Artificial intelligence and its ethical and legal implications: an integrative review

Sabrina Maciel Nascimento [1], Tábita Maika Gama de Paiva [1], Marcos Paulo Maciel Kasuga [1], Thiago de Assis Furtado Silva [1], Conceição Maria Guedes Crozara [1], Jonas Byk [1], Silvania da Conceição Furtado [1]

1. Universidade Federal do Amazonas, Manaus/AM, Brasil.

Abstract

Artificial intelligence has recently been incorporated into medicine to improve patient care, speeding up processes and allowing for greater precision. However, there is a global concern regarding the ethical and legal consequences and commitments involved in using this tool. In this context, this integrative review aimed to map evidence about the ethical and legal implications of using artificial intelligence in medicine based on studies on the subject. The use of this technology in the health area resulted in a disruptive scenario, and, therefore, it urgently needs to be regulated in all subareas ethically and responsibly. More studies are required from stakeholders to continue addressing the issue of data privacy and protection through global agreements and transdisciplinary pacts.

Keywords: Artificial intelligence. Bioethics. Jurisprudence.

Resumo

Inteligência artificial e suas implicações éticas e legais: revisão integrativa
Recentemente, a inteligência artificial passou a ser incorporada na medicina para melhorar o atendimento ao paciente, acelerando processos e permitindo maior precisão. Entretanto, há uma preocupação global quanto às consequências e comprometimentos éticos e legais que envolvem o uso dessa ferramenta. Nesse contexto, esta revisão integrativa teve o objetivo de mapear evidências acerca de implicações éticas e legais do emprego de inteligência artificial na medicina com base em estudos sobre o tema. Constatou-se que o uso dessa tecnologia na área da saúde resultou em um cenário disruptivo e que, portanto, ela precisa ser urgentemente regulamentada em todas as subáreas de forma ética e responsável. É necessário mais trabalho das partes interessadas para continuar a abordar a questão da privacidade e da proteção dos dados por meio de acordos mundiais e pactos transdisciplinares.

Palavras-Chaves: Inteligência artificial. Bioética. Jurisprudência.

Resumen

La inteligencia artificial y sus implicaciones éticas y legales: revisión integradora
Recientemente, la inteligencia artificial se ha incorporado a la medicina para mejorar la atención al paciente, agilizando los procesos y permitiendo una mayor precisión. Sin embargo, existe una preocupación mundial respecto a las consecuencias y compromisos éticos y legales que implica el uso de esta herramienta. En este contexto, esta revisión integradora tuvo el objetivo de mapear evidencias acerca de las implicaciones éticas y legales del empleo de inteligencia artificial en la medicina con base en estudios sobre el tema. Se constató que el uso de esta tecnología en el ámbito de la salud resultó en un escenario disruptivo y que, por lo tanto, urge regularla en todas las subáreas de forma ética y responsable. Es necesario que las partes interesadas trabajen más para seguir abordando la cuestión de la privacidad y de la protección de datos por medio de acuerdos mundiales y de pactos transdisciplinarios.

Palabras clave: Inteligencia artificial. Bioética. Jurisprudencia.

Research

Artificial intelligence (AI), a term coined in 1955[1], is described by computer scientist John McCarthy as *the combination of science and engineering to make intelligent devices for human welfare*, as Rupali and Amit[2] pointed out. During the 20th century, AI began to be adopted in several areas, including healthcare, such as the Robodoc in 1992[3] and the first surgical robot, the DaVinci, in 1997[4].

Therefore, technological progress in healthcare presents correspondence, i.e., reciprocity between the actors in the process. On the one hand, it benefits the patient by providing solutions to a problem or disease. On the other hand, it enables healthcare professionals to optimize protocols, standardize methods and technical instruments, and analyze data more efficiently[5].

However, to implement an innovation in the healthcare field, access to personal patient records is necessary for training models, which, since it does not please people who value their privacy, acts as an obstacle to the practical implementation of AI. Therefore, the correct use of data, privacy, and biases fall under the ethics of AI[6].

A significant challenge that the medical community faces when using AI is that the very tools used to generate models associated with healthcare to improve public health are also used by different industries. Such use by the industry can negatively affect public health by influencing human behavior through machine learning (ML) algorithms to comb through databases generated by human interaction with computers in everyday life. Therefore, it is essential to understand the strengths, limitations, opportunities, ethical challenges, and risks of AI related to health[7].

Bioethics establishes universal and mandatory ethical principles, standards, and rules for all humanity. It is an interdisciplinary and second-order discipline that broadens its theme, considers complex and interdisciplinary problems, and welcomes diverse perspectives[8]. Thus, considering the advancement of research using AI in the health area, the question arises: what are the main ethical and legal implications of AI in the health area? Therefore, this integrative review aims to map the evidence regarding the ethical and legal implications of AI in health.
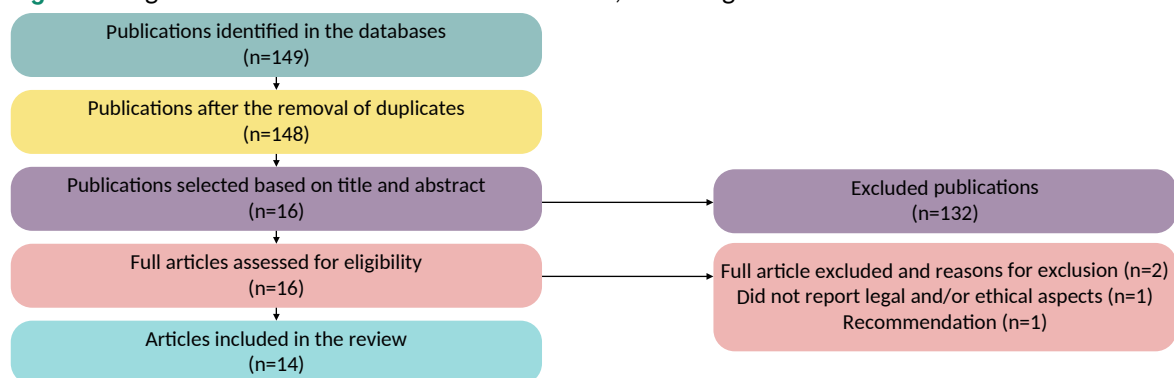
Method

