Reliability of Motor Evoked Potentials Induced by Transcranial Magnetic Stimulation: The Effects of Initial Motor Evoked Potentials Removal

Fahimeh Hashemirad1*, Maryam Zoghi2, Paul B Fitzgerald3, Shapour Jaberzadeh1

1. Department of Physiotherapy, School of Primary Health Care, Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia.
2. Department of Medicine at Royal Melbourne Hospital, University of Melbourne, Melbourne, Australia.
3. Monash Alfred Psychiatry Research Centre, Alfred and Monash University Central Clinical School, Melbourne, Australia.

ABSTRACT

Introduction: Transcranial magnetic stimulation (TMS) is a useful tool for assessment of corticospinal excitability (CSE) changes in both healthy individuals and patients with brain disorders. The usefulness of TMS-elicited motor evoked potentials (MEPs) for the assessment of CSE in a clinical context depends on their intra-and inter-session reliability. This study aimed to evaluate if removal of initial MEPs elicited by using two types of TMS techniques influences the reliability scores and whether this effect is different in blocks with variable number of MEPs.

Methods: Twenty-three healthy participants were recruited in this study. The stimulus intensity was set at 120% of resting motor threshold (RMT) for one group while the stimulus intensity was adjusted to record MEPs up to 1 mV for the other group. Twenty MEPs were recorded at 3 time points on 2 separate days. An intra-class correlation coefficient (ICC) reliability with absolute agreement and analysis of variance model were used to assess reliability of the MEP amplitudes for blocks with variable number of MEPs.

Results: A decrease in ICC values was observed with removal of 3 or 5 MEPs in both techniques when compared to all MEP responses in any given block. Therefore, removal of the first 3 or 5 MEPs failed to further increase the reliability of MEP responses.

Conclusion: Our findings revealed that a greater number of trials involving averaged MEPs can influence TMS reliability more than removal of the first trials.

1. Introduction

Transcranial magnetic stimulation (TMS) is a useful tool for assessment of corticospinal excitability (CSE) changes in both healthy individuals and patients with brain disorders (Barker et al., 1987; Rossini et al., 1994; Liepert et al., 2000). The magnetic pulses induced by TMS over the contralateral primary motor cortex (M1) can pass through the scalp and induce a response known as “motor evoked potential” (MEP) in the target muscle. This response is recorded using surface electromyography...
phy (EMG) electrodes placed over the muscle of interest (Malcolm et al., 2006). The peak-to-peak amplitude of the elicited MEPs is an indication of changing CSE. Smaller amplitudes indicate lower excitability, while larger amplitudes suggest higher CSE (Chipchase et al., 2012).

Literature review indicates that there is a high degree of variability in the TMS-induced resting MEPs (Kiers et al., 1993; Ellaway et al., 1998). This variability could result from technical factors such as orientation, location, and stability of the TMS coil (Barker et al., 1987; Hill et al., 2000; Chipchase et al., 2012). However, variability in MEP responses remains even after controlling these factors. This inherent variability could result from neurophysiological changes in the CSE pathway (Truccolo et al., 2002). More variability might be expected in the amplitude of the first few MEPs due to changes in regional cerebral flow (Mochizuki et al., 2006) and changes in excitatory synaptic drive to corticospinal neurons (Ellaway et al., 1998).

The first few MEP responses might be larger than the subsequent MEPs (Brasil-Neto et al., 1994), and the increased variability in initial MEPs can affect TMS reliability (Schmidt et al., 2009). Therefore, removal of the first, more fluctuating MEPs might increase the averaged reliability scores. In TMS studies, CSE could be assessed using two different techniques. In the first technique, the test stimulus is calculated as a ratio of a resting motor threshold (RMT) such as 120% RMT. In the second technique, the test stimulus is adjusted to produce MEP responses up to 1 mV, which is commonly used in paired-pulse TMS studies. Since there is an inverse relationship between variability of MEP responses and TMS stimulus intensity (Kiers et al., 1993), the MEPs evoked by the 1 mV technique are less subject to variability, which may be less affected by more variable and fluctuating initial MEPs.

