Agreement Between Orthopedic Surgeons and Neurosurgeons Regarding a New Algorithm for the Treatment of Thoracolumbar Injuries

A Multicenter Reliability Study

Y. Raja Rampersaud, MD, FRCSC, Charles Fisher, MD, MHSc,† Jared Wilsey, PhD,‡ Paul Arnold, MD, PhD,§ Neel Anand, MD,‖ Chris M. Bono, MD,¶ Andrew T. Dailey, MD,# Marcel Dvorak, MD,† Michael G. Fehlings, MD, PhD,** James S. Harrop, MD,†† F. C. Oner, MD, PhD,‡‡ and Alexander R. Vaccaro, MD§§*

Introduction: Considerable variability exists in the management of thoracolumbar (TL) spine injuries. Although there are many influences, one significant factor may be the treating surgeon's specialty and training (ie, orthopedic surgery vs. neurosurgery). Our objective was to assess the agreement between spinal orthopedic and neurologic surgeons in rating the severity of TL spine injuries with a new treatment algorithm. This information could be important in establishing consensus-based protocols for managing these challenging injuries.

Methods: Twenty-eight spinal surgeons (8 neurosurgeons and 20 orthopedic surgeons) reviewed 56 TL injury case histories. Each case was classified and scored according to the TL injury severity score (TLISS). The case histories were reordered and the physicians repeated the exercise 3 months later. At both intervals the surgeons were asked if they agreed with the final treatment recommendation of the TLISS algorithm. The reliability and decision validity of the TLISS was compared.

Results: Between-group interrater reliability was similar to within group reliabilities. Intrarater reliability was also similar between groups. The between speciality interrater reliability of the TLISS management recommendation was moderate (74% agreement, $\kappa = 0.532$). Orthopedic and neurosurgeons agreed with the TLISS management recommendation 91.4% and 94.4% of the time, respectively.

Conclusions: The TLISS demonstrated good reliability in terms of intraobserver and interobserver agreement on the algorithmic treatment recommendations. The recommendation for operation seems to be consistent between fellowship-trained orthopedic and neurosurgical spine surgeons. This type of classification system may reduce the existing variability and initial management decision for treatment of TL injuries.

Key Words: thoracolumbar injury classification, reliability, orthopedic, neurosurgery

(*J Spinal Disord Tech* 2006;19:477–482)

Few areas in the treatment of spinal disorders are as controversial as the management of thoracolumbar (TL) trauma. This is illustrated by the variability in treatment approaches for one specific injury (Grauer et al[1]). In an attempt to reduce variability and improve outcome Vaccaro et al[2,3] introduced a TL injury classification and severity score (TLISS) and a recent modification. This system attempts to improve on existing TL injury classification systems by simplifying diagnostic variables and categorically directing treatment. Although in theory this has addressed some of the uncertainty and limitations of previous classifications, it remains to be seen if other factors influencing treatment decision-making have been addressed; most specifically, surgeon training and practice setting.[4–13]

One of the most dramatic areas of change over the last decade has been the evolution of an interdisciplinary dimension in the treatment of people with spine disorders where specialists now move through the boundaries of their governing disciplines. Both orthopedic and neurosurgeons manage spinal injuries, yet their training is different coming from 2 distinct disciplines; neurosurgery

Received for Publication February 25, 2006; accepted May 14, 2006.
From the *Division of Orthopaedic and Neurosurgery; **Division of Neurosurgery, University of Toronto, Toronto, Ontario; †Division of Spine, University of British Columbia, Vancouver, British Columbia, Canada; ‡Medtronic Sofamor Danek, Memphis, TN; §Division of Neurosurgery, Kansas University, Kansas City, KS; ‖Cedars Sinai Medical Center, Los Angeles, CA; ¶Department of Orthopaedic Surgery, Boston University Medical Center, Boston, MA; #Rocky Mountain Neurosurgical Alliance, Englewood, CO; ††Department of Neurosurgery; §§Department of Orthopaedic Surgery, Thomas Jefferson University, Philadelphia, PA; and ‡‡Department of Orthopedics, University Hospital Utrecht, Utrecht, The Netherlands.
Supported by members of the Spine Trauma Study Group and funded by an educational grant from Medtronic Sofamor Danek.
Reprints: Y. Raja Rampersaud, MD, FRCSC, Toronto Western Hospital, University Health Network, 399 Bathurst St, East Wing 1-441, Toronto, Ontario, Canada, M5T-2S8 (e-mail: raja.rampersaud@uhn.on.ca).

