Paulina Szaniawska

Faculty of Political Science and Journalism of the Maria Curie-Skłodowska University
E-mail: paulina.military02@gmail.com
ORCID: 0009-0009-9421-6307

ETHICS AND SECURITY RELATED TO AI IN MEDICINE

Abstract: Artificial intelligence (AI) in medicine is becoming one of the most important tools to support diagnosis, therapy and management of healthcare systems. However, its rapid development comes with a number of ethical challenges and risks related to related to safety. Key issues include algorithmic biases, medical data privacy, accountability for AI decisions and the availability of these technologies. The article discusses issues related to transparency, inequalities in access to AI, and the impact of these technologies on the doctor-patient relationship. Particular attention is given to regulations that should keep up with the rapid pace of AI implementation in medicine. Potential solutions that can help the sustainable and responsible development of this technology are also proposed.

Keywords: artificial intelligence (AI), medicine, ethics, data security, privacy, transparency.

INTRODUCTION

Artificial intelligence (AI) is defined as the ability of machines to mimic human intellectual processes such as analysis, learning and decision-making. Its rapid development over the past few years has led to its increasing application in in medicine. AI not only assists doctors in diagnosis and treatment, but also sets new standards in the management of medical facilities. However, the introduction of this technology comes with ethical, technological and social challenges that require attention to ensure AI can safely and effectively support the healthcare system.

Artificial intelligence systems are now being used worldwide in various areas of medicine, such as diagnostics, analysing imaging results and supporting the selection of the most appropriate therapy for patients. Machine learning algorithms also play a key role in developing new drugs, identifying potential therapeutic targets and virtually testing the effectiveness of new molecules.

The COVID-19 pandemic has accelerated the use of AI in public health activities, including health systems management and infectious disease monitoring. These technologies are also being used to preselect patients with suspected SARS-CoV-2 virus infection. In turn, restrictions on social interactions and impeded access to health services have contributed to the growing popularity of combining mobile technologies with AI systems, which are increasingly being used to monitor the health of formally healthy people who want to take an active interest in their wellbeing.

Support from AI algorithms also enables better access to health services for residents in non-urbanised areas and remote from large medical centres, the World Health Organisation reminds. A paper published by the WHO entitled 'Ethics and governance of artificial intelligence for health' (WHO, 2021) indicates that the potential of AI should not be overestimated, especially when the implementation of such technology comes at the expense of investment to universalise access to health care.

Like any innovative technology, artificial intelligence has great potential to significantly improve access to medical services for millions of people around the world. At the same time, like other technologies, it can be misused, leading to harm to people, noted Dr Tedros Adhanom Ghebreyesus (WHO, 2021), Director-General of the World Health Organisation. – This report is a valuable guide for governments that want to maximise the benefits of from the use of AI, while minimising risks and avoiding potential problems.

According to the authors of the WHO report, the document is mainly aimed at public policy makers, such as ministries of health, regulators or legislators, responsible for implementing technology to improve the quality of life of citizens. However, the guidelines and examples it contains can also be a valuable resource for technology companies, scientists, researchers and health professionals.

SIX PRINCIPLES OF WHO

Public and private investment in the development of artificial intelligence technologies is key to improving health services, the report's authors note. However, a lack of adequate regulation in this area could jeopardise the rights of patients and communities, exposing them to commercial efforts by large technology corporations focused on maximising profits. For this reason, the World Health Organisation (WHO) calls for systemic oversight by governments. Such oversight should encompass every stage – from design to implementation to operation of AI systems.

Experts involved by WHO in the development of the report formulated six basic principles that can help in the evaluation and implementation of artificial intelligence in the health sector (WHO, 2021):

1. Protecting human self-determination. Final clinical decisions must belong to humans and not to AI systems. It is crucial to ensure privacy and the protection of medical data, and to obtain informed consent from patients for its use in accordance with applicable regulations.

2. Promoting wellbeing, safety and public interest. Systems for measuring and controlling quality in AI-enabled environments are needed. Algorithm developers should comply with regulations regarding safety and effectiveness in specific medical applications.

3. Transparency and comprehensibility. Information about AI systems should be public and available in a way that it can be understood. A public debate on the use of AI in medicine is essential to build public trust.

