

ASSESSMENT OF SCORES IN DECISION MAKING IN METASTASES OF THE SPINE

AValiação de Escores na Tomada de Decisão em Metástases na Coluna

Evaluación de las Puntuaciones en la Toma de Decisiones en las Metástasis de Columna

MARCUS VINICIUS DE OLIVEIRA FERREIRA¹, RENATO HIROSHI SALVIONI UETA¹, DAVID DEL CURTO¹, EDUARDO BARROS PUERTAS¹

ABSTRACT

Objective: The aim of this study is to assess the intra- and interobserver concordance of SINS, Harrington, Tokuhashi and Tomita scores among general orthopedic surgeons and spine surgeons with experience above 5 and 10 years in the evaluation of patients with spinal metastasis. **Methods:** Twenty cases of patients with metastatic lesion of the spine were presented to 10 examiners and the scores aforementioned have been applied. After six weeks, the cases were reintroduced in a different order and data were analyzed. **Results:** The intraobserver reliability showed better agreement in SINS score among examiners with less experience and Harrington and Tomita scores among those who had more than 10-year experience. The interobserver reliability of the examiners of the group with over 10-year experience showed higher precision when using these scores, especially Harrington and Tomita. The SINS score was the choice for daily practice and was able to modify the management more often. **Conclusions:** This study demonstrated that the use of predictive scores of instability, Harrington, and prognosis, Tomita, had a higher intra- and interobserver reliability particularly among spine surgeons with experience above 10 years.

Keywords: Prognosis; Neoplasm metastasis; Spinal neoplasms; Life expectancy.

RESUMO

Objetivo: O objetivo deste estudo é avaliar a concordância intra e interobservador dos escores SINS, Harrington, Tokuhashi e Tomita entre ortopedistas gerais e cirurgiões de coluna com experiência acima de cinco e 10 anos na avaliação de pacientes com metástase vertebral. **Métodos:** Foram apresentados 20 casos de pacientes com lesões metastáticas da coluna vertebral a 10 examinadores e os escores citados acima foram aplicados. Após seis semanas, os casos foram reapresentados em ordem diferente e os dados foram analisados. **Resultados:** A confiabilidade intraobservador apresentou melhor concordância no escore SINS entre os examinadores com menor experiência, e nos escores de Harrington e Tomita entre os que tinham experiência maior que 10 anos. A confiabilidade interobservador dos examinadores do grupo com experiência maior que 10 anos apresentou maior precisão na utilização desses escores, e com destaque para Harrington e Tomita. O escore SINS foi o de eleição para a prática diária e foi capaz de modificar a conduta com mais frequência. **Conclusões:** Este estudo demonstrou que o uso de escores preditores de instabilidade, Harrington, e prognóstico, Tomita, apresentam maior confiabilidade intra e interobservador, principalmente entre os cirurgiões de coluna com experiência superior a 10 anos.

Descritores: Prognóstico; Metástase neoplásica; Neoplasias da coluna vertebral; Expectativa de vida.

RESUMEN

Objetivo: El objetivo de este estudio es evaluar la concordancia intra e interobservador en las puntuaciones SINS, Harrington, Tokuhashi y Tomita entre los ortopedistas generales y cirujanos de la columna vertebral con experiencia superior a 5 y 10 años en la evaluación de pacientes con metástasis vertebrales. **Métodos:** Se presentaron 20 casos de pacientes con lesiones metastáticas de la columna vertebral a 10 examinadores, quienes aplicaron las puntuaciones mencionadas. Luego de seis semanas, los casos fueron reintroducidos en un orden diferente y se analizaron los datos. **Resultados:** La fiabilidad intraobservador mostró mejor concordancia en la puntuación SINS entre los examinadores con menos experiencia y en los puntajes Harrington y Tomita en los que tenían experiencia de más de 10 años. La fiabilidad interobservador de los examinadores del grupo con más de 10 años de experiencia mostró una mayor precisión en el uso de estas puntuaciones, especialmente las puntuaciones Harrington y Tomita. La puntuación SINS fue de elección para la práctica diaria y ha sido capaz de modificar la conducta con más frecuencia. **Conclusiones:** Este estudio demostró que el uso de las puntuaciones de predicción de inestabilidad, Harrington, y del pronóstico, Tomita, presentan mayor fiabilidad intra e interobservador principalmente entre los cirujanos de columna con más de 10 años de experiencia.

