Diagnostic accuracy of motor evoked potentials to detect neurological deficit during idiopathic scoliosis correction: a systematic review

Parthasarathy D. Thirumala, MD, Donald J. Crammond, PhD, Yoon K. Loke, MBBS, Hannah L. Cheng, Jessie Huang, and Jeffrey R. Balzer, PhD

OBJECTIVE The goal of this study was to evaluate the efficacy of intraoperative transcranial motor evoked potential (TcMEP) monitoring in predicting an impending neurological deficit during corrective spinal surgery for patients with idiopathic scoliosis (IS).

METHODS The authors searched the PubMed and Web of Science database for relevant lists of retrieved reports and/or experiments published from January 1950 through October 2014 for studies on TcMEP monitoring use during IS surgery. The primary analysis of this review fit the operating characteristic into a hierarchical summary receiver operating characteristic curve model to determine the efficacy of intraoperative TcMEP-predicted change.

RESULTS Twelve studies, with a total of 2102 patients with IS were included. Analysis found an observed incidence of neurological deficits of 1.38% (29/2102) in the sample population. Of the patients who sustained a neurological deficit, 82.8% (24/29) also had irreversible TcMEP change, whereas 17.2% (5/29) did not. The pooled analysis using the bivariate model showed TcMEP change with sensitivity (mean 91% [95% CI 34%–100%]) and specificity (mean 96% [95% CI 92–98%]). The diagnostic odds ratio indicated that it is 250 times more likely to observe significant TcMEP changes in patients who experience a new-onset motor deficit immediately after IS correction surgery (95% CI 11–5767). TcMEP monitoring showed high discriminant ability with an area under the curve of 0.98.

CONCLUSIONS A patient with a new neurological deficit resulting from IS surgery was 250 times more likely to have changes in TcMEPs than a patient without new deficit. The authors’ findings from 2102 operations in patients with IS show that TcMEP monitoring is a highly sensitive and specific test for detecting new spinal cord injuries in patients undergoing corrective spinal surgery for IS. They could not assess the value of TcMEP monitoring as a therapeutic adjunct owing to the limited data available and their study design.

https://thejns.org/doi/abs/10.3171/2015.7.SPINE15466

KEY WORDS idiopathic scoliosis; motor evoked potentials; paraplegia; spinal cord injury; intraoperative monitoring; surgical technique

Neurological deficits can range from loss of sensation and paralysis of voluntary muscles to chronic pain, fatigue, and mental health dysfunction. Potential debilitating influences on various body systems can further reduce a patient’s quality of life, leading to depression, anxiety, and low self-esteem. Studies have predicted that 20%–40% of people with spinal cord injuries are at risk for a depressive disorders while in rehabilitation, with about 15%–60% at risk 1 year postdischarge. The use of intraoperative neurophysiological monitoring of spinal cord...
function has been shown to reduce the risk of motor deficit or paraplegia\(^6\,\)and is now standard and recommended by the Scoliosis Research Society and the guidelines by the American Academy of Neurology\(^6\) during surgical procedures which incur a risk of damaging the spinal cord.\(^{35,46}\)

Somatosensory evoked potential (SSEP) monitoring has been widely recognized as a means to reduce the incidence of spinal cord injury during corrective scoliosis surgery.\(^{29}\) However, the use of SSEPs alone can only provide indirect evidence of injury to the motor system.\(^{38,45,47}\) In recognition of this risk, a variety of electrophysiological monitoring techniques that assay the descending motor pathways have been developed, including direct cortical stimulation, transcranial magnetic stimulation, and transcranial electrical stimulation.\(^{29}\) The most commonly used stimulation technique in the operating room, however, is transcranial electrical stimulation.\(^{29}\) Transcranial motor evoked potential (TcMEP) monitoring during corrective idiopathic scoliosis (IS) surgery thus plays an important role in reducing the incidence of neurological complications by directly monitoring the descending corticospinal motor tracts.\(^{23}\) Significant changes in SSEPs and TcMEPs can be immediately communicated to a surgeon. Though significant changes in SSEPs are universally accepted, there is no established or agreed upon “alarm criteria” for TcMEPs that associate a significant “change” in TcMEPs from the baseline with a postoperative neurological deficit.\(^3\,\)\(^{22,45}\) 

