Machine Learning and Artificial Intelligence: Definitions, Applications, and Future Directions

J. Matthew Helm[1] · Andrew M. Swiergosz[1] · Heather S. Haeberle[2] · Jaret M. Karnuta[1] · Jonathan L. Schaffer[1] · Viktor E. Krebs[1] · Andrew I. Spitzer[3] · Prem N. Ramkumar[1]

Abstract

Purpose of Review With the unprecedented advancement of data aggregation and deep learning algorithms, artificial intelligence (AI) and machine learning (ML) are poised to transform the practice of medicine. The field of orthopedics, in particular, is uniquely suited to harness the power of big data, and in doing so provide critical insight into elevating the many facets of care provided by orthopedic surgeons. The purpose of this review is to critically evaluate the recent and novel literature regarding ML in the field of orthopedics and to address its potential impact on the future of musculoskeletal care.
Recent Findings Recent literature demonstrates that the incorporation of ML into orthopedics has the potential to elevate patient care through alternative patient-specific payment models, rapidly analyze imaging modalities, and remotely monitor patients.
Summary Just as the business of medicine was once considered outside the domain of the orthopedic surgeon, we report evidence that demonstrates these emerging applications of AI warrant ownership, leverage, and application by the orthopedic surgeon to better serve their patients and deliver optimal, value-based care.

Keywords Artificial intelligence · Machine learning · Patient-specific payment models · Remote patient monitoring systems · Value-based care · Big data

Introduction

Originally coined in the 1950s, the term "artificial intelligence" initially began as the simple theory of human intelligence being exhibited by machines [1•]. In 1976, Jerrold S. Maxmen foretold that artificial intelligence (AI) would bring about the "post-physician era" in the twenty-first century [2, 3]. In today's era of rapid technological advancement and

exponential increases in extremely large data sets ("big data"), AI has transitioned from mere theory to tangible application on an unprecedented scale [4]. From evaluating extraordinarily large data sets in near real-time, autonomous driving cars and stream history-influenced video viewing recommendations (Netflix, Los Gatos, CA, USA), to online purchase recommendations, advertisements, and fraud detection (Amazon, Seattle, WA, USA), AI has become fundamentally ingrained within many facets of society and often functions invisibly in the background of our personal electronic devices.

Considered a subset of AI, machine learning (ML) exhibits the experiential "learning" associated with human intelligence, while also having the capacity to learn and improve its analyses through the use of computational algorithms [1•, 3]. These algorithms use large sets of data inputs and outputs to recognize patterns and effectively "learn" in order to train the machine to make autonomous recommendations or decisions. After sufficient repetitions and modification of the algorithm, the machine becomes able to take an input and to predict an output [1•, 3]. Outputs are then compared with a set

of known outcomes in order to judge the accuracy of the algorithm, which is then iteratively adjusted to perfect the ability to predict further outcomes [5••].

"Deep neural networks" are more complex versions of these models that make use of hierarchical tiers to segregate and manage the final output. The network begins with an input tier that then progresses to a number of "hidden tiers" that each respond to different features of the input (Fig. 1) [4, 6••]. These tiers allow for an increase in understanding as the input ascends "deeper," permitting the development of models without explicitly programmed directions [7••]. As the machine studies a specific concept on multiple tiers, the existing algorithm is then able to successively refine itself as new data is available. Similar to the way the human brain functions, the machine is able to make "neuronal" connections from "dendritic" connections on multiple hierarchical data levels. These networks have thus given rise to a new form of AI, known as "deep learning" [5••].

The predictive abilities of ML are being increasingly applied in the field of healthcare. As a convergence between health and data sciences, ML models have been proposed and tested as potential solutions to a variety of issues involving diagnostic errors, treatment mistakes, workflow inefficiencies, and impediments to value-based care [3, 4]. One

prominently theorized application of automated machine learning involves the automation of "clicks" in the electronic health record (EHR) to combat the "world of shallow medicine" we currently live in with "insufficient time, insufficient context, and insufficient presence," as Dr. Eric Topol has described [4]. Specifically, a single patient in the EHR is associated with approximately 32,000 discrete data elements for analysis, necessitating an AI-based algorithm to process patient data [8]. Presently, efforts to automate actions in the electronic medical record are limited due to regulated proprietary access issues. However, when the eventual automation capabilities are realized, AI offers physicians the gift of time to finally execute the tasks for which they were trained, and "to make healthcare human again" [4].

