SPINE Volume 35, Number 9S, pp S37–S46

The Evidence for Intraoperative Neurophysiological Monitoring in Spine Surgery

Does It Make a Difference?

Michael G. Fehlings, MD, PhD, FRCSC, FACS,* Darrel S. Brodke, MD,†
Daniel C. Norvell, PhD,‡ and Joseph R. Dettori, PhD‡

Objective. The objective of this article was to undertake a systematic review of the literature to determine whether IOM is able to sensitively and specifically detect intraoperative neurologic injury during spine surgery and to assess whether IOM results in improved outcomes for patients during these procedures.

Summary and Background Data. Although relatively uncommon, perioperative neurologic injury, in particular spinal cord injury, is one of the most feared complications of spinal surgery. Intraoperative neuromonitoring (IOM) has been proposed as a method which could reduce perioperative neurologic complications after spine surgery.

Methods. A systematic review of the English language literature was undertaken for articles published between 1990 and March 2009. MEDLINE, EMBASE, and Cochrane Collaborative Library databases were searched, as were the reference lists of published articles examining the use of IOM in spine surgery. Two independent reviewers assessed the level of evidence quality using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) criteria, and disagreements were resolved by consensus.

Results. A total of 103 articles were initially screened and 32 ultimately met the predetermined inclusion criteria. We determined that there is a high level of evidence that multimodal IOM is sensitive and specific for detecting intraoperative neurologic injury during spine surgery. There is a low level of evidence that IOM reduces the rate of new or worsened perioperative neurologic deficits. There is very low evidence that an intraoperative response to a neuromonitoring alert reduces the rate of perioperative neurologic deterioration.

Conclusion. Based on strong evidence that multimodality intraoperative neuromonitoring (MIOM) is sensitive and specific for detecting intraoperative neurologic injury during spine surgery, it is recommended that the use of MIOM be considered in spine surgery where the spinal cord or nerve roots are deemed to be at risk, including procedures involving deformity correction and procedures that require the placement of instrumentation. There is a need to develop evidence-based protocols to deal with intraoperative changes in MIOM and to validate these prospectively.

Key words: SSEP, MEP, neuromonitoring, spine surgery, spinal cord injury. **Spine 2010;35:S37–S46**

Although relatively uncommon, perioperative neurologic injury, in particular spinal cord injury, is one of the most feared complications of spinal surgery.[1] It has been proposed that intraoperative neuromonitoring may reduce the risk of neurologic complications of spine surgery. The purpose of intraoperative neuromonitoring (IOM) in patients undergoing spine surgery is to: (1) identify neural irritation or injury at a time when the surgeon can take steps to reduce or reverse it and (2) define the nature of the injury in a way that will allow the surgeon to complete the procedure while minimizing further injury.[2] There is a known theoretical risk or intraoperative neurologic deterioration associated with surgery of the spinal column and its accompanying neural elements. Maneuvers such as the Stagnara wake-up test were implemented to detect and correct problems before they become irreversible.[3] Multiple monitoring methods have been developed each with a different mechanism of feedback including ascending sensory tracts, descending motor tracts, and individual nerve root monitoring.[4]

Somatosensory-evoked potentials (SSEPs) are cortical responses elicited by peripheral nerve stimulations and are the most commonly used form of intraoperative spinal monitoring tools.[5,6] Although valuable, SSEPs typically require several hundred averages and hence may involve minutes of sampling time. Thus SSEPs do not provide true real time feedback. Despite this, Nuwer *et al* reported 90% of all clinically relevant neurologic events were detected with SSEP monitoring.[6] However, there have been several reports of false-positive and false-negative SSEP studies.[7–11] Primary explanations for this are that only ascending pathways are being monitored.

This has led to efforts to supplement SSEP monitoring with feedback from descending motor pathways. Spinal motor-evoked potentials (MEPs) apply a stimulus at the spinal level and record peripheral motor responses. The primary limitation of such testing is that crossover of signals between the ascending and descending spinal

From the *Neural Repair and Regeneration, University of Toronto Spine Program, University of Toronto Neuroscience Program; †Department of Orthopaedics, University of Utah; and ‡Spectrum Research, Inc.
The manuscript submitted does not contain information about medical device(s)/drug(s).
Supported by AOSpine North America. Analytic support for this work was provided by Spectrum Research, Inc. with funding from AOSpine North America. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.
M.G.F. oversaw the overall conception and data analysis and took primary responsibility for the writing of this article; D.S.B. contributed to the conception, execution, and editing of the article; D.C.N. and J.R.D. compiled the systematic reviews and assisted with data analysis and manuscript execution.
Address correspondence and reprints to Michael G. Fehlings, MD, PhD, FRCSC, FACS, Division of Neurosurgery, University of Toronto, 399 Bathurst St, Suite 4WW-449, Toronto, ON M5T2S8; E-mail: Michael.fehlings@uhn.on.ca

	Inclusion	Exclusion
Patient	• Cervical, thoracic, lumbar surgery • Adults	• Animal studies • <18 years of age
Diagnostic test	• SSEP, MEP, EMG	• No neuromonitoring
Reference standard	• New or worsening neurological deficit	• No reference standard
Measures	• Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV)	• No report of sensitivity, specificity, PPV, NPV

Figure 1. Inclusion and exclusion criteria for diagnostic study questions 1 and 2.

pathways can lead to incorrect findings.[4] Alternatively, transcranial MEPs (tcMEPs) have been developed, applying a stimulus at the cortical level. An additional reported benefit of MEP monitoring is that recording electrodes are in place for monitoring from muscles which may facilitate the monitoring of ongoing spontaneous electromyographic (EMG) activity. Nonetheless, the necessity for more complex intravenous anesthetic techniques is a real consideration when weighing the relative merits and drawbacks of MEPs. Free-running EMG activity can be monitored continuously when peripheral nerves or roots are at risk for potential injury.[2] EMG is most commonly administered at the lumbar level, but can also be performed in cervical level procedures.

