



# Establishment and evaluation of a prognostic model for surgical outcomes of patients with atlanto-axial dislocations

Shuai Guo, Jie Chen, Baohui Yang and Haopeng Li

## Abstract

**Objective:** Atlanto-axial dislocations (AADs) are potentially fatal disturbances with high spinal cord compression syndrome. As surgeons are still uncertain who is likely to benefit the most from surgery, a prediction tool is needed to provide decision-making support.

**Methods:** The model was established based on 108 patients with AADs using multiple binary logistic regression analysis and evaluated by calibration plot and the area under the receiver operating curve (AUC). Bootstrapping was used for internal validation.

**Results:** The prognostic model can be expressed as:  $\text{logit}(P) = -2.2428 + 0.3168\text{SCOPE} - 2.0375\text{SIGNAL}$ , in which two covariates were accepted (SCORE represents the preoperative modified Japanese Orthopedic Association (mJOA) score and SIGNAL represents the intramedullary hyperintense T<sub>2</sub>-weighted imaging (T<sub>2</sub>WI) with AUC = 0.8081).

**Conclusions:** The model was internally valid, and the preoperative mJOA score and hyperintense T<sub>2</sub>WI were important predictors of outcomes. The threshold was defined as  $\text{logit}(P) = -0.7282$  according to the receiver operating curve (ROC).

## Keywords

atlanto-axial dislocations, prognostic model, retrospective study

Date received: 3 May 2016; accepted: 28 July 2016

## List of abbreviations

AAD: atlanto-axial dislocation

AUC: area under the receiver operating curve

mJOA: modified Japanese Orthopaedic Association

T<sub>2</sub>WI: T<sub>2</sub>-weighted imaging

The Health Science Center and the Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, China

### Corresponding author:

Haopeng Li, The Health Science Center and the Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an 710004, Shaanxi, China.

Email: lhp3993@126.com



95% CI: 95% confidence interval

CT: computed tomography

## Introduction

Atlanto-axial dislocations (AADs) are rare and potentially fatal disturbances in the normal occipital-cervical anatomy that may cause a series of symptoms collectively referred to as high spinal cord compression syndrome. Possible symptoms include local pain, numb limbs with dyskinesia, and bladder and rectal dysfunction. Moreover, AADs can result in respiratory failure, circulatory collapse, and death if not appropriately treated.<sup>1</sup> Although sufficient decompression of the vertebral canal and consistent reconstruction of spinal stability have been recognized as possible methods of surgical treatment,<sup>1,2</sup> accurate prediction of which patients are most likely to benefit from such interventions remains difficult because AADs have several unique features that distinguish them from subaxial cervical spine injuries. The cranio-vertebral junction is very mobile because of its unique bony configuration and ligamentous attachments, which allow for various complex movements without compromising the neural tissue that passes through the area.<sup>3</sup> Thus, a prediction tool that provides decision-making support would be valuable in the treatment of AADs.

## Patients and methods

The study protocol was approved by the Biomedicine Ethics Committee of Xi'an Jiaotong University (clearance number M152039). This was a retrospective study, and the patients had already been discharged; therefore, with the permission and supervision of the Biomedicine Ethics Committee of Xi'an Jiaotong University, verbal (but not written) consent was obtained from each patient by calling them individually. All patients were verbally informed that only their clinical data were

needed and that their real names and identities would not be shared without their permission. Those who did not want to share their clinical data were excluded. The entire process, including the verbal consent, was recorded by the Biomedicine Ethics Committee of Xi'an Jiaotong University. This consent procedure was approved by the Biomedicine Ethics Committee of Xi'an Jiaotong University.

## Sample for internal validation

From September 2005 through April 2015, 149 consecutive patients with clinically diagnosed and image-confirmed AADs were admitted to the authors' department. Patients were eligible if they had no previous spine surgery, presented symptomatically with at least one clinical sign of myelopathy, and were radiographically confirmed to have complete relief from the spinal cord compression after surgery (i.e., the operation was successful and the outcome was favourable). Patients with a perfect preoperative modified Japanese Orthopaedic Association (mJOA) score of 18 were excluded because they had no room for improvement.

