

# Ethical forethoughts on the use of artificial intelligence in medicine

Artificial  
intelligence in  
medicine

Bassem T. Elhassan

*Department of Orthopaedic Surgery, Massachusetts General Hospital,  
Boston, Massachusetts, USA, and*

Alya A. Arabi

*Department of Biochemistry and Molecular Biology, College of Medicine and  
Health Sciences, United Arab Emirates University, Al Ain, United Arab Emirates*

Received 12 September 2023  
Revised 11 December 2023  
Accepted 27 January 2024

## Abstract

**Purpose** – The purpose of this paper is to illuminate the ethical concerns associated with the use of artificial intelligence (AI) in the medical sector and to provide solutions that allow deriving maximum benefits from this technology without compromising ethical principles.

**Design/methodology/approach** – This paper provides a comprehensive overview of AI in medicine, exploring its technical capabilities, practical applications, and ethical implications. Based on our expertise, we offer insights from both technical and practical perspectives.

**Findings** – The study identifies several advantages of AI in medicine, including its ability to improve diagnostic accuracy, enhance surgical outcomes, and optimize healthcare delivery. However, there are pending ethical issues such as algorithmic bias, lack of transparency, data privacy issues, and the potential for AI to deskill healthcare professionals and erode humanistic values in patient care. Therefore, it is important to address these issues as promptly as possible to make sure that we benefit from the AI's implementation without causing any serious drawbacks.

**Originality/value** – This paper gains its value from the combined practical experience of Professor Elhassan gained through his practice at top hospitals worldwide, and the theoretical expertise of Dr. Arabi acquired from international institutes. The shared experiences of the authors provide valuable insights that are beneficial for raising awareness and guiding action in addressing the ethical concerns associated with the integration of artificial intelligence in medicine.

**Keywords** Ethics, Artificial intelligence, Challenges, Solutions

**Paper type** Technical paper

## Introduction

Artificial intelligence (AI) is rapidly paving its way into many sectors including medicine. The promise of AI in medicine is undeniable starting from the research and development to the surgical rooms in hospitals, and even the management aspects. Despite the substantial

© Bassem T. ElHassan and Alya A. Arabi. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

In the cover page preserve the blind-review conditions.

*Statements and declarations:* No competing interests to declare. The authors have no relevant financial or nonfinancial interests to disclose.



---

advancement and accuracy in AI, there is still public hesitance in blindly accepting it. In terms of surgeries, up until 2020, more than half and up to 63% of the respondents across multiple studies did not feel comfortable with undergoing a robotic surgery (Stai *et al.*, 2020). If the public perception of AI in medicine is predominantly influenced by an admixture of fear from the incursion of this nascent technology and its ethical pitfalls (Arabi, 2021), then there will be continued hesitance toward embracing AI for invasive surgeries despite its advancements and proven accuracy in various medical applications. This emphasizes the importance of bridging the gap between experts in AI technology and the public, to guide the advantageous use of this technology and foster trust without compromising the safety and rights of end users.

### Artificial intelligence advantages

AI is a predictive and automated decision-making tool based on fed data input. Examples of data fed into surgery-related models include electronic health/medical records, biographical information, laboratory results, livestreaming vital signs, biometrics, surgical videos, radiomics, and radiological imaging. AI has the potential to save lives, money, and time (Pushkaran and Arabi, 2024). Massachusetts Institute of Technology (MIT) algorithms can analyze 3D images *ca.* thousand times quicker than standard techniques. Despite its infancy, AI is powerful in early interventions, clinical diagnoses, and surgeries. Postoperatively, AI robotics reduce complications after orthopedic surgeries by fivefold and inpatient stays by 21%. Intraoperatively, AI helps keep surgeries minimally invasive and supports proactive and preventive actions rather than reactions to unexpected changes. AI's accuracy in image analysis can abolish many invasive (*e.g.* brain) surgeries, saving pain and adverse surgery outcomes.

