VOLUME 29 · NUMBER 22 · AUGUST 1 2011

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Spinal Instability Neoplastic Score: An Analysis of Reliability and Validity From the Spine Oncology Study Group

Daryl R. Fourney, Evan M. Frangou, Timothy C. Ryken, Christian P. DiPaola, Christopher I. Shaffrey, Sigurd H. Berven, Mark H. Bilsky, James S. Harrop, Michael G. Fehlings, Stefano Boriani, Dean Chou, Meic H. Schmidt, David W. Polly, Roberto Biagini, Shane Burch, Mark B. Dekutoski, Aruna Ganju, Peter C. Gerszten, Ziya L. Gokaslan, Michael W. Groff, Norbert J. Liebsch, Ehud Mendel, Scott H. Okuno, Shreyaskumar Patel, Laurence D. Rhines, Peter S. Rose, Daniel M. Sciubba, Narayan Sundaresan, Katsuro Tomita, Peter P. Varga, Luiz R. Vialle, Frank D. Vrionis, Yoshiya Yamada, and Charles G. Fisher

Author affiliations appear at the end of this article.

Submitted January 16, 2011; accepted May 18, 2011; published online ahead of print at www.jco.org on June 27, 2011.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Daryl R. Fourney MD, FRCSC, FACS, Associate Professor of Neurosurgery, Director, Neurosurgery Residency Training Program, University of Saskatchewan, Royal University Hospital, 103 Hospital Dr, Saskatoon, Saskatchewan, Canada S7N 0W8; e-mail: daryl.fourney@usask.ca.

0732-183X/11/2922-3072/$20.00

DOI: 10.1200/JCO.2010.34.3897

ABSTRACT

Purpose
Standardized indications for treatment of tumor-related spinal instability are hampered by the lack of a valid and reliable classification system. The objective of this study was to determine the interobserver reliability, intraobserver reliability, and predictive validity of the Spinal Instability Neoplastic Score (SINS).

Methods
Clinical and radiographic data from 30 patients with spinal tumors were classified as stable, potentially unstable, and unstable by members of the Spine Oncology Study Group. The median category for each patient case (consensus opinion) was used as the gold standard for predictive validity testing. On two occasions at least 6 weeks apart, each rater also scored each patient using SINS. Each total score was converted into a three-category data field, with 0 to 6 as stable, 7 to 12 as potentially unstable, and 13 to 18 as unstable.

Results
The κ statistics for interobserver reliability were 0.790, 0.841, 0.244, 0.456, 0.462, and 0.492 for the fields of location, pain, bone quality, alignment, vertebral body collapse, and posterolateral involvement, respectively. The κ statistics for intraobserver reliability were 0.806, 0.859, 0.528, 0.614, 0.590, and 0.662 for the same respective fields. Intraclass correlation coefficients for inter- and intraobserver reliability of total SINS score were 0.846 (95% CI, 0.773 to 0.911) and 0.886 (95% CI, 0.868 to 0.902), respectively. The κ statistic for predictive validity was 0.712 (95% CI, 0.676 to 0.766).

Conclusion
SINS demonstrated near-perfect inter- and intraobserver reliability in determining three clinically relevant categories of stability. The sensitivity and specificity of SINS for potentially unstable or unstable lesions were 95.7% and 79.5%, respectively.

J Clin Oncol 29:3072-3077. © 2011 by American Society of Clinical Oncology

INTRODUCTION

Spinal cord compression from epidural tumor is often discussed as an indication for operation. A prospective randomized trial has demonstrated superiority of surgery and radiation therapy compared with radiation alone in the treatment of high-grade spinal cord compression for solid tumors.[1] Spinal instability is a separate indication for surgery[2-7] or percutaneous cement augmentation,[8,9] but it has not received the same degree of scrutiny in the literature as spinal cord compression. This paucity of data may reflect the controversy that exists regarding instability resulting from neoplastic destruction of spinal elements, as evidenced by a wide variety of criteria published in the literature[2,8-16] and significant differences of opinion suggested by spine surgeons.[2,12,13]

The Spine Oncology Study Group (SOSG) defines spine instability as the "loss of spinal integrity as a result of a neoplastic process that is associated with movement-related pain, symptomatic or progressive deformity and/or neural compromise under physiological loads."[14] The development of a standard and valid classification with easily assigned radiographic and patient factors was championed to aid communication and appropriate referral between oncologists, radiologists, and spine surgeons and facilitate prompt, optimized treatment plans. Furthermore, a classification system could lead to a more consistent therapeutic approach among spine surgeons and aid in education and scientific study.