In 2018, the Cleveland Clinic's Department of Orthopaedic Surgery established the Machine Learning Arthroplasty Laboratory (MLAL) in an effort to further the study of these seminal issues and assess the ability to apply machine learning to musculoskeletal medicine, specifically as it relates to patient-specific, value-based care and human mobility [9••]. These research efforts are laying the groundwork for the future design and application of increasingly accurate ML algorithms that may directly improve patient outcomes and the practice of orthopedics in general [6••, 9••, 10••, 11••, 12••,

Fig. 1 Example of input, hidden, and output layers used to predict value-based metrics prior to elective primary total hip or knee arthroplasty from Ramkumar et al. [6••]. Dr. Ramkumar retains the rights to this figure
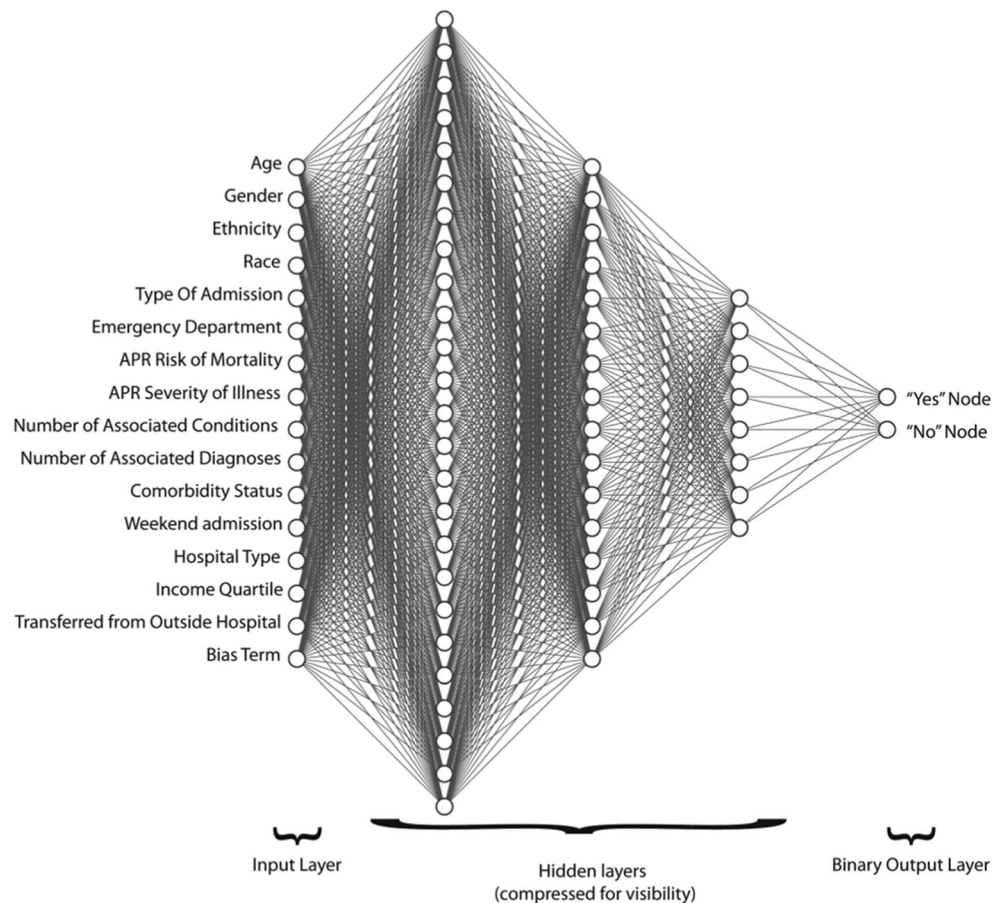
13•]. This review serves to stand on the shoulders of those coming before it and critically evaluate the literature published over the last 5 years regarding ML in the field of orthopedic surgery, as well as to address its impact on the future of medicine.