Descriptores: Pronóstico; Metástasis de la neoplasia; Neoplasias de la columna vertebral; Esperanza de vida.

1. Universidade Federal de São Paulo (UNIFESP), Escola Paulista de Medicina, Department of orthopedics and Traumatology, Spine Group, São Paulo, SP, Brazil.

Study conducted at Universidade Federal de São Paulo (UNIFESP), Escola Paulista de Medicina, Department of orthopedics and Traumatology, Spine Group, São Paulo, SP, Brazil.
Correspondence: Spine Group of The Department of Orthopedics and Traumatology of Escola Paulista de Medicina of Universidade Federal de São Paulo (Unifesp), Rua Borges Lagoa, 783 – 5º. Andar, Vila Clementino, São Paulo, SP, Brazil. 04038-03. ferreiramvo@gmail.com

INTRODUCTION

The spine is the most common site of metastatic disease.^{1,2} Patients with cancer present spinal metastases in 70% of cases and up to 10% develop spinal cord compression.^{2,3} The ratio of these lesions to primary tumors is 40:1 and skeletal dissemination should be considered in the differential diagnosis of a patient with a spinal cord lesion.^{1,4,5}

Metastatic involvement of the spine is more common in primary tumors of the breast, lung, kidney, thyroid gland, and the prostate, in that order, according to Papastefanou et al.⁶ In our practice, Valesin Filho et al.⁷ described the prevalence in 55 patients: breast (32.7%), multiple myeloma (25.4%), prostate (14.5%), gastric carcinoma (5.4%), neoplasia of the lung (3.6%), neoplasia of the kidney (3.6%), and others (14.5%).

The involvement of the spine is a common problem and its incidence is increasing because the methods of detection, screening, and treatment of primary cancer are allowing patients with active disease to live longer.^{6,8}

The development of classifications help with communication and proper referrals between oncologists, radiologists, and spine surgeons, and facilitate optimized treatment plans.⁹⁻¹² The main classification systems take into account instability (the Spine Instability Neoplastic Score [SINS]⁸ and Harrington¹³) and the prognosis (Tokuhashi et al.¹⁰ and Tomita et al.¹⁴), and help reduce uncertainty in the decision-making process.⁶

This study therefore attempts to evaluate the intra- and interobserver concordances in the SINS,⁸ Harrington,¹³ Tokuhashi et al.,¹⁰ and Tomita et al.¹⁴ scores between general orthopedic surgeons and spine surgeons with more than 5 and 10 years of experience in the evaluation of patients with vertebral metastasis.

MATERIAL AND METHODS

We retrospectively evaluated the medical records of 20 patients with metastatic lesions in the spine treated at the Spine Clinic of Hospital São Paulo, of the UNIFESP Department of Orthopedics, Escola Paulista de Medicina. The study was submitted for the approval of the Research Ethics Committee of this institution, under number 959680. The patients' identities were not revealed to the evaluating physicians.

The study included patients with metastatic disease of the spine, verified in the case histories, clinical exams, and imaging exams. Five patients with multiple myeloma were also included, as we believe that these patients have natural histories, behaviors, and treatments similar to those of the patients with metastases. Insufficient information was adopted as the criterion for exclusion from the study.

The cases were presented to four general orthopedists, four spine surgeons with more than 5 years of experience, and two spine surgeons with more than 10 years of experience, and displayed in Power Point with a list of symptoms, laboratory and imaging (radiography, tomography, and resonance) exams, and biopsy results. The examiners were then given the SINS⁸ (Tables 1 and 2), Harrington¹³ (Table 3), Tokuhashi¹⁰ (Tables 4 and 5), and Tomita¹⁴ (Tables 6 and 7) scores for their assessment of the cases. After six weeks the cases were reviewed again in a different order, to eliminate any memory bias, with a restatement of the scores presented in the tables below and a self-assessment questionnaire consisting of three questions: (1) Do you use these scores routinely in treating patients with vertebral tumors? (2) If you do not use them, what is your method of choice for daily practice? and (3) If the treatment indicated by the score is different from yours, which classification system would change your conduct? The responses were stored, for the calculation of intra- and interobserver reproducibility.

