

Perspective

Open Access



Applications of machine learning in surgery: ethical considerations

Nikdokht Rashidian¹, Mohammed Abu Hilal²

¹Department of Human Structure and Repair, Ghent University, Ghent 9000, Belgium.

²Department of Surgery, Poliambulanza Foundation Hospital, Brescia 25100, Italy.

Correspondence to: Prof. Mohammed Abu Hilal, Department of Surgery, Poliambulanza Foundation Hospital, Via Bissolati, Brescia 25100, Italy. E-mail: abuhilal9@gmail.com

How to cite this article: Rashidian N, Abu Hilal M. Applications of machine learning in surgery: ethical considerations. *Art Int Surg* 2022;2:18-23. <https://dx.doi.org/10.20517/ais.2021.13>

Received: 15 Dec 2021 **First Decision:** 15 Feb 2022 **Revised:** 6 Mar 2022 **Accepted:** 14 Mar 2022 **Published:** 18 Mar 2022

Academic Editors: Andrew A. Gumbs, Zbigniew Nawrat **Copy Editor:** Xi-Jun Chen **Production Editor:** Xi-Jun Chen

INTRODUCTION

Artificial intelligence (AI) is defined as the study of algorithms to develop methodologies to simulate and extend human intelligence. Machine learning (ML) is a sub-type of AI that can learn, identify, predict, and make decisions by recognizing linear as well as novel nonlinear patterns within large datasets. Since “What Computers Can’t Do” was first published in 1972 by Dreyfus^[1], a critic philosopher of AI, the field of AI research has expanded rapidly, and the ability of a machine to display human-like capabilities has been proven in many domains. Accessibility to big data enables the cognitive computer to scan billions of unstructured data, extract the relevant information, and recognize intricate patterns with increasing confidence. In turn, trained algorithms can predict an outcome based upon its “experience” when presented with novel yet similar data. These ML algorithms have the capability of optimizing data via training, processing, and segmentation of imaging that can classify medical data based on complex patterns not detected by the human eye^[2,3].

As a high-stakes data-intensive process in surgery, there is a tremendous opportunity to benefit from the ML domain. Examples include ML models created to identify clinical diagnoses, support decision-making, improve surgical training, interpret medical images, and navigate preoperative planning and intraoperative guidance^[4,5]. Along with potential benefits to healthcare delivery, using these new technologies in the



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



surgical field raises several ethical concerns. Therefore, it is crucial to safeguard ethical measures in the development, deployment, and use of ML in surgery. While researchers must respect fundamental rights, applicable regulations, and core principles and values to ensure serving ethical purposes, reliability and technical robustness of each innovation must be critically appraised before introducing it to clinical practice^[6,7].

Ethics has been a matter of debate among philosophers for many centuries, and there are various well-known ethical principles. Immanuel Kant's categorical imperative is probably one of the best-known moral principles ever to be formulated: "act in such a way that you treat humanity...never merely as a means to an end, but always at the same time as an end". When it comes to medicine, Kantian ethics means that the physician has a duty to treat her/his patients as an end and not merely a means^[8]. This article aims to review the ethical considerations, the responsibility of surgeons, and some of the salient issues that arise when ML technology is used in the surgical field.

ETHICAL CONSIDERATIONS

Innovation has always been an essential part of surgical practice. While there are highly regulated and structured models for the assessment of innovative pharmaceuticals, the development of innovative surgical practice remains largely unregulated and variable^[9]. To fill the gap, the IDEAL framework and recommendations have been designed for evaluating the safety and efficacy of innovative surgical interventions and have been increasingly used since their introduction^[10,11]. The IDEAL framework describes five stages of innovation in surgery: idea, development, exploration, assessment, and long-term outcomes. It mainly focuses on appropriate methodology, data transparency, and rigorous outcome reports^[10]. For instance, transanal total mesorectal excision, as an innovative approach to distal rectal dissection for low rectal cancer, has been carefully adopted into clinical practice with a systematic approach using IDEAL framework principles^[12]. The complexity of AI and ML algorithms poses extra challenges for conducting a rigorous and comprehensive evaluation. The key aspect of ML is that the specific parameters for defining algorithms are not determined by human intelligence. Instead, they are learned from data using a general-purpose learning procedure to obtain the desired output in response to specific input data^[13]. The data-rich world of ML and the frequent absence of causal links in detected patterns make these new technologies different from previous innovations in surgery. Therefore, more and more recommendations and guidelines have been developed to address ethical AI^[14,15]. According to the European ethics guidelines for AI, a trustworthy AI should respect all applicable laws and regulations, consider ethical principles, and be robust, both from a technical perspective and considering the social environment^[14]. The medical ethics generally recommend that medico-moral decisions should be based on four philosophical principles formulated by two American philosophers, Beauchamp and Childress: (1) justice; (2) patient's autonomy; (3) beneficence; and (4) non-maleficence^[16,17]. To translate these ethical principles and values into practice in AI development, several requirements should be considered.