This integrative literature review used the PubMed and IEEE Xplore databases and included articles published between 2018 and 2023. The search strategy adopted combined the descriptors "artificial intelligence and ethical issues and legal aspects and health" for the PubMed database and "artificial intelligence and ethical issues and health" for the IEEE Xplore database.

Results

The search identified 149 results, 98 from PubMed and 51 from IEEE Xplore. After reading the 149 titles and abstracts, 16 articles were pre-selected for full-text reading. Two were excluded, one for not reporting ethical and/or legal aspects and the other for being a recommendation (Figure 1). After the exclusions, 14 articles from 10 countries were selected for qualitative analysis, as illustrated in Table 1.

Figure 1. Diagram of the articles included and excluded, according to the criteria established in the search

Publications identified in the databases
(n=149)

Publications after the removal of duplicates
(n=148)

Publications selected based on title and abstract
(n=16)

→ Excluded publications
(n=132)

Full articles assessed for eligibility
(n=16)

→ Full article excluded and reasons for exclusion (n=2)
Did not report legal and/or ethical aspects (n=1)
Recommendation (n=1)

Articles included in the review
(n=14)

Table 1. Characteristics of the studies included in the review

Authorship; year	Country	Objective	Type of study	Type of AI	Ethical implications	Legal implications	Summary of main findings
Drabiak and collaborators; 2023 [9]	United States	Describe the new ethical challenges posed by AI/ML for clinical care and identify specific considerations for its medical practice.	Literature review	ML	There are concerns that AI/ML-driven technologies may exacerbate racial and gender inequity due to inherent bias in the training process and lack of prediction transparency.	Data privacy requirements in the United States are less extensive than in the United Kingdom or the European Union.	Realizing AI/ML's promise of improving clinical care requires addressing the emerging ethical challenges and concerns regarding patient privacy, legal responsibility, trustworthiness, and fairness.
O'Sullivan and collaborators; 2019 [10]	Brazil	Advance the debate on the potential of AI and autonomous robotic surgery, focusing on ethics, regulation, and legal aspects.	Literature review	Robotic surgery	Current approaches lack explicit declarative knowledge and, therefore, lack transparency.	While strong legal arguments support the use of autonomous surgical robots in hostile military environments, they are a potential target for liability claims, as there are currently no laws governing liability for errors in robotic surgery.	In the near future, a surgical robot may learn and perform routine surgical tasks that a human surgeon can then supervise.
Morris and collaborators; 2023 [11]	United States	Recognize and address AI's ethical, financial, and legal implications in patient care.	Observational study	Robotic surgery	The many ethical challenges associated with AI should be informed by preemptive efforts to develop regulations to combat the lack of trust and transparency.	Legality, liability, and accountability contribute to the hesitancy to integrate AI into surgical practice and need to be carefully considered to keep pace with AI advances	Technology will undoubtedly influence surgery, but ethical standards, regulatory policies, and financial forces must be addressed to maximize its potential.

Research

Table 1. Continuation

Authorship; year	Country	Objective	Type of study	Type of AI	Ethical implications	Legal implications	Summary of main findings
Amann and collaborators; 2020 [12]	Germany, Switzerland, and the United Kingdom	Provide a comprehensive assessment of the role of explainability in medical AI and an ethical assessment of the significance of explainability for adopting AI-based tools in clinical practice.	Observational study	Artificial neural networks	Personal health data can only be legally processed after the individual's consent. Without general laws facilitating the use of personal data and information, informed consent is the standard for data use in AI.	From a legal perspective, data – acquisition, storage, transfer, processing, and analysis – will need to comply with all applicable laws, regulations, and other legal requirements. In addition, the law and its interpretation will need to adapt to the evolving state of the art of technology.	The omission of explainability in clinical decision-support systems threatens medicine's fundamental ethical values and may have detrimental consequences for individual and public health.
Wang and collaborators; 2023 [13]	China and the United States	Provide a comprehensive examination of the ethical implications of ChatGPT in healthcare.	Exploratory	ChatGPT	To prevent potential human harm, ChatGPT must adhere to principles such as beneficence, justice, fairness, medical integrity, non-maleficence, privacy, responsibility, and transparency.	Factors to consider: determining legal responsibility in cases where ChatGPT's advice leads to harm or adverse outcomes, collecting and storing sensitive patient information, and considering regulatory and licensing requirements for healthcare professionals when incorporating ChatGPT into clinical practice.	Comprehensive ethical guidelines could provide essential legal, ethical, algorithmic, and informational support for patients, physicians, and health researchers. By establishing a robust ethical framework, it is possible to harness AI's potential while safeguarding individuals' well-being.
Carter and collaborators; 2020 [14]	Australia	Assess outcomes and issues of bias, data ownership, confidentiality, consent, and legal, moral, and professional responsibility.	Literature review	DL	A significant and well-observed problem for ML systems is bias in outcomes. Human choices can distort AI systems to operate in discriminatory or exploitative ways.	Virtually no courts have developed rules specifically addressing who should be held legally responsible if AI causes harm.	Once AI becomes institutionalized, it can be difficult to reverse its use and consequences: this is why due diligence is needed before, not after, implementation. A solid and proactive role from the government, regulators, and professional groups will help to ensure that AI is introduced safely.