Unlike humans, AI can retain big data and acquire experience from a vast number of surgeries. Thus, despite humans' cognitive reasoning, AI may distinguishably compete in common/unexceptional cases. Examples include the following:

- accurate detection of skin cancer by an AI system developed at Stanford (Esteva *et al.*, 2017);
- Baidu's outperformance in detecting breast cancer metastasis;
- Corti's outperformance by 22% in detecting cardiac arrests; and
- superior accuracies in predicting life expectancies while humans overestimate them by fivefold (Avati *et al.*, 2018).

Although not fully elaborate yet, virtual humans built using AI can handle clinical trials better than humans.

AI is peculiarly influential in surgical rooms. The AI system, *Prescience*, increases the early detection of hypoxemia by 15%. Based on real-time vital data, *Prescience* alerts the anesthetist about the risk of hypoxemia, 5 min before its occurrence (Lundberg *et al.*, 2018). Based on personal experience with surgeries, when performing shoulder replacement, the bony deficiency of the glenohumeral joint can be significant, which makes the surgery very complex and, in many instances, difficult to plan appropriately. The use of virtual reality can help visualize the scapula with the glenoid and humerus in 3D during surgery which makes the intraoperative planning and surgical steps much easier. Also, we can use the same technique to connect with expert surgeons, while doing the surgery, around the world. For instance, from our location, we may instantaneously guide surgeons in surgical rooms worldwide. This significantly improves the patients' outcome.

AI bypasses corporeal limitations of humans, *e.g.* fatigue from prolonged surgeries that may extend up to 16h in orthopedics or microsurgeries, marginal delays in judgments

during surgeries where each minute can save a life, flawed individual judgments under stress (Loftus *et al.*, 2020), emotional perplexity while operating on beloved ones and limited time for availability around patients.

AI can save billions of dollars annually on administrative and medical functions through virtual receptionists or nursing assistance. *Care Angel* saves superfluous visits and admissions to hospitals by better orchestrating care delivery, management, and coordination between patients and physicians. However, this cost reduction is contradicted by an arguably higher cost of running supercomputers and their excessive consumption of energy and storage for AI applications (Strubell *et al.*, 2019). This is not the sole drawback, there are serious ethical and legal issues associated with the use of AI as will be discussed below.

### Ethical pitfalls and challenges

As a result of errors in the algorithms-data-practice trichotomy, ample deontological and teleological ethical concerns, commensurate with the use of AI in healthcare, have begun to emerge (Kohli and Geis, 2018). The pervasive ethical ambiguity of AI (at the design, implementation, and evaluation stages) (Buchlak *et al.*, 2020) impairs its adoption. The intricacies of the AI invasion of privacy, fairness, and transparency are highlighted here along with the consequences engendered at various levels: socially, culturally, legally, and professionally.

#### *Algorithms*

Most machine learning (ML) algorithms suffer from a lack of transparency. They are rather “black boxes” (more so for the end users than the developers) where the input data goes through many layers of deep neural networks and the analysis is completed without revealing the intermediate steps or chain of predictions. The following questions are then raised: “Are the results understandable, explainable, and interpretable by diligent physicians?”, “Based on what, how, or why was a prediction made?” and subsequently, other questions pertaining to validity, scientific reliability, medical trust, and safety.

Bias and inscrutability are outcomes of technical inadequacies, inaccuracies in deploying algorithms, and overfitting (*e.g.* an algorithm trained on *in vitro* data cannot be used for *in vivo* predictions; or a heart-failure prediction model trained on cardiovascular risk factors of adult Caucasian males (Landry *et al.*, 2018) cannot be used on adult African populations). This renders the applications of such models fraught with unfairness.

There is a vicious cycle where people have the right to information and explanation, but AI makes nonexplainable nor interpretable automated decisions. Thus, informed consents have their hitches (Daniel Schiff, 2019). They reflect either superficial overconfidence or anchoring complexity (out of medical staff’s fear).

#### *Data*

In AI, decisions are autonomous based on the analysis of the input and the predicted output. Quixotically, if these predictions are consistently congruent with the expected outcome, the model is 100% accurate. However, it is impracticable to have such models. Even exceptionally accurate models have a small margin for erroneous predictions, depending on the inaccuracies in the training sets (*e.g.* outdated, incomplete, nonstandardized, noninclusive, nondiverse, incorrectly labeled or flawed data). Biased models seriously amplify existing health discrepancies and beget unfairness, disparities, and inequalities in decisions related to *e.g.* patient triage, preoperative risk stratification, and prioritization to intensive-care-unit resources postoperatively. Thus, without adequate data, any AI model would be worthless, if not even detrimental.