Though unimodal SSEP monitoring has been the standard of practice for those who apply IOM in pain surgery, the application of multimodal intraoperative monitoring (MIOM) has become routine in several spine centers for a myriad of surgical procedures.[6,12–16] Despite the success reported in the literature; there have been multiple anecdotal accounts of apparent false-positive alerts in which wake-up tests were performed or surgery was aborted because of a TcMEP alert and no significant neurologic deficit was observed. These false alerts may compel surgeons to take unreasonable risks or unnecessarily change the operative plan.[17]

The following questions determined the focus of this systematic review:

1. For patients undergoing instrumented spine surgery with intraoperative neuromonitoring, what are the diagnostic test characteristics (*i.e.*, sensitivity, specificity, positive predictive value, negative predictive value) of unimodal IOM (SSEPs or MEPs independently) and MIOM (combination) methods?
2. For patients undergoing instrumented spine surgery with intraoperative neuromonitoring, which methods are superior with respect to diagnostic test characteristics in the same patient population?
3. For patients undergoing instrumented spine surgery, does intraoperative neuromonitoring reduce the rate of a new or worsening neurologic event?
4. For patients undergoing instrumented spine surgery with intraoperative neuromonitoring, does the reaction to a positive intraoperative neuromonitoring alert reduce the rate of a new or worsening neurologic event? Are there specific intraoperative remedial measures employed as a result of a positive alert more effective than others?

■ Materials and Methods

Electronic Literature Database

The literature search is outlined in detail elsewhere.[17a] Briefly, a systematic search was conducted in MEDLINE, EMBASE, and the Cochrane Collaboration Library for literature published from 1990 through December 2008. We limited our results to humans and to articles published in the English language. Reference lists of key articles were also systematically checked. For our first question, we identified all articles that were designed to evaluate the diagnostic test characteristics (*i.e.*, measures of validity such as sensitivity, specificity, positive predictive value, negative predictive value) of unimodal methods (SSEPs or MEPs independently) or MIOM methods (combinations) in cervical, thoracic, and lumbar spine surgeries. For our second question, we identified all articles that compared the diagnostic test characteristics between different neuromonitoring methods (*e.g.*, SSEP, MEP, EMG, multimodal) in the same patient population in cervical, thoracic, and lumbar spine surgeries. For our third question, we identified all articles that compared new or worsening postoperative neurologic events between patients who underwent intraoperative neuromonitoring and patients who did not undergo intraoperative neuromonitoring. For our fourth question, we identified all articles that compared new or worsening postoperative neurologic events between patients where the surgeon reacted to a positive intraoperative alert with a remedial measure to patients where the surgeon did not react to a positive intraoperative alert. If possible, we attempted to determine which remedial measures were more effective. Articles were excluded if they did not calculate and report at least 1 of the 4 primary diagnostic test characteristics: sensitivity, specificity, positive predictive value (PPV), or negative predictive value (NPV). Other exclusions included reviews, editorials, case reports, non-English written studies, and animal studies. Questions number 1 and 2 were evaluated as diagnostic studies and 3 and 4 as therapy studies, Figures 1 and 2.

Data Extraction

Each retrieved citation was reviewed by 2 independently working reviewers (D.C.N. and J.R.D.). Most articles were excluded on the basis of information provided by the title or abstract. Citations that appeared to be appropriate or those that could

	Inclusion	Exclusion
Patient	• Cervical, thoracic, lumbar surgery • Adults	• Animal studies • < 18 years of age
Intervention	• SSEP, MEP, EMG (study question 3) • Response to an intraoperative alert (study question 4)	• No neuromonitoring
Comparison	• No neuromonitoring (study question 3) • No response to an intraoperative alert (study question 4)	• Neuromonitoring (study question 3)
Outcome	• Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) • Rate of a new or worsening neurological deficit	• No report of sensitivity, specificity, PPV, NPV or rate of a new or worsening neurological deficit.

Figure 2. Inclusion and exclusion criteria for therapy study questions 3 and 4.

not be excluded unequivocally from the title and abstract were identified, and the corresponding full text reports were reviewed by the 2 reviewers. Any disagreement between them was resolved by consensus. From the included articles, the following data were extracted: patient demographics, diagnosis, spine surgical intervention, monitoring methods and corresponding alert thresholds, intraoperative management (*i.e.*, remedial measures) reference standards, and results.

Study Quality

Level of evidence ratings were assigned to each article independently by 2 reviewers using criteria set by *The Journal of Bone and Joint Surgery, American Volume* (*J Bone Joint Surg Am*)[18] for diagnostic studies and therapeutic studies and modified to delineate criteria associated with methodologic quality and described elsewhere (See Supplemental Digital Content 1, individual study ratings, tables, individual study ratings, available at: http://links.lww.com/BRS/A424).

Analysis

Postoperative neurologic events rates were reported as the proportion of patients experiencing a new or worsening neurologic deficit as defined by each author. Diagnostic characteristics were based on author calculations. Data were summarized in Tables 1–4 and qualitative analysis[19] was performed considering the following 3 domains: quality of studies (level of evidence), quantity of studies (the number of published studies similar in patient population, condition treated, and outcome assessed), and consistency of results across studies (whether the results of the different studies lead to a similar conclusion).[20] We judged whether the body of literature represented a minimum standard for each of the 3 domains using the following criteria: for study quality, at least 80% of the studies reported needed to be rated as a level of evidence I or II; for study quantity, at least 3 published studies were needed which were adequately powered to answer the study question; for study consistency, at least 70% of the studies had to have consistent results. The overall strength of the body of literature was expressed in terms of the impact that further research may have on the results. An overall strength of "HIGH" means that further research is very unlikely to change our confidence in the estimate of effect. The overall strength of "MODERATE" is interpreted as further research is likely to have an important impact on our confidence in the estimate of effect and *may* change the estimate. A grade of "LOW" means that further research is very likely to have an important impact on our confidence in the estimate of effect and is *likely* to change the estimate, whereas "VERY LOW" means that any estimate of effect is very uncertain.[17a]

■ Results

We identified 103 articles from our literature search evaluating rates of new or worsening neurologic complications among patients undergoing spine surgery with neuromonitoring. From these potential articles, we judged 42 adequate to undergo full text review. After full text review, we excluded 10 of the articles for the following reasons: 9 articles did not report any of the following diagnostic test characteristics: sensitivity, specificity, PPV, or NPV. The other was a review article and did not report patient specific data, Figure 3. Table 1 summarizes the rates of new neurologic deficits and diagnostic test characteristics for each type of neuromonitoring mo-

Table 1. Summary of Rates of New or Worsening Neurological Injury and Diagnostic Test Characteristics

	SSEPs		TcMEPs		EMG			
Outcomes	No. Studies	Results (Range) (%)	No. Studies	Results (Range) (%)	No. Studies	Results (Range) (%)	Multimodal No. Studies	Results (Range) (%)
Rate of new neurological deficit	18	0.09–28.5	6	0.8–3.2	1	3.2	11	4.9–28.5
Sensitivity	15	0–100	6	81–100	1	46	11	70–100
Specificity	15	27–100	6	81–100	1	73	11	52.7*–100
Positive predictive value	9	15–100	3	17–96	1	3	3	5.2–100
Negative predictive value	9	95–100	3	97–100	1	97	3	96–100

*This represents an outlier. Nine of 11 studies reported specificity between 90% and 100%.

