen R, Cros D, Curra A, Di Lazzaro V, Lefaucheur JP, Magistris MR, et al. The clinical diagnostic utility of transcranial magnetic stimulation: report of an IFCN committee. *Clin Neurophysiol.* 2008; 119:504–32. [PubMed: 18063409]
- Cobo E, Cortes J, Ribera JM, Cardellack F, Selva-OCallaghan A, Kostov B, et al. Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial. *BMJ.* 2011; 343:1–11.
- Conforto AB, Z'Graggen WJ, Kohl AS, Rosler KM, Kaelin-Lang A. Impact of coil position and electrophysiological monitoring on determination of motor thresholds to transcranial magnetic stimulation. *Clin Neurophysiol.* 2004; 115:812–9. [PubMed: 15003761]
- Cook C, Brismée JM, Pietrobon R, Sizer P Jr, Hegedus E, Riddle DL. Development of a quality checklist using Delphi methods for prescriptive clinical prediction rules: the QUADCPR. *J Manipulative Physiol Ther.* 2010; 33:29–41. [PubMed: 20114098]
- Edwards MJ, Talelli P, Rothwell JC. Clinical applications of transcranial magnetic stimulation in patients with movement disorders. *Lancet Neurol.* 2008; 7:827–40. [PubMed: 18703005]
- EQUATOR Network. Library for health research reporting. 2011. <www.equator-network.org/resource-centre/library-of-health-research-reporting>
- Eysenbach G. Improving the quality of web surveys: the checklist for reporting results of Internet e-surveys (CHERRIES). *J Med Internet Res.* 2004; 6:12–6.
- Ferreri F, Pasqualetti P, Maatta S, Ponzio D, Guerra A, Bressi F, et al. Motor cortex excitability in Alzheimer's disease: a transcranial magnetic stimulation followup study. *Neurosci Lett.* 2011; 492:94–8. [PubMed: 21281700]
- Graham B, Regehr G, Wright JG. Delphi as a method to establish consensus for diagnostic criteria. *J Clin Epidemiol.* 2003; 56:1150–6. [PubMed: 14680664]
- Greatorex J, Dexter T. An accessible analytical approach for investigating what happens between the rounds of a Delphi study. *J Adv Nurs.* 2000; 32:1016–24. [PubMed: 11095243]
- Groppa S, Oliviero A, Eisen A, Quartarone A, Cohen LG, Mall V, et al. A practical guide to diagnostic transcranial magnetic stimulation: report of an IFCN committee. *Clin Neurophysiol.* 2012; 123:858–82. [PubMed: 22349304]
- Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs.* 2000; 32:1008–15. [PubMed: 11095242]
- Hill AC, Davey NJ, Kennard C. Current orientation induced by magnetic stimulation influences a cognitive task. *Neuroreport.* 2000; 11:3257–9. [PubMed: 11043559]

