

Spinal Instability Neoplastic Score (SINS): Reliability Among Spine Fellows and Resident Physicians in Orthopedic Surgery and Neurosurgery

Shandy Fox, MD¹ , Michael Spiess, MD, FRCSC¹,
Luke Hnenny, MD, FRCSC¹, and Daryl R. Fourney, MD, FRCSC, FACS¹

Abstract

Study Design: Reliability analysis.

Objectives: The Spinal Instability Neoplastic Score (SINS) was developed for assessing patients with spinal neoplasia. It identifies patients who may benefit from surgical consultation or intervention. It also acts as a prognostic tool for surgical decision making. Reliability of SINS has been established for spine surgeons, radiologists, and radiation oncologists, but not yet among spine surgery trainees. The purpose of our study is to determine the reliability of SINS among spine residents and fellows, and its role as an educational tool.

Methods: Twenty-three residents and 2 spine fellows independently scored 30 de-identified spine tumor cases on 2 occasions, at least 6 weeks apart. Intraclass correlation coefficient (ICC) measured interobserver and intraobserver agreement for total SINS scores. Fleiss's kappa and Cohen's kappa analysis evaluated interobserver and intraobserver agreement of 6 component subscores (location, pain, bone lesion quality, spinal alignment, vertebral body collapse, and posterolateral involvement of spinal elements).

Results: Total SINS scores showed near perfect interobserver (0.990) and intraobserver (0.907) agreement. Fleiss's kappa statistics revealed near perfect agreement for location; substantial for pain; moderate for alignment, vertebral body collapse, and posterolateral involvement; and fair for bone quality (0.948, 0.739, 0.427, 0.550, 0.435, and 0.382). Cohen's kappa statistics revealed near perfect agreement for location and pain, substantial for alignment and vertebral body collapse, and moderate for bone quality and posterolateral involvement (0.954, 0.814, 0.610, 0.671, 0.576, and 0.561, respectively).

Conclusions: The SINS is a reliable and valuable educational tool for spine fellows and residents learning to judge spinal instability.

Keywords

tumor, metastases, oncology, orthopedic, neurosurgery, instability

Introduction

Spinal metastases are identified in 10% to 20% of patients with cancer.¹⁻³ Of these patients, 10% to 20% develop neural compression that may require surgical intervention.² Surgical decompression in this setting has been well established in the literature.⁴ Patients without neural compression but with significant instability due to tumor may also benefit from stabilizing surgery⁵; mechanical instability being an independent indication for stabilization regardless of grade or radiosensitivity of the tumor.^{6,7} Until recently, guidance for the determination of spinal stability in these patients was minimal. Scoring

systems utilized for instability in a traumatic setting do not apply and much of the decision making was left to the surgeon's clinical judgment. A standardized, validated classification system for spinal instability in the setting of neoplasia was needed to guide more consistent therapeutic approaches among

¹ University of Saskatchewan, Saskatoon, Saskatchewan, Canada

Corresponding Author:

Shandy Fox, 103 Hospital Drive, Saskatoon, Saskatchewan, S7N 0W8, Canada.
Email: slf672@mail.usask.ca



Table 1. Spinal Instability Neoplastic Score (SINS) System.^a

Component	Score
Location	
Junctional (O-C2; C7-T2; T11-L1; L5-S1)	3
Mobile spine (C3-6; L2-4)	2
Semirigid (T3-10)	1
Rigid (S2-S5)	0
Mechanical pain	
Yes	3
No	2
Pain free lesion	1
Bone lesion	
Lytic	2
Mixed (lytic/blastic)	1
Blastic	0
Radiographic spinal alignment	
Subluxation/translation present	4
Deformity (kyphosis/scoliosis)	2
Normal	0
Vertebral body collapse	
>50% collapse	3
<50% collapse	2
No collapse with >50% body involved	1
None of the above	0
Posterolateral involvement	
Bilateral	3
Unilateral	1
None of the above	0

^aData adapted from Fischer et al.⁹

spine surgeons, aid communication and appropriate referral between oncologists, radiologists, and spine surgeons, and facilitate more organized and prompt treatment plans for these patients.⁸ Such a system, known as the Spinal Instability Neoplastic Score (SINS), was devised by the Spine Oncology Study Group (SOSG)⁹ in 2010 (Table 1). It assesses and scores 6 variables: location of lesion, characterization of pain, type of bony lesion, radiographic spinal alignment, degree of vertebral body destruction, and involvement of posterolateral spinal elements. The scores for each variable are added, and a final score is obtained. The minimum score is 0, and the maximum score is 18. A score of 0 to 6 denotes stability, a score of 7 to 12 denotes indeterminate (possibly impending) instability, and a score of 13 to 18 denotes instability. For scores greater than 7, a surgical consultation is recommended⁹ (Table 2).