and orthopedic residency training programs, where there is usually minimal integration. It is critical therefore that new classifications or treatment paradigms are reliable for both neuro and orthopedic spine surgeons to reduce treatment variability in an already highly controversial treatment arena. Hence the objective of this study was to assess the degree of agreement between spinal orthopedic surgeons and neurosurgeons on the decision for management of TL injuries using the TLISS system. This information could be important in establishing consensus-based protocols for managing these challenging spinal injuries.

METHODS

Subjects

Twenty-eight academic spinal surgeons, all members of an international Spine Trauma Study Group (STSG) who participated in the development of the TLISS classification system, reviewed 56 clinical TL injury case histories and relevant imaging studies. Surgeons were provided a brief clinical vignette that included the patient's age, description of the traumatic injury, and neurologic status. Representative imaging (combinations of x-ray, computed tomography and sagittal T2 magnetic resonance images) was provided. Each case was classified and scored to determine treatment recommendations according to the TLISS system (Table 1). After 3 months, the case histories were presented in a different order and the physicians repeated the exercise. For interval 1, there were 20 orthopedic and 8 neurosurgeons. For interval 2, there were 16 orthopedic and 6 neurosurgeons. At both assessments surgeons were asked if they independently agreed or disagreed with the final TLISS treatment recommendation.

TABLE 1. Thoracolumbar Injury Severity Score

	Qualifier	Points
I. Mechanism of injury		
Compression	—	1
	Lateral angulation > 15 degrees	1
	Burst	1
Translational/rotational	—	3
Distraction	—	4
II. Neurologic involvement		
Intact	—	0
Nerve root	—	2
Cord (including the conus medullaris)	Incomplete	3
	Complete	2
Cauda equina	—	3
III. Posterior ligamentous complex		
Intact	—	0
Injury suspected/indeterminate	—	2
Injured	—	3

TL injuries are assigned points in each of the 3 subsections. An aggregate score is of ≤3 points recommends the injury should be managed nonoperatively and ≥5 points operatively. Injuries scoring 4 points are intermediate and can be managed operatively or nonoperatively.

TLISS Classification

The TLISS is based upon 3 major categories with points assigned for each specific variable within a category (Table 1): (1) the mechanism of injury determined primarily from the observed or imaged fracture pattern (1 to 4 points), (2) the integrity of the posterior ligamentous complex (PLC) (0 to 3 points), and (3) the neurologic status of the patient (0 to 3 points).[2] The assigned points are summated and a management recommendation for nonoperative treatment is given for injuries scoring ≤3 points and operative for those with ≥5 points. Injuries scoring 4 points are intermediate and can be managed either way.

Statistical Methods

The reliability and decision validity of the TLISS were compared between groups. The data were analyzed using SPSS and Analyze IT to determine percent agreement, linear weighted and unweighted Cohen κ, and Spearman rank order correlation. The Cohen κ value was defined as the observer agreement (Pa) minus the chance agreement (Pc) divided by the maximum possible agreement that is not related to chance (1-Pc): $\kappa = (Pa-Pc)/(1-Pc)$. The κ values obtained may range from (−) 1.0 (complete disagreement) through 0 (chance agreement) to 1.0 (perfect agreement).[14–16] All interval 1 data were used to calculate interrater reliability. The 16 orthopedic surgeons and 6 neurosurgeons with intervals 1 and 2 data were used for intrarater statistics. For significance tests, all unweighted coefficients were converted into Fisher z-scores, and the difference in z-scores was divided by standard error. A level was set at 0.05 ($\Delta Z/SE \geq 1.96$). A power analysis was also conducted, using a bench mark of interrater agreement of 75% or greater *between* orthopedic and neurosurgeons on the critical management component.

