**Guideline**

# White paper: ethics and trustworthiness of artificial intelligence in clinical surgery

Giulia Capelli[1,2,#], Daunia Verdi[3,#], Isabella Frigerio[4], Niki Rashidian[5], Antonella Ficorilli[6], S. Vincent Grasso[7], Darya Majidi[8], Andrew A Gumbs[9], Gaya Spolverato[1], *Artificial Intelligence Surgery* Editorial Board Study Group on Ethics

[1]Department of Surgical, Oncological and Gastroenterological Sciences (DiSCOG), University of Padova, Padova 35122, Italy.
[2]Department of Surgery, ASST Bergamo Est, Seriate (Bergamo) 24068, Italy.
[3]Department of Surgery, Mirano Hospital, Mirano (Venezia) 30035, Italy.
[4]Department of Hepato-Pancreato-Biliary Surgery, Pederzoli Hospital, Peschiera 37019, Italy.
[5]Department of General, HPB Surgery & Liver Transplantation, Ghent University Hospital, Gent 109000, Belgium. [6]Department of Cardiac, Thoracic, Vascular Sciences and Public Health, University of Padua, Padua 35128, Italy.
[7]Research Professor, Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM 87131, USA.
[8]President "Associazione Donne 4.0", Livorno 57123, Italy.
[9]Departement de Chirurgie, Centre Hospitalier Intercommunal de Poissy/Saint-Germain-en-Laye, Poissy 78300, France.
[#]Authors contributed equally.

**Correspondence to:** Prof. Gaya Spolverato, Department of Surgical, Oncological and Gastroenterological Sciences (DiSCOG), University of Padua, Via Giustiniani 235128, Padova 35122, Italy. E-mail: gaya.spolverato@unipd.it

## Abstract

This white paper documents the consensus opinion of the *Artificial Intelligence Surgery* (*AIS*) task force on Artificial Intelligence (AI) Ethics and the *AIS* Editorial Board Study Group on Ethics on the ethical considerations and current trustworthiness of artificial intelligence and autonomous actions in surgery. The ethics were divided into 6 topics defined by the Task Force: Reliability of robotic and AI systems; Respect for privacy and sensitive data; Use of complete and representative (i.e., unbiased) data; Transparencies and uncertainties in AI; Fairness: are we exacerbating inequalities in access to healthcare?; Technology as an equalizer in surgical education. Task Force members were asked to research a topic, draft a section, and come up with several potential consensus statements. These were voted on by members of the Task Force and the Study Group, and all proposals that received > 75 % agreement were adopted and included in the White Paper.

## INTRODUCTION

There is a growing concern in the surgical community about the ethical implications of artificial intelligence (AI) in clinical surgery[1].

While AI has been shown to improve both surgical training and the management of surgical patients, questions have been raised about the reliability of robotic and AI systems, leading to various uncertainties and ultimately limiting the applicability of these tools in clinical practice. Moreover, the nowadays ease of gathering data from different unregulated and not opportunely anonymized data sources causes serious privacy issues, while concerns have been raised about potential bias due to the use of unrepresentative data. Finally, AI should help to address inequalities in access to healthcare by improving patient access to treatment and equalizing surgical training through telementoring, virtual reality (VR), and simulation platforms.

Such issues have been repeatedly raised in many different areas of application of AI[2], to the point of becoming an object of study in philosophy and sociology[3].

Based on these premises, this consensus project aimed to clarify the ethical implications of AI in clinical surgery by summarizing current evidence and collecting expert opinions on the topic.

## METHODS

### Definitions

Surgery is defined as all interventional medical disciplines, including surgery, endoscopy, and interventional fields such as interventional radiology and cardiology. AI in surgery is defined as the use of ad-hoc algorithms involved in learning procedures for prediction and/or classification in the pre-operative and post-operative phases. *Artificial Intelligence Surgery* (*AIS*) is defined as the utilization of such algorithms during an actual intervention and covers autonomous actions during a procedure.

Levels of autonomy in surgery have been defined in detail in a previous paper[4].

### Consensus questionnaire

The methods for this study were partially described in a previous paper[4].

