Review Article

Challenges to AI use in anesthesia and healthcare: An anesthesiologist's perspective

Sakshi Gandotra [1,2]*, **Slomi Gupta**[3]

[1]Dept. of Anesthesia, St Michael's Hospital, Toronto, Ontario, Canada
[2]Dept. of Anesthesia and Pain Medicine, University of Toronto, Toronto, Canada
[3]Dept. of Cardiac Anaesthesiology, Sri Jayadeva Institute of Cardiovascular Sciences and Research, Bangalore, Karnataka, India

ARTICLE INFO

ABSTRACT

Anesthesiology is well positioned to benefit from applications of artificial intelligence on multiple elements such as monitoring the depth of anesthesia, control of anesthetic machine functions, ultrasound guidance for procedures and diagnosis, adverse event prediction, pain assessment and management, and optimising the operating room workflow. The ethical concerns can arise from multiple aspects of AI research and deployment such as the nature and source of the data, data collection methodologies, AI models design, output interpretation and inappropriate use. AI solution can have the unintended consequences like perpetuation of systematic biases and discrimination towards under-represented sections of society. There could be conflicts about data protection, intellectual property rights and economic gains. Also, the research must be transparent and solutions feasible. The clinician's role is ever changing in this landscape. We will discuss the broad ethical frameworks that are applicable to developing and using AI in medicine.

1. Introduction

The use of Artificial Intelligence in healthcare is a promising prospect and its applications in medicine continue to grow.

We can think about AI applications in medicine in areas of clinical practice, biomedical research, and translational medicine and research. Broad list of AI applications demonstrate how AI solutions can advance patient care and medical research:

1. Image analysis: AI applicability for image analysis is used in intraoperative echocardiography, vascular access, interventional procedures and focussed scans. AI solutions can reduce errors due to human fatigue.
2. Patient risk stratification (or population level primary prevention)
3. Risk of readmission (usually 30 days)
4. Medical research: AI can help with novel trial design, analytics and novel patient recruitment strategies. AI based chatbots can be used for trial screening.
5. Home videos for diagnosis of autism or learning disabilities.
6. Quality improvement: OR "blackbox" platform is already in use. AI systems can help quantify blood loss by analysing the photos of the sponges in OR or delivery suites.
7. Drug discovery: AI use in research has a major role in precision medicine. Developing targeted drugs based on phenotypes is being widely studies. Key features are gene identification, RNA expression, DNA mutations, and protein-protein interaction.
8. Medical education: AI competency is being considered in training programes. AI based simulation

---

* Corresponding author.
E-mail address: doctor.sakshigupta@gmail.com (S. Gandotra).

programs are being used for training and evaluation purposes.

AI can potentially provide high performance data-driven medicine, optimize patient care trajectories, suggest the right therapy for the right patient, improve diagnostics, and improve the process of clinical decision-making. Furthermore, AI can improve clinical reliability, reduce errors related to human fatigue, reducing over all healthcare costs and help physicians understand patients' values and goals.[1]

Artificial intelligence can have a major impact on anesthesiology; on multiple elements such as monitoring the depth of anesthesia, control of anesthetic machine functions, ultrasound for procedures and diagnosis, adverse event prediction, pain assessment and management, and optimising the operating room workflow.[2–9] AI use in anesthesia is still under investigation, but potential applications include automated recognition of anatomical structures during image-guided (US/CT/X-ray) regional anesthesia blocks or chronic pain procedures or for intravenous or arterial access; prediction of hemodynamic adverse events based on graph analysis, help in interpretation of ECHO, FAST images, recognition of the glottic opening and vocal cords during laryngoscopy and identification of endobronchial or oesophageal intubation; prediction of difficult airway using facial images. It is important that clinicians understand how AI solutions can be leveraged to deliver more efficient, safer and cost-effective care.

There are many ethical, legal, and regulatory factors that determine and constrain the financing and delivery of healthcare in ways that may not apply to other commercial products and services.

We will focus on the broad ethical issues that are applicable to developing and using AI, particularly in medicine. These include ethical frameworks that not only guide medicine, but also the development and application of AI.[10,11] We will examine these issues especially in relation to AI applications in anesthetic practice. Broadly speaking, we ask the questions: Whether AI tools help or harm patients or healthcare providers? Is the medical community ready to accept the AI solutions? And whether these tools perpetuate the social inequities?