RESULTS

Overall Agreement Between Orthopedic and Neurosurgeons With the TLISS Management Recommendation

The overall agreement between specialties with the TLISS management recommendation for surgical versus nonsurgical approach was 92% (91.1%-orthopedic vs. 94.4%-neurosurgeons).

Interrater Analysis Between Orthopedic and Neurosurgeons (Table 2)

The interrater reliability of the TLISS assessment between neurosurgeons and orthopedic surgeons was evaluated for each category of the TLISS. The mechanism subscore, interrater reliability between groups was fair (Cohen κ = 0.317) with 52.2% agreement. The neurologic status subscore (neurologic findings were given) interrater reliability between groups was almost perfect (Cohen κ = 0.927) with 95.8% agreement. The PLC integrity subscore, interrater reliability between groups was fair (Cohen κ = 0.347) with 60.5% agreement. The total

TABLE 2. Interrater Agreement Within and Between Orthopedic and Neurosurgeons

	Percent Agreement	Cohen κ	Weighted κ	Spearman Rank Order
Mechanism				
Neurosurgeons (within group)	54.0*	0.341	0.318	0.433*
Orthopedic surgeons (within group)	50.8	0.294	0.291	0.354
Between neurosurgeons and orthopedic surgeons	52.2	0.317	0.295	0.395
Neurologic status				
Neurosurgeons (within group)	94.6	0.906	0.952	0.975
Orthopedic surgeons (within group)	97.2*	0.95	0.975	0.988*
Between neurosurgeons and orthopedic surgeons	95.8	0.927	0.963	0.981
PLC integrity				
Neurosurgeons (within group)	59.8	0.348	0.457	0.546
Orthopedic surgeons (within group)	62.8*	0.371	0.455	0.534
Between neurosurgeons and orthopedic surgeons	60.5	0.347	0.44	0.524
Total TLISS				
Neurosurgeons (within group)	35.8	0.27	0.531	0.717
Orthopedic surgeons (within group)	34.4	0.255	0.540	0.709
Between neurosurgeons and orthopedic surgeons	35.0	0.263	0.529	0.713
Management				
Neurosurgeons (within group)	75.0	0.555	0.515	0.545
Orthopedic surgeons (within group)	74.5	0.538	0.496	0.521
Between neurosurgeons and orthopedic surgeons	73.9	0.532	0.491	0.520

*$P < 0.05$ for difference between orthopedic and neurosurgeons.

TLISS score interrater reliability between groups was fair (Cohen κ = 0.263) with 35.0% agreement. The TLISS management recommendation interrater reliability between groups was moderate (Cohen κ = 0.532) with 73.9% agreement. The between-group interrater reliability was similar to the within group reliabilities (Table 2), suggesting that neurosurgeons and orthopedic surgeons agree with each other as consistently as they agree among themselves. Among the 3 subscores, orthopedic and neurosurgeons agreed the most on neurologic status and the least on mechanism.

Interrater Analysis Within Orthopedic and Neurosurgeons (Table 2)

The mechanism subscore, interrater reliability and percent agreement were slightly higher among neurosurgeons (N) versus orthopedic (O) surgeons [agreement of N-54% vs. O-50.8% ($P = 0.02$) and Cohen κ of 0.341 vs. 0.294]. Orthopedic surgeons, on the other hand, had slightly higher interrater reliability and percent agreement for neurologic status [agreement of N-94.6% vs. O-97.2% ($P < 0.0001$) and Cohen κ of 0.906 vs. 0.95] and PLC integrity [agreement of 59.8% vs. 62.8% ($P = 0.02$) and Cohen κ of 0.348 vs. 0.371]. There was no significant difference between neurosurgeons versus orthopedic surgeons for the total TLISS score interrater reliability [agreement of 35.8% vs. 34.4% ($P = 0.3$) and Cohen κ of 0.27 vs. 0.255] and with regard to the extent of agreement with the management recommendation [agreement of 75.0% vs. 74.5% ($P = 0.6$) and Cohen κ of 0.555 vs. 0.538].