The literature suggests that increasing the number of evoked MEPs increases the TMS reliability (Ellaway et al., 1998; Truccolo et al., 2002; Kamen, 2004; Bastani & Jaberzadeh, 2012). Little is known about how removal of the first few MEPs affects the reliability scores of TMS techniques. In this study, we investigate the effects of removal of the initial elicited MEPs on reliability scores, and also whether this effect is different in blocks with different MEP numbers. We hypothesised that removal of three or five initial MEPs should increase reliability. We also hypothesised that the removal of the initial MEPs should have more profound effects on enhancement of reliability than the number of MEPs in each block.

2. Methods

2.1. Participants

Twenty three healthy participants were recruited in this study and divided into two groups to assess the reliability of MEPs responses induced by two types of TMS techniques. Thirteen participants (11 females and two males with the mean [SD] age of 26.5[9.9] y) were included in one group where the test stimulus was considered at 120% RMT. In the other group (11 females and two males, with the mean [SD] age of 24[3.7] y), the test stimulus was adjusted at 1 mV. Handedness of the participants was assessed using the Edinburgh Handedness Questionnaire (Oldfield, 1971). The dominant hand was tested in each participant. Of 23 participants, 21 were right-hand dominant. Participants were screened for contraindication to TMS applications. They provided their written informed consent prior to the experiments. All protocols used were approved by the Human Research Ethics Committees at Monash University and conformed to the Declaration of Helsinki.

2.2. Measurements

2.2.1. Electromyography

Participants were tested in a sitting position with forearm supported in a pronated position. A standard skin preparation (Gilmore & Meyers, 1983) procedure was performed for each electrode placement site. EMG electrodes were placed on the first dorsal interosseous (FDI) muscle of the dominant hand with an inter-electrode distance of 2 cm. A ground electrode was placed ipsilaterally over the styloid process of the ulna bone. All EMG signals were filtered, amplified (10 Hz–500 Hz x 1000), and sampled at 1000 Hz. All data were recorded on a PC via a commercially available software (Chart™ software, ADInstrument, Australia) and a laboratory analogue-digital interface (The Power Lab 8/30, ADInstrument, Australia) for later off-line analysis.

2.2.2. Motor evoked potentials

Single pulse magnetic stimuli were delivered using two stimulators with a figure-of-eight coil. A Magstim 2002 (Magstim Company Limited, UK) stimulator was used for recording MEPs with intensity of 120% RMT in group 1, and a MagPro R30 (MagOption) stimulator (MagVenture Denmark) was used for recording MEPs using the second technique in group 2. In both groups, the coil was placed over the dominant M1, i.e. contralateral to the muscle of interest. The orientation of the coil...
was set at an angle of 45° to the midline and tangential to the scalp. In this orientation, the induced current flow is directed from posterior to anterior. The coil was moved around the M1 of the FDI muscle to determine the optimal site of stimulation. After localizing this site, known as a hot spot, the coil position was marked on the scalp as a reference. Coil position and orientation were constantly assessed throughout the experiment to minimize technical inconsistencies.

After localizing the hot spot, RMT was measured. RMT is defined as the lowest intensity to induce at least 5 MEPs larger than 50 µV in peak-to-peak amplitude out of 10 consecutive stimuli to find RMT, also the intensity of the stimulator was decreased in steps of 2% of the maximum stimulator output. The test stimulus was set at 120% of each individual’s RMTs in group 1 and adjusted up to produce MEP responses of about 1 mV in group 2.

2.3. Procedure

Each participant was tested in two separate testing sessions. The first session involved two sets of data collection. FDI muscle MEPs were recorded before and immediately after a 20-minute break in which subjects were recommended to do activities such as reading books or magazines. During each testing session, 20 MEPs with interpulse intervals of 10 seconds (Vaseghi et al., 2015) were recorded. A follow-up session was held at least 72 hours after the first session. All participants were assessed at the same time of day in both sessions to avoid diurnal variations.