Table 1. SINS

SINS Component	Score
Location	
Junctional (occiput-C2, C7-T2, T11-L1, L5-S1)	3
Mobile spine (C3-C6, L2-L4)	2
Semirigid (T3-T10)	1
Rigid (S2-S5)	0
Pain*	
Yes	3
Occasional pain but not mechanical	1
Pain-free lesion	0
Bone lesion	
Lytic	2
Mixed (lytic/blastic)	1
Blastic	0
Radiographic spinal alignment	
Subluxation/translation present	4
De novo deformity (kyphosis/scoliosis)	2
Normal alignment	0
Vertebral body collapse	
> 50% collapse	3
< 50% collapse	2
No collapse with > 50% body involved	1
None of the above	0
Posterolateral involvement of spinal elements†	
Bilateral	3
Unilateral	1
None of the above	0

NOTE. Data adapted.[14]
Abbreviation: SINS, Spinal Instability Neoplastic Score.
*Pain improvement with recumbency and/or pain with movement/loading of spine.
†Facet, pedicle, or costovertebral joint fracture or replacement with tumor.

An evidence-based process using the best available literature and expert-opinion consensus was used to develop the Spine Instability Neoplastic Score (SINS; Table 1).[12-14] In this classification system, tumor-related instability is assessed by adding together six individual component scores: spine location, pain, lesion bone quality, radiographic alignment, vertebral body collapse, and posterolateral involvement of the spinal elements. The minimum score is 0, and the maximum is 18. A score of 0 to 6 denotes stability, 7 to 12 denotes indeterminate (possibly impending) instability, and 13 to 18 denotes instability. A surgical consultation is recommended for patients with SINS scores greater than 7.[14]

With face and content validity evaluated, the next phase of psychometric evaluation is to determine the reliability and predictive validity of the classification. The objective of this study is to determine the intraobserver and interobserver reliability of SINS. A secondary objective is a preliminary assessment of the predictive validity of SINS.

METHODS

Patient Case Selection and Evaluation

The SOSG is an international group of 30 spine oncology experts and thought leaders from North America, Europe, South America, and Asia who meet to discuss research, assess the best evidence for current practices, and formulate clinical trials to advance the field of spine oncology. SOSG members were asked to contribute patient case examples with imaging and clinical information for the purpose of testing SINS reliability and validity.

Table 2. Patient Cases

Level	Stable	Potentially Unstable	Unstable	Total
Cervical	3	2	5	10
Thoracic	2	5	3	10
Lumbar	3	3	4	10
Total	8	10	12	30

NOTE. Final case series was selected to represent range of spinal levels and grades of stability. Stability was determined by anonymous voting by panel of experts (consensus opinion).

A total of 50 de-identified patient cases were obtained. Patient cases that did not contain sufficient history or quality imaging were excluded. To obtain a SINS score, the history must include a description of pain, especially as it relates to patient movement. Imaging must include computed tomography (CT) scan or magnetic resonance imaging; however, if it is the latter, x-ray films (or preferably CT) are also required to determine bone lesion quality (ie, lytic, blastic, or mixed). In the case of multiple spinal lesions, contributors identified the specific lesion they intended for scoring. Thirty patient cases were chosen, with roughly equal representation of cervical, thoracic, and lumbar spinal levels as well as a broad range of neoplastic instability (Table 2).

Patient cases were classified as stable, potentially unstable, or unstable on the basis of anonymous voting by SOSG members. The median category for each patient case was termed the consensus opinion and was used as the gold standard for reference in the predictive validity analysis of SINS. Next, each SOSG member was provided with a CD-ROM that included the case series, a scoring sheet, and instructions on SINS scoring. Twenty-four members independently applied SINS in the 30 patient cases. Scoring was repeated at least 6 weeks later using the same patient cases, presented in different order.

