Efficacy of a machine learning-based approach in predicting neurological prognosis of cervical spinal cord injury patients following urgent surgery

Tomoaki Shimizu (✉ tomoaki.shimizu.96@gmail.com)  
Hokkaido Spinal Cord Injury Center

Kota Suda  
Hokkaido Spinal Cord Injury Center

Satoshi Maki  
Chiba University

Masao Koda  
University of Tsukuba

Satoko Matsumoto Harmon  
Hokkaido Spinal Cord Injury Center

Miki Komatsu  
Hokkaido Spinal Cord Injury Center

Masahiro Ota  
Hokkaido Spinal Cord Injury Center

Hiroki Ushirozako  
Hokkaido Spinal Cord Injury Center

Akio Minami  
Hokkaido Spinal Cord Injury Center

Masahiko Takahata  
Hokkaido University

Norimasa Iwasaki  
Hokkaido University

Hiroshi Takahashi  
University of Tsukuba

Masashi Yamazaki  
University of Tsukuba

Article

Keywords: cervical spinal cord injury, machine learning, multiclass classification, prediction models, neurological outcomes, urgent surgery
Abstract

Prediction of the neurological prognosis of cervical spinal cord injury (CSCI) is useful for setting treatment goals. This study aimed to develop a machine learning (ML) model for predicting neurological outcomes of CSCI. We retrospectively analyzed 135 patients with CSCI who underwent surgery within 24 hours after injury. American Spinal Injury Association impairment scales (AIS; grades A to E) were analyzed 6 months after injury as primary outcome measures. A total of 33 features extracted from demographic variables, surgical factors, laboratory variables, neurological status, and radiological findings were analyzed. The multiclass ML model was created using Light GBM, XGBoost, and CatBoost. The accuracy, recall, precision, and F1 score were calculated for each classification model. We evaluated Shapley Additive Explanations (SHAP) values to determine the variables that contributed the most to the prediction models. Of the ML models used, CatBoost had the highest accuracy (0.807), recall (0.574), precision (0.808), and F1 score (0.622). AIS grade at admission, intramedullary hemorrhage, longitudinal extent of intramedullary T2 hyperintensity, total motor index score of the lower extremity, and HbA1c were identified as important features for the prediction model. The ML models successfully predicted neurological outcomes for five grades of AIS 6 months after injury.

Introduction

Cervical spinal cord injury (CSCI) can result in neurological impairment and may impact daily activities and mental health. Prediction of the neurological prognosis of CSCI is useful for setting treatment goals and for reintegration planning and support. The elucidation of predictive factors is useful for improving treatment efficiency and cost-effectiveness. However, the neurological prognosis of CSCI is not easy to predict due to the complexity of various factors involved in clinical practice.

Machine learning (ML) is a subset of artificial intelligence that can identify confounding rules in a database to predict outcomes. When compared to traditional prognostic models, like logistic regression, ML has advantages of being able to determine complicated relationships in databases and is less constrained by the number of features that can be used in the prediction model. Previous studies reported that ML models were useful for predicting the outcomes of orthopedic conditions, such as hip fracture,1 carpal tunnel syndrome,2 ossification of the posterior longitudinal ligament (OPLL),3 and degenerative cervical myelopathy.4 ML can predict intensive care unit (ICU) stay5 and 1-year mortality6 in patients with spinal cord injuries. A previous study reported that ML models were useful for predicting neurological improvements in patients with CSCI.7 However, it included cases that received both surgical and conservative treatment, and the outcome of the prediction was a binary classification of the American Spinal Injury Association (ASIA) Impairment Scale (AIS), grades A/B/C or D/E. Several studies reported that early surgery within 24 hours after injury has a significant impact on the neurological prognosis of spinal cord injury.8,9 Therefore, the prognosis of a mixed population of surgically treated and conservatively treated patients may differ significantly at the initial stage of treatment selection. We hypothesized that prediction of the prognosis
of CSCI using ML can be achieved with high accuracy, even with a 5-level outcome of AIS grade A to E, if the target population is limited to surgically treated patients in acute phase. To our knowledge, there has been no report on multiclass classification prediction models for neurological prognosis limited to surgically treated patients with CSCI. The purpose of this study was to construct a ML model for predicting neurological prognosis based on AIS grade following early surgery in patients with CSCI. It also sought to determine factors that influence the prognosis of CSCI.