TcMEPs are sensitive to tissue ischemia and should be able to detect potential motor deficits earlier than SSEPs, which require averaging over a longer time period, thus potentially enabling more rapid identification and reversal of impending spinal cord injury.\(^{33,45}\) A limitation of TcMEP monitoring compared with SSEP recording, however, is that TcMEPs are less reliable and record variable responses from moment to moment that preclude quantification; TcMEPs are also more sensitive to routinely used anesthetic agents, which suppress cortical and spinal motor neuron excitability.\(^{23}\) Though TcMEP sensitivity has previously been believed to be 100%, recent studies have shown that there is a possibility of false-positive TcMEP changes related to a patient’s obesity and increased duration of surgery.\(^{3,17}\) Nevertheless, the predictive value of TcMEP changes during idiopathic scoliosis procedures related to neurological deficit, could provide an additional monitoring modality for surgeons to increase diagnostic accuracy.

The objective of this study is to perform a systematic review of available peer-reviewed literature to evaluate the efficacy of TcMEP monitoring in predicting new neurological complications in patients undergoing IS surgery. The aim of this review is to assess the sensitivity, specificity, diagnostic odds ratio, and area under receiver operating characteristic (ROC) curves of intraoperative TcMEP changes in relation to neurological outcome in patients undergoing surgical procedures for IS.

**Methods**

**Search Criteria**

The PRISMA 2009 guidelines were followed.\(^{31}\) A systematic search of peer-reviewed publications, using the MEDLINE/PubMed database, was conducted to determine eligible studies published before October 2014. The following key words were used to locate studies based on patients with IS: “scoliosis,” “spinal deformity,” and “correction spinal deformity.” The search was further refined to select for patients who underwent corrective scoliosis surgery with TcMEP monitoring, using the key words: “intraoperative neurophysiological monitoring,” “motor evoked potentials,” “motor evoked potential,” and “intraoperative neurophysiological monitoring.” Motor evoked potential monitoring during surgical procedures for IS was used as the index test and postoperative analysis of TcMEP monitoring information as the reference standard.

**Study Selection**

Studies were incorporated in the systematic review if they satisfied the following inclusion criteria: 1) were randomized controlled trials, prospective, or retrospective cohort reviews, 2) conducted in patients with any type of IS, including infantile, juvenile, and adolescent scoliosis, 3) had intraoperative TcMEP monitoring performed during corrective procedures, 4) had immediate postoperative assessment, 5) has ≥ 25 patients as the total sample size, 6) published in English, and 7) included the absence of postoperative neurological deficits.

All titles and abstracts were independently screened, by 3 authors (H.L.C., P.D.T., and J.H.), against the inclusion criteria to identify relevant studies. Studies that did not meet the inclusion criteria were rejected and the reason for rejection recorded on an Excel spreadsheet, indicated by the corresponding inclusion criteria (1–7). Discrepancies between evaluators were resolved by discussion, and a final list of eligible publications was generated.

**Data Extraction**

Data were extracted independently by the authors to ensure consistency. The extracted information contained the following: first author’s name, year of publication, study design, intraoperative neurophysiological monitoring modality (TcMEP and others), time the baseline values were obtained, study data (total sample size, idiopathic sample size, TcMEP changes, reversible and irreversible changes to TcMEP), and outcome data (reversible and irreversible neuromuscular deficits). A postoperative deficit was defined as any novel persistent neurological deficit (weakness, paraplegia) that lasted at least 1–24 hours, excluding sensory deficits. TcMEP change was classified as at least a 65%–80% reduction in amplitude compared with the baseline. An irreversible TcMEP change was defined as any change that did not return to baseline despite increase in blood pressure and/or increase in stimulus intensity or pulse number. A reversible TcMEP change was defined as any intraoperative change that resolved.