On the basis of preliminary analysis results and after further discussion among SOSG members, SINS was modified to improve reliability by simplifying the scoring method so that the minimum score in each category was 0. In addition, regions of the spine were defined more clearly: junctional levels were occiput-C2, C7-T2, T11-L1, and L5-S1; mobile levels were C3-6 and L2-4; semi-rigid spine was T3-T10; and rigid spine was S2-S5. On the basis of data from the evidence-based reviews[12,13] and expert consensus, SINS was also modified to include consideration of lesion bone quality and nonmechanical back pain.[14]

Six months later, 24 SOSG members scored the 30 patient cases again using the revised SINS classification system via the same methods (Appendix Figs A1, A2, online only). Once completed, the results were sent to an independent central study coordinator. Scoring was repeated at least 6 weeks later by the same observers, with the patient cases presented in a different order to limit recall bias.

Statistical Analysis

Three statistical tests were used to assess inter- and intraobserver reliability. The intraclass correlation coefficient (ICC) was used to measure both inter- and intraobserver agreement for total SINS scores (two-way mixed effect model, in which people effects are random, and measures effects are fixed).[17] For each of the six components of SINS (ie, location, pain, bone quality, radiographic alignment, vertebral body collapse, and posterolateral involvement), Fleiss's κ for multiple raters was used to measure interobserver agreement, and Cohen's κ was used to evaluate intraobserver agreement.[18,19]

Each total SINS score was collapsed into three categories, with 0 to 6 as stable, 7 to 12 as potentially unstable, and 13 to 18 as unstable. Predictive validity was assessed using Cohen's κ for agreement between SINS categorization and consensus score.[19] Analysis was performed with SPSS version 15.0 (SPSS, Chicago, IL). Level of agreement for κ was determined as per Landis et al[20] (Table 3).

RESULTS

Interobserver Reliability

The interobserver ICC reliability for total SINS score was 0.846 (95% CI, 0.773 to 0.911). The analysis of SINS components revealed

Table 3. Level of Agreement for κ Statistic Levels

κ Value	Level of Agreement
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
> 0.80	Near perfect

NOTE. Data adapted.[20]

interobserver reliabilities (Fleiss κ) of 0.790, 0.841, 0.244, 0.456, 0.462, and 0.492 for the fields of location, pain, bone quality, alignment, vertebral body collapse, and posterolateral involvement, respectively (Table 4). The level of agreement was substantial for location, near perfect for pain, fair for bone quality, and moderate for the radiographic criteria of alignment, vertebral body collapse, and posterolateral involvement.

Intraobserver Reliability

The intraobserver ICC reliability for total SINS score was 0.886 (95% CI, 0.868 to 0.902). Analysis of SINS components revealed intraobserver reliabilities (Cohen's κ) of 0.806, 0.859, 0.528, 0.614, 0.590, and 0.662 for the fields of location, pain, bone quality, alignment, vertebral body collapse, and posterolateral involvement, respectively (Table 4). The level of agreement was near perfect for location and pain, substantial for alignment and posterolateral involvement of the spine, and moderate for bone quality and vertebral body collapse.

Validity

The Cohen's κ statistic for agreement between collapsed SINS category and consensus score was 0.712 (95% CI, 0.676 to 0.766), which represents substantial agreement. Table 5 lists the scores determined by SINS raters, collapsed into the three categories (0 to 6, stable; 7 to 12, potentially unstable; and 13 to 18, unstable) and cross-tabulated with the a priori consensus opinion (gold standard). All unstable patient cases (by consensus opinion) obtained scores by SINS of 7 or greater, which suggests referral for assessment by a spine surgeon.

By analyzing results as negative for stable lesions (scores 0 to 6) and positive for potentially unstable or unstable lesions (scores ≥ 7), SINS could be subjected to a binary classification test. Sensitivity was $563 \div 588 = 95.7\%$, and specificity was $171 \div 215 = 79.5\%$. The type II error (false-negative rate) was 4.3%, and the type I error (false-positive rate) was 20.5%.