Results

Baseline patient characteristics

The flowchart of the study population was shown in Fig. 2. Of 262 patients identified, 10 were excluded due to AIS grade E at the initial examination, 50 were excluded due to receipt of conservative treatment, 36 were excluded due to a follow-up period < 6 months, four were excluded due to surgery at another hospital, 21 were excluded due to surgery more than 24 hours after injury, and six were excluded due to disturbed consciousness. Finally, a total of 135 patients met all inclusion criteria and were analyzed. Demographic variables and surgical factors are shown in Table 1. Laboratory variables, neurological status, and radiological findings on admission are shown in Table 2.
Table 1
Patient characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N = 135</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean ± SD), years</td>
<td>67.0 ± 12.7</td>
</tr>
<tr>
<td>Sex male / female, n</td>
<td>115 / 20</td>
</tr>
<tr>
<td>Height (mean ± SD), cm</td>
<td>165.6 ± 8.7</td>
</tr>
<tr>
<td>Body weight (mean ± SD), kg</td>
<td>67.3 ± 16.1</td>
</tr>
<tr>
<td>BMI (mean ± SD), kg/m²</td>
<td>24.4 ± 4.7</td>
</tr>
<tr>
<td>Follow up period (mean ± SD), months</td>
<td>14.4 ± 9.7</td>
</tr>
<tr>
<td>Time to initial evaluation (mean ± SD), hours</td>
<td>8.5 ± 6.2</td>
</tr>
<tr>
<td>Comorbidities, n</td>
<td></td>
</tr>
<tr>
<td>Alcohol</td>
<td>64</td>
</tr>
<tr>
<td>Smoking</td>
<td>27</td>
</tr>
<tr>
<td>Hypertension</td>
<td>55</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>16</td>
</tr>
<tr>
<td>Ischemic heart disease</td>
<td>11</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>28</td>
</tr>
<tr>
<td>Malignancy</td>
<td>26</td>
</tr>
<tr>
<td>Mental disease</td>
<td>6</td>
</tr>
<tr>
<td>Mechanism of injury, n</td>
<td></td>
</tr>
<tr>
<td>Fall on ground</td>
<td>52</td>
</tr>
<tr>
<td>Fall from high place</td>
<td>29</td>
</tr>
<tr>
<td>Fall down stairs</td>
<td>23</td>
</tr>
<tr>
<td>Traffic injury</td>
<td>25</td>
</tr>
<tr>
<td>Sports</td>
<td>6</td>
</tr>
<tr>
<td>Surgical procedure, n</td>
<td></td>
</tr>
<tr>
<td>Posterior fixation</td>
<td>133</td>
</tr>
<tr>
<td>Laminoplasty</td>
<td>2</td>
</tr>
</tbody>
</table>

BMI, body mass index; SD, standard deviation.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N = 135</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to surgery (mean ± SD), hours</td>
<td>10.3 ± 6.1</td>
</tr>
</tbody>
</table>

BMI, body mass index; SD, standard deviation.
Table 2  
Laboratory variables, neurological status, and radiological findings on admission

