Revista Bioética
Print version ISSN 1983-8042 | On-line version ISSN 1983-8034

Rev. Bioét. vol.31 Brasília 2023

Artificial intelligence in health and bioethical implications: a systematic review

Mariele Abadia Elias [1], *Luciana Arruda Faversani* [1], *Josiane Aparecida Vieira Moreira* [1], *Anelise Viapiana Masiero* [2], *Natalia Veronez da Cunha* [1]

1. Universidade do Planalto Catarinense, Lages/SC, Brasil. **2**. University of Iowa, Iowa City, United States of America.

Abstract

The presence of artificial intelligence in healthcare is growing, helping in diagnosis and decision making. However, its application raises doubts, mostly related to ethics. This study aimed to identify its uses in health and its bioethical implications from a systematic literature review using the PRISMA guidelines. The ScienceDirect and Scopus databases were searched, using the descriptors "artificial intelligence," "bioethics" and "health." Works in English, published between 2017 and 2021 were considered, resulting in 102 articles found and, after applying the established criteria, 11 were selected. The studies reported on the bioethical principles of beneficence, non-maleficence, autonomy and justice, adding an element, explainability. Relationships were found between artificial intelligence in health and unpredictability, predictability, trust, physicians' role, systems development, privacy, data security, financial and social aspects. Developers, healthcare professionals and patients must maximize the benefits and limit the risks of tools that use this technology.

Keywords: Health care. Machine learning. Big data. Ethics.

Resumo

Inteligência artificial em saúde e implicações bioéticas: uma revisão sistemática

A presença de inteligência artificial na saúde vem crescendo, ajudando em diagnósticos e tomadas de decisão, mas suas implicações geram dúvidas relacionadas à ética. Esta revisão sistemática, baseada nas diretrizes Prisma, identificou os usos de inteligência artificial na saúde e suas implicações bioéticas. Foi realizada busca nas bases de dados Science Direct e Scopus usando os descritores "*artificial intelligence*", "*bioethics*" e "*health*". Trabalhos em inglês, publicados entre 2017 e 2021 foram considerados, resultando em 102 artigos. Após aplicação dos critérios estabelecidos, 11 foram selecionados. Os estudos discutiram os princípios bioéticos da beneficência, não maleficência, autonomia e justiça, adicionando o elemento explicabilidade. Inteligência artificial mostrou correlação com imprevisibilidade, previsibilidade, confiança, papel do médico, desenvolvimento de sistemas, privacidade, segurança de dados, e aspectos sociais e financeiros. Desenvolvedores, profissionais da saúde e pacientes devem maximizar os benefícios e limitar os riscos das ferramentas que usam essa tecnologia.

Palavras-chave: Atenção à saúde. Aprendizado de máquina. Big data. Ética.

Resumen

Inteligencia artificial en salud y sus implicaciones bioéticas: una revisión sistemática

El uso de la inteligencia artificial en salud va en aumento por facilitar el diagnóstico y la toma de decisiones, pero sus implicaciones plantean dudas relacionadas con la ética. Esta revisión sistemática desde las directrices Prisma identificó los usos de la inteligencia artificial en salud y sus implicaciones bioéticas. Las búsquedas se realizaron en Science Direct y Scopus utilizando los descriptores "*artificial intelligence*", "*bioethics*" y "*health*". De los trabajos en inglés publicados entre 2017 y 2021, se obtuvo 102 artículos. Aplicados los criterios, quedaron 11. Los estudios abordaron los principios bioéticos de beneficencia, no maleficencia, autonomía y justicia, añadiendo el elemento explicabilidad. La inteligencia artificial se correlacionó con la imprevisibilidad, previsibilidad, confianza, papel de los médicos, desarrollo de sistemas, privacidad, seguridad de los datos y aspectos financieros y sociales. Los desarrolladores, los profesionales sanitarios y los pacientes deben maximizar los beneficios y limitar los riesgos que involucra esta tecnología.

Palabras clave: Atención a la salud. Aprendizaje automático. Macrodatos. Ética.

The authors declare no conflict of interest.

Pesquisa

Artificial intelligence (AI) is being increasingly adopted in different areas. The term itself is difficult to define since this phenomenon depends on different factors [1]. However, despite the difficulty in defining it and its various concepts, the common understanding is that AI is associated with machines and computers to help humanity solve problems and facilitate work processes [2].