Table 5. Cross-Tabulation of Scores Determined by SINS Raters and Median Categorization of Expert Panel

	Collapsed SINS			
Consensus Opinion	Stable (0-6)	Potentially Unstable (7-12)	Unstable (13-18)	Total
Stable	171	44	0	215
Potentially unstable	25	322	52	399
Unstable	0	47	142	189
Total	196	413	194	803

Abbreviation: SINS, Spinal Instability Neoplastic Score.

DISCUSSION

A major goal in the treatment of metastatic disease is the preservation or restoration of spinal stability.[1-16] Prediction of fracture risk in oncology is an important goal of classification systems. Mirels[21] proposed a scoring system that is useful for predicting fractures in the peripheral skeleton. In the spine, criteria to define instability, which were originally developed for use in trauma, are not directly applicable in the setting of neoplasia.[2] Before the development of SINS, there were no accepted evidence-based guidelines for the classification of spinal instability secondary to tumor.[12-14] This significant void left oncologists, radiologists, and other health care providers who manage these patients without tools to assess this critical component of care. Patients who develop spinal instability have or are at high risk of having neurologic deficit, severe pain, and progressive deformity.[1-11]

SOSG envisions that SINS will allow clinicians from multiple disciplines who care for patients with spine tumors to make informed decisions about instability and when to refer patients for surgical assessment. SINS should also help spine surgeons make decisions regarding optimal management for specific patients who present with metastatic spine disease. In this initial analysis of reliability and predictive validity, we chose ourselves as the evaluators; however, future work is required to evaluate the performance of SINS among other groups of raters, particularly oncologists and radiologists.

It must be emphasized that the mechanical integrity of the spine is only one component of the evaluation process when considering surgery. If patients have neural deficit or if imaging reveals

Table 4. Reliability Analysis for Six Component Variables of SINS

	Interobserver Reliability (n = 23)				Intraobserver Reliability (n = 21)			
			Agreement				Agreement	
SINS Component	κ	95% CI	Level	%	κ	95% CI	Level	%
Location	0.790	0.774 to 0.806	Substantial	86.3	0.806	0.767 to 0.845	Near perfect	87.4
Pain	0.841	0.823 to 0.859	Near perfect	91.0	0.859	0.822 to 0.896	Near perfect	92.1
Bone quality	0.244	0.227 to 0.260	Fair	68.1	0.528	0.454 to 0.602	Moderate	82.6
Alignment	0.456	0.438 to 0.474	Moderate	71.6	0.614	0.559 to 0.669	Substantial	79.8
Vertebral body	0.462	0.449 to 0.475	Moderate	59.7	0.590	0.541 to 0.639	Moderate	69.3
Posterolateral involvement	0.492	0.476 to 0.508	Moderate	66.5	0.662	0.613 to 0.711	Substantial	77.7

NOTE. $P < .001$ for all.
Abbreviation: SINS, Spinal Instability Neoplastic Score.

spinal cord compression (with or without neural deficit), surgery is often indicated regardless of SINS score.[1] Other important considerations include tumor histology, response to nonoperative treatment, prognosis, patient medical fitness, and patient preference as part of informed choice.[3,9] To our knowledge, this is the first study to examine, in a large, international cohort of neurosurgical and orthopedic spine surgeons, the interobserver and intraobserver reliability and predictive validity of a classification system for tumor-related spinal instability.

We calculated percentage of agreement as a way of gauging agreement among different raters and agreement by the same raters in different timing. To account for the agreement that would be expected purely by chance, κ coefficients and ICC[18] were calculated according to types of variables assessed, as discussed in Methods.

