Review Article

Predictive modeling in spine surgery

Azeem Tariq Malik, Safdar N. Khan

Department of Orthopaedics, The Ohio State University Wexner Medical Center, Columbus, OH, USA

Contributions: (I) Conception and design: All authors; (II) Administrative support: SN Khan; (III) Provision of study materials or patients: AT Malik; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Safdar N. Khan, MD. Department of Orthopaedics, The Ohio State University Wexner Medical Center, Columbus, OH, USA. Email: Safdar.Khan@osumc.edu.

Abstract: As the cost of healthcare in the United States increases at an unsustainable rate, health-policy leaders are looking towards innovative ways to maximize value in delivery of care. Incorporating technology, such as artificial intelligence/machine-learning, to assist physicians in decision-making and predicting outcomes, on a real-time basis, is a major topic of discussion. While machine learning is gradually pulling traction in the medical community, it still remains a nascent field in the realm of spine surgery. The current review aims to gather current literature discussing the validity and applicability of machine-learning models in spine surgery.

Keywords: Predictive modeling; spine surgery; spinal fusions

Submitted Jul 16, 2019. Accepted for publication Jul 19, 2019.
doi: 10.21037/atm.2019.07.99
View this article at: http://dx.doi.org/10.21037/atm.2019.07.99

Introduction

The cost of healthcare in the United States is the highest in the world, accounting for more than 17% of the nation's gross domestic product in 2016 (1,2). Despite the high costs, there is wide variation in the delivery of quality care across the nation resulting in healthcare disparities and barriers to access of care. Due to a lack of value-driven insight in the healthcare system, 21–47% of the annual healthcare cost was attributable to avoidable events, such as readmissions, medical errors, inappropriate use of antibiotics and fraud. In order to combat the rising healthcare cost as well as introduce value in the healthcare system, health-policy makers, surgeons and public health scientists are increasingly applying artificial intelligence (AI) on "big data" to better understand short-falls in delivery of care. According to recent estimates, application of AI-based data mining can save the healthcare system more than $400 billion annually. Despite AI gaining a lot of traction in the medical community (3-5), the use and adoption of predictive analytic modeling in spine surgery still remains a nascent field. Contrary to other popular means of analysis, wherein the choice of variables going into a statistical model is dependent on the hypothesis of an investigator/researcher/statistician, predictive analytics are based on the principle of handing over the control of the data to the program. Predictive analytic algorithms are designed to interpret data using complex modeling to identify co-relations that would often be invisible to a researcher's eyes when looking at the data. The ability to extract information from complex "big data," and produce real-time clinical information that can be translated into decision-making offers physicians a potential avenue to launched value-based approaches in care.

This review aims to summarize current literature revolving around the use of predictive analytical modeling in spine surgery. To facilitate readers to better understand the applications of various models, we have focused on reviewing literature relevant to complications/readmissions, opioid dependence/usage and patient-reported outcomes following spine surgical procedures involving the cervical or thoracolumbar region.

General spine surgery complications

Around 2% to 23% of individuals undergoing spine surgery experience at least one adverse event and/or complication following spine surgery (6,7). While the exact rate of adverse event varies, based on surgical complexity, it is important to note that such complications/adverse events pose additional financial burdens on the healthcare system. Furthermore, as the healthcare system shifts away from traditional fee-for-service reimbursements to alternative payment models, complications will ultimately be tied to quality metrics that will determine prospective payments (8-10). Using "big data" to risk-stratify patients who may be at risk of developing complications will allow providers to launch appropriate peri-operative optimization protocols to mitigate the risk of experiencing these costly events. In addition, harnessing the power of AI to identify patients who are at a significantly higher risk of contracting complications can also benefit health administrators and surgeons in demanding higher prospective payments for caring for sick and vulnerable patients.In a recent paper published in *The Spine Journal*, Han *et al.* carried out a comprehensive machine-learning based analysis to identify which patients are at risk of experiencing 30-day adverse events (11). The authors utilized two national datasets (MarketScan and Medicare) to answer their research questions. The databases were queried for patients undergoing cervical or thoracolumbar surgery for four distinct cohorts: (I) degenerative spinal pathology, (II) malignancy/tumor, (III) tumor and (IV) infections. The authors found that Medicaid status was the single most important factor in predicting the risk of various complications after spine surgery. Among other factors, a diagnosis of infection, undergoing instrumentation in the presence of a pulmonary disorder, neurological disorder, a combined anterior-posterior fusion, and undergoing a fusion with Medicare insurance was associated with significantly higher rates of adverse events.