continues…

Table 1. Continuation

Authorship; year	Country	Objective	Type of study	Type of AI	Ethical implications	Legal implications	Summary of main findings
Morley and collaborators; 2020 [15]	United Kingdom	Summarize current AI ethics debates and identify open questions for future research.	Literature mapping review	AI algorithms	Physicians and patients need to be trained to understand and control AI tools better, which can produce predictions or recommendations used in treatments, access plans, and the protection of patient data.	The main legal challenge is allocating accountability in cases of medical error. This legislative and regulatory discussion depends on understanding the issues at each stage of AI implementation and making normative decisions about how these risks and burdens will be distributed across society.	A transparent and well-distributed system of accountability is expected to develop, where all stakeholders in the supply chain of clinical algorithms can be held accountable proportionately while respecting patient safety.
Stewart, Wong and Sung; 2021 [16]	Australia and China	Lay the foundation for an ethical framework that can be used to resolve difficult choices regarding the use of AI in clinical practice.	Literature mapping review	DL/ML	The OECD, the European Commission, and the International Conference on Ethics and Data Protection have created statements of ethical principles for using AI.	AI's use of private and confidential information should always be based on consent or legally justified alternatives that consider people's rights and protect them against privacy-based harms.	A principle-based approach is best used in problem-solving. All information should be considered sensitively and carefully, and decisions should be made, communicated, and documented.
Masoumian and collaborators; 2023 [17]	Canada and Iran	Map the scope of AI applications in emergency medicine, identify ethical issues related to the use of AI, and propose an ethical framework for its use.	Scoping review	DL/ML	Despite promising applications, AI is still not widely used in hospital departments or healthcare systems. Some of the reasons include lack of transparency standards, lack of better data protection, and human-AI relationships.	Data collection following the Personal Data Protection and Electronic Documents Act and the General Data Protection Regulation is recommended but may be detrimental to AI technologies.	Several studies have demonstrated AI's potential in various contexts, particularly in improving patient outcomes through predictive models. However, according to the synthesis of studies in this review, AI-based decision-making lacks transparency.

continues…

Research

Research

Table 1. Continuation

Authorship; year	Country	Objective	Type of study	Type of AI	Ethical implications	Legal implications	Summary of main findings
Reddy; 2023[18]	Australia	Argue that the healthcare domain, due to its complexity, requires specialized approaches to regulate AI.	Not cited	Not cited	Adapting general AI regulations to healthcare requires a deep understanding of the ethical principles that guide the sector, such as autonomy, beneficence, non-maleficence, and justice.	Ensure the safety and quality of clinical AI applications, bias toward the use, medical liability, data protection, privacy, and transparency.	Accurate regulation, tailored to the specific needs and challenges of AI in healthcare, can ensure this technology's ethical and responsible use.
Upreti and collaborators; 2023[19]	India	Promote the responsible implementation of these innovations and provide an overview of the ethical, legal, and sociological ramifications of AI use in healthcare.	Not cited	Not cited	Data security and confidentiality. Bias and discrimination. Understanding authorization. Transparency and accountability. Diversity and accessibility.	From an ethical perspective, AI technologies should be developed and used in a way that respects patient autonomy, confidentiality, and respectability.	
Wang and collaborators; 2023[20]	China	Identify commonalities and differences in understandings of fairness in medical AI, critical influencing factors, and key measures to implement fairness within different disciplinary perspectives and from English and Chinese literature.	Scoping review	Not cited	It is recommended that a medical AI ethics committee be established for ethical review and governance with clarification. The committee should aim to value orientation, embed human values and norms in medical AI, and translate ethical theories and norms into regulated algorithms and operational protocols.	Dynamic regulatory frameworks that accommodate algorithmic iteration characteristics, safety standards, operational rules for oversight, and management and risk prediction in medical AI are needed.	There is consensus on the importance of data fairness as a foundation for achieving fairness in medical AI across multidisciplinary perspectives. However, there are substantial discrepancies in critical aspects such as the concept, influencing factors, and implementation measures of fairness in medical AI.