- Horvath JC, Perez JM, Forrow L, Fregni F, Pascual-Leone A. Transcranial magnetic stimulation: a historical evaluation and future prognosis of therapeutically relevant ethical concerns. *J Med Ethics*. 2011; 37:137–43. [PubMed: 21106996]
- Hsu C-H, Sandford BA. The Delphi technique – making sense of consensus. *Pract Assess Res Eval*. 2007; 12:1–8.
- Inghilleri M, Conte A, Curra A, Frasca V, Lorenzano C, Berardelli A. Ovarian hormones and cortical excitability. An rTMS study in humans. *Clin Neurophysiol*. 2004; 115:1063–8. [PubMed: 15066531]
- Keeney S, Hasson F, McKenna HP. A critical review of the Delphi technique as a research methodology for nursing. *Int J Nurs Stud*. 2001; 38:195–200. [PubMed: 11223060]
- Livingston S, Goodkin H, Ingersoll C. The influence of gender, hand dominance, and upper extremity length on motor evoked potentials. *J Clin Monit Comput*. 2010; 24:427–36. [PubMed: 21110222]
- Manganotti P, Storti SF, Formaggio E, Acler M, Zoccatelli G, Pizzini FB, et al. Effect of median-nerve electrical stimulation on BOLD activity in acute ischemic stroke patients. *Clin Neurophysiol*. 2012; 123:142–53. [PubMed: 21741301]
- Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet*. 1999; 354:1896–900. [PubMed: 10584742]
- Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Clin Epidemiol*. 2010; 63:e1–e37.
- Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010; 19:539–49. [PubMed: 20169472]
- Oliviero A, Profice P, Tonali PA, Pilato F, Saturno E, Dileone M, et al. Effects of aging on motor cortex excitability. *Neurosci Res*. 2006; 55:74–7. [PubMed: 16584795]
- Palter VN, MacRae HM, Grantcharov TP. Development of an objective evaluation tool to assess technical skill in laparoscopic colorectal surgery: a Delphi methodology. *Am J Surg*. 2011; 201:251–9. [PubMed: 20832048]
- Pitcher JB, Ogston KM, Miles TS. Age and sex differences in human motor cortex input–output characteristics. *J Physiol*. 2003; 546:605–13. [PubMed: 12527746]
- Raskin M. The Delphi study in field instruction revisited: results of a Delphi study. *J Soc Work Educ*. 1994; 30:75–89.
- Rayens M, Hahn E. Building consensus using the policy Delphi method. *Policy Polit Nurs Prac*. 2000; 1:308–15.
- Rijntjes M, Hamzei F, Glauche V, Saur D, Weiller C. Activation changes in sensorimotor cortex during improvement due to CIMT in chronic stroke. *Restor Neurol Neurosci*. 2011; 29:299–310. [PubMed: 21697588]
- Rossini PM, Berardelli A, Deuschl G, Hallett M, Maertens de Noordhout AM, Paulus W, et al. Applications of magnetic cortical stimulation. The International Federation of Clinical Neurophysiology. *Electroencephalog Clin Neurophysiol*. 1999; (Suppl 52):171–85.
- Rossini PM, Rossi S. Transcranial magnetic stimulation. *Neurology*. 2007; 68:484–8. [PubMed: 17296913]
- Rothwell JC. Using transcranial magnetic stimulation methods to probe connectivity between motor areas of the brain. *Hum Mov Sci*. 2011; 30:906–15. [PubMed: 21056490]
- Sale MV, Ridding MC, Nordstrom MA. Factors influencing the magnitude and reproducibility of corticomotor excitability changes induced by paired associative stimulation. *Exp Brain Res*. 2007; 181:615–26. [PubMed: 17487476]
- Smith AE, Sale MV, Higgins RD, Wittert GA, Pitcher JB. Male human motor cortex stimulus-response characteristics are not altered by aging. *J Appl Physiol*. 2011; 110:206–12. [PubMed: 21071590]
- Sommer M, Alfaro A, Rummel M, Speck S, Lang N, Tings T, et al. Half sine, monophasic and biphasic transcranial magnetic stimulation of the human motor cortex. *Clin Neurophysiol*. 2006; 117:838–44. [PubMed: 16495145]

- Thompson M. Considering the implication of variations within Delphi research. *Fam Pract.* 2009; 26:420–4. [PubMed: 19684122]
- Vandenbroucke JP, Von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Gac Sanit.* 2009; 23:158. [PubMed: 19249134]
- Vernon W. The Delphi technique: a review. *Int J Ther Rehabil.* 2009; 16:69–76.
- Voti PL, Conte A, Suppa A, Iezzi E, Bologna M, Aniello MS, et al. Correlation between cortical plasticity, motor learning and BDNF genotype in healthy subjects. *Exp Brain Res.* 2011; 212:91–9. [PubMed: 21537966]
- Wassermann EM. Variation in the response to transcranial magnetic brain stimulation in the general population. *Clin Neurophysiol.* 2002; 113:1165–71. [PubMed: 12088713]
- Wunsch G. Confounding and control. *Demogr Res.* 2007; 16:97–120.

HIGHLIGHTS

- A TMS methodological checklist was developed for the reporting and interpreting of TMS studies of the motor system.
- An international expert panel participated in a Delphi study to establish consensus on items in the TMS methodological checklist.
- Use of the TMS methodological checklist should improve the quality of data collection and reporting in TMS studies of the motor system.

Fig. 1.

Mean and standard deviation (SD) for each question from round one (1A) and round two (1B). A mean of 5 represents a mean response of ‘very important’ or ‘always’ on a 5-point Likert scale. A greater number of mean responses between 4 and 5 in round two (1B) indicate greater support for the importance of the factors provided in this questionnaire. The smaller SDs following round two (1B) indicates a convergence of opinion and greater consensus of opinion in round 2.

Table 1

Profile of the expert panel participating in round 1.

Characteristic	N (%)
Male	32 (76.2)
<i>Discipline background</i>	
Physiology	7 (16.7)
Medicine	23 (54.8)
Physiotherapy	5 (11.9)
Neuroscience	5 (11.9)
Other	2 (4.8)
<i>Highest qualification</i>	
Diploma	1 (2.4)
Bachelor degree	1 (2.4)
Masters degree (coursework)	1 (2.4)
PhD	31 (73.8)
MD	5 (11.9)
Other	3 (7.1)
<i>Place of work</i>	
University	29 (69.0)
Research institute	7 (16.7)
Hospital	5 (11.9)
Other	1 (2.4)
<i>Current role</i>	
Post doctoral researcher	7 (16.7)
Academic researcher e.g. professor	27 (64.3)
Clinical researcher	6 (14.3)
Other	2 (4.8)

Table 2

Factors excluded after round 1.