## Analytical methods

Extensive data were collected for each patient, including their age at surgery; sex; smoking status; preoperative and postoperative mJOA scores; co-morbidities in the spine or nervous system, such as syringomyelia or subaxial cervical spondylotic myelopathy; duration of symptoms; and intramedullary hyperintense T<sub>2</sub>-weighted imaging (T<sub>2</sub>WI) findings. According to Tschy and Benzel, 12-months period was chosen as the clinically threshold of the symptom durations, and patients were then divided into two groups.<sup>4</sup> The mJOA score evaluated at the 6-month follow-up visit was dichotomized for use as a dependent

variable for logistic regression. A successful outcome was defined as a final mJOA score of  $\geq 16$ , and failure was defined as a score of  $< 16$ . This 18-point scale was modified from the validated JOA assessment scale of Hirabayashi et al.,<sup>5</sup> as described in Benzel et al.,<sup>6</sup> and measures the motor and sensory function of the upper extremities, the motor function of the lower extremities, and the sphincter function. The threshold of 16 points was deemed clinically appropriate because it is within the range of mild impairment according to Tetreault et al.<sup>7</sup> The other items listed above were considered to be independent variables.<sup>8</sup>

Missing data were assumed to be missing completely at random and were replaced with the third set of plausible values using a multiple imputation procedure with five iterations. This procedure is more efficient than removing patients with incomplete variables.<sup>9</sup>

Internal validation was used as a strategy in the present study. Specifically, the model was developed based on the entire datasets, and bootstrapping, a technique for data reuse, was applied to assess performance considering the limited data available.<sup>10</sup> There were 10,000 individual bootstrap replicates in the bootstrap re-sampling procedure.

The model's calibration was assessed by plotting the observed proportions of successful outcomes against the predicted probabilities for each of 10 groups that were defined to have ranges of predicted risk of equal size. A calibration plot with all points on a diagonal line indicated perfect calibration, and positions relative to this line indicated whether the predictions were too high or too low. The model's discriminatory power was assessed by integrating the area under the receiver operating curve (AUC) using the trapezoidal rule. An AUC of 1.0 reflects perfect discriminatory power, and 0.5 indicates zero discriminatory power. Generally, the AUC is sufficiently high if it is  $\geq 0.8$ .

Calculation of the AUC was performed by calculating the average area of 10,000 replicates. The area of the raw data and the standard error obtained via bootstrapping are helpful for computing the 95% confidence interval (95% CI) of a normal distribution, whereas the percentile method is more suitable for a skewed distribution.<sup>10,11</sup>

Data entry and multiple imputation were implemented using IBM SPSS Statistics Version 22 (IBM Corp., Armonk, NY, USA). The additional processes were executed through a self-compiled program in SAS 9.3 (SAS Institute Inc., Cary, NC, USA).

## Results

### Sample

A total of 108 patients met all of the inclusion criteria and underwent surgical treatment for AAD (Table 1). The sample included 46 males and 62 females, and the patients' ages ranged from 12 to 76 years (mean = 43.7 years,  $P_{25} = 23$  years,  $P_{75} = 54$  years). The duration of their symptoms ranged from 8 hours to 33 years. The symptoms comprised neck pain in 75 patients, dizziness in 4, limited cervical movement in 56, pathologic reflexes in 71, sensory disturbances in 75, impaired gait in 68, quadriplegia after trauma in 13, defecation disorder in 32, and respiratory dysfunction in 1. The causes of AAD were congenital malformation or maldevelopment in 80 patients, trauma in 17, and pathological dislocation in 11 (of whom 6 had an infection, 2 had rheumatoid arthritis, 2 had ankylosing spondylitis, and 1 had an atlanto-axial tumour). No AADs were caused by idiopathic dislocation, which generally occurs during early childhood. According to the practical classification described by Wang et al.,<sup>12</sup> 36 AADs were classified as instabilities (type I), 19 as reducible dislocations (type II), 33 as irreducible dislocations (type III), and 20 as

