Clin Orthop Relat Res (2019) 477:1777-1780
DOI 10.1097/CORR.0000000000000873

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

Value-based Healthcare

Published online: 17 July 2019

Value-based Healthcare: Can Artificial Intelligence Provide Value in Orthopaedic Surgery?

Prakash Jayakumar MBBS, BSc(Hons), MRCS(Eng), FSEM, DPhil (PhD), Meredith L. G. Moore BS, Kevin J. Bozic MD, MBA

Artificial intelligence (AI), broadly defined as a "branch of computer science that simulates intelligent behavior in computers" [9], has expanded into the healthcare sector with the promise of enhanced predictive, diagnostic, and decision-making capabilities [11]. The accelerating allure of AI in health care has been fueled by growing datasets, algorithmic innovation, storage capacity, and the steep rise in affordable computational power [11].

But can AI-related applications provide value in orthopaedic surgery?

Famed physicist, author, and cosmologist Stephen Hawking stated that AI will either be the best or worst thing to happen to humanity [3]. While we may be a long way off from a Hollywood-styled "rise of the machines," focusing our AI efforts on value may help us to transform orthopaedic health care. We believe value-oriented AI-related applications in orthopaedics might best focus on three domains: (1) Advanced data discovery and extraction, (2) improved diagnostics and prediction, and (3) enhanced clinical and decision support.

Data Discovery and Extraction

Deep-learning predictive models use advanced algorithms and multi-layered artificial neural networks to recognize, learn, and improve upon patterns and correlations from massive datasets [1]. Each successive layer builds upon the output from the previous layer, and this multi-layered structure allows for powerful conclusions based on unstructured data. The complex mapping of individual linguistic elements and idiomatic phrases for natural language processing tools used for physician dictation is one example. These "neural network" models can also take raw, unorganized data from electronic health records (EHRs) and stratify patients by risk potential, or predict adverse events such as in-hospital mortality, sepsis, multi-organ failure,

A note from the Editor-in-Chief: We are pleased to present to readers of Clinical Orthopaedics and Related Research® the latest Value-based Healthcare column (formerly Orthopaedic Healthcare Worldwide). Value-based Healthcare *explores strategies to enhance the value of musculoskeletal care by improving health outcomes and reducing the overall cost of care delivery. We welcome reader feedback on all of our columns and articles; please send your comments to* eic@clinorthop.org.

One author (KJB) certifies that he, or a member of his immediate family, has received or may receive payments or benefits as a consultant, during the study period, an amount of USD (USD 10,000 to USD 100,000) from The Centers for Medicare & Medicaid Services (Baltimore, MD, USA).
The coauthors (PJ and MGM) certify that they, or any members of their immediate families, have no commercial associations (such as consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted article.

K. J. Bozic MD, MBA (✉), Dell Medical School at the University of Texas at Austin, 1701 Trinity Street, Austin, TX 78712 USA, Email: kevin.bozic@austin.utexas.edu

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

P. Jayakumar, UK Harkness Fellow in Health Care Policy and Practice Innovation, Value Institute for Health and Care/Department of Surgery and Perioperative Care, Dell Medical School at The University of Texas at Austin, Austin, TX, USA.M. L. G. Moore, Value Based Care Fellow, Department of Surgery and Perioperative Care, Dell Medical School at The University of Texas at Austin, Austin, TX, USAK. J. Bozic, Chair, Department of Surgery and Perioperative Care, Dell Medical School at The University of Texas at Austin, Austin, TX, USA.

Wolters Kluwer

unplanned readmissions, or prolonged length of stay [13]. While these applications offer clinical benefits in terms of forecasting the risk of adverse events for acutely ill and surgical patients, future studies should examine whether AI can project pain, function, and health-related quality of life based on deep learning from big clinical data. Operational factors like clinical work flow, resource utilization, and costs of integrating such systems into electronic health platforms may also serve as fodder for examining the use of AI for data discovery.

Diagnostics and Prediction

When applied to digital-imaging data, machine and deep-learning AI approaches have demonstrated improved image acquisition and disease detection compared to conventional imaging for detecting long-bone and fragility fractures, differentiating between benign and malignant bone tumors, determining prognosis for patients with cancer [16], determining the risk of death after arthroplasty [2], and mapping disease progression for developmental dysplasia of the hip, or degenerative disease of the spine and lower extremities [1]. While many of these applications are still experimental [15], the growing evidence base for their technical and clinical efficacy suggests we are approaching a tipping point for more widespread implementation. Future studies should examine how these solutions can move the needle on process-level metrics such as minutes required for analysis and detection rate. These studies should also evaluate the impact on patient-focused outcomes and cost-effectiveness, with special consideration given to upfront costs and care delivery implications like clinical workflows and workforce utilization.