<table>
<thead>
<tr>
<th>Laboratory data, mean ± SD</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>White blood cell (WBC) count, ×10^9/L</td>
<td>9149 ± 2811</td>
</tr>
<tr>
<td>Lymphocyte count, ×10^9/L</td>
<td>1023 ± 372</td>
</tr>
<tr>
<td>ALB, g/dL</td>
<td>3.71 ± 0.36</td>
</tr>
<tr>
<td>CRP, mg/dL</td>
<td>0.63 ± 1.80</td>
</tr>
<tr>
<td>e-GFR, mL/min/1.73m²</td>
<td>75.9 ± 22.2</td>
</tr>
<tr>
<td>Hb, g/dL</td>
<td>12.7 ± 1.6</td>
</tr>
<tr>
<td>HbA1c, %</td>
<td>5.88 ± 0.65</td>
</tr>
<tr>
<td>AIS grade on admission, n</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>35</td>
</tr>
<tr>
<td>B</td>
<td>17</td>
</tr>
<tr>
<td>C</td>
<td>48</td>
</tr>
<tr>
<td>D</td>
<td>35</td>
</tr>
<tr>
<td>NLI, n</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>22</td>
</tr>
<tr>
<td>C3</td>
<td>30</td>
</tr>
<tr>
<td>C4</td>
<td>59</td>
</tr>
<tr>
<td>C5</td>
<td>11</td>
</tr>
<tr>
<td>C6</td>
<td>3</td>
</tr>
<tr>
<td>C7</td>
<td>4</td>
</tr>
<tr>
<td>C8</td>
<td>6</td>
</tr>
<tr>
<td>Total MIS, mean ± SD</td>
<td></td>
</tr>
<tr>
<td>Upper extremity</td>
<td>18.5 ± 15.1</td>
</tr>
<tr>
<td>Lower extremity</td>
<td>16.5 ± 18.0</td>
</tr>
</tbody>
</table>

AIS, ASIA impairment scale; ALB, serum albumin; ASIA, American Spinal Injury Association; CRP, C-reactive protein; CT, computed tomography; e-GFR, estimated glomerular filtration rate; Hb, hemoglobin; HbA1c, glycosylated hemoglobin; MCC, maximum canal compromise; MIS, motor index score; MRI, magnetic resonance imaging; MSCC, maximum spinal cord compression; NLI, neurological level of injury; SD, standard deviation.
Laboratory data, mean ± SD

<table>
<thead>
<tr>
<th></th>
<th>CT</th>
<th>MRI, mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPLL</td>
<td>42 / 135</td>
<td>25.8 ± 16.2 mm</td>
</tr>
<tr>
<td>DISH</td>
<td>22 / 135</td>
<td>29.0 ± 16.4%</td>
</tr>
<tr>
<td>MCC</td>
<td>24.2 ± 13.5%</td>
<td>24.2 ± 13.5%</td>
</tr>
<tr>
<td>MSCC</td>
<td>25 / 156</td>
<td></td>
</tr>
</tbody>
</table>

AIS, ALB, serum albumin; ASIA, American Spinal Injury Association; CRP, C-reactive protein; CT, computed tomography; e-GFR, estimated glomerular filtration rate; Hb, hemoglobin; HbA1c, glycosylated hemoglobin; MCC, maximum canal compromise; MIS, motor index score; MRI, magnetic resonance imaging; MSCC, maximum spinal cord compression; NLI, neurological level of injury; SD, standard deviation.

Changes in neurological status on AIS grade

Of 135 patients, 35 were AIS grade A, 17 were AIS grade B, 48 were AIS grade C and 35 were AIS grade D at admission. Six months after injury, 16 were AIS grade A, 9 were AIS grade B, 17 were AIS grade C, 87 were AIS grade D and 6 were AIS grade E. Changes in the neurological status of AIS grade from admission to 6 months after injury are shown in Table 3.

<table>
<thead>
<tr>
<th>AIS grade 6 months after injury</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIS grade at admission</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>16</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>0</td>
<td>35</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>10</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>43</td>
<td>0</td>
<td>48</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>29</td>
<td>6</td>
<td>35</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>9</td>
<td>17</td>
<td>87</td>
<td>6</td>
<td>135</td>
</tr>
</tbody>
</table>

AIS, ASIA impairment scale; ASIA, American Spinal Injury Association.
The accuracy, recall, precision, and F1 score for each classification model are shown in Table 4. CatBoost had the highest accuracy (0.807), the highest recall (0.574), the highest precision (0.808), and the highest F1 score (0.622). The confusion matrix between the actual and the predicted AIS grade with the CatBoost model 6 months after injury is shown in Table 5.

