Editorial

The Promise of Artificial Intelligence in Cardiothoracic Surgery

Michael Salna, M.D., M.B.A.

Division of Cardiac, Thoracic, and Vascular Surgery, Department of Surgery, Columbia University Irving Medical Center, New York, NY, USA

Corresponding author
Michael Salna
Tel 1-212-305-5156, **Fax** 1-212-305-2633, **E-mail** mp2191@cumc.columbia.edu, **ORCID** https://orcid.org/0000-0002-0565-6485

Introduction

Artificial intelligence (AI) powers technologies we rely on every day—from movie streaming recommendations to social media content and credit card fraud alerts. What makes AI unique is the nature of its "intelligence"—specifically, its ability to improve its own predictive accuracy through learning by experience.

With its ability to rapidly process immense amounts of data, AI can help overcome fundamental limitations of the more classic researcher-driven strategy: missing out on potentially meaningful relationships between disregarded variables, failing to identify more complex nonlinear relationships, or incorrectly discarding variables deemed to be less clinically relevant.

As pressures mount for greater operational efficiency within hospitals and patients demand more personalized medicine, healthcare systems can employ AI to harness the wealth of data they generate to enhance the effectiveness and efficiency of care they provide. In cardiothoracic surgery, AI can be implemented to enhance all phases of a patient's experience, from preoperative risk counseling to cardiopulmonary bypass (CPB) optimization and postoperative complication prevention. Table 1 outlines explanations of examples of commonly used AI terms and techniques (Figs. 1, 2).

Preoperative applications of artificial intelligence

From the moment a patient walks into a surgeon's office, AI can improve the quality, experience, and efficiency of care. Proof-of-concept studies have already demonstrated the potential of AI to recognize and extract speech in clinic visits and generate consult notes summarizing such conversations through natural language processing [1]. While these technologies are still being developed, what may be more relevant to patients is surgical candidacy and risk.

Currently, prognostic calculators based on complex multivariable regression are used to determine a patient's risk of death, dialysis, prolonged ventilation, and other unfavorable outcomes (Fig. 3). Typically, these scores are restricted to a handful of variables generated from human data entry and computation.

Machine learning (ML), however, can analyze data on a scale that would have previously been impossible, such as automatically drawing data directly from electronic health records (EHRs) or claims data. Algorithms can then be generated by assessing the relevance of all available variables and identifying complex, nonlinear relationships between predictors. These predictions are then iteratively improved over time as the algorithm is fed more data.

Several studies have demonstrated the superiority of ML-generated risk scores over traditional regression-based scores, including the European System for Cardiac Operative Risk Evaluation (EuroSCORE) II, and modest improvements compared with the Society of Thoracic Surgeons Predicted Risk of Mortality. For example, in a prospective study of 18,362 cardiac surgical patients, Nilsson et al. [2] used artificial neural networks to select 34 of the most relevant mortality predictors out of 72 variables, yielding an area under the receiver operating curve (AUROC) of 0.81 compared with the 0.79 generated by the EuroSCORE II (p=0.001).

Table 1. Key terms in artificial intelligence and machine learning with potential applications in cardiothoracic surgery

Term	Definition	Potential applications
Artificial intelligence	A field of computer science where machines are trained to do tasks typically requiring human intelligence	Everything from identifying congenital cardiac gene targets to predicting probability of a wound infection from photographs to all the examples listed below
Machine learning	Subset of AI where computers can learn to identify data patterns without explicit programming	Generating more accurate necessary OR case durations for a given surgeon by combining historical OR data, prior patient data, and hospital capability data
Supervised learning	Subcategory of ML where labeled data are used to train classification or prediction algorithms	Training a computer to identify the presence or absence of a pneumothorax on X-ray
Unsupervised learning	Subcategory of ML where the machine uses unlabeled data to identify patterns and make conclusions	Identifying clusters of mitral valve features that would most benefit from multiple artificial chords in the repair strategy
Ensemble learning	An approach to ML where multiple models are combined (e.g., multiple decision trees combined for random forest classification). See Fig. 1.	An automated ventilator determining the need to increase inspiratory time based on measured expiratory volume, flow, and pressure
Deep learning	Branch of ML using neural networks with many layers to analyze and make representations of data, similar to how a human would think. See Fig. 2.	A more accurate and continuously improving surgical mortality prediction score
Computer vision	Field of AI enabling computers to recognize, interpret, and respond to visual input data	A camera that recognizes when drapes are coming down to automatically alert cleaning staff for room turnover
Natural language processing	Field of AI enabling computers to recognize, inter pret, and respond to text or voice data	Extracting speech and generating clinic notes from dictations