Intrarater Analysis Within Orthopedic and Neurosurgeons (Table 3)

Orthopedic surgeons had greater intrarater reliability in all subcategories and TLISS score, as shown in Table 3. The difference in percent agreement was statistically significant for neurologic status, PLC integrity, and the TLISS score. The difference was most profound in rating PLC integrity, where, on intervals 1 and 2, neurosurgeons gave the same answer 61.4% of the time and orthopedic surgeons did so 71% of the time ($P = 0.001$). This corresponds to Cohen κ scores of 0.38 and 0.51, respectively. There was no significant difference between neurosurgeons versus orthopedic surgeons in agreement with the management recommendation [agreement of 76.1% vs. 79.1% ($P = 0.3$) and Cohen κ of 0.59 vs. 0.63].

Power

On the basis of an earlier assessment of the TLISS reliability (Harrop et al[15]), we established an interrater agreement of 75% or greater *between* orthopedic and neurosurgeons on the critical management component would be desirable. Our present sample size has greater than 80% power to detect a 2.5% difference from this benchmark of 75% interrater agreement.

DISCUSSION

Well-designed and carefully developed classification systems are paramount to achieving more consistent approaches and better outcomes in the treatment of TL trauma. To ensure proposed classifications such as TLISS evolve they must undergo appropriate psychometric scrutiny. Precision of the measurement of the variables critical to the TLISS could be significantly influenced by the training and background of the surgeons utilizing it. Because both neurosurgeons and orthopedic spine surgeons are involved in the care of TL trauma it is important that the TLISS provides reliability in both disciplines.

TABLE 3. Intrarater Agreement Within Orthopedic and Neurosurgeons

	Percent Agreement	Cohen κ	Weighted κ	Spearman Rank Order
Mechanism				
Neurosurgeons	56.8	0.39	0.42	0.54
Orthopedic surgeons	60.6	0.44	0.52	0.59
Neurologic status				
Neurosurgeons	87.4	0.78	0.85	0.88
Orthopedic surgeons	92.3*	0.86	0.88	0.90
PLC integrity				
Neurosurgeons	61.4	0.38	0.48	0.54
Orthopedic surgeons	71.0*	0.51	0.59	0.65*
Total TLISS				
Neurosurgeons	37.6	0.29	0.54	0.73
Orthopedic surgeons	44.2*	0.37	0.65	0.78
Management				
Neurosurgeons	76.1	0.59	0.56	0.59
Orthopedic surgeons	79.1	0.63	0.59	0.61

*$P < 0.05$ for difference between orthopedic and neurosurgeons.

The results of this study demonstrate good overall interspeciality reliability of the TLISS management recommendation when applied to 56 clinical TL injuries (74% agreement between neurosurgeons and orthopedic surgeons). Although there were statistically significant differences between the percent agreements within orthopedic and within neurosurgeons for the TLISS subcategories and total score, these differences were less then 5% for interrater assessments and most of the intrarater differences. Thus, these small differences are unlikely to be clinically significant. The interrater and intrarater reliability for the TLISS system injury mechanism, PLC integrity, and subsequent total score was only fair (κ 0.21–0.40), regardless of specialty.[14]

Although not the focus of this paper, the low intraobserver and interobserver reliability of the mechanism of injury and PLC integrity subscores as discussed by Harrop et al[17] has led to a detailed reassessment and modification of TLISS.