Using data from the ACS-NSQIP registry, Goyal *et al.* employed 7 distinct machine learning algorithms to create a predictive model to estimate the odds of non-home discharge and unplanned readmissions within 30 days of a cervical or lumbar fusion (12). The seven different classification algorithms used included a logistic regression (GLM), an elastic-net penalized GLM, a penalized linear discriminant analysis (pLDA), a naïve Bayesian model, artificial neural networks (ANN), random forest (RF) and gradient boosting machines (GBM). The authors used area under curve (AUC) to evaluate superiority and/or inferiority of different algorithms. For non-home discharge, the AUC was at least 0.85 across all algorithms with simple GLM having an AUC of 0.87. The authors noted that a simple GLM showed predictive performance similar to other remaining algorithms (GLMnet =0.87, ANN =0.87, GBM =0.87 and Bayesian =0.87). However, for unplanned readmissions the predictive performance was generally lower for all algorithms. Out of all models, ANN had the most superior performance with an accuracy of 0.71.

In another NSQIP-based analysis, Kim *et al.* compared the performance of ANN, a multivariate logistic regression (LR) and American Society of Anesthesiologists (ASA) class to predict cardiac complications, wound complications, thromboembolic events and mortality within 30 days of posterior lumbar fusions (13). On the basis of AUC values, the authors found that machine learning through LR or ANNs was more accurate at identifying risk factors for complications rather than using standard benchmark ASA scoring systems.

In a separate study, Kim *et al.* identified more than 4,000 patients undergoing surgery for adult spinal deformity (ASD) and used the data to train ANN, LR and ASA models for predicting complications and mortality within 30 days of the surgery (14). The authors found that the ANN and LR outperformed ASA scoring in predicting every complication. Furthermore, the ANN performed better than LR in predicting cardiac complications, wound complications and mortality. The authors concluded that as more granular clinical data would become available in these registries, training these machine learning algorithms will be useful for improving risk-stratification and subsequent clinical decision-making.

Opioid dependence/usage

Opioid prescription rates for the management of low back pain have doubled over the past decade. According to a recent report, direct and indirect costs associated with opioid usage for back pain have increased by 660% (15). Furthermore, given the increasing utilization of spinal fusions in elderly patients, health-policy makers are beginning to caution the long-term effects of opioid dependence and/or usage following spine surgery. Nearly 8% to 9% of opioid-naïve and 42% to 45% opioid-dependent individuals undergoing lumbar fusion continue to use opioids at 12 months (16). Pre-operative identification of patients who may be at a risk of long-term opioid dependence following spine

surgery will be essential in curbing the cost-burden on the healthcare system.

In two separate analyses, Karhade *et al.* employed machine-learning models to identify risk factors for sustained opioid use following surgery for lumbar disc herniation (17) or anterior cervical discectomy and fusions (ACDFs) (18). Sustained opioid usage was defined as the presence of continuous prescription opioid from day 0 of surgery up to at least day 90–180 post-operatively. The authors used a multi-institutional database of electronic health records to retrieve the study samples. For both ACDF and lumbar disc herniation cohorts, the authors employed 5 different machine learning models (RF, stochastic gradient boosting, neural network, support vector machine, and elastic-net penalized regression) for more than 2,500 patients and compared overall performance between each model using C-statistics. The authors found that 1/10th of patients following ACDF experience sustained opioid use. The overall c-statistics for all models ranged from 0.63 to 0.81, with stochastic gradient boosting algorithm having the best performance (0.81) and good calibration. The model demonstration that the duration of pre-operative opioid use, anti-depressant medication use, tobacco/smoking disorder and presence of Medicaid insurance were the most important predictors of sustained opioid prescription. For lumbar disc herniation, the authors employed the same machine learning models for more than 5,000 patients. Sustained post-operative opioid use was seen in 7.7% of patients. The elastic-net penalized logistic regression model had the best discrimination (c-statistic =0.81), good calibration and overall performance. The three most important predictors of sustained opioid use, based on the model, were concurrent instrumentation, duration of pre-operative opioid prescription and presence of a depressive disorder at the time of the surgery.