4. Promoting accountability. It is crucial to implement control mechanisms to assess the effectiveness of AI and eliminate harm from its use. These technologies should only be used by appropriately trained personnel and within the guidelines set.

5. Ensuring equality and inclusivity. AI algorithms should be designed so that their benefits extend to the widest possible audience, regardless of characteristics such as age, gender, socio-economic status or ethnicity.

6. Promoting sustainability and environmental responsibility. AI systems must be designed to minimise their environmental

impact and optimise energy consumption. In addition, stakeholders should take measures to mitigate the effects of automation, such as job reductions, by providing training and support for medical staff adapting to new technologies.

The WHO report points to the need for responsible deployment of AI to maximise the public health benefits while mitigating the risks associated with its use.

APPLICATION OF AI IN MEDICINE

Artificial intelligence (AI) is a technology that is playing an increasingly important role in various areas of life, including medicine. With its ability to process vast amounts of data, learn from it and make decisions, AI is becoming an invaluable support for medical professionals. Since the first attempts to use this technology in diagnostics in the 1960s, AI has come a long way, evolving from simple algorithms to advanced systems based on deep machine learning (CNN) and machine learning (ML) (Medidesk, 2024). Currently, its applications include areas such as medical image analysis, personalisation of therapy, patient monitoring or clinical decision support.

AI is not only increasing the precision of diagnosis and the effectiveness of treatment, but is also revolutionising the daily work of doctors, enabling a more efficient use of time and resources. Despite the huge potential of this technology, its integration into the healthcare system comes with a number of challenges, such as the need to ensure the security of patient data, transparency in decision-making and the creation of appropriate regulations. Nevertheless, AI promises to be a key part of the future of medicine.

1. The first applications of AI in medicine took place as early as the 1960s, but it is only the current technological developments that allow its potential to be put into in practice. Contemporary application examples include (WHO, 2021).
2. Imaging diagnostics: programmes can analyse X-ray images, CT scans and MRI scans, often with a precision that rivals that of specialists.

3. Oncology: IBM's Watson assists in the selection of oncology therapy, taking into account individual patient characteristics, which reduces diagnosis time and increases treatment effectiveness.
4. Telemedicine: DeepMind Health has developed Streams, an app that allows real-time monitoring of a patient's health status, making doctors more efficient.

BENEFITS OF AI

The application of artificial intelligence in medicine opens up new possibilities that were previously beyond the reach of traditional diagnostic and therapeutic methods. AI can analyse medical data quickly and accurately, enabling doctors to make more accurate decisions in less time (Medidesk, 2024). With this technology, it is also possible to identify subtle patterns in test results that may be missed by humans. This in turn allows for earlier detection of diseases and more effective treatment of patients.

Additionally, AI can automate many tedious administrative tasks and analytical tasks, relieving the burden on medical staff and allowing them to focus on the individual needs of the patient. The introduction of AI also promotes greater accessibility to healthcare, reducing waiting times for appointments or test results. The benefits of this technology have the potential not only to improve the work of doctors, but also to increase the comfort and safety for patients, making healthcare more efficient and modern.

The introduction of artificial intelligence into medicine brings many benefits such as reduced diagnostic and treatment times, where AI algorithms can process medical data in a fraction of the time needed by a human, increased efficiency of diagnoses thanks to AI's ability to detect subtle changes that can be missed by a human, AI also simulates the development of clinical conditions, enabling precise treatment planning, but also helps to automate administrative processes and analytical processes, allowing doctors to focus on patients (WHO, 2021). With these advantages, AI can help to improve the quality of medical services, as well as reduce the operational costs of healthcare facilities.

ETHICAL AND SOCIAL CHALLENGES

The development of artificial intelligence (AI) in medicine is generating not only excitement about the of the potential benefits, but also deep reflection on the ethical and social implications of its implementation. As a technology capable of autonomously analysing data, making decisions and predicting clinical outcomes, AI raises questions about the limits of its responsibility, its impact on the doctor-patient relationship and the transformation of professional roles in the medical sector. These dilemmas are particularly relevant in the context of the complexity and unpredictability of clinical processes, where every decision can have a direct impact on patients' lives and health.