### Table 4
Comparison of performance of the multiclass classification models to predict neurological status based on AIS grade at 6 months after injury

<table>
<thead>
<tr>
<th>Model</th>
<th>Prediction of AIS grade at discharge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
</tr>
<tr>
<td>CatBoost</td>
<td>0.807</td>
</tr>
<tr>
<td>Light GBM</td>
<td>0.777</td>
</tr>
<tr>
<td>XGBoost</td>
<td>0.770</td>
</tr>
</tbody>
</table>

AIS, ASIA impairment scale; ASIA, American Spinal Injury Association.

### Table 5
Confusion matrix between actual AIS grade and predicted AIS grade with the CatBoost model 6 months after injury

<table>
<thead>
<tr>
<th>Predicted AIS grade with CatBoost model</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>E</td>
</tr>
</tbody>
</table>

AIS; ASIA impairment scale; ASIA, American Spinal Injury Association.

### Features affecting prediction models

The model's performance did not increase despite the use of recursive feature selection. To maximize performance, the ML model utilized all 33 independent variables. We evaluated SHAP values in the CatBoost model, which showed the highest accuracy, to determine the variables that contributed the most. The 5 most important features for predicting neurological outcomes 6 months after injury were AIS grade at admission, intramedullary hemorrhage, longitudinal extent of intramedullary T2 hyperintensity, total MIS of the lower extremity, and HbA1c (Fig. 3).

### Discussion
We used ML models to construct prediction models based on 33 parameters extracted from demographic, surgical, laboratory, neurological, and radiological factors. Among the models used, CatBoost showed the highest accuracy (0.807) for prediction of AIS grade 6 months after injury. Moreover, AIS grade at admission, intramedullary hemorrhage, longitudinal extent of intramedullary T2 hyperintensity, total MIS of the lower extremity, and HbA1c were identified as important predictors.

The present study showed that the accuracy of prediction of AIS at 6 months after injury was 0.807 using CatBoost, which demonstrated that ML could predict neurological prognosis following surgery in patients with CSCI. Inoue and colleagues proposed models predicting neurological outcomes 6 months after injury based on the binary classification of AIS grade A/B/C or D/E, and XGBoost had the highest accuracy (81.1%). For predicting reacquisition of walking ability, the predictive model with a binary classification of improvement to more than AIS grade D is useful. However, not only predicting motor ability but also predicting recovery of one AIS grade or two AIS grades (eg, AIS grade A to B or AIS grade A to C), is considered clinically meaningful neurological recovery from the patient’s perspective. Therefore, we proposed multiclass classification prediction models for neurological outcomes based on AIS grade. Despite the complexity of factors that contribute to CSCI, we were able to achieve accuracy on five grades of multiclass classification prediction models equivalent to the ML models that predicted binary classifications.

The multiclass classification prediction models identified AIS grade at admission, intramedullary hemorrhage, longitudinal extent of intramedullary T2 hyperintensity, total MIS of the lower extremity, and HbA1c as important features for the model predicting neurological outcomes based on AIS grade 6 months after injury. Although the method to identify these features was to evaluate their contribution to the prediction model, these results were consistent with the factors identified in previous studies using conventional statistical methods. Longitudinal extent of intramedullary T2 hyperintensity was a significant predictor of long-term neurological prognosis, according to Arabi and colleagues, and length of less than 30 mm was associated with a high likelihood of AIS grade improvement. Konomi and coworkers reported that the presence of intramedullary hemorrhage at initial MRI was a major risk factor for complete paralysis in the chronic phase. Both have been reported as negative predictive factors. Similarly, intramedullary MRI findings, such as longitudinal extent of intramedullary T2 hyperintensity and intramedullary hemorrhage, were identified as important features in this study. In contrast to these intramedullary findings, Haefeli and colleagues reported that neither MSCC nor MCC significantly correlated with AIS grade at discharge, and they were likewise less correlated with AIS grade 6 months after injury in our study. Based on these results, intramedullary findings may be more important than the severity and degree of spinal cord compression in predicting the prognosis of spinal cord injury.