Table 2. Rating of Overall Strength of Evidence for Each Key Question

	Strength of Evidence	Conclusions/Comments	Quality	Quantity	Consistency
Question 1: What are the diagnostic test characteristics of unimodal (SSEPs or MEPs) and multimodal (combination) modalities?					
1. Unimodal SSEP and MEP	Very low	There is very low evidence from the literature that unimodal SSEPs or MEPs are valid diagnostic tests for measuring intraoperative neurological injury.	−	+	−
2. Multimodal	High	There is high evidence from the literature that multimodal neuromonitoring is sensitive and specific for measuring intraoperative neurological injury.	+	+	+
	Moderate	The evidence is moderate with respect to NPV.	+	−	+
	Low	The evidence is low with respect to PPV.	+	−	−
Question 2: Which modalities are superior with respect to diagnostic test characteristics in the same patient population?					
1. Unimodal SSEP *vs.* MEP	High	There is high evidence that unimodal MEPs outperform SSEPs and EMG with respect to test sensitivity, but are similar with respect to specificity. There were not enough studies assessing PPV and NPV for these comparisons.	+	+	+
2. Unimodal *vs.* multimodal	No evidence	There is no evidence from studies directly comparing unimodal with multimodal neuromonitoring.	None	None	None
Questions 3: Does intraoperative neuromonitoring reduce the rate of a new or worsening neurological event?					
1. Monitoring *vs.* no monitoring	Low	There is low evidence that intraoperative neuromonitoring reduces the rate of new or worsening neurological deficits. The quantity and consistency appeared to favor neuromonitoring; however, study quality was poor.	−	+	+
Question 4: Does the reaction to a positive intraoperative neuromonitoring alert reduce the rate of a new or worsening neurological event?					
1. Response *vs.* no response	Very low	There is very low evidence that an intraoperative response to a neuromonitoring alert reduces the rate of new or worsening neurological deficits since only one study was identified evaluating this comparison.	—	—	—
2. Remedial measure *vs.* no remedial measure	No evidence	There is no evidence from studies directly comparing specific intraoperative remedial measures.	None	None	None

dality for all studies reviewed in this manuscript. Details for each study can be found in Supplemental Digital Content 2, Figures, available at: http://links.lww.com/BRS/A425; Supplemental Digital Content 3, text, available at: http://links.lww.com/BRS/A426.

For Patients Undergoing Instrumented Spine Surgery With Intraoperative Neuromonitoring, What Are the Diagnostic Test Characteristics of Unimodal IOM (SSEPs or MEPs Independently) and MIOM (Combination) Methods?

The majority of studies in the literature evaluating IOM and MIOM techniques assess their diagnostic test characteristics (*e.g.*, sensitivity, specificity, PPV, NPV) in a series of patients. Diagnostic accuracy does not establish efficacy of neuromonitoring unless the method is compared with another method; however, it does provide a quantitative assessment of the technique's validity measured by sensitivity, specificity, PPV, and NPV, Table 2.

We identified 10 studies that were designed to evaluate the diagnostic characteristics of unimodal monitoring. Eight studies evaluated SSEPs and 2 studies evaluated TcMEPs. Four of the SSEP studies evaluated the diagnostic test characteristics in all spine procedures (cervical, thoracic, and lumbar),[21–24] 2 in only cervical procedures,[25,26] and 2 in only thoracic procedures.[27,28] Both TcMEPs studies evaluated the test characteristics in all spine procedures[29,30] (Supplemental Digital Content 1, Table 1, available at: http://links.lww.com/BRS/A424). All studies were retrospective cohort studies. Four of 10 evaluated a consecutive series of patients[22,23,25,26] and only 2 blinded the evaluator of new neurologic deficits from the intraoperative neuromonitoring findings.[25,26] The method for measuring new neurologic deficits varied across studies. The threshold for declaring a positive intraoperative neuromonitoring alert for SSEPs was a decrease in amplitude of >50% in

Table 3. Definitions for Diagnostic Tests Relative to Their Application With Intraoperative Neuromonitoring

Term	Definition
True positive	A true positive (TP) result equals the no. individuals with a disease who test positive. Examples from the literature: Any case in which the changes in evoked potentials were reversed immediately by an intervention or in which the changes persisted and the patient awoke with a neurological deficit.[7] An alert followed by observation of a new neurologic motor deficit during a wake-up test or at the end of the procedure.[17]
False positive	A false positive (FP) result equals the no. of individuals without a disease who test positive. Examples from the literature: Any case in which changes in the evoked potentials did not respond to an intervention and the patient awoke neurologically intact.[7] An alert that persisted despite corrective measures and the absence of any observable new deficit during a wake-up test (if one was performed) and/or the absence of a new deficit at the conclusion of the procedure (with persistent loss of potentials).[17]
True negative	A true negative (TN) result (cell d) equals the no. individuals without a disease who test negative. Examples from the literature: A case in which monitoring revealed no changes and there was no new postoperative deficit.[7] The absence of any alert in a patient without a new postoperative deficit.[17]
False negative	A false negative (FN) result equals the no. individuals with a disease who test negative. Examples from the literature: New onset of a neurological deficit in a patient who had had no change in the neurophysiological monitoring data or in whom the change had resolved to a value within 2.0 standard deviations of the original baseline value following an intervention or by the end of the surgical procedure.[7] The absence of an alert in a patient with a new postoperative motor deficit.[17]
Sensitivity	The probability of a positive intraoperative neuromonitoring test among patients with a postoperative neurological injury.
Specificity	The probability of a negative intraoperative neuromonitoring test among patients without a postoperative neurological injury.
Positive predictive value	The percentage of patients with a positive intraoperative neuromonitoring test who have a postoperative neurological injury.
Negative predictive value	The percentage of patients with a negative intraoperative neuromonitoring test who do NOT have a postoperative neurological injury.

