Surgical Innovation

Explainable Artificial Intelligence for Safe Intraoperative Decision Support

Lauren Gordon, MD, MSc; Teodor Grantcharov, MD, PhD; Frank Rudzicz, PhD

What Is the Innovation?

Intraoperative adverse events are a common and important cause of surgical morbidity.[1,2] Strategies to reduce adverse events and mitigate their consequences have traditionally focused on surgical education, structured communication, and adverse event management. However, until now, little could be done to anticipate these events in the operating room. Advances in both data capture in the operating room and explainable artificial intelligence (XAI) techniques to process these data open the way for real-time clinical decision support tools that can help surgical teams anticipate, understand, and prevent intraoperative events.

In a systematic review, 64% of studies reported improvements in clinical decisions with automated decision support, especially if suggestions were provided at the same time as the task.[3] Machine learning (ML) techniques can provide this real-time decision support, estimating risk automatically from patient and intraoperative data. However, there has been hesitation to adopt ML techniques in health care[4] because these systems can have rare catastrophically incorrect predictions, and high accuracies can be achieved in unexpected ways, such as recognizing patterns in the manner of data recording, rather than in the content of the data themselves.

Explainable artificial intelligence is a collection of algorithms that improve on traditional ML techniques by providing the evidence behind predictions. For example, while a traditional ML algorithm in radiology may predict that an image contains evidence of cancer, an XAI system will indicate what and where that evidence is (eg, 3 cm, right lower lobe nodule).

In 2018, Lundberg et al[5] developed an XAI-based warning system called *Prescience* that predicts hypoxemia during surgical procedures up to 5 minutes before it occurs. This system monitors vital signs and provides the clinician with a risk score that updates in real time. It also continuously updates the clinician with reasons for its predictions, listing risk factors such as vital sign abnormalities and patient comorbidities. This can act like an additional vital sign, regularly updating information to warn the anesthetist in real time about upcoming risk.

With XAI, surgeons can receive similar warnings about upcoming intraoperative events to augment their clinical judgement, helping to avoid complications. Our team is currently working in surgical XAI to use laparoscopic videos to warn surgeons about upcoming bleeding events in the operating room and explain this risk in terms of patient and surgical factors. By anticipating and avoiding adverse events, surgical teams may be able to reduce operative times and improve outcomes for patients.

What Are the Key Advantages Over Existing Approaches?

Currently, risk prediction is done predominantly in the preoperative setting.[6] Intraoperatively, surgical teams rely almost exclusively on clinical judgement to predict patient risk and physiologic disturbance. However, many risky situations are unanticipated[2,5] and are managed reactively rather than preventively.[5]

Machine learning, by making fewer assumptions about parameter relationships, can use unstructured data sources such as text, audio, and video to provide more accurate predictions than traditional statistical techniques. However, the accuracy of ML often comes at the cost of explainability.

Predictions by ML systems have traditionally been difficult to interpret: the prediction itself is output to users, but the logic underlying that output is not. New work in XAI opens this black box. Both algorithm-specific and broadly applicable techniques have been developed to understand how different types of ML models make their predictions. For example, each risk factor can be consecutively removed to see how this absence affects the prediction. This essentially produces a relative risk calculation for each factor. By presenting both a real-time risk estimate and its underlying reasoning, XAI will allow surgeons to take advantage of complex ML-based prediction without losing the interpretability of logistical regression.

How Will This Affect Clinical Care?

Explainable artificial intelligence could be a powerful tool for intraoperative decision support, used in warning systems to help clinicians predict and avoid adverse events that may lead to complications. The overall risk estimates and contributing factors could be shown to the surgeon and updated with vital signs.

Using multiple data sources, including audio and video, XAI-based decision support may be able to provide these interpretable early warnings about intraoperative events such as bleeding and flag action points such as hypothermia, aberrant anatomy, or tool use that may be linked to the risk. With these factors provided to surgical teams in the operating room, XAI can significantly augment clinician judgement.

Clinicians' attention can then be drawn to warnings and they can decide whether to take action based on the XAI-provided explanations: to modify risk factors by, for instance, reviewing imaging, changing the surgical approach, or requesting different instruments.

Is There Evidence Supporting the Benefits of the Innovation?

Explainable artificial intelligence, as an emerging technology, has yet to be broadly implemented to provide decision support for surgeons. Prescience, a clinical decision support tool for anesthesiologists, is one example of how XAI might be used intraoperatively.

When provided with decision support from Prescience, anesthesiologists predicted hypoxemia with greater accuracy than by clinical judgement alone (area under the curve, 0.78 vs 0.66; $P < .001$). The authors estimate that anesthesiologists predict approximately 15% of intraoperative events and may be able to predict 30% when using Prescience. One-fifth of the time, hypoxemic risk was potentially related to medications provided intraoperatively–a highly modifiable risk factor.

