Diagnostic accuracy of motor evoked potentials to detect neurological deficit during idiopathic scoliosis correction: a systematic review
Parthasarathy D. Thirumala, Donald J. Crammond, Yoon K. Loke, Hannah L. Cheng, Jessie Huang, and Jeffrey R. Balzer
Journal of Neurosurgery: Spine

Abstract

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.

INTRODUCTION

Iatrogenic spinal cord injury leading to paraplegia is an uncommon, but devastating complication. The prevalence of such neurological deficits during corrective spinal surgery has been estimated by the Scoliosis Research Society (SRS) to be at least 1% [1-3]. Though rare, the economic, physical, social, and psychological burdens that result are significant. Economic loss is reckoned to be upwards of $0.65 million to $4.6 million for any person paraplegic or tetraplegic at the age of 25 [4, 5]. Neurological damage can range from loss of sensation and paralysis of voluntary muscles to chronic pain, fatigue, and mental health dysfunction [6-8]. Potential debilitating influences on various body systems can further reduce a patient's quality of life, leading to depressive moods, anxiety, and low self-efficacy [7, 8]. Studies have predicted that 20% to 40% of people with spinal cord injuries are at risk of a depressive disorder while in rehabilitation [8], with about 15% to 60% at risk post 1-year discharge [8, 9]. The use of intraoperative neurophysiological monitoring (IONM) of spinal cord function has been shown to reduce risk of motor deficit or paraplegia [10] and is now standard and recommended during surgical procedures which bear a risk of damaging the spinal cord [10, 11].

Somatosensory evoked potential (SSEP) monitoring has been widely recognized to reduce the prevalence of spinal cord injury during corrective scoliosis surgery [12]. However, the use of SSEPs alone can only provide indirect evidence of injury to the motor system [13-16]. In recognition of this risk, a variety of MEP monitoring techniques have been devised, including direct cortical stimulation (DCS) and transcranial magnetic stimulation (TMS) [12]. The most commonly used stimulation technique, however, is transcranial electric stimulation (TES) [12]. Transcranial motor evoked potential (TcMEP) monitoring during corrective IS surgery thus plays an increasingly important role in reducing the incidence of neurological complications through direct monitoring of the corticospinal motor tracts. TES is usually applied to cross scalp (C3/C4) and midline (C3Cz/C4Cz) positions [17]. There is no officially established “alarm”; reductions in MEPs varying from 65% - 80% compared to the baseline have been used as a neurophysiological alert [13, 18, 19]. MEPs are highly sensitive, and have been shown to be able to detect potential motor deficits sooner and more accurately than SSEPs, enabling more rapid identification and reversal of impending spinal cord injury [13, 20]. A major drawback of MEP monitoring, however, is that it may be difficult to obtain reliable signals, particularly in the lower extremities, due to anesthetic agents, which suppress cortical and spinal motor neuron excitability [17]. Though MEP sensitivity has previously been believed to be 100%, recent studies have shown that there is a possibility of false-positives resulting from obesity and increased length of surgery [21, 22]. Nevertheless, the predictive value of MEP changes during idiopathic scoliosis procedures could offer a
helpful avenue for surgeons to increase diagnostic accuracy during the IS procedure to detect global spinal cord problems as well other weaknesses which can lead to post-operative paraplegia.

The objective of this paper is to perform a systematic review of available scientific literature to evaluate the efficacy of motor evoked potentials in reducing neurological complications in patients undergoing corrective IS surgery. By assessing the sensitivity, specificity, diagnostic odds ratio, and area under receiver operating characteristic (ROC) curves of intraoperative MEP changes in relation to neurological outcome in patients undergoing surgical procedures for idiopathic scoliosis.

METHODS
Search criteria

The PRISMA 2009 guidelines were followed. A systematic literature search, using the MEDLINE/PubMed database, was conducted to determine eligible studies published before October 2014. The following keywords were used to locate studies based on patients with idiopathic scoliosis: “scoliosis”, “spinal deformity”, and “correction spinal deformity”. The search was further refined to select for patients who underwent corrective scoliosis surgery with MEP monitoring, using the keywords: “intraoperative neurophysiological monitoring”, “motor evoked potentials”, “motor evoked potential”, and “intraoperative neurophysiological monitoring”. Motor evoked potential monitoring during surgical procedures for idiopathic scoliosis was used as the index test and post-operative analysis of MEP monitoring information as the reference standard.