Machine Learning in Orthopedics

Although ML is relatively new to the field of orthopedic surgery, it is vital for tomorrow's orthopedic surgeons intending to practice in the future to be conscious of and fully understand ML and the paradigm shift it represents within medicine in the delivery of optimal care [7••]. In 2018, Cabitza et al. conducted one of the first literature reviews of ML in orthopedics, demonstrating a 10-fold increase in ML publications over the previous two decades [14••]. Many of these studies have very practical applications to orthopedics. For example, a study by Thong et al. created a collection of algorithms known as an "artificial neural network" (ANN) to optimize encoding-decoding of 3D spine model vectors for the automatic detection of adolescent idiopathic scoliosis [15]. ANNs are modeled loosely after the human brain and designed to recognize patterns in order to generate an output. Using a specific type of ANN known as a stacked auto-encoder (SAE), the authors demonstrated the successful ability to cluster patients by deformity and simplify the complex nature of 3D spine models. With separate encoding and decoding stages, the SAE was able to learn the parameters needed to map an input vector from a latent representation, and to use that input vector to regenerate an output vector consisting of a 3D spine model [15]. Another study by Olczak et al. made use of an ANN trained to identify fractures from plain radiographic images [16]. The network was benchmarked against a gold standard for fractures that was created by reviewing each image in full resolution with its multiple views in addition to the radiologist report. All of their networks exhibited an accuracy of at least 90% in identifying laterality, body part, and exam view when compared with the "gold standard" radiology report, as well as the interpretation of the image by two additional senior orthopedic surgeons. Possible implications of this kind of image-based analysis include improved ability to detect pathology, and the ability to identify orthopedic implants for the purpose of facilitating hardware removal and/or modular revision [17]. With the use of ML, a specific manufacturer's implant may be identified via a simple radiographic film, eliminating the costly impact of delays in care or misidentifications leading to lack of appropriate equipment during the operation, and potential complications imposed by work-around options [17].

In a different application, Shah et al. used a validated ML segmentation model to automate the measurement and segmentation of articular cartilage thickness (ACT) in healthy knees on MR images [18]. The algorithm analyzed 3910 MRIs and accurately identified which pixels represented which tissue type. This demonstrated how ML could potentially be used as a tool for automated tracking of the impact of medical intervention on the progression of cartilage degeneration [18]. Using dual-energy X-ray absorptiometry (DEXA), Kruse et al. built 24 statistical models to apply ML principles to the prediction of hip fractures over time in 4722 women and 717 men with 5 years of follow-up [19•]. The best "female model" performed with a test area under the curve (AUC) of 0.92, while the best "male model" performed with a test AUC of 0.89. Identifying predictive factors such as bone mineral density, glucose measurements, and osteoarthritis diagnosis, the model demonstrated that ML can improve hip fracture prediction beyond logistic regression [19•].

While many of these original ML applications in orthopedics focused primarily on imaging-based pathology detection and prediction, more current and advanced applications of these predictive models have become increasingly centered on outcome-based measures and patient-specific, value-based care models [6••, 7••, 10••, 11••, 12••]. Recently, Haeberle et al. conducted a review of these more novel applications, specifically exploring imaging-based analyses, value-based payment models, and mobile health technologies in the setting of lower extremity arthroplasty [5••]. Demonstrating the novelty and effectiveness of many of these models, this review concluded that the incorporation of ML is critical for orthopedic surgeons in order to elevate their level of care and advocate for their patients through alternative payment models and preoperative expectation management [5••]. Specifically, with the advent of bundled care programs by the Center for Medicare and Medicaid Services, ML has been used to assess the financial feasibility of these models. After analyzing 98,562 patients from the New York Statewide Planning and Research Cooperative System database, Karnuta et al. found bundling hip fractures represented an unsustainable business model as the cost of delivering care was based on non-modifiable patient-specific factors [12••].

One of the first studies to apply these ML algorithms in orthopedics was that of Kuo et al. [20•]. Using a model that produces multiple decision trees from randomly selected features known as a "random forest model," they demonstrated the successful prediction of medical costs in 532 spinal fusion cases over the span of 3 years. With an accuracy of 84.30%, a sensitivity of 71.4%, a specificity of 92.2%, and an AUC of 0.904, the model showed promise in its ability to inform hospital strategy regarding financial management and decision-making. Navarro et al. took this a step further by using a similar model to explore the impact and predictive ability of patient-specific factors such as race, gender, age, and comorbidities on postoperative length of stay (LOS) and cost for patients undergoing total knee arthroplasty (TKA) [11••]. Using 141,446 patients, they demonstrated that as patient complexity increased, cost incrementally increased in tiers of 3%, 10%, and 15% for moderate, major, and extreme

mortality risks, respectively [11••]. Ramkumar et al. applied the same model to patients undergoing primary total hip arthroplasty (THA), demonstrating continued successful predictive ability with the same postoperative factors despite a different joint and a different operation [10••].