AI, artificial intelligence; ML, machine learning; OR, operating room.
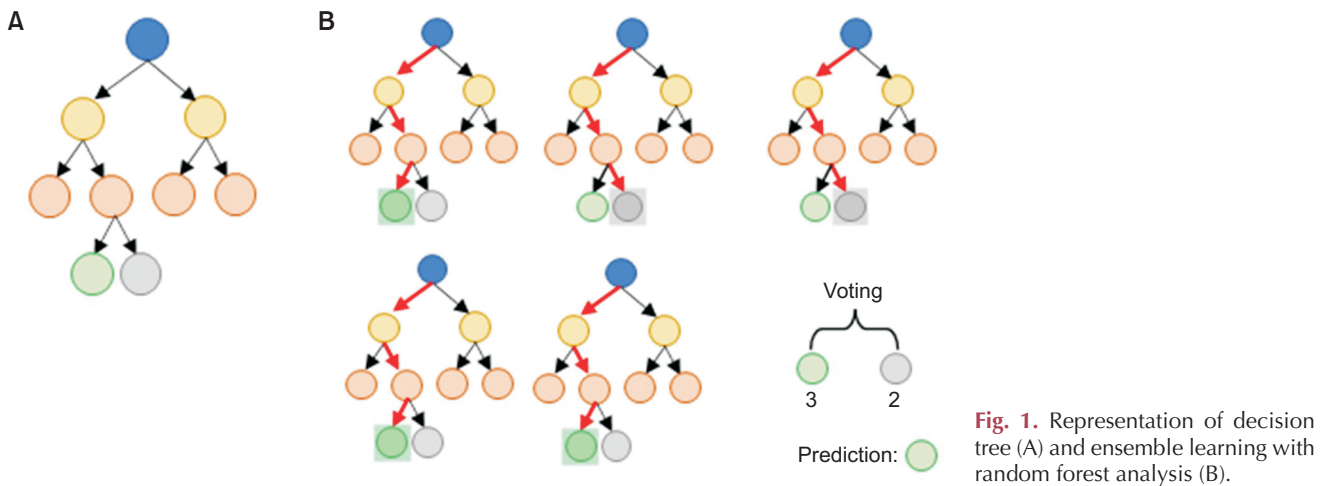

Fig. 1. Representation of decision tree (A) and ensemble learning with random forest analysis (B).

Interestingly, a 2019 systematic review of 72 studies comparing logistic/linear regression (LR) and ML models for binary outcomes ultimately found no evidence of superior performance of ML over LR [3]. This does not signal a death knell for ML, but rather suggests that the 2 types of models may work together synergistically. What ML lacks in clinical knowledge, it makes up for in being able to parse enormous data sets and pick up on nuanced data associations. Thus, LR and ML models can work together, with data scientists developing LR models informed by ML results and honed by providers' clinical expertise.

Intraoperative applications of artificial intelligence

ML has the potential to optimize the entire operating room (OR) workflow, from perfusionists managing CPB and surgeons performing the operation to hospital-wide initiatives at improving OR efficiency.
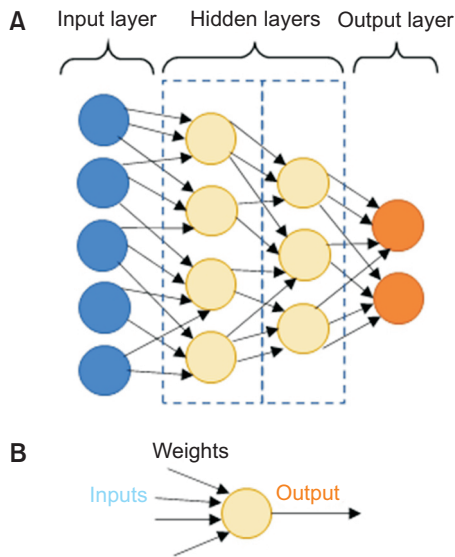
Fig. 2. Representation of an artificial neural network (A) composed of individual "neurons" (B) receiving multiple inputs to generate an output.

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$$

Fig. 3. Graphical representation of linear regression, a traditional statistical prediction approach, with a predicted result based on the data-derived regression equation.