---

The need to access massive volumes of patients' data for training ML models is contradicted by the need to protect privacy and autonomy. Despite exercising prudence in data collection, AI models are assembled at the cost of risking the invasion of sensitive data or the right to stringent confidentiality. It is worth noting that synthetic data can now be generated through generative adversarial network (GAN) and conditional GAN, where fake/dummy data can be generated out of true datasets. This saves the hassle of extensive data collection and protects the anonymity of the data better than traditional tools such as k-anonymity and differential privacy. The question is then about the utility of the fake data.

On another note, the data-hungry AI often pushes for extensive endless questionnaires. This can be time-exhausting and frustrating for patients and physicians. Physicians allocate *ca.* 44% of their time on data entry versus only 28% on direct contact with patients (Hill *et al.*, 2013). For AI in surgery, data should be recorded in the operating rooms which can affect the performance of the surgeons who, under ubiquitous surveillance, may feel uncomfortable, face subtle nuances that allude to a breach of professional secrecy, or fear legal and jurisprudence risks (Henken *et al.*, 2012).

The use of AI dictates, in addition to hardware malfunctioning/glitches, the cybersecurity risks encountered with regular software such as hacking, encrypting data viruses, bugs, malware and, coding defects.

Last, data scientists do not have the skills of clinical relevance, thus, despite the input from medical experts, their models need to be built recursively for self-improvement until full automation at expert levels.

### *Practice*

Certain AI decisions may be replete with uncertainties, especially if medical doctors themselves are uncertain about the "right" decision, especially in the surgical rooms where peculiar *ad hoc* decisions are frequently made. The morals in surgical ethics (Rudzicz and Saqur, 2020) are prone to amplifications and complexities when AI technology is engaged. In the surgery rooms, the existing challenges related to pragmatically applying the *prima facie* duty theory of clinical ethics are exaggerated as follows:

- AI-assisted surgeries cannot guarantee *beneficence* in the stipulated sense of maximizing the patients' benefits while minimizing the harm. This is because the benefits are gained at the cost of the high risks of AI's unavoidable challenge of explainability and interpretability. In addition, should the AI robot decide that the mission is impossible and the surgery will be futile, it may choose not to rescue the patient. Such scenarios impose challenges for judges to adjudicate;
- intraoperative AI requires interactive actions/decisions. However, patients under anesthesia cannot have the *autonomy* of choosing to accept or reject the AI's intraoperative recommendations;
- the paucity of data resources causes biased AI algorithms to abolish the pillar of *justice* and fairness in medicine; and
- *nonmaleficence* is inevitable even with the minimally invasive surgeries assisted by AI robots. In addition, AI robots may decide to execute euthanasia.

No matter how technically-balanced and well-trained ML algorithms are, they lack human intuition. For example:

- ML would never meet the doctors' capability of handling unfamiliar cases;
- AI may offer unwelcome surprising suggestions;

- AI lacks intuition and the urgency of beneficence to rescue patients;
- AI crumbles with erratic patient behaviors (although this criticism can equally apply to humans);
- AI models are not mature enough to respond to statements like “I want to commit suicide.”;
- AI cannot admit mistakes nor advocate for colleagues and justice; and
- AI has no humanistic interactions or perceptions (Vergheese *et al.*, 2018) such as eye contact, authenticity, creativity, love, empathetic rather than stoic approaches, caring for patients (Stokes and Palmer, 2020), kindness keeping in mind that much of the pain is psychologically fought, and trust toward medical staff often boosts quick recovery (Israni and Vergheese, 2019).

Culturally, AI sets the risks of eroding humanism in healthcare, threatening patient-physician synergy (although, with the assistance of AI, physicians will have more free time to interact with their patients). Professionally, AI may deskill physicians and jeopardize their jobs. The real ethical dilemma is in AI models replacing the physicians who have offered their own expertise to improve these models. In a nutshell, AI must be implemented only under the surveillance of human intelligence until full autonomy in medical, legal, and cultural contexts, which is still a long way away from now.