**Table 1.** Patient characteristics.

<b>General characteristics</b>	
Age <sup>a</sup> (years)	43.7 ± 17.6 (12–76)
Sex (Male/Female)	46/62
Symptoms for > 1 year (Yes/No)	80/28
Baseline mJOA score <sup>a</sup>	10.9 ± 3.6 (0–17)
Six-month operative mJOA score ≥ 16 (Yes/No)	66/42
Smoking (Yes/No)	10/98
<b>Symptoms and signs (Yes/No)</b>	
Neck pain	75/33
Dizziness	4/104
Cervical movement limitation	56/52
Pathologic reflex	71/37
Sensory disturbance	75/33
Impaired gait	68/40
Quadriplegia after trauma	13/95
Defecation disorder	32/76
Respiratory dysfunction	1/107
<b>Co-morbidities in spine or nervous system (Yes/No)</b>	42/66
<b>Causes</b>	
Congenital malformation or maldevelopment	80
Trauma	17
Pathological dislocation	11
Infection	6
Rheumatoid arthritis	2
Ankylosing spondylitis	2
Atlanto-axial tumour	1
<b>Classifications</b>	
Instability (type I)	36
Reducible dislocation (type II)	19
Irreducible dislocation (type III)	33
Bony dislocations (type IV)	20
<b>Intramedullary hyperintensity on T<sub>2</sub>WI (Yes/No)</b>	34/74

<sup>a</sup>Values are presented as mean ± standard deviation, with the range in parentheses. mJOA, modified Japanese Orthopaedic Association; T<sub>2</sub>WI, T<sub>2</sub>-weighted imaging

bony dislocations (type IV). Notably, however, not all patients with type IV AADs had osseous fusion between C1 and C2. Thus, we modified the classification: AADs for which we could not achieve complete reduction with normal traction but for which we could achieve complete reduction with high-weight traction under general anaesthesia were classified as type III (irreducible dislocations), and AADs for which we could not achieve complete reduction with high-weight traction under general anaesthesia or

that exhibited osseous fusion between C1 and C2 on the reconstructive computed tomography (CT) scan were classified as type IV.

The patients subsequently underwent various types of operations according to their surgical classifications.<sup>12</sup> Types I and II AADs were fused in the reduced position from a posterior approach. Type III AADs, which are irreducible dislocations, were converted to reducible dislocations using a transoral atlanto-axial release and

subsequently treated by posterior fusion. Type IV AADs exhibited bony dislocations and required transoral osseous decompression prior to posterior fusion.<sup>12</sup> All of the operations were smoothly performed by the corresponding author (a chief physician) to ensure that the operation induced minimal damage to the spinal cord or nerves.

Posteroanterior, lateral, and dynamic lateral flexion and extension radiographic views; CT scans; and magnetic resonance images of the cervical vertebrae were routinely examined for all patients before the operation. Open-mouth plain films were obtained in patients with odontoid fractures. The preoperative baseline mJOA scores were also measured.

After the operations, the patients' necks were fixed with splints for 3 months, and their mJOA scores were measured 6 months later. During this period, we evaluated the effect of reduction and internal fixation using X-rays; rehabilitation exercises were under the direction of rehabilitation physiatrists, and early ambulation was actively encouraged.

### Model

The prediction model consisted of two statistically significant predictors: the baseline mJOA score and intramedullary hyperintensity on T<sub>2</sub>WI. The following equation was used to quantify the probability of a mild impairment by combining these two clinical variables (Eq. 1):

$$P = 1/[1 + \exp(2.2428 - 0.3168\text{SCORE} + 2.0357\text{SIGNAL})] \quad (1)$$

Equation 1 can also be expressed as follows (Eq. 2):

$$\text{logit}(P) = -2.2428 + 0.3168\text{SCORE} - 2.0357\text{SIGNAL} \quad (2)$$

In both equations, SCORE represents the baseline mJOA score (0–17) and SIGNAL

represents the presence of intramedullary hyperintensity on T<sub>2</sub>WI (0 = No, 1 = Yes).