It is not simple to assign a definite interpretation to κ coefficients, because it is dependent on prevalence, number of categories, possible weighting, and presence of bias. Landis et al[20] proposed standards for interpretation of the strength of reliability with the κ coefficient (Table 3). Similar formulations and adaptations exist with slightly different descriptors.[22]

The minimum acceptable value of κ coefficient depends on clinical context and choice of such benchmark is inevitably arbitrary. Many medical journals consider a κ coefficient of less than 0.41 to be clinically unacceptable for reliability.[23] The inter- and intraobserver ICC reliability of final SINS categorization achieved near-perfect agreement. All subcategories had moderate to near-perfect agreement, with the exception of bone quality, which only had fair interobserver reliability and moderate intraobserver reliability. The decision to include this category was made based on strong biomechanical literature suggesting that the cross-sectional area of a defect combined with bone mineral density are excellent predictors of vertebral body failure and pathologic fracture risk.[24-26] In those patients with metastatic lesions of the spine, low bone density is associated with greater fracture risk.[27,28] Removing this component of SINS might improve reliability but would be detrimental to content validity. We may have been able to improve reliability by providing raters with multislice CT for every patient case.

The issue of validity is far more complex than that of reliability. Content and face validity of SINS were facilitated by integrating the best evidence provided by two systematic reviews with expert consensus from the SOSG.[12,13] Quantitative scores were assigned based on the relative importance of particular factors gleaned from the literature and refined by expert consensus.[14] One objective of the current study was an analysis of predictive validity. Obviously, the best method to determine the validity of a classification system is with prospective data; however, a preliminary assessment of the ability of SINS to predict instability was required before proceeding with a prospective protocol.

In medicine, a gold-standard test refers to a diagnostic test or benchmark regarded as definitive. The lack of a gold standard for measuring neoplastic spinal instability has hampered the development of a classification system.[2] For this initial study, we considered two possible gold standards: contributor opinion and consensus opinion. Initially, we thought that the surgeon contributor of each patient case example would be the best to judge instability, because that person would presumably have knowledge of the outcome of the case. On the other hand, this would allow each contributor to impart significant personal bias. For example, if the outcome of a case was surgical stabilization, does that mean the spine was unstable? We believed using the median categorization of a panel of experts would bring us closer to a gold standard based on expert clinical opinion and the knowledge obtained by undertaking the creation of two systematic reviews.[12,13] If consensus opinion and contributor opinion were equal gold standards, they should have a high level of agreement. In a previous abstract, we found only fair agreement between the three-tier category of stability (stable, potentially unstable, unstable) provided by the contributor and the consensus opinion (Cohen's κ, 0.38).[29]

SINS demonstrated high sensitivity (95.7%) for detecting lesions that were unstable or potentially unstable. This was based on combining the potentially unstable and unstable categories as disease positive in a binary classification test. All type II errors (4.3%) were potentially unstable patient cases categorized by SINS as stable. No unstable patient cases were incorrectly categorized as stable by SINS. Likewise, all type I error (20.5%) were stable but categorized by SINS as potentially unstable. In summary, for all raters, SINS did not classify any unstable patient cases as stable or vice versa. The positive predictive value of SINS cannot be calculated with the current data, because this value is directly proportional to prevalence of disease, which is not likely to be the case in a selected, retrospective case series.[30]

Each patient case included a clinical history regarding pain, which is an important component of SINS. In the real world, history taking in this area may be more complex because of multiple sources of pain in patients with metastatic disease. We suspect this has artificially increased the interobserver reliability of pain categorization. Despite being provided the pattern of pain and location of the lesions, the panel still did not have perfect agreement on scores for pain and location, demonstrating some degree of imprecision with scoring and understanding of instructions.

Lack of the ability to view the images in multiple slices in each plane may have reduced the ability to determine the full extent of neoplastic involvement of the spine, thus potentially diminishing reliability of the imaging components of SINS. Bone quality in particular was difficult to score in some patient cases, because although magnetic resonance imaging was provided for every patient case, CT images were not always available. Although plain x-ray is helpful, the sensitivity of CT for assessing bony characteristics is greater.[31]

Total SINS scores had near-perfect inter- and intraobserver reliability when collapsed into three clinically relevant assessments of tumor-related instability, which can be described as stability (scores 0 to 6), indeterminate (possibly impending) instability (7 to 12), and instability (13 to 18). Each component of SINS, with the exception of bone quality, demonstrated clinically acceptable reliability. The false-negative rate was low (4.3%), and all type II errors resulted from distinguishing stable and potentially unstable patient cases. No unstable patient cases were classified as stable by any SINS rater or vice versa.