Table 4. List of Consistent Remedial Measures Performed as a Result of a Positive Intraoperative Alert

Surgical	Physiological	Pharmaceutical
Dissection	Perfusion pressure	Vasospasm
Limitation of resection	Warm saline irrigation	Local instillation of paparverine
Vessel clipping	Induced hypertension	Dexamethasone intravenously
Retraction	Wake-up test	Volume and pharmacologic resuscitation
Abandonment of procedure	Relieving vena cava pressure	Verify anesthetic concentrations
Adjusting operative position	Patient warming to 36°C	High-dose methylprednisolone
Releasing distraction	Changing stimulation parameters and repositioning electrodes	
Retractor repositioning	Hypothermia (systemic or local)	
Removing or replacing hardware	Correction of anemia	
Bone graft removal	Correction of hypotension	
Surgical decompression		
Arm neural injury prevention		
Wound exploration		
Removal of hemostatic sponges and epidural packing agents		

all studies. The study by Langeloo *et al* evaluated 3 different TcMEP criteria for an intraoperative alert (A: 1 of 6 recordings >80% decrease in amplitude; B: 2 of 6 recordings >80% decrease in amplitude; C: 1 of 2 anterior tibial muscle recordings >80% decrease in amplitude[30]).

All SSEP studies reported sensitivity and specificity with the exception of 1 which only reported sensitivity.[27] Six of the 8 SSEP studies reported PPV, NPV, or both. Sensitivity ranged from 0%[27] to 100%.[23,24,28] Specificity ranged from 27%[26] to 100%.[25] The PPV ranged from 15%[22] to 100%.[25] Leung *et al* evaluated these characteristics in patients with a "normal" cord and those with a cord "at risk" with a greater PPV in those with a cord "at risk" (11.1% and 23.5%, respectively).[28] Both TcMEP studies reported sensitivity and specificity. Only Langeloo[30] reported the PPV and NPV. Lang *et al*[29] reported a sensitivity of 100% and specificity of 81%. For criteria A, Langeloo *et al* reported a sensitivity, specificity, PPV, and NPV of 100%, 91%, 61%, and 100%, respectively; for criteria B, 81%, 97%, 97%, and 76%, respectively; for criteria C, 88%, 95%, 98%,

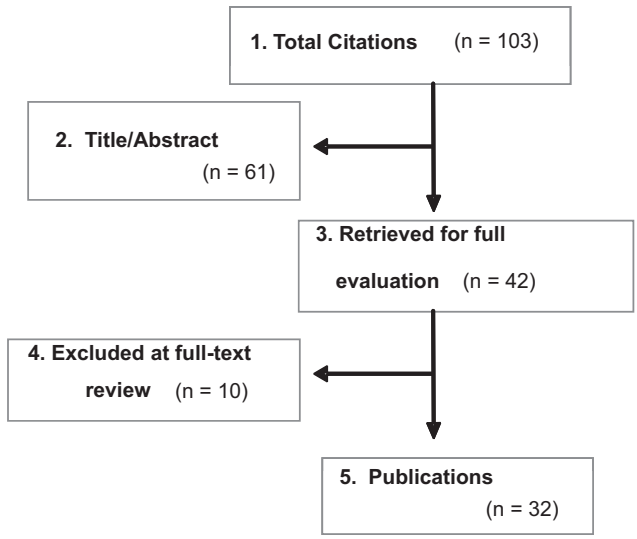

Figure 3. Flow chart showing results of literature search.

and 70%, respectively. Ranges for unimodal diagnostic test characteristics are summarized in Table 1.

The overall strength of the evidence for unimodal SSEP and MEP studies is VERY LOW, that is, any estimate of effect is very uncertain, based on the following summary of criteria: For SSEP studies, the quality was poor ("−"), the quantity was high ("+"), and the consistency was poor ("−"). For MEP studies, the quality was poor ("−"), the quantity was poor ("−"), and the consistency was high ("+"), Table 2.

We identified 11 studies designed to evaluate the diagnostic characteristics of MIOM techniques (Supplemental Digital Content 1, Table 2, available at: http://links.lww.com/BRS/A424). All 11 studies included MEPs in combination with another modality. Diagnostic characteristics were reported based on multimodal findings not by individual methods. Nine of these combined SSEPs with MEPs. Other methods included EMG and compound muscle action potentials (CMAPs). Four of these studies were pediatric evaluations,[8,31–33] three studies evaluated adult spine surgeries for various diagnoses (scoliosis, deformity multiple causes, and tumors),[34–37] 2 evaluated adult lumbar surgeries,[38,39] 1 evaluated adult cervical surgeries,[40] and 1 evaluated adult thoracic surgeries.[41] The methods for measuring a new postoperative neurologic deficit varied widely across studies with many not reporting specifically how a new deficit was measured or quantified. The thresholds for declaring a positive intraoperative neuromonitoring alert were less consistent than observed in the unimodal studies. None of these studies blinded the postoperative evaluator to the MIOM findings. MacDonald *et al* reported an alert as a focal decrement defined as an amplitude decrement unequivocally exceeding trial-to-trial variation for SSEPs and requiring disappearance of a response for MEPs. He compared diagnostic characteristics between protracted focal decrements (>40 minutes) and transient focal decrements (quickly resolved).[37]

All studies reported sensitivity and specificity with the exception of one.[42] Only the studies by Accadbled *et al*, MacDonald *et al*, and Quraishi *et al* reported the PPV and NPV. Sensitivity ranged from 70%[8] to 100%.[31,32,37,38] The other studies, with the exception of cervical surgery study reported by Eggspuehler *et al* (83.3%),[40] reported a sensitivity above 90%. Accadbled *et al* and Quraishi *et al* reported a specificity of 52.7%[31] and 84.3%,[37] respectively. The remaining 9 studies reported specificity at or above 90%. Accadbled *et al* reported a PPV of 5.4% and an NPV of 100%.[31] MacDonald *et al* reported a PPV of 100% and an NPV of 98% when using a protracted focal decrement and 13% and 96% when using a transient focal decrement. Quraishi *et al* reported a PPV of 13.9% and an NPV of 97% for all patients. In a subset of patients with corrective surgeries involving osteotomies, the sensitivity was less (67%) then the overall group; however, the PPV (80%) was greater. Lieberman *et al*, evaluating MIOM in patients undergoing lumbar osteotomy and fusion for fixed sagittal plane deformities, calculated the overall sensitivity and specificity and the muscle specific characteristics.[38] The sensitivity was weakest when assessing the tibialis anterior (50%) and strongest in the vastus medialis (90%). The specificity was weakest when assessing the vastus medialis (75%) and strongest in the adductor muscles (93%). However, when all muscles were relied on, the sensitivity was 100% and specificity was 90%. Ranges for multimodal diagnostic test characteristics are summarized in Table 1.