Study Selection

Studies were incorporated in the meta-analysis if they satisfied the following inclusion criteria: (1) were randomized controlled trials, prospective, or retrospective cohort reviews, (2) conducted in patients with idiopathic scoliosis, (3) had intraoperative MEP monitoring performed during corrective procedures, (4) had immediate post-operative assessment, and (5) ≥ 25 patients as the total sample size. Studies published in languages other than English were excluded.

All titles and abstracts were independently screened, by the authors (H.C., P.D.T, J.E.H), against the inclusion criteria to identify relevant studies. Studies that did not meet the specific criteria were rejected and the reason for rejection recorded on an Excel spreadsheet, indicated by a number corresponding to one of the inclusion criteria (0-6). Additional criteria include the absence of post-operative neurological deficits. Discrepancies between evaluators were resolved by discussion, and a final list of eligible articles was generated.

Data Extraction

Data was extracted independently by the authors to ensure consistency. The extracted information contained: first author’s name, year of publication, study design, IONM modality (MEP and others), time the baselines were obtained, study data (total sample size, idiopathic sample size, MEP changes, reversible and irreversible changes to MEP), and outcome data (neuromuscular deficits,
reversible and irreversible). Post-operative deficit was defined as any persistent neurological deficit (weakness, paraplegia) that was present post-operatively (post-op) and lasted at least 1-24 hours, but excluding sensory deficit. MEP change was classified as a 65%-80% reduction in amplitude compared to the baseline. Irreversible MEP change was defined as any change that did not return to baseline despite increase in blood pressure and/or transient abortion of procedure. Reversible MEP change was defined as any intraoperative change that returned to baseline after increase in blood pressure, and/or temporary cessation of the operation.

The number of true positives, false negatives, false positives, and true negatives in patients with idiopathic scoliosis were extracted and tabulated for each study.


Assessment of Methodological quality

The review authors used the QUADAS 2 tool to assess the susceptibility to bias of the included studies [23]. The four domains assessed by the QUADAS 2 tool were patient selection, index test, reference standard, and flow and timing. Patient selection refers to avoiding nonconsecutive or nonrandom sampling, case-control, or inappropriate exclusion. The index test refers to proper MEP monitoring. The reference standard refers to proper testing for post-operative neurological deficits. Flow and timing refers to the interval between the index and reference tests, whether all patients received the same reference test and whether all patients were included in the analysis. 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 where the reported data was 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 (StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP). The primary analysis of this review was to fit the data into a hierarchical summary receiver operating curve (HSROC) model using the bivariate model, which has been demonstrated to yield useful summary measures of diagnostic test performance[24]. We were also able to obtain area under the receiver operating curve (AUROC), pooled sensitivity, specificity and pooled diagnostic odds
ratio (DOR) through the same bivariate model used in generating the HSROC. We were unable to integrate datasets where TP+FN = 0, or TN+FP = 0, into our meta-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.

RESULTS

Literature Search

A total of 522 studies were initially identified through our electronic search of the MEDLINE/PubMed database, of which 466 studies were excluded after screening titles and abstracts (Figure 1). After assessing the full text of the remaining 56 studies, 30 papers were removed for failing to meet the inclusion criteria, and 13 studies had insufficient data. A publication by Padberg et al.37 was excluded after peer review. The remaining 12 studies were included in the systematic review, and we were able to conduct meta-analysis with the bivariate model in nine studies. All selected studies used MEP monitoring as a modality during corrective scoliosis surgery.

Study Characteristics

Baseline recordings were obtained either before or after incision. Preestablished alarm criteria for significant changes in MEP were classified as a 50% - 80% decrease in amplitude and 10% increase in latency from baseline values (Table 1). Table 2 shows the patient demographics. The 12 eligible studies evaluated involved 2,102 patients with idiopathic scoliosis. 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 ($I^2 = 89$, 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 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 MEP monitoring is a promising and reliable method of assessing spinal cord integrity during corrective scoliosis surgery, with a specificity of 0.96 and sensitivity of 0.91. The diagnostic odds ratio indicated that it was 250 times more likely to observe significant MEP changes in patients with paraplegia after idiopathic scoliosis correction. 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%.