The number of true positives, false negatives, false positives, and true negatives in patients with IS were extracted and tabulated for each study. True positives (TP) were defined as patients with TcMEP changes and with a new nonsensory postoperative neurological deficit; false negatives as patients with no TcMEP changes and with a new nonsensory postoper-
ative neurological deficit; true negatives as patients with no TcMEP changes and with no new nonsensory postoperative neurological deficits; and false positive as patients with TcMEP changes and without a new nonsensory postoperative neurological deficit.

Assessment of Methodological Quality

The QUADAS-2 tool was used to assess the susceptibility to bias of the included studies. The 4 domains assessed by the QUADAS-2 tool were patient selection, index test, reference standard, and flow and timing. The patient selection category assesses the presence of non-consecutive or nonrandom sampling, case-control, or inappropriate exclusion. The index test refers to TcMEP monitoring. The reference standard refers to the postoperative analysis of TcMEP results. Signaling questions aid in assessing the potential risk of bias introduced by the conduction, interpretation, or applicability of the index test and the reference standard. Flow and timing refer to the interval between the index and reference tests. Delay or treatment of patients between the index test and the reference standard can cause misclassification and introduce bias. Signaling questions for this domain help to judge the possibility of verification bias, which may occur if a portion of the patient population does not receive the reference standard or index test, or if a portion does not receive the same reference standard or index test. If the answers to all signaling questions in a domain are “yes,” then the “low” risk grade is given. If the answer to any signaling question is “no,” then a “high” risk grade is given. The “unclear” category was only used when the reported data were insufficient to permit a judgment. The methodological quality of the included studies was assessed independently by two review authors and disagreement was resolved by reexamination of primary literature.

Statistical Analysis

We used Stata 13 for the statistical analyses (Stata Statistical Software, Release 13, StataCorp). The primary analysis of this review served to fit data into a hierarchical summary ROC curve model using a bivariate model, which has been demonstrated to yield useful summary measures of diagnostic test performance, such as sensitivity and specificity.[44] We were also able to obtain area under the ROC curve, pooled sensitivity, specificity, and pooled diagnostic odds ratio through the same bivariate model used in generating the hierarchical summary ROC curve. We were unable to integrate data sets where TP+FN ≥ 0, or TN+FP = 0 (TP, true positive; FN, false negative; TN, true negative; FP, false positive), into our statistical analysis because we could not accurately estimate either sensitivity or specificity. A Fagan nomogram was drawn to show the positive and negative likelihood ratios and the post EEG change probability of perioperative stroke. A funnel plot was constructed to check for publication bias.[7]

Results

Literature Search

A total of 522 peer-reviewed publications were initially identified through our electronic search of the MEDLINE/PubMed database, of which 466 studies were excluded after screening titles and abstracts (Fig. 1). After assessing the full text of the remaining 56 studies, 30 publications were removed for failing to meet the inclusion criteria, and 13 studies were excluded because the data present were not sufficient for calculating sensitivity or specificity. A publication by Padberg et al. was excluded after peer review. The remaining 12 studies were included in the systematic review, and we were able to conduct a statistical analysis with the bivariate model in 8 studies. All selected studies used TcMEP monitoring as a modality during corrective IS surgery.

Study Characteristics

Table 1 illustrates the study characteristics. All studies included used MEP monitoring for corrective surgical procedures in IS patients. Baseline recordings were obtained either before or after incision. The alarm criterion defining significant change in TcMEP was at least a 50%–80% decrease in amplitude. Figure 2 shows a graphic display of the proportion of studies with low, high, or unclear risk of bias assessed using the QUADAS-2 tool.