Table 1. Continuation

Authorship; year	Country	Objective	Type of study	Type of AI	Ethical implications	Legal implications	Summary of main findings
Vayena, Blasimme and Cohen; 2018[21]	Switzerland, the United Kingdom, and the United States	Argue that machine learning in medicine must offer data protection, algorithmic transparency, and accountability to gain the trust of patients and physicians.	Not cited	ML	Not cited	Privacy, data fairness, accountability, transparency, and liability.	To earn patient trust and provider adoption, ML must be fully aligned with data protection requirements, minimize the effects of bias, be effectively regulated, and achieve transparency.
Gallese and collaborators; 2022[22]	United Kingdom	Analyze public administrative data from the United Kingdom's National Health Service (NHS) related to pre- and post-pandemic patient complaints, analyzing legal and ethical issues related to using AI systems.	Unsupervised approach	Computational intelligence (Fuzzy C-Means and swarm intelligence)	Human agency and oversight; technical robustness and security; privacy and data governance; transparency; diversity, non-discrimination, and justice; social and environmental well-being; responsibility.		Computer intelligence-based systems can be employed in the healthcare sector to support a wide range of decisions that will ultimately significantly impact patients. While most decisions are made to improve patient well-being, they may also prevent economic losses or justify management choices that adversely affect patients.

AI: *artificial intelligence; ML: machine learning; DL: deep learning; OECD: Organization for Economic Cooperation and Development*

Research

Type of technology using artificial intelligence

AI and its machine learning (ML) algorithms offer new promise for personalized biomedicine and more cost-effective healthcare, with impressive technical capabilities to mimic human cognitive capabilities[9]. AI can help in disease diagnosis, drug discovery, epidemiology, individual care, and operational efficiency in medical and hospital settings[23].

The included publications point to different types of AI in different healthcare fields. Regarding the use of AI in robotic surgery, O'Sullivan and collaborators[10] state that the successful development of reliable, state-of-the-art autonomous surgical robots is a challenge, and seeking acceptance and approval for their use is a multifaceted issue. For the authors, a more obvious concern concerns the consequences involving patient death or disability resulting directly from a surgical error.

Morris and collaborators[11] add that transparency and responsibility are fundamental to implementing AI in surgical decision-making and robotic surgery. The authors emphasize that while the influence of technology applied to surgery is undeniable, discussing ethical standards, regulatory policies, and financial forces to maximize its potential is required.

Another type of AI inserted into people's daily lives is ChatGPT (OpenAI), a large language model and AI chatbot. Wang and collaborators[13] state that its remarkable ability to access and analyze large amounts of information allows it to generate, categorize, and summarize texts with high coherence.

Machine learning is an AI branch in which computational algorithms are built from data learning[24]. Recent advances in ML software algorithms and the availability of powerful computing hardware resources have resulted in a significant expansion in the potential application of AI/ML in several areas of medicine. Despite the enthusiasm and these anticipated potentials, the application of AI/ML in daily clinical practice remains scarce, which can be attributed to technical challenges, such as integration into the clinical workflow[9].

Naik and collaborators[23] present an artificial neural network (ANN) as a conceptual framework for developing AI algorithms. It is a model of the human brain composed of an interconnected network of neurons connected by weighted communication channels. AI uses various algorithms to find complex non-linear correlations in massive data sets.

Deep learning (DL) is a type of AI that allows an algorithm to classify and group data independently. Rather than being explicitly programmed to pay attention to specific attributes or variables, these algorithms can develop the ability to recognize patterns[15] with exposure to data.

Although these technologies are promising, their widespread application has been limited in the medical domain, and ethical challenges and concerns regarding patient privacy, legal liability, reliability, and fairness have altered expectations.

Ethical aspects

According to Morris and collaborators[11], the use of AI should be based on the ethical principles and challenges involved in the physician-patient relationship, health systems, and local communities. In this sense, Drabiak and collaborators[9] state that, as it becomes increasingly prominent in health care, AI/ML raises a multitude of concerns about its impact on how physicians practice medicine.

These concerns are centered on issues related to automation bias and deskilling. Automation bias occurs when physicians over-rely on AI/ML and reduce their personal efforts to verify machine results, creating risks to patient safety[9]. Another issue regarding current ML tools in healthcare is cases in which algorithms advise surgical decisions in real-time in the operating room. Hence, doubts arise as to whether the liability lies with the surgeon or the developer of the ML tool[11].

Indeed, according to data found in a bibliometric analysis, the main aspects found in the literature on ethics in AI are those generally identified in AI-based systems, with safety assuming prominent relevance. Two major application sectors appear mainly associated with ethics in AI: medical and autonomous driving (including robotics). As a moderate attitude, adopting traditional mechanisms to improve ethical and legal aspects in cyber-physical systems engineering is recommended,

such as auditing, voluntary labeling, certification by standards organizations, responsible licensing of AI, and third-party auditing[25].

Morley and collaborators[15] point out that the legal challenge of allocating liability in medical error cases is an example of a frequent problem. Thus, legislative and regulatory discussions depend directly on understanding the issues at each stage of AI implementation—for example, data collection, training, and deployment—and how these risks and burdens will be distributed across society.

The ability to trace blame back to the manufacturer is allegedly threatened by machines that can operate according to non-fixed rules and the ability to learn new behavior patterns. This difficulty is cause for alarm, as it threatens both the moral framework of society and the foundation of the idea of liability in law. Thus, using AI may result in no one being held accountable for any harm caused[10].

Fairness is a fundamental concern of equity and equality and constitutes another issue involving ethics in the use of AI. Stewart, Wong and Sung[16] highlight the risk of unrepresentative data that consolidate and exacerbate health disparities, underestimating or overestimating risks in certain patients and populations. Reducing bias in AI is therefore necessary to promote better and more equitable health outcomes.