2.4. Data analysis

In both groups, 20 stimuli were delivered, with 10 seconds interstimulus interval. The averaged MEPs at each time point were calculated for the first 10 (Block 1), first 15 (Block 2), and all 20 MEPs (Block 3). Then the averaged MEPs were also calculated after removal of the first 3 and the first 5 MEPs in each block. The effects of removal of the first 3 and the first 5 MEPs in each block were evaluated using intraclass correlation coefficients (ICCs) with absolute agreement and a 2-way mixed model. Repeated measures analysis of variance (ANOVA) was used to detect any differences between the averaged MEPs across 3 time points at any given block.

SPSS (version 20) was used for the data analysis. A significance level of P<0.05 was adopted for all conditions. Post hoc tests (Student t test with Bonferroni correction) were performed where indicated.

3. Results

A total of 23 individuals were recruited for this study. Three subjects took part in both groups while the rest of the subjects participated in only one group. In group 1 (n=13), stimulus intensity was delivered at 120% RMT. In group 2 (n=13), the average stimulus intensity required to produce MEPs of about 1 mV was 139% RMT (with Min and Max 104% and 185%.

<table>
<thead>
<tr>
<th>Test Intensity</th>
<th>Block 1a (1-10 MEPs)</th>
<th>Block 1b (4-10 MEPs)</th>
<th>Block 1c (6-10 MEPs)</th>
<th>Block 2a (1-15 MEPs)</th>
<th>Block 2b (4-15 MEPs)</th>
<th>Block 2c (6-15 MEPs)</th>
<th>Block 3a (1-20 MEPs)</th>
<th>Block 3b (4-20 MEPs)</th>
<th>Block 3c (6-20 MEPs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>120% RMT</td>
<td>0.78±0.47</td>
<td>0.71±0.39</td>
<td>0.71±0.43</td>
<td>0.74±0.41</td>
<td>0.69±0.35</td>
<td>0.68±0.37</td>
<td>0.72±0.40</td>
<td>0.67±0.35</td>
<td>0.67±0.38</td>
</tr>
<tr>
<td>N=13</td>
<td>T1 Session 1 (Mean±SD)</td>
<td>T2 Session 1 (Mean±SD)</td>
<td>T1 Session 2 (Mean±SD)</td>
<td>F (2, 24)</td>
<td>P</td>
<td>ICCs P</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.70±0.66</td>
<td>0.65±0.69</td>
<td>0.64±0.68</td>
<td>0.71±0.64</td>
<td>0.71±0.66</td>
<td>0.69±0.66</td>
<td>0.70±0.37</td>
<td>0.77±0.65</td>
<td>0.76±0.67</td>
<td>0.77±0.67</td>
</tr>
<tr>
<td>0.68±0.50</td>
<td>0.63±0.45</td>
<td>0.68±0.49</td>
<td>0.70±0.41</td>
<td>0.67±0.36</td>
<td>0.69±0.40</td>
<td>0.69±0.40</td>
<td>0.68±0.37</td>
<td>0.69±0.38</td>
<td>0.69±0.38</td>
</tr>
<tr>
<td>0.325</td>
<td>0.197</td>
<td>0.134</td>
<td>0.069</td>
<td>0.022</td>
<td>0.009</td>
<td>0.397</td>
<td>0.514</td>
<td>0.521</td>
<td>0.521</td>
</tr>
<tr>
<td>0.726</td>
<td>0.864</td>
<td>0.875</td>
<td>0.934</td>
<td>0.978</td>
<td>0.991</td>
<td>0.677</td>
<td>0.605</td>
<td>0.601</td>
<td>0.601</td>
</tr>
<tr>
<td>0.851</td>
<td>0.754</td>
<td>0.830</td>
<td>0.897</td>
<td>0.839</td>
<td>0.881</td>
<td>0.922</td>
<td>0.893</td>
<td>0.895</td>
<td>0.895</td>
</tr>
<tr>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Significant results are bold.
RMT). The average (SD) handedness scores were 79.4(25.2) and 86.7(9.8) in groups 1 and 2, respectively.