Predicting outcomes in spinal metastases and epidural abscesses

Despite significant advancements in the field of spine surgery, mortality and morbidity following operative or non-operative management of spinal epidural abscesses (SEAs) remain high. Providers have recently advocated for the construction of well-tuned prediction models that would allow the filtering/identification of high-risk vulnerable patients at the time of diagnosis to mitigate the risks of experiencing costly adverse events. Karhade *et al.* utilized a single-institution registry to create a comprehensive prediction model for in-hospital and 90-day mortality following SEAs (19). In addition to including demographics, the authors included granular clinical data such as presence of motor deficit, sensory changes, American Spinal Injury Association (ASIA) score and lab parameters. The authors employed a number of different machine-learning approaches across more than 1,000 patients presenting with SEA and found that older age, hypoalbuminemia, thrombocytopenia, elevated neutrophil:lymphocyte ratio, hemodialysis use, presence of active malignancy and diabetes increased the risk of 90-day mortality. In another study by Shah *et al.* (20), the authors employed a machine-learning approach to predict failure of non-operative management of SEA in a cohort of 367 patients. They found the presence of motor deficit, diabetes, location of ventral component of abscess relative to thecal sac, history of compression/pathologic vertebrae, presence of sensory deficit, active malignancy and involvement of 3 or more vertebral levels were associated with failure of non-operative management.

More than 40% of cancer patients present with spinal metastases. Though surgery for spinal metastases has increased over time, the short-term mortality following the surgery itself still remains a cause of concern. Karhade *et al.* carried out a machine-learning analysis to predict 30-day mortality following surgery for spinal metastases using the ACS-NSQIP database (21), through 4 different approaches: neural networks, support vector machine, Bayes point machine and decision tree modeling. The authors found that album, functional health status, white blood counts, hematocrit levels, alkaline phosphatase, location of abscess (cervical *vs.* thoracic *vs.* lumbosacral) and ASA class predicted 30-day mortality, with Bayes point machine modeling having the best calibration (21). The same authors conducted another multi-institutional study to predict 90-day and 1-year mortality on a cohort of 732 patients undergoing surgery for spinal metastatic disease (22). The authors found that albumin, primary tumor histology and Eastern Cooperative Oncology Group (ECOG) performance status were the three most important predictors of 90-day mortality. Furthermore, the authors found that a stochastic gradient boosting machine model had higher predictability as compared to the current existing scoring systems, such as SORG (Skeletal Oncology Research Group), New England Spine Metastasis and Tokuhashi.

While machine learning is still a relatively new topic in healthcare, evidence regarding the ability to predict outcomes using this technique is promising. As healthcare

systems begin to prioritize value and quality over volume, integrating machine learning models in electronic medical record (EMR) systems will be an effective way to help physicians make the best choices for their patients to ensure excellent outcomes.

