

The Impact of Artificial Intelligence on Quality and Safety

Michelle S. Lee, BA¹, Matthew M. Grabowski, MD¹ ,
Ghaith Habboub, MD¹, and Thomas E. Mroz, MD¹

Global Spine Journal
2020, Vol. 10(15) 99S-103S
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2192568219878133
journals.sagepub.com/home/gsj



Abstract

As exponential expansion of computing capacity converges with unsustainable health care spending, a hopeful opportunity has emerged: the use of artificial intelligence to enhance health care quality and safety. These computer-based algorithms can perform the intricate and extremely complex mathematical operations of classification or regression on immense amounts of data to detect intricate and potentially previously unknown patterns in that data, with the end result of creating predictive models that can be utilized in clinical practice. Such models are designed to distinguish relevant from irrelevant data regarding a particular patient; choose appropriate perioperative care, intervention or surgery; predict cost of care and reimbursement; and predict future outcomes on a variety of anchored measures. If and when one is brought to fruition, an artificial intelligence platform could serve as the first legitimate clinical decision-making tool in spine care, delivering on the value equation while serving as a source for improving physician performance and promoting appropriate, efficient care in this era of financial uncertainty in health care.

Keywords

artificial intelligence, machine learning, spine predictive modeling, spine quality, spine safety

In the most recent version of the projections of the Office of the Actuary in the Centers for Medicare & Medicaid Services, national health spending growth is expected to average 5.5% per year for 2017-2026 in the United States (US): approximately 1.0% higher than projected gross domestic product (GDP) growth.¹ This results in an increase in the health-related costs as a percentage of GDP from 17.9% in 2016, to nearly 20% by 2026, reaching a total of \$5.7 trillion by 2026.¹ When looking granularly at specific medical fields, a significant portion of the rising costs of health care in the US relates to the diagnosis and treatment of spinal pathology. It is estimated that 12% to 30% of US adults have an active back problem with approximately 6% having made a visit to a physician for these conditions at one point in their lives, costing upward of \$100 billion to the system each year.^{2,3} Specifically with regard to spine surgery, fusions and laminectomies were the third and fifth most commonly performed surgical procedures in the United States in 2015, respectively.⁴ Given the rising costs associated with spine surgery and an aging population, it becomes increasingly clear that the current trajectory is not sustainable, and further scrutiny will be placed on the field in assessing the effectiveness, efficiency, and safety of care delivered. As more healthcare systems invest in healthcare analytics and “big data” (large, complex datasets such as those found in

electronic medical records), the opportunity arises to employ predictive analytics via machine learning (ML)/artificial intelligence (AI) approaches to improve quality, reduce waste and error, and minimize cost.^{5,6}

Recent developments in the technologies related to healthcare data collection and analytics have led to a rapid rise in the application of AI within health-related fields. One such application is ML, a branch of AI that involves the construction and application of statistical algorithms that continuously learn and make observations from existing data, and then create a predictive model based on that data.⁷ With advances in computer processing capability, data storage, and networking, these computer-based algorithms can perform the intricate and extremely complex mathematical operations of classification or regression (specifically nonlinear regression) on immense amounts of data to detect intricate and potentially previously

¹ Cleveland Clinic Foundation, Cleveland, OH, USA

Corresponding Author:

Thomas E. Mroz, Director, Center for Spine Health, Spine Research Lab, Departments of Orthopaedic and Neurological Surgery, Cleveland Clinic, 9500 Euclid Avenue, S-40, Cleveland, OH 44195, USA.
Email: mrozte@ccf.org



unknown patterns in that data.⁸ ML algorithms have been able to analyze complex and large volumes of electronic medical record data to produce predictions for a wide range of clinical problems.⁹ For example, Rajkomar et al⁹ demonstrated that ML models outperformed traditional, clinically used models in predicting mortality, unexpected readmission, and increased length of stay (LOS) in a study cohort of all admissions in 2 major hospitals from 2009 to 2016. Various investigators have been developing image analysis methods using ML algorithms that have shown promising results in fields such as dermatology, radiology, and ophthalmology. For example, Esteva et al¹⁰ have trained ML algorithms to classify skin cancer with a level of competence comparable to dermatologists. These early examples provide insight into early contemporary use of AI in medicine and provide a view of technology that may transform the medical field over the decades to come.