Data bias, fairness, and equity

The principle of justice in medical ethics deals with the equal distribution of resources within a society and avoiding discrimination of individuals^[16]. Considering the data-driven nature of ML, the machine's learning process can be compromised by the specific characteristics of training data. Systematic biases in societies are reflected in clinical data collections and have limited the inclusion of underrepresented minorities in databases^[18,19]. Developing ML algorithms based on biased datasets not only potentially perpetuates systematic inequities in societies but also can limit the performance of ML as a diagnostic and treatment tool due to the lack of generalizability^[20]. As with other surgical innovations, it is crucial that innovation in surgical ML is complemented by efforts to reduce the risk of bias. We must ensure that the benefits of these

new technologies are broadly shared across different countries and among people with diverse gender, race, ethnicity, culture, and socioeconomic status. A possible approach to prevent these already recognized biases is to expand surgeons' involvement and supervision on clinical data collection. Before providing data scientists with clinical data to develop ML algorithms, it is our responsibility to critically assess the quality of the data registries and ensure the inclusion of the underrepresented minorities into the dataset.

With the rapid development of data sciences and the expansion of data-driven clinical research, the publication of freely available datasets has become invaluable. Many public databases collect heterogeneous and multidimensional medical and surgical data on open platforms. The success of clinical research based on open access data is highly dependent on the quality of the data. As researchers, it is important to remember the value of making data publicly available in a standardized machine-readable format. Moreover, it is recommended to publish both negative and positive result data to reduce the data bias^[21].

To establish ground truth in surgical data sciences that involve supervised ML and computer vision, the need for correctly labeled data is a key challenge. The process of annotating surgical data such as surgical videos is a time-consuming task and requires skilled personnel, as wrong annotations can dramatically affect the performance of the trained model. In this regard, the Society of Gastrointestinal and Endoscopic Surgeons has provided a set of recommendations on a general framework to standardize the annotation of surgical video data. Standardizing the video annotation process may allow the combination or concatenation of heterogeneous datasets from different sources^[22].

Data privacy and security

Given that the ML algorithms must be trained with a sizeable amount of personal health information, data privacy and security are prominent challenges^[20]. Anonymization (the removal of private data) and pseudonymization (replacement of sensitive data by one or more artificial identifiers or pseudonyms allowing re-attribution using a look-up table) are currently the most common privacy preservation techniques for medical datasets^[23]. As a high-dimensional data domain in surgery, we handle different types of personal information, including patients' electronic files, medical images, and surgical videos. Anonymization of medical images and surgical videos requires removing all pertinent metadata entries (e.g., patient name, gender, date, camera type, *etc.*). Pseudonymization of such high-dimensional data poses additional difficulties, as it requires data deletion and data manipulation as well as the safekeeping of the look-up tables for reversing the process^[23]. While all the de-identification methods aim to reduce the risk of re-identification and minimize the loss of data utility, even the most complex and sophisticated of solutions never offer informative data with zero risk of re-identification^[24]. As a consequence, complex surgical datasets are prone to technical errors, resulting in re-identification or inaccuracy of data, and, therefore, must be processed more rigorously. Since appropriate data collection is a crucial step toward developing surgical ML algorithms, the trade-offs among data accuracy, interpretability, and privacy need to be further researched.