AI systems work using complex algorithms and large datasets that generate conclusions, thus replacing human reasoning with routine analysis [3]. To achieve human-level intelligence, AI needs guidance as a model of reality [1]. Machine learning (ML) is an AI system that can learn from models and eventually become autonomous, making decisions and generating conclusions that were previously considered only within the competence of the human mind [3].

The expansion of computing for storing, managing, accessing and processing data through a network of remote servers—such as cloud storage—has led to the expansion of AI applications for healthcare [4]. AI and ML have the potential to revolutionize the provision of service in this health [3]. These technologies can be used in health informatics, which is the business that describes the acquisition, storage, retrieval and use of information to improve patient care by interactions with the system [4]. Big data tools, that is, the storage and analysis of voluminous data, such as those used in healthcare, can also be used in association with AI [5].

AI has improved clinical diagnosis and decision-making performance in several domains of medical work [6]. In fact, these tools can help adapt public health programs, ensuring that relevant information is available for solid policy and decision making [4]. Automated medical imaging diagnosis is arguably the most successful domain of AI use in the medical field today. Many medical specialties, including radiology, ophthalmology, dermatology, and pathology, rely on image-based diagnoses [6].

However, it is important to note that patients recognize and are beginning to address the many issues raised by AI applications in healthcare [7]. These demands are legal, commercial, social and, mainly, ethical [8]. Designing these systems and using them is not a merely technical challenge, as it requires attention to bioethical principles [3].

Bioethics focuses on the relationship between living beings and, as AI emerges, humans must ethically engage with something that is not natural by itself, that is, with its own creation [2]. Concerns about the potential loss of control in the human-AI relationship are growing, such as the extent to what AI can or should support medical decisions or even make them on its own [1].

Although a key technology today, in many cases, it will be necessary to understand how a machine's decision was made and to evaluate the explanation for such a choice [1]. As AI advances, bioethical frameworks need to be adapted to address the problems these systems may pose, just as the development of these technologies needs to be adapted to incorporate bioethical principles [3]. Thus, this study aimed to carry out a systematic literature review to identify the applications of artificial intelligence in health and its bioethical implications.

Method

This study is a systematic review conducted according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA Statement) [9]. The Population, Intervention, Comparison, Outcome (PICO) strategy was used. This method is used in evidence-based practice and is recommended to structure the bibliographic search for evidence for reviews [10]. The PICO description used in this review is presented in Table 1 and was used to answer the question: what are the applications of artificial intelligence in health and its bioethical implications?

Table 1. Description of the PICO strategy

P	individuals, people
I	the use of artificial intelligence in health
C	not applicable
O	bioethical implications

The search for articles was performed manually in the ScienceDirect and Scopus databases in October 2021. The health sciences descriptors (DeCS/MeSH) "artificial intelligence," "bioethics" and "health" were used, interrelated by the Boolean locator ''AND''.

The following inclusion criteria were considered: scientific articles published between 2017 and 2021, in English and that had online access.

Books, book chapters, theses, dissertations, papers presented and published at events, review articles and editorial notes were excluded.

For the first selection, the titles and abstracts of the articles were read. The selected articles were read in full, strictly observing the inclusion and exclusion criteria, verification of duplicity and if they met the research theme. The articles selected for review were systematized in a framework for the analysis of the results. The stages of selection and full reading of the articles were performed by three independent reviewers.

Results

In total, 102 articles were identified in the databases following the search criteria. Of these, 72 were found in the ScienceDirect database and 30 in Scopus. After reading the titles and abstracts, 86 articles were excluded due to not meeting the proposed subject. One article was excluded because it was a duplicate. Thus, 15 potentially eligible articles were selected, of which 11 were selected after full reading. The selection procedure is shown in Figure 1.

Of the total selected publications, seven (63.64%) were accessed in ScienceDirect and four (36.36%) in Scopus. As for the publication year, seven (63.64%) of the articles were published in 2021, and one (9.09%) article was published each year for the remainder of the period included in the search, that is, between 2017 and 2020. In the analysis of the origin of the studies, most (45.45%) were developed in North America or in collaboration with institutions in that continent.