The present study also identified HbA1c as important feature for the model predicting neurological outcomes. Hyperglycemia has been reported to increase the inflammatory response by inducing overactivation of NF-kB in microglial cells and exacerbating secondary injury. Kobayakawa reported that the HbA1c level exhibited negative relationships with both the ASIA motor score at discharge and the
recovery rate.\textsuperscript{14} Park reported that individuals with HbA1c < 6.5\% had lower concentrations of inflammatory biomarkers, such as IL-6 and IL-8, and better functional outcomes than those with HbA1c ≥ 6.5\%.\textsuperscript{15} Taking these findings into consideration, HbA1c negatively affected neurological prognosis after CSCI; hence, this feature may have been extracted as important for predicting neurological outcomes.

Early reduction of cervical spine dislocation was reported to be associated with favorable neurological outcomes in patients with complete motor paralysis.\textsuperscript{16} \textsuperscript{17} Therefore, early surgery for CSCI would have a positive impact on neurological prognosis, but was not extracted as an important factor in the prediction model of this study. There are two possible reasons for this result. First, there were a few variations in time to surgery in this study (10.3 ± 6.1 hours), because the study population was limited to patients operated on within 24 hours from injury. Second, this study included individuals with incomplete motor paralysis, such as AIS grade C or D, which may have offset the effect of early surgery. To investigate the true impact of time to surgery on patients with CSCI, further studies in patients with complete motor paralysis may be warranted.

This study has several limitations. First, our models were validated using cross-validation, although the validated cohort was still from the same database. Second, the number of patients was relatively small for a ML study. To address these limitations, further studies with large numbers of patients with CSCI will be warranted. Third, our prediction models were conducted with a database from a single institution. Therefore, the results may have limited generalizability because of different patient backgrounds, indications for surgery, and surgical procedures. Finally, the timing of the initial assessment was different for each patient. As previously reported, 72 hours after injury was considered more reliable following SCI to diagnose whether patients were in spinal shock.\textsuperscript{18} The timing of the initial assessment of this study was 8.5 hours after injury, which was much earlier on average. Therefore, it cannot be ruled out that some patients with spinal shock were included as patients with complete motor paralysis in this study. However, we thought that implementation of surgical interventions as early as possible may be more important for the prevention of secondary injury than an accurate diagnosis of spinal shock.

In conclusion, the present study showed that ML could predict neurological outcomes in five grades of AIS (A to E) 6 months after injury following surgery in patients with CSCI. Moreover, AIS grade at admission, intramedullary hemorrhage, longitudinal extent of intramedullary T2 hyperintensity, total MIS of the lower extremity, and HbA1c were identified as important features for the prediction model. The present study suggests the efficacy of the ML-based approach for the clinical management of patients with CSCI.

\textbf{Methods}

\textbf{Patient population}

This retrospective analysis of patient data was approved by the institutional review board of the Hokkaido Spinal Cord Injury Center. All procedures involving human participants were in accordance with...
the Declaration of Helsinki. The requirement for consent was waived by the institutional review board of the Hokkaido Spinal Cord Injury Center because of the retrospective design. Consecutive patients with traumatic CSCI admitted to a single institution from April 2017 to June 2021 were candidates. Based on the International Standards for Neurological Classification of Spinal Cord Injury,\(^\text{19}\) neurological classification was performed for all patients at admission using the AIS, neurological level of injury (NLI) and total motor index score (MIS) of the upper extremity and lower extremity. Two physicians conducted clinical assessments, one who had more than five years of experience examining patients with spinal cord injuries.

We excluded patients who met the following criteria: (1) AIS grade E at initial examination, (2) underwent conservative treatment, (3) follow-up period < 6 months, (4) underwent surgery at another hospital, (5) underwent surgery after 24 hours from injury, (6) neurological status not evaluable because of disturbed consciousness, such as brain injury or severe mental disorder.