A 24-item electronic questionnaire (reported in the Supplementary Materials) was sent to the members of the Editorial Board (EB) of *Artificial Intelligence Surgery* (*AIS*, www.aisjournal.net) to reach a consensus on the following issues: (a) reliability of robotic and AI systems; (b) respect for privacy and use of sensitive data; (c) use of complete and representative (i.e., unbiased) data; (d) need for transparency and current uncertainties in AI; (e) fairness and access to healthcare; (f) and finally, the role of technology as an equalizer in surgical training. EB members who expressed an interest in the topic and completed the survey became part of the *Artificial Intelligence Surgery* Editorial Board Study Group on Ethics.

The questions and statements were drafted by a Task Force on AI Ethics composed of the Editor-in-Chief of *AIS*, members of the Associate Editorial Board, members of Women in Surgery (WIS) Italia, a surgeon and expert in machine learning (ML), two AI pioneers and entrepreneurs, and a bioethicist and philosophy graduate.

The members of the panel were required to express either their approval (i.e., "Yes, I agree") or non-approval (i.e., "No, I don't agree") of each proposed statement. Statements from the survey were accepted if they received more than 75 % consensus from the *Artificial Intelligence Surgery* Editorial Board Study Group on Ethics and the Task Force on AI Ethics.

### Demographics

The demographics of the 60 *AIS* EB members have been described previously[4]. A total of 45 members participated in this survey, 2 members of WIS, 1 surgeon and 1 entrepreneur, both with experience in AI. There were 28 (62.2 %) males and 17 (37.8 %) females, of whom 39 (86.7 %) were surgeons and 6 (13.3 %) were non-surgeons with Ph.D. level experience in either mechanical engineering, statistics, robotics, ethics, or AI. Although the composition of the panel is fairly representative of the different realities involved in the development and use of AI in clinical surgery, the unbalance concerning their gender and skills should be taken into account and could be considered a possible limitation of our study.

## RESULTS

### Reliability of robotic and AI systems

The impact of AI and AI-enabled robotics in surgery is growing rapidly. As the role of AI in surgery becomes more prominent with the emergence of autonomous and intelligent robots, there is a need to build trust in these technologies by their human counterparts[1]. Beneficence and non-maleficence for the individual patient are considered to be important principles in medical ethics. In innovation, the principle of beneficence refers to developing technologies that can "do good" and improve patient care, while the principle of non-maleficence states "do no harm"[5,6]. As the technology develops, the use of AI-powered robots in clinical practice must be rigorously assessed to reduce the risk to individual patients. To this end, the European Medical Device Directive (EU MDD) and the Food and Drug Administration (FDA) in the United States have established strict regulations for AI-powered medical devices[7-9]. It is reasonable to assume that the more complex the function of the robot, the greater the risk to the patient in the event of a malfunction, and therefore the more stringent the regulation of these devices. However, surgical devices, ranging from complete robotic surgical systems to handheld surgical devices such as energy devices or power staplers, are currently classified as Class I, representing low-risk devices. It seems that a new class of medical devices needs to be defined to cover AI-powered robots with higher levels of autonomy[4]. Then, to assess the reliability of the new AI-powered robots, verification and validation methods are needed to assess and document the reliability and reproducibility of their performance. Finally, implementation into daily clinical practice should follow evidence-based medicine, similar to the introduction of other surgical innovations[10].

### Respect for privacy and sensitive data

The healthcare vertical continues to evolve from its original analog working environment to an increasingly digital one. The inherent logistical complexities associated with the transfer of analog medical data have been virtually eliminated by the ubiquitous availability of high-speed digital networks. These networks enable access to and transfer of medical data in one-to-one, one-to-many and many-to-many relationships, creating new complexities associated with traditional auditing practices.

The volume of healthcare data generated each year is staggering, representing approximately 30 % of the world's data volume, and its compound annual growth rate (CAGR) will reach 36 % by 2025. In 2021, the amount of data generated globally will be approximately 79 zettabytes and is expected to double by 2025. Interestingly, 90 % of this data is currently replicated rather than new[11].

Whether digital health data is new or legacy, data in a digital format presents new challenges in terms of policies and procedures related to its security and privacy. Data security refers specifically to policies and procedures that focus on protecting information from unauthorized access, including breaches, cyber-attacks, and other related violations. Data privacy refers specifically to policies and procedures that govern the use of data, particularly personal health information (PHI), and ensure that it is used in accordance with owner preferences and prohibitions.

