HHS Public Access
Author manuscript
JAMA Surg. Author manuscript; available in PMC 2024 April 01.

Published in final edited form as:
JAMA Surg. 2023 April 01; 158(4): 337–338. doi:10.1001/jamasurg.2022.5444.

Predictive Analytics and Artificial Intelligence in Surgery—Opportunities and Risks

Kathryn Colborn, PhD, MSPH,
University of Colorado Denver, Anschutz Medical Campus, Aurora.

Gabriel Brat, MD,
Beth Israel Deaconess Medical Center, Boston, Massachusetts.

Rachael Callcut, MD
University of California, Davis, Sacramento.

Predictive analytics and artificial intelligence (AI) are being applied in hospitals across the world to aid in clinical decision support, discuss risk of certain procedures with patients, and identify patients whose clinical status is deteriorating. Researchers are using massive amounts of data and recent advances in machine learning to improve surgical quality and patient outcomes. The Surgical Outcomes Club, a consortium of surgeons and health services researchers who advance the science of surgical outcomes research, convened a panel of 4 experts at the 2022 meeting who highlighted their research and experience using predictive analytics and AI in surgical research. Three main areas of AI in surgery were discussed: computer vision, digital transformation at the point of care, and electronic health records data. They discussed the opportunities and risks of these areas of AI in surgery, and in this Viewpoint, we expand on that discussion.

Computer Vision

Video is a by-product of minimally invasive and robotic surgery and is increasingly being collected for open procedures. The potential exists for real-time annotation of video streams by automated algorithms to track surgical performance, identify complex anatomy, and provide feedback to reduce technical errors. In parallel to the growth of video-based analysis by human reviewers, computer vision for surgery is also useful for surgical skills training[1] and review of surgeon behavior. Annotation of surgical phases and tracking of surgical tools

Corresponding Author: Kathryn Colborn, PhD, MSPH, Department of Surgery, University of Colorado, 12631 E 17th Ave, C-305, Aurora, CO 80045 (kathryn.colborn@cuanschutz.edu).
Additional Contributions: The authors would like to acknowledge panelist Robert Meguid, MD, MPH (University of Colorado), for his invaluable contribution to the Surgical Outcomes Club and the content included in this article.

Conflict of Interest Disclosures: Dr Colborn reported grants from the Agency for Healthcare Research and Quality (AHRQ). Dr Callcut reported technology licensed to GE Healthcare, being a cofounder of BeeKeeperAI, a patent for 62/824,183-Federated Machine Learning Techniques for Highly Curated Health-Care Datasets licensed to BeeKeeperAI with royalties paid from University of California, San Francisco, a patent for 051392-Method and Systems for the Automated Detection of Free Fluid Using Artificial Intelligence for the Focused Assessment Sonography for Trauma (“FAST”) Examination for Trauma Care pending University of California, San Francisco hold patent, and grants from GE Healthcare, National Institutes of Health, Philips Health Care to their institution. Some of the work described in this Viewpoint is supported by grant 1R01HS027417 from the AHRQ. No other disclosures were reported.

Disclaimer: This content is solely the responsibility of the authors and does not necessarily represent the official views of the AHRQ.

and hands[2] will allow coaches to evaluate this skill with quantifiable metrics that would otherwise require tedious human video annotation and qualitative review. Furthermore, advances in statistical methods applied to video data, such as convolutional neural networks, have tremendously improved the capability of robotic procedures—essentially giving computers the ability to see.[3]

Surgical vision tools can augment the trajectory of surgeons in training and capabilities of human operators. However, the high difficulty of operations makes AI for surgical computer vision difficult. Simulated or toy examples where performance is high do not yet generalize to real-world complexity. Overcoming these limitations of generalizability is hampered by inconsistent video annotation and lack of large and diverse data sets. Institutional interests and data sharing limitations have prevented creation of a large open data set drawn from many institutions that would be necessary to train robust algorithms. Surgical video data from publicly available sites are being used for training models, but the quality of these video sources is variable. A recent consensus statement[4] suggested that a 10-year time horizon is likely for real-time implementation of these algorithms, and retrospective training tools may be available in the next 2 years. Appreciating the nascent nature of surgical vision will allow it to develop as a collaboration that will ensure that tools work in concert with surgeons and ultimately transform surgical care.

Digital Transformation at the Point of Care

Advances in AI, including the ability to harness large data sets, have now made it possible to apply advanced analytics to a range of optimizing care provided in the intensive care unit and at the point of care in high-intensity environments, including the trauma bay. These time-pressured situations, where complex information streams are plentiful and the stakes high for making the right decision, have significant promise for the application of AI technology. Signal processing approaches can theoretically now convert multimodality data streams coming from monitors, point-of-care testing devices, imaging, laboratory, and real-time documentation by bedside clinicians into early warning signals. These signals can draw attention of the clinicians to patients who seemingly look well but are in fact, off trajectory. The net effect is to augment clinician decision-making through leveraging technology advances. Although still in its infancy, there are several demonstration projects that have been recently published showing AI assistance in detecting early respiratory deterioration, guiding fluid resuscitation, earlier detection of sepsis or critical illness, and identification of potentially life-threatening findings on radiographic imaging. Each of these requires careful implementation to avoid unintended consequences, ongoing monitoring when in clinical environments to ensure appropriate outcomes, and external validation before reaching clinical production deployment.