#### *Unanswered questions*

Questions stemming from the use of AI in medicine are plentiful (McGreevey *et al.*, 2020):

- What are the guidelines for patients and physicians to accept and trust AI machines to operate? How would higher levels of autonomy affect the ethics paradigm and its challenges? (Bertoncini and Serafim, 2023) Shall AI machines acquire certifications and licenses like physicians do to practice medicine? What is the equivalent of the oath “First, do no harm.” that physicians swear? Could AI robots one day overrule humans and eschew instructions/procedures or decide to conduct voluntarily euthanasia for patients? How can one prohibit a machine from recidivism, causing deleterious effects, or making harmful decisions? What levels of moral culpability and legal liability would machines have, if any? Are they accountable for their behavior? Is it possible to claim that a machine inadvertently harmed a patient? Could they explain their actions? The question of liability is “a problem of many hands” where blame attribution can be easily obfuscated. This can leave the judges’ hands tied as it challenges the possibilities of making sharp decisions.
- How can a machine understand the patient’s cultural background and implement cultural norms? Does AI discern mood, intent, gestures, disabilities, and bleeding, to rigorously assess the severity of the case? Nonetheless, AI is progressing in this direction, there is now a Danish firm that can detect the tone of voice and use it in its prediction of cardiac arrests over the phone.
- Who has the ownership and control over the data used in AI? Can insurance companies base their quotes based on AI predictions?
- Who owns the patent for a new surgical procedure developed by an AI machine? The law in some countries, such as the USA, does not allow granting patents to machines.

---

*Rules and legislation*

With AI development/implementation, ethical concerns will further elaborate in numbers and severity. Updated legally binding regulations and legislative norms need to be issued as traditional laws would no longer be suitable for scenarios where AI machines make fully or partially autonomous decisions. There have been efforts put in this direction. Most of the AI-related regulations started to be released after 2016. At first, more than half of these regulations (approximately 55%) were set by the USA, the UK, Japan, and countries in the European Union. Roughly half of the regulations are led by private sectors and government agencies. The majority of the rest of the documents are released by scientific organizations, institutes, foundations, and associations, with minimal contributions from worker unions and political parties (Jobin *et al.*, 2019). Committees of experts left their fingerprints on the (nonlegislative) policies or soft laws concerning:

- the ethics of the use of AI; and
- the guidance on the degrees of autonomy of medical equipment and on the safe and effective implementation of AI technologies through the European Commission, the European Parliament Resolution-Civil law rules on robotics (February 2017), the Organization for Economic Co-operation and Development, Vatican's AI ethics plan (The Rome Call for AI Ethics), which includes six fundamental principles: transparency, inclusion, responsibility, impartiality, reliability, security and privacy, the Advisory Council on the Ethical Use of Artificial Intelligence and Data in Singapore, the UK House of Lords, the International Electrotechnical Commission (IEC 60601), the Institute of Electrical and Electronics Engineers, the National Academy of Medicine, the Association of Computing Machinery, Access Now and Amnesty International, EU-US Privacy Shield Framework, General Data Protection Regulation (GDPR), United Nations Educational, Scientific and Cultural Organization (UNESCO)'s Recommendation on the Ethics of Artificial Intelligence, The European Union's Artificial Intelligence Act, The Montreal Declaration for Responsible AI in Health and The World Health Organization (WHO)'s Ethical Principles for Artificial Intelligence in Health.

Companies like Google developed their own guidelines. There are also risk assessments on the use of AI, *e.g.* the (mandatory) Algorithmic Impact Assessment by the Canadian government and the Harm Assessment Risk Tool framework used by the police forces in the UK. Legislative bodies can benefit from such technical papers by having comprehensive summaries of the flaws that need to be addressed.

