

# Agreement in Metastatic Spinal Cord Compression

Estanislao Arana, MD, MHE, PhD<sup>a,b,c</sup>; Francisco M. Kovacs, MD, PhD<sup>c,d</sup>; Ana Royuela, PhD<sup>e,f</sup>; Beatriz Asenjo, MD, PhD<sup>c,g</sup>; Úrsula Pérez-Ramírez, MSc<sup>c,h</sup>; Javier Zamora, PhD<sup>c,e,f,i</sup>; and the Spanish Back Pain Research Network Task Force for the Improvement of Inter-Disciplinary Management of Spinal Metastasis\*

## Abstract

**Background:** Metastatic epidural spinal cord compression (ESCC) is a devastating medical emergency. The purpose of this study was to determine the reliability of the 6-point ESCC scoring system and the identification of the spinal level presenting ESCC. **Methods:** Clinical data and imaging from 90 patients with biopsy-proven spinal metastases were provided to 83 specialists from 44 hospitals. The spinal levels presenting metastases and the ESCC scores for each case were calculated twice by each clinician, with a minimum of 6 weeks' interval. Clinicians were blinded to assessments made by other specialists and their own previous assessment. Fleiss kappa ( $\kappa$ ) statistic was used to assess intraobserver and interobserver agreement. Subgroup analyses were performed according to clinicians' specialty (medical oncology, neurosurgery, radiology, orthopedic surgery, and radiation oncology), years of experience, and type of hospital. **Results:** Intraobserver and interobserver agreement on the location of ESCC was substantial ( $\kappa > 0.61$ ). Intraobserver agreement on the ESCC score was "excellent" ( $\kappa = 0.82$ ), whereas interobserver agreement was substantial ( $\kappa = 0.64$ ). Overall agreement with the tumor board classification was substantial ( $\kappa = 0.71$ ). Results were similar across specialties, years of experience and hospital category. **Conclusions:** The ESCC score can help improve communication among clinicians involved in oncology care.

*J Natl Compr Canc Netw* 2016;14(1):70–76

## Background

Malignant or metastatic spinal cord compression of the thecal sac is a devastating medical emergency presented by 5% to 20% of patients with spinal metastases.<sup>1</sup> It can be caused by vertebral collapse, but is usually provoked by soft tissue causing epidural spinal cord compression (ESCC).<sup>2</sup>

Clinical symptoms and the ESCC grade are the major determinants in the decision to operate or irradiate.<sup>3,4</sup> The ESCC score system is a 6-point scale for diagnosing and reporting ESCC based on imaging findings (Figure 1).<sup>3</sup> It was developed by oncologic spine surgeons and proven to be reliable among a small sample of these specialists.<sup>3</sup> However, managing spinal cord compression requires a multidisciplinary approach,<sup>5,6</sup> and the lack of nomenclature standardization prevents agreement

in decision-making,<sup>7</sup> delays appropriate treatment, and hinders treatment effectiveness.<sup>8,9</sup>

Therefore, the purpose of this study was to assess intraobserver and interobserver agreement in identification of spine level involved in each patient and in the calculation of the ESCC score among a large sample of clinicians from different specialties with varied degrees of experience and working in different settings and locations.

## Methods

This study was approved by the Institutional Review Boards of the participating hospitals, and complied with the Guidelines for Reporting Reliability and Agreement Studies (GRRAS).<sup>10</sup>

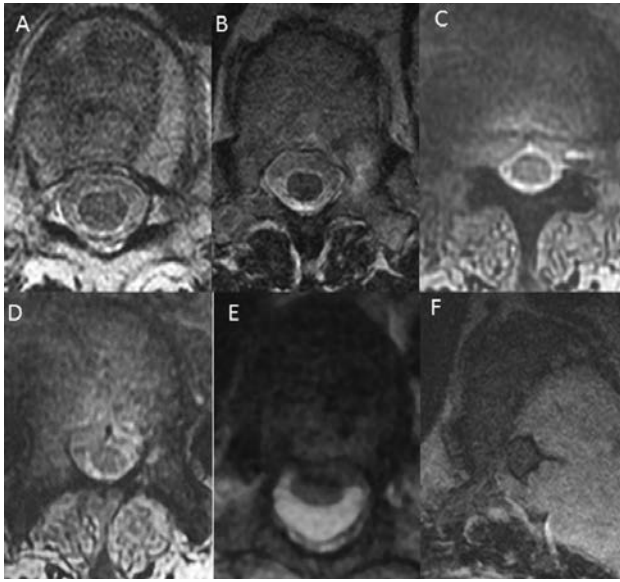
From the <sup>a</sup>Department of Radiology, Valencian Oncology Institute Foundation, Valencia, Spain; <sup>b</sup>Research Institute in Health Services Foundation, Valencia, Spain; <sup>c</sup>Spanish Back Pain Research Network, Kovacs Foundation, Palma de Mallorca, Spain; <sup>d</sup>Scientific Department, Kovacs Foundation, Palma de Mallorca, Spain; <sup>e</sup>CIBER Epidemiology and Public Health (CIBERESP), Madrid, Spain; <sup>f</sup>Clinical Biostatistics Unit, Hospital Ramón y Cajal, IRYCIS, Madrid, Spain; <sup>g</sup>Department of Radiology, Hospital Regional Universitario Carlos Haya, Málaga, Spain; <sup>h</sup>Center for Biomaterials and Tissue Engineering, Universitat Politècnica de València, Valencia, Spain; and <sup>i</sup>Barts and the London School of Medicine & Dentistry, Queen Mary University of London, London, United Kingdom.