Clinical Decision Support

Clinicians and patients can use AI to engage in shared decision-making [14]. Clinicians can apply patient-reported outcomes (PROs), paired with demographic and clinical data, to machine-learning algorithms, which then provide patient-specific risk-benefit ratios for, as an example, the likelihood of achieving benefit following arthroplasty surgery [7, 14]. We have also seen machine-learning data applied to shared decision-making in knee osteoarthritis, ACL deficiency, and complex degenerative spinal disease [1]. Still, prospective, randomized studies will be needed to assess the impact of these advanced decision aids on decision quality, patient experience, patient limitations, and episode of care costs.

AI-enabled systems are also enhancing procedural efficiency in the operating room by guiding surgical teams through orthopaedic procedures. Although these functions have the potential to reduce variation on the operating table, they may simultaneously put clinical acumen and relationship-building opportunities in jeopardy with an inherent focus on numbers rather than the individual patient [8]. For example, AI in telehealth for virtual consultations and remote diagnostic support provides clear benefits for managing patients remotely (supporting earlier discharge home), but it also decreases human contact; this may at times be detrimental given that face-to-face interaction between patient and provider offers a valuable chance to pick up nuanced verbal and non-verbal cues [5].

Ethical, Legal, Policy, and Practice Considerations

A broad set of ethical, legal, policy, and practice-related aspects of AI implementation must be considered if we are to truly gauge whether this technology will deliver greater value to orthopaedic patients.

Ethical Issues

There are several ethical challenges to the use of AI during surgical procedures that should be identified and mitigated given the variable levels of risk to patient privacy and confidentiality, consent, and patient autonomy. For instance, how should we consent patients when performing AI-enabled robotic-assisted orthopaedic procedures?

We believe that surgeons should disclose exactly how the procedure will be done, emphasizing differential risk profile potentially associated with an AI-assisted operation. There are still some unknowns associated with the potential for robots to autonomously adapt and alter the course of a procedure [10]. The current regulatory frameworks and ethical guidelines appear to be falling behind the pace of technological progress in health care [12, 17]. Market approval on devices based on predictive analytic solutions and AI rests on developing the requisite regulatory framework for integrating these tools into clinical practice. Parikh and colleagues [12] recently recommended that predictive analytic solutions include meaningful endpoints (those driving changes relevant and beneficial to clinicians and patients), appropriate benchmarks (assuring these are dynamic based on evolving sets of real-world data and performance), and longitudinal audit mechanisms (incorporating post-market surveillance of a solution inherently shifting over time as the algorithm learns) [12]. As they suggest, the true impact for these tools is observed at scale and these sweeping advancements require inter-operable solutions that are both generalizable and genuinely reflective of the populations utilized to build algorithms [12].

But first, regulatory and professional bodies should provide guidance on interventions aligned with these AI solutions [12]. Indeed, we need guidelines for clinically oriented AI tools; the need seems especially pressing in light of the number of innovations seeking market access (or achieving it; in 2018, the FDA granted premarket clearance for a machine-learning system that predicts vital-sign instability [4]).

Understanding the Machine

Surgeons should understand the population characteristics used to develop algorithms and work with programmers to calibrate their outputs to match the needs of the patient populations being treated. Algorithms generated using certain data may not be accurate or fully translatable to populations with different sociodemographic characteristics. A poor understanding of the data used to build the "black box" algorithms, how a given output was reached, and what the output actually "means" may generate potential legal, medical malpractice, and product-liability issues (placing proper informed consent in jeopardy). Access to the data may be challenging, since AI technology companies commonly withhold datasets as proprietary information. Surgeons should take it upon themselves to understand the nature and resolution of data inputs to the model, and monitor global validation efforts specific to the AI tool in consideration.

Impact on Medical Education and Clinical Practice

If an AI computer model can diagnose orthopaedic conditions more accurately and efficiently than a board-certified musculoskeletal radiologist, or if an AI-enabled robot can perform certain techniques more precisely than an orthopaedic surgeon, extensive capital investments in medical education and clinical training are avoided. How does this impact the physician workforce? AI technology may signal a need for redesign of the current medical education paradigm and the service industry orientation of health care, where future professionals spend their time interacting with and adapting to this technology for safe use in clinical settings. A key aspect of medical education would become the interaction between providers, patients, and technology. Based on the complexity of clinical care, it is likely such technologies will complement physicians rather than act as substitutes. An approach of "working with AI" will surely uncover complexities in relation to the reactions and interactions of patients and caregivers with technology. Clear boundaries will need to be set for the differential roles of machines and physicians in patient care.

