



Machine Learning for Outcome Prediction and Prognostication in Spine Trauma: A Systematic Review of Current Evidence

Zheqing Zhang¹ , Caitlyn J. Smith², Jeremiah Oluwatomi Itodo Daniel³, Joanna Fedjo Tsague⁴, Kamala Thompson⁵ and Najib El Tecle⁶

OPEN ACCESS

This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

¹Medical Sciences Division, University of Oxford, Oxford, UK

²Department of Neurosurgery, University of Pennsylvania, Philadelphia, Pennsylvania, USA

³College of Health Sciences, Obafemi Awolowo University, Ile-Ife, Osun, Nigeria

⁴Georgia State University, Atlanta, Georgia, USA

⁵Dartmouth College, Hanover, New Hampshire, USA

⁶Department of Neurological Surgery, Feinberg School of Medicine, Northwestern University, Illinois, Chicago, USA

Correspondence to:

Zheqing Zhang, zheqing119@gmail.com

Additional material is published online only. To view please visit the journal online.

Cite this as: Zhang Z, Smith CJ, Daniel JOI, Tsague JF, Thompson K and Tecle NE. Machine Learning for Outcome Prediction and Prognostication in Spine Trauma: A Systematic Review of Current Evidence. Premier Journal of Science 2025;14:PJS.100188

DOI: <https://doi.org/10.70389/PJS.100188>

Peer Review

Received: 23 November 2025

Accepted: 23 November 2025

Version accepted: 1

Published: 27 December 2025

Ethical approval: N/a

Consent: N/a

Funding: No industry funding

Conflicts of interest: N/a

Author contribution:

Various as shown

Guarantor: Zheqing Zhang

Provenance and peer-review: Unsolicited and externally peer-reviewed

Data availability statement: N/a

ABSTRACT

Introduction: Spine trauma represents a major cause of disability and substantial healthcare costs. Despite classification systems and risk scores, outcome prediction following spine trauma remains challenging. Machine learning (ML) approaches offer multidimensionality but vary widely in methodology, performance, and clinical application. This systematic review aims to characterize the use of ML for outcome prediction and prognostication in spine trauma.

Methods: A systematic search of PubMed, Embase, Scopus, Web of Science, and Cochrane Library was performed for publications from January 2015 to May 2025. Studies reporting ML for predicting clinical outcomes in spine trauma patients were identified using pre-defined criteria. Primary outcomes were types and performance metrics of ML models for spine trauma prognostication. Secondary outcomes included cost-effectiveness, clinical utility and implementation feasibility.

Results: Twenty-seven studies were included, of which 26 were published from 2020 onwards. Only one study was prospective, while the rest were retrospective. Commonly used ML algorithms for spine trauma prognosis were logistic regression (LR), extreme gradient boosting (XGBoost), Random Forest (RF), light gradient boosting machine (LightGBM), and support vector machine (SVM). Outcomes predicted include mortality, neurological function or recovery, ambulation and walking, self-care and independence, complications, treatment failure, length of stay, and discharge destination. Most studies employed internal testing and validation; only four reported external validation. Seven studies reported direct comparison with non-ML models or risk scores, demonstrating comparable or superior performance of ML models. Performance metrics include area under the curve (AUC), accuracy, sensitivity, specificity, precision, F1-score, Brier score, and others. The highest AUC was reported by a nomogram constructed using consensus clustering to predict postoperative pulmonary infection in patients with acute cervical spinal cord injury (internal validation; AUC 0.993; 95% CI 0.981–1.000). Quality of evidence varied between studies, but risk of bias was a concern.

Conclusions: Existing evidence on ML for spine trauma prognosis is characterized by retrospective study designs, homogeneous cohorts, and lack of external validation, limiting their generalizability and real-world application. While ML approaches demonstrate encouraging predictive performance for spine trauma outcomes, their clinical adoption is limited by methodological heterogeneity. More robust validation, standardized reporting, and comparison with traditional tools are needed before ML models can reliably inform clinical decision-making.

Keywords: Spine trauma prognostication, Machine learning outcome prediction, Neurological recovery modeling, Retrospective cohort studies, External validation gaps