Regarding the type of study, most are cross-sectional studies (45.45%), but prospective studies (9.09%), case studies (9.09%) and descriptive studies (9.09%) were also found, as well as exploratory studies (9.09%), randomized clinical trials (9.09%) and a multidimensional approach study (9.09%).

Based on the 11 selected publications, a table was created to analyze the objectives, type of study, main results and conclusions (Table 2).

Figure 1. Flowchart of the study selection process for the systematic review

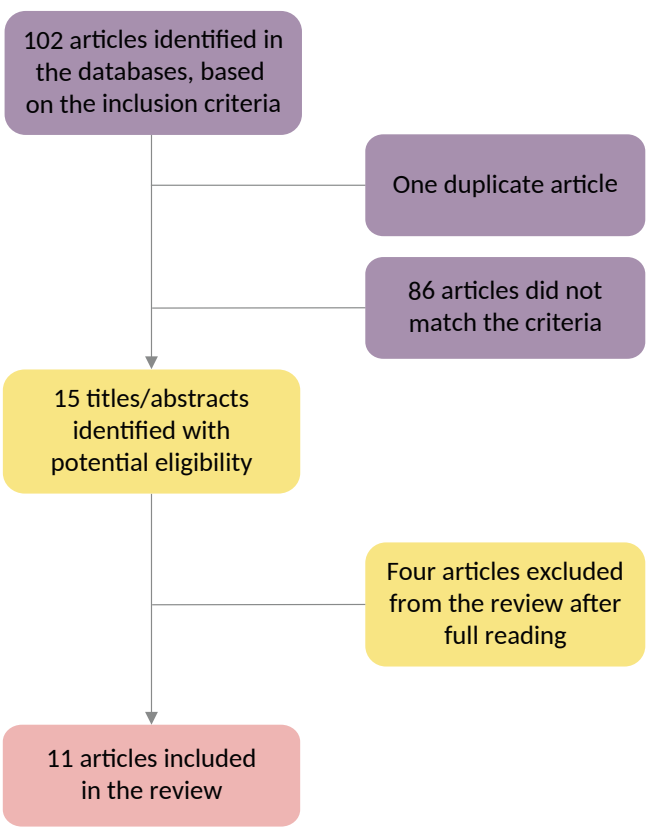

Table 2. Analysis of the objective, type of study, main results and conclusions of the selected publications

Authors; year	Main objective	Type of study	Main results	Conclusions
Fernandes and collaborators; 2019 [11]	Provide results that can support the study of medical research in the decision-making process in clinical bioethics, particularly in cases of euthanasia.	Cross-sectional	Data processed by trait selection methods were used to create models capable of predicting the euthanasia decision using ML and an eye tracker. Statistical experiments showed that the predictive model resulting from the multilayer perceptron (MLP) algorithm led to the best performance. Interesting results (standards and rules) for bioethical decision making were extracted from simulations with MLP models. Some participants made a rational decision, respecting the code of ethics of nursing professionals and the Brazilian penal code, where euthanasia is considered homicide. Others considered the emotional aspect, linking this decision to the patient's suffering.	The good performance presented by the predictive model demonstrates that the proposed investigation approach can be used to test scientific hypotheses related to visual attention and decision making, verifying the extent to which vision is a determining factor in decision making, particularly in clinical bioethics when dealing with end-of-life issues.
Silva, Lehoux, Hagemeister; 2018 [12]	Assess whether an innovation qualifies as responsible innovation in health using a tool developed in three stages: screening, evaluation and classification; and discuss the political aspects of using the tool.	Prospective	The screening and assessment tool for responsible innovation in health that was developed was judged by experts and, after the second round of comments on the topic, a consensus was reached for 16 of the 20 questions regarding the importance, clarity and adequacy of the tool structure. The sustainability of health systems is harmed by the current way in which health innovations are designed and brought to market. Consensus was reached on most of the criteria, attributes and scales of the tool. The future use of the tool can contribute to the development of innovations that provide greater social value.	The development of this tool will help fill an important knowledge and policy gap by clarifying decisions made at an early stage by innovation stakeholders such as investors, technology developers, research funding agencies and policymakers.
Lysaght and collaborators; 2019 [13]	Address and analyze the ethics framework for big data in health and research to demonstrate how decision-making can be based on it for the development and implementation of AI-assisted support systems in health ethically and responsibly [5].	Case study	Clinical decision support systems (CDSS) are programs that generate health information. The ones that use ML and AI are complex and the reason why physicians must be trained to improve their clinical decision-making skills when using this type of resource. CDSS AI algorithms can reinforce social prejudices but also bring benefits as more efficient public health systems. The inclusion and analysis of data happens on an ongoing basis, contributing with information about appropriate future practice. Feeding the CDSS with patient data may cause conflicts about the dual role of physicians—care and research. The substantive values listed are professional integrity and fairness; and the procedural values are transparency and accountability. AI-assisted CDSS must be explainable. The final decision must be made by the health professional and it must be considered that there may be moral judgments that the program is uncapable of making. The case study addresses the use of these programs in an intensive care unit. Despite bringing financial benefits, this tool could generate ethical issues such as economy over health, distrust in recommendations and concern about responsibility.	Given the rising costs of healthcare, the development and implementation of AI assistance in clinical decision-making is likely to be unavoidable. Values of professional integrity and responsibility will play a more prominent role at patient care level, whereas values of fairness and the potential for harm to the group must be balanced with imperatives of public benefit at the societal level. Public benefit imperatives at the social level must be used for balance. Transparency affects both trust in the medical profession and health systems.