Mechanism of injury has been redefined as injury morphology (less subjective) and clearer guidelines regarding scoring have been implemented. The STSG has also established a more precise definition and means of diagnosing posterior ligamentous integrity. These recommendations have resulted in evolution of the TLISS to the thoracolumbar injury classification and severity score (TLICS).[3]

Several studies have shown the influence of surgeon specialty on management decisions, surgical treatment, and preferences between orthopedic versus neurosurgeons.[18–23] However, only a few have focused on spinal trauma. In a survey of Canadian neurologic (n = 44) and orthopedic surgeons (n = 25), Findlay et al[19] reported that the neurologic status of the patient and the surgeon's specialty were significant determinates in the management of TL burst fractures. Glaser et al[20] reported only slight agreement (κ ranged from 0.09 to 0.14) among 31 orthopedic surgeons and neurosurgeons regarding the appropriateness of several management techniques for 5 cervical spine trauma cases. Unfortunately no specific analysis was conducted to directly compare orthopedic surgeons to neurosurgeons.[20] In contrast to the Glaser study, Grauer et al[24] reported an average agreement (yes or no response) of 91% in the decision to operate or not. This high degree of agreement, however, may be due to selection bias in the clinical vignettes. In the cases where significant agreement to operate was noted (6 out of 8), a good general consensus on the indication for surgery exists. In comparison, the current study achieved 92% overall agreement (yes or no response) and 75% interrater agreement with the management recommendations derived from applying the TLISS to 56 clinical vignettes that were representative of the full spectrum of TL injuries.

In addition to the current study, the work by Grauer et al, and recent studies assessing variation in surgical decision making for degenerative spinal diseases by Irwin et al, have also shown a high degree of consensus between orthopedic and neurosurgeons.[21,22,24] This trend may reflect the increasing cross-pollination of spinal education between orthopedic and neurosurgical training programs with combined scientific meetings, multispecialty spinal societies and collaborative research.

Although the trend toward consensus in management of spinal disorders is encouraging, the significant regional variations in rates of spine surgery suggest that further work is required to develop evidence-based therapeutic strategies. This is evidenced by the fact that spinal surgery has a higher degree of regional variation compared with other types of surgery.[12] A large number of factors have been reported to contribute to this variation. Some factors include, overall spine surgeon density, patient factors, surgeon factors (age and type of spine surgery training), and clinical uncertainty with several apparently equivalent treatment options.[5–13] The latter (ie, lack of clear diagnostic and management decisions) is highlighted by the recent attempts in the development of standards or guidelines for the management of cervical trauma and spinal cord injury and for lumbar disorders.[25,26] Although there is insufficient evidence to support treatment standards and guidelines for many spinal conditions including TL trauma, there is an opportunity to develop consensus-based approaches and then to validate these prospectively. Given that there is a high degree of consensus among neurosurgeons and orthopedic surgeons regarding the TLISS classification, we believe that this presents a unique opportunity to undertake prospective evidence-based validation studies. These in turn could lead to management guidelines that would influence practice and streamline care delivery.

Currently, the Denis and Arbeitsgemeinschaft für Osteosynthesefragen TL fracture classification systems are the most commonly used schemes to classify a TL injury, but both systems are reported to have inconsistent intraobserver and interobserver agreement (slight to borderline substantial agreement).[27–30] Both these systems have only shown moderate reliability and repeatability at the simplest level of the classification system.[29,30] Neither system has been assessed regarding interrater reliability between orthopedic and neurosurgeons. Although the study by Wood et al,[30] also used members of the STSG, the Arbeitsgemeinschaft für Osteosynthesefragen or Denis system does not directly give a recommendation for surgical versus nonsurgical treatment. Consequently, a direct comparison of these systems to the TLISS regarding the degree of agreement between spinal orthopedic surgeons and neurosurgeons on the initial decision for management of TL injuries is not possible. The TLISS (now TLICS) has been proposed as an attempt to improve fracture classification and provide guidance toward ultimate management. It is essentially an adaptation of the strengths of the aforementioned classification systems. The genesis of the 3 major categories is further based on input from a large group (STSG) of both spinal orthopedic and spinal neurosurgeons. Compared with similar studies, the greatest strength of this study lies in the large number of participants and the broad spectrum of cases assessed. Consequently, although the TLISS system had only fair interrater agreement for specific subscores, the authors can state that the TLISS system at least seems to provide an algorithm whereby moderate to substantial interobserver and intraobserver agreement regarding the management recommendation for TL injuries can be achieved within and between spinal orthopedic surgeons and neurosurgeons.