The overall strength of the evidence for multimodal studies with respect to sensitivity and specificity is HIGH, that is, further research is very unlikely to change our confidence in the estimate of effect, based on the following summary of criteria: the quality was high ("+"), the quantity was high ("+"), and the consistency was high ("+"), Table 2. For NPV it was MODERATE, meaning that further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate, as there were only 2 studies identified (despite high quality and consistency) and for PPV it was LOW, that is, further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate, as only two studies were identified and findings were inconsistent.

For Patients Undergoing Instrumented Spine Surgery With Intraoperative Neuromonitoring, Which Methods Are Superior With Respect to Diagnostic Test Characteristics in the Same Patient Population?

We identified 6 studies that compared different neuromonitoring methods in the same patients during the same surgical procedure with respect to diagnostic test characteristics (Supplemental Digital Content 1, Table 3, available at: http://links.lww.com/BRS/A424). Three studies evaluated cervical spine surgeries,[7,17,43] 1 evaluated thoracolumbar surgery in adults,[9] 1 evaluated pediatric scoliosis correction surgery,[44] and 1 evaluated various spine surgeries in adults.[34] These studies compared SSEPs, MEPs, and EMG. No studies were identified that compared multimodal techniques with unimodal techniques. All 6 studies reported sensitivity and specificity while two included the PPV and NPV. The sensitivities for SSEPs ranged from 0%[17] to 100%,[44] for EMG 46%[43] to 100%,[9] and for MEPs (both regular and transcranial) 100%.[7,17,34,43,44] The specificity for SSEPs ranged from 95%[9] to 100%,[7,17,43,44] for EMG 24%[9] to 73%,[43] and for MEPs (both regular and transcranial) 90%[17] to 100%.[7,34,44] Kelleher *et al* calculated a PPV for SSEPs of 100%, for MEPs 96%, and for EMG 3%. The NPV for SSEPs was 97%, for MEPs 100%, and for EMG 97%.[43] Gunnarson calculated a PPV of 8.5% for EMG and 28.6% for SSEPs; a NPV of 100% for EMG and 94.7% for SSEPs. This group suggests the low PPV and high sensitivity for EMG represents a tool that provides a warning in time before actual neural injury occurs sug-

gesting that a high PPV would lose its sensitivity to detect early injury and therefore not as useful for neuromonitoring. The study by Kim *et al* was designed to examine the performance of TcMEPs in patients undergoing surgery for cervical myelopathy in an effort to identify risk factors for false-positive results.[17] They also evaluated SSEPs and reported that the false positive rate of TcMEP monitoring in cervical myelopathy patients may be as high as 83%, and the PPV of an isolated TcMEP alert without a corresponding SSEP alert may be as low as 17%. They concluded that in these instances, the surgeon is justified in continuing surgery without performing a wake-up test or initiating a spinal cord injury steroid protocol.

The overall strength of the evidence for comparing unimodal methods is HIGH, meaning further research is very unlikely to change our confidence in the estimate of effect, based on the following summary of criteria: the quality was high ("+"), the quantity was high ("+"), and the consistency was high ("+"), Table 2. This suggests that unimodal MEPs outperform SSEPs and EMG with respect to test sensitivity. SSEPs and MEPs are similar with respect to specificity. There were not enough studies assessing PPV and NPV for these comparisons. There were no comparisons between multimodal and unimodal to suggest superiority of multimodal over unimodal neuromonitoring.

For Patients Undergoing Instrumented Spine Surgery, Does Intraoperative Neuromonitoring Reduce the Rate of a New or Worsening Neurologic Event?

To establish the efficacy of intraoperative neuromonitoring, patients treated with neuromonitoring should be compared with patients who do not receive neuromonitoring with respect to an outcome such as a new postoperative neurologic deficit or postoperative neurologic changes. We identified 4 observational studies that compared patients with and without neuromonitoring (Supplemental Digital Content 1, Table 4, available at: http://links.lww.com/BRS/A424). No randomized trials or systematic reviews were identified. Three of the 4 studies comparing monitoring *versus* no monitoring used historical control groups from the same institution during periods before neuromonitoring were initiated.

Two studies evaluated neuromonitoring for cervical procedures which included various degenerative diseases including stenosis, myelopathy, spondylosis, ossified posterior longitudinal ligaments, and nonunion.[45,46] The other 2 studies evaluated adult thoracolumbar patients[47] and spinal cord tumors for all areas of the spine,[16] respectively. Sala *et al* evaluated a combination of SSEPs and MEPs (*i.e.*, multimodal monitoring) whereas Smith *et al* and Epstein *et al* evaluated SSEPs only. Sala *et al* and Epstein *et al* measured efficacy by a change in McCormick and Ranawat classification schemes, respectively. Both found better improvement in patients who underwent IOM testing. Sala *et al* reported that the MIOM group had an overall improvement in neurologic status whereas the control group exhibited an overall deterioration. Similar findings were reported by Epstein *et al* who reported a 5.4% neurologic deterioration rate in the control group (compared with 1% in the unimodal SSEP group) and a 3.7% (n = 8 quadriplegics) rate of quadriplegia in the control group (compared with 0% in the unimodal SSEP group). Smith *et al* reported only one neurologic deficit in 1039 subjects and this occurred in the monitored group. No neurologic deficits were reported in the control group. The monitored group displayed six transient SSEP changes which resolved after intervention. Both Sala *et al* and Epstein *et al* concluded that intraoperative neuromonitoring was a practical and useful tool for their respective cervical surgeries. Smith *et al* noted that an intraoperative neurologic deficit is possible despite normal SSEP signals. Meyer *et al* evaluated patients with acute spinal cord injury to the thoracolumbar region.[47] Patients monitored with intraoperative SSEPs (n = 150) were compared to patients who underwent wakeup tests or no monitoring (n = 145). Monitoring was performed in patients with incomplete spinal cord injuries whereas no monitoring was performed in subjects treated before monitoring was used in the institution and in patients with complete spinal cord injuries. Postoperative neurologic changes were assessed using the Trauma Motor Index and the Frankel Grade. The presence of new neurologic deficits was greater in the control group (6.9%) than the monitored group (0.7%).