Face and content validity has been evaluated for SINS,⁹ as has reliability and predictive validity among expert spine surgeons in the SOSG.⁸ Since the introduction of SINS by the SOSG, a few independent reliability studies have evaluated its use in varying specialties involved in the care of these patients. Campos et al¹⁰ found excellent inter/intraobserver reliability among 6 physicians; 1 radiotherapy oncologist, 1 palliative oncologist, 3 spine orthopedic surgeons, and 1 general orthopedic surgeon. Fisher et al^{11,12} reported SINS as a highly reliable, reproducible, and valid tool among 33 radiation oncologists and 37 radiologists from international sites.

Table 2. Total Spinal Instability Neoplastic Score: Determination of Stability.^a

	Score (Total = 0-18)		
	1-6	7-12	13-18
Clinical categories	Stable	Potentially unstable	Unstable
Binary scale	Stable	Current or potentially unstable = possible surgical intervention	

^aData adapted from Fischer et al.⁹

Teixeira et al evaluated the scoring system based on experience of the evaluator. They demonstrated that experience has a significant impact on the reliability of the SINS score; interobserver agreement was only fair among non-spine surgeons and spine surgeons not experienced in the treatment of vertebral metastatic disease.¹³ All of these articles assessed the scoring system based on practicing physicians in varying fields. To date, none have included trainees, either residents or fellows, in the participant groups and thus the role of SINS has not been evaluated in an education setting.

The purpose of the present study was to establish the intraobserver reliability and interobserver reliability of the SINS among spine fellows and residents in neurosurgery and orthopedic surgery. Fellows and residents of all levels of training participate in the initial assessment, treatment planning, surgical management, and postoperative care of patients with spinal neoplastic disease. For this reason, any classification system that claims to be reliable needs to be shown to be so amongst residents and fellows as well. Additionally, a system shown to be reliable among trainees can serve as a learning tool, providing a validated framework for determining instability in spinal neoplastic disease.

Material and Methods

A total of 47 cases were reviewed and scored using SINS by the authors. Cases with insufficient history or inadequate imaging were excluded. Of the remaining cases, 30 were selected. Care was taken to have equal numbers of stable, potentially unstable, and unstable cases, and to have an anatomical frequency of distribution of cervical, thoracic, and lumbar cases (Table 3). Sacral cases were not included as these lesions are commonly stable and were also not included in the reference study by the SOSG.^{8,9}

Each case had appropriate demographic data and history provided, with emphasis on the characterization of the patient's pain and how it related to movement. This information is important for the surgeon to help determine mechanical versus oncologic type pain and resultant instability. Images provided included select slices from either computed tomography (CT) scan or magnetic resonance imaging (MRI) combined with X-rays.

To determine validity for a representative sample group of trainees, residents in neurosurgery and orthopedic surgery at the Universities of Saskatchewan and Calgary were approached

Table 3. Patient Cases.

Level	Stable	Potentially Unstable	Unstable	Total
Cervical	1	2	5	8
Thoracic	4	8	3	15
Lumbar	2	3	2	7
Total	7	13	10	30

to participate in the study. A total of 23 residents (orthopedic surgery = 18; neurosurgery = 5) were recruited, along with 2 spine fellows from the University of Calgary. Although more orthopedic surgery residents were enrolled compared with neurosurgery, this is a reflection that more residents are admitted to the orthopedic programs per year. Residents of all levels, from year 1 to 5 in orthopedics and 1 to 6 in neurosurgery, were included to avoid bias toward more experienced participants. Prior to evaluating the 30 selected cases, all participants underwent an introduction to the SINS. This included reviewing the original article, "A Novel Classification System for Spinal Instability in Neoplastic Disease," along with written instructions about the SINS. A descriptive presentation was given outlining the process and working through a number of case examples in order to familiarize them with the scoring system.