Recognizing the unique nature of healthcare data and how best to ensure its security and privacy, government bodies within the global community have formalized efforts to address these issues. As of 2018, the General Data Protection Regulation (GDPR) imposes its data security and privacy standards on any organization involved in the collection and use of the data of individuals residing within the EU[12].

In the US, legislation on the privacy and security of patients' private health information was enacted in the Health Insurance Portability and Accountability Act of 1996 (HIPAA). It was updated in 2013 by the HIPAA Omnibus Rule, which introduced new requirements mandated by the Health Information Technology for Economic and Clinical Health (HITECH) Act. Additional changes have been proposed through the 2020 CARES Act, which aligned 42 CFR Part 2 to address substance abuse data[13].

In the pre-AI environment that healthcare experienced as it transitioned from an entirely analog to an increasingly digital delivery model, the privacy and security challenges were complex but manageable. Data deidentification of direct identifiers (e.g., name and biometrics) and indirect identifiers (e.g., race, religion, and age) was accomplished using a variety of tools and techniques, including the use of algorithms. The challenge of data deidentification has been, and remains, the risk of re-identification, which increases with the use of both direct and indirect identifiers.

Although there has been considerable experimentation with the use of AI in healthcare over the last 50 years, the last 10 years have seen an explosion in the development of software for healthcare applications, fueled by the growth and availability of extremely powerful computing resources and other related services provided by various cloud providers such as Microsoft, Google, and Amazon. The promise of AI outcomes in healthcare, including early disease diagnosis, tailored cancer treatment protocols, disease risk factor detection, mental health insights, and many others, requires access to vast pools of objective and subjective data residing in a plethora of information systems not limited to electronic medical records (EMRs).

The benefits of using AI in healthcare bring with them new risks to data privacy and security, including the ability to re-identify data that was not available in the pre-AI era. As a result, a new era of patient risk has emerged concerning the use of their data beyond what they originally consented to. In fact, regulatory and compliance efforts related to data privacy with audits are severely hampered by the lack of full transparency across the entire ecosystem related to PHI use. Algorithms used to aggregate and analyze vast amounts of deidentified PHI, with data sources residing in information systems not traditionally used in clinical settings (media, telecoms), pose profound challenges to regulators.

In the future, the promise of improved healthcare for the global community through the use of AI ecosystem assets will be recognized and will continue to result in increased outcomes. The use of AI requires access to a wide range of data residing in disparate information systems and involves data transformation and analysis at many levels. This complexity creates significant challenges in ensuring not only the privacy of patient data but also the governance of its intended use. It is imperative that healthcare organizations, vendors, regulators, policymakers, patient advocates, and all other relevant third parties work

together to further develop policies and procedures that mitigate the risk of data misuse while enabling innovation driven by advances in AI use.

**Use of complete and representative (i.e., bias-free) data**

AI is a technology designed by humans to replicate, augment or replace human intelligence[14]. These tools typically rely on large amounts of data to generate insights. Poorly designed projects based on biased data can have unintended, potentially harmful consequences. Moreover, the rapid advancement of algorithmic systems means that in some cases, it is not clear how AI reached its conclusions, so we are essentially relying on systems that we cannot explain to make decisions that could affect society. The word "biased" means unfairly prejudiced for or against someone or something[15], usually reflecting widespread societal biases[16]. The EU Charter of Fundamental Rights lists the following characteristics as morally relevant in cases of discrimination: sex, race, color, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age and sexual orientation[17].

There are two types of bias in AI:

(a) Data bias: AI algorithms are trained on a set of data that is used to inform or build the algorithm. If the algorithm makes inaccurate or unethical decisions, it may mean that there was not enough data to train the model, or that the learning reinforcement was not appropriate for the desired outcome. One of the more damaging risks of AI is the introduction of bias into decision-making algorithms[18]. AI systems learn from the dataset on which they have been trained, and depending on how that dataset has been assembled, it may reflect assumptions or biases[19].

(b) Human bias: This is where our assumptions and norms as a society cause us to have blind spots or certain expectations in our thinking. We tend to seek out or interpret evidence in a way that supports our own beliefs. Given access to the same set of data and information, different people can come to very different conclusions. The risk of AI introducing and amplifying existing human biases into models is significant[20], but the benefits include a fairer decision-making scenario when the algorithms are trained to avoid bias[21].