By creating these more complex "deep learning" algorithms that are capable of unmonitored learning through multiple layers of connected "neurons," Karnuta et al. built upon this work by comparing the performance of two different ANN designs (MLP and DenseNet) along with multiple techniques for modifying the two algorithms known as "regularization" techniques, in predicting outcomes for primary THA and TKA in 295,605 patients [4, 7••]. Demonstrating the DenseNet neural network model's ability to achieve a higher AUC above 0.80, the study established important relationships to consider when choosing a specific model to predict surgical cost. The study was also one of the most comprehensive to date, validating ML as a valuable predictive tool beyond that of disease detection.

These ANNs were created and specifically applied to THA and TKA to create dynamic models that proposed patient-specific payment models (PSPMs) that risk-stratified patients based on their preoperative complexity to support reimbursement commensurate with complexity. While quantifying global patient risk was previously based on individual experience and therefore non-quantifiable, Ramkumar et al. established risk increases of 2.5%, 8.9%, and 17.3% for moderate, major, and severe comorbidities in the primary THA population and risk increases of 2.0%, 21.8%, and 82.6% in the primary TKA population [6••, 21••]. Armed with this data, a PSPM may be used to communicate case complexity to the payers, and reverse the narrative that THA or TKA are presently "misvalued" commodities [22]. As an example of the application of such an algorithm, THA patients with extreme comorbidities and severity of illness incur $10,200 in additional costs (17.3% greater risk) over those with minor risk in the Medicare population. This analysis provides value-based insights and evidence-based support to inform and hopefully influence policy advocates and decision-makers to provide more equitable reimbursement via a patient-specific tiered bundled payment model. Moreover, such a patient-specific payment model (PSPM) provides surgeons the long-sought ability to prospectively characterize and quantify non-modifiable patient-level factors.

Total hip arthroplasty (THA) has been called "the operation of the century," because it can reliably restore quality of life, physical activity and function, and mobility, while improving overall health status, in part, by reducing obesity and the impact of diabetes mellitus [23]. However, the present bundled payment model that offers a singular fee for an entire episode of care remains agnostic to patient variation, and indirectly encourages exclusion of those patients with risk factors for potential complications and increased costs of care. It incentivizes one of the healthiest patients, who in fact may demonstrate less overall benefit from THA. In contrast, a PSPM informed by the ANN described above represents a dynamic framework which can be utilized by patient, surgeon, care coordinator, and administrator alike, anywhere in the world, to analyze relevant patient-specific data and obtain meaningful, evidence-based output assessing patients' risk and profile for complications and increased costs. In turn, reimbursement could be adjusted to more realistically account for the most costly care provided to medically or surgically complex patients and create an equitable, cost-sharing relationship between payer, patient, hospital, and surgeon.

As another example, Karhade et al. used ML to develop a preoperative algorithm for the prediction of postoperative opioid use following total hip arthroplasty [24•]. Five algorithms analyzed 5507 patients, 345 of which had prolonged postoperative opioid use, and determined the predictive factors for prolonged postoperative opioid prescriptions. The best model achieved fair discrimination (AUC of 0.77) and exhibited higher net benefit than the default strategies of changing management for all patients or no patients [24•]. Once externally validated, this algorithm will allow clinicians to preoperatively identify specific THA patients at highest risk for developing postoperative opioid use dependence and alter their perioperative management, and hopefully outcome, accordingly.

In addition to these value-based payment and perioperative predictive models, ML is beginning to be used in clinical workflow to evolve our remote patient monitoring systems. Notably, a mobile and wearable knee technology has been developed (FocusMotion, Santa Monica, CA, USA) to create ML-driven recovery solutions that track and assess pre- and postoperative progress remotely [17]. With real-time, patient-specific feedback sent from a "smart" orthotic to a smartphone application, the surgeon is able to use movement data to monitor the patient's mobility and activity compliance. The patient is then provided with an individually optimized interactive recovery platform, utilizing notifications and real-time feedback to encourage compliance, guiding them throughout each step of their personalized recovery process. Ramkumar et al. validated the feasibility of the remote patient monitoring (RPM) system by successfully capturing uninterrupted, passive data for mobility, knee range of motion, patient-reported outcome measures, opioid use, and home exercise program compliance in 25 patients undergoing primary TKA for osteoarthritis [25••]. By establishing the ability to remotely acquire continuous data, the study introduces a scalable RPM system in orthopedics and challenges the current paradigm of capturing patient data only through standardized, geographically dependent, inefficient, and costly processes [25••].