Following the introduction, participants then reviewed the 30 deidentified cases. All 25 raters independently scored each case based on the six individual components of the SINS classification. The total score (0-18) was calculated and classified as stable, potentially unstable, or unstable according to the 3 clinical categories. At least 6 weeks following the initial evaluation, the participants once again evaluated the same set of cases. The cases were presented in a different order from the initial scoring to limit recall bias. The total scores were once again calculated and classified for level of stability.

Statistical Analysis

Three statistical tests were used to assess inter- and intraobserver reliability. These were the same statistical tests used by the SOSG to evaluate reliability amongst Spine Surgeons. The intraclass correlation coefficient (ICC) was used to measure both the inter- and intraobserver agreement for the total SINS scores. Each of the 6 components (location, pain, bone quality, radiographic alignment, vertebral body collapse, and posterolateral involvement) then underwent Fleiss's kappa analysis for multiple raters to measure interobserver agreement and Cohen's kappa to evaluate intraobserver agreement. The level of agreement for kappa values was calculated using the Landis grading system¹⁴ (Table 4).

Results

The interobserver reliability for the total SINS score revealed an ICC value of 0.990 (95% CI, 0.984-0.955). This corresponds with near perfect level of agreement using the Landis grading system. The analysis of interobserver reliability for the six individual SINS components revealed values

Table 4. Level of Agreement for κ Statistic Level.^a

κ	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

^aData adapted.¹⁴

(Fleiss κ) of 0.948, 0.739, 0.382, 0.427, 0.550, and 0.435 for the areas of location, pain, bone quality, alignment, vertebral body collapse, and posterolateral involvement, respectively (Table 5). Using the Landis grading system, these values correspond to near perfect agreement for location, substantial for pain, fair for bone quality, and moderate for alignment, vertebral body collapse, and posterolateral involvement.

The intraobserver reliability for the total SINS score revealed an ICC value of 0.907 (95% CI, 0.893-0.930). This corresponds with near perfect level of agreement. The analysis of intraobserver reliability for the 6 individual SINS components revealed values (Cohens κ) of 0.954, 0.814, 0.576, 0.610, 0.671, and 0.561 for the areas of location, pain, bone quality, alignment, vertebral body collapse, and posterolateral involvement, respectively (Table 5). These values correspond to near perfect agreement for location and pain, moderate for bone quality and posterolateral involvement, and substantial for alignment and vertebral body collapse.

Discussion

An assessment of spinal instability is an essential component of decision making in spine surgery. Spinal instability was classically defined by Panjabi and White¹⁵ as "the loss of the ability of the spine under physiological loads to maintain relationships between the vertebrae in such a way that there is neither damage nor subsequent irritation to the spinal cord or nerve roots and, in addition, there is no development of incapacitating deformity or pain due to structural changes." This elegant definition has been used as the basis for multiple systems that classify spinal stability in the context of various pathologic states. Stability of traumatic spinal injury has been particularly well-classified by a number of different systems, including the AO classification of thoracolumbar injuries,¹⁶ the Thoracolumbar Injury Severity Score,¹⁷ and Subaxial Cervical Spine Injury Classification System.¹⁸ These classification systems not only provide a framework for the clinician to decide whether a particular injury is stable or unstable but also serve as a guide to management of these injuries. In the past, assessment of tumor-related instability has lacked an accepted classification system such as these and decision making was largely guided by clinical experience.⁸ This raises a number of issues for nonspine surgeons, particularly those in training, when assessing patients and determining appropriate intervention.

Table 5. Reliability Analysis for Component Variables of Spinal Instability Neoplastic Score (SINS).^a

SINS Component	Interobserver Reliability			Intraobserver Reliability		
	Fleiss's κ	95% CI	Agreement	Cohen's κ	95% CI	Agreement
Location	0.948	0.933-0.964	Near perfect	0.954	0.936-0.974	Near perfect
Pain	0.739	0.690-0.788	Substantial	0.814	0.775-0.853	Near perfect
Bone quality	0.382	0.345-0.420	Fair	0.576	0.523-0.629	Moderate
Alignment	0.427	0.383-0.471	Moderate	0.610	0.570-0.668	Substantial
Vertebral body	0.550	0.527-0.573	Moderate	0.671	0.630-0.712	Substantial
Posterolateral involvement	0.435	0.417-0.453	Moderate	0.561	0.512-0.610	Moderate

^aP < .001 for all.