Artificial intelligence for perfusionists and cardiopulmonary bypass

Metabolic management is crucial during CPB, but there are currently no integrated, automated, systems to optimize oxygen delivery and reduce anaerobic respiration. Emerging evidence supports the use of an algorithmic technique known as goal-directed perfusion (GDP) in preventing anaerobic metabolism–related complications such as post-cardiac surgery acute kidney injury (AKI) [4]. In GDP, CPB parameters are adjusted to maintain sufficient oxygen delivery above a critical value through real-time changes in variables such as flow and hematocrit rather than simply flowing to achieve a certain cardiac index. This is an excellent opportunity for ML integration, as patterns of hemodynamics and tissue oxygen delivery can be identified and used to develop algorithms that adjust parameters such as flow, sweep, and temperature continuously in response to changes in patient physiology. Perfusionists could then fine-tune CPB parameters and manage the algorithms in real-time as they receive feedback from the surgical team.

Artificial intelligence for surgeons

As AI offers training opportunities to perfusionists, so too does it provide surgeons with opportunities to hone their skills.

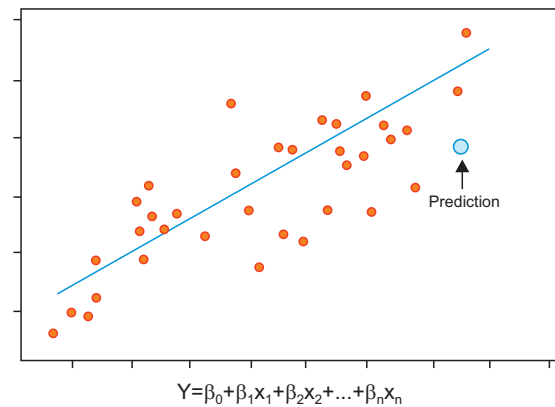Computer vision using neural networks that mimic the

structure of the visual cortex has given computers human-like recognition ability. Harnessing this technology, ML algorithms capable of integrating motion analysis and force usage will make it possible to acquire automated and accurate skills assessments [5]. Several studies in robotic surgery have used ML methods to evaluate robotic surgery skills by extracting data on kinematic data such as task completion time, motion speed, smoothness, and tortuosity of needle angles to differentiate expert from novice performance [6].

Scaling up the assessment of surgical skills to an entire procedure, however, is no small feat. AI must not only learn to recognize a particular instrument, but also must recognize that instrument within the dynamic surgical field filled with visual disturbances such as smoke, blood, smudged cameras, and varying angles. Computers must also understand the various phases of a procedure (i.e., knowing which steps are occurring and which are next).

The prospect of a truly autonomous robotic surgeon—one capable of the classic robot control architectures: sensing, planning, and acting—is still a very long way away. However, Tanwani et al. [7] developed a semi-supervised AI deep-learning system named Motion2Vec that, through watching publicly available surgical videos, has broken down the key movements of suturing (needle insertion, extraction, transfer) and has mimicked these movements to a high degree of accuracy.

Through the development of an ever-growing database of different sets of surgical maneuvers, researchers have continuously decomposed skills such as suturing and knot tying. While autonomous robotic surgery is nowhere near ready for prime time, it is possible that within our lifetime we will see robotic assistants emerging to help with simple tasks such as suture following and video-assisted thoracic

surgery camera navigation, enabling surgeons to focus on more demanding parts of the case.

Artificial intelligence for surgical logistics

The OR is one of the highest revenue generators for a hospital, but also one of the most expensive. AI technologies can be integrated into the OR workflow to predict case durations more accurately and minimize unnecessary delays.

These delays often arise from poor synchronization between workflows within and outside of the OR, where poor communication between teams and potentially inaccurate scheduling exacerbates an already inefficient system. ML approaches can refine case duration predictions by analyzing procedure type, patient factors, and surgical team expertise. The University of Colorado began implementing ML in its transition away from traditional block scheduling and has already achieved a $10 million increase in revenue as a result [8].

Using telemonitoring, context-aware AI could further optimize OR efficiency and turnover time (and, through enhancing OR utilization, improve profitability) through a variety of other ways such as automatically alerting the intensive care unit (ICU) that the surgeon is closing, cleaning staff that the drapes are coming off, and staff in the preoperative area that the room is being cleaned to begin preparing for the next patient.