\*The view members of the Spanish Back Pain Research Network Task Force, see supplemental eAppendix 1 (available with this issue, at JNCCN.org). Submitted June 22, 2015; accepted for publication September 4, 2015.

The authors have disclosed that they have no financial interests, arrangements, affiliations, or commercial interests with the manufacturers of any products discussed in this article or their competitors. Funding for this work was provided by The Kovacs Foundation.

Correspondence: Estanislao Arana, MD, MHE, PhD, Fundación Instituto Valenciano de Oncología, C/ Beltrán Báuena, 19, 46009 Valencia, Spain. E-mail: aranae@uv.es

## Spinal Cord Compression Agreement



**Figure 1.** The 6-point ESCC classification according to Bilsky et al.<sup>3</sup> (A) Grade 0 indicates bone-only disease. (B) Grade 1a, epidural impingement, without deformation of the thecal sac. (C) Grade 1b, deformation of the thecal sac, without spinal cord abutment. (D) Grade 1c, deformation of the thecal sac with spinal cord abutment, but without cord compression. (E) Grade 2, spinal cord compression, but with cerebrospinal fluid (CSF) visible around the cord. (F) Grade 3, spinal cord compression, no CSF visible around the cord.

### Selection of Hospital Departments and Clinicians

Sample size was calculated assuming that the prevalence of ESCC in a particular spinal level would be 10%, and that at least 5 readers would be recruited per medical specialty and 5 per hospital category. In order to ensure that, should  $\kappa$  values reflect an “almost perfect” agreement ( $\kappa=0.81$ ), the lower limit of the 95% CI would lie within the range classified as reflecting a substantial agreement ( $\kappa=0.61$ – $0.80$ ), the sample size was established at 90 patients. Sample size calculations were performed using  $\kappa$  size package of the R library.<sup>11</sup>

The 61 hospital departments specializing in radiology, medical oncology, radiation oncology, neurosurgery, and orthopedic surgery, which had previously participated in studies undertaken by the Spanish Back Pain Research Network (REIDE) or had expressed interest in doing so, were invited to participate in this study. Twelve departments were located in 6 private hospitals and the other 49 in 38 nonprofit hospitals, belonging to or working for the Spanish National Health Service (SNHS). The SNHS is the tax-funded, government-owned organization that provides free health care to every resident in Spain.

The SNHS classifies hospitals into 5 categories based on the size of the catchment area; number of beds; number of clinicians; availability of high-tech medical equipment and procedures; education, training, and academic activity; and clinical complexity of the cases treated (ie, being the reference hospital for specific diseases or procedures).<sup>12</sup> Category 1 is the simplest and category 5 is the most complex. Departments invited to participate in this study were located in hospitals belonging to categories 2, 3, 4, and 5.

All clinicians who had finished their residency and worked at the participating departments were invited to act as readers in this study, and asked to report the number of years they had been working in clinical practice after their residency. The departments and clinicians did not receive any compensation for participating in this study.

### Patients and Images Selection

A radiologist at a category 4 hospital, who did not act as reader, was responsible for selecting patients and images for study inclusion. He identified patients who had undergone an MRI in his department for spinal cord compression and whose ESCC scores had been rated by a tumor board.

The tumor board comprised a medical oncologist, a radiation oncologist, a radiologist, a pathologist, an orthopedic surgeon, and a neurosurgeon. None of its members acted as readers for the study.

For each case, demographic data, histopathology, and a pain description with an emphasis on neurologic signs were provided to simulate information typically provided to any physician in routine practice. All MR imaging had been performed with a 1.5-T unit (Magnetom Symphony; Siemens, Erlangen, Germany) with a spinal matrix coil. The recruiting radiologist selected 2 images per patient: a sagittal T2-weighted turbo spin-echo sequence (4000/115; section thickness, 4 mm) and an axial T2-weighted turbo spin-echo sequence (4500/110; section thickness, 5 mm) at maximal ESCC grade.<sup>3</sup> The sagittal MRI image included at least 2 spine segments.<sup>13</sup>

The first 90 cases that complied with inclusion criteria were selected. Inclusion criterion was presentation with stage IV (AJCC classification, 7th Edition, 2010) metastatic spine disease confirmed with biopsy. Exclusion criteria included clinical history lacking data required to assess ESCC or imaging of insufficient quality to assess the spinal levels affected.

Arana et al

## Procedure

The recruiting radiologist prepared an information pack corresponding to each patient, comprising 2 images and a clinical vignette that included the patient's age, oncologic history, clinical signs, and symptoms (Figure 2).<sup>3</sup> Patient identity was masked and a code was assigned to each information pack. All of the information packs were uploaded to an online platform designed for this study (<http://www.typeform.com/>).

Each reader was provided with a personal password to assess the information packs online. Readers were asked to indicate all the spinal segments in which they identified metastases for each patient (cervical, thoracic, lumbar, and/or sacral) and to calculate the ESCC score. They were only provided with definitions included in the ESCC (Figure 1). No attempt was made to further define or standardize the meaning of the terms included in the scoring systems or to homogenize the diagnostic criteria, and readers did not receive any instructions regarding the interpretation of images. They were told to use their own clinical judgment when in doubt.