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. 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. 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. In calculating the sensitivity, patients with changes in MEP but without postoperative neurological deficit were presumed to be false positives instead of true positives, hence a lower sensitivity than that reported by current literature. It is well known that MEPs are highly sensitive compared to SSEPs, but while they seem to be influenced by the same systemic factors, MEPs are more vulnerable to ischemic injuries, and thus experience more changes in amplitude than SSEPs due to the nature of their blood supply [12, 21, 25]. The anterior spinal artery (ASA) supplies around 75% of the spinal cord, which includes gray matter and anterior horn cells [26, 27]. The ASA receives a rather limited flow from the radicular arteries compared to the posterior spinal artery (PSA), which supplies the sensory tracts [27]. Adequate blood pressure between 50 and 150 mmHg is thus vital in maintaining normal perfusion in the brain and spinal cord [12, 27]. Studies on baboons have shown that MEPs were depressed when cerebral blood flow was reduced to less than 16 mL/min/100 g [28, 29]. Autoregulation will be lost and hypoperfusion may occur if cerebral perfusion pressure (CPP) and oxygen delivery decrease [27]. In addition, compressive-contusion-type injuries in animal models, similar to spinal cord injuries during scoliosis fusion, appear to more severely affect the gray matter than the white matter [30]. As a result, anterior horn cells could potentially be affected; MEPs will thus be more sensitive than SSEPs during scoliosis fusion. Swelling and hemorrhaging of white matter accompanies necrosis of gray matter, and may lead to spinal cord ischemia [31]. It is believed that hyper/hypocapnia, hypoxemia, and anemia affect MEP waveforms, but these effects are minimal and require further analysis [29]. Animal studies on rats have shown that slight latency and reduction of amplitude of MEP waveforms occur when subjected to moderate hypoxia (15.75% O2), with deviation from baseline becoming more pronounced as the level of hypoxia increases [32]. It should be noted that in patients with severe scoliosis, there is a chance that spinal cord blood flow may already be compromised [33, 34]. The vulnerability of the motor pathways to changes in blood flow make MEPs a better indicator of spinal cord integrity than SSEPs, which are more resilient to ischemia and have been known to remain unchanged despite significant spinal cord injury.

There is no widely accepted criteria for detecting an impending neurological deficit by MEP monitoring. Alarm criteria as defined by the studies included in the meta-analysis ranged from 50% - 80% decreases in amplitude. It has been widely reported that amplitudes vary considerably from trial to trial [12, 35]. Motor units have an all-or-nothing behavior, and though compound muscle responses are more graduated, they still exhibit non-linearity [26]. This characteristic, while allowing for high sensitivity, makes it challenging to clearly differentiate between a minor degree of deterioration of the motor tract and a complete loss of response.

The low incidence of false-negatives in 3 out of 2102 patients (0.14%) is concurrent with the current literature and is likely due to the high sensitivity of MEPs. There were 76 (2.92%) cases of false-positives, though it is likely because we presumed that patients
with MEP changes and without postoperative deficit were false positives rather than true positives. However, other studies have found relatively high rates of false-positives in MEP monitoring [22, 36]. It is hypothesized that the cause of such high incidences is the use of inhalation anesthetics, obesity, prolonged length of surgery, or failure to adjust anesthetic regimen for fade [12, 21, 22, 36]. Another factor may be the lack of standard alarm criteria for MEP monitoring [12]. We recommend that proper criteria be selected based on published evidence and highlight the importance of experience and proper methodology in reducing the frequency of false-positives.

Although our meta-analysis has significant strengths in its comprehensive literature search and quality assessment with QUADAS-2, it is important to note that our study was subject to limitations, and that while efforts were made to identify all relevant published data, some search bias may exist. 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 meta-analysis, we were limited by the available data published by the individual studies. It is plausible that some of the heterogeneity can be attributed to the reversibility of MEP waveforms, which is desirable but not always achieved.

Acknowledgements, competing interests, etc.
Figure 1: PRISMA chart exhibiting the elimination process for study analysis.