In Table 1, the results of the ICCs and F tests values in all blocks with different number of trials are shown for group 1. The ICC values ranged from 0.75 to 0.92 in blocks 1, 2, and 3 indicating that increasing the number of trials can lead to an increase in ICC values. The results of ICCs in all blocks with removal of the first 3 or 5 MEPs revealed slightly decreased reliability for the FDI MEP responses. More reduction in ICC values was obtained with removing the first 3 MEPs in all blocks, compared to removal of the first 5 MEPs (Table 1). There were no significant differences in the average MEP size at any time points in any given block (Table 2).

### Table 2. The results of ICCs and f tests in three blocks 10, 15 and 20 MEPs in three types of conditions (all trials, after removal of the first three or five MEPs) at three time points across the two sessions (MEPs=1 mV).

<table>
<thead>
<tr>
<th>Stimulus Intensity=MEPs 1mV N=13</th>
<th>T1 Session 1 (Mean±SD)</th>
<th>T2 Session 1 (Mean±SD)</th>
<th>T1 Session 2 (Mean±SD)</th>
<th>F (2, 24)</th>
<th>P</th>
<th>ICCs</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block 1a (1-10 MEPs)</td>
<td>1.09±0.24</td>
<td>1.12±0.35</td>
<td>1.01±0.19</td>
<td>0.791</td>
<td>0.465</td>
<td>0.533</td>
<td>0.056</td>
</tr>
<tr>
<td>Block 1b (4-10 MEPs)</td>
<td>1.05±0.37</td>
<td>1.10±0.39</td>
<td>1.05±0.28</td>
<td>0.094</td>
<td>0.911</td>
<td>0.422</td>
<td>0.135</td>
</tr>
<tr>
<td>Block 1c (6-10 MEPs)</td>
<td>1±0.42</td>
<td>1.17±0.48</td>
<td>1.07±0.42</td>
<td>0.699</td>
<td>0.510</td>
<td>0.564</td>
<td>0.043</td>
</tr>
<tr>
<td>Block 2a (1-15 MEPs)</td>
<td>1.06±0.20</td>
<td>1.07±0.32</td>
<td>1.05±0.25</td>
<td>0.073</td>
<td>0.930</td>
<td>0.721</td>
<td>0.005</td>
</tr>
<tr>
<td>Block 2b (4-15 MEPs)</td>
<td>1.03±0.34</td>
<td>1.03±0.38</td>
<td>1.11±0.32</td>
<td>0.267</td>
<td>0.768</td>
<td>0.609</td>
<td>0.029</td>
</tr>
<tr>
<td>Block 2c (6-15 MEPs)</td>
<td>0.99±0.30</td>
<td>1.08±0.38</td>
<td>1.09±0.36</td>
<td>0.527</td>
<td>0.597</td>
<td>0.694</td>
<td>0.008</td>
</tr>
<tr>
<td>Block 3a (1-20 MEPs)</td>
<td>1.03±0.21</td>
<td>1.06±0.29</td>
<td>1.04±0.25</td>
<td>0.082</td>
<td>0.94</td>
<td>0.770</td>
<td>0.002</td>
</tr>
<tr>
<td>Block 3b (4-20 MEPs)</td>
<td>1.02±0.31</td>
<td>1.02±0.34</td>
<td>1.11±0.30</td>
<td>4.37</td>
<td>0.651</td>
<td>0.684</td>
<td>0.009</td>
</tr>
<tr>
<td>Block 3c (6-20 MEPs)</td>
<td>0.98±0.303</td>
<td>1.06±0.33</td>
<td>1.07±0.33</td>
<td>0.56</td>
<td>0.578</td>
<td>0.733</td>
<td>0.003</td>
</tr>
</tbody>
</table>

**Figure 1.** Comparison of MEPs amplitude in blocks of 20 MEPs in 3 conditions (all trials, after removal of the first 3 and 5 MEPs at 3 time points across two sessions. a) group 120% RMT, b) group 1 mV.
The ICC values for this group ranges from 0.42 to 0.77 in blocks 1, 2, and 3 indicating that raising the number of trials can lead to an increase in ICC values.

Figure 1 shows the results of comparison of MEPs amplitude in block 20 MEP responses in 3 conditions (all trials, after removal of the first 3 and 5 MEPs) for two types of TMS methods, 120% RMT and intensity to elicit 1 mV MEPs.