2. Discussion

Autonomy, beneficence, non maleficence and justice are the four cornerstones of medical ethics.[12] The ethical concerns can arise from multiple aspects of AI research and deployment such as the nature and source of the data, data collection methodologies, AI models design, output interpretation and inappropriate use. AI solution can have the unintended consequences like perpetuation of systematic biases and discrimination towards under-represented sections of society. In addition, there can be broad social implications if the AI systems are intended to replace humans at healthcare related jobs.

2.1. Consent and potential harm

AI models use large datasets, mostly from electronic health records (EHRs) or digitised form of paper data, to produce mathematical models that can prospectively interpret complex multifactorial relationships. As an example, this data may be used to develop predictive models to anticipate adverse events and provide clinicians with sufficient lead times to intervene. This can help in predicting and taking actions in advance for the optimum outcome, for example, advising the perioperative team of the volume and type of blood product to order for a patient depending on the type, duration and complexity of the procedure.[13]

Patients have the right to be know about the use of AI in their healthcare, receive a full disclosure and be able to provide informed consent. Ethical concerns arise when patients are uncomfortable with AI assistance or are unaware of AI involvement in their care.

The benefit to clinicians is that of clinical decision support to supplement clinical judgement.[14]

2.2. Privacy and security of data

Hospitals, insurance companies and government programmes generate huge amounts of data. That patients' information, the quality of care received, outcomes and the associated costs are also vast valuable data. Much of these digital data is usable for analysis. A big general concern here is about the privacy and security of digital data. Healthcare data are sensitive and may have implications for patients' personal and professional life.[15] The patients may not consent to their data being used for algorithmic training. In addition, there are jurisdiction specific laws and regulations pertaining to the use and transmission of healthcare data.

2.3. Introduction of bias and inequities

Bias can occur during almost any stage from data collection to AI programing, data collection to model development. AI systems are prone to biases namely historical bias, representation bias, measurement bias, aggregation bias, evaluation bias and deployment bias.[16]

AI algorithms are trained on data. Historically, healthcare data is extremely male and extremely white. This historical bias inherent to real world data cannot be overcome, even by perfect sampling and randomisation. When such models are applied to the final use population, it creates representation bias. If the attributes differ across groups and lead to differential performance, it gives rise to measurement bias. For example, a model trained to predict myocardial infarction when used in women may fail to predict,

resulting in missed diagnosis. To mitigate aggregation bias, developers need to identify and understand the distinction between the groups and reasons why they are different from each other.

AI fairness and unintended bias alluded to maybe a larger problem in countries like India where health disparities exist based on patient demographics, gender, and geographical distribution.

Another example could be genomic data and precision medicine. These databases represent European ancestry and are missing population specific information particularly from Asian and African populations. Therefore, genomic test results for persons of non-European ancestry could be less accurate and therefore, will have limited applicability to Indian population.[17]

Deployment bias is the intersection of AI solution with the general applicability, that is, how society or the medical community uses the AI solution and derives its output. For example, algorithms for predicting whether an individual should receive a spinal cord stimulator for chronic pain may be biased towards those who are able to access and afford this expensive procedure.

AI models typically do not include social factors such as income, education, or social standing. These factors are well-known to have strong effects on hospital admissions and other long term health outcomes. Vulnerable populations such as the poor, people with disabilities, and people from far away geographical regions tend to incur disproportionately high healthcare costs. That could be because these populations are sicker when they do seek healthcare or because they have low access to medical care. AI models analyze insurance claims, health records to predict which of its patients are likely to incur the highest cost of care over the next year.

The deployment of this model is biased when used to predict healthcare need instead of health care cost. This deployment bias may also result in a rural or a low socio-economic group patient needing to be almost twice as sick as an urban patient to qualify for the same beneficial care program.[18] This could also affect the insurance premiums or simply the access to insurance. Models need to distinguish patients that incur high costs from patients that are sicker or have more medical needs so that a large vulnerable section is not denied the care they deserve.