How can we fight back against AI bias? We can start by reading widely, engaging with progressive ideas, and sharing helpful articles and research that can be used to educate others. By understanding the sources of algorithmic and data bias, we can diversify our data sets. By being more aware of the societal biases we live with every day, we can mitigate them in our work.

**Transparencies and uncertainties in AI**

The development, deployment, and use of reliable AI in surgery require a transparent approach to both the opacity of AI algorithms and the uncertainties associated with using these types of techniques as decision-support tools in clinical settings.

As the computational complexity of AI algorithms and predictions increases, it becomes difficult for humans to understand and explain why a model has produced a particular output or arrived at a particular clinical decision. This is known as the "black box effect". The opacity of the process raises ethical issues[22-24].

First, the outcomes/decisions generated by AI algorithms cannot be properly reviewed by human experts, including surgeons, to determine whether or not there is bias.

Second, surgeons cannot adequately use their experience and clinical judgment to evaluate the results and recommendations provided by an AI method.

Third, surgeons may not be able to fully inform patients about how certain results or recommendations were derived. As a result, patient autonomy may be undermined and a new AI-based paternalistic medicine may become widespread.

Given the enormous potential of AI systems to improve pre-operative and post-operative care[25], it is critical that all these issues are considered and managed, especially where there is a significant impact on patients' lives.

A key element is to make every effort to improve our understanding of how AI algorithms work. In this regard, the early involvement of surgeons in the design of AI algorithms could help to improve the accountability and interpretability of big data analyses. Furthermore, their suggestions could ensure transparent documentation of how results and recommendations were derived based on the information available at any given time. Another key aspect relates to the fact that AI systems have some limitations, such as questions about the type and accuracy of available data. In addition, AI systems are poor at providing the appropriate clinical context in which to interpret the data. Therefore, surgeons must critically evaluate the predictions generated by AI and interpret the data in a clinically meaningful way[26].

Uncertainties remain! Indeed, working with AI systems generates uncertainty because it increases the complexity of the process. Therefore, even more than in the past, it is necessary to be transparent about uncertainties and to share "uncertain information" with patients. Only in this way will it be possible to continue to meet the key principles of the ethical framework outlined in the 1970s, which assesses biomedical ethical challenges: autonomy, beneficence, non-maleficence, and justice[6].

**Fairness: are we exacerbating inequities in access to healthcare?**
According to the Ethical Guidelines for Trustworthy AI, published by the European Commission's High-Level Expert Group on AI in 2019[1], AI should be beneficial for all people, preserving and promoting democratic processes concerning the plurality of values. To achieve this, AI should facilitate inclusivity and diversity of values throughout its lifecycle by promoting equal access to services through inclusive design processes and equal treatment. AI systems should be user-centered and conceived in such a way that all people can use AI products, regardless of their age, gender, ability, economic power, or other characteristics[27-29].

When applied to healthcare systems and surgical services, the value of fairness encompasses manufacturers, healthcare organizations, clinicians, and patients.

AI can address inequities or perpetuate them if not used properly, such as when relying on biased data[30]. For example, the use of deep neural network algorithms can enhance the bias of surgical data that suffers from historical inequity, incompleteness, and poor governance models. Biased data can later be translated into biased information, affecting the decision-making process and clinical outcomes. As the Council of Europe reminds us, non-discrimination and data protection laws have been created to mitigate the risks of AI-driven discrimination and enable the dynamics of shared decision-making, co-production, and tailored approaches, but AI itself may create new types of bias that escape these laws. Extracting data from different backgrounds, cultures and disciplines can ensure diversity and should be encouraged to create more inclusive data management processes. Trustworthiness of AI systems is an essential requirement that can be

achieved by involving stakeholders who may be directly or indirectly affected by the system[31], such as surgeons or patients, and an active role for them should be advocated to improve the whole process of implementing AI systems.

Economic factors limit the accessibility of AI services and thus affect the equitable distribution of AI-based systems in developing countries, where inequalities in access to clinical care and quality of services are real: solutions are needed to avoid widening the existing gap in data access, surgical training, and use of technology. However, the integration of AI into healthcare can offer a wide range of benefits to diverse populations who would otherwise not have access to highly skilled and specialized care[32]. AI systems designed for high-income countries may not meet the needs of developing regions, making dedicated funding from industry and governments a necessary investment. Meanwhile, cloud computing and a simple smartphone represent an already existing way for such technology to reach the most remote parts of the globe[33].