continues...

http://dx.doi.org/10.1590/1983-80342023542PT

Table 2. Continuation

Authors; year	Main objective	Type of study	Main results	Conclusions
Cawthorne, Robbins-Van Wynsberghe; 2020 [14]	Create an ethical framework to be applied during the design, development, implementation and evaluation of drones in public health.	Descriptive	The hierarchy of values used consists of ethical principles, human values, standards and design requirements. For the creation of the structure, the four principles of bioethics were considered, in addition to a fifth principle of AI ethics: explainability. Benevolence in the field of health drones can be translated into values of human (and non-human animal) well-being, human jobs and skills, and environmental sustainability. Nonmaleficence encompasses privacy, security, protection, tranquility, jobs and human skills, and environmental sustainability. Autonomy includes free will, human values, responsibility and trust. Justice includes the equitable distribution of benefits and damages. The adoption of health drones can lead to the reduction of local health infrastructure, reducing personal assistance. However, they can also connect people in remote places to modern services. Explainability deals with the ease with which systems can be understood. The use of an ethical framework is especially useful for those with limited experience in technology ethics.	Ethical principles are abstract and need more contextualization and specification for reflection. The creation of this ethical framework reinforces the value of integrating ethics into practice and serves as a model for design and development in drone and non-drone domains. The framework helped to identify and refine potential benefits and mitigate risks.
Antes and collaborators; 2021 [15]	Develop a new measure and assess the openness and extent of concerns and perceived benefits regarding AI-based health technologies in a sample of adults in the United States.	Cross-sectional	Participants were moderately open to AI-based healthcare technologies, but there was variation depending on the type of application. The trust in healthcare system and technology were the strongest and most consistent correlates of openness, concern and perceived benefit. Older participants were less open to technologies and men were more open than women. Full-time employment was associated with greater openness and less concern. The two technologies that made predictions about serious illness—heart attack risk and the probability of survival from cancer—were the technologies best evaluated.	The openness of the participants seems tenuous, suggesting that early promotion strategies and experiments with new AI technologies can strongly influence opinions on the subject. Addressing trust in targeting the acceptance of these innovations in healthcare may be needed.
Batlle and collaborators; 2021 [16]	Understand the best practices in sharing patient data in healthcare institutions.	Exploratory	Five broad domains of important activities for collaboration using patient data were identified by a working group: privacy, informed consent, standardization of data elements, supplier contracts, and data evaluation. The methods and ethical understanding of commonly used legal frameworks for these purposes were presented, as well as data flow design that can help inform how permissions are created and revoked. A description of the careful preparation and annotation of datasets is needed when discussing anonymity and de-identification in the zeal for privacy, technically pointing out the difficulties. The volume of data required for the preparation of the AI algorithm is very high and therefore the premise is that the preparation of such data takes place in a safe and shareable way with its owners (patients).	Creating a data sharing relationship involves ethical and information technology complexity. Patient anonymity and privacy maintain trust and protect entities seeking to safely share data.

continues...