These findings indicate that intraoperative neuromonitoring may be effective; however, the overall strength of the evidence to suggest that intraoperative neuromonitoring reduces the rate of new or worsening neurologic deficits is LOW, that is, further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate, based on the following summary of criteria: the quality was low ("−"), the quantity was high ("+"), and the consistency was high ("+"), Table 2. The quality of the evidence from these studies is poor because of the retrospective nature, historical control groups, potential for uncontrolled selection bias, and lack of blinding of intraoperative neuromonitoring findings when performing postoperative neurologic evaluations.

For Patients Undergoing Instrumented Spine Surgery With Intraoperative Neuromonitoring, Does the Reaction to a Positive Intraoperative Neuromonitoring Alert Reduce the Rate of a New or Worsening Neurologic Event? Are There Specific Intraoperative Remedial Measures Employed as a Result of a Positive Alert More Effective Than Others?

The only comparative study identified evaluating the effects of responding to an intraoperative alert was by Wiedemayer *et al*[48] (Supplemental Digital Content, Table 4, available at: http://links.lww.com/BRS/A424). Al-

though the study by Wiedemayer *et al* examined a mixed series of patients undergoing cranial and spinal procedures, this study was judged by the authors of this systematic review to be the most useful article with respect to establishing the efficacy of intraoperative neuromonitoring. The importance of this article is based on the fact that it compared the rate of a new neurologic deficit between patients in which the surgeon reacted with an intraoperative alert and performed an intervention to patients where surgeons did not react to an intraoperative alert. A new neurologic deficit was judged by a surgeon who was blinded to the neuromonitoring results. The overall rate of a new neurologic deficit was 20% (n = 84 of 423). The overall rate of an intervention among all patients was 10% (n = 42 of 423). The rate of a new neurologic deficit among those patients who received an intervention was 4.7% and among those who did not receive an intervention was 15.1%. The authors reported that those patients who did not receive an intervention may have had alerts that did not warrant an intervention. The authors reported that 5.2% of the patients benefited from monitoring based on those patients who awoke without neurologic deficits (n = 22 of 423) when there was a positive intraoperative alert.

There were no studies evaluating the efficacy of specific intraoperative or remedial measures performed in response to an intraoperative neuromonitoring alert. The study by Wiedemeyer *et al* was not designed to evaluate a specific remedial measure; rather, it combined all measures together and evaluated the rate of new postoperative deficits in patients who received a remedial measure of any kind compared to no remedial measures. Nearly all studies in this systematic review described the various responses to intraoperative alerts; however, because of the various responses to the various conditions, it is not possible to evaluate any one type of response or remedial measure. The various intraoperative responses from the reviewed studies are listed in Table 4.

The overall strength of the evidence to suggest that responding to an intraoperative neuromonitoring alert reduces the rate of a new or worsening neurologic deficit is VERY LOW, that is, any estimate of effect is very uncertain, based on the following summary of criteria: the quality was low ("−"), the quantity was low ("−"), and the consistency was low ("−") because only 1 study addressing this question was identified, Table 2. The study by Weidemayer *et al* appeared to indicate that patients who have a positive alert with an intervention have a lower rate of a new neurologic deficit; however, selection bias (without control of possible confounding) between those who received an intervention and those who did not cannot be ruled out.

■ Conclusion

The purpose of this literature review was to answer the following questions regarding the use of intraoperative neuromonitoring of patients undergoing instrumented spine surgery: What are the diagnostic test characteristics (*i.e.*, sensitivity, specificity, positive predictive value, negative predictive value) of unimodal IOM (SSEPs or MEPs independently) and MIOM (combination) methods? Which methods are superior with respect to diagnostic test characteristics in the same patient population? Does intraoperative neuromonitoring reduce the rate of a new or worsening neurologic event? Does the reaction to a positive intraoperative neuromonitoring alert reduce the rate of a new or worsening neurologic event? Are there specific intraoperative remedial measures employed as a result of a positive alert more effective than others?

There is *very low evidence* from the literature supporting unimodal SSEPs or MEPs as valid diagnostic tests for measuring intraoperative neurologic injury; however, there is *high evidence* that multimodal neuromonitoring is sensitive and specific for detecting intraoperative neurologic injury. The evidence for multimodal monitoring is *moderate* with respect to NPV and *low* with respect to PPV. There is *high evidence* that unimodal MEPs outperform SSEPs and EMG with respect to test sensitivity, but are similar with respect to specificity. There were not enough studies assessing PPV and NPV for these comparisons. There is *no evidence* from studies directly comparing unimodal with multimodal neuromonitoring. There is *low evidence* that intraoperative neuromonitoring reduces the rate of new or worsening neurologic deficits. The quantity and consistency appeared to favor neuromonitoring; however, study quality was poor.

There is *very low evidence* that an intraoperative response to a neuromonitoring alert reduces the rate of new or worsening neurologic deficits. Only one study was identified evaluating this comparison and was judged by the authors as potentially the most useful with respect to establishing the efficacy of intraoperative neuromonitoring because it compared the rate of a new neurologic deficit between patients in which the surgeon reacted to an intraoperative alert and performed an intervention to patients where surgeons did not react to an intraoperative alert.[48] The rate of a new neurologic deficit among those patients who received an intervention was 4.7% and among those who did not receive an intervention was 15.1%. The authors reported that those patients who did not receive an intervention may have had alerts that did not warrant an intervention. The authors reported that 5.2% of the patients benefited from monitoring based on those patients who awoke without neurologic deficits when there was a positive intraoperative alert. To date, there is *no evidence* from studies directly comparing specific intraoperative remedial measures, although the preclinical animal literature has reported a number of promising approaches to treat spinal cord injury including sodium-glutamate blockers such as riluzole, anti-inflammatory drugs (minocycline) and inhibitors of Rho (Cethrin).[49] As previously reported by

the senior author (MGF), a number of approaches have been used in the clinical setting in an effort to mitigate the impact of perioperative spinal cord injury, including hypertensive therapy, corticosteroids, CSF drainage, minimization of distraction or retraction of neural tissue, and avoidance of hyperthermia.[1] However, there is a lack of evidence to support these approaches. Clearly, further research is required to optimize protocols for the prevention and treatment of perioperative spinal cord injury.