Recently, Bini et al. conducted a pilot study designed to demonstrate the accuracy and feasibility of coupling ML with these wearable technologies to predict outcomes of total joint arthroplasty (TJA) [13•]. Wrist-based Fit Bit (FitBit, San

Francisco, CA) activity trackers and Mio Activity Tracker (Mio Technology, Taipei, Taiwan) and the Lumo Run waist-based posture sensory (Lumo Bodytech, Mountain View, CA) were used to collect 3 million data points from 15 patients undergoing total joint arthroplasty (TJA). ML algorithms were then able to successfully cluster sensor data from as early as 11 days postoperatively into groups that correlated to 6-week patient-reported outcomes measures (PROMs) (Hip and Knee Disability and Osteoarthritis Outcome Score, Knee Osteoarthritis Outcome Score, and Veterans RAND 12-Item Health Survey Physical Component Score). Bloomfield et al. also made use of a ML algorithm to successfully separate 68 postoperative unilateral TKA patients into clinically relevant groups based on wearable sensor-instrumented performance tests [26]. Using the instrumented timed-up-and-go test, a custom wearable system extracted 55 metrics for analysis that were separated by ML into functionally distinguished groups with significantly different test completion times. By identifying these different cohorts, functional outcomes and PROMs could potentially be predicted, and those patients destined for negative result could be identified early. Armed with this information, the surgeon could potentially intervene in a timely fashion in order to change the negative trajectory and improve the overall outcomes [26]. In addition, this technology enables the surgeon to more carefully monitor patients postoperatively, potentially reducing the number of clinical visits and post-operative compilations, and also providing the means to harness the full potential of virtual medicine and remote patient monitoring [5••, 25••].

As the availability of information grows and modern data science improves, predictive models relying on AI and ML will continue to be incorporated into the orthopedic care of our patients. With its large patient volumes, reliance on radiologic and imaging studies for diagnosis, and robust outcomes data-driven nature, orthopedics is poised to embrace and propagate transformative technological advancements in an increasingly value-based era of medicine.

Conclusions

It is critical that we as musculoskeletal specialists understand, embrace, and apply AI and ML to maximize the quality and cost-effectiveness of our care. We must not consider this explosive area of research and knowledge outside of our scope. However, it is mandatory that we exercise caution when incorporating new and potentially revolutionary technology into our healthcare system. We must understand the limitations and pitfalls of implementing AI and ML, especially the impact of skewed databases in creating potentially erroneous conclusions and the inability to dissect decision-making from these "black box" algorithms. Nevertheless, as data is being aggregated at an unprecedented rate, AI and ML will provide

critical insight into so many facets of the care we as orthopedic surgeons provide. Thus far, the Machine Learning Arthroplasty Laboratory at Cleveland Clinic and other similar endeavors have generated preoperative tools to predict opioid use and cost of care, insights into imaging analysis, and advanced remote patient monitoring systems.

By harnessing the power of big data, AI and ML are poised to transform orthopedics in particular and medicine in general. It can enable automation of redundant tasks and save time, assist in evidence-based decisions by using algorithms to identify risk factors for complications or other outcomes of concern, and is uniquely capable of predicting outcomes based on patient-specific algorithms. This empowers the clinician to intervene and intercede in order to change the course of treatment and favorably impact and transform an undesired outcome into a more acceptable or even positive overall clinical and surgical result.

While Maxmen predicted the emergence of a "post-physician era," Topol more pragmatically predicts AI may offer the profound opportunity to usher in a "physician-aided" era whereby human and artificial intelligence converge to offer physicians more time to paradoxically make healthcare human again [2, 4].

Compliance with Ethical Standards

References

Papers of particular interest, published recently, have been highlighted as:
• Of importance
•• Of major importance

1.• Bini SA. Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care? J Arthroplast. 2018;33(8):2358–61. https://doi.org/10.1016/j.arth.2018.02.067 **The purpose of this article was to discuss the fundamentals of artificial intelligence so that practicing surgeons may better understand**

how and where to apply it. The article touches on the origins of artificial intelligence and its progression into machine learning, as well as the development of artificial neural networks and deep learning algorithms. It provides a foundation for knowledge of machine learning and its use in healthcare.

2. Maxmen JS. The post-physician era : medicine in the twenty-first century. Hoboken: Wiley; 1976.

3. Naylor CD. On the prospects for a (deep) learning health care system. JAMA. 2018;320(11):1099–100. https://doi.org/10.1001/jama.2018.11103.

4. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. https://doi.org/10.1038/s41591-018-0300-7.