The assessment, diagnosis, and management of spine patients is an integral part of resident training and surgical practice. Among the patients and disease states trainees are exposed to, spine oncology patients are often complex due to medical comorbidity, functional limitations, limited life expectancy and variable treatment strategies.^{1,3-7} Although many factors are involved in the treatment of these patients, one of the key components is that of tumor-related instability.^{1,3-7,19} As spine oncology patients require timely multidisciplinary care, it is imperative that a clear framework is established to avoid variability in interpretation, referral patterns, and consequently patient management.¹¹ SINS is the only tool for scoring instability in spine tumors that has repeatedly demonstrated reliability among varying cohorts. Our study is the first, to our knowledge, to assess reliability among spine surgery trainees.

The inter- and intraobserver ICC reliability of final SINS scores achieved near perfect agreement among trainees. In practice, this correlates with an appropriate categorization of patients as being unstable, potentially unstable, or unstable. For spine surgical trainees, SINS should safely and accurately identify those patients for which surgical intervention should be considered. As an educational tool, it can act as a guide for trainees in learning how to evaluate for instability.

All 6 subcategories had moderate to near perfect agreement in both inter- and intraobserver reliability, with the exception of bone quality, which only had fair agreement in interobserver and moderate agreement in intraobserver reliability. This correlates with clinically acceptable reliability in all areas other than bone quality. Comparing the analysis of intraobserver reliabilities of the 6 SINS components, the kappa values in both studies were very comparable. We acknowledge that agreement in each category may in part be due to the design of the study and the limited information provided in each case. Although the cases attempt to recreate clinical scenarios, information provided by patients in real life can vary and images provided are often much more detailed than was possible in these studies.

Some areas of the scoring system have little variability by nature of objectivity. An example of this is location. There is little subjective interpretation to location based on imaging and thus it is not surprising that we had near perfect agreement in both inter and intrarater reliability. In all of the studies evaluating reliability, location consistently has excellent reproducibility.¹⁰⁻¹³ The subcategory of pain is another area

with little variability based on the study design. In our cases, participants were presented with a well-documented pain history leading to near perfect agreement. This level of agreement was seen across the board with the other studies in all physicians and all levels of experience.¹⁰⁻¹³ An argument could be made that in real life, this category may not be so straightforward. Patients often have a difficult time describing their pain and may have multiple source of pain that can cloud the clinician's judgment.

On the other hand, determining bone quality, alignment, vertebral body collapse, and posterolateral involvement is more variable as it relies on image interpretation. Each case provided in our study included only a selection of representative images (4-6). This may be one explanation for only achieving fair-moderate agreement on bone quality and substantial in the rest. In clinical practice, the clinician has the ability to navigate through multiple slices of each diagnostic imaging modality, which likely facilitates a more accurate assessment of bone quality and overall bony changes. To eliminate this, we could have created short MPEG files of the complete CT/MRI scans. As our study was referenced around the original article published by the SOSG, we provided single images in order to keep the study designs consistent.

Overall, the SINS has fared well in reliability and validity studies. Classifying patients as having unstable or potentially unstable spinal disease can help determine those patients in whom surgical intervention should be considered and thus improve patient outcomes. This is especially the case in those patients without obvious spinal cord compression or neurological compromise. Fourny et al⁸ reported that no unstable spinal cases were classified as stable in their analysis of the SINS system, thus no at risk patients were missed.

Retrospective reviews further support the use of the SINS system in the decision-making process.²⁰⁻²³ Zadnik et al²⁰ looked at 31 patients with multiple myeloma who underwent surgery for stabilization. Twenty-five of the cases were classified according to the SINS system based on information provided. All cases were either potentially unstable or unstable, suggesting the SINS system is an accurate tool to simulate clinical judgement.²⁰

Conclusion

Spinal neoplasia is a complex condition that may lead to spinal instability and neurologic compromise. Assessing stability can

be challenging and overwhelming for healthcare providers who lack the experience of spine specialists. Our study shows that the SINS is a reliable educational tool for orthopedic and neurosurgery trainees of all levels, providing a validated framework for diagnosing instability in spinal neoplastic disease.