While not the first medical field to adopt an “AI approach” to problem solving, the spine surgery field has recently seen an outpouring of publications related to research in this area. An initial topic of focus by researchers was related to the cost of spine care, as there has been heightened emphasis on moving to a value-based (quality/cost) health care market. For example, as Medicare payments are standardized by procedures performed regardless of hospital LOS, ML systems have been designed with the ability to accurately predict spine surgery-related LOS, discharge to nonhome facility, and early unplanned readmissions using only presurgical or pre-discharge variables.¹¹⁻¹⁴ These models can help identify/target certain high-risk patients and the variables that contribute to that risk status, allowing hospitals to allocate specific clinical and social resources to reduce costly LOS and readmissions. This can help to maximize efficiency of care delivered, while also keeping constant or even increasing the quality of care delivered.

As spinal surgery has evolved with an explosion of new techniques and technologies in recent decades, there still remains a lack of quality, high-level evidence to support much of the spine care rendered in the US, especially with the cost associated with many of the treatments and devices. As there are numerous surgical treatments in spine surgery that do not easily lend themselves to traditional randomized controlled trials (due to either cost or ethical considerations, among other reasons), an opportunity arises that is ripe for solutions derived from ML approaches. Multiple clinical registries are being collected that contain large quantities of high-quality, spine health care data, such as the 1000-patient Spinal Laminectomy versus Instrumented Pedicle Screw (SLIP) II study.¹⁵⁻¹⁷ These registries contain demographics, surgery-related variables, patient-reported and complication outcome measures, and notably, they even contain digital imaging with metadata. Leveraging of these vast data repositories can help develop predictive algorithms that are able to incorporate the full range of variables (including complex imaging) in order to guide treatment recommendations.

Because of this lack of high-level evidence, there remains much heterogeneity in the current surgical treatment of spinal disorders, with significant clinical and economic

implications.¹⁸⁻²⁰ For instance, national surveys of US spine surgeons conducted by Mroz et al²¹ found 69% disagreement for recurrent lumbar disk herniation, while another study demonstrated 75% disagreement among surgeons on the approach to treat patients with lower back pain,²² implying that 2 similar patients with the same pathology could receive entirely different care. Furthermore, a cost analysis based on the results of the national survey mentioned above revealed that there is also a variation in costs based on spine surgeon specialty, practice type, surgical volume and geographical location.²³ Recent ML/AI approaches to this problem have been published that attempt to assist surgeons' decisions with predictions of patient outcomes. Utilizing data from repositories created from AOSpine prospective, multicenter studies, Merali et al¹⁷ developed a supervised ML model that accurately predicts a positive outcome on an individual patient after surgery for degenerative cervical myelopathy, with an average area under the curve of 0.70, classification accuracy of 77%, and sensitivity of 78% on an independent testing cohort. Shah et al²⁴ were able to build an ML model that predicts probability of failure of nonoperative management in spinal epidural abscess, while Karhade et al²⁵ successfully developed an ML algorithm that predicts in-hospital and 90-post discharge mortality in this patient group. The same group was able to predict short-term postoperative mortality in individual patients with spinal metastatic disease with an ML model, aiding in decision-making and informed discussions with the patient regarding surgical intervention this challenging patient population.²⁶ All of these previously mentioned studies have now published their prognostic tools in an open-access, digital interface to be integrated into practice, supporting clinicians in developing treatment plans that are more standardized across the world.

Along with prediction of positive patient outcomes, clinician researchers have also used AI/ML to forecast negative outcomes as well, as recent publications have explored the likelihood of complications from spine surgery. In multiple articles, the same group led by Cho et al utilized an artificial neural network-based ML algorithm to predict surgical complications in patients undergoing elective anterior cervical discectomy and fusion, posterior lumbar fusion, and adult spinal deformity surgeries. Their models were able to specifically predict the risk of cardiac-related, wound-related, venous thromboembolism-related, and mortality in these patients, outperforming the American Society of Anesthesiologists Physical Status Classification scoring in predicting individual risk prognosis.^{27,28} Another publication by Sheer et al²⁹ describes their method to create a ML model that successfully predicts major intraoperative/perioperative complications following adult spinal deformity surgery with an accuracy of 87%. Utilizing large databases of patient information, Han et al³⁰ were able to analyze over 1 000 000 patients that had previously undergone spine surgery and developed multiple ML predictive models that identify risk factors for postoperative complications. Karhade et al²⁴ were even able to predict prolonged opioid prescription after surgery for lumbar disc herniation in an ML algorithm. These surgery- and patient-specific models can help

to aid in surgical planning, as well as patient counseling and shared decision making. If these models identify modifiable risk factors in the preoperative setting of a nonurgent surgery, time and effort could be dedicated to improved medical management of that comorbidity prior to surgery, in effect reducing the risk of complications and increasing the probability of a good outcome.