Readers assessed the information pack on their own and uploaded the resulting report directly onto the online platform. They assessed the same clinical sets twice, with a minimum interval period of 6 weeks. The platform software ensured that the minimum interval period was observed, and that readers had no access to their own previous reports or to their colleagues' uploaded reports.

Data introduced into the platform were automatically converted into a spreadsheet. The software engineer in charge of developing the platform cross-checked that the data in the database

matched the information introduced into the platform by the readers.

## Statistical Analysis

At the analysis phase, grades 1a, 1b, and 1c were grouped, resulting in a 4-point ESCC: 0, 1 (including 1a, 1b, and 1c), 2, and 3. Data on the spine level affected in each patient was classified as cervical, thoracic, lumbar, or sacral, and rated as yes or no.

To assess intraobserver and interobserver agreement, ratings from each reader were cross-tabulated and the Fleiss  $\kappa$  statistic was calculated. A weighted  $\kappa$  approach, with a bisquared weighting scheme, was used to analyze the agreement when using the ESCC scoring system. Kappa values were categorized as “almost perfect” (0.81–1.00), “substantial” (0.61–0.80), “moderate” (0.41–0.60), “fair” (0.21–0.40), “slight” (0.00–0.20), and “poor” (<0.00).<sup>14</sup>

To assess intraobserver agreement for each variable (ESCC score and levels involved), a  $\kappa$  index was calculated for each of the 83 readers, and median, 5th and 95th percentiles values were calculated.

To assess interobserver agreement, the corresponding  $\kappa$  index was calculated and the 95% CI was determined following the jackknife resampling method.<sup>15</sup>

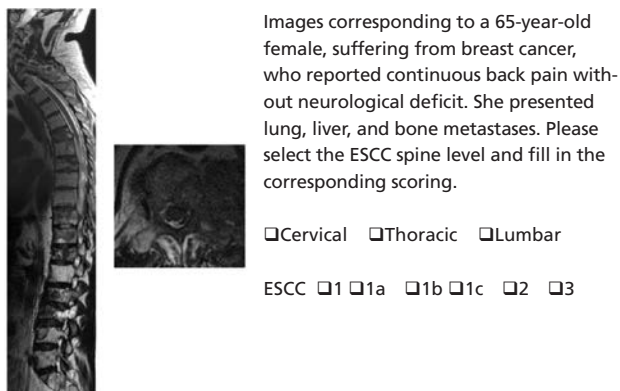
Subgroup analyses for each variable were performed, in which  $\kappa$  values were calculated separately depending on medical specialty, hospital category, and professional experience. Professional experience was classified as “junior” ( $\leq 7$  years in practice, after residency), “experienced” (8–13 years), and “senior specialist” ( $\geq 14$  years).

The ESCC scores established by the tumor board were subsequently classified into grades 0, 1, 2, and 3. These grades were used as the gold standard to assess overall agreement. The agreement between this gold standard and the median score for each image among the 83 readers was calculated through the  $\kappa$  statistic.

Stata 13 software was used (StataCorp 2013; Stata Statistical Software: Release 13, College Station, TX).

## Results

Of the 132 clinicians invited to act as readers, 83 (62.87%) participated in this study. The first 90 patients selected by the recruiting radiologist (51 women and 39 men; mean age, 60.8 years) complied with the inclusion criteria, and none were excluded. These 90 patients presented metastases in 182 spinal segments. Table 1 shows sample characteristics.



**Figure 2.** An example of the information pack provided to readers for each patient.

## Spinal Cord Compression Agreement

Table 1. Sample Characteristics	
Characteristic	n (%)
<b>Hospitals</b>	44
Degree of complexity <sup>a</sup>	
Category 2	3 (6.8)
Category 3	11 (25.0)
Category 4	9 (20.4)
Category 5	21 (47.7)
Management	
Nonprofit <sup>b</sup>	38
For profit <sup>c</sup>	6
<b>Departments</b>	61
Radiology	19 (31.1)
Radiation oncology	11 (18.0)
Orthopedic surgery	12 (19.7)
Neurosurgery	12 (19.7)
Medical oncology	7 (11.5)
<b>Readers</b>	83
Specialty	
Radiology	23 (27.7)
Radiation oncology	22 (26.5)
Orthopedic surgery	16 (19.3)
Neurosurgery	14 (16.9)
Medical oncology	8 (9.6)
Years in practice (post-residency)	
Junior	27 (32.5)
Experienced	25 (30.1)
Senior specialist	31 (37.4)
Setting	
Category of hospital in which they work <sup>a</sup>	
Category 2	3 (3.6)
Category 3	25 (30.1)
Category 4	19 (22.9)
Category 5	36 (43.4)
Hospital management	
Nonprofit <sup>b</sup>	71
For profit <sup>c</sup>	12
<b>Patients</b>	90
Age, y [mean (SD)]	60.8 (12.3)
Male sex	39 (43.3)
<b>Spine segments showing ESCC<sup>d</sup></b>	
Cervical	12 (13.3)
Thoracic	52 (57.8)
Lumbar	26 (28.9)
<b>ESCC grades<sup>d</sup></b>	
0	14 (15.6)
1	45 (50.0)
1a	7 (7.8)
1b	20 (22.2)
1c	18 (20.0)
2	21 (23.3)
3	10 (11.1)

Abbreviation: ESCC, epidural spinal cord compression.