RESULTS

A convenience sampling was taken, consisting of 20 patients with an average age of 54.5 years, mostly women (70%). The diagnosis was

Table 1. SINS score.

SINS element	Score
Location	
Junctional (occiput-C2, C7-T2, T11-L1, L5-S1)	3
Mobile spine (C3-C6, L2-L4)	2
Semi-rigid (T3-T10)	1
Rigid (S2-S5)	0
Pain relief with recumbency and/or pain with movement/loading of the spine	
Yes	3
No (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 the spinal elements (facet joints, pedicle or costovertebral joint fracture or replacement with tumor)	
Bilateral	3
Unilateral	1
None of the above	0

Table 2. Interpretation of SINS.

Total	Condition	Treatment
≤6	Stable	Observation
7-12	Indeterminate	Consider surgery
≥13	Unstable	Surgery

Table 3. Harrington classification.

Class	Description	Treatment
I	No significant neurological involvement	Clinical
II	Involvement of bone without collapse or instability	Clinical
III	Major neurological impairment (sensory or motor), without significant involvement of bone	Clinical or surgical
IV	Vertebral collapse with pain due to mechanical causes or instability, but without significant neurological impairment	Surgical
V	Collapse of the vertebral body or instability combined with major neurological impairment	Surgical

Table 4. Tokuhashi score.

Characteristic	Score
General Condition (Karnofsky Performance Status)	
Poor (PS 10-40%)	0
Moderate (PS 50-70%)	1
Good (PS 80-100%)	2
Number of extra-spinal bone metastases	
≥3	0
1-2	1
0	2
Number of vertebral body metastases	
≥3	0
2	1
1	2
Metastases to major internal organs	
Unresectable	0
Resectable	1
No metastases	2
Location of primary cancer	
Lung, osteosarcoma, stomach, bladder, esophagus, pancreas	0
Liver, gall bladder, unidentified	1
Others	2
Kidney, uterus	3
Rectum	4
Thyroid, breast, prostate, carcinoid tumor	5
Paralysis	
Complete (Frankel A, B)	0
Incomplete (Frankel C, D)	1
Absent (Frankel E)	2

Table 5. Strategy for the treatment of metastases according to the Tokuhashi score.

Score	Survival forecast	Treatment
0-8	≤6 months	Conservative treatment
9-11	≥6 months	Palliative treatment or excisional surgery if there is a single lesion with no metastases to major internal organs
12-15	≥1 year	Excisional surgery

Table 6. Tomita score.

Prognostic factor	Score
Primary tumor	
Slow growth (ex. breast, prostate, thyroid)	1
Moderate growth (ex. kidney, uterus)	2
Rapid growth (ex. lung, liver, stomach, colon, primary site unknown)	4
Visceral metastases to vital organs (lung, liver, kidney, and brain)	
None	0
Treatable	2
Untreatable	4
Bone metastases	
Solitary or isolated	1
Multiple	2

myeloma in 25% of the cases (5/20), followed by 15% breast (3/20), 10% adenocarcinoma (2/20), lymphoma (2/20), lung (2/20), and indeterminate (2/20), and 5% kidney (1/20), liver (1/20), and prostate (1/20).

Intraobserver reliability (Stability over time and repeatability)¹⁵

Figure 1 shows, according to the Bland and Altman method,¹⁶ the average difference between each pair of observations made by the ten examiners (y axis) and how much these differences deviate from zero (ideal) and from the line of average difference (real bias)

Table 7. Strategy for the Tomita score.

Score	Treatment	Surgical strategy
2-3	Long term Local control	Wide or marginal excision
4-5	Medium term Local control	Marginal or intralesional excision
6-7	Short term Palliative	Palliative surgery
8-10	Terminal care	Supportive care