Electronic Health Records Data

Electronic health record (EHR) data are frequently used for development of AI in quality improvement and surgical outcomes research. Preoperative risk estimation using EHR data is currently implemented in some institutions to inform discussions between surgeons and patients prior to undergoing an operation. Moreover, given the high cost and resources required for postoperative complication surveillance and reporting, such as that of the National Surgical Quality Improvement Program and the National Health and Safety Network, some institutions have begun to replace manual medical record review with AI. This affords the opportunity to report complication rates at the level of the surgeon specialty and individual surgeon on a frequent basis—a task that would be impossible to accomplish without AI. There are many advantages to using HER data. They almost always use common data models and universal codes, such as *International Classification of Diseases* or *Current Procedural Terminology*. These universal coding systems allow researchers to produce generalizable results; however, due to their high dimensionality, they can be difficult to work with. Recently, mappings of these codes to larger disease classes, such as PheCodes,[5] have greatly improved the ability to work with HER data and produce generalizable results. Additionally, when structured EHR data are limited, unstructured or narrative text data can be used to fill in the gaps. There are many examples of applying natural language processing (NLP) to the EHR notes, for example to identify postoperative complications or symptoms of a urinary tract infection. NLP methods have advanced tremendously over the last decade, and pretrained models for EHR data have proven to be extremely useful, such as clinical Bidirectional Encoder Representations From Transformers (BERT).[6] However, there are some risks with using HER data for AI. The data can introduce bias in the AI system. If data are missing due to lack of access to care, poor recording, or lost to follow-up of a patient, conclusions will be incorrect and ungeneralizable. Furthermore, sicker patients tend to be included in hospital EHR data, so collider bias can introduce distorted associations between risk factors and clinical outcomes. It is therefore paramount to familiarize yourself with the source data prior to developing the AI. Moreover, the limitations of EHR data are relevant to most clinical data sources, including those used in computer vision and digital transformation at the point of care. These limitations should always be addressed in publications and the conclusions should be appropriately tempered. Finally, not all data from the HER may be used for development of AI, and there is some ethical debate about whether informed consent should be sought for the use of EHR data for AI.[7]

Conclusions

Predictive analytics and AI in surgery have improved processes and quality of care. Significant impacts have been in early detection of patient deterioration, reduction in costs and manual labor of surveillance and reporting of complications, and training algorithms to analyze surgical execution. However, because the development of AI in surgery relies heavily on limited data sets with inherent biases and limitations, a critical evaluation of sources of bias in the data should be carried out and monitored prospectively. Finally, conclusions should always reflect the precision in the findings when developing AI in surgical outcomes and health services research.

REFERENCES

1. Fazlollahi AM, Bakhaidar M, Alsayegh A, et al. Effect of artificial intelligence tutoring vs expert instruction on learning simulated surgical skills among medical students: a randomized clinical trial. JAMA Netw Open. 2022;5(2):e2149008–e2149008. doi:10.1001/jamanetworkopen.2021.49008 [PubMed: 35191972]

2. Goodman ED, Patel KK, Zhang Y, et al. A real-time spatiotemporal AI model analyzes skill in open surgical videos. arXiv. Published online December 14, 2021. https://arxiv.org/abs/2112.07219
3. Ward TM, Mascagni P, Ban Y, et al. Computer vision in surgery. Surgery. 2021;169(5):1253–1256. doi:10.1016/j.surg.2020.10.039 [PubMed: 33272610]
4. Vedula SS, Ghazi A, Collins JW, et al. ; Collaborative for Advanced Assessment of Robotic Surgical Skills. Artificial intelligence methods and artificial intelligence-enabled metrics for surgical education: a multidisciplinary consensus. J Am Coll Surg. 2022;234(6):1181–1192. doi:10.1097/XCS.0000000000000190 [PubMed: 35703817]
5. Wei WQ, Bastarache LA, Carroll RJ, et al. Evaluating phecodes, clinical classification software, and ICD-9-CM codes for phenome-wide association studies in the electronic health record. PLoS One. 2017;12(7):e0175508. doi:10.1371/journal.pone.0175508 [PubMed: 28686612]
6. Alsentzer E, Murphy J, Boag W, et al. Publicly Available Clinical BERT Embeddings. Proceedings of the 2nd Clinical Natural Language Processing Workshop. Published online June 2019. doi:10.18653/v1/W19-1909
7. Naik N, Hameed BMZ, Shetty DK, et al. Legal and ethical consideration in artificial intelligence in healthcare: who takes responsibility? Front Surg. 2022;9:862322. doi:10.3389/fsurg.2022.862322 [PubMed: 35360424]