There are limitations of this study that require consideration. First, although a strength, the large number of cases may have contributed to reduced reliability secondary to rater fatigability. Second, although there was a slight reduction in the number of surgeons at the second interval, the results were not significantly affected. In fact the agreement at the second interval was slightly better then the first interval. Third, many of the participants of this study were significantly involved in the development of the TLISS and subsequently are biased to the key elements of the algorithm. Consequently, the "generalize-ability" of this system (TLISS/TLICS) to other academic and nonacademic spinal orthopedic and neurosurgeons must be assessed.

Although this system, as with any other, does not take into account all other clinical factors (eg, patient factors such as medical comorbidities or polytrauma) that influence surgical decision making, the high degree of agreement achieved is very encouraging. The TLICS will require rigorous validity assessments and may yet undergo further refinements before recommendation for wide spread clinical use. It will require reliability and validity (decision to operate) assessment by a broader cross-section of spinal surgeons as well as trauma and emergencyphysicians with appropriate education on the use of the system. The development and evaluation of this system by both orthopedic and neurosurgeons and the reliability between them in an efficacy setting will hopefully improve acceptance by both specialties and enable more widespread dissemination and education regarding cotemporary thinking about TL injury classification and management. By improving surgeon agreement regarding the management of TL injuries, TLISS/TLICS may assist in improving the regional variability that exists in the decision for surgical management of TL injuries. Although consistency is definitely a step in the right direction, this does not automatically translate into improved patient care.[31] This system has not been validated with respect to the clinical outcome of operative versus nonoperative treatment for a given injury or the various surgical techniques available once a decision to operate has been made. Furthermore, if the decision to operate, even if we all agree, is not supported by the best available evidence, then the TLISS remains nothing more than a therapeutic guideline.

CONCLUSIONS

The TLISS system demonstrated good reliability between orthopedic and neurosurgeons and in terms of intraobserver and interobserver agreement on the algorithm's treatment recommendations for the management of 56 TL trauma cases. Surgeons were in agreement (92%) with the systems recommendation for operative versus nonoperative management and thus TLISS seems to have high validity for both orthopedic and neurosurgeons regarding the initial management decision for treatment of TL injuries. This type of classification system (decision making algorithm) is a step in the right direction to reduce the variability that exists for the management of TL injuries.

REFERENCES

1. Grauer J, Vaccaro A, Lim M, et al. Surgical decision making for unstable thoracolumbar spine injuries: results of a consensus panel review by the Spine Trauma Study Group. *J Spinal Disord Tech*. 2006. In press.
2. Vaccaro AR, Zeiller SC, Hulbert RJ, et al. The thoracolumbar injury severity score: a proposed treatment algorithm. *J Spinal Disord Tech*. 2005;18:209–215.
3. Vaccaro AR, Lehman RA Jr, Hurlbert RJ, et al. A new classification of thoracolumbar injuries: the importance of injury morphology, the integrity of the posterior ligamentous complex, and neurologic status. *Spine*. 2005;30:2325–2333.