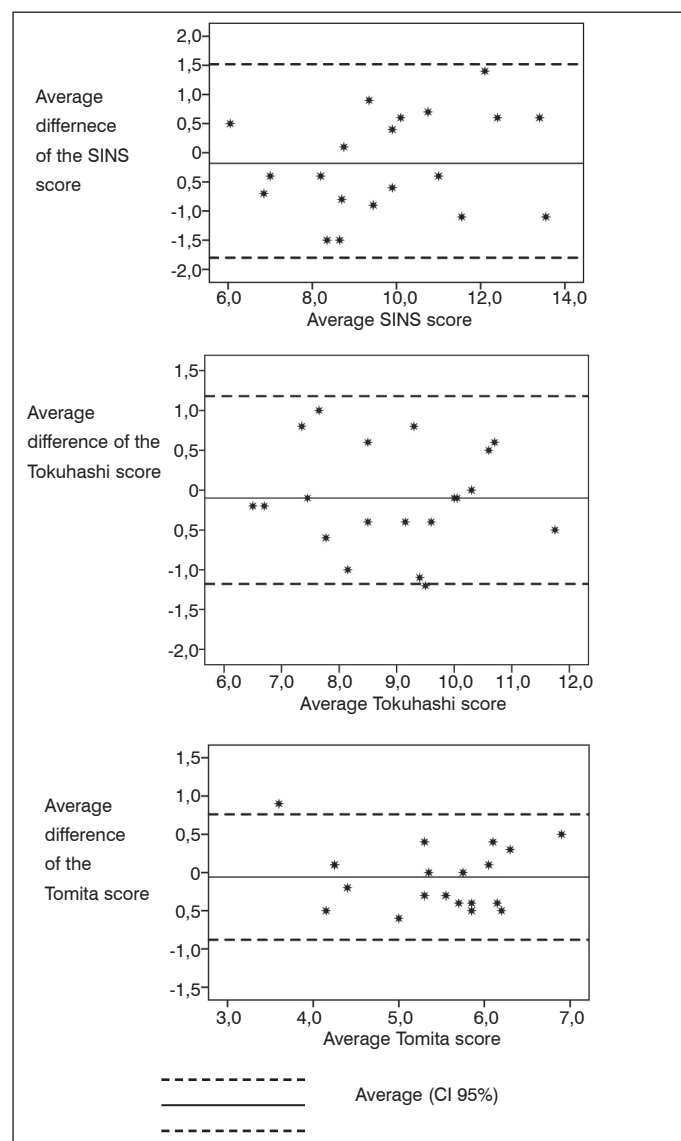


Figure 1. Bland and Altman¹⁶ method for analysis of the repeatability of the stability/prognosis scores between two different observations. São Paulo, 2014.

for all observations. In the Tomita and Tokuhashi tools, only one pair of sets of observations showed a difference beyond the limits of confidence of the bias line.

One can see that the real bias line in all three tools approximates zero, denoting a small average difference between the two application times of the tools. The distance between the bias line and zero was tested under the hypothesis of being different, with a negative result accepting the null hypothesis.

No significant correlations were observed between the bias and the magnitude of the measurement, i.e. there is no association between the average distance of the pairs of observations and the average score obtained.

To summarize, we can say that the tools analyzed using the Bland and Altman¹⁶ method are stable during the time interval between the two measurements and there is an acceptable level of concordance between the scores obtained by the same examiner on two separate occasions.

Table 8 shows the degree of concordance between the stability/prognosis categories of the 10 examiners, stratified according to the level of expertise and hands-on experience using the tools.

The SINS tool yielded one pair of observations with perfect intraobserver concordance, another pair with substantial concordance, and three pairs with moderate concordance, while the other pairs did not demonstrate reasonable concordance between the two time periods. The group of general orthopedic surgeons presented the highest degrees of concordance, with a Kappa average of 0.44 and half of the observations with substantial concordance or better.

In the evaluation using the Tomita tool, the group with more than five years of experience had a substantially higher frequency of concordant observations, though the highest Kappa average (0.62) was observed in the group with more than 10 years of experience.

In general, the Tokuhashi tool presented low intraobserver concordance with 60% of the observations at a reasonable degree of concordance or worse. The Kappa average was 0.34 and was higher among the examiners with more than 10 years of experience.

The Harrington tool had a higher proportion of observers with almost perfect or substantial concordance (30%), while the other 70% of the observations obtained reasonable or moderate concordance. The Kappa averages for the groups by expertise/experience were 0.45, 0.52, and 0.59 for general orthopedic surgeons, spine surgeons with more than 5 years of experience, and spine surgeons with more than ten years of experience, respectively.

Interobserver Reliability (Reproducibility)¹⁵

Table 9 gives an overview of reproducibility among the scores obtained using the tools of the study. For the SINS tool, Intraclass Correlation Coefficients closer to 1 were observed among the examiners with experience suggesting better concordance between the diagnoses. This finding was not observed in the analysis of the ICC of the interobserver scores for the Tomita and Tokuhashi tools.