5.•• Haeberle HS, Helm JM, Navarro SM, et al. Artificial intelligence and machine learning in lower extremity arthroplasty: a review. J Arthroplast. 2019. https://doi.org/10.1016/j.arth.2019.05.055 **The purpose of this review was to (1) summarize and review the most recent applications of artificial intelligence and machine learning–specific to lower extremity arthroplasty, (2) discuss the origins and model-specifics of machine learning, and (3) examine the progression of machine learning into healthcare. This review specifically examines osteoarthritis gait models, joint-specific imaging, and value-based payment models. It is one of the few reviews to look at the advancement and application of machine learning within the field of orthopaedics.**

6.•• Ramkumar PN, Karnuta JM, Navarro SM, et al. Preoperative prediction of value metrics and a patient-specific payment model for primary total hip arthroplasty: development and validation of a deep learning model. J Arthroplast. 2019. https://doi.org/10.1016/j.arth.2019.04.055 **The purpose of this study was to (1) test an artificial neural network capable of predicting variables such as length of stay, inpatient charges, and discharge disposition for total hip arthroplasty, and (2) create a patient-specific payment model using this artificial neural network to account for patient complexity. Using 78,335 primary total hip arthroplasty cases, the model demonstrated area under the curves of 82.0%, 83.4%, and 79.4% for the three variables. The patient-specific payment model established risk increases of 2.5%, 8.9%, and 17.3% for moderate, major, and severe comorbidities. This study validated the use of machine learning in its prediction of patient-centered outcomes and tiering of payments based on case complexity.**

7.•• Karnuta JM, Navarro SM, Haeberle HS, et al. Predicting inpatient payments prior to lower extremity arthroplasty using deep learning: which model architecture is best? J Arthroplast. 2019. https://doi.org/10.1016/j.arth.2019.05.048 **The purpose of this study was to compare two deep learning models with different regularization techniques in predicting outcomes for primary total hip arthroplasty and total knee arthroplasty. Using 295,605 patients, the study demonstrated an improvement in performance when the DenseNet model was used with regularization. This study established an important set of features when using deep learning to more accurately predict inpatient surgical costs, and provided a foundation for other studies seeking to use machine learning.**

8. Milinovich A, Kattan MW. Extracting and utilizing electronic health data from epic for research. Ann Transl Med. 2018;6(3):42. https://doi.org/10.21037/atm.2018.01.13.

9.•• Ramkumar PN, Haeberle HS, Bloomfield MR, et al. Artificial intelligence and arthroplasty at a single institution: real-world applications of machine learning to big data, value-based care, mobile health, and remote patient monitoring. J Arthroplast. 2019. https://doi.org/10.1016/j.arth.2019.06.018 **The purpose of this review**

was to discuss the objectives of Cleveland Clinic's newly created Machine Learning Arthroplasty Laboratory: patient-specific, value-based care, and human movement. This review highlights an important step in the application of machine learning in order to improve patient outcomes in the field of orthopaedics.**

10.•• Ramkumar PN, Navarro SM, Haeberle HS, et al. Development and validation of a machine learning algorithm after primary total hip arthroplasty: applications to length of stay and payment models. J Arthroplast. 2019;34(4):632–7. https://doi.org/10.1016/j.arth.2018.12.030 **The purpose of this study was to (1) develop and validate a machine learning algorithm to predict length of stay and patient-specific inpatient payments after primary total hip arthroplasty, and (2) to propose a risk-adjusted patient-specific payment model taking patient comorbidity into account. Applying 122,334 patients to the model, the algorithm used race, age, gender, and comobidity scores to demonstrate area under the curve's of 0.87 and 0.71 for length of stay and payment. As patient complexity increased, predicting payment increased in tiers of 3%, 12%, and 32% for moderate, major, and extreme comorbidities, respectively. This study demonstrated the potential of using a risk-based patient-specific payment model that offers reimbursement commensurate with case complexity.**

11.•• Navarro SM, Wang EY, Haeberle HS, et al. Machine learning and primary total knee arthroplasty: patient forecasting for a patient-specific payment Model. J Arthroplast. 2018;(12):33, 3617–3623. https://doi.org/10.1016/j.arth.2018.08.028 **The purpose of this study was to (1) develop a machine learning algorithm to predict length of stay and inpatient costs after total knee arthroplasty, and (2) to propose a tiered patient-specific payment model commensurate with case complexity for reimbursement. Using 141,446 patients, a Bayesian model was created and trained to demonstrate an area under the curve of 0.7822 and 0.7382 for length of stay and cost. Cost ad-ons increased in teirs of 3%, 10%, and 15% for moderate, major, and extreme morality risks, respectively. This study supported the use of machine learning in predicting length of stay and cost in total knee arthroplasty, and demonstrated the poential of this application in the creation of a patient-specific payment model.**