In deciding if a patient is indicated for surgery, one area where a surgeon's subjectivity may still reign supreme is review of the spine imaging. Utilizing classification techniques from radiology literature, new research is revealing the applicability of AI and ML algorithms to the analysis of spine imaging. One technique involves the use of ML models utilizing natural language processing to distinguish specific words and phrases from unstructured radiology reports in order to classify patients by imaging findings, as Tan et al³¹ were able to do in a cohort of patients with low back pain. However, in more recent publications, other groups were able to utilize the imaging itself to detect and classify a variety of pathologies. Hopkins et al³² were able to predict both the diagnosis of cervical spondylotic myelopathy and its severity with high sensitivity and specificity (90.25% and 85.05%, respectively), utilizing magnetic resonance imaging alone in an artificial neural network model. Further, work has been done to develop ML models in the detection and grading of lumbar spinal stenosis³³ and fracture detection and classification with various types of imaging modalities.³⁴ AI/ML imaging analysis can even assist real time in the outpatient clinic, where Sharif Bidabadi et al³⁵ were able to accurately identify foot drop of an L5 origin and classify patients into various recover stages with an 85% accuracy. While there is much work to be done, this initial work which was all published in the past year, shows the feasibility of using AI/ML-based approaches to analyzing spine imaging.

Common themes among large institutions and large spine centers are tighter financial margins, less resources, and heightened payer scrutiny on indications, outcomes, and postprocedural treatments. This collectively creates real strain on the departmental workforce (ie, secretaries, advanced practice providers, physicians). An AI platform that successfully predicts patient and surgeon performance from financial, outcome, and electronic medical record databases across an entire book of business stands to provide the leverage to homogenize outcome and cost. This, in turn, positions said organization optimally for contract negotiations and population health initiatives. Furthermore, a fully integrated AI platform can also automate much of what currently strains department assets. Postsurgical checks, ordering medications and imaging, patient reminders, and scheduling follow-up visits, are all some examples of how such a platform can enhance overall spine center efficiencies and performance, patient satisfaction (eg, more automated touch points), and employee engagement.

Challenges Ahead

Even though current research described above highlights the promise and potential of AI in spine surgery, the field as a

whole still face many challenges. First, in order to create an AI-driven decision platform, very large and appropriately labeled data sets are required, which the majority of centers in the United States still lack. This becomes even more difficult with imaging-based analysis. Second, some ML models require manual labeling of the data for classification and learning to occur. This presents a clear challenge in the analysis of spine surgery pathologies, where there is still widespread disagreement about what constitutes normal versus abnormal with regard to certain exam/imaging findings, and subsequently the appropriate treatment(s). This can be circumvented by allowing the model itself to do the analysis and classification, such as is the case with unsupervised algorithms. Given the vast quantity of data analyzed, this can reveal links between variables that experts would not have otherwise expected. However, it is difficult to backtrack and get precise information regarding the specifics of the data sorting in these types of models. And with poor quality or quantity of data to learn from, the model may make erroneous associations and/or can be "overfitted" to the training dataset, producing a lack of external validity. Furthermore, many ML algorithms thus far are typically trained and validated internally within one institution. Further work needs to be carried out to examine if a predictive model is transferable from one site to another, and what implications this holds as a "live" ML model undergoes continuous calibration and evolution based on new sets of data.

As exponential expansion of computing capacity converges with unsustainable healthcare spending, a hopeful opportunity has emerged: the use of AI to enhance healthcare quality and safety. AI-based, ML approaches to spinal pathologies are already distinguishing relevant from irrelevant data regarding a particular patient, assisting with appropriate hospital-based care, interventions or even surgeries, predicting cost of care, and predicting future outcomes on a variety of anchored measures. While many shortcomings still exist as the technology is in early development, extrapolating from today's progress and fully implemented into the healthcare system, AI could help solve a number of problems in spine surgery by improving outcomes, minimizing cost, standardizing care for a given pathology, and driving efficiencies within a spine service line in large centers. These types of approaches could deliver on the value equation while serving as a resource for improving physician performance and promoting appropriate, efficient care in this era of financial uncertainty in health care.