Table 8. Intraobserver concordance according to the Kappa method.

	SINS				Tomita				Tokuhashi				Harrington			
	Examiners															
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
General Orthopedics	0.17	0.62	1.0	0.11	0.38	0.29	0.13	0.44	0.31	0.19	0.20	0.31	0.39	0.91	0.21	0.31
Experience > 5 years	0.23	0.45	0.45	0.17	0.32	0.68	0.66	0.20	0.31	0.11	0.55	0.46	0.28	0.37	0.55	0.90
Experience > 10 years	0.56	0.38	—	—	0.55	0.70	—	—	0.45	0.60	—	—	0.72	0.44	—	

Table 9. Reproducibility according to the intraclass correlation coefficient (ICC) and the average coefficient of variation (CV), calculated for the scores obtained.

	SINS		Tomita		Tokuhashi	
	ICC (CI95%)	CV	ICC (CI95%)	CV	ICC (CI95%)	CV
All the examiners	0.54 (0.86-0.96)	23%	0.32 (0.18 -0.54)	24%	0.40 (0.25 -0.61)	19%
General Orthopedics	0.41 (0.18 -0.65)	23%	0.29 (0.08 -0.55)	19%	0.45 (0.22 -0.68)	19%
Experience > 5 years	0.68 (0.49 -0.95)	17%	0.30 (0.08 -0.56)	25%	0.39 (0.16 -0.44)	17%
Experience > 10 years	0.67 (0.33 -0.85)	22%	0.49 (0.13 -0.86)	7%	0.35 (0.09 -0.68)	13%

The examiners in the group with more than 10 years of experience had lower average coefficients of variation (CV), especially for the Tomita tool, denoting greater precision in its use.

For the analysis of concordance of stability/prognosis among the professional groups, Table 10 demonstrates greater precision among the examiners for the Harrington tool.

Table 10. Interobserver concordance according to the Kappa method.

	Sins	Tomita	Tokuhashi	Harrington
General Orthopedics	0.45	0.28	0.18	0.38
Experience > 5 years	0.31	0.23	0.38	0.71
Experience > 10 years	0.36	0.56	0.21	0.81

There was better intra-examiner concordance (k=0.45) in the group of general orthopedic surgeons using SINS. The group with more than 5 years of experience performed better than the other groups using the Tokuhashi tool. The group of examiners with more than 10 years of experience seemed to have better precision using the Tomita and Harrington tools.

Regarding the self-assessment questions, we observed that the general orthopedists did not use the scores to treat cases (100%) and they were more inclined to use the SINS tool. The examiners with experience used the scores routinely, most often using the Tokuhashi and SINS. The latter is the more frequently used, and is liable to modify the conduct in relation to a case (83%).

DISCUSSION

Our study was developed to evaluate the use of scores in patients with vertebral metastases. The treatment decision in these patients may be modified through the use of classifications that are capable of evaluating instability (SINS⁸ and Harrington¹³) and prognosis (Tokuhashi et al.¹⁰ and Tomita et al.¹⁴). The development of simple classification systems with easy radiographical attributes and patient factors that contribute to and facilitate communication and appropriate referrals between oncologists, radiologists, orthopedic and spine surgeons, and neurosurgeons help to ensure that treatment plans are faster and better optimized.

Establishing criteria for a surgical indication is very difficult due to the variety of symptoms and survival prognoses.^{10,17} Early detection and proper intervention are critical to minimizing the sequelae from spine metastases, reestablishing function, and maximizing the quality of life.¹⁸

There was almost perfect inter- and intraobserver reliability in the total SINS scores for the three clinically relevant evaluations of tumor-related instability, which can be described as stable (score from 0 to 6), indeterminate (imminent from 7 to 12), and unstable (13 to 18).^{8,17,19} We observed that all the examiners rated it as important in daily practice and capable of changing their conduct in relation to a case. One criticism of the SINS score was that the neurological status of the patient, a potential modifier of the treatment approach, is not included in the evaluation.