After standardization, the disharmony of the various metric units was reconciled, and the corrected partial regression coefficients (i.e., b<sub>SCORE</sub>' and b<sub>SIGNAL</sub>') were as follows: b<sub>SCORE</sub>' = 0.0137 and b<sub>SIGNAL</sub>' = -0.6018.

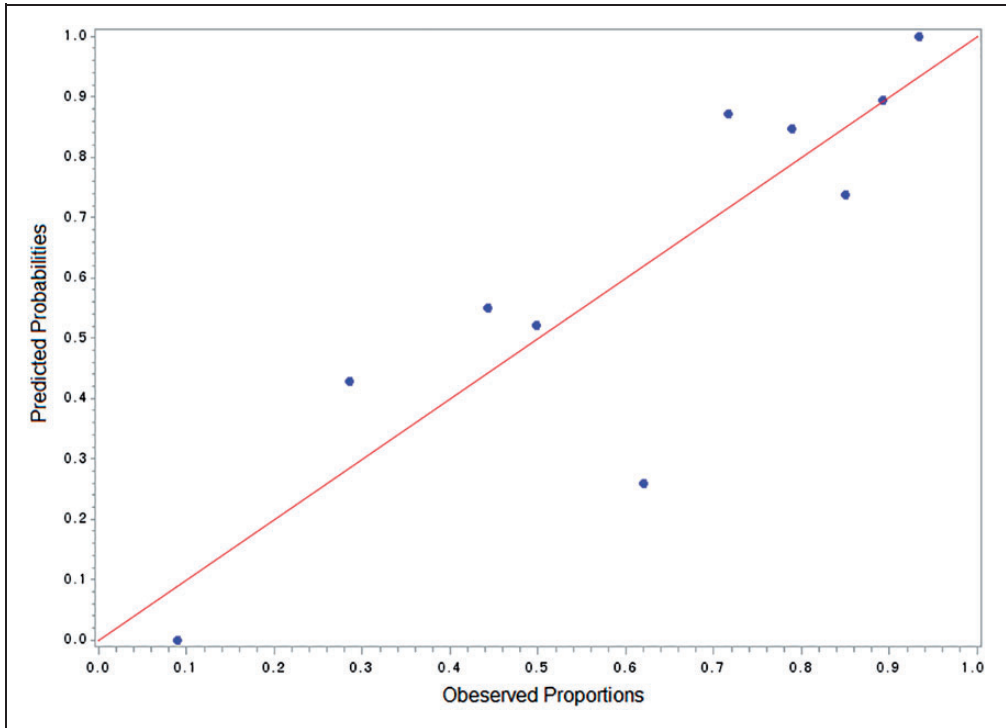
The calibration plot and receiver operating curve are illustrated in Figures 1 and 2. The AUC was 0.8081 (95% CI 0.7259, 0.8903) because the bootstrap replicates obeyed a normal distribution.

### Discussion

AAD refers to a loss of stability between the atlas and axis, which results in a loss of normal articulation. This condition can occur because of traumatic, inflammatory, idiopathic, or congenital abnormalities. High spinal cord compression syndrome is commonly associated with AADs. If this abnormality occurs, it can have devastating functional consequences, and the distinctive anatomy of the cranio-vertebral junction leads to various patterns of injury of the unique osseous structures or ligamentous connections.<sup>1,13–15</sup>

Given the variability among patients, a single variable rarely provides an adequate characterization of a patient's prognosis. Therefore, surgeons use multiple predictors. Prognostic models are valuable for identifying the most important predictors and providing outcome probabilities for various combinations of predictors. However, predicting outcomes is not synonymous with explaining their cause. Additionally, the calibration and discriminatory power of a multivariable model are highly relevant in prognostic research but meaningless in aetiological research.<sup>16</sup>

In the present study, the key outcome predictors were the baseline mJOA score and intramedullary hyperintensity on T<sub>2</sub>WI. Specifically, the baseline mJOA score was a protective factor with a standardized odds



**Figure 1.** Calibration plot for the model. A calibration plot shows the observed proportions of successful outcomes versus the predicted probabilities. A calibration plot with all points on a diagonal line indicates perfect calibration, and the distribution of these points indicates whether the predictions are too high or too low. This figure indicates that the model was well-calibrated but slightly pessimistic.