Acknowledgments

None.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Papanicolas I, Woskie LR, Jha AK. Health Care Spending in the United States and Other High-Income Countries. JAMA 2018;319:1024-39.
2. Branning G, Vater M. Healthcare Spending: Plenty of Blame to Go Around. Am Health Drug Benefits 2016;9:445-7.
3. Reddy S, Fox J, Purohit MP. Artificial intelligence-enabled healthcare delivery. J R Soc Med 2019;112:22-8.
4. Seetharam K, Shrestha S, Sengupta PP. Artificial Intelligence in Cardiovascular Medicine. Curr Treat Options Cardiovasc Med 2019;21:25.
5. Johnson KW, Torres Soto J, Glicksberg BS, et al. Artificial Intelligence in Cardiology. J Am Coll Cardiol 2018;71:2668-79.
6. Yeramaneni S, Robinson C, Hostin R. Impact of spine surgery complications on costs associated with management of adult spinal deformity. Curr Rev Musculoskelet Med 2016;9:327-32.
7. Nasser R, Yadla S, Maltenfort MG, et al. Complications in spine surgery. J Neurosurg Spine 2010;13:144-57.
8. Damberg CL, Sorbero ME, Lovejoy SL, et al. Measuring Success in Health Care Value-Based Purchasing Programs: Findings from an Environmental Scan, Literature Review, and Expert Panel Discussions. Rand Health Q 2014;4:9.
9. Conrad DA. The Theory of Value-Based Payment Incentives and Their Application to Health Care. Health Serv Res 2015;50 Suppl 2:2057-89.
10. Dietz N, Sharma M, Alhourani A, et al. Bundled Payment Models in Spine Surgery: Current Challenges and Opportunities, a Systematic Review. World Neurosurg 2019;123:177-83.
11. Han SS, Azad TD, Suarez PA, et al. A machine learning approach for predictive models of adverse events following spine surgery. Spine J 2019. [Epub ahead of print].
12. Goyal A, Ngufor C, Kerezoudis P, et al. Can machine learning algorithms accurately predict discharge to nonhome facility and early unplanned readmissions following spinal fusion? Analysis of a national surgical registry. J Neurosurg Spine 2019. [Epub ahead of print].
13. Kim JS, Merrill RK, Arvind V, et al. Examining the Ability of Artificial Neural Networks Machine Learning Models to Accurately Predict Complications Following Posterior Lumbar Spine Fusion. Spine (Phila Pa 1976) 2018;43:853-60.
14. Kim JS, Arvind V, Oermann EK, et al. Predicting Surgical Complications in Patients Undergoing Elective Adult Spinal Deformity Procedures Using Machine Learning. Spine Deform 2018;6:762-70.
15. Martin BI, Turner JA, Mirza SK, et al. Trends in health care expenditures, utilization, and health status among US adults with spine problems, 1997-2006. Spine (Phila Pa 1976) 2009;34:2077-84.
16. Kalakoti P, Hendrickson NR, Bedard NA, et al. Opioid Utilization Following Lumbar Arthrodesis: Trends and Factors Associated With Long-term Use. Spine (Phila Pa 1976) 2018;43:1208-16.
17. Karhade AV, Ogink PT, Thio Q, et al. Development of Machine Learning Algorithms for Prediction of Prolonged Opioid Prescription after Surgery for Lumbar Disc Herniation. Spine J 2019. [Epub ahead of print].
18. Karhade AV, Ogink PT, Thio Q, et al. Machine learning for prediction of sustained opioid prescription after anterior cervical discectomy and fusion. Spine J 2019;19:976-83.
19. Karhade AV, Shah AA, Bono CM, et al. Development of Machine Learning Algorithms for Prediction of Mortality in Spinal Epidural Abscess. Spine J 2019. [Epub ahead of print].
20. Shah AA, Karhade AV, Bono CM, et al. Development of a machine learning algorithm for prediction of failure of nonoperative management in spinal epidural abscess. Spine J 2019. [Epub ahead of print].
21. Karhade AV, Thio Q, Ogink PT, et al. Development of Machine Learning Algorithms for Prediction of 30-

Day Mortality After Surgery for Spinal Metastasis. Neurosurgery 2019;85:E83-E91.
22. Karhade AV, Thio Q, Ogink PT, et al. Predicting 90-Day and 1-Year Mortality in Spinal Metastatic Disease: Development and Internal Validation. Neurosurgery 2019. [Epub ahead of print].

Cite this article as: Malik AT, Khan SN. Predictive modeling in spine surgery. Ann Transl Med 2019;7(Suppl 5):S173. doi: 10.21037/atm.2019.07.99