Our team is currently working to developing applications of XAI in surgical practice: identifying intraoperative events like bleeding using patient, team, and surgeon factors. We aim to show that if surgeons are warned about bleeding risk and associated risk factors, bleeding events can potentially be avoided.

What Are the Barriers to Implementing This Innovation More Broadly?

Barriers to implementing XAI include data collection, technical development, and clinician trust. Explainable artificial intelligence requires a large volume of high-quality data for algorithm training. It can be particularly challenging to obtain this volume of data in the operating room. In response, we developed the OR Black Box recording platform,[7] now implemented internationally, that makes this high-quality, high-volume data collection feasible.

Explainable artificial intelligence is still in its relative infancy, and no standard approach has yet emerged. The relevance of technological development to surgeons may depend on whether researchers focus on explainablity or interpretability. Explainability of ML models describes the means of decision-making generally, allowing for high-level oversight and auditing of accuracy. The result can be mathematically complex and may not always be understandable to clinicians. By contrast, interpretability provides understandable reasons behind individual predictions that clinicians can use to make judgements. The broader discussion of what qualifies as an acceptable explanation is the subject of ongoing debate.

Clinicians have been hesitant to trust traditional ML systems because the internal workings are unclear: while the algorithm may generally have high accuracy, it can be difficult to evaluate if any individual prediction is correct. Clinicians may be more willing to trust an ML system when its logic is made transparent. Ideally, physicians will use XAI to augment their clinical judgement and evaluate whether the algorithm's logic makes sense within the clinical context prior to action. Further research on XAI-based decision support, demonstrating improved clinical judgement and reduced intraoperative events, would help build clinician trust and facilitate adoption of XAI.

In What Time Frame Will This Innovation Likely Be Applied Routinely?

While we expect XAI to become more widespread within the next 5 years, its adoption in surgery will depend on technological, cultural, and regulatory factors. Culturally, it will depend on trust, which will increase if XAI can generate demonstrably clear and actionable interpretations in, for example, intraoperative risk prediction. Similarly, clinical trials using XAI, with increasing regulatory acceptance of artificial intelligence in health care, will help spread the use of data-driven, realtime analytics in the operating room within the coming years.

ARTICLE INFORMATION

Author Affiliations: International Centre for Surgical Safety, Li Ka Shing Knowledge Institute of St Michael's Hospital, Toronto, Ontario, Canada (Gordon, Grantcharov, Rudzicz); Department of Surgery, University of Toronto, Toronto, Ontario, Canada (Gordon, Grantcharov); Department of Computer Science, University of Toronto, Toronto, Ontario, Canada (Rudzicz).

Corresponding Author: Teodor Grantcharov, MD, PhD, International Centre for Surgical Safety, Li Ka Shing Knowledge Institute, 209 Victoria St, Toronto, ON M5B 1T8 , Canada (grantcharovt@smh.ca).

Section Editor: Justin B. Dimick, MD, MPH.

Published Online: September 11, 2019. doi:10.1001/jamasurg.2019.2821

Conflict of Interest Disclosures: Dr Gordon reports grants from Canadian Institutes of Health Research during the conduct of the study. Dr Grantcharov reports IP ownership in SST Inc and grants from Medtronic Canada, Ethicon Canada, Intuitive Surgical, Olympus Surgical, and Baxter outside the submitted work. Dr Rudzicz reports personal fees from Surgical Safety Technologies outside the submitted work.

Submissions: Authors should contact Justin B. Dimick, MD, MPH, at jdimick@med.umich.edu if they wish to submit Surgical Innovation papers.

REFERENCES

1. Jung JJ, Jüni P, Lebovic G, Grantcharov T. First-year analysis of the operating room black box study [published online June 18, 2018]. *Ann Surg*. doi:10.1097/SLA.0000000000002863

2. Ehrenfeld JM, Funk LM, Van Schalkwyk J, Merry AF, Sandberg WS, Gawande A. The incidence of hypoxemia during surgery: evidence from two institutions. *Can J Anaesth*. 2010;57(10):888-897. doi:10.1007/s12630-010-9366-5

3. Garg AX, Adhikari NKJ, McDonald H, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA*. 2005;293(10):1223-1238. doi:10.1001/jama.293.10.1223

4. Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. *JAMA*. 2017;318(6):517-518. doi:10.1001/jama.2017.7797

5. Lundberg SM, Nair B, Vavilala MS, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat Biomed Eng*. 2018;2(10):749-760. doi:10.1038/s41551-018-0304-0

6. Moonesinghe SR, Mythen MG, Das P, Rowan KM, Grocott MP. Risk stratification tools for predicting morbidity and mortality in adult patients undergoing major surgery: qualitative systematic review. *Anesthesiology*. 2013;119(4):959-981. doi:10.1097/ALN.0b013e3182a4e94d

7. Goldenberg MG, Jung J, Grantcharov TP. Using data to enhance performance and improve quality and safety in surgery. *JAMA Surg*. 2017;152(10):972-973. doi:10.1001/jamasurg.2017.2888