Table 2 shows patient demographics. The 12 eligible studies evaluated included 2102 patients with IS. The total incidence of neurological deficits in these patients was 1.38% (29/2102). No TcMEP change was observed in 2007 patients (95.5%). TcMEP change indicative of a new neurological deficit was observed in 95 (4.52%) of 2102 patients. Of this subgroup, 38 deficits (40.0%) were reversible, 33 (34.7%) were irreversible, and data were not reported for the remaining 24 (25.3%) for which a TcMEP change was observed. In the population of patients who sustained a neurological deficit, 24 (82.8%) of 29 deficits were preceded by an irreversible TcMEP change, while the remaining 5 (17.2%) were not.

Statistical Analysis Results

Figure 3 shows a forest plot of sensitivities and specificities for each publication. The combined specificity of the studies was 0.96 (95% CI 0.92–0.98) and the combined sensitivity was 0.91 (95% CI 0.34–1.00). There was substantial heterogeneity in these pooled analyses (I² = 89%, 95% CI 77–100%). The pooled diagnostic odds ratio for MEP monitoring was 250 (95% CI 11–5767), shown in Fig. 4. A summary ROC curve was graphed to show the overall test performance (Fig. 5). The bivariate model yielded an area under the ROC curve for TcMEP monitoring of 0.98 (95% CI 0.98–0.99), which indicates excellent ability to distinguish between patients who develop complications and those who are unharmed. The subgroup analyses were performed for reversibility of MEP changes. No major differences in the diagnostic performance were noted, and we were not able to fully account for the heterogeneity.

A Fagan nomogram (Fig. 6) was drawn to determine the posttest probability of neurological deficit in a patient based on the result of the diagnostic test (TcMEP monitoring) and the pretest probability. The pretest probability was assumed to be equal to the incidence of deficits in our cohort (1.38%). The positive likelihood ratio for TcMEP change in patients with postoperative neurological deficit...
Diagnostic accuracy of motor evoked potentials

J Neurosurg Spine Volume 26 • March 2017

377

was calculated to be 26 and the negative likelihood ratio was estimated to be 0.11. Using the line drawn from the pretest probability of 1.38% through the positive likelihood ratio of 23, the posttest probability of a neurological deficit was found to be 26.31%. The probability of no neurological deficit after a negative test (no TcMEP change) was estimated to be 99.85%.

Discussion

The results suggest that TcMEP monitoring is a reliable method of assessing the integrity of the corticospinal pathways during corrective scoliosis surgery, with a specificity of 0.96 and sensitivity of 0.91. The diagnostic odds ratio indicated that it is 250 times more likely to observe significant TcMEP changes in patients who experience a new-onset motor deficit immediately after IS correction surgery. Twenty-nine (1.38%) of the 2102 patients included in this systematic review developed a neurological deficit postoperatively, a rate that is comparable to previously published rates of iatrogenic injury during these procedures, which have ranged from 0.6% to 3.5%.21

The high specificity (0.96) is characteristic of TcMEPs and confirms the value of TcMEP monitoring as a gold standard for neuromonitoring of the motor tracts.45 In calculating the sensitivity, patients with irreversible changes in TcMEP but without postoperative neurological deficits were presumed to represent false positives instead of true positives. These results reflect a lower sensitivity (0.91) compared with the sensitivity (1.0) reported previously.17