12.•• Karnuta JM, Navarro SM, Haeberle HS, Billow DG, Krebs VE, Ramkumar PN. Bundled care for hip fractures: a machine-learning approach to an untenable patient-specific payment model. J Orthop Trauma. 2019;33(7):324–30. https://doi.org/10.1097/BOT.0000000000001454 **The purpose of this study was to (1) develop a supervised naïve Bayes machine learning algorithm to pedict length of stay and cost after hip fracture, and (2) propose a patient-specific payment model to adjust reimbursement based on patient comborbidities. Studying 98,562 Medicare patients who underwent operative management for hip fracture, a model was built that used age, sex, ethnicity, race, type of admission, risk of mortality, and severity of illness to demonstrate an accuracy of 76.5% and 79.0% for length of stay and cost. Performance was 88% for length of stay and 89% for cost. Analysis showed increasing model error with increasing risk of mortality, thus lending validity to increases in risk-adjusted payment for each risk of mortality. This study concluded that bundled care is an implausible payment model for hip fractures due to the cost of delivery of hip fracture care being dependent on non-modifiable patient-specific factors.**

13.• Bini SA, Shah R, Bendich I, Patterson J, Hwang K, Zaid M. Machine learning algorithms can use wearable sensor data to accurately predict 6-week patient reported outcome scores following

joint replacement in a prospective trial. J Arthroplast. 2019. https://doi.org/10.1016/j.arth.2019.07.024 **The aim of this study was to demonstrate the feasibility of combining machine learning with patient-generated health data in total joint arthroplasty to predict patient-reported outcome measures. The pilot study used 22 patients and used 3 activity trackers to collect 35 features from 4 weeks before to 6 weeks following surgery. Of the 22, 15 patients completed the study and 3 million data points were collected. The machine learning algorithm grouped patients into 3 clusters predictive of 6-week patient-reported outcome measure results. This study served to prove the concept that machine learning can be used in combination with patient-generated health data to predict patient-reported outcome measures following surgery.**

14.•• Cabitza F, Locoro A, Banfi G. Machine learning in orthopedics: a literature review. Front Bioeng Biotechnol. 2018;6. https://doi.org/10.3389/fbioe.2018.00075 **The purpose of this review was to provide a systematic literature review of articles published in the last two decades in which the application of machine learning was described in relation to an orthopaedic problem or pupose. The content of 70 articles was screened and analyzed in order to outline the articles' content in terms of main machine learning technique used, the orthopaedic application domain, and the source of data and quality of their predictive performance. This review is critical as it serves as one of the first to examine the use of artificial intelligence and machine learning in the field of orthopaedics specifically.**

15. Thong W, Parent S, Wu J, Aubin C-E, Labelle H, Kadoury S. Three-dimensional morphology study of surgical adolescent idiopathic scoliosis patient from encoded geometric models. Eur Spine J. 2016;25(10):3104–13. https://doi.org/10.1007/s00586-016-4426-3.

16. Olczak J, Fahlberg N, Maki A, Razavian AS, Jilert A, Stark A, et al. Artificial intelligence for analyzing orthopedic trauma radiographs. Acta Orthop. 2017;88(6):581–6. https://doi.org/10.1080/17453674.2017.1344459.

17. Artificial Intelligence and Machine Learning Across the Care Continuum in Lower Extremity Arthroplasty. Cleveland Clinic. https://consultqd.clevelandclinic.org/artificial-intelligence-and-machine-learning-across-the-care-continuum-in-lower-extremity-arthroplasty/. Accessed 16 Aug 2019

18. Shah RF, Martinez AM, Pedoia V, Majumdar S, Vail T, Bini S. Variation in the thickness of knee cartilage. The use of a novel machine learning algorithm for cartilage segmentation of MRI images. J Arthroplast. 2019. https://doi.org/10.1016/j.arth.2019.07.022.