Although SINS was developed in part to allow non–spine surgeon clinicians to make informed decisions about spinal stability, the raters in this study were a panel of experts. Further determination of SINS reliability is required among other groups of raters, particularly oncologists and radiologists. The prospective application of SINS in different patient populations is required to better determine validity and assess the application of this tool in therapeutic decision making.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked

with a "U" are those for which no compensation was received; those relationships marked with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Employment or Leadership Position: David W. Polly, Scoliosis Research Society (U) **Consultant or Advisory Role:** Christopher I. Shaffrey, Medtronic (C), Biomet (C), DePuy Spine (C); Sigurd H. Berven, Medtronic (C), DePuy Spine (C), Stryker (C), Biomet (C), Orthovita (C); David W. Polly, Medtronic (C); Shane Burch, Medtronic (C) **Stock Ownership:** Ziya L. Gokaslan, Spinal Kinitics, US Spine **Honoraria:** Christopher I. Shaffrey, Stryker; Sigurd H. Berven, Medtronic, Stryker, Biomet, Orthovita; James S. Harrop, Medtronic; Shane Burch, Medtronic **Research Funding:** Christopher I. Shaffrey, National Institutes of Health, Department of Defense, AOSpine; Sigurd H. Berven, Medtronic, AOSpine, Orthopaedic Research and Education Foundation; David W. Polly, Department of Defense, Scoliosis Research Society, Pediatric Orthopaedic Society of North America **Expert Testimony:** None **Other Remuneration:** Mark B. Dekutoski, Medtronic, Broadwater Associates, Synthes, DePuy Spine, Stryker

AUTHOR CONTRIBUTIONS

Conception and design: Daryl R. Fourney, Evan M. Frangou, Timothy C. Ryken, Christian P. DiPaola, Christopher I. Shaffrey, Sigurd H. Berven, Mark H. Bilsky, James S. Harrop, Michael G. Fehlings, Stefano Boriani, Meic H. Schmidt, David W. Polly, Shane Burch, Mark B. Dekutoski, Aruna Ganju, Peter C. Gerszten, Ziya L. Gokaslan, Michael W. Groff, Norbert J. Liebsch, Ehud Mendel, Scott H. Okuno, Shreyaskumar Patel, Laurence D. Rhines, Peter S. Rose, Daniel M. Sciubba, Narayan Sundaresan, Katsuro Tomita, Peter P. Varga, Luiz R. Vialle, Frank D. Vrionis, Yoshiya Yamada, Charles G. Fisher

Collection and assembly of data: Daryl R. Fourney, Evan Mark Frangou, Timothy C. Ryken, Christian P. Dipaola, Christopher I. Shaffrey, Sigurd H. Berven, Mark H. Bilsky, James S. Harrop, Michael G. Fehlings, Stefano Boriani, Dean Chou, Meic H. Schmidt, David W. Polly, Roberto Biagini, Shane Burch, Mark B. Dekutoski, Aruna Ganju, Peter C. Gerszten, Ziya L. Gokaslan, Michael W. Groff, Norbert J. Liebsch, Ehud Mendel, Scott H. Okuno, Shreyaskumar Patel, Laurence D. Rhines, Peter S. Rose, Daniel M. Sciubba, Narayan Sundaresan, Katsuro Tomita, Peter P. Varga, Luiz R. Vialle, Frank D. Vrionis, Yoshiya Yamada, Charles G. Fisher

Data analysis and interpretation: Daryl R. Fourney, Evan M. Frangou, Timothy C. Ryken, Christian P. DiPaola, Christopher I. Shaffrey, Sigurd H. Berven, Mark H. Bilsky, James S. Harrop, Michael G. Fehlings, Stefano Boriani, Dean Chou, Meic H. Schmidt, David W. Polly, Mark B. Dekutoski, Peter C. Gerszten, Ziya L. Gokaslan, Michael W. Groff, Norbert J. Liebsch, Ehud Mendel, Scott H. Okuno, Shreyaskumar Patel, Laurence D. Rhines, Peter S. Rose, Daniel M. Sciubba, Narayan Sundaresan, Katsuro Tomita, Peter P. Varga, Luiz R. Vialle, Frank D. Vrionis, Yoshiya Yamada, Charles G. Fisher