An interesting issue is raised by Cobianchi et al. manufacturers and AI industries that improve their products and revenues by using data generated during surgical procedures, should they have any obligation to patients[34]? Or, additionally, do patients or the institution paying for the surgery have a right to demand financial compensation for providing useful data to improve technology? These are questions that remain to be answered.

Given that global health industries operate for profit, they may be able to interfere with the national health policies of countries with limited resources and health needs. It is crucial to ensure the fairness of AI by defining the boundaries of its operation and involvement by governments and international institutions.

### Technology as an equalizer in surgical training

Surgical training is still burdened by the negative effects of discrimination based on either sex or race/ethnicity[35,36]. Gender bias and stereotypes in surgical training have been linked to anxiety and burnout symptoms, ultimately leading to poorer performance and limiting career engagement and opportunities for female surgeons[37]. Similar detrimental effects have been documented among residents from racial/ethnic minorities[36].

Cognitive bias has also been studied in surgical residents during self-assessment of their skills and performance, with several disadvantages identified in their practice and surgical and/or academic careers[38,39]. Less accuracy in self-assessment has been described in female residents and has been associated with less improvement in surgical skills during residency. In addition to these difficulties, students and residents training in low- and middle-income countries face several additional challenges, both in terms of opportunities and quality of training, which further exacerbate inequalities in surgical training around the world[40].

In recent years, AI has been proposed as a means of optimizing surgical training through various technologies, including telementoring, virtual reality, and simulation platforms[41,42]. Such technologies allow (a) the possibility of objective assessment of surgical performance; (b) the possibility of access to training even in a rural/low-income situation. Several technologies have proven to be effective in providing an objective assessment of psychomotor skills in both open and laparoscopic surgery[43]. Moreover, ML technologies have been proposed as a means to improve intraoperative judgment and decision-making[44]. The ability to extract a large amount of data about an individual's technical performance during a simulated task could be used to provide the trainee with personalized training based on their particular struggles and

weaknesses, ultimately eliminating the negative effects of both implicit and explicit bias.

Remote telepresence proved effective in training surgeons residing in rural areas while conducting complex laparoscopic procedures[42], and could be offered to trainees in low- and middle-income countries as a means of providing access to world-class training while avoiding the need for costly overseas training. In addition, surgical simulation has been shown to significantly improve trainees' surgical skills when used in locally taught programs in low-income countries[45,46]. The ability to train remotely and have access to surgical simulation platforms could also offer significant benefits to pregnant residents, allowing them to manage their schedules more flexibly and ultimately have a positive impact on both fetal and maternal health[47].

While AI may offer unlimited opportunities to surgical trainees around the world, some limitations should be acknowledged. Most available studies on the objective assessment of surgical skills are based on limited data. In addition, most technologies do not provide information on how to improve psychomotor skills, limiting their usefulness in actual training[47,48]. Finally, some may argue that surgeons need a whole range of psychological and emotional attributes that cannot be measured or objectified, but which contribute significantly to patient outcomes (e.g., empathy and/or the ability to perform well in stressful situations). When discussing telementoring and simulation, a word of caution is needed regarding low-income countries where budgets are limited and the need for surgical care is still unmet[49]. In this environment, the investment of money in technologies that do not provide care for patients may be an unethical choice.

### Consensus statements

The statements listed below were approved by the majority of panelists, achieving a consensus rate of > 75 %. Twenty-two (91.7 %) of 24 statements were endorsed.

1. In the development phase, verification and validation methods must be in place to assess and document the reliability and reproducibility of AI-powered robots.

2. The introduction of AI-powered robots into clinical practice needs to be rigorously evaluated in an evidence-based manner, similar to the introduction of other surgical innovations.

3. A new and higher class of medical device risk needs to be defined to cover AI-powered robots/devices with high levels of autonomy (i.e., level 4 and 5 autonomous surgery).

4. A new and lower class of medical device risk needs to be defined to cover AI-powered robots/devices with low levels of autonomy (i.e., autonomous surgery levels 2 and 3) to encourage the development of intelligent surgical technology.

5. Healthcare data bloom, fueled by excessive replication of legacy data, needs to be addressed as an unnecessary risk factor for privacy breaches of patient health information (PHI).