19.• Kruse C, Eiken P, Vestergaard P. Machine learning principles can improve hip fracture prediction. Calcif Tissue Int. 2017;100(4):348–60. https://doi.org/10.1007/s00223-017-0238-7 **The purpose of this study was to apply machine learning principles in the prediction of hip fractures and estimate the predictive importance of dual-energy X-ray absorptiometry. Data from 4722 women and 717 men with 5 years of follow-up was collected and 24 statistical models were built, In women, the bootstrap aggregated flexible discriminant analysis performed best with an area under the curve of 0.91. For men, eXtreme Gradient Boosting performed best with an area under the curve of 0.89. This study's importance is highlighted by its conclusion that machine learning can improve fracture prediction beyond logistic regression using ensemble methods.**

20.• Kuo CY, Yu LC, Chen HC, Chan CL. Comparison of models for the prediction of medical costs of spinal fusion in Taiwan diagnosis-related groups by machine learning algorithms. Healthc Inform Res. 2018;24(1):29–37. https://doi.org/10.4258/hir.2018.24.1.29 **The aims of this study were (1) to compare the performance of machine learning methods for the prediction of medical costs associated with spinal fusion in terms of profit or loss, and (2) to apply these methods to explore the factors associated with the medical costs of spinal fusion. Through the categorization of 532 cases, the random forest method was found to be the most accurate predictive model, achieving an accuracy of 84.30%, a sensitivity of 71.4%, and a specificity of 92.2%, and an area under the curve of 0.904. This study demonstrated the utility of machine learning in the prediction of medical costs, a potential tool to inform hospital strategy in terms of increasing the efficiency of financial management.**

21.•• Ramkumar PN, Karnuta JM, Navarro SM, et al. Deep learning preoperatively predicts value metrics for primary total knee arthroplasty: development and validation of an artificial neural network model. J Arthroplast. 2019. https://doi.org/10.1016/j.arth.2019.05.034 **The purpose of this study was to (1) develop and validate an artificial neural network to learn and predict length of stay, inpatient charges, and discharge disposition in patients undergoing total knee arthroplasty, and (2) to apply the artificial neural network to propose a risk-based, patient-specific payment model commensurate with case complexity. Using 175,042 cases, a neural network was developed that demonstrated an area under the curve of 74.8%, 82.8%, and 76.1% for length of stay, charges, and discharge disposition. The subsequent patient-specific payment model demonstrated an increase in risk of 2.0%, 21.8%, and 82.6% as patient comorbidity increased from moderate to major to severe. This study demonstrated a model that is able to be applied to patient-specific payment models when tiering reimbursement based upon case complexity.**

22. AAHKS. 2019 Medicare physician fee schedule rule includes potentially misvalued codes for THA/TKA. http://www.aahks.org/newsroom/press-releases/2019-medicare-physician-fee-schedule-rule-includes-potentially-misvalued-codes-for-tha-tka/. Accessed 25 Nov 2018.

23. Learmonth ID, Young C, Rorabeck C. The operation of the century: total hip replacement. Lancet. 2007;370(9597):1508–19. https://doi.org/10.1016/S0140-6736(07)60457-7.

24.• Karhade AV, Schwab JH, Bedair HS. Development of machine learning algorithms for prediction of sustained postoperative opioid prescriptions after total hip arthroplasty. J Arthroplast. 2019. https://doi.org/10.1016/j.arth.2019.06.013 **The purpose of this study was to develop machine learning algorithms for the preoperative prediction of prolonged opioid prescriptions after total hip arthroplasty. Reviewing the records of 5507 patients, five machine learning algorithms were developed and determined the factors for prolonged addiction to be age, duration of opioid exposure, preoperative hemoglobin, and preoperative medications. The importance of this study is that it demontrates a utility of machine learning that can be used to improve preoperative screening and prediction of patients at risk for prolonged postoperative opioid prescriptions.**

25.•• Ramkumar PN, Haeberle HS, Ramanathan D, et al. Remote patient monitoring using mobile health for total knee arthroplasty: validation of a wearable and machine learning–based surveillance platform. J Arthroplast. 2019. https://doi.org/10.1016/j.arth.2019.05.021 **The purpose of this study was to validate the feasibility of remote patient monitoring systems in terms of the frequency of data interruptions and patient acceptance. Using a pilot cohort of 25 patients undergoing total knee arthroplasty, a mobile application and wearable knee sleeve was used to report data**

for mobility, knee range of motion, patient-reported outcome measures, opioid usage, home exercise program, and compliance. Of the 25 patients, 100% had uninterrupted passive data collection, establishing the ability to remotely acquire continuous data in postoperative patients.

26. Bloomfield RA, Williams HA, Broberg JS, Lanting BA, McIsaac KA, Teeter MG. Machine learning groups patients by early functional improvement likelihood based on wearable sensor instrumented preoperative timed-up-and-go tests. J Arthroplast. 2019. https://doi.org/10.1016/j.arth.2019.05.061.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.