Manuscript writing: All authors

Final approval of manuscript: All authors

REFERENCES

1. Patchell RA, Tibbs PA, Regine WF, et al: Direct decompressive surgical resection in the treatment of spinal cord compression caused by metastatic cancer: A randomised trial. Lancet 366:643-648, 2005
2. Fourney DR, Gokaslan ZL: Spinal instability and deformity due to neoplastic conditions. Neurosurg Focus 14:E8, 2003
3. Fourney DR, Abi-Said D, Lang FF, et al: Use of pedicle screw fixation in the management of malignant spinal disease: Experience in 100 consecutive procedures. J Neurosurg 94:25-37, 2001 (suppl 1)
4. Gokaslan ZL, York JE, Walsh GL, et al: Transthoracic vertebrectomy for metastatic spine tumors. J Neurosurg 89:599-609, 1988
5. DeWald RL, Bridwell KH, Prodromas C, et al: Reconstructive spinal surgery as palliation for metastatic malignancies of the spine. Spine 10:21-26, 1985
6. Falicov A, Fisher CG, Sparkes J, et al: Impact of surgical intervention on quality of life in patients with spinal metastases. Spine (Phila Pa 1976) 31: 2849-2856, 2006
7. Thomas KC, Nosyk B, Fisher CG, et al: Cost-effectiveness of surgery plus radiotherapy versus radiotherapy alone for metastatic epidural spinal cord compression. Int J Radiat Oncol Biol Phys 66:1212-1218, 2006
8. Fourney DR, Schomer DF, Nader R, et al: Percutaneous vertebroplasty and kyphoplasty for painful vertebral body fractures in cancer patients. J Neurosurg 98:21-30, 2003 (suppl 1)
9. Bilsky M, Smith M: Surgical approach to epidural spinal cord compression. Hematol Oncol Clin North Am 20:1307-1317, 2006
10. Taneichi H, Kaneda K, Takeda N, et al: Risk factors and probability of vertebral body collapse in metastases of the thoracic and lumbar spine. Spine (Phila Pa 1976) 22:239-245, 1997
11. Asdourian PL, Mardjetko S, Rauschning W, et al: An evaluation of spinal deformity in metastatic breast cancer. J Spinal Disord 3:119-134, 1990
12. Fehlings MD, David KS, Furlan JC, et al: Oncologic instability of the cervical spine: A systematic review, Can J Surg 51:11, 2008 (suppl)
13. Weber MH, Burch M, Buckley J, et al: Instability and impending instability of the thoracolumbar spine in patients with spinal metastases: A systematic review. Int J Oncol 38:5-12, 2011
14. Fisher CG, DiPaola CP, Ryken TC, et al: A novel classification system for spinal instability in neoplastic disease: An evidence-based approach and expert consensus from the Spine Oncology Study Group. Spine (Phila Pa 1976) 35:E1221-E1229, 2010
15. Cybulski GR: Methods of surgical stabilization for metastatic disease of the spine. Neurosurgery 25:240-252, 1989
16. Kostuik J, Weinstein J: Differential diagnosis and surgical treatment of metastatic spine tumors, in Frymoyer JW, Ducker TB, Hadler MN, et al (eds): The Adult Spine: Principles and Practice. New York, NY, Raven Press, 1991, pp 861-888
17. Shrout PE, Fleiss JL: Intraclass correlations: Uses in assessing rater reliability. Psycho Bull 86: 420-428, 1979
18. Fleiss JL: Measuring nominal scale agreement among many raters. Psycho Bull 76:378-381, 1971
19. Cohen J: A coefficient of agreement for nominal scales. Educ Psychol Meas 20:37-46, 1960
20. Landis JR, Koch GG: The measurement of observer agreement for categorical data. Biometrics 33:159-174, 1977
21. Mirels H: Metastatic disease in long bones: A proposed scoring system for diagnosing impending pathologic fractures. Clin Orthop 249:256-264, 1989
22. Shrout PE: Measurement reliability and agreement in psychiatry. Stat Methods Med Res 7:301-317, 1998
23. Sim J, Wright CC: The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. Phys Ther 85:257-268, 2005
24. Dimar JR 2nd, Voor MJ, Zhang YM, et al: A human cadaver model for determination of pathologic fracture threshold resulting from tumorous destruction of the vertebral body. Spine (Phila Pa 1976) 23:1209-1214, 1998
25. Windhagen H, Hipp JA, Hayes WC, et al: Postfracture instability of vertebrae with simulated defects can be predicted from computed tomography data. Spine (Phila Pa 1976) 25:1775-1781, 2000
26. Windhagen HJ, Hipp JA, Silva MJ, et al: Predicting failure of thoracic vertebrae with simulated and actual metastatic defects. Clin Orthop 344:313-319, 1997
27. Tschirhart CE, Finkelstein JA, Whyne CM, et al: Biomechanics of vertebral level, geometry, and transcortical tumors in the metastatic spine.J Biomech 40:46-54, 2007
28. Whyne CM, Hu SS, Klisch S, et al: Effect of the pedicle and posterior arch on vertebral body strength predictions in finite element modeling. Spine (Phila Pa 1976) 23:899-907, 1998
29. Fourney DR, Ryken T, Fisher CG: A novel approach to define instability in patients with spine tumors: Validity and inter-rater reliability of the Spinal Instability Neoplastic Score (SINS). Can J Surg 52:14, 2009 (suppl)
30. Altman DG, Bland JM: Diagnostic tests 2: Predictive values. BMJ 309:102, 1994
31. White AP, Kwon BK, Lindskog DM, et al: Metastatic disease of the spine. J Am Acad Orthop Surg 14:587-598, 2006