<sup>a</sup>Category of hospital; complexity (eg, based on size, availability of high-tech medical equipment and procedures, education activity) ranges from category 1 (the simplest; none of this type were included in this study) to category 5 (the most complex). See text for details.

<sup>b</sup>Hospitals belonging to the Spanish National Health Service (SNHS) or to charities working for the SNHS.

<sup>c</sup>Hospitals under private ownership and management.

<sup>d</sup>According to tumor board and where therapeutic treatment was planned.

There were more than 5 readers for each specialty and degree of professional experience. However, only 3 readers worked at category 2 hospitals; therefore, agreement for this subgroup was not calculated (Tables 2 and 3).

Regarding the identification of spine levels showing ESCC, intraobserver and interobserver agreements were substantial ( $\kappa$  value: median, 0.772 [5th, 95th percentiles: 0.541, 0.948], and  $\kappa$  value: 0.610 [95% CI, 0.531, 0.696], respectively) (Table 2). Subgroup analyses showed that interobserver agreement was only moderate among junior specialists, those working in category 3 hospitals, and in all specialties except radiation oncology (Table 2).

Regarding ESCC score, intraobserver agreement was almost perfect ( $\kappa$  value: median, 0.819 [5th, 95th percentiles: 0.636, 0.923]), and interobserver agreement was substantial ( $\kappa$  value: 0.635 [95% CI, 0.578, 0.699]). Subgroup analyses showed that intraobserver agreement was only substantial among junior specialists, orthopedic surgeons, medical oncologists, and radiation oncologists, whereas interobserver agreement was only moderate among junior and experienced specialists, orthopedic surgeons, and medical oncologists, and among those working in category 3 and 4 hospitals (Table 3).

The agreement between the median of the scores calculated by the readers, and the ESCC grades based on the scores established by the tumor board, was 0.713 (95% CI, 0.596–0.835). Classification by readers and by the tumor board coincided in all of the 31 patients in whom compressive findings were observed (ESCC 2 and 3) (Table 4).

## Discussion

Results from this study show that there is a substantial interobserver agreement in determining the ESCC score. These results are generally consistent; differences across specialties, number of years of experience, and type of hospital are small. Results from this study show that there is substantial interobserver agreement in determining the ESCC score. Although some differences in  $\kappa$  values across hospitals, specialties, and number of years of experience were documented, the 95% CI of these values overlap, and differences are small and likely to be clinically meaningless (Tables 2 and 3).<sup>10</sup>

Table 2. Intraobserver and Interobserver Agreement on Level of Cord Compression		
	Intraobserver Agreement <sup>a</sup>	Interobserver Agreement <sup>b</sup>
<b>Global agreement</b>	0.772 (0.541; 0.948)	0.610 (0.531; 0.696)
<b>Subgroup analyses</b>		
By specialty		
Orthopedic surgery	0.767 (0.541; 0.882)	0.479 (0.221; 0.781)
Neurosurgery	0.768 (0.589; 0.996)	0.547 (0.319; 0.821)
Medical oncology	0.612 (0.522; 0.806)	0.448 (0.329; 0.667)
Radiation oncology	0.747 (0.508; 0.884)	0.720 (0.597; 0.867)
Radiology	0.841 (0.672; 0.959)	0.576 (0.377; 0.791)
By years of practice		
Junior	0.751 (0.522; 0.903)	0.513 (0.361; 0.689)
Experienced	0.752 (0.541; 0.963)	0.673 (0.552; 0.816)
Senior specialist	0.790 (0.646; 0.943)	0.609 (0.464; 0.770)
By setting (category of hospital) <sup>c</sup>		
Category 2 <sup>d</sup>	---	---
Category 3	0.752 (0.522; 0.959)	0.591 (0.458; 0.752)
Category 4	0.722 (0.541; 0.996)	0.667 (0.523; 0.843)
Category 5	0.780 (0.590; 0.943)	0.626 (0.499; 0.768)

<sup>a</sup>K values: median (5th; 95th percentiles).

<sup>b</sup>K values (95% CI).

<sup>c</sup>Complexity (eg, based on size, availability of high tech medical equipment and procedures, education activity) ranges from category 1 (the simplest; none of this category were included in this study) to category 5 (the most complex). See text for details.

<sup>d</sup>Only 3 specialists working in category 2 hospitals participated in this study. Therefore, agreement was not calculated for this subgroup.

The substantial agreement in identifying the spinal level showing ESCC based on MRI is reassuring, because clinical symptoms are unreliable for selecting the target level.<sup>13</sup>

This study assessed the reliability of the ESCC score across the different specialties involved in the

management of ESCC, in conditions as close as possible to routine practice. All patients showed lesions at 2 or more spine levels and clinicians had to identify the target vertebra based on clinical judgment, as is often the case in clinical practice.<sup>16</sup> A high number of readers participated, and they had differ-