4. Chassin MR, Brook RH, Park RE, et al. Variations in the use of medical and surgical services by the Medicare population. *N Engl J Med*. 1986;314:285–290.
5. Ciol MA, Deyo RA, Howell E, et al. An assessment of surgery for spinal stenosis: time trends, geographic variations, complications, and reoperations. *J Am Geriatr Soc*. 1996;44:285–290.
6. Cherkin DC, Deyo RA, Loeser LD, et al. An international comparison of back surgery rates. *Spine*. 1994;19:1201–1206.
7. Davis H. Increasing rates of cervical and lumbar spine surgery in the United States, 1979–1990. *Spine*. 1994;19:1117–1124.
8. Deyo RA, Gray DT, Kreuter W, et al. United States trends in lumbar fusion surgery for degenerative conditions. *Spine*. 2005; 30:1441–1445; discussion 1446–1447.
9. Einstadter D, Kent DL, Fihn SD, et al. Variation in the rate of cervical spine surgery in Washington state. *Med Care*. 1993; 31:711–718.
10. Keller RB, Soule DN, Wennburg JE, et al. Dealing with geographic variations in the use of hospitals. The experience of the Maine Medical Assessment Foundation Orthopaedic Study Group. *J Bone Joint Surg Am*. 1990;72:1286–1293.
11. Nilasena DS, Vaughn RJ, Mori M, et al. Surgical trends in the treatment of diseases of the lumbar spine in Utah's Medicare population 1984–1990. *Med Care*. 1995;33:585–587.
12. The Center for the Evaluative Clinical Sciences, Dartmouth Medical School. The Dartmouth Atlas of Musculoskeletal Health Care. The Trustees of Dartmouth College, 2000.
13. Volinn E, Mayer F, Diehr P, et al. Small area analysis of surgery for low back pain. *Spine*. 1992;17:575–579.
14. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005;37:360–363.
15. Hripcsak G, Heitjan DF. Measuring agreement in medical informatics reliability studies. *J Biomed Inform*. 2002;35:99–110.
16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
17. Harrop JS, Vaccaro AR, Hurlbert RJ, et al. Intrarater and interrater reliability and validity in the assessment of the mechanism of injury and integrity of the posterior ligamentous complex: a novel injury severity scoring system for thoracolumbar injuries. *J Neurosurg Spine*. 2006;4:118–122.
18. Drew B, Bhandari M, Orr D, et al. Surgical preference in anterior cervical discectomy: a national survey of Canadian spine surgeons. *J Spinal Disord Tech*. 2002;15:454–457.
19. Findlay JM, Grace MG, Saboe LA, et al. A survey of vertebral burst-fracture management in Canada. *Can J Surg*. 1992; 35:407–413.
20. Glaser JA, Jaworski BA, Cuddy BG, et al. Variation in surgical opinion regarding management of selected cervical spine injuries. A preliminary study. *Spine*. 1998;23:975–983.
21. Irwin ZN, Hilibrand A, Gustavel M, et al. Variation in surgical decision making for degenerative spinal disorders. Part I: Lumbar spine. *Spine*. 2005;30:2208–2213.
22. Irwin ZN, Hilibrand A, Gustavel M, et al. Variation in surgical decision making for degenerative spinal disorders. Part II: Cervical spine. *Spine*. 2005;30:2214–2219.
23. Pickett GE, Van Soelen J, Duggal N. Controversies in cervical discectomy and fusion: practice patterns among Canadian surgeons. *Can J Neurol Sci*. 2004;31:478–483.
24. Grauer JN, Vaccaro AR, Beiner JM, et al. Similarities and differences in the treatment of spine trauma between surgical specialties and location of practice. *Spine*. 2004; 29:685–696.
25. Hadley MN, Walter BC, et al. Guideline Committee. Guidelines for the management of acute cervical spine and spinal cord injuries. *Neurosurgery*. 2002;50:S1–S199.
26. Resnick DK, Choudhri TF, Dailey AT, et al. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 1: Introduction and methodology. *J Neurosurg Spine*. 2005;2:637–638.
27. Denis F. The three column spine and its significance in the classification of acute thoracolumbar spinal injuries. *Spine*. 1983;8:817–831.
28. Magerl F, Aebi M, Gertzbein SD, et al. A comprehensive classification of thoracic and lumbar injuries. *Eur Spine J*. 1994; 3:184–201.
29. Oner FC, Ramos LM, Simmermacher RK, et al. Classification of thoracic and lumbar spine fractures: problems of reproducibility. A study of 53 patients using ct and mri. *Eur Spine J*. 2002; 11:235–245.
30. Wood KB, Khanna G, Vaccaro AR, et al. Assessment of two thoracolumbar fracture classification systems as used by multiple surgeons. *JBJS*. 2005;87A:1423–1429.
31. Woolf SH, Grol R, Hutchinson A, et al. Clinical guidelines: potential benefits, limitations and harms of clinical guidelines. *BMJ*. 1999;318:527–530.