**Solutions and suggestions**

Ethics must be embedded as early as the research stages and at the practical levels. Solutions already implemented, currently at a nascent stage of development, or to be considered in the future are incorporated in this section. Before allowing its implementation in medicine, an AI model must prioritize the benefit of patients by design and protect their emotional fulfillment. It should encompass best clinical practice, be transparent, and fulfill privacy laws and cultural norms. In addition, AI models have to adhere to existing rules and regulations. They should be strictly rejected if they have the potential to make unsafe or unethical decisions, under any situation or circumstance. AI models must specify the scope of proficiencies and the certainty of their safe use across multifarious subgroups according to available risk assessment measures. Finally, detailed instructions on how to use the machine should be provided with fully extensive documentation on the predictive accuracy,

the limitations of the model, the types of errors along with their frequency or rates of occurrences, and the severity of the side effects stemming from these errors.

### *Transparency*

Transparency helps improve the field, foster trust, reduce damages, clarify matters for legal issues, and fulfill principles of democracy (Buhmann and Fieseler, 2022). Transparency can be addressed by providing source codes, data used, lists of limitations, and potential consequences using nontechnical terms to bridge between the developers, investors, service providers, and the end-users of AI. The use of blockchain in healthcare also helps transparency and provides a clear audit trail for AI decision-making (Abad-Segura *et al.*, 2021). A blockchain network in healthcare is also useful for securing accounting management, exchanging patient data, and for avoiding serious mistakes (Haleem *et al.*, 2021).

### *Techniques*

It is mandatory to diminish the knowledge gap between the developers and the end-users of AI in medicine. This can be done through training. Technical support should be provided to supervise the proper use of the model, monitor the correct implementations of protocols, and ensure the correct remediation of errors and appropriate data analysis. The healthcare systems shall undergo recurrent thorough audits. In addition, education institutes shall inaugurate multidisciplinary programs that interwind computer sciences with medicine.

### *Liability*

There is an urgent need to set explicit international (standardized) laws about the accountability and liability for each type of error stemming from the use of AI (O'Sullivan *et al.*, 2019).

### *Justice and fairness*

It is imperative to block any nonmaleficence that may foreseeably harm society, even if unintentionally. It is equally mandatory to protect privacy, freedom, trust, dignity, solidarity, sustainability, beneficence, emotional and psychological well-being, socio-economic opportunities, socio-democratic rights and diversity, equity, and inclusion. As important as logic and strategies are in ML, value alignment holds even greater significance because it incorporates a major additional component: moral intelligence (Bertoncini and Serafim, 2023). To understand the preferences of a society or the ethics inherent to humans, inverse reinforcement learning can be applied.

### *Data privacy*

For data privacy, federated learning can be used: the downloaded ML on a device uses data in it without having to upload the data to a cloud. Differential privacy is another solution that adds noise that hinders the private info without altering the precision of the model. *K*-anonymity keeps the data anonymous without the risk of guessing the identity of the patient from the collective input. Synthetic data is among the most recent techniques for data anonymity, where fake data can be generated from real data, but the utility of the former is debatable. The American Health Insurance Portability and Accountability Act and the European GDPR laws also impose the strict application of data privacy for patient care and research. It is worth noting, however, that the language used in such laws may lack precision, which opens doors for ambiguities.

*Bias*

*Vis-à-vis* prejudices, we need to focus on the quality and representativeness of data used in training AI. Algorithms should be thoroughly trained, properly calibrated, trustfully validated, and appropriately selected for the right application. Astute clinicians can provide or comment on the quality of data used in AI models. It is important to obstruct forced and blind acceptance of biased and discriminative AI decisions which often emanate from building ML models using flawed datasets. AI decisions shall be challenged and appealed. Patients shall not surrender themselves to AI, they should register complaints and provide postimplantation feedback.

*Interpretability and explainability*

Clinicians can better trust AI if the bottlenecks of explainability (Theunissen and Browning, 2022) and interpretability are eluded. For interpretability, data scientists can use a linear-gradient boosting tree with labeled data so that the feature inputs are assigned importance scores. For explainability (Bertoncini and Serafim, 2023), the AI's behavior shall be simultaneous, decomposable, and transparent with textual or visualized explanations. Explainable artificial intelligence (XAI) opens the "black box" of explainability and interpretability in AI models and elucidates its opacity (Gordon *et al.*, 2019). It computes the relative risk for each risk factor by exclusively removing it to monitor the effects of its absence on the predictions. XAI would, based on laparoscopic videos, predict an event of bleeding while explaining the reasons for its occurrence with respect to the patient, team, and surgical factors; along with risk-connected alerts such as the patient's blood pressure or existing anatomical abnormalities.