Harrington designed a classification scheme with five categories for metastatic tumors of the spine based on bone destruction and neurological impairment.¹³ In this score, surgery is indicated only in the presence of vertebral instability or mechanical pain. The Harrington classification had the best interobserver precision among the spine specialists, almost perfect among those with more than 10 years of experience (K 0.81). However, this system is excessively oversimplified, resulting in broad categories of patients who can have very different prognoses.² For example, a patient with radicular pain but good function can be allocated to the same group as a patient with complete paralysis from a large tumor.²

In 1990 Tokuhashi et al.¹⁰ elaborated a treatment and procedure selection strategy based on life expectancy. In 2005 they revised the tool to improve the precision of this system. The score interval for the parameter "primary site" was changed to 0 to 5 points and the total score possible was increased to a maximum of 15 points. Pre-treatment prognostic evaluation is the most important factor in determining the selection of treatment methods, including surgical procedures. Using this system, life expectancy was consistent with real post-treatment survival time in 86.4% of cases in a prospective series of 118 patients and in 82.5% of cases in a retrospective series of all 246 cases.¹⁰ Our study showed low intraobserver concordance with 60% of the observations with a reasonable degree of concordance or worse and a Kappa average of 0.34. Despite these conclusions, in the self-assessment questionnaires this score was included in the preferences of the spine specialists as one of the scores of choice, and as a score capable of changing the treatment approach in 80% of the cases. This fact may be attributed to the presence of a large number of variables and the inclusion of neurological status.

Tomita et al.¹⁴ recommend an alternative prognosis scoring system, taking the histology of the tumor and its biological behavior into account. This system was built based on the retrospective data of 67 patients between 1987 and 1991, and weighted point values were given to prognostic factors following an evaluation of their statistical rates of risk. The histology of the primary tumor is closely correlated with survival in both surgical patients and medical cohorts, with greater survival times observed in patients with breast, prostate, and thyroid cancers. Therefore, the type of primary tumor played a predominant role in the scoring of the Tomita *apud* Choi et al.² system. This score had the lowest coefficient of variation (7%) among the examiners with more than 10 years of experience, demonstrating better

interobserver concordance, though it was neither the tool of choice for daily practice nor was it selected as a modifier of approach.

There is controversy in the literature concerning the use of the Tokuhashi and Tomita scores in patients with myeloma.^{20,21} Leithner et al.²⁰ proposed the inclusion of multiple myeloma among the malignancies with longer survival times, despite its being a hematological disease rather than a metastatic disease spread from a solid tumor, and suggested that these patients be included in a group with a better prognosis. Thus, in the Tokuhashi score, the primary site would receive 5 points, and in the Tomita score, a classification as a slow-growing primary site would receive 1 point. The study by Majeed et al.²¹ also included patients with myeloma in the Tokuhashi and Tomita scores, given that they are allocated to the same category as patients with metastasis of the prostate and breast and because a similar survival was observed in the study by these authors. In our study we observed great variability in our attempt to include myeloma in these scores. In general, the patients in poorer clinical condition were allocated to the groups with poorer prognoses. In the Tokuhashi score, patients were classified by primary site in the other and unidentified categories, while in the Tomita classification, most cases were placed in the rapid growth group. This put these patients in categories with less aggressive or conservative treatments, thereby modifying this form of conduct in the cases. In our practice, Avanzi et al.²² studied the correlation between spinal fractures and survival in patients with multiple myeloma using the Tokuhashi and Tomita scores. These authors were not able to predict survival using these scores. The inclusion of myeloma in well-defined categories may help to better identify these patients.

CONCLUSION

This study demonstrated that the use of scores that predict instability, that of Harrington in particular, and prognosis, mainly Tomita, have a higher level of intra- and interobserver reliability among spine surgeons with more than 10 years of experience. The SINS score was the instrument of choice for daily practice, and the one that most often led to a change of conduct.

All authors declare no potential conflict of interest concerning this article.