Postoperative applications of artificial intelligence

Artificial intelligence for imaging and pathology

ML pattern recognition algorithms will inevitably become the standard platform for interpreting imaging and pathology slides. By eliminating the variability in interpretations between readers and having the capability to function 24/7, these algorithms will soon exceed human accuracy, efficiency, and performance. This has already been illustrated by Yala et al. [9], whose deep learning algorithm was significantly more accurate at predicting the presence or absence of cancer in mammograms than general radiologists.

Though there are still challenges to overcome—processing huge, digitized files of histopathology slides or computed tomography (CT)/magnetic resonance imaging series, interpreting different stains of the same tissue type, and overcoming noisy 2-dimensional slices of ultrasound im-

ages—these are all being worked on [10].

The introduction of these disruptive technologies will be on the order of years, not decades. A future of nearly instant preliminary frozen section reads during lung cancer resections and CT diagnostics will likely be within our lifetime. This does not mean that radiologists and pathologists will be out of work, but rather will pivot from primarily reading scans/slides to interpreting digitized images, auditing ML-based reads, and feeding more data into algorithms to enhance their learning and accuracy.

Artificial intelligence in the intensive care unit

Effective ICU management is essential for cardiothoracic surgical patients. The ICU generates immense amounts of data, continuously streaming in for each patient and often requiring real-time changes to be made in therapies such as vasoactive medications or ventilator settings, making it a rich target for ML. For example, ML methods have the potential to detect and classify patient-ventilator asynchrony, direct weaning more precisely, and change ventilator parameters in real time in response to dynamic changes in patient effort, compliance, or airway resistance [11].

AI may also aid in predicting impending complications. For example, Lin et al. [12] developed a random forest-based model to predict mortality in AKI patients requiring ICU care with an AUROC of 0.87 based on urine output, systolic blood pressure, age, serum bicarbonate, and heart rate. Algorithms have also been developed to predict a range of ICU parameters, from impending septic shock and AKI to acute hypotensive episodes and volume responsiveness. Integrating AI into electronic medical records could potentially flag impending complications based on collating already available data and alerting providers to clinical patterns they may not have recognized.

Limitations

Despite the immense promise of AI-based technologies in cardiothoracic surgery, AI is not a panacea. Instead, effective AI use hinges on asking the right questions and providing the appropriate type and amount of data to answer those questions.

Models can overfit based on noise rather than the true signal

ML algorithms employ whatever signals are available in large datasets to generate the best-performing model. As a result, they are subject to overfitting and spurious correla-

tions leading to overly optimistic estimates of model accuracy. For example, a deep neural network algorithm was more likely to classify a skin lesion as malignant if an image had a ruler in it, given the correlation between the presence of a ruler and a higher likelihood of cancer in the training dataset [13].

This overfitting portends future issues with generalizability to other clinical sites given differences in equipment, EHR systems, and laboratory assays. To overcome this, external validation testing of AI algorithms with large multi-institutional datasets—that is, data from institutions other than those that supplied data for the initial model development—will be necessary to assess real-world performance potential.

Algorithmic interpretability

Despite AI and ML advancements in healthcare, their usefulness is still limited by difficulty in understanding the algorithms' decision-making processes. For example, the best performing AI models (e.g., deep learning artificial neural networks) also tend to be the least explainable, with intermediate steps being shrouded in the black box of hidden layers (Fig. 2). This is of particular concern in healthcare where explainability and transparency are paramount for patient applications given the potentially dangerous consequences of unrecognized system errors, inappropriately used variables, or illicitly manipulated algorithmic components.

AI's dependency on data networks makes it vulnerable to cyberattacks

AI applications and devices integrated into hospital data networks can all serve as doorways for cyberattacks, from stealing personal health information to wirelessly taking over ICU syringe pumps [14]. Healthcare systems will need to invest in improving the visibility of all devices in their network and work collaboratively with device manufacturers and EHR providers to develop sophisticated security measures to prevent breaches in data and devices.

Challenges with implementation

Healthcare organizations will be challenged to implement many of these technologies due to high upfront capital requirements. In addition to investing in costly software and maintenance, as well as the need for huge amounts of data to develop meaningful predictive algorithms, hospitals must compete with "Big Tech" to attract and retain AI talent. The solution may lie in larger professional organizations such as the Society for Thoracic Surgeons acting as the primary data repository and investing in AI talent on a project-to-project basis.