The predictions based on theses systematic errors become a self-fulfilling prophecy and perpetuate the biases in further studies.

2.4. Algorithmic fairness

AI solutions must not propagate inequities. We must take a view of AI fairness in terms of justice that centers on the health and lives of people, not the outputs alone.[19] AI solutions that use datasets that are under representative of certain groups, may need specific training data to improve

decision-making and identify and reduce unfair results. Classification parity asks for equal predictive performance; Anti-classification requires the exclusion of any protected attributes in the outcome modeling. It is not always possible to exclude the "protected attributes" (like gender, ethnicity) as they may be essential for correct prediction, and reporting these is required for maintain transparency of AI models, as we will discuss later.

If an AI model is used to triage high-risk patients, poorly calibrated risk estimates will lead to false expectations. For instance, proposed models for triaging trauma patients for resuscitation can be in direct conflict with the fundamental rights of the individuals. If an AI model predicts poor outcome based on age or co-morbidities, should the clinician stop the resuscitative efforts?

In medical research, the available data rarely reflect the variable of interest. Hence, much of our research is by proxy. For instance, using race as a proxy for genetic ancestry to explain difference in disease prevalence or severity. Proxies are imperfect; this introduces the possibility of systematic error (thereby violating beneficence and nonmaleficence). When the possibility of systematic error is introduced, there's always the possibility of discrimination, and as discussed, leads to negative consequences for certain patient groups.

2.5. Lack of transparency

Machine learning models are constantly changing and updating themselves based on the data on which they are operating. Machine learning operates in a sort of black box, which lacks clinical explanations and accountability. It is often difficult to predict how AI is arriving at its decisions. Clinicians are reliant on technologies believing them to be safe and effective for use on their patients. There are very few standards or regulations for evaluating safety and effectiveness of AI based products. Should the clinicians and healthcare providers be held ethically and legally responsible for the decisions that may be informed by AI?

Most machine learning studies do not report the demographic breakdown of the training data used to develop and train their models. Some may argue this is to promote algorithmic fairness. It thus, becomes exceedingly difficult to evaluate the bias and fairness, and its applicability across populations. In an analysis, looking at reporting of demographics and population representativeness in AI research, race and ethnicity of data sample were not reported in 64% studies, gender and age were not reported in 75% of the studies, and social economic status was not reported in over 90% of the studies.[20] Detailed information on the data used to develop and train the model are necessary for unbiased and appropriate application of the AI solution.

The MINIMAR (Minimum information for medical AI reporting) has been proposed as a solution for transparent

reporting.[21]

2.6. Regulatory concerns and human oversight

Physicians have as fiduciary responsibilities and are ethically bound to serve the best interests of their patients.[22] They rely on a lot of other support systems to serve the best interests of their patients. For example, clinicians rely on researchers, scientists, and scientific societies to provide valid evidence to support and guide clinical practices. The regulatory systems to evaluate drugs and devices and the certification systems to evaluate hospitals and laboratories are also critical. As of now, there is no regulatory guidelines and clear directives on the responsibilities of physicians in the context of AI use to influence clinical decision making. The clinicians use their own judgement without directives. It will be necessary to modify and evolve the current ethical frameworks pertaining to patient confidentiality and the fiduciary duty clinicians have toward their patients as such systems become commonplace in clinical practice.[23]

Recently, some regulatory authorities have approved the clinical use of devices that incorporate machine learning algorithms for nerve identification during regional anesthesia. The ScanNav Anatomy Peripheral Nerve Block system, NerveBlox and NerveTrack are being considered for clinical practice.[24] There exists probability of medical harm associated with the use of AI solutions. This type of harm could come from medical error in diagnosis or treatment, or from undertreatment or overtreatment.

Maintaining human oversight in AI-assisted anesthesia is essential to avoid over-reliance on AI systems and to ensure that clinical judgment remains central to patient care. The use of AI predictions for depth of anesthesia or target controlled infusions are as good as the data they have, which may in turn be influenced by the patient condition and human factors. Every patient has their own set of challenges and thus anesthesia is highly individualised. Blanket use of AI models may be inappropriate in every scenario.