4. Discussion

In this study, we assessed the reliability of TMS induced MEP, using two types of TMS techniques (120% RMT and 1 mV), and considering removal of data for the first few trials in each block. The hypothesis that the removal of initial MEPs would increase the MEP reliability was refuted by the results. Our results have shown that reliability scores decrease with removal of the first 3 or 5 MEPs in each block, except for block 10 at 1 mV intensity in which removing the first 5 trials slightly increased ICCs compared to all 10 MEPs. In both techniques, we observed more reduction in ICC values with removing the first 3 MEPs in all blocks, compared to removal of the first 5 MEPs. The results also indicate that, compared to removal of the first few MEPs, the number of MEPs in each block has a more profound effect on the enhancement of reliability in both techniques.

The patterns of variability of MEP size and the mechanisms responsible for this variability have not been completely determined. Changes in the level of synchrony of neuronal pulse activity and spontaneous changes in motor neuron excitability are often identified as the sources of such variability (Srinivasan et al., 1999; Sankarasubramanian et al., 2015; Livingston & Ingersoll, 2008; Lopez-Alonso et al., 2015; Mőks et al., 1987; Truccolo et al., 2001). Large changes in CSE might result in greater fluctuations in MEP amplitude during the first few trials of TMS (Brasil-Neto et al., 1994; Ellaway et al., 1998), which can affect overall reliability of elicited MEPs. However, our finding demonstrated that removal of the first few trials resulted in lower values of MEP reliability when compared to removing all trials in any given block. The ICC values recorded for all three blocks of 10, 15, and 20 MEPs showed a rise in reliability score with increasing the number of trials, which is in agreement with those results suggesting that there is a relationship between the number of trials and reliability score (Kiers et al., 1993; Kamen, 2004; Christie et al., 2007; Bastani and Jaberzadeh, 2012).

In the current study, different impacts on reliability scores are achieved by removing the first 3 or 5 MEPs. Different values of ICCs in a given block with removal of the first 3 or 5 trials indicated that not only the number of MEPs, but also the number of removed initial trials can influence reliability of this response. In the current study, a slight increase in ICC values was observed in blocks with the first 5 trials removed, compared to exclusion of the first 3 trials. This finding can be explained by the increased homogeneity in MEP amplitudes being expected after the first 5 MEPs, which is line with some studies that reported ICC values above 0.6 for blocks of 5 MEPs (Kamen, 2004; Christie et al., 2007; Bastani and Jaberzadeh, 2012).

Similar patterns in reliability scores were found between two types of TMS techniques. The only difference was found in block 10 MEPs using TMS technique 1 mV. In this case, by removing the first 5 trials, ICCs slightly increased. In addition, there is a clear trend that, after removal of first few trials, the SD of MEP increased in the 1 mV technique more than that in the 120% RMT technique. This increase was larger for the removal of 5 trials than the removal of 3 trials, indicating the first 3 or 5 MEPs were very close to the mean value of all trials.

Taken together, to receive reliable responses, increasing the number of trials might be more effective than removing the first few trials. Therefore, using 20 MEPs allows us to accurately measure mean MEP amplitude as a valid outcome. More studies are needed to find out factors which contribute to MEP variability and the reliability of MEP responses.

There are some limitations in this research. Healthy young participants were assessed in this study, therefore, our results cannot be extrapolated to other populations such as patients or elderly people. Furthermore, the intensity of the stimuli was set at 120% of RMT or 1 mV at rest condition, therefore the findings could not be generalized to other TMS intensities and active conditions. Future studies must be conducted on patients, on other age ranges, and for active and rest conditions at different TMS intensities.

This study demonstrated that a greater number of trials involving averaged MEPs can influence TMS reliability more than removal of the first few trials in a given block. On the other hand, removal of more variable and fluctuating initial MEPs did not have a significant impact on overall reliability of TMS-induced MEPs between two techniques (1 mV and 120% RMT).

Acknowledgments
The authors gratefully thank Dr. Allie Ford for her proofreading on this manuscript. This paper had no financial support.

Conflict of Interest

The authors declared no conflict of interests.

References