**Measurements**

Demographic variables included age, sex, height, body weight, body mass index, comorbidities (hypertension, hyperlipidemia, ischemic heart disease, diabetes, malignancy, smoking history, mental disease, dementia), and mechanism of injury. Comorbidities were recorded based on patient self-report at admission. Mechanism of injury was classified as a fall on the ground, fall from a high place, fall down stairs, traffic injury, and sports injury. Laboratory variables were white blood cell (WBC) count, lymphocyte count, serum albumin (ALB), C-reactive protein (CRP), estimated glomerular filtration rate (e-GFR), hemoglobin (Hb), and HbA1c. Neurological status at admission included AIS grade, NLI, total MIS of the upper extremity, and total MIS of the lower extremity. Surgical timing was defined as the time from injury to the start of surgery.

Computed tomography (CT) and magnetic resonance imaging (MRI) data were obtained at admission before any treatment. Two spine surgeons assessed CT and MRI findings for all metrics while blinded to neurological outcomes. For CT evaluation, we assessed the presence of diffuse idiopathic skeletal hyperostosis (DISH)\(^\text{20}\) and ossification of the posterior longitudinal ligament (OPLL)\(^\text{21}\) at the level of the injury. Pre-operative MRI included T1- and T2-weighted imaging (WI) of the cervical spine in both the axial and sagittal views. All MRI studies were performed on a 3.0 T GE Discovery MR 750 (GE Healthcare). We assessed the longitudinal extent of intramedullary T2 hyperintensity,\(^\text{22}\)\(^\text{23}\) maximum canal compromise (MCC),\(^\text{13}\)\(^\text{24}\) maximum spinal cord compression (MSCC),\(^\text{13}\) and presence of intramedullary-confined low-intensity changes within diffuse high-intensity areas, which represents an intramedullary hemorrhage\(^\text{12}\) (Fig. 1).

**ML algorithms**

Among numerous ML algorithms, gradient boosting decision trees are reported to be effective in prediction models.\(^\text{25}\)\(^\text{26}\)\(^\text{27}\) The gradient boosting algorithm gradually combines weak learners such that
each new learner fits the residuals from the preceding phase, hence improving the model. We used three different gradient boosting tree-based classification models (Light GBM, XGBoost, CatBoost). All algorithms were implemented in Python 3.7 using scikit-learn library in a Google Collaboratory environment. The model’s hyperparameters were automatically optimized using Optuna.

Five grades (A, B, C, D, E) of the neurological outcome six months after injury were analyzed as primary outcome measures, based on previous studies indicating that the neurological recovery of CSCI reaches a plateau after six months.\textsuperscript{12,28}

With stratified 5-fold cross-validation, patients were placed into 5 evenly populated groups and randomly assigned to a training and validation set in a 4:1 ratio. The training set was used to train the ML model, and the validation set was used to evaluate ML model performance. Missing values were imputed by the mean value for the base models. To avoid overfitting and to ensure the generalization of the ML model, this cross-validation process was repeated five times.

**Performance of ML models**

The accuracy (number of correctly classified data instances over total number of data instances), recall (number of correct positive predictions made from all positives in a dataset), precision (metric for determining the accuracy of a positive prediction), and F1 score (computed as the harmonic mean of precision and recall) were calculated for each classification model.

**Feature selection**

We entered 33 variables into the prediction models, as listed below: demographic variables (age, sex, height, body weight, body mass index, eight comorbidities, mechanism of injury), laboratory variables on admission (WBC, LYM, ALB, CRP, e-GFR, Hb, HbA1c), surgical factors (surgical procedure, time to surgery), neurological status on admission (AIS, NLI, total MIS of upper extremity, total MIS of lower extremity), and radiological findings (OPLL, DISH, longitudinal extent of intramedullary T2 hyperintensity, MCC, MSCC, intramedullary hemorrhage).