The Cost Perspective

The functions achieved by AI technologies (data discovery, diagnostics and prediction, and clinical decision support) aim to reduce overall spending through reducing time, resource utilization, manpower, and computational power. However, these potential savings should be considered alongside the up-front investment, plus the challenges of obtaining clinical buy-in regarding the requisite culture change necessary for success. While the drive to adopt such new technologies is strong in orthopaedics and is reflected by an impressive legacy of successful innovations (advanced analysis of musculoskeletal pathology in imaging, useful pattern-seeking within the "big data" of the EHR), there are also some catastrophic failures in the portfolio—metal-on-metal joint replacement stands as a prominent example.

We appear to be at point where there is a critical need for well-designed AI technologies to be tested in clinical workflows, validated in clinical decision-making processes, and evaluated for impact on patient clinical outcomes (whether patient-reported or system-level) [12]. There is also a need to demonstrate the impact of AI on value with high-quality cohort and randomized controlled trials utilizing patient-reported outcome measures, cost-effectiveness analyses, and rapid quality improvement studies compliant with evolving research standards in this arena such as the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis [12]. This type of research needs to be conducted while working out ethical, policy, and practice guidelines. Given the expense associated with development of AI-based tools, we must focus on value: The demonstrated benefits of any new tools divided by their costs.

References

1. Cabitza F, Locoro A, Banfi G. Machine learning in orthopedics: A literature review. *Front Bioeng Biotechnol*. 2018;6:75.

2. Harris AHS, Kuo AC, Weng Y, Trickey AW, Bowe T, Giori NJ. Can machine learning methods produce accurate and easy-to-use prediction models of 30-day complications and mortality after knee or hip arthroplasty? *Clin Orthop Relat Res*. 2019;477:452-460.
3. Hern A. Stephen Hawking: AI will be 'either best or worst thing' for humanity. Available at: https://www.theguardian.com/science/2016/oct/19/stephen-hawking-ai-best-or-worst-thing-for-humanity-cambridge. Accessed June 3, 2019.
4. Hravnak M, Edwards L, Clontz A, Valenta C, DeVita MA, Pinsky MR. Defining the incidence of cardiorespiratory instability in patients in step-down units using an electronic integrated monitoring system. *Arch Intern Med*. 2008;168:1300–1308.
5. Jacobs J. Doctor on video screen told a man he was near death, leaving relatives aghast. Available at: https://www.nytimes.com/2019/03/09/science/telemedicine-ethical-issues.html?smid=nytcore-ios-share. Accessed June 10, 2019.
6. Rail Leap. Streamline operating room management. Available at: https://leaprail.com/. Accessed June 3, 2019.
7. Leopold SS. Editor's Spotlight/Take 5: Can machine learning algorithms predict which patients will achieve minimally clinically important differences from total joint arthroplasty? *Clin Orthop Relat Res*. 2019;477:1262-1266.
8. Makhoul S. Digital surgery. Available at: https://digitalsurgery.com/. Accessed June 3, 2019.
9. Webster Merriam. Artificial intelligence. Available at: https://www.merriam-webster.com/dictionary/artificial%20intelligence. Accessed June 3, 2019.
10. O'Sullivan S, Nevejans N, Allen C, Blyth A, Leonard S, Pagallo U, Holzinger K, Holzinger A, Sajid MI, Ashrafian H. Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. *Int J Med Robot*. 2019;15:e1968.
11. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. *N Engl J Med*. 2016; 375:1216–1219.
12. Parikh RB, Obermeyer Z, Navathe AS. Regulation of predictive analytics in medicine. *Science*. 2019;363:810–812.
13. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. Available at: https://www.nature.com/articles/s41746-018-0029-1. Accessed June 3, 2019.
14. Sambare T, Uhler L, Bozic KJ. Shared decision making: Time to get personal. Available at: https://catalyst.nejm.org/shared-decision-making/. Accessed June 3, 2019.
15. Schmaranzer F, Helfenstein R, Zeng G, Lerch TD, Novais EN, Wylie JD, Kim YJ, Siebenrock KA, Tannast M, Zheng G. Automatic MRI-based three-dimensional models of hip cartilage provide improved morphologic and biochemical analysis. *Clin Orthop Relat Res*. 2019;477:1036-1052.
16. Thio QCBS, Karhade AV, Ogink PT, Raskin KA, De Amorim Bernstein K, Lozano Calderon SA, Schwab JH. Can machine-learning techniques be used for 5-year survival prediction of patients with chondrosarcoma? *Clin Orthop Relat Res*. 2018;476:2040-2048.
17. US Food and Drug Administration. Statement from FDA Commissioner Scott Gottlieb, M.D. on steps toward a new, tailored review framework for artificial intelligence-based medical devices. Available at: https://www.fda.gov/news-events/press-announcements/statement-fda-commissioner-scott-gottlieb-md-steps-toward-new-tailored-review-framework-artificial. Accessed June 10, 2019.