Records identified through database searching (n = 522)

Records screened (n = 522)

Records excluded (n = 466)

Records excluded, with reasons (n = 30)

365-Not conducted with idiopathic patient
58- Sample size less than 35 patients
19- No abstract found
23- No post-operative data
1- Not randomized study

Studies included in qualitative synthesis (n = 26)

Studies included in systematic review (n = 12), with n=8 for bivariate meta-analysis
<table>
<thead>
<tr>
<th>Author (pub yr)</th>
<th>Study Design</th>
<th>Modality</th>
<th>Wakeup Test</th>
<th>Alarm Criteria *</th>
<th>Baseline SSEP#</th>
<th>Length of Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accadbled, 2006</td>
<td>Prospective Cohort</td>
<td>SSEP, NMEP</td>
<td>Yes</td>
<td>60% decrease in N20-P25 or 10% increase latency</td>
<td>Yes (after anesthesia)</td>
<td>Immediately Post-Op</td>
</tr>
<tr>
<td>El-Hawary, 2006</td>
<td>Retrospective cohort</td>
<td>nMEP, MEP, SSEP</td>
<td>Yes</td>
<td>50% decrease in N20-P25 or 10% increase latency</td>
<td>Yes (after anesthesia)</td>
<td>Immediately Post-Op</td>
</tr>
<tr>
<td>Feng, 2012</td>
<td>Retrospective Cohort</td>
<td>tceMEP, SEP, MEP</td>
<td>Yes</td>
<td>75% decrease in N20-P25</td>
<td>Yes (after anesthesia)</td>
<td>Immediately Post-Op</td>
</tr>
<tr>
<td>Kundnani, 2010</td>
<td>Prospective cohort</td>
<td>SSEP, NMEP</td>
<td>Yes</td>
<td>65% decrease in N20-P25 or 10% increase latency</td>
<td>Yes (before anesthesia)</td>
<td>Immediately Post-Op</td>
</tr>
<tr>
<td>Lo, 2008</td>
<td>Retrospective cohort</td>
<td>MEP</td>
<td>Yes</td>
<td>50% decrease in N20-P25 or 10% increase latency</td>
<td>Yes (before anesthesia)</td>
<td>Immediately to discharge and 12 weeks Post-Op</td>
</tr>
<tr>
<td>Luk, 2001</td>
<td>Prospective Cohort</td>
<td>CMEP, SSEP, SCEP</td>
<td>Yes</td>
<td>50% decrease in N20-P25 or 10% increase latency</td>
<td>Yes (after anesthesia)</td>
<td>Immediately Post-Op</td>
</tr>
</tbody>
</table>
- EEG; electroencephalogram, SSEP; Somatosensory Evoked Potential, NMEP; Neurogenic Motor-Evoked Potential

Table 2. Patient Demographics

<table>
<thead>
<tr>
<th>Author (pub yr)</th>
<th>Sample size for analysis</th>
<th>Idiopathic population</th>
<th>Patient MEP Change</th>
<th>Reversible MEP Change</th>
<th>Irreversible MEP Change</th>
<th>Neurologic deficit</th>
<th>Deficit w/ reversible</th>
<th>Deficit w/ irreversible</th>
</tr>
</thead>
<tbody>
<tr>
<td>MacDonald, 2007</td>
<td>191</td>
<td>90</td>
<td>89</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Noonan, 2002</td>
<td>217</td>
<td>60</td>
<td>60</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Pastorelli, 2011</td>
<td>177</td>
<td>136</td>
<td>80</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pereon, 1998</td>
<td>176</td>
<td>63</td>
<td>63</td>
<td>3</td>
<td>N/A</td>
<td>N/A</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Schwartz, 2007</td>
<td>354</td>
<td>354</td>
<td>354</td>
<td>13</td>
<td>9</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Lo, 2008</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>9</td>
<td>N/A</td>
<td>N/A</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Luk, 2001</td>
<td>30</td>
<td>30</td>
<td>24</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Summary Operating Point:
- SENS = 0.91 [0.34 - 1.00]
- SPEC = 0.96 [0.92 - 0.98]

SROC Curve:
- AUC = 0.98 [0.96 - 0.99]

95% Confidence Contour

95% Prediction Contour

SROC with Prediction & Confidence Contours
REFERENCES