Table 3. Intraobserver and Interobserver Agreement on Spinal Cord Compression Score		
	Intraobserver Agreement <sup>a</sup>	Interobserver Agreement <sup>b</sup>
<b>Global agreement</b>	0.819 (0.636; 0.923)	0.635 (0.578; 0.699)
<b>Subgroup analyses</b>		
By specialty		
Orthopedic surgery	0.788 (0.567; 0.972)	0.484 (0.328; 0.692)
Neurosurgery	0.828 (0.723; 0.991)	0.689 (0.571; 0.861)
Medical oncology	0.697 (0.498; 0.840)	0.486 (0.334; 0.726)
Radiation oncology	0.766 (0.639; 0.884)	0.626 (0.533; 0.753)
Radiology	0.859 (0.806; 0.928)	0.682 (0.572; 0.823)
By years of practice		
Junior	0.789 (0.567; 0.885)	0.594 (0.495; 0.720)
Experienced	0.827 (0.615; 0.923)	0.595 (0.501; 0.717)
Senior specialist	0.828 (0.654; 0.969)	0.678 (0.582; 0.799)
By setting (category of hospital) <sup>c</sup>		
Category 2 <sup>d</sup>	---	---
Category 3	0.816 (0.567; 0.871)	0.593 (0.493; 0.720)
Category 4	0.817 (0.615; 0.991)	0.564 (0.442; 0.726)
Category 5	0.819 (0.645; 0.923)	0.687 (0.598; 0.798)

<sup>a</sup>K values: median (5th; 95th percentiles).

<sup>b</sup>K values (95% CI).

<sup>c</sup>Complexity (eg, based on size, availability of high tech medical equipment and procedures, education activity) ranges from category 1 (the simplest; none of this category were included in this study) to category 5 (the most complex). See text for details.

<sup>d</sup>Only 3 specialists working in category 2 hospitals participated in this study. Therefore, agreement was not calculated for this subgroup.

## Spinal Cord Compression Agreement

**Table 4. Cross-Tabulation of Scores Determined by ESCC Board Tumor, and Median Categorization of Readers**

Score	Board Tumor				Total
	0	1	2	3	
Median ESCC					
0	14	6	0	0	20
1	0	33	1	0	34
2	0	5	18	3	26
3	0	1	2	7	10
Total	14	45	21	10	90

Abbreviation: ESCC, epidural spinal cord compression.

ent backgrounds and worked in different hospitals located in different regions; most readers had never met their colleagues in person. Contrary to some previous studies, the present one did not implement any measures to improve agreement,<sup>17</sup> such as training, consensus, offering a stipend to readers, agreeing on diagnostic criteria, or using standardized nomenclature linked to examples available online.<sup>18</sup> However, results are consistent with those from studies in which spine surgeons used the same scoring system,<sup>3</sup> and those in which radiosurgery experts used an ad hoc version of the Weinstein-Boriani-Biagini scoring system.<sup>19</sup> These results support the use of scoring systems, because the intraobserver and interobserver agreement on the size, location, and shape of tumors is very low when MRI images are analyzed without using such systems, irrespective of physicians' specialty.<sup>20,21</sup>

The similarity of results obtained by physicians, irrespective of years of experience and hospital category, is also generally consistent with previous studies and supports current routine practice<sup>19</sup>; patients with cancer in whom ESCC is suspected undergo MRI at their hospital and are referred to surgery when deemed appropriate.<sup>1</sup> It is reassuring that all of the patients who experienced compressive grades of ESCC (grades 2 and 3), and therefore required urgent clinical management, were correctly identified in this study (Table 4).<sup>19,22</sup> This is consistent with previous studies that showed the sensitivity of the ESCC scoring system for detecting such cases.<sup>3</sup> However, this does not necessarily imply a perfect external validity, because gold standards are difficult to define for metastatic spine disease,<sup>23,24</sup> and discrepancies between imaging and real surgical outcomes do exist.<sup>7</sup>

Good communication among clinicians involved in the management of spine metastatic dis-

ease leads to consistency of care, which is a prerequisite for effectiveness.<sup>22</sup> For instance, good communication between surgeons and radiation oncologists facilitates rapid identification of patients with epidural disease in whom surgical resection improves results from postoperative stereotactic body radiotherapy (SBRT).<sup>4</sup> Ensuring that the diagnostic instruments used are reliable, is probably the most effective means of decreasing inappropriate variability in health care.<sup>25</sup> Results from this study suggest that using the ESCC score can be useful to ensure accurate communication among multidisciplinary team members and, therefore, should be used routinely.<sup>6</sup> However, it should be kept in mind that the intrinsic characteristics of certain tumors make it impossible to reach good agreement when it comes to their assessment and management, even after repeated training.<sup>26</sup> Furthermore, agreement when using a scoring system does not necessarily mean that the resulting recommended treatment is appropriate, because clinicians sometimes agree on measures that are not evidence-based or effective,<sup>27</sup> and an improvement in the quality of oncologic care does not necessarily translate immediately into better clinical results or improved survival rates.<sup>28</sup> In fact, no current scoring system is robust enough to establish a solid prognosis for all patients with spinal metastases.<sup>29</sup>

