VIEWPOINT

Does Intraoperative Artificial Intelligence Decision Support Pose Ethical Issues?

Charles E. Binkley, MD
Markkula Center for Applied Ethics at Santa Clara University, Santa Clara, California.

Brian P. Green, PhD
Markkula Center for Applied Ethics at Santa Clara University, Santa Clara, California.

Artificial intelligence (AI) holds great promise for many aspects of health care, including surgery. When used intraoperatively, AI clinical decision support systems (AI CDSSs) may reduce errors and increase clinical accuracy.[1] For example, AI CDSSs may assist surgeons in correctly identifying structures in the critical view of safety (CVS) during laparoscopic cholecystectomy (LC).[2] However, little consideration has been given to ethical issues that arise from the use of an intraoperative AI CDSS, aside from those of bias and privacy.[1] Adoption of this technology without recognizing and addressing these and other ethical issues risks long-term effects such as loss of public trust, overly restrictive regulation of AI systems, and rejection of the technology by patients and surgeons.

Explainability of Underlying Algorithm

A central theme in AI and especially CDSSs is the need for explainability.[3] Explainability is the ability of the system to convey to the surgeon the reasoning behind a particular decision: to explain not only how it arrives at its decision, but also why it arrived at that decision. For instance, take correct identification of the structures in the CVS during an LC. An AI CDSS, or those speaking for it, may be able to point to its algorithm and video training data to explain how it arrives at decisions in general. The AI CDSS may not, however, be able to provide the reasoning behind the specific recommendation it is making for an individual patient, in this case identification of the cystic duct, artery, and plate. This lack of transparency has profound ethical implications.

Conflicting Priorities of Developers and Surgeons

One issue is the degree to which the central moral requirement to do good and avoid harm is encoded into the system. While surgeons are bound to prioritize this inherent moral obligation, neither AI systems nor scientists and developers who created them have the same formal societal or professional obligation. One reason a surgeon wants to correctly identify the CVS during LC is their moral obligation to the patient. Developers of AI CDSSs may inadvertently, or intentionally, incorporate into the system certain outcomes and goals that conflict with the surgeon's primary ethical obligation.

For example, developers might prioritize increased efficiency and revenue. In the case of LC there may be an algorithmic trade-off between precisely identifying the CVS and completing the operation within a specified time. How much weight is given to each of the values may not be transparent. For instance, in prioritizing efficiency, the system may be programmed to decrease its precision if the CVS has not been identified within 30 minutes. Although multiple values can harmoniously coexist and be mutually beneficial, they can also be in conflict. It is important to identify these values, determine how much each influences the final decision, and resolve conflicts.

A second critical ethical issue arises from the shared agency between AI CDSSs and surgeons. When these systems provide intraoperative decision support, they necessarily augment, inform, or supplant surgeons' decision-making authority. Yet because of the opacity of the system's decision-making process, surgeons have no logical basis for evaluating the decision of AI CDSSs. If surgeons and AI CDSSs agree, surgeons can ethically rely on their own reasoning to explain the decision. Disagreement, however, places surgeons in an untenable ethical quandary.

For instance, during an LC, a surgeon may believe that a structure is the right hepatic duct, whereas the AI CDSS may identify it as the cystic duct. Because of the explainability problem, the surgeon must choose arbitrarily between their own identification of the structure and that of the machine. The surgeon cannot have a discussion with the algorithm as they would with a colleague. Automation bias may lead the surgeon to subordinate their own judgment to that of the algorithm simply because of inferred superior accuracy of the program. There may be external pressure from payers and health systems to prioritize the algorithm's decision. Or, as the accuracy of AI CDSS is further validated, these systems may be accepted as the legal standard of care. Thus, surgeons must weigh the unexplained decision of the system against their own reasoned decision to execute the clinical action.

Moral Responsibility Assigned to Surgeon

Regardless of conflict or agreement, the surgeon alone accepts full moral responsibility for the ultimate action. This puts the surgeon in an ethical quandary, even if the right action is ultimately taken. Imagine a situation where the machine identifies a duct as the cystic duct, and the surgeon believes that the same duct is the right hepatic duct. The surgeon disregards the system's recommendation and divides what they believe to be the cystic duct, only to find that the surgeon was mistaken and divided the incorrect structure. Alternatively, imagine that the surgeon accepts the system's decision and divides what the machine identified as the cystic duct only to find that the surgeon was correct and that the structure is the right hepatic duct. Moral responsibility in either case is assigned to the surgeon.