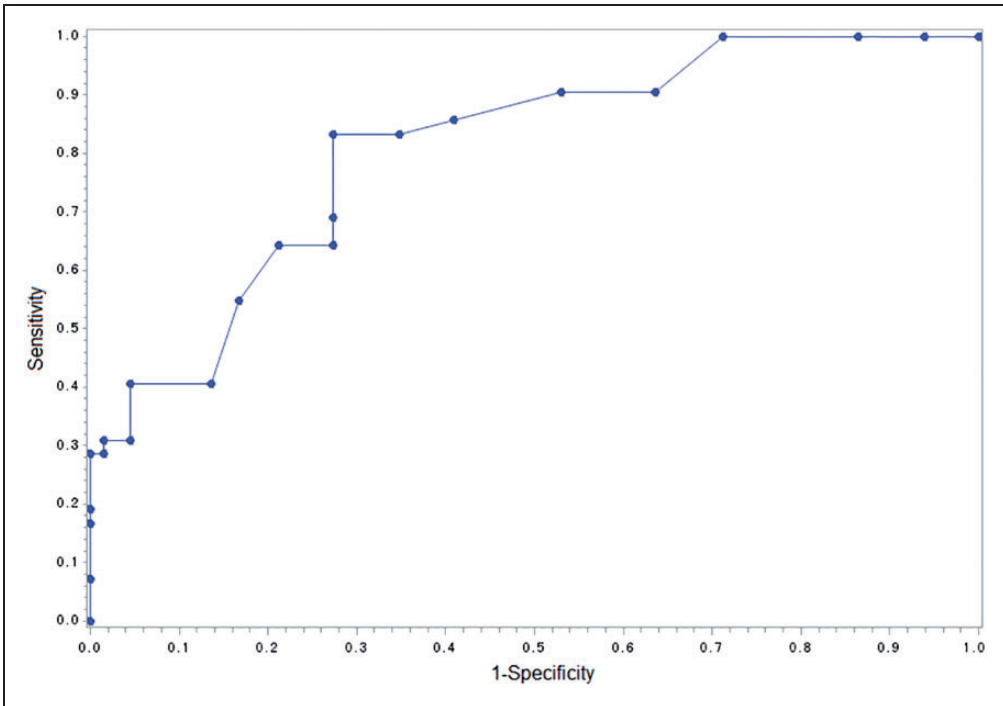
ratio of  $e^{0.0137}$  (e.g.,  $\text{logit}(P)$  increased by 0.0137 for an additional standard unit of the baseline mJOA score). In contrast, intramedullary hyperintensity on T<sub>2</sub>WI seemed to be a risk factor, with an odds ratio of  $e^{-0.6018}$  (i.e., if hyperintensity on T<sub>2</sub>WI existed,  $\text{logit}(P)$  decreased by 0.6018).

The predictive performance of the model was excellent in terms of both its calibration and discriminatory power, and the observed proportions and predicted probabilities agreed across the entire range of probabilities. The model assigns a higher probability of an mJOA score of  $\geq 16$  to patients who will develop AADs compared with those who will not. A defect of this model is that the forecast may be slightly pessimistic. Patients with mild myelopathy or

substantial residual neurological impairment can be discriminated if  $\text{logit}(P)$  is greater than or equal to  $-0.7282$  based on the receiver operating curve, which is the fifth value in ascending order.

The present study revealed that both the baseline severity score and hyperintensity on T<sub>2</sub>WI are important clinical factors for predicting surgical outcomes. The rationale is that either severe or chronic compression of the spinal cord may lead to irreversible damage due to demyelination and necrosis of the grey matter.<sup>4,7,17-19</sup>

The lack of significance of the symptom duration, another accepted predictor, may have been due to the interaction between the duration and hyperintensity. A total of 53% of researchers chose 65 years as the