It is possible that the lower positive predictive value is a result of the corrective steps taken following a significant TcMEP change, which may have prevented neurological deficit. The positive likelihood ratio indicated that a patient who experienced a neurological deficit was 26 times more likely have a positive test result (TcMEP change). The prevalence obtained in our study (1.38%) was used for the Fagan nomogram, which estimated that the probability of experiencing a postoperative neurological deficit after a positive TcMEP change was 26.31%. As expected, a negative test result (no TcMEP change) indicated that the probability of no postoperative neurological deficit was 99.85%. TcMEPs have been shown to be particularly sensitive to ischemia and compressive injuries, due in part to the tenuous and less redundant nature of the anterior column’s blood supply.4,13,29,49 Adequate blood pressure between 50 and 150 mm Hg is thus vital in maintaining normal perfusion in the brain and spinal cord.20,29 Animal studies have shown that TcMEPs were depressed when cerebral blood flow was reduced to less than 16 ml/min/100 g.25 Examination of compressive-contusion–type injuries in animal models similar to spinal cord injuries that can occur during scoliosis corrective maneuvers have shown that vascular insults affect the metabolically active gray matter in the anterior horn more than the white matter.15,20,27 Since it is believed that most postoperative paraplegia is related to ischemia, the sensitivity of TcMEPs to ischemic insult enables TcMEPs to be a better and earlier indicator of impending neurological damage than SSEPs, which are relatively resilient to ischemia and have been
known to remain unchanged despite significant motor spinal cord injury.\textsuperscript{2,8,40,45} There are no accepted criteria for detecting impending neurological deficits using TcMEP monitoring. Alarm criteria as defined by the studies included in the present systematic review ranged from 50% to 80% decreases in TcMEP amplitude despite reports that TcMEP amplitudes can vary considerably from trial to trial within patients.\textsuperscript{16,29} Motor units demonstrate an all-or-nothing electrophysiological characteristic, and although compound muscle responses are more graduated, they still exhibit nonlinearity.\textsuperscript{27} The fluctuations in corticospinal and motor neuron excitability levels to consistent stimuli can cause variability in TcMEP response.

The low incidence of false negatives (0.14%) is consistent with rates reported in the peer-reviewed literature. There were 70 cases (3.33%) of false positives, although this rate is likely because we defined the presence of TcMEP changes and the absence of a postoperative deficit as a false-positive result rather than a true positive. However,
### TABLE 2. Patient demographics

<table>
<thead>
<tr>
<th>Authors &amp; Year</th>
<th>Sample Size</th>
<th>MEP Change*</th>
<th>MEP Change Reversible†</th>
<th>MEP Change Irreversible‡</th>
<th>Neurological Deficit</th>
<th>New Deficit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accadbled et al., 2006</td>
<td>191</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Eggspuehler et al., 2007</td>
<td>217</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>El-Hawary et al., 2006</td>
<td>177</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Feng et al., 2012</td>
<td>176</td>
<td>3</td>
<td>NA</td>
<td>NA</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Kundnani et al., 2010</td>
<td>354</td>
<td>13</td>
<td>9</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Lo et al., 2008</td>
<td>25</td>
<td>9</td>
<td>NA</td>
<td>NA</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Luk et al., 2001</td>
<td>30</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MacDonald et al., 2007</td>
<td>206</td>
<td>7</td>
<td>6</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Noonan et al., 2002</td>
<td>134</td>
<td>10</td>
<td>NA</td>
<td>NA</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Pastorelli et al., 2011</td>
<td>172</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Péreson et al., 1998</td>
<td>112</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Schwartz et al., 2007</td>
<td>1121</td>
<td>38</td>
<td>12</td>
<td>26</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Total no.</td>
<td>2915</td>
<td>95</td>
<td>38</td>
<td>33</td>
<td>29</td>
<td>24</td>
</tr>
</tbody>
</table>

NA = not available.

* Greater than 50% decrease in N20-P25 and/or 10% increase in latency.

† MEP change that returned to baseline.

‡ MEP change that did not return to baseline despite increase in blood pressure and/or increase in stimulus intensity or pulse number.