One of the key ethical challenges is the problem of the so-called 'black box', or lack of transparency in the operation of AI algorithms. Although the technology can achieve impressive results in diagnosis and treatment, the decision-making processes on which AI is based often remain incomprehensible to users – including doctors. This causes difficulties in assessing the reliability of results and limits the ability of specialists to control the system's work. This raises the question of responsibility for possible AI errors: should it be borne by the algorithm developer, the supervising physician, or perhaps the medical facility that decided to implement the technology?

Another issue is the impact of AI on human relationships in medicine. Patients often stress the importance of empathy and the doctor's involvement in the treatment process – values that are difficult to replace with technology. The automation of some aspects of medical work, while streamlining many processes, runs the risk of objectifying the patient, which can lead to a loss of trust in the healthcare system. At the same time, AI is changing the way professional roles in medicine are perceived. Concerns about automation and the replacement of humans by machines are causing anxiety among doctors and medical students who fear losing their jobs or marginalising their role.

There is also the issue of data security. Artificial intelligence requires the processing of huge amounts of sensitive medical information, which poses the risk of privacy breaches and

unauthorised access to patient data. With the popularisation of AI, it becomes necessary to develop rigorous data protection standards and encryption methods to prevent misuse (Medidesk, 2024). The potential impact of AI on medical education and professional practice is also not insignificant. There is a risk that over-reliance on technology may lead to the disappearance of some key skills, such as physical examination and intuitive clinical decision-making. In the long term, this could reduce the competence of future generations of doctors, who will depend more on technology than on their own knowledge and experience.

These and other societal and ethical challenges require comprehensive analysis and appropriate action. It will be crucial to develop clear regulations that define the rules for the use of AI in medicine (WHO, 2021), and the creation of educational systems that prepare both doctors and patients to cooperate with this technology. The development of artificial intelligence carries enormous potential, but its effective and safe use requires a thoughtful approach that takes into account the needs of both patients and and medical staff.

POTENTIAL RISKS AND BARRIERS

Artificial intelligence (AI) in medicine raises great hopes but, at the same time, is a source of major concerns and challenges. Despite its enormous technological potential, the implementation of AI faces various barriers and poses significant risks that could affect both patients and healthcare systems. The adoption of this technology requires not only advanced technical solutions, but also a rethinking of legal, ethical and organisational issues that can significantly affect the effectiveness and safety of its use.

One of the most commonly cited risks is the potential for AI systems to make diagnostic and therapeutic errors. While algorithms can achieve a high level of efficiency in analysing medical data, their decisions are based on patterns in available databases. If this data is incomplete, inaccurate or subject to errors, AI can draw misleading conclusions. This can lead to serious health consequences, especially in situations where time and precision are of the essence, such as emergencies.

Another barrier is the 'black box effect' (Grace *et al.*, 2018), i.e. the lack of transparency in how AI systems make decisions. Even for their creators, it is often difficult to explain exactly on what basis an algorithm has reached a particular conclusion. This lack of understanding can create distrust among doctors and patients, and make it difficult to assess the correctness of the system. As a result, the question arises as to who should be responsible for the consequences of any errors – the technology developer, the medical facility or the supervising physician.

The high cost of implementing artificial intelligence is also not insignificant. The purchase of equipment, the creation of the relevant algorithms and the training of staff all require significant financial resources, which limits the availability of this technology, especially in smaller facilities or countries with limited healthcare budgets. Furthermore, the integration of new technological solutions involves a lengthy testing and implementation process, further increasing costs and adaptation time.

From a societal perspective, resistance from medical staff is also a significant challenge. Many doctors fear that artificial intelligence could replace them in their day-to-day duties, which could lead to a reduction in demand for certain specialties, such as radiology or pathomorphology. Such fears, although often unfounded in the short term, can negatively affect the willingness of staff to use the technology.

The issue of data security cannot be overlooked either. AI requires the processing of vast amounts of sensitive medical information, which poses risks to patient privacy and the possibility of unauthorised access to data. Developing effective methods to secure this information is a priority, but even the best systems can be vulnerable to cyber-attacks.