The surgeon then must explain to the patient that the decision to divide the structure was guided, correctly or incorrectly, by intraoperative support from an AI CDSS. This could lead to a significant loss of trust be-

Corresponding Author: Charles E. Binkley, MD, Santa Clara University, Markkula Center for Applied Ethics, 500 El Camino Real, Santa Clara, CA 95053 (cbinkley@scu.edu).

tween the patient and the surgeon. Should the patient have been informed before the surgeon used an intraoperative AI CDSS? One argument is that if the shared agency between the surgeon and the algorithm resembles that between an attending physician and first assistant, where the attending physician's judgment takes precedence, then the patient may not need to be informed of the participation of an AI CDSS. However, if—either because of internal bias or external pressure—a surgeon prioritizes the algorithm's judgment over their own, then the roles are reversed, and the algorithm is essentially the attending physician, and the surgeon is the first assistant. This reversal would diverge from the expectations of the patient, and the patient should be informed.[4]

Proposed Process for Ethical Implementation

Explainability and shared agency are but 2 of the most pressing ethical issues posed by intraoperative AI CDSSs. One proposed solution for the responsible implementation of AI in health care systems is to form teams of stakeholders, including patients and families, who, with input from ethicists, would guide the development, evaluation, and deployment of AI systems.[5] This is consistent with models used by other companies, such as Google, that partnered with ethics experts to implement their AI principles.[6] Like other companies, health care systems must institute mechanisms for auditing and reporting not only safety and accuracy concerns, but also ethical issues. This includes ensuring that the AI CDSS is being applied justly and that certain patients or groups are not being prioritized for or excluded from its use based on nonclinically important data such as insurance status, race, or sex.

Another way of addressing some of the ethical issues raised by AI is further research into how to make the systems' decisions understandable to surgeons and patients. Artificial intelligence systems should be subjected to rigorous evaluation through clinical trials to verify improved clinical outcomes.[7] In the absence of explainable decisions, this level of scrutiny would at least give both surgeons and patients some empirical basis to trust AI CDSSs. Recognizing these key ethical issues and setting a research agenda to address them will help ensure the fair, validated, and ethical adoption of intraoperative AI CDSSs.

ARTICLE INFORMATION

Published Online: June 16, 2021. doi:10.1001/jamasurg.2021.2055

Conflict of Interest Disclosures: None reported.

REFERENCES

1. Loftus TJ, Tighe PJ, Filiberto AC, et al. Artificial intelligence and surgical decision-making. *JAMA Surg*. 2020;155(2):148-158. doi:10.1001/jamasurg.2019.4917

2. Mascagni P, Fiorillo C, Urade T, et al. Formalizing video documentation of the critical view of safety in laparoscopic cholecystectomy: a step towards artificial intelligence assistance to improve surgical safety. *Surg Endosc*. 2020;34(6):2709-2714. doi:10.1007/s00464-019-07149-3

3. Gordon L, Grantcharov T, Rudzicz F. Explainable artificial intelligence for safe intraoperative decision support. *JAMA Surg*. 2019;154(11):1064-1065. doi:10.1001/jamasurg.2019.2821

4. Cohen I. Informed consent and artificial intelligence: what to tell the patient? *Georgetown Law J*. 2020;108(6):1425-1469. doi:10.2139/ssrn.3529576

5. Wiens J, Saria S, Sendak M, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med*. 2019;25(9):1337-1340. doi:10.1038/s41591-019-0548-6

6. Walker K. Google AI Principles updates, six months in. December 18, 2018. Accessed February 4, 2021. https://www.blog.google/technology/ai/google-ai-principles-updates-six-months

7. Topol EJ. Welcoming new guidelines for AI clinical research. *Nat Med*. 2020;26(9):1318-1320. doi:10.1038/s41591-020-1042-x