---

**FIG. 3.** A bivariate linear model of the sensitivity and specificity of MEPs in included studies, with associated confidence intervals and summary measures. Figure is available in color online only.
other studies have found relatively high false-positives rates in TcMEP monitoring.\textsuperscript{17,48} It is hypothesized that the cause of such high rates of false-positive TcMEP changes is the use of inhalational anesthetics, obesity, prolonged length of surgery, and failure to adjust anesthetic regimen for degradation of TcMEP response; additionally, modest TcMEP amplitude changes can be subclinical and may not reflect a new neurological deficit.\textsuperscript{4,17,29,48} Another factor may be the lack of standard alarm criteria for TcMEP monitoring.\textsuperscript{29} A recent multicenter study by institutions of the Japanese Society for Spine Surgery and Research has suggested a 70% decrease in amplitude as an alarm point during surgery for spinal deformity.\textsuperscript{19} Other published studies have used a reduction in amplitude of ≥ 50%.\textsuperscript{18} The only generally accepted warning sign is the complete disappearance of a consistently present response.

While the use of intraoperative monitoring has steadily increased since its introduction in the early 20th century, there remains some debate regarding the level of evidence available to support the idea that action taken in response to a neurophysiological alert can improve neurological outcome. To the best of our knowledge, the study by Wie-
demayer et al. is the only study to compare the rate of new postoperative neurological deficits in patients in whom an intervention was performed in response to an intraoperative alert with rates in patients in whom action was not taken despite an intraoperative alert. The modalities the authors used were SSEP and brainstem auditory evoked potential (BAEPs) monitoring. They found that the rate of new neurological deficits was 4.7% in patients who received an intervention and 15.1% in patients who did not receive an intervention. The authors concluded that interventions during intraoperative neuromonitoring aided in the prevention of postoperative deficit in 5.2% cases (n = 22/423). Unfortunately, the study only offers conclusive evidence regarding the efficacy of interventions based on SSEPs and BAEPs in the operating room. We suggest that future studies of this nature done after determining appropriate alarm criteria include TcMEPs as well.

Although our systematic review has significant strengths in its comprehensive peer-reviewed literature search and quality assessment with QUADAS-2, it is important to note that our study was subject to limitations. We evaluated the use of TcMEPs during scoliosis surgery as a diagnostic adjunct; our study was not designed to assess the role of TcMEPs as a therapeutic adjunct, and it offers no definite data to support a correlation between an intervention after TcMEP waveform changes during scoliosis surgery and altered postoperative neurological outcome. Some search bias may exist due to the difficulty of obtaining all relevant published studies that assessed the use of TcMEP monitoring for patients with IS. Significant heterogeneity was observed in the sensitivity and specificity of the studies. Causes of heterogeneity were explored in the analyses; however, due to the nature of the systematic review, we were limited by the available data published in the individual studies. It is plausible that some of the heterogeneity can be attributed to the reversibility of TcMEP waveforms, which is desirable but not always achieved.

Conclusions

The findings from our systematic review indicate that intraoperative TcMEP monitoring is a highly sensitive and specific test for predicting neurological deficits in patients undergoing corrective spinal surgery for IS. It was 250 times more likely to observe significant TcMEP changes in patients who experienced postoperative neurological deficits. TcMEPs can be an effective biomarker for spinal cord injury during scoliosis fusion. Experimental and clinical studies are necessary to evaluate the alarm criteria needed to warn the surgeons during the surgical procedure.
FIG. 6. The posttest probability was obtained by drawing a line from the estimated pretest probability (1.38%) through the center axis (likelihood ratio). prob. = probability. Figure is available in color online only.

References
Disclosures
The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Author Contributions
Conception and design: Thirumala. Acquisition of data: Cheng, Huang. Analysis and interpretation of data: Loke, Huang. Drafting the article: Cheng, Huang. Critically revising the article: Cheng, Huang. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Thirumala. Statistical analysis: Loke. Study supervision: Thirumala.

Correspondence
Parthasarathy D. Thirumala, Center for Clinical Neurophysiology, Department of Neurologic Surgery, University of Pittsburgh Medical Center, UPMC Presbyterian, 200 Lothrop St., Ste. B-400, Pittsburgh, PA 15213. email: thirumalapd@upmc.edu.