6. The role of AI in deidentifying PHI with re-identification of PHI needs expanded transparency.

7. The public needs to better understand the trade-off between putting their PHI at risk and tangible benefits to their overall health outcomes.

8. An ethical AI system must be accountable and have a positive purpose.

9. An inclusive AI system is unbiased and works equally well across all spectrums of society.

10. An AI system that uses data responsibly respects privacy rights.

11. In the future, we need to make every effort to reduce the current opacity in AI systems to increase surgeon and patient confidence in situations where AI is used in the decision-making process.

12. In the future, we need to make every effort to reduce the current opacity in AI systems so that surgeons can make an informed decision about whether or not to rely on the system's recommendations.

13. In the future, we need to disclose the use of opaque AI systems when asking for informed consent, so that patients can exercise full autonomy.

14. In the future, we need to share uncertainties and limitations with patients about an AI-assisted care pathway to manage potential adverse outcomes through a transparent approach.

15. In the future, we will need to share the decision-making process with patients to share responsibility for choosing one care pathway over another.

16. AI systems should be based on unbiased data collected through transparent and non-discriminatory processes.

17. AI systems should be affordable and equally accessible around the world.

18. Industry and AI developers should consider funding AI systems tailored to the needs of developing countries.

19. AI can provide some valuable help in addressing the bias still documented in surgery, which particularly affects women and racial/ethnic minorities.

20. The ability to objectively assess the psychomotor performance of surgical residents can provide them with a bias-free evaluation, ultimately enhancing their training and career opportunities and improving their mental health.

21. Remote telepresence and surgical simulation can be a valuable tool to provide high-level training opportunities to trainees residing in low- and middle-income countries; however, the decision to invest significant budgets in training technologies should be weighed against the unmet medical and surgical care needs of low-income country populations.

22. Caution should be exercised in considering AI as the sole means of assessing trainee performance, as it does not take into account some relevant psychological characteristics that may ultimately affect patient outcomes, including empathy and the ability to work under stressful conditions.

## CONCLUSIONS

We are at the dawn of a new technological revolution that has the potential to dramatically change the way healthcare is delivered. Unlike previous periods of innovation in medicine, many of the innovations are being developed by non-medical people because the number of doctors with expertise in computer science and AI is limited. One of the reasons we started this journal is to help organize the physicians who are interested in this topic and to disseminate the concepts in a form that all modern surgeons can understand and digest. As economic inequalities continue to grow globally and within countries, the potential for AI to fill gaps and improve access to healthcare must be tempered by the potential for these technologies to be misused and ultimately exacerbate inequalities in healthcare delivery and access. The sheer volume of data that can now be collected has probably become too overwhelming for surgical, interventional, and endoscopic departments alone to generate clean and accurate data without the appointment of a dedicated data scientist.

The complexity of ML algorithms has created such a "black box" effect that it could be argued that all surgical departments already need a dedicated member of their team to accurately and safely integrate AI into patient care. As the cost of these innovations continues to rise, universities, healthcare systems, and surgical societies may not have the resources to continue to control and monitor the introduction of more and more AI and autonomous robots. This is highlighted by the observation that the Fundamentals of Laparoscopic Surgery (FLS) is run by the Society of American Gastrointestinal and Endoscopic Surgeons, but the Fundamentals of Robotic Surgery (FRS) is run primarily by Intuitive Surgical (Intuitive Surgical, Sunnyvale, CA, USA), and no similar certification is run by a surgical society. With the explosion of new telemanipulators on the market, it should be surgeons who dictate training, certification, and maintenance of competency in artificial intelligence surgery, not industry alone. As the price of these innovations continues to rise, governments may need to play a more active role in overseeing and validating the development of artificial intelligence in surgery.