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This supplement was supported by funding from AO Spine North America.

ORCID iD

Matthew M. Grabowski  <https://orcid.org/0000-0001-8550-0124>

References

1. Cuckler GA, Sisko AM, Poisal JA, et al. National Health Expenditure Projections, 2017-26: despite uncertainty, fundamentals primarily drive spending growth. *Health Aff (Millwood)*. 2018; 37:482-492. doi:10.1377/hlthaff.2017.1655
2. McGirt MJ, Resnick D, Edwards N, Angevine P, Mroz T, Fehlings M. Background to understanding value-based surgical spine care. *Spine (Phila Pa 1976)*. 2014;39(22 suppl 1):S51-S52. doi:10.1097/BRS.0000000000000544
3. Davis MA, Onega T, Weeks WB, Lurie JD. Where the United States spends its spine dollars: expenditures on different ambulatory services for the management of back and neck conditions. *Spine (Phila Pa 1976)*. 2012;37:1693-1701. doi:10.1097/BRS.0b013e3182541f45
4. Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project (HCUP). <https://www.ahrq.gov/data/hcup/index.html>. Published April 2, 2015. Accessed February 27, 2019.
5. Wang Y, Kung L, Wang WYC, Cegielski CG. An integrated big data analytics-enabled transformation model: application to health care. *Inf Manag*. 2018;55:64-79. doi:10.1016/j.im.2017.04.001
6. Ristevski B, Chen M. Big data analytics in medicine and healthcare. *J Integr Bioinform*. 2018;15:20170030. doi:10.1515/jib-2017-0030
7. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA: MIT Press; 2017.
8. Chollet F. *Deep Learning With Python*. Shelter Island, NY: Manning; 2018.
9. Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning for electronic health records. *NPJ Digit Med*. 2018;1:18. doi:10.1038/s41746-018-0029-1
10. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017; 542:115-118. doi:10.1038/nature21056
11. Ogink PT, Karhade AV, Thio QCBS, et al. Predicting discharge placement after elective surgery for lumbar spinal stenosis using machine learning methods. *Eur Spine J*. 2019;28:1433-1440. doi: 10.1007/s00586-019-05928-z
12. Goyal A, Ngufor C, Kerezoudis P, Mccutcheon B, Storlie C, Bydon M. Can machine learning algorithms accurately predict discharge to nonhome facility and early unplanned readmissions following spinal fusion? Analysis of a national surgical registry [published online June 7, 2019]. *J Neurosurg Spine*. doi:10.3171/2019.3.spine181367
13. Stopa BM, Robertson FC, Karhade AV, et al. Predicting nonroutine discharge after elective spine surgery: external validation of machine learning algorithms [published online July 26, 2019]. *J Neurosurg Spine*. doi:10.3171/2019.5.spine1987
14. Kalagara S, Eltorai AEM, Durand WM, Depasse JM, Daniels AH. Machine learning modeling for predicting hospital readmission following lumbar laminectomy. *J Neurosurg Spine*. 2019;30: 344-352. doi:10.3171/2018.8.spine1869
15. Ghogawala Z, Dunbar MR, Essa I. Lumbar spondylolisthesis: modern registries and the development of artificial intelligence. *J Neurosurg Spine*. 2019;30:729-735. doi:10.3171/2019.2.spine18751
16. Ghogawala Z, Dunbar M, Essa I. Artificial intelligence for the treatment of lumbar spondylolisthesis. *Neurosurg Clin N Am*. 2019;30:383-389.
17. Merali ZG, Witiw CD, Badhiwala JH, Wilson JR, Fehlings MG. Using a machine learning approach to predict outcome after surgery for degenerative cervical myelopathy. *PLoS One*. 2019;14: e0215133. doi:10.1371/journal.pone.0215133
18. Irwin ZN, Hilibrand A, Gustavel M, et al. Variation in surgical decision making for degenerative spinal disorders. Part I: lumbar spine. *Spine (Phila Pa 1976)*. 2005;30:2208-2213. https://journals.lww.com/spinejournal/Fulltext/2005/10010/Variation_in_Surgical_Decision_Making_for.14.aspx.
19. Arnold PM, Brodke DS, Rampersaud YR, et al; Spine Trauma Study Group. Differences between neurosurgeons and orthopedic surgeons in classifying cervical dislocation injuries and making assessment and treatment decisions: a multicenter reliability study. *Am J Orthop (Belle Mead NJ)*. 2009;38:E156-E161. <http://europepmc.org/abstract/MED/20011745>.
20. Rampersaud YR, Fisher C, Wilsey J, et al. Agreement between orthopedic surgeons and neurosurgeons regarding a new algorithm for the treatment of thoracolumbar injuries: a multicenter reliability study. *J Spinal Disord Tech*. 2006;19:477-482. doi:10.1097/01.bsd.0000211237.14211.21
21. Mroz TE, Lubelski D, Williams SK, et al. Differences in the surgical treatment of recurrent lumbar disc herniation among spine surgeons in the United States. *Spine J*. 2014;14: 2334-2343. doi:10.1016/j.spinee.2014.01.037
22. Lubelski D, Williams SK, O'Rourke C, et al. Differences in the surgical treatment of lower back pain among spine surgeons in the United States. *Spine (Phila Pa 1976)*. 2016;41:978-986. doi:10.1097/BRS.0000000000001396
23. Alvin MD, Lubelski D, Alam R, et al. Spine surgeon treatment variability: the impact on costs. *Global Spine J*. 2018;8:498-506. doi:10.1177/2192568217739610
24. Shah AA, Karhade AV, Bono CM, Harris MB, Nelson SB, Schwab JH. Development of a machine learning algorithm for prediction of failure of nonoperative management in spinal epidural abscess [published online May 4, 2019]. *Spine J*. doi:10.1016/j.spinee.2019.04.022
25. Karhade AV, Shah AA, Bono CM, et al. Development of machine learning algorithms for prediction of mortality in spinal epidural abscess [published online June 27, 2019]. *Spine J*. doi:10.1016/j.spinee.2019.06.024
26. Karhade AV, Thio QCBS, 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.
27. 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: 762-770. doi:10.1016/j.jspd.2018.03.003
28. Arvind V, Kim JS, Oermann EK, Kaji D, Cho SK. Predicting surgical complications in adult patients undergoing anterior