The accuracy of a diagnostic test consists of 2 general components: the accuracy of classifying patients with respect to their disease status (validity), and the degree to which repeated measures yield the same results (reliability). However, regardless of how accurate or predictive a test may be, health policy and public health perspectives assert that a diagnostic test should only be performed if it leads to the use of interventions that are likely to improve patient outcomes or if it prevents the use of interventions that are not likely to improve outcomes.[50] Sensitivity and specificity are the traditional measures of diagnostic tests used in validation to describe the accuracy of classification. They do not, however, describe the probability that a patient actually has the disease if the test is positive or does not have it if the test is negative. These are positive and negative predictive value, respectively. These values are important to consider when employing intraoperative neuromonitoring as they can help guide decision-making with respect to false positives and negatives. However, if one is trying to decide if neuromonitoring is necessary, it really requires a comparison of postoperative neurologic injuries in patients who do and do not receive neuromonitoring. Further, in those who do receive neuromonitoring, is there a difference in neurologic injury for those in which a positive alert led to a remedial measure *versus* positive alerts where no response is rendered? These should be further explored in future studies to assist the surgeon in decision-making.

In conclusion, our systematic review of the literature indicates that there is strong evidence that multimodality neuromonitoring is sensitive and specific for detecting intraoperative neurologic injury during spine surgery. Although the level of evidence supporting the conclusion that intraoperative neuromonitoring reduces the rate of new or worsening neurologic deficits is low, the available evidence does consistently support this conclusion. There is a need to define evidence-based protocols to deal with intraoperative changes in neuromonitoring and to prospectively evaluate the impact of these protocols.

Clinical Recommendations

Based on the available evidence, it is recommended that the use of multimodality neuromonitoring be considered in complex spine surgery where the spinal cord or nerve roots are deemed to be at risk, including procedures involving deformity correction and some procedures that require placement of instrumentation.

■ Key Points

- Based on strong evidence that multimodality intraoperative neuromonitoring (MIOM) is sensitive and specific for detecting intraoperative neurologic injury during spine surgery, it is recommended that the use of MIOM be considered in spine surgery where the spinal cord or nerve roots are deemed to be at risk, including procedures involving deformity correction and some procedures that require placement of instrumentation.
- There is a need to develop evidence-based protocols to deal with intraoperative changes in MIOM and to validate these prospectively.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.spinejournal.com).

Acknowledgments

The authors are indebted to Ms. Nancy Holmes, RN, for her administrative assistance, and to Mr. Jeff Hermsmeyer, BS, for his assistance in searching the literature, abstracting data, and proofing.

References

1. Ahn H, Fehlings MG. Prevention, identification, and treatment of perioperative spinal cord injury. *Neurosurg Focus* 2008;25:E15.
2. Pajewski TN, Arlet V, Phillips LH. Current approach on spinal cord monitoring: the point of view of the neurologist, the anesthesiologist and the spine surgeon. *Eur Spine J* 2007;16:10.
3. Vauzelle C, Stagnara P, Jouvinroux P. Functional monitoring of spinal cord activity during spinal surgery. *Clin Orthop Relat Res* 1973;93:173.
4. Magit DP, Hilibrand AS, Kirk J, et al. Questionnaire study of neuromonitoring availability and usage for spine surgery. *J Spinal Disord Tech* 2007; 20:282.
5. Dawson EG, Sherman JE, Kanim LE, et al. Spinal cord monitoring. Results of the Scoliosis Research Society and the European Spinal Deformity Society survey. *Spine* 1991;16:S361.
6. Nuwer MR, Dawson EG, Carlson LG, et al. Somatosensory evoked potential spinal cord monitoring reduces neurologic deficits after scoliosis surgery: results of a large multicenter survey. *Electroencephalogr Clin Neurophysiol* 1995;96:6.
7. Hilibrand AS, Schwartz DM, Sethuraman V, et al. Comparison of transcranial electric motor and somatosensory evoked potential monitoring during cervical spine surgery. *J Bone Joint Surg Am* 2004;86:1248.
8. MacDonald DB, Al Zayed Z, Khoudeir I, et al. Monitoring scoliosis surgery with combined multiple pulse transcranial electric motor and cortical somatosensory-evoked potentials from the lower and upper extremities. *Spine* 2003;28:194.
9. Gunnarsson T, Krassioukov AV, Sarjeant R, et al. Real-time continuous intraoperative electromyographic and somatosensory evoked potential recordings in spinal surgery: correlation of clinical and electrophysiologic findings in a prospective, consecutive series of 213 cases. *Spine* 2004;29:677.
10. Lesser RP, Raudzens P, Luders H, et al. Postoperative neurological deficits may occur despite unchanged intraoperative somatosensory evoked potentials. *Ann Neurol* 1986;19:22.
11. Minahan RE, Sepkuty JP, Lesser RP, et al. Anterior spinal cord injury with preserved neurogenic 'motor' evoked potentials. *Clin Neurophysiol* 2001; 112:1442.
12. Deletis V, Sala F. Intraoperative neurophysiological monitoring of the spinal