Conclusion

Exciting advances in AI and ML hold immense promise in cardiothoracic surgery, offering value across a patient's entire journey through surgery. Ongoing work and interdisciplinary collaborations between surgeons, data and computer scientists, and engineers will continue to be essential to finetune these algorithms to provide more productive and effective care in an environment that protects patient safety and privacy. Cardiothoracic surgeons have a history of pushing the boundaries of the "possible," and AI is yet another frontier where our specialty can be at the forefront of surgical innovation.

Article information

ORCID

Michael Salna: https://orcid.org/0000-0002-0565-6485

Author contributions

Conceptualization: MS. Methodology: MS. Project administration: MS. Visualization: MS. Writing–original draft: MS. Writing–review & editing: MS.

Conflict of interest

No potential conflict of interest relevant to this article was reported.

Funding

References

1. Rajkomar A, Kannan A, Chen K, et al. *Automatically charting symptoms from patient-physician conversations using machine learning.* JAMA Intern Med 2019;179:836-8. https://doi.org/10.1001/jamainternmed.2018.8558

2. Nilsson J, Ohlsson M, Thulin L, Hoglund P, Nashef SA, Brandt J. *Risk factor identification and mortality prediction in cardiac surgery using artificial neural networks.* J Thorac Cardiovasc Surg

2006;132:12-9. https://doi.org/10.1016/j.jtcvs.2005.12.055

3. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. *A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models*. J Clin Epidemiol 2019;110:12-22. https://doi.org/10.1016/j.jclinepi.2019.02.004

4. Ranucci M, Johnson I, Willcox T, et al. *Goal-directed perfusion to reduce acute kidney injury: a randomized trial*. J Thorac Cardiovasc Surg 2018;156:1918-27. https://doi.org/10.1016/j.jtcvs.2018.04.045

5. Andras I, Mazzone E, van Leeuwen FW, et al. *Artificial intelligence and robotics: a combination that is changing the operating room*. World J Urol 2020;38:2359-66. https://doi.org/10.1007/s00345-019-03037-6

6. Wang Z, Majewicz Fey A. *Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery*. Int J Comput Assist Radiol Surg 2018;13:1959-70. https://doi.org/10.1007/s11548-018-1860-1

7. Tanwani AK, Sermanet P, Yan A, Anand R, Phielipp M, Goldberg K. *Motion2vec: semi-supervised representation learning from surgical videos*. Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA); 2020 May 31-Aug 31; Paris, France. Piscataway (NJ): IEEE; 2020. https://doi.org/10.1109/ICRA40945.2020.9197324

8. LaPointe J. *OR efficiency, machine learning boosts UCHealth's revenue by $10M* [Internet]. Danvers (MA): Recyle Intelligence; 2018 [cited 2022 Jan 25]. Available from: https://revcycleintelligence.com/news/or-efficiency-machine-learning-boosts-uchealths-revenue-by-10m.

9. Yala A, Schuster T, Miles R, Barzilay R, Lehman C. *A deep learning model to triage screening mammograms: a simulation study*. Radiology 2019;293:38-46. https://doi.org/10.1148/radiol.2019182908

10. Singh SP, Wang L, Gupta S, Goli H, Padmanabhan P, Gulyas B. *3D deep learning on medical images: a review*. Sensors (Basel) 2020;20:5097. https://doi.org/10.3390/s20185097

11. Ossai CI, Wickramasinghe N. *Intelligent decision support with machine learning for efficient management of mechanical ventilation in the intensive care unit: a critical overview*. Int J Med Inform 2021;150:104469. https://doi.org/10.1016/j.ijmedinf.2021.104469

12. Lin K, Hu Y, Kong G. *Predicting in-hospital mortality of patients with acute kidney injury in the ICU using random forest model*. Int J Med Inform 2019;125:55-61. https://doi.org/10.1016/j.ijmedinf.2019.02.002

13. Esteva A, Kuprel B, Novoa RA, et al. *Dermatologist-level classification of skin cancer with deep neural networks*. Nature 2017;542:115-8. https://doi.org/10.1038/nature21056

14. Luz E. *CyberMDX discovers vulnerability in the Becton Dickinson Alaris TIVA syringe pump* [Internet]. New York (NY) CyberMDX; 2018 [cited 2022 Jan 25]. Available from: https://www.cybermdx.com/research/vulnerability-bd-alaris-pump-180508.