After all 33 variables were entered into the LightGBM model, recursive feature selection was performed to find a subset of features for use in the final modeling.

**Features affecting prediction models**

Shapley Additive Explanations (SHAP) is a strategy for explaining individual predictions based on the game’s theoretically best Shapley values. Unlike conventional statistical methods, such as logistic regression, which evaluate the significance of variables for the outcome, SHAP values show the distribution of the effect that each feature has on the model’s output.\textsuperscript{29,30} We evaluated SHAP values in the prediction models with the highest accuracy to determine the factors that contributed the most for higher model confidence levels.

**Data availability**
The datasets analyzed during the current study are not publicly available due to their containing information that could compromise the privacy of research participants but are available from the corresponding author on reasonable request.

**Abbreviations**

AIS, ASIA impairment scale; Alb, serum albumin; ASIA, American Spinal Injury Association; BMI, body mass index; CRP, C-reactive protein; Hb, hemoglobin; HbA1c, glycosylated hemoglobin; L/E, lower extremity; MCC, maximum canal compromise; MIS, motor index score; MSCC, maximum spinal cord compression; NLI, neurological level of injury; T2WI, T2-weighted imaging; U/E, upper extremity; WBC, white blood cell

**Declarations**

The authors declare no competing interests.

No funds were received in support of this work.

No relevant financial activities outside the submitted work.

**Acknowledgements**

The authors would like to thank JAM Post (https://www.jamp.com) for English language editing.

**Author contributions**

T.S. wrote the manuscript. T.S., S.M.H, M.Kom., M.O., and H.U. collected the patient data. T.S. and S.M. developed the ML source code. T.S. and H.U. reviewed the radiological findings. T.S. performed data analysis, and prepared the figures and tables. K.S., S.M., and M.Kod. reviewed and revised the entire manuscript based on the results. A.M., M.T., N.I., H.T. and M.Y. supervised the work. All authors have read and approved the final manuscript.

**References**


Figures
Figure 1

MRI findings.

(A) Sagittal T2-weighted MR imaging of the cervical spine of patients with acute SCI. The longitudinal extent of intramedullary T2 hyperintensity (mm, black double-headed arrow) and intramedullary-confined low-intensity changes within diffuse high-intensity areas (white arrow) are demonstrated.

(B) Maximum spinal cord compression (MSCC) evaluated with mid-sagittal T2-weighted MR imaging. MSCC=\(1-\frac{(d_i/[(d_a + d_b)/2])}{100}\). \(d_i\) demonstrates the diameter of the spinal cord at the level of injury. \(d_a\) demonstrates the diameter of the spinal cord at one level above the level of injury. \(d_b\) demonstrates the diameter of the spinal cord at one level below the level of injury.

(C) Maximum canal compromise (MCC) evaluated with mid-sagittal T1-weighted MR imaging. MCC=\(1-\frac{(D_i/[(D_a + D_b)/2])}{100}\). \(D_i\) demonstrates the diameter of the spinal canal at the level of injury. \(D_a\) demonstrates the diameter of the spinal canal at one level above the level of injury. \(D_b\) demonstrates the diameter of the spinal canal at one level below the level of injury.
Figure 2

Flowchart of the study population.

AIS, ASIA impairment scale; ASIA, American Spinal Injury Association; CSCI, cervical spinal cord injury
Figure 3

Shapley Additive Explanations (SHAP) values obtained from the CatBoost model.

Variables are arranged by magnitude of the SHAP value. Features are coded as sex (0=female, 1=male); AIS grade at admission (A=1, B=2, C=3, D=4); mechanism of injury (1=fall on ground, 2=fall from stairs, 3=spot, 4=fall from high place, 5=traffic injury); surgical procedure (0=laminoplasty, 1=posterior fixation); and (0=no, 1=yes) for the following: alcohol, smoking, hypertension, hyperlipidemia, ischemic heart disease, diabetes mellitus, malignancy, mental disease, OPLL, DISH, and intramedullary hemorrhage.