Legal aspects

AI/ML systems require a large volume of data for training and validation, and ownership and consent for using and protecting this data are critical issues. According to Stewart, Wong and Sung[16], using AI and confidential information should always be based on consent or legally justified alternatives, which require that people's rights be considered and protected against privacy-related harms, such as reputational damage and exposure to ridicule or hatred.

Considering that AI is being introduced exponentially in healthcare, Carter and collaborators conclude that there is *no clear regulator, no clear trial process, and no clear accountability trail*[14] on the subject. The authors view this situation as a regulatory vacuum, i.e., virtually no court has developed rules specifically addressing who should be held legally responsible if AI causes harm.

Despite this, there is an ongoing debate about whether AI *fits within existing legal categories or whether a new category with its special features and implications should be developed*[23]. The limitation of algorithmic transparency is a concern that has dominated most legal discussions on AI.

Data protection

In addition to the general legal aspects of using AI, it is essential to highlight that there is also a need to control the use of personal data, including in the health area, especially with the use of ML. An adequate data representation in this type of AI provides better performance[26]. DL, in turn, focuses on creating neural network models capable of making data-driven decisions and is particularly suitable for contexts with a large set of data available[27].

Dourado and Aith[28] point out that Law 13,709/2018 was enacted in the Brazilian scenario, establishing the legal system's explanation and review of automated decisions. This seeks to guarantee data security and protection, the fundamental rights of freedom and privacy, and the free formation of individual personality. Likewise, other supranational entities have created specific legislation for this issue.

In the case of the European Union, Goddard[29] points out that the General Data Protection Regulation (GDPR) was established. This regulation applies to all member states of the bloc and covers all residents, regardless of the data processing location. It has the following general principles: fairness and lawfulness, purpose limitation, data minimization, accuracy, storage limitation, and integrity and confidentiality.

In the case of Asian countries, while South Korea has had laws and regulations on data privacy for two decades, it has received greater legislative attention with the Personal Information Protection Act, which has boosted issues related to data privacy[30]. The structure of this law is generally similar to the GDPR of the European Union[31].

In Oceania, Australia's privacy principles are based on the Organization for Economic

Cooperation and Development's Guidelines on the Protection of Privacy and Transborder Flows of Personal Information. However, there is no specific law on privacy in the country [32]. In Africa, there is variation between countries, so some offer little or no protection policy, while others have extensive digital governance frameworks [33].

Concerning the American continent, the United States follows a sectoral approach to data protection and privacy, with no comprehensive legislation but legislation that protects within specific contexts, such as in health, education, communications, and others [34].

Brazilian Artificial Intelligence Strategy

Given the development of AI and the emergence of concerns among countries about this new technology and its potential advantages and consequences, in 2020, Brazil defined AI as a priority concerning innovations, research projects, and technology development [35].

In this context, aiming to guide the actions of the Brazilian State in the development and stimulation of research, innovation, solutions, actions, and the conscious and ethical use of AI, the Brazilian Artificial Intelligence Strategy (EBIA) was created, which *aims to enhance the development and use of technology to promote scientific advancement and solve specific problems in the country, identifying priority areas in which there is greater potential for obtaining benefits* [36].

The document is based on nine thematic axes [35]:
- Four transversal axes, which refer to legislation, regulation, and ethics; AI governance; and international aspects; and
- Five vertical axes, comprising education, workforce and training, RD&I (research, development, and innovation projects) and entrepreneurship, application in the productive sectors, and application in the public power and public security.

The strategy is based on five principles defined by the Organization for Economic Cooperation and Development (OECD), whose guidelines are ensured by Brazil, aiming at the responsible management of AI systems [35]:

- Inclusive growth, sustainable development, and well-being;
- Human-centered values and equity;
- Transparency and explainability;
- Robustness, security, and protection; and
- Accountability.

In addition, the document states that, depending on the application of AI and the associated risks, the idea of accountability imposes the need to establish governance structures in AI. This seeks to ensure the adoption of principles for trustworthy AI and the establishment of mechanisms for its observance, which are, according to the document [35] itself:

- Designating specific individuals or groups within an organization to promote compliance with the principles;
- Adopting measures to increase internal awareness of the need for such compliance, including through guidance and training; and
- Implementing an escalation process through which employees can raise compliance concerns and resolve those concerns.

In addition, there is a challenge in structuring a governance ecosystem for using AI, both in the public and private sectors. To address this, EBIA mentions the following strategic actions [35]:

- Structure governance ecosystems for the use of AI in the public and private sectors;
- Encourage data sharing, as per the General Data Protection Law (LGPD);
- Promote the development of voluntary and consensual standards to manage risks associated with AI applications;
- Encourage organizations to create data review boards or ethics committees concerning AI;
- Create an AI observatory in Brazil that can connect to other international observatories;
- Encourage the use of representative datasets to train and test models;
- Facilitate access to open government data;
- Improve the quality of available data to facilitate the detection and correction of algorithmic biases;
- Encourage the dissemination of open-source code capable of verifying discriminatory trends in datasets and ML models;

- Develop guidelines for drafting the Data Protection Impact Reports (RIPD);
- Share the benefits of AI development to the greatest extent possible and promote equal development opportunities for different regions and industries;
- Develop educational and awareness campaigns;
- Stimulate social dialogue with multisectoral participation;
- Leverage and encourage accountability practices related to AI in organizations and
- Define general and specific indicators by sector (agriculture, finance, health, etc.).