**Conclusions**

AI can span a vast swath of scenarios from highly utopian to dangerously dystopian. AI can make the world a better place to live in with fewer medical problems, but it may also set humanism at risk. In reality, we are in a transhumanism phase (Massotte, 2017) where we shall, as Dan Brown discusses in his book "Origin", vigilantly and judiciously maximize the benefit of the hybrid model of humans and technology. Physicians and AI will inherently be two sides of the same coin. We just need to leverage the benefits of AI in medicine by scrutinizing the technology through multiple layers and selectively and controllably using it. It is crucial to proactively control the adoption of AI as a complementary tool in the healthcare system to augment the intellectual and practical functions of the physician without sacrificing humanism, the indispensable existence of the practitioner's unwavering cognitive reasoning, and the ethics of medicine. The flourishing of AI in our era has sparked myriad debates about the ethical facets and profound dilemmas juxtaposed with the integration of beneficial AI decision-making systems into medical practices. Therefore, there is a need to urgently develop a mechanism to wisely, yet promptly, generate up-to-date rules (with close attention to the abundant allies in this field) as the pace of AI evolution and implementation is exceptionally dynamic. Filling the gap between experts in AI and the public, including decision-makers, has certainly assisted in better articulating these rules, although there is always more room for further transparency to allow sharp decisions, especially in court cases.

**References**

- Abad-Segura, E., Infante-Moro, A., González-Zamar, M.-D. and López-Meneses, E. (2021), "Blockchain technology for secure accounting management: research trends analysis", *Mathematics*, Vol. 9 No. 14, p. 1631.

- Arabi, A.A. (2021), "Artificial intelligence in drug design: algorithms, applications, challenges and ethics", *Future Drug Discovery*, Vol. 3 No. 2, p. FDD59.
- Avati, A., Jung, K., Harman, S., Downing, L., Ng, A. and Shah, N.H. (2018), "Improving palliative care with deep learning", *BMC Medical Informatics and Decision Making*, Vol. 18 No. S4.
- Bertoncini, A.L.C. and Serafim, M.C. (2023), "Ethical content in artificial intelligence systems: a demand explained in three critical points", *Frontiers in Psychology*, Vol. 14
- Buchlak, Q.D., Esmaili, N., Leveque, J.-C., Bennett, C., Piccardi, M. and Farrokhi, F. (2020), "Ethical thinking machines in surgery and the requirement for clinical leadership", *The American Journal of Surgery*, Vol. 220 No. 5, pp. 1372-1374.
- Buhmann, A. and Fieseler, C. (2022), "Deep learning meets deep democracy: deliberative governance and responsible innovation in artificial intelligence", *Business Ethics Quarterly*, Vol. 33 No. 1, pp. 1-34.
- Daniel Schiff, M.J.B. (2019), "How should clinicians communicate with patients about the roles of artificially intelligent team members?", *AMA Journal of Ethics*, Vol. 21 No. 2, pp. 138-145.
- Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M. and Thrun, S. (2017), "Dermatologist-level classification of skin cancer with deep neural networks", *Nature*, Vol. 542 No. 7639, pp. 115-118.
- Gordon, L., Grantcharov, T. and Rudzicz, F. (2019), "Explainable artificial intelligence for safe intraoperative decision support", *JAMA Surgery*, Vol. 154 No. 11, p. 1064.
- Haleem, A., Javaid, M., Singh, R.P., Suman, R. and Rab, S. (2021), "Blockchain technology applications in healthcare: an overview", *International Journal of Intelligent Networks*, Vol. 2, pp. 130-139.
- Henken, K.R., Jansen, F.W., Klein, J., Stassen, L.P., Dankelman, J. and Van Den Dobbelen, J.J. (2012), "Implications of the law on video recording in clinical practice", *Surgical Endoscopy*, Vol. 26 No. 10, pp. 2909-2916.
- Hill, R.G., Sears, L.M. and Melanson, S.W. (2013), "4000 Clicks: a productivity analysis of electronic medical records in a community hospital ED", *The American Journal of Emergency Medicine*, Vol. 31 No. 11, pp. 1591-1594.
- Israni, S.T. and Verghese, A. (2019), "Humanizing artificial intelligence", *JAMA*, Vol. 321 No. 1, pp. 29-30.
- Jobin, A., Ienca, M. and Vayena, E. (2019), "The global landscape of AI ethics guidelines", *Nature Machine Intelligence*, Vol. 1 No. 9, pp. 389-399.
- Kohli, M. and Geis, R. (2018), "Ethics, artificial intelligence, and radiology", *Journal of the American College of Radiology*, Vol. 15 No. 9, pp. 1317-1319.
- Landry, L.G., Ali, N., Williams, D.R., Rehm, H.L. and Bonham, V.L. (2018), "Lack of diversity in genomic databases is a barrier to translating precision medicine research into practice", *Health Affairs*, Vol. 37 No. 5, pp. 780-785.
- Loftus, T.J., Tighe, P.J., Filiberto, A.C., Efron, P.A., Brakenridge, S.C., Mohr, A.M., Rashidi, P., Upchurch, G.R. and Bihorac, A. (2020), "Artificial intelligence and surgical decision-making", *JAMA Surgery*, Vol. 155 No. 2, pp. 148-158.
- Lundberg, S.M., Nair, B., Vavilala, M.S., Horibe, M., Eisses, M.J., Adams, T., Liston, D.E., Low, D.K.-W., Newman, S.-F., Kim, J. and Lee, S.-I. (2018), "Explainable machine-learning predictions for the prevention of hypoxaemia during surgery", *Nature Biomedical Engineering*, Vol. 2 No. 10, pp. 749-760.
- McGreevey, J.D., Hanson, C.W. and Koppel, R. (2020), "Clinical, legal, and ethical aspects of artificial intelligence-assisted conversational agents in health care", *JAMA*, Vol. 324 No. 6, pp. 552-553.
- Massot, P. (2017), "Ethics and transhumanism: control using robotics and artificial intelligence", *Ethics in Social Networking and Business 2*, John Wiley and Sons, Inc., pp. 57-80.
- O'Sullivan, S., Nevejans, N., Allen, C., Blyth, A., Leonard, S., Pagallo, U., Holzinger, K., Holzinger, A., Sajid, M.I. and Ashrafian, H. (2019), "Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery", *The International Journal of Medical Robotics and Computer Assisted Surgery*, Vol. 15 No. 1, p. e1968.