**Figure 2.** Receiver operating curve for the model. A receiver operating curve plots the true-positive rate (sensitivity) against the false-positive rate ( $1 - \text{specificity}$ ). The discriminatory power of a prediction model can be assessed by calculating the area under the receiver operating curve. An area of 1.0 indicates a perfect test, whereas an area of 0.5 indicates no discriminatory power. This calculated area of 0.8081 indicates good discriminatory power (with “good” defined as an area of  $>0.8$ ).

threshold age, above which there is a negative impact on surgical outcomes. Although this variable was ignored in a limited amount of older people (14 of the 108 patients were aged  $\geq 65$  years), surgeons should be aware that older adult patients do not translate neurological recovery to functional improvement as well as younger patients. The roles of sex, smoking status, and spine or nervous system co-morbidities remain unclear; these variables also lacked statistical significance in other studies.<sup>4,7,20</sup>

The group of patients in the present study may have been too heterogeneous, considering that their duration of symptoms ranged from 8 hours to 33 years; however, the purpose of this study was to establish a pervasive model. Additionally, some readers

may question why a simple “yes” or “no” response was used to evaluate high signal intensity in the spinal cord on T<sub>2</sub>WI. It is true that imaging examinations do not always provide clear answers; as previously mentioned, however, the calibration and discriminatory power of a multivariable model are highly relevant in prognostic research but meaningless in aetiological research. But still the authors admit that the quantitative analysis is a precise method to reflect the extent of T<sub>2</sub>WI signal variation and it will be used in future studies.

A notable case in the present study involved a 76-year-old non-smoking woman with a 2-year history of persistent symptoms and no hyperintensity on T<sub>2</sub>WI. Surgeons predicted that the patient’s

condition would likely remain unchanged according to her 6-point preoperative mJOA score. The model yielded a logit(P) that exceeded the threshold level, and the patient did indeed improve. The same situation occurred in a 45-year-old woman. However, a similar case involving a 15-year-old girl resulted in a poor outcome; her preoperative mJOA score was 13. These examples demonstrate discrepancies between intuitive perceptions and predicted probabilities and emphasize the urgent need for further improvements to the predictive model. Moreover, four surgeries were postponed because of complications from diabetes, and one patient with a 40-pack-year history of smoking cigarettes died within 2 years after surgery because of respiratory failure. Therefore, future studies should focus on how systemic diseases affect recovery.

The major strength of the proposed model is that it can be used to determine which patients are likely to benefit the most from surgical treatment of upper cervical spine injuries. This method can aid surgeons in their predictions.

As mentioned previously, two limitations of the study are that the sample size was insufficient and that the data were retrospective. The findings from the authors' department should be evaluated in a wider population using more stringent validation strategies, such as temporal and external validation.<sup>10</sup> Another limitation was the use of a threshold for the postoperative mJOA score; this measurement remains controversial. This limitation should be addressed in future studies.<sup>4</sup>

## Conclusions

A clinical prediction model for assessing the surgical outcome of patients with AADs has been developed, and it was determined to be internally valid but slightly pessimistic. The baseline mJOA score and intramedullary

hyperintensity on T<sub>2</sub>WI were two important predictors. Patients whose logit(P) values were greater than  $-0.7282$  had a greater chance of achieving an mJOA score of  $\geq 16$ . Despite some disadvantages, this model can be of great value in the clinical setting.

## Acknowledgments

The authors thank Dr. Jie Qin for his aid during the research process and Chunsu Bai for providing comments on earlier versions of the manuscript. The authors also thank Lindsay R. and Christopher H., two staff members of American Journal Experts, for editing this manuscript. Finally, the authors thank Ms. Lanxuan Huang for her hospitality and logistical support during the project.

## Authors' contributions

GS and LH developed the study concept and design. GS, YB, and CJ acquired the data. GS programmed the bootstrapping algorithm. LH analysed the data. GS, YB, CJ, and LH drafted the manuscript. All authors were involved in the revision of the manuscript. All authors read and approved the final manuscript.

## Availability of data and materials

The authors would like to share the data corresponding to this research. The dataset supporting the conclusions of this study is included within the additional file of this article. For requirements, please contact the authors using the email address below.

## Declaration of conflicting interest

The authors declare that there is no conflict of interest.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.