In 2021, the World Health Organization (WHO)[37] published guidance establishing six fundamental principles to promote the ethical use of AI for health: 1) protect autonomy; 2) promote human well-being, human safety, and the public interest; 3) ensure transparency, explainability, and intelligibility; 4) promote responsibility and accountability; 5) ensure inclusiveness and equity; and 6) promote AI that is responsive and sustainable.

To implement these principles and obligations, all stakeholders, whether designers and programmers, providers and patients, or ministries of health and information technology, must work together to integrate ethical norms into all phases of technology development (design, development, and deployment)[38].

Discussion

The use of new technologies raises concerns about the risks associated with their implementation, with the possibility of data breaches and inaccuracies being a significant concern. As technology continues to evolve, there is an increasing need to ensure data protection and accuracy. Security breaches can significantly impact individuals and organizations, leading to financial losses, image damage, and privacy compromise. Thus, it is essential to understand the risks of using new technologies and take appropriate measures to minimize them.

According to O'Sullivan and collaborators[10], current approaches lack explicit declarative knowledge and, therefore, transparency. In the medical domain, such a lack of transparency does not promote trust and acceptance of robotic surgery among physicians. Thus, there is a need for work among stakeholders to continue to address privacy and data protection issues, as there is a gap concerning the clinical application of AI.

In addition, existing instruments, such as tort or privacy laws, products, etc., could be adapted to mitigate harm. The conclusions of Morris and collaborators[11] corroborate issues involving legal aspects, liability, and accountability. According to the authors, legal gaps contribute significantly to the hesitancy in integrating AI into surgical practice.

Regarding data protection, Amann and collaborators[12] state that personal health data can only be legally processed after the individual's consent. In the absence of general laws that facilitate the use of personal data and information, informed consent is the standard for the current use of patient data in AI applications. However, Masoumian Hosseini and collaborators[17] state that data collection, as recommended by the Personal Data Protection and Electronic Documents Act of Canada and the GDPR, may harm AI technologies.

From a legal perspective, data acquisition, storage, transfer, processing, and analysis must comply with all laws, regulations, and other legal requirements. In addition, the law and its interpretation and implementation must constantly adapt to the evolution of the state of the art of technology[12].

For Carter and collaborators[14], a commonly observed problem in ML systems is bias in outcomes stemming from bias in the training data. Human choices can distort AI systems to operate in discriminatory or exploitative ways, as it is well recognized that both healthcare and evidence-based medicine are biased against disadvantaged groups.

Similarly, Drabiak and collaborators[9] cite concerns that AI/ML-driven technologies may exacerbate racial and gender inequity due to inherent bias in the training process and lack of prediction transparency. Carter and collaborators[14]

Research

argue that once AI becomes institutionalized in systems, it can be difficult to reverse its use and consequences, and therefore, due diligence is needed before implementation.

Wang and collaborators [20] recommend the creation of a medical AI ethics committee to review ethics and governance with clarification to guide values, incorporate human values and norms, and translate ethical theories and norms into algorithms and operational protocols that can be regulated. In this context, Reddy [18] adds that adapting general AI regulations to healthcare requires a thorough understanding of the ethical principles that guide the sector, such as autonomy, beneficence, non-maleficence, and justice.

According to Wang and collaborators [13], there should be consensus on the importance of data fairness as a basis for achieving fairness in medical AI from multidisciplinary perspectives. However, there are substantial discrepancies in fundamental aspects such as concept, influencing factors, and implementation measures of fairness in medical AI. Consequently, future research should facilitate interdisciplinary discussions to bridge the cognitive gaps between different fields and improve the practical implementation of fairness in medical AI.

From an ethical perspective, AI technologies must be developed and used to respect patient autonomy, confidentiality, and respectability. This can only be done if the processes are transparent, fair, and impartial and individuals have control over their knowledge [19].

In this context, the WHO highlights that, although technical designers and developers play a crucial role in developing AI tools for use in health, no certification or licensing process like that is required for health professionals. Therefore, it is not enough to simply appeal to individuals to uphold ethical values such as repeatability, transparency, fairness, and human dignity. Thus, a process-oriented approach is required, not solutions-oriented, that meets the needs of stakeholders following the moral and social values embodied in human rights [29].

As a limitation of the study, it can be mentioned that no scientific databases other than PubMed and IEEE Xplore were investigated, as well as the gray literature, which may have contributed to excluding studies and documents relevant to the research. However, the two databases used were sufficient to map important information on the subject and point out the current insufficiency of laws and guidelines that regulate the use of AI in the health area with the efficiency and security necessary for the proper protection of managers and users of this technology.

Final considerations

AI in healthcare can already be considered a disruptive scenario and, therefore, needs to be regulated ethically and responsibly in all sub-areas. A possible solution in the short term would be individual training through the training of professionals in the field and patients to improve the understanding and control of AI tools.

Appropriate algorithms based on unbiased real-time data must avoid data bias. Thus, stakeholders must continue to work on addressing data privacy and protection through global agreements and transdisciplinary pacts.

Robust evidence on the ethical and legal implications of using AI in healthcare is still insufficient to guarantee the safe implementation of this technology in the various sub-areas of AI in healthcare settings in a globalized world.