- Pushkaran, A.C. and Arabi, A.A. (2024), "From understanding diseases to drug design: can artificial intelligence bridge the gap?", *Artificial Intelligence Review*, Vol. 57, p. 86.
- Rudzicz, F. and Saqur, R. (2020), "Ethics of artificial intelligence in surgery", *Artificial Intelligence in Surgery: A Primer for Surgical Practice*, McGraw Hill, New York, ISBN: 978-1260452730.
- Stai, B., Heller, N., McSweeney, S., Rickman, J., Blake, P., Vasdev, R., Edgerton, Z., Tejpal, R., Peterson, M., Rosenberg, J., Kalapara, A., Regmi, S., Papanikolopoulos, N. and Weight, C. (2020), "Public perceptions of artificial intelligence and robotics in medicine", *Journal of Endourology*, Vol. 34 No. 10, pp. 1041-1048.
- Stokes, F. and Palmer, A. (2020), "Artificial intelligence and robotics in nursing: ethics of caring as a guide to dividing tasks between AI and humans", *Nursing Philosophy*, Vol. 21 No. 4, p. e12306.
- Strubell, E., Ganesh, A. and McCallum, A. (2019), "Energy and policy considerations for deep learning in NLP", In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Florence, pp. 3645-3650.
- Theunissen, M. and Browning, J. (2022), "Putting explainable AI in context: institutional explanations for medical AI", *Ethics and Information Technology*, Vol. 24 No. 23.
- Verghese, A., Shah, N.H. and Harrington, R.A. (2018), "What this computer needs is a physician", *JAMA*, Vol. 319 No. 1, pp. 19-20.

---

**Corresponding author**

Alya A. Arabi can be contacted at: [alya.arabi@uaeu.ac.ae](mailto:alya.arabi@uaeu.ac.ae)