This study has some limitations. Readers were only provided with 2 selected images per case. It is possible that providing all of the readers with all of the images available for each patient might have changed the degree of agreement. However, providing a selection of images ensures that all of the readers assess the same ones, and is consistent with the procedure followed by high-quality studies assessing reliability.<sup>3,30</sup> Readers were volunteers from each of the hospital departments participating in this study, and were not randomly selected. Therefore, selection bias may exist; it is possible that physicians who agreed to participate in this study were the most motivated or interested in metastatic spine disease.<sup>31</sup> However, clinicians involved in management of spinal metastases in routine clinical practice are usually highly specialized, and this does not challenge the results from this study. The prevalence of patients with grades 1a, 1b, and 1c ESCC made it necessary to merge these categories into a single category (grade 1). Maintaining the 3 subcategories would have led to groups too small for the  $\kappa$  statistic to reliably re-

flect the degree of agreement.<sup>32</sup> In fact, subscale analysis of grade 1 ESCC was not performed in this study or in its original design.<sup>3</sup> Moreover, there are no guidelines on the dosage of irradiation suitable for such cases<sup>19</sup>; for instance, more careful planning of radiotherapy is required for grade 1c ESCC than for grades 1a and 1b in order to avoid reaching the dosage above which the risk of spinal cord radiation overdose and myelopathy increases significantly.<sup>3,33</sup>

## Conclusions

This study suggests that there is substantial agreement among radiologists, medical oncologists, radiation oncologists, orthopedic surgeons, and neurosurgeons when identifying the spinal level affected by metastases and when using the ESCC scoring system. Therefore, although there is room for improvement, the use of the ESCC score in clinical practice could improve communication among specialists involved in the management of spine metastases.

## Acknowledgments

The authors wish to thank the following investigators of Spanish Back Pain Research Network Task Force for the improvement of interdisciplinary management of spinal metastasis (see supplemental eAppendix 1). The authors are also grateful to Prof. David Moratal for his continuous collaboration.

## References

1. Sutcliffe P, Connock M, Shyangdan D, et al. A systematic review of evidence on malignant spinal metastases: natural history and technologies for identifying patients at high risk of vertebral fracture and spinal cord compression. *Health Technol Assess* 2013;17:1–274.
2. Schmidt MH, Klimo P, Vronis FD. Metastatic spinal cord compression. *J Natl Compr Cancer Netw* 2005;3:711–719.
3. Bilsky MH, Laufer I, Fourny DR, et al. Reliability analysis of the epidural spinal cord compression scale. *J Neurosurg Spine* 2010;13:324–328.
4. Al-Omar A, Masucci L, Masson-Cote L, et al. Surgical resection of epidural disease improves local control following postoperative spine stereotactic body radiotherapy. *Neuro Oncol* 2013;15:1413–1419.
5. L'esperance S, Vincent F, Gaudreault M, et al. Treatment of metastatic spinal cord compression: CEPO review and clinical recommendations. *Curr Oncol* 2012;19:e478–490.
6. Arana E, Kovacs FM, Royuela A, et al. Agreement in the assessment of metastatic spine disease using scoring systems. *Radiother Oncol* 2015;115:135–140.

7. Putz C, Gantz S, Bruckner T, et al. Preoperative scoring and limits of prognostication: functional outcome after surgical decompression in metastatic spinal cord compression. *Oncology* 2014;86:177–184.
8. Brooks FM, Ghatahora A, Brooks MC, et al. Management of metastatic spinal cord compression: awareness of NICE guidance. *Eur J Orthop Surg Traumatol* 2014;24(Suppl 1):S255–259.
9. McGivern UM, Drinkwater KJ, Clarke JI, Locke I. A Royal College of Radiologists national audit of radiotherapy in the treatment of metastatic spinal cord compression and implications for the development of acute oncology services. *Clin Oncol (R Coll Radiol)* 2014;26:453–460.
10. Kottner J, Audigé L, Brorson S, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol* 2011;64:96–106.
11. Donner A, Rotondi MA. Sample size requirements for interval estimation of the kappa statistic for interobserver agreement studies with a binary outcome and multiple raters. *Int J Biostat* 2010;6:Article 31.
12. [Spanish hospitals classification by cluster analysis.] 2007. Available at: <http://www.icmbd.es/docs/resumenClusterHospitales.pdf>. Accessed December 7, 2015.
13. Venkitaraman R, Sohaib SA, Barbachano Y, et al. Frequency of screening magnetic resonance imaging to detect occult spinal cord compromise and to prevent neurological deficit in metastatic castration-resistant prostate cancer. *Clin Oncol* 2010;22:147–152.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159.
15. Efron B. Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods. *Biometrika* 1981;68:589–99.
16. Keil S, Barabach A, Dirrichs T, et al. Target lesion selection: an important factor causing variability of response classification in the Response Evaluation Criteria for Solid Tumors 1.1. *Invest Radiol* 2014;49:509–517.
17. Obuchowski NA. How many observers are needed in clinical studies of medical imaging? *AJR Am J Roentgenol* 2004;182:867–869.
18. 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.
19. Cox BW, Spratt DE, Lovelock M, et al. International Spine Radiosurgery Consortium consensus guidelines for target volume definition in spinal stereotactic radiosurgery. *Int J Radiat Oncol Biol Phys* 2012;83:e597–605.
20. Wu DH, Mayr NA, Karatas Y, et al. Interobserver variation in cervical cancer tumor delineation for image-based radiotherapy planning among and within different specialties. *J Appl Clin Med Phys* 2005;6:106–110.
21. Dobbs HJ. Defining the radiation target on a daily basis. *Cancer Imaging* 2006;6:30–32.
22. Terezakis SA, Heron DE, Lavigne RF, et al. What the diagnostic radiologist needs to know about radiation oncology. *Radiology* 2011;261:30–44.
23. Bankier AA, Levine D, Halpern EF, Kressel HY. Consensus interpretation in imaging research: is there a better way? *Radiology* 2010;257:14–17.
24. Harel R, Angelov L. Spine metastases: current treatments and future directions. *Eur J Cancer* 2010;46:2696–707.
25. Tuijn S, Janssens F, Robben P, van den Bergh H. Reducing interrater variability and improving health care: a meta-analytical review. *J Eval Clin Pract* 2012;18:887–895.
26. Rock JB, Washington MK, Adsay NV, et al. Debating deposits: an interobserver variability study of lymph nodes and pericolic tumor deposits in colonic adenocarcinoma. *Arch Pathol Lab Med* 2014;138:636–642.
27. Dea N, Fisher CG. Evidence-based medicine in metastatic spine disease. *Neurol Res* 2014;36:524–529.
28. Keating NL, Landrum MB, Lamont EB, et al. Tumor boards and the quality of cancer care. *J Natl Cancer Inst* 2013;105:113–121.
29. Tokuhashi Y, Uei H, Oshima M, Ajiro Y. Scoring system for prediction of metastatic spine tumor prognosis. *World J Orthop* 2014;5:262–271.
30. 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.
31. Lim C, Cheung MC, Franco B, et al. Quality improvement: an assessment of participation and attitudes of medical oncologists. *J Oncol Pract* 2014;10:e408–414.
32. Lantz CA, Nebenzahl E. Behavior and interpretation of the kappa statistic: Resolution of the two paradoxes. *J Clin Epidemiol* 1996;49:431–434.
33. Sahgal A, Weinberg V, Ma L, et al. Probabilities of Radiation Myelopathy Specific to Stereotactic Body Radiation Therapy to Guide Safe Practice. *Int J Radiat Oncol* 2013;85:341–347.