Affiliations

Daryl R. Fourney and Evan M. Frangou, University of Saskatchewan, Royal University Hospital, Saskatoon, Saskatchewan; Christian P. DiPaola and Charles G. Fisher, University of British Columbia, Vancouver, British Columbia; Michael G. Fehlings, University of Toronto, Toronto, Ontario, Canada; Timothy C. Ryken, Iowa Spine and Brain Institute, Waterloo, IA; Christopher I. Shaffrey, University of Virginia Medical Center, Charlottesville, VA; Sigurd H. Berven, Shane Burch, and Dean Chou, University of California at San Francisco, San Francisco, CA; Mark H. Bilsky and Yoshiya Yamada, Memorial Sloan-Kettering Cancer Center; Narayan Sundaresan, Mount Sinai Hospital, New York, NY; James S. Harrop, Thomas Jefferson University, Philadelphia; Peter C. Gerszten, University of Pittsburgh Medical Center, Pittsburgh, PA; Stefano Boriani, Ospedale Maggiore, Ausl Bologna, Bologna; Roberto Biagini, Regina Elena Institute, Rome, Italy; Meic H. Schmidt, Huntsman Cancer Institute and Clinical Neurosciences Center, University of Utah, Salt Lake City, UT; David W. Polly, University of Minnesota, Minneapolis; Mark B. Dekutoski, Peter S. Rose, and Scott H. Okuno, Mayo Clinic, Rochester, MN; Aruna Ganju, Northwestern University, Chicago, IL; Ziya L. Gokaslan and Daniel M. Sciubba, Johns Hopkins University, Baltimore, MD; Michael W. Groff, Beth Israel Deaconess Medical Center; Norbert J. Liebsch, Massachusetts General Hospital, Boston, MA; Ehud Mendel, The Ohio State University, Columbus, OH; Shreyaskumar Patel and Laurence D. Rhines, The University of Texas MD Anderson Cancer Center, Houston, TX; Katsuro Tomita, Kanazawa University School of Medicine, Kanazawa, Japan; Peter P. Varga, National Center for Spinal Disorders, Budapest, Hungary; Luiz R. Vialle, Catholic University of Paraná, Curitiba, Paraná, Brazil; and Frank D. Vrionis, H. Lee Moffitt Cancer Center and Research Institute and University of South Florida College of Medicine, Tampa, FL.