Technical robustness and safety

The principles of beneficence and non-maleficence imply that the application of ML technologies in surgery must be beneficial and non-maleficent for the individual patient^[16]. It is important to remember that some ML use cases have the potential to violate these principles. For example, a falsely optimistic prognosis tool, trained based on an unsuitable dataset for training an ML model, could trigger potentially inappropriate surgical intervention. When applied wisely, there is a tremendous opportunity for surgeons to use ML technologies pre-, intra-, and post-operatively to improve patient care^[25]. The optimal outcome for patients can only be expected with surgeons who can make informed decisions on when to apply an ML algorithm and how to interpret its results. Therefore, using ML for high-stakes decision-making in surgery must be

assessed rigorously and implemented in an evidence-based fashion similar to the introduction of other surgical innovations^[10].

While medical device manufacturers and healthcare institutions increasingly apply ML to innovate their products, specific regulatory requirements need to be met before these can be used in clinical practice. To this end, the European Medical Device Regulation and the Food and Drug Administration in the United States have imposed stringent requirements in filling regulatory gaps to develop ML-based medical devices and software tools^[26-28].

Autonomy, transparency, and explainability

The principles of patient autonomy refer to the rights of patients to make decisions about their medical care based on their own values and beliefs^[16]. Patients should be offered options along with sufficient relevant medical information and allowed to make voluntary choices about potentially life-changing healthcare interventions. While applying ML as a decision-aid tool, lack of explainability or so-called “black-box” design is a controversial topic with implications extending beyond AI’s technical properties. Explainability can be defined as a characteristic of an ML-driven tool allowing human intelligence to reconstruct why a certain algorithm came up with the presented prediction or decision^[29]. Concerning autonomy, the lack of explainable AI has implications for both patients and surgeons. Before signing informed consent for surgical intervention, patients must receive comprehensive and understandable information regarding the nature and risks of the procedure and alternative options. To date, disclosing the use of medical ML algorithms is not a mandatory requirement of a medical informed consent. The often-unexplainable nature of algorithms makes it potentially difficult to discuss with patients and explain the pros and cons. However, failure to disclose the use of an ML algorithm in the process of patient management may undermine patients’ autonomy, be considered a betrayal of trust, and jeopardize the doctor-patient relationship^[29]. Surgeons’ involvement early in the designing process of ML algorithms may help improve the interpretability of data-driven analyses. Despite advances in ML, one should realize that machines still cannot provide clinically meaningful context unless being interpreted by human physicians^[25]. Perhaps shared decision-making with patients or their surrogates is the best way to finally decide upon the consequences of using ML-based technologies.

CONCLUSION

The incorporation of ML into surgical practice holds promise for augmenting pre-, intra-, and post-operative patient care. To make the ML-based algorithms accountable and trustworthy, we need to regulate the development of ML algorithms according to the medical ethics and values of humanity. The development process must be aligned with data privacy and transparency requirements and focus on minimizing data bias. To achieve this, we must actively promote ethical education for ML-research stakeholders in the surgical community, enhance their awareness of ethics, and promote general practices towards robust and trustworthy ML algorithm development. Surgeons should partner with computer scientists to develop and assess ML applications in clinically meaningful ways. It remains the responsibility of surgeons to critically assess the quality of the data registries and ensure the inclusion of the underrepresented minorities into the datasets. Using ML for high-stakes decision-making in surgery must be assessed rigorously and implemented in an evidence-based fashion similar to introducing other surgical innovations.

DECLARATIONS

Authors’ contributions

Made substantial contributions to conception and design of the study and drafted and revised the manuscript: Rashidian N

Made substantial contributions to conception and design of the study and critically revised the manuscript:
Abu Hilal M

Availability of data and materials

Not applicable.

Financial support and sponsorship

None.

Conflicts of interest

Both authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2022.