REFERENCES

- Biermann JS, Holt GE, Lewis VO, Schwartz HS, Yaszemski MJ. Metastatic bone disease: diagnosis, evaluation, and treatment. *Instr Course Lect.* 2010;59:593-606.
- Choi D, Crockard A, Bunger C, Harms J, Kawahara N, Mazel C, et al. Review of metastatic spine tumour classification and indications for surgery: the consensus statement of the Global Spine Tumour Study Group. *Eur Spine J.* 2010;19(2):215-22.
- Yamashita T, Siemionow KB, Mroz TE, Podichetty V, Lieberman IH. A prospective analysis of prognostic factors in patients with spinal metastases: use of the revised Tokuhashi score. *Spine (Phila Pa 1976).* 2011;36(11):910-7.
- Kim HJ, Buchowski JM, Moussallem CD, Rose PS. Modern techniques in the treatment of patients with metastatic spine disease. *Instr Course Lect.* 2013;62:375-82.
- Klimo P Jr, Kestle JR, Schmidt MH. Clinical trials and evidence-based medicine for metastatic spine disease. *Neurosurg Clin N Am.* 2004;15(4):549-64.
- Papastefanou S, Alpentaki K, Akra G, Katonis P. Predictive value of Tokuhashi and Tomita scores in patients with metastatic spine disease. *Acta Orthop Traumatol Turc.* 2012;46(1):50-6.
- Valesin Filho ESV, Tardini R, Abreu LC, Motter BV, Adams F, Rodrigues LM. Estudo epidemiológico de 55 pacientes portadores de doença vertebral metastática sintomática em santo André-SP, Brasil. *Coluna/Columna.* 2013;12(1):32-5.
- Fisher CG, DiPaola CP, Ryken TC, Bilsky MH, Shaffrey CI, Berven SH, 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(22):E1221-9.
- Bilsky MH, Lauder I, Burch S. Shifting paradigms in the treatment of metastatic spine disease. *Spine (Phila Pa 1976).* 2009;34(22 Suppl):S101-7.
- Tokuhashi Y, Matsuzaki H, Oda H, Oshima M, Ryu J. A revised scoring system for preoperative evaluation of metastatic spine tumor prognosis. *Spine (Phila Pa 1976).* 2005;30(19):2186-91.
- Tokuhashi Y, Ajiro Y, Umezawa N. Outcome of treatment for spinal metastases using scoring system for preoperative evaluation of prognosis. *Spine (Phila Pa 1976).* 2009;34(1):69-73.
- Fourney DR, Frangou EM, Ryken TC, DiPaola CP, Shaffrey CI, Berven SH, et al. Spinal instability neoplastic score: an analysis of reliability and validity from the spine oncology study group. *J Clin Oncol.* 2011;29(22):3072-7.
- Harrington KD. Metastatic disease of the spine. *J Bone Joint Surg Am.* 1986;68(7):1110-5.
- Tomita K, Kawahara N, Kobayashi T, Yoshida A, Murakami H, Akamaru T. Surgical strategy for spinal metastases. *Spine (Phila Pa 1976).* 2001;26(3):298-306.
- Fleiss JL. *Statistical methods for rates and proportion.* New York: John Wiley; 1981.
- Bland JM, Altman DG. *Statistical methods for assessing agreement between two methods of clinical measurement.* *Lancet.* 1986;1(8476):307-10.
- Teixeira WG, Coutinho PR, Marchese LD, Narazaki DK, Cristante AF, Teixeira MJ, et al. Interobserver agreement for the spine instability neoplastic score varies according to the experience of the evaluator. *Clinics (Sao Paulo).* 2013;68(2):213-8.
- Harel R, Angelov L. Spine metastases: current treatments and future directions. *Eur J Cancer.* 2010;46(15):2696-707.
- Silva HJSS, Neto MIR, Pratali RR, Zuiani GR, Cavali PTM, Veiga IG, et al. Avaliação da reprodutibilidade interobservadores de uma nova escala para orientação da conduta terapêutica nas metástases vertebrais- escore SINS. *Coluna/Columna.* 2012;11(4):287-9.
- Leithner A, Radl R, Gruber G, Hochegger M, Leithner K, Welkerling H, et al. Predictive value of seven preoperative prognostic scoring systems for spinal metastases. *Eur Spine J.* 2008;17(11):1488-95.
- Majeed H, Kumar S, Bommireddy R, Klezl Z, Calthorpe D. Accuracy of prognostic scores in decision making and predicting outcomes in metastatic spine disease. *Ann R Coll Surg Engl.* 2012;94(1):28-33.
- Avanzi O, Landim E, Meves R, Caffaro MFS, Lima MV. Fratura na coluna vertebral por mieloma múltiplo: correlação de sobrevida e índices de Tokuhashi e Tomita. *Coluna/Columna.* 2009;8(1):73-9.