cord during spinal cord and spine surgery: a review focus on the corticospinal tracts. *Clin Neurophysiol* 2008;119:248.
13. Iwasaki H, Tamaki T, Yoshida M, et al. Efficacy and limitations of current methods of intraoperative spinal cord monitoring. *J Orthop Sci* 2003;8:635.
14. Owen JH. The application of intraoperative monitoring during surgery for spinal deformity. *Spine* 1999;24:2649.
15. Sala F, Krzan MJ, Deletis V. Intraoperative neurophysiological monitoring in pediatric neurosurgery: why, when, how? *Childs Nerv Syst* 2002;18:264.
16. Sala F, Palandri G, Basso E, et al. Motor evoked potential monitoring improves outcome after surgery for intramedullary spinal cord tumors: a historical control study. *Neurosurgery* 2006;58:1129.
17. Kim DH, Zaremski J, Kwon B, et al. Risk factors for false positive transcranial motor evoked potential monitoring alerts during surgical treatment of cervical myelopathy. *Spine* 2007;32:3041.
17a. Dettori JR, Norvell DC, Dekutoski M, et al. Methods for systematic reviews on patient safety during spine surgery. *Spine* 2010;35:S22–S27.
18. Wright JG, Swiontkowski MF, Heckman JD. Introducing levels of evidence to the journal. *J Bone Joint Surg Am* 2003;85:1.
19. van Tulder M, Furlan A, Bombardier C, et al. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine* 2003;28:1290.
20. West S, King V, Carey TS, et al. *Systems to Rate the Strength of Scientific Evidence*. Rockville, MD: Agency for Healthcare Research and Quality; 2002. Evidence Report/Technology Assessment No. 47 (Prepared by the Research Triangle Institute-University of North Carolina Evidence-based Practice Center, Contract No. 290–97–0011).
21. Dinner DS, Luders H, Lesser RP, et al. Intraoperative spinal somatosensory evoked potential monitoring. *J Neurosurg* 1986;65:807.
22. Manninen PH. Monitoring evoked potentials during spinal surgery in one institution. *Can J Anaesth* 1998;45:460.
23. Papastefanou SL, Henderson LM, Smith NJ, et al. Surface electrode somatosensory-evoked potentials in spinal surgery: implications for indications and practice. *Spine* 2000;25:2467.
24. Stechison MT, Panagis SG, Reinhart SS. Somatosensory evoked potential. Monitoring during spinal surgery. *Acta Neurochir (Wien)* 1995;135:56.
25. Khan MH, Smith PN, Balzer JR, et al. Intraoperative somatosensory evoked potential monitoring during cervical spine corpectomy surgery: experience with 508 cases. *Spine* 2006;31:E105.
26. May DM, Jones SJ, Crockard HA. Somatosensory evoked potential monitoring in cervical surgery: identification of pre- and intraoperative risk factors associated with neurological deterioration. *J Neurosurg* 1996;85:566.
27. Deutsch H, Arginteanu M, Manhart K, et al. Somatosensory evoked potential monitoring in anterior thoracic vertebrectomy. *J Neurosurg* 2000;92:155.
28. Leung YL, Grevitt M, Henderson L, et al. Cord monitoring changes and segmental vessel ligation in the "at risk" cord during anterior spinal deformity surgery. *Spine* 2005;30:1870.
29. Lang EW, Beutler AS, Chesnut RM, et al. Myogenic motor-evoked potential monitoring using partial neuromuscular blockade in surgery of the spine. *Spine* 1996;21:1676.
30. Langeloo DD, Lelivelt A, Louis Journee H, et al. Transcranial electrical motor-evoked potential monitoring during surgery for spinal deformity: a study of 145 patients. *Spine* 2003;28:1043.
31. Accadbled F, Henry P, de Gauzy JS, et al. Spinal cord monitoring in scoliosis surgery using an epidural electrode. Results of a prospective, consecutive series of 191 cases. *Spine* 2006;31:2614.
32. Hsu B, Cree AK, Lagopoulos J, et al. Transcranial motor-evoked potentials combined with response recording through compound muscle action potential as the sole modality of spinal cord monitoring in spinal deformity surgery. *Spine* 2008;33:1100.
33. Padberg AM, Wilson-Holden TJ, Lenke LG, et al. Somatosensory- and motor-evoked potential monitoring without a wake-up test during idiopathic scoliosis surgery. An accepted standard of care. *Spine* 1998;23:1392.
34. Calancie B, Harris W, Brindle GF, et al. Threshold-level repetitive transcranial electrical stimulation for intraoperative monitoring of central motor conduction. *J Neurosurg* 2001;95:161.
35. Eggspuehler A, Sutter MA, Grob D, et al. Multimodal intraoperative monitoring during surgery of spinal deformities in 217 patients. *Eur Spine J* 2007;16:14.
36. Sutter M, Eggspuehler A, Grob D, et al. The diagnostic value of multimodal intraoperative monitoring (MIOM) during spine surgery: a prospective study of 1,017 patients. *Eur Spine J* 2007;16:31.
37. Quraishi NA, Lewis SJ, Kelleher MO, et al. Intraoperative multimodality monitoring in adult spinal deformity: analysis of a prospective series of one hundred two cases with independent evaluation. *Spine* 2009;34:1504.
38. Lieberman JA, Lyon R, Feiner J, et al. The efficacy of motor evoked potentials in fixed sagittal imbalance deformity correction surgery. *Spine* 2008;33:E414.
39. Sutter MA, Eggspuehler A, Grob D, et al. Multimodal intraoperative monitoring (MIOM) during 409 lumbosacral surgical procedures in 409 patients. *Eur Spine J* 2007;16:3.
40. Eggspuehler A, Sutter MA, Grob D, et al. Multimodal intraoperative monitoring (MIOM) during cervical spine surgical procedures in 246 patients. *Eur Spine J* 2007;16:4.
41. Eggspuehler A, Sutter MA, Grob D, et al. Multimodal intraoperative monitoring (MIOM) during surgical decompression of thoracic spinal stenosis in 36 patients. *Eur Spine J* 2007;16:4.
42. Wilson-Holden TJ, Padberg AM, Lenke LG, et al. Efficacy of intraoperative monitoring for pediatric patients with spinal cord pathology undergoing spinal deformity surgery. *Spine* 1999;24:1685.
43. Kelleher MO, Tan G, Sarjeant R, et al. Predictive value of intraoperative neurophysiological monitoring during cervical spine surgery: a prospective analysis of 1055 consecutive patients. *J Neurosurg Spine* 2008;8:215.
44. Schwartz DM, Auerbach JD, Dormans JP, et al. Neurophysiological detection of impending spinal cord injury during scoliosis surgery. *J Bone Joint Surg Am* 2007;89:2440.
45. Epstein NE, Danto J, Nardi D. Evaluation of intraoperative somatosensory-evoked potential monitoring during 100 cervical operations. *Spine* 1993;18:737.
46. Smith PN, Balzer JR, Khan MH, et al. Intraoperative somatosensory evoked potential monitoring during anterior cervical discectomy and fusion in nonmyelopathic patients–a review of 1,039 cases. *Spine J* 2007;7:83.
47. Meyer PR Jr, Cotler HB, Gireesan GT. Operative neurological complications resulting from thoracic and lumbar spine internal fixation. *Clin Orthop Relat Res* 1988;125.
48. Wiedemayer H, Fauser B, Sandalcioglu IE, et al. The impact of neurophysiological intraoperative monitoring on surgical decisions: a critical analysis of 423 cases. *J Neurosurg* 2002;96:255.
49. Baptiste DC, Fehlings MG. Emerging drugs for spinal cord injury. *Expert Opin Emerg Drugs* 2008;13:63.
50. Weiss NS. *Clinical Epidemiology: The Study of the Outcome of Illness* 3rd ed. New York: Oxford University Press; 2006:178.
51. Armstrong B, White E, Saracci R. *Principles of Exposure Measurement in Epidemiology*. Oxford University Press; 1992.
52. Koepsell T, Weiss N. *Epidemiologic Methods: Studying the Occurrence of Illness*. New York: Oxford University Press, Inc.; 2003.
53. Alberg AJ, Park JW, Hager BW, et al. The use of "overall accuracy" to evaluate the validity of screening or diagnostic tests. *J Gen Intern Med* 2004;19:460.