## DECLARATIONS

**Financial support and sponsorship**
None.

**Conflicts of interest**
All 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) 2023.

## REFERENCES

1.   Ethics guidelines for trustworthy AI. Available from: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai. [Last accessed on 15 May 2023]. DOI
2.   Top 9 ethical issues in artificial intelligence. Available from: https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/. [Last accessed on 15 May 2023].
3.   Bostrom N. Ethical issues in advanced artificial intelligence. In: Routledge, ed. Machine Ethics and Robot Ethics 2017. Available from: https://www.taylorfrancis.com/chapters/edit/10.4324/9781003074991-7/ethical-issues-advanced-artificial-intelligence-nick-bostrom. [Last accessed on 15 May 2023].
4.   Gumbs AA, Alexander F, Karcz K, et al. White paper: definitions of artificial intelligence and autonomous actions in clinical surgery. *Art Int Surg* 2022;2:93-100. DOI
5.   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
6.   Holm S. Principles of biomedical ethics, 5th edn. Beauchamp T L, Childress J F. Oxford University Press, 2001, £19.95, pp 454. ISBN 0-19-514332-9. 2002;28:332-332. Available from: https://jme.bmj.com/content/28/5/332.2. [Last accessed on 15 May 2023].
7.   Regulation (EU) 2017/745 of the european parliament and of the council of 5 April 2017. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R0745. [Last accessed on 15 May 2023].
8.   Proposal for a regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206. [Last accessed on 15 May 2023].
9.   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 15 May 2023].
10.  McCulloch P, Altman DG, Campbell WB, et al; Balliol Collaboration. No surgical innovation without evaluation: the IDEAL recommendations. *Lancet* 2009;374:1105-12. DOI
11.  Amin F, Abbasi R, Mateen A, Ali Abid M, Khan S. A step toward next-generation advancements in the internet of things technologies. *Sensors* 2022;22:8072. DOI PubMed PMC
12.  Complete guide to GDPR compliance. Available from: https://gdpr.eu/. [Last accessed on 15 May 2023].
13.  HIPAA & your health rights. Available from: https://www.hhs.gov/programs/hipaa/index.html. [Last accessed on 15 May 2023].
14.  AI ethics (AI code of ethics). Available from: https://www.techtarget.com/whatis/definition/AI-code-of-ethics. [Last accessed on 15 May 2023].
15.  Unethical use of artificial intelligence. Available from: https://lasserouhiainen.com/unethical-use-of-artificial-intelligence/. [Last accessed on 15 May 2023].
16.  Bias in AI and machine learning: some recent examples (OR cases in point). Available from: https://www.lexalytics.com/blog/bias-in-ai-machine-learning/. [Last accessed on 15 May 2023].
17.  Should AI be fair and non-discriminative. Available from: https://ethics-of-ai.mooc.fi/chapter-6/3-discrimination-and-biases. [Last accessed on 15 May 2023].
18.  3 essential steps for AI ethics. Available from: https://www.sas.com/en_us/insights/articles/analytics/artificial-intelligence-ethics.html. [Last accessed on 15 May 2023].
19.  What are the risks of artificial intelligence? Available from: https://www.linkedin.com/pulse/what-risks-artificial-intelligence-philip-mckeown. [Last accessed on 15 May 2023].
20.  Top 7 challenges in artificial intelligence in 2023. Available from: https://www.upgrad.com/blog/top-challenges-in-artificial-

intelligence/. [Last accessed on 15 May 2023].

21.  Is the way you think biased? Here are 3 common mistakes to watch out for. Available from: https://www.weforum.org/agenda/2022/04/cognitive-biases-three-common-types-illustrated/. [Last accessed on 15 May 2023].

22.  Sparrow R, Hatherley J. The promise and perils of AI in medicine. Available from: https://www.researchgate.net/publication/344879321_The_promise_and_perils_of_AI_in_medicine. [Last accessed on 15 May 2023].

23.  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

24.  Rashidian N, Hilal MA. Applications of machine learning in surgery: ethical considerations. *Art Int Surg* 2022;2:18-23.  DOI

25.  Gumbs AA, Perretta S, d'Allemagne B, Chouillard E. What is artificial intelligence surgery? *Art Int Surg* 2021;1:1-10.  DOI

26.  Boutros C, Singh V, Ocuin L, Marks JM, Hashimoto DA. Artificial intelligence in hepatopancreaticobiliary surgery - promises and perils. *Art Int Surg* 2022;2:213-23.  DOI

27.  European Commission. The 2030 agenda for sustainable development and SDGs. Available from: https://commission.europa.eu/strategy-and-policy/international-strategies/sustainable-development-goals/eu-and-united-nations-common-goals-sustainable-future_en. [Last accessed on 18 May 2023].