2.7. Competing interests and ownership of data

While clinicians and health care organizations are trying to serve the best interests of their patients, they have interests of their own and are subject to conflicts.[20] There are financial interests and intellectual interests. The healthcare system being an economic and commercial entity might prioritize cost effectiveness even though the physician is the one dispensing the care.

The specific health data belongs to the patient and its physical form belongs to the healthcare organisation. The questions then need to be addressed like who owns the data, the traceability of specific data elements from each individual patient into the "big" datasets, and where does patients' rights to privacy stand?

Data sets, involving image or biopsy interpretations, clinical interventions, may also reflect the significant intellectual contributions of clinicians. The subsequent work to curate this data and develop an algorithm certainly adds value to this raw data, but not all the value. How should we adjudicate claims regarding the value of the data, the value of each individual's data contribution to aggregate dataset, the value of the intellectual contribution from each provider and the pricing of the AI system? Should the clinicians who provided the intellectual contribution or the patients' whose data was used to train the models be compensated? There has also been ongoing patient activism for inclusion in recognition for specimen contribution to scientific advances.[24]

2.8. Challenges to role of physicians

The introduction of AI in anesthesia challenges traditional professional roles and responsibilities, raising questions about job displacement and professional autonomy. There is significant anxiety in the healthcare community to implement AI systems without proper validation and explainability. There is also concern about skill degradation.

AI systems risk becoming something followed either blindly or poorly. The output from an AI system may take on an unintended authority. Individuals who challenge an AI-based recommendation have frequently been required to provide significantly more robust evidence to refute the recommendation than the evidence which the AI recommendation was based on.[25] This has been observed in non healthcare contexts.

3. Conclusion

AI tools and clinical assessments have a common goal of improving patient outcomes. Nevertheless, balance must always be maintained. Not embracing the improvements that new technologies could provide and only favouring the traditional medical practices risks limiting the possible benefits of this technological revolution for our patients.

There are more challenges and opportunities for fair AI research. It is nearly impossible that one solution could address all challenges. It is vitally important that an AI model focuses on equity, meaning that each individual or group is given the same number of resources, attention, or outcomes. Many of these biases are systematic and we're often unaware they exist.

There are also many opportunities. The increasing number of papers discussing bias in AI solutions related to healthcare is evidence that we are forging a new path for data-driven, evidence-based health care. Transparency in reporting, deployment and use of AI solutions necessary for populations to trust medical system, particularly where AI is involved. The publisher of the New England Journal of Medicine is introducing a new journal, NEJM AI, to identify

and evaluate the applications of artificial intelligence to clinical medicine. JAMA has a series of videos, podcasts and series on AI. The key is to use AI in a way that does benefit all groups, which requires thoughtful evaluations and human interpretations.

4. Conflicts of Interest and Source of Funding

The authors have disclosed they do not have any conflicts of interest.

5. Author Contributions

Drafting of the manuscript: All authors. Revising the manuscript critically for important intellectual content: All authors. Final approval of the version to be published: All authors.