References

1. McCarthy J, Minsky ML, Rochester N, Shannon CE. A proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. AI Magazine [Internet]. 2006 [acesso 23 out 2023];27(4):12-12. DOI: 10.1609/aimag.v27i4.1904

2. Rupali M, Amit P. A review paper on general concepts of "Artificial intelligence and machine learning". Int Adv Res J Sci Eng Technol [Internet]. 2017 [acesso 23 out 2023].;4(4):79-82. DOI: 10.17148/IARJSET/NCIARCSE.2017.22

3. Cowley G. Introducing "robodoc": a robot finds his calling: in the operating room. Newsweek [Internet]. 1992 [acesso 23 out 2023];120(21):86. Disponível: https://pubmed.ncbi.nlm.nih.gov/10122477

4. George EI, Brand TC, LaPorta A, Marescaux J, Satava RM. Origins of robotic surgery: from skepticism to standard of care. JSLS [Internet]. 2018 [acesso 2 nov 2023];22:e2018.00039. DOI: 10.4293/JSLS.2018.00039

5. Silva Neto BRD. Medicina e adesão à inovação: a cura mediada pela tecnologia. Ponta Grossa: Atena; 2021.

6. Christensen CM, Baumann H, Ruggles R, Sadtler TM. Disruptive innovation for social change. Harv Bus Rev [Internet]. 2006 [acesso 16 nov 2023];84(12):96-101. Disponível: https://pubmed.ncbi.nlm.nih.gov/17183796

7. Sher T, Sharp R, Wright RS. Algorithms and bioethics. Mayo Clin Proc [Internet]. 2020 [acesso 16 nov 2023];95(5):843-4. DOI: 10.1016/j.mayocp.2020.03.020

8. Rhodes R, Ostertag G. Bioethics is philosophy. Am J Bioeth [Internet]. 2022 [acesso 16 nov 2023];22(12):22-5. DOI: 10.1080/15265161.2022.2134499

9. Drabiak K, Kyzer S, Nemov V, El Naqa I. AI and machine learning ethics, law, diversity, and global impact. Br J Radiol [Internet]. 2023 [acesso 16 nov 2023];96(1150):20220934. DOI: 10.1259/bjr.20220934

10. O'Sullivan S, Nevejans N, Allen C, Blyth A, Leonard S, Pagallo U *et al*. Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. Int J Med Robot [Internet]. 2019 [acesso 16 nov 2023];15:e1968. DOI: 10.1002/rcs.1968

11. Morris MX, Song EY, Rajesh A, Asaad M, Phillips BT. Ethical, legal, and financial considerations of artificial intelligence in surgery. Am Surg [Internet]. 2023 [acesso 25 nov 2023];89:55-60. DOI: 10.1177/00031348221117042

12. Amann J, Blasimme A, Vayena E, Frey D, Madai VI. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak [Internet]. 2020 [acesso 25 nov 2023];20:310. DOI: 10.1186/s12911-020-01332-6

13. Wang C, Liu S, Yang H, Guo J, Wu Y, Liu J. Ethical Considerations of using ChatGPT in health care. J Med Internet Res [Internet]. 2023 [acesso 25 nov 2023];25:e48009. DOI: 10.2196/48009

14. Carter SM, Rogers W, Win KT, Frazer H, Richards B, Houssami N. The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. Breast [Internet]. 2020 [acesso 2 nov 2023];49:25-32. DOI: 10.1016/j.breast.2019.10.001

15. Morley J, Machado CCV, Burr C, Cowls J, Joshi I, Taddeo M, Floridi L. The ethics of AI in health care: a mapping review. Soc Sci Med [Internet]. 2020 [acesso 5 out 2023];260:113172. DOI: 10.1016/j.socscimed.2020.113172

16. Stewart C, Wong SKY, Sung JJY. Mapping ethico-legal principles for the use of artificial intelligence in gastroenterology. J Gastroenterol Hepatol [Internet]. 2021 [acesso 15 out 2023];36:1143-8. DOI: 10.1111/jgh.15521

17. Masoumian Hosseini M, Masoumian Hosseini ST, Qayumi K, Ahmady S, Koohestani HR. The aspects of running artificial intelligence in emergency care; a scoping review. Arch Acad Emerg Med [Internet]. 2023 [acesso 15 out 2023];11:e38. DOI: 10.22037/aaem.v11i1.1974

18. Reddy S. Navigating the AI revolution: the case for precise regulation in health care. J Med Internet Res [Internet]. 2023 [acesso 27 out 2023];25:e49989. DOI: 10.2196/49989

19. Upreti K, Vats P, Nasir MS, Alam MS, Shahi FI, Kamal MS. Managing the ethical and sociologically aspects of ai incorporation in medical healthcare in India: unveiling the conundrum [Internet]. In: Proceedings of the 2023 IEEE World Conference on Applied Intelligence and Computing (AIC); 29-30 jul 2023; Sonbhadra, India. Sonbhadra: IEEE; 2023 [acesso 2 nov 2023];p. 293-9. DOI: 10.1109/AIC57670.2023.10263931

20. Wang Y, Song Y, Ma Z, Han X. Multidisciplinary considerations of fairness in medical AI: A scoping review. Int J Med Inform [Internet]. 2023 [acesso 10 nov 2023];178:105175. DOI: 10.1016/j.ijmedinf.2023.105175

21. Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: addressing ethical challenges. PLoS Med [Internet]. 2018 [acesso 10 nov 2023];15(11):e1002689. DOI: 10.1371/journal.pmed.1002689

Research

22. Gallese C, Fuchs C, Riva SG, Foglia E, Schettini F, Ferrario L *et al*. Predicting and characterizing legal claims of hospitals with computational intelligence: the legal and ethical implications [Internet]. In: Proceedings of the 2022 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB); 15-17 ago 2022; Ottawa. Ottawa: IEEE; 2022 [acesso 10 nov 2023]. DOI: 10.1109/CIBCB55180.2022.9863033

23. Naik N, Hameed BMZ, Shetty DK, Swain D, Shah M, Paul R *et al*. Legal and ethical consideration in artificial intelligence in healthcare: who takes responsibility? Front Surg [Internet]. 2022 [acesso 18 nov 2023];9:862322. DOI: 10.3389/fsurg.2022.862322

24. Paixão GMDM, Santos BC, Araujo RMD, Ribeiro MH, Moraes JLD, Ribeiro AL. Machine learning na medicina: revisão e aplicabilidade. Arq Bras Cardiol [Internet]. 2022 [acesso 4 nov 2023];118:95-102. DOI: 10.36660/abc.20200596

25. Leitao P, Karnouskos S. The emergence of ethics engineering in Industrial Cyber-Physical Systems [internet]. In: 2022 IEEE 5th International Conference on Industrial Cyber-Physical Systems (ICPS); 24-26 maio 2022; Coventry. Coventry: IEEE; 2022 [acesso 15 out 23]. DOI: 10.1109/ICPS51978.2022.9816931

26. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O *et al*. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data [Internet]. 2021 [acesso 15 nov 2023];8:53. DOI: 10.1186/s40537-021-00444-8

27. Kelleher JD. Deep learning. Cambridge: MIT Press; 2019.

28. Dourado DDA, Aith FMA. A regulação da inteligência artificial na saúde no Brasil começa com a Lei Geral de Proteção de Dados Pessoais. Rev Saúde Pública [Internet]. 2022 [acesso 4 nov 2023];56:80. DOI: 10.11606/s1518-8787.2022056004461

29. Goddard M. The EU General Data Protection Regulation (GDPR): European Regulation that has a Global Impact. International Journal of Market Research [Internet]. 2017 [acesso 14 out 2023];59(6):703-5. DOI: 10.2501/IJMR-2017-050

30. Ko H, Leitner J, Kim E, Jeong J. Structure and enforcement of data privacy law in South Korea. International Data Privacy Law [Internet]. 2017 [acesso 27 out 2023];7:100-14. DOI: 10.2139/ssrn.2904896

31. Ko H. Pseudonymization of healthcare data in South Korea. Nat Med [Internet]. 2022 [acesso 20 out 2023];28:15-6. DOI: 10.1038/s41591-021-01580-7

32. Watts D, Casanovas P. Privacy and data protection in Australia: a critical overview (extended abstract). W3C Workshop on Privacy and Linked Data [Internet]. 2017 [acesso 27 out 2023]. Disponível: https://bit.ly/3XlDUuH

33. Daigle B. Data Protection Laws in Africa: A Pan-African Survey and Noted Trends. J Int'l Com & Econ [Internet]. 2021 [acesso 2 nov 2023]. Disponível: https://bit.ly/3ZjCoM5

34. Boyne SM. Data Protection in the United States. Am J Comp Law [Internet]. 2018 [acesso 27 out 2023];66(supl 1):299-343. DOI: 10.1093/ajcl/avy016

35. Brasil. Ministério da Ciência, Tecnologia e Inovação. Estratégia Brasileira de Inteligência Artificial – EBIA [Internet]. Brasília: MCTI; 2021 [acesso 22 nov 2023]. Disponível: https://bit.ly/3zfe8jD

36. Brasil. Op. cit. p. 4.

37. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance [Internet]. Geneva: World Health Organization; 2021 [acesso 11 set 2024]. Disponível: https://bit.ly/4ekS71S

38. Taeihagh, A. Governance of artificial intelligence. Policy and Society [Internet]. 2021 [acesso 22 nov 2023];40(2):137-57. DOI: 10.1080/14494035.2021.1928377

Research

Research

Sabrina Maciel Nascimento – Master's student – nascimentosabrina94@gmail.com
 0000-0002-8214-3421

Tábita Maika Gama de Paiva – Master's student – tabitamayka@hotmail.com
 0009-0002-0268-7443

Marcos Paulo Maciel Kasuga – Undergraduate student – markasuga@gmail.com
 0009-0008-6365-1392

Thiago de Assis Furtado Silva – Undergraduate student – thiagodeassis.fs@gmail.com
 0009-0008-8834-2051

Conceição Maria Guedes Crozara – PhD – concaguedes@hotmail.com
 0000-0002-2213-972X

Jonas Byk – PhD – jonas.byk@hotmail.com
 0000-0001-5854-4316

Silvania da Conceição Furtado – PhD – silvaniafurtado@ufam.edu.br
 0000-0003-0065-3119

Correspondence
Sabrina Maciel Nascimento – Rua Juazeirinho, 21, Novo Aleixo CEP 69098-194. Manaus/AM, Brasil.

Participation of the authors
All authors participated together in the preparation of this article.