## References

1. Yang SY, Boniello AJ, Poorman CE, et al. A review of the diagnosis and treatment of atlantoaxial dislocations. *Global Spine J* 2014; 4: 197–210.
2. Tan MS. Thoughts of surgical therapeutic principle and operative indicatio in atlantoaxial dislocation. *Zhongguo Gu Shang* 2012; 25: 1–3. [in Chinese, English Abstract].
3. Menezes AH and Traynelis VC. Anatomy and biomechanics of normal craniovertebral junction (a) and biomechanics of stabilization (b). *Childs Nerv Syst* 2008; 24: 1091–1100.
4. Techy F and Benzel EC. Predictors of outcome in patients with cervical spondylotic myelopathy undergoing surgical treatment: the evidence and the international common practice. *World Neurosurg* 2014; 81: 503–507.
5. Hirabayashi K, Miyakawa J, Satomi K, et al. Operative results and postoperative progression of ossification among patients with ossification of cervical posterior longitudinal ligament. *Spine (Phila Pa 1976)* 1981; 6: 354–364.
6. Benzel EC, Lancon J, Kesterson L, et al. Cervical laminectomy and dentate ligament section for cervical spondylotic myelopathy. *J Spinal Disord* 1991; 4: 286–295.
7. Tetreault LA, Kopjar B, Vaccaro A, et al. A clinical prediction model to determine outcomes in patients with cervical spondylotic myelopathy undergoing surgical treatment: data from the prospective, multi-center AOSpine North America study. *J Bone Joint Surg Am* 2013; 95: 1659–1666.
8. Royston P, Moons KG, Altman DG, et al. Prognosis and prognostic research: developing a prognostic model. *BMJ* 2009; 338: b604.
9. Tetreault LA, Côté P, Kopjar B, et al. A clinical prediction model to assess surgical outcome in patients with cervical spondylotic myelopathy: internal and external validations using the prospective multicenter AOSpine North American and international datasets of 743 patients. *Spine J* 2015; 15: 388–397.
10. Altman DG, Vergouwe Y, Royston P, et al. Prognosis and prognostic research: validating a prognostic model. *BMJ* 2009; 338: b605.
11. Gu J, Ghosal S and Roy A. Bayesian bootstrap estimation of ROC curve. *Stat Med* 2008; 27: 5407–5420.
12. Wang S, Wang C, Yan M, et al. Novel surgical classification and treatment strategy for atlantoaxial dislocations. *Spine (Phila Pa 1976)* 2013; 38: E1348–E1356.
13. Syre P, Petrov D and Malhotra NR. Management of upper cervical spine injuries: a review. *J Neurosurg Sci* 2013; 57: 219–240.
14. Joaquim AF, Ghizoni E, Tedeschi H, et al. Upper cervical injuries - a rational approach to guide surgical management. *J Spinal Cord Med* 2014; 37: 139–151.
15. Joaquim AF, Ghizoni E, Tedeschi H, et al. Upper cervical injuries: clinical results using a new treatment algorithm. *J Craniovertebr Junction Spine* 2015; 6: 16–20.
16. Moons KG, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: what, why, and how? *BMJ* 2009; 338: b375.
17. Tetreault LA, Dettori JR, Wilson JR, et al. Systematic review of magnetic resonance imaging characteristics that affect treatment decision making and predict clinical outcome in patients with cervical spondylotic myelopathy. *Spine (Phila Pa 1976)* 2013; 38(22 Suppl 1): S89–S110.
18. Zeitoun D, El Hajj F, Sariali E, et al. Evaluation of spinal cord compression and hyperintense intramedullary lesions on T2-weighted sequences in patients with cervical spondylotic myelopathy using flexion-extension MRI protocol. *Spine J* 2015; 15: 668–674.
19. Karadimas SK, Erwin WM, Ely CG, et al. Pathophysiology and natural history of cervical spondylotic myelopathy. *Spine (Phila Pa 1976)* 2013; 38(22 Suppl 1): S21–S36.
20. Tracy JA and Bartleson JD. Cervical spondylotic myelopathy. *Neurologist* 2010; 16: 176–187.