See [JNCCN.org](http://JNCCN.org) for supplemental online content.

Supplemental online content for:

## Agreement in Metastatic Spinal Cord Compression

*Estanislao Arana, MD, MHE, PhD; Francisco M. Kovacs, MD, PhD; Ana Royuela, PhD; Beatriz Asenjo, MD, PhD; Úrsula Pérez-Ramírez, MSc; Javier Zamora, PhD; and the Spanish Back Pain Research Network Task Force for the Improvement of Inter-Disciplinary Management of Spinal Metastasis*

J Natl Compr Canc Netw 2016;14(1):70–76

- **eAppendix 1:** Members of the Spanish Back Pain Research Network Task Force for the Improvement of Inter-Disciplinary Management of Spinal Metastasis



## eAppendix 1: Members of the Spanish Back Pain Research Network Task Force for the Improvement of Inter-Disciplinary Management of Spinal Metastasis (in alphabetical order)

Víctor Abaira<sup>1,2,3</sup>; Lucía Alcázar<sup>1,4</sup>; Ana Alonso<sup>1,5</sup>; Luis Álvarez<sup>1,6</sup>; Marco Antonio Álvarez<sup>1,7</sup>; Guillermo Amengual<sup>1,8</sup>; Aida Antuña<sup>1,9</sup>; Fernando Aparici<sup>1,10</sup>; Andrés Barriga<sup>1,11</sup>; María Barrios<sup>1,12</sup>; Paloma Bas<sup>1,10</sup>; José Begara<sup>1,13</sup>; Francisco Bravo-Rodríguez<sup>1,14</sup>; Alberto Cabrera<sup>1,15</sup>; Carlos Casillas<sup>1,16</sup>; Gregorio Catalán<sup>1,17</sup>; Antonio José Conde<sup>1,16</sup>; Ramón de las Peñas<sup>1,16</sup>; Laura Díaz<sup>1,18</sup>; Diego Dualde<sup>1,19</sup>; Ana Estremera<sup>1,8</sup>; Joaquín Fenollosa<sup>1,12</sup>; Carlos Fernández<sup>1,20</sup>; Eva Fernández<sup>1,21</sup>; Nicomedes Fernández-Baillo<sup>1,22,23</sup>; Pilar Ferrer<sup>1,24</sup>; Salvador Fuster<sup>1,25</sup>; María Isabel Galarraga<sup>1,26</sup>; Cristina García-Villar<sup>1,18</sup>; Luis García-Ferrer<sup>1,27</sup>; Sara García-Duque<sup>1,23</sup>; Javier Garde<sup>1,27</sup>; Andrés González<sup>1,28</sup>; Rafael González-Díaz<sup>1,29</sup>; Alberto Hernández-Fernández<sup>1,30</sup>; Ovidio Hernando<sup>1,23</sup>; Raúl Hernanz<sup>1,21</sup>; Asunción Hervás<sup>1,21</sup>; Esther Holgado<sup>1,23</sup>; María José Juan<sup>1,12</sup>; Javier Lavernia<sup>1,12</sup>; Antonio Lazo<sup>1,31</sup>; Ana Lersundi<sup>1,30</sup>; Escarlata López<sup>1,32</sup>; Margarita Majem<sup>1,33</sup>; Antonio Martín-Benlloch<sup>1,34</sup>; María Isabel Martín<sup>1,12</sup>; Javier Martínez<sup>1,21</sup>; Julia Montoya<sup>1,6</sup>; Paloma Moreno<sup>1,13</sup>; Arturo Navarro<sup>1,35</sup>; Esther Noguerón<sup>1,9</sup>; Ana Ortiz de Mendivil<sup>1,23</sup>; Julio César Palomino<sup>1,36</sup>; Juan Carlos Paniagua<sup>1,37</sup>; David Pereira<sup>1,25</sup>; Luis A. Pérez-Romasanta<sup>1,37</sup>; Rocío Pérez<sup>1,38</sup>; Ángel Ramón Piñera<sup>1,6</sup>; Pilar Piñero<sup>1,39</sup>; Julio Plata-Bello<sup>1,40</sup>; José Poblete<sup>1,26</sup>; José Ramírez<sup>1,9</sup>; Daniel Rivas<sup>1,32</sup>; Héctor Roldán<sup>1,40</sup>; Fernando Ruiz<sup>1,41</sup>; José Miguel Sánchez<sup>1,22,23</sup>; Helena Sarasíbar<sup>1,8</sup>; Juan Manuel Sepúlveda<sup>1,42</sup>; Antonio Silvestre<sup>1,19</sup>; Beatriz Sobrino<sup>1,6</sup>; Félix Tomé-Bermejo<sup>1,6</sup>; Isabel Tovar<sup>1,41</sup>; María del Carmen Vallejo<sup>1,21</sup>; Vicente Vanaclocha<sup>1,43</sup>; Asunción Villanueva<sup>1,38</sup>; Joaquín Zamarro<sup>1,44</sup>; and Idoya Zazpe<sup>1,45</sup>