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Shandy Fox, MD  <http://orcid.org/0000-0001-5447-6379>

References

1. Delank KS, Wedntner C, Eich HT, Eysel P. The treatment of spinal metastases. *Dtsch Arztebl Int*. 2011;108:71-84.
2. Greenlee RT, Murray T, Bolden S, Wingo PA. Cancer statistics 2000. *CA Cancer J Clin*. 2000;50:7-33.
3. Klimo P Jr, Schmidt MH. Surgical management of spinal metastases. *Oncologist* 2004;9:188-196.
4. 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*. 2005;366:643-648.
5. Hare R, Angelov L. Spine metastases: current treatments and future directions. *Eur J Cancer*. 2010;46:2696-2707.
6. Filis AK, Aghayev KV, Doulgeris JJ, Gonzalez-Blohm SA, Vriornis FD. Spinal neoplastic instability: biomechanics and current management options. *Cancer Control*. 2014;21:144-150.
7. Bilsky MH, Azeem S. The NOMS framework for decision making in metastatic cervical spine tumors. *Curr Opin Orthop*. 2007;18:263-269.
8. Fourny DR, Frangou EM, Ryken TC, et al. Spinal instability neoplastic score: an analysis of reliability and validity from the spine oncology study group. *J Clin Oncol*. 2011;29:3072-3077.
9. 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)*. 2010;35:E1221-E1229.
10. Campos M, Urrutia J, Zamora T, et al. The Spine Instability Neoplastic Score: an independent reliability and reproducibility analysis. *Spine J*. 2014;14:1466-1469.
11. Fisher CG, Schouten R, Versteeg AL, et al. Reliability of the Spinal Instability Neoplastic Score (SINS) among radiation oncologists: an assessment of instability secondary to spinal metastases. *Radiat Oncol*. 2014;9:69. doi:10.1186/1748-717X-9-69.
12. Fisher CG, Versteeg AL, Schouten R, et al. Reliability of the Spinal Instability Neoplastic Scale among radiologists: an assessment of instability secondary to spinal metastases. *AJR Am J Roentgenol*. 2014;203:869-874.
13. Teixeira WG, Coutinho PR, Marchese LD, et al. Interobserver agreement for the spine instability neoplastic score varies according to the experience of the evaluator. *Clinics*. 2013;68:213-218.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
15. Panjabi MM, White AA 3rd. Basic biomechanics of the spine. *Neurosurgery*. 1980;7:76-93.
16. Magerl F., Aebi M, Gertzbein SD, Harms J, Nazarian S. A comprehensive classification of thoracic and lumbar injuries. *Eur Spine J*. 1994;3:184-201.
17. 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.
18. Vaccaro AR, Hulbert RJ, Patel AA, et al. The subaxial cervical spine injury classification system: a novel approach to recognize the importance of morphology, neurology, and integrity of the disco-ligamentous complex. *Spine (Phila Pa 1976)*. 2007;32:2365-2374.
19. Fourny DR, Gokaslan ZL. Spinal instability and deformity due to neoplastic conditions. *Neurosurg Focus*. 2003;14:E8.
20. Zadnik PL, Goodwin CR, Karami K, et al. Outcomes following surgical intervention for impending and gross instability caused by multiple myeloma in the spinal column. *J Neurosurg Spine* 2015;22:301-309. doi:10.3171/2014.9.SPINE14554.
21. Huisman M, van der Velden JM, van Vulpen M, et al. Spinal instability as defined by the spinal instability neoplastic score is associated with radiotherapy failure in metastatic spinal disease. *Spine J*. 2014;14:2835-2840. doi:10.1016/j.spinee.2014.03.043
22. Sahgal A, Atenafu EG, Chao S, et al. Vertebral compression fracture after spine stereotactic body radiotherapy: a multi-institutional analysis with a focus on radiation dose and the Spinal Instability Neoplastic Score. *J Clin Oncol*. 2014;31:3426-3431. doi:10.1200/JCO.2013.50.1411.
23. Ivanishvili Z, Fourny DR. Incorporating the spine instability neoplastic score into a treatment strategy for spinal metastasis: LMNOP. *Global Spine J*. 2014;4:129-136.