Section	Factor
Participant factors	Subject genetics e.g. siblings
	Subject ethnicity
	Hormonal/menstrual cycle of female subjects
	EEG differences between subjects
	Nicotine intake on the day of testing
	Caffeine intake on the day of testing
	Recent exercise history (prior 1 h)
Methodological factors	Number of stimuli given when determining the hotspot
	Number of stimuli given when determining threshold
	Time of day tested
Analysis factors	Expressing the MEP as a percentage of the M wave

Table 3

IQD and percentage of positive respondents for each factor from the round 2.

	Control of factor IQD	'Important' and 'very important' to control % respondents	Reporting of factor IQD	Factor should be reported 'always' and 'most of the time' % respondents
<i>Subject factors</i>				
Age of subjects	0.5	92	0	100
Gender of subjects	0.5	46	0	97
Handedness	0.25	77	0	97
Prescribed medication	0.5	98	0.5	87
Use of CNS active drugs (e.g. anti-convulsants)	0	100	0	97
Presence of neurological/psychiatric disorders when studying healthy subjects	0	100	0	100
Medical condition	0.5	95	0.5	90
History of specific repetitive motor activity	0.5	64	1	62
<i>Methodological factors</i>				
Position and contact of EMG electrodes	0.5	95	0.5	90
Amount of relaxation/contraction of target muscles	0	100	0	97
Prior motor activity of the muscle to be tested	0	95	0.5	68
Level of relaxation of muscles other than those being tested	0	80	0.5	51
Coil type (size and geometry)	0	100	0.0	100
Coil orientation	0	100	0	100
Direction of induced current in the brain	0	100	0	98
Coil location and stability (with a neuronavigation system)	0	79	0.5	82
Coil location and stability (without a neuronavigation system)	0.5	95	0.5	92
Type of stimulator used (e.g. brand)	0.5	87	0.5	95
Stimulation intensity	0	100	0	97
Pulse shape (monophasic or biphasic)	0	97	0	98
Determination of optimal hotspot	0.5	95	0.5	97
The time between MEP trials	0.5	82	0.5	87
Time between days of testing	0.5	61	0.5	79
Subject attention (level of arousal) during testing	0	92	0.5	82
Method for determining threshold (active/resting)	0	100	0	100
Number of MEP measures made	0.5	90	0	95
<i>Paired pulse only: intensity of test pulse</i>	0	100	0	97
<i>Paired pulse only: intensity of conditioning pulse</i>	0	100	0	97
<i>Paired pulse only: inter-stimulus interval</i>	0	100	0	97
<i>Analytical factors</i>				
Method for determining MEP size during analysis			0	100
Size of unconditioned MEP			0.5	90

Table 4

Final checklist (N/A = Not applicable).

Were the following participant factors	Reported?	Controlled?
Age of subjects	☐	☐
Gender of subjects	☐	N/A
Handedness of subjects	☐	☐
Subjects prescribed medication	☐	☐
Use of CNS active drugs (e.g. anti-convulsants)	☐	☐
Presence of neurological/psychiatric disorders when studying healthy subjects	☐	☐
Any medical conditions	☐	☐
History of specific repetitive motor activity	☐	☐
<i>Were the following methodological factors</i>		
Position and contact of EMG electrodes	☐	☐
Amount of relaxation/contraction of target muscles	☐	☐
Prior motor activity of the muscle to be tested	☐	☐
Level of relaxation of muscles other than those being tested	N/A	☐
Coil type (size and geometry)	☐	☐
Coil orientation	☐	☐
Direction of induced current in the brain	☐	☐
Coil location and stability (with or without a neuronavigation system)	☐	☐
Type of stimulator used (e.g. brand)	☐	☐
Stimulation intensity	☐	☐
Pulse shape (monophasic or biphasic)	☐	☐
Determination of optimal hotspot	☐	☐
The time between MEP trials	☐	☐
Time between days of testing	☐	☐
Subject attention (level of arousal) during testing	☐	☐
Method for determining threshold (active/resting)	☐	☐
Number of MEP measures made	☐	☐
<i>Paired pulse only: Intensity of test pulse</i>	☐	☐
<i>Paired pulse only: Intensity of conditioning pulse</i>	☐	☐
<i>Paired pulse only: Inter-stimulus interval</i>	☐	☐
<i>Were the following analytical factors</i>		
Method for determining MEP size during analysis	☐	☐
Size of unconditioned MEP	☐	☐