<sup>1</sup>Spanish Back Pain Research Network, Kovacs Foundation, Palma de Mallorca, Spain

<sup>2</sup>CIBER Epidemiology and Public Health (CIBERESP), Madrid, Spain

<sup>3</sup>Clinical Biostatistics Unit, Hospital Ramón y Cajal, IRYCIS, Madrid, Spain

<sup>4</sup>Hospital Universitario la Princesa, Madrid, Spain

<sup>5</sup>Hospital Universitario Rey Juan Carlos, Móstoles, Madrid, Spain

<sup>6</sup>Hospital Universitario Fundación Jiménez Díaz, Madrid, Spain

<sup>7</sup>Hospital Universitario Central de Asturias, Asturias, Spain

<sup>8</sup>Hospital Son Llàtzer, Palma de Mallorca, Spain

<sup>9</sup>Complejo Hospitalario Universitario de Albacete, Albacete, Spain

<sup>10</sup>Hospital Universitario y Politécnico La Fe, Valencia, Spain

<sup>11</sup>Hospital Nacional de Paraplégicos, Toledo, Spain

<sup>12</sup>Fundación Instituto Valenciano de Oncología, Valencia, Spain

<sup>13</sup>Instituto Oncológico Xanit, Benalmádena, Málaga, Spain

<sup>14</sup>Hospital Universitario Reina Sofía, Córdoba, Spain

<sup>15</sup>OSATEK. Hospital de Galdakao, Vizcaya, Spain

<sup>16</sup>Consorcio hospitalario Provincial de Castellón, Castellón, Spain

<sup>17</sup>Hospital Universitario Cruces, Barakaldo, Vizcaya, Spain

<sup>18</sup>Hospital Universitario Puerta del Mar, Cádiz, Spain

<sup>19</sup>Hospital Clínico Universitario de Valencia, Valencia, Spain

<sup>20</sup>Hospital General Universitario Gregorio Marañón, Madrid, Spain

<sup>21</sup>Hospital Universitario Ramón y Cajal, Madrid, Spain

<sup>22</sup>Hospital Universitario la Paz, Madrid, Spain

<sup>23</sup>HM Universitario Sanchinarro, Madrid, Spain

<sup>24</sup>Hospital Intermutual de Levante, San Antonio de Benagéber, Valencia, Spain

<sup>25</sup>Hospital Clinic de Barcelona, Barcelona, Spain

<sup>26</sup>Hospital de Manacor, Manacor, Islas Baleares, Spain

<sup>27</sup>Hospital Arnau de Vilanova, Valencia, Spain

<sup>28</sup>Hospital Universitario Marqués de Valdecilla, Santander, Spain

<sup>29</sup>Hospital Universitario Fundación Alcorcón, Alcorcón, Madrid, Spain

<sup>30</sup>Hospital Universitario Donostia, Donostia, Spain

<sup>31</sup>Clínica Oncosur, Córdoba, Spain

<sup>32</sup>Clínica Oncosur, Granada, Spain

<sup>33</sup>Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

<sup>34</sup>Hospital Doctor Peset, Valencia, Spain

<sup>35</sup>Hospital Duran i Reynals, l'Hospitalet de Llobregat, Barcelona, Spain

<sup>36</sup>Hospital POVISA, Vigo, Spain

<sup>37</sup>Hospital Universitario de Salamanca, Salamanca, Spain

<sup>38</sup>Hospital Regional Universitario Carlos Haya, Málaga, Spain

<sup>39</sup>Hospital Universitario Virgen del Rocío, Sevilla, Spain

<sup>40</sup>Hospital Universitario de Canarias, Islas Canarias, Spain

<sup>41</sup>Hospital Universitario Virgen de las Nieves, Granada, Spain

<sup>42</sup>Hospital Universitario 12 de Octubre, Madrid, Spain

<sup>43</sup>Hospital 9 de Octubre, Valencia, Spain

<sup>44</sup>Hospital Universitario Virgen de la Arrixaca, Murcia, Spain

<sup>45</sup>Complejo Hospitalario de Navarra, Navarra, Spain