REFERENCES

1. Dreyfus HL. What computers can't do: the limits of artificial intelligence. New York: Harper & Row; 1972.
2. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng* 2018;2:719-31. DOI PubMed
3. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542:115-8. DOI PubMed PMC
4. Madani A, Namazi B, Altieri MS, et al. Artificial intelligence for intraoperative guidance: using semantic segmentation to identify surgical anatomy during laparoscopic cholecystectomy. *Ann Surg* 2020. DOI PubMed PMC
5. Birkhoff DC, van Dalen ASHM, Schijven MP. A review on the current applications of artificial intelligence in the operating room. *Surg Innov* 2021;28:611-9. DOI PubMed PMC
6. Grote T, Berens P. On the ethics of algorithmic decision-making in healthcare. *J Med Ethics* 2020;46:205-11. DOI PubMed PMC
7. Beil M, Proft I, van Heerden D, Sviri S, van Heerden PV. Ethical considerations about artificial intelligence for prognostication in intensive care. *Intensive Care Med Exp* 2019;7:70. DOI PubMed PMC
8. Donaldson CM. Using Kantian ethics in medical ethics education. *Med Sci Educ* 2017;27:841-5. DOI
9. Barkun JS, Aronson JK, Feldman LS, Maddern GJ, Strasberg SM. Evaluation and stages of surgical innovations. *Lancet* 2009;374:1089-96. DOI PubMed
10. Mcculloch P, Altman DG, Campbell WB, et al. No surgical innovation without evaluation: the IDEAL recommendations. *Lancet* 2009;374:1105-12. DOI PubMed
11. Hirst A, Philippou Y, Blazeby J, et al. No surgical innovation without evaluation: evolution and further development of the IDEAL framework and recommendations. *Ann Surg* 2019;269:211-20. DOI PubMed
12. Robertson RL, Karimuddin A, Phang T, Raval M, Brown C. Transanal versus conventional total mesorectal excision for rectal cancer using the IDEAL framework for implementation. *BJS Open* 2021;5:zrab002. DOI PubMed PMC
13. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436-44. DOI PubMed
14. High-Level Expert Group on Artificial Intelligence. Ethics guidelines for trustworthy AI. Available from: <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1S> [Last accessed on 17 Mar 2022].
15. Winfield AF, Michael K, Pitt J, Evers V. Machine ethics: the design and governance of ethical AI and autonomous systems [Scanning the Issue]. *Proc IEEE* 2019;107:509-17. DOI
16. Kessel R. British Medical Association: 1988, Philosophy & Practice of Medical Ethics, B.M.A., London, 94 pp. plus appendices, etc., £9.50 (paper). *J Med Philos* 1989;14:709-10. DOI
17. Holm S. Principles of Biomedical Ethics, 5th edn. *J Med Ethics* 2002;28:332. DOI
18. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database Syst Rev* 2009;2009:MR000006. DOI PubMed PMC
19. Jüni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ* 2001;323:42-6. DOI PubMed PMC
20. Murphy K, Di Ruggiero E, Upshur R, et al. Artificial intelligence for good health: a scoping review of the ethics literature. *BMC Med Ethics* 2021;22:14. DOI PubMed PMC

21. Boulton G, Rawlins M, Vallance P, Walport M. Science as a public enterprise: the case for open data. *Lancet* 2011;377:1633-5. DOI PubMed
22. Meireles OR, Rosman G, Altieri MS, et al; SAGES Video Annotation for AI Working Groups. SAGES consensus recommendations on an annotation framework for surgical video. *Surg Endosc* 2021;35:4918-29. DOI PubMed
23. Kaissis GA, Makowski MR, Rückert D, Braren RF. Secure, privacy-preserving and federated machine learning in medical imaging. *Nat Mach Intell* 2020;2:305-11. DOI
24. Avraam D, Jones E, Burton P. A deterministic approach for protecting privacy in sensitive personal data. *BMC Med Inform Decis Mak* 2022;22:24. DOI PubMed PMC
25. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. *Ann Surg* 2018;268:70-6. DOI PubMed PMC
26. Regulation (EU) 2017/745 of the European Parliament and of the Council. Available from: <https://www.legislation.gov.uk/eur/2017/745/contents> [Last accessed on 17 Mar 2022].
27. Proposal for a Regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. Available from: <https://www.europeansources.info/record/proposal-for-a-regulation-laying-down-harmonised-rules-on-artificial-intelligence-artificial-intelligence-act-and-amending-certain-union-legislative-acts/> [Last accessed on 17 Mar 2022].
28. Statement from FDA Commissioner Scott Gottlieb, M.D. on steps toward a new, tailored review framework for artificial intelligence-based medical devices. Available from: <https://www.fda.gov/news-events/press-announcements/statement-fda-commissioner-scott-gottlieb-md-steps-toward-new-tailored-review-framework-artificial> [Last accessed on 17 Mar 2022].
29. Amann J, Blasimme A, Vayena E, Frey D, Madai VI; Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak* 2020;20:310. DOI PubMed PMC