28.  Stypińska J, Franke A. AI revolution in healthcare and medicine and the (re-)emergence of inequalities and disadvantages for ageing population. *Front Sociol* 2022;7:1038854.  DOI  PubMed  PMC

29.  d'Elia A, Gabbay M, Rodgers S, et al. Artificial intelligence and health inequities in primary care: a systematic scoping review and framework. *Fam Med Community Health* 2022;10:e001670.  DOI  PubMed  PMC

30.  Gupta M, Parra CM, Dennehy D. Questioning racial and gender bias in AI-based recommendations: do espoused national cultural values matter? *Inf Syst Front* 2022;24:1465-81.  DOI  PubMed  PMC

31.  Elwyn G, Nelson E, Hager A, Price A. Coproduction: when users define quality. *BMJ Qual Saf* 2020;29:711-6.  DOI  PubMed  PMC

32.  Kerasidou A. Ethics of artificial intelligence in global health: Explainability, algorithmic bias and trust. *J Oral Biol Craniofac Res* 2021;11:612-4.  DOI  PubMed  PMC

33.  Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? *BMJ Glob Health* 2018;3:e000798.  DOI  PubMed  PMC

34.  Cobianchi L, Verde JM, Loftus TJ, et al. Artificial intelligence and surgery: ethical dilemmas and open issues. *J Am Coll Surg* 2022;235:268-75.  DOI

35.  Myers SP, Dasari M, Brown JB, et al. Effects of gender bias and stereotypes in surgical training: a randomized clinical trial. *JAMA Surg* 2020;155:552-60.  DOI  PubMed  PMC

36.  Yuce TK, Turner PL, Glass C, et al. National evaluation of racial/ethnic discrimination in US surgical residency programs. *JAMA Surg* 2020;155:526-8.  DOI  PubMed  PMC

37.  Ferrari L, Mari V, De Santi G, et al. Early barriers to career progression of women in surgery and solutions to improve them: a systematic scoping review. *Ann Surg* 2022;276:246-55.  DOI

38.  Karnick A, Limberg J, Bagautdinov I, et al. Can general surgery interns accurately measure their own technical skills? Analysis of cognitive bias in surgical residents' self-assessments. *Surgery* 2021;170:1353-8.  DOI  PubMed

39.  Dill-Macky A, Hsu CH, Neumayer LA, Nfonsam VN, Turner AP. The role of implicit bias in surgical resident evaluations. *J Surg Educ* 2022;79:761-8.  DOI  PubMed

40.  Voss M, Swart O, Abel L, Mahtani K. Capacity-building partnerships for surgical post-graduate training in low- and middle-income countries: a scoping review of the literature with exploratory thematic synthesis. *Health Policy Plan* 2021;35:1385-412.  DOI  PubMed

41.  Mirchi N, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. The virtual operative assistant: an explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PLoS One* 2020;15:e0229596.  DOI  PubMed  PMC

42.  Anvari M, Manoharan B, Barlow K. From telementorship to automation. *J Surg Oncol* 2021;124:246-9.  DOI  PubMed

43.  Reiley CE, Lin HC, Yuh DD, Hager GD. Review of methods for objective surgical skill evaluation. *Surg Endosc* 2011;25:356-66.  DOI  PubMed

44.  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* 2022;276:363-9.  DOI  PubMed  PMC

45.  Asghar MS, Zaman BS, Zahid A. Past, present, and future of surgical simulation and perspective of a developing country: A narrative review. *J Pak Med Assoc* 2021;71:2770-6.  PubMed

46.  Martinerie L, Rasoaherinomenjanahary F, Ronot M, et al. Health care simulation in developing countries and low-resource situations. *J Contin Educ Health Prof* 2018;38:205-12.  DOI

47.  Rangel EL, Smink DS, Castillo-Angeles M, et al. Pregnancy and motherhood during surgical training. *JAMA Surg* 2018;153:644-52.  DOI  PubMed  PMC

48.  Bissonnette V, Mirchi N, Ledwos N, Alsidieri G, Winkler-Schwartz A, Del Maestro RF, Neurosurgical Simulation & artificial intelligence learning centre. Artificial intelligence distinguishes surgical training levels in a virtual reality spinal task. *J Bone Joint Surg Am* 2019;101:e127.  DOI  PubMed  PMC

49.  Hasan O, Ayaz A, Jessar M, Docherty C, Hashmi P. The need for simulation in surgical education in developing countries. The wind of change. Review article. *J Pak Med Assoc* ;69:S62-s68.  PubMed