References

1. Álvarez Machancoses Ó, Fernández-Martínez JL. Using artificial intelligence methods to speed up drug discovery. *Expert Opin Drug Discov*. 2019;14(8):769–77.
2. Hashimoto DA. Artificial Intelligence in Anesthesiology. Current Techniques, Clinical Applications, and Limitations. *Anesthesiology*. 2020;132(2):379–94.
3. Shalbaf A, Saffar M, Sleigh JW, Shalbaf R. Monitoring the depth of anesthesia using a new adaptive neurofuzzy system. *IEEE J Biomed Health Inform*. 2018;22(3):671–7.
4. Zaouter C, Hemmerling TM, Lanchon R, Valoti E, Remy A, Leuillet S, et al. The feasibility of a completely automated total IV anesthesia drug delivery system for cardiac surgery. *Anesth Analg*. 2016;123(4):885–93.
5. Hatib F, Jian Z, Buddi S, Lee C, Settels J, Sibert K, et al. Machine-learning algorithm to predict hypotension based on high-fidelity arterial pressure waveform analysis. *Anesthesiology*. 2018;129(4):663–74.
6. Lin CS, Chang CC, Chiu JS, Lee YW, Lin JA, Mok MS, et al. Application of an artificial neural network to predict postinduction hypotension during general anesthesia. *Med Decis Making*. 2011;31(2):308–14.
7. Smistad E, Løvstakken L. Vessel Detection in Ultrasound Images Using Deep Convolutional Neural Networks. In: Carneiro G, editor. Deep Learning and Data Labeling for Medical Applications. DLMIA LABELS 2016 2016. Lecture Notes in Computer Science(10008). Cham: Springer; 2016.
8. Wingert T, Lee C, Cannesson M. Machine Learning, Deep Learning, and Closed Loop Devices-Anesthesia Delivery. *Anesthesiol Clin*. 2021;39(3):565–81.
9. Gkikas S, Tsiknakis M. Automatic assessment of pain based on deep learning methods: A systematic review. *Comput Methods Programs Biomed*. 2023;231:107365.
10. Riddick FA. The code of medical ethics of the American Medical Association. *Ochsner J*. 2003;5(2):6–10.
11. Lawler PA. The ACHE Code of Ethics: Its Role for the Profession. *J Continuing High Educ*. 2000;48(3):31–4.
12. Hope T, Dunn M. Medical Ethics: A Very Short Introduction. Oxford, England: Oxford University Press; 2018.
13. Braun M, Hummel P, Beck S, Dabrock P. Primer on an ethics of AI-based decision support systems in the clinic. *J Med Ethics*. 2021;47(12):e3.
14. Lin CS, Chiu JS, Hsieh MH, Mok MS, Li YC, Chiu HW. Predicting hypotensive episodes during spinal anesthesia with the application of artificial neural networks. *Comput Methods Programs Biomed*. 2008;92(2):193–7.
15. Smith H. Clinical AI: opacity, accountability, responsibility and liability. *AI Soc*. 2020;36:535–45.
16. Suresh H, Guttag J. A Framework for Unde A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cyclerstanding Sources of Harm throughout the Machine Learning Life Cycle. In: EAAMO '21: Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization. United States: Association for Computing Machinery; 2021. p. 1–9.
17. Kessler MD, Yerges-Armstrong L, Taub MA, Shetty AC, Maloney K, Jeng LJB, et al. Challenges and disparities in the application of personalized genomic medicine to populations with African ancestry. *Nat Commun*. 2016;7:12521. doi:10.1038/ncomms12521.
18. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447–53.
19. Corbett-Davies S, Gaebler JD, Nilforoshan H, Goel S, Shroff R, Goel S. The measure and mismeasure of fairness: A critical review of fair machine learning. *arXiv*. 2018;doi:10.48550/arXiv.1808.00023.
20. Bozkurt S, Cahan EM, Seneviratne MG, Sun R, Lossio-Ventura JA, Ioannidis JPA, et al. Reporting of demographic data and representativeness in machine learning models using electronic health records. *J Am Med Inform Assoc*. 2020;27(12):1878–84.
21. Hernandez-Boussard T, Bozkurt S, Ioannidis JPA, Shah NH. MINIMAR (MINimum Information for Medical AI Reporting): Developing reporting standards for artificial intelligence in health care. *J Am Med Inform Assoc*. 2020;27(12):2011–5.
22. Smith H, Downer J, Ives J. Clinicians and AI use: where is the professional guidance? *J Med Ethics*. 2023;doi:10.1136/jme-2022-108831.
23. Char DS, Shah NH, Magnus D. Implementing Machine Learning in Health Care-Addressing Ethical Challenges. *N Engl J Med*. 2018;378:981–3.
24. Bowness JS, El-Boghdadly K, Woodworth G, Noble J, Higham H, Lauren DBS, et al. Exploring the utility of assistive artificial intelligence for ultrasound scanning in regional anesthesia. *Reg Anesth Pain Med*. 2022;47(6):375–9.
25. O'Neil C. The Ivory Tower Can't Keep Ignoring Tech; 2017. Available from: https://www.nytimes.com/2017/11/14/opinion/academia-tech-algorithms.html.

Author biography

Sakshi Gandotra, Clinical Fellow https://orcid.org/0000-0001-6213-3528

Slomi Gupta, Assistant Professor