The implementation of artificial intelligence in medicine is a complex process and multifaceted. In order to avoid potential risks and overcome existing barriers, extensive measures are required, such as creating regulations, investing in staff education and ensuring data security. Only the combination of these elements will allow AI to be used to its full potential in a safe and responsible manner.

THE FUTURE OF AI IN MEDICINE

Artificial intelligence (AI) in medicine is on the threshold of a disruptive change that has the potential to change healthcare delivery forever. Advances in technology and the increasing availability of medical data are enabling increasingly sophisticated applications of AI, from diagnosing diseases to personalising treatments and managing processes in medical facilities (Kaplan & Haenlein, 2019, pp. 15–25.). As the technology develops, it is becoming clear that its impact on medicine goes far beyond the current framework, offering entirely new opportunities to improve the quality and efficiency of treatment.

In the coming years, AI has the potential to become an integral part of healthcare systems (Grace *et al.*, 2018, p. 62). Algorithms based on machine learning and deep neural networks will not only be able to analyse test results faster and more accurately than humans, but also predict disease risk based on genetic and environmental data. This approach will enable even more effective disease prevention, the development of personalised treatment plans and the minimisation of the risk of complications.

The future of AI in medicine also includes the development of technologies to support the daily work of doctors. AI-enabled intelligent medical decision support systems, surgical robots or telemedicine platforms can not only ease the burden on medical staff, but also increase the availability of healthcare in regions where there is a shortage of specialists. What's more, virtual medical assistants can help patients monitor their health and remind them to take their medication, increasing patient engagement in the treatment process.

However, the future of AI in medicine is not just about promises. The implementation of these technologies will require addressing a number of challenges, such as ensuring the transparency of algorithms, the protection of patient data, and regulations that balance the development of technology with an ethical approach to healthcare (Khan *et al.*, 2017, pp. 8–13). Also key will be a change in the education of future doctors, who will need to learn how to work effectively with AI-based systems.

While there is still a long way to go before AI is fully integrated into medical systems, the possibilities that this technology brings are inspiring. If implemented thoughtfully and ethically, it can contribute to a healthcare system that is more precise, accessible and patient-centric than ever before.

SUMMARY

Artificial intelligence (AI) is one of the fastest growing areas of technology that has the potential to revolutionise medicine on many levels. Its ability to process vast amounts of data, learn from it and make real-time decisions is opening up new possibilities in diagnosis, therapy and healthcare management. The use of AI, from analysing medical images to personalising treatment and assisting surgery, is already bringing tangible benefits such as more effective diagnoses, faster implementation of treatments and easing the burden on medical staff.

However, the development of AI in medicine also brings with it numerous challenges and potential risks. Issues such as the lack of transparency of algorithms, data security or liability for diagnostic errors require special attention. Ensuring appropriate regulation and education of medical personnel so that the technology can be used ethically and safe.

The future of AI in medicine seems promising, but its full exploitation requires a balanced approach that combines advanced technological capabilities with empathy and human experience. Only a harmonious collaboration between doctors and AI systems will allow the creation of an effective and sustainable healthcare model that serves both patients and society as a whole. Introducing AI into medicine is a complex and challenging process, but its potential benefits can significantly outweigh the challenges if done thoughtfully and responsibly.

BIBLIOGRAPHY

Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2018). When will AI exceed human performance? Evidence from AI experts. *Journal of Artificial Intelligence Research*, *62*, 729–754.

Kaplan, A., & Haenlein, M. (2019). Siri, Siri in my Hand, who's the Fairest in the Land? On the Interpretations, Illustrations and Implications of Artificial Intelligence. *Business Horizons*, *62*(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004

Khan, O., Comm, B., Bebb, G., & Alimohamed, N. A. (2017). Artificial intelligence in medicine – What oncologists need to know about its potential and its limitations. *Oncology Exchange*, *16*(4), 8–13.

Medidesk, M. (2024). *Ethical aspects of the use of artificial intelligence in medicine.* https://medidesk.pl/etyczne-aspekty-stosowania-sztucznej-inteligencji-w-medycynie/ (15-11-2024).

WHO (2021.) *Ethics and governance of artificial intelligence for health.* https://www.who.int/publications/i/item/9789240029200 (15-11-2024).