- cervical discectomy and fusion using machine learning. *Neurospine*. 2018;15:329-337. doi:10.14245/ns.1836248.124
29. Scheer JK, Smith JS, Schwab F, et al; International Spine Study Group. Development of a preoperative predictive model for major complications following adult spinal deformity surgery. *J Neurosurg Spine*. 2017;26:736-743. doi:10.3171/2016.10.spine16197
30. Han SS, Azad TD, Suarez PA, Ratliff JK. A machine learning approach for predictive models of adverse events following spine surgery [published online June 20, 2019]. *Spine J*. doi:10.1016/j.spinee.2019.06.018
31. Tan WK, Hassanpour S, Heagerty PJ, et al. Comparison of natural language processing rules-based and machine-learning systems to identify lumbar spine imaging findings related to low back pain. *Acad Radiol*. 2018;25:1422-1432. doi:10.1016/j.acra.2018.03.008
32. Hopkins BS, Weber KA 2nd, Kesavabhotla K, Paliwal M, Cantrell DR, Smith ZA. Machine learning for the prediction of cervical spondylotic myelopathy: a post hoc pilot study of 28 participants. *World Neurosurg*. 2019;127:e436-e442. doi:10.1016/j.wneu.2019.03.165
33. Huber FA, Stutz S, de Martini IV, et al. Qualitative versus quantitative lumbar spinal stenosis grading by machine learning supported texture analysis—experience from the LSOS study cohort. *Eur J Radiol*. 2019;114:45-50. doi:10.1016/j.ejrad.2019.02.023
34. Langerhuizen DWG, Janssen SJ, Mallee WH, et al. What are the applications and limitations of artificial intelligence for fracture detection and classification in orthopaedic trauma imaging? A systematic review [published online June 21, 2019]. *Clin Orthop Relat Res*. doi:10.1097/corr.0000000000000848
35. Sharif Bidabadi S, Tan T, Murray I, Lee G. Tracking foot drop recovery following lumbar-spine surgery, applying multiclass gait classification using machine learning techniques. *Sensors*. 2019;19:E2542. doi:10.3390/s19112542