http://dx.doi.org/10.1590/1983-80342023542PT

Table 2. Continuation

Authors; year	Main objective	Type of study	Main results	Conclusions
Green and collaborators; 2021 [17]	Develop tools integrated into digital health systems to support shared decision-making and optimize preparation for treatment in chronic kidney disease.	Randomized clinical trial	Using the tools, 243 (24%) of 1,032 patients in four nephrology clinics were identified as high risk for progressing to renal failure within two years. Kidney transition specialists enrolled 117 (48%) high-risk patients until the end of the first year of research. Nurses used the app for 100% of patients to document 287 planning steps for renal replacement therapy. All kidney transition experts (100%) rated the ease of use and usefulness of the tool, agreeing or strongly agreeing with all items.	Nurses reported that the tools developed facilitated the identification of patients who need support and their navigation activities. And the fast identification of patients who need shared and informed decision-making and their preparation for renal replacement treatments.
Martinho, Kroesen, Chorus; 2021 [18]	Obtain information about patterns of reasoning and moral opinions about health AI from people involved in medical practice.	Cross-sectional	Based on physicians' questions about ethics around health AI, four main perspectives were identified: 1. AI is a useful tool: let physicians do what they are trained to do. 2. Rules and regulations are crucial: private companies are all about money. 3. Ethics is enough: private companies can be trusted. 4. Explainable AI tools: learning is necessary and inevitable. All perspectives consider that physicians should participate in the design process of AI health technologies, contributing to explainability. Physicians are more concerned with the role of large companies in the health area and less aware or concerned with issues such as equity, prejudice and inequalities in health.	Each perspective provides valuable and often contrasting insights into ethical issues that must be operationalized and taken into consideration in the design and development of AI in health.
Shen and collaborators; 2021 [19]	Analyze highly portable magnetic resonance imaging (MRI) surveys in remote and resource-limited international settings for creating ethical and legal guidance in a complex global landscape.	Cross-sectional	It is necessary to ensure that local communities are partners in the research enterprise and to guarantee the local social value of the research. Field MRI studies need to be responsible for the safety of participants and all around them. It is important to pay attention to data privacy and security regulations (local and international). It is necessary to identify if the sample where the AI model was trained was diverse so that the predictions are more accurate, considering a diversity of factors. Exam results need to be communicated to the participants in an enlightening way and when it comes to incidental findings, a challenging issue is how to provide clinical support and referral in these remote communities.	More affordable and portable MRI scanners provide opportunities to address unmet research needs and health inequities in remote and resource-limited international settings. Local communities must be continuous partners in the co-creation of knowledge. Research must produce local value to justify the risks and minimize the possibility of abuse.

continues...

http://dx.doi.org/10.1590/1983-80342023542PT

Table 2. Continuation

Authors; year	Main objective	Type of study	Main results	Conclusions
Spiegel, Barker, Kistnasamy; 2021[20]	Describe and evaluate the application of AI in the development of computer-aided diagnostics to support the more efficient adjudication of claims for former gold miners with occupational lung disease in Southern Africa.	Cross-sectional	The results were correlated with the principles of bioethics. Beneficence: AI could provide more consistent judgment than multiple professionals with varying skill levels. Nonmaleficence: maintain data privacy and security and avoid the infiltration of certain trends in decision-making systems. Autonomy: AI can lead to losses in professional qualification, as professionals come to trust technology more. Protocols must be established and professionals trained, so false negatives or false positives are identified. Justice: when market demand is weak for investments in AI and where public institutions have not responded to its use, failure to handle the technology can be a way of maintaining inequalities. This highlights the timely use of innovations to benefit those in need.	Efforts to overcome technical challenges in the application of AI must be followed from the beginning to ensure its ethical use.
Stahl and collaborators; 2021[21]	Theoretically capture and empirically measure the benefits and disadvantages of AI for human progress, beyond the principles for machine learning; counterbalance the technical and economic benefits of AI and its legal, social and ethical aspects.	Multidimensional approach	The AI ethics discourse is discussed in three streams: 1. Issues related to ML application: it is difficult to predict to what degree data will be used for the given purpose—as a personal profile leads to a classification for purposes other than the initial one. 2. Social and political issues arising in a digital society: these systems require access to large data amounts for training and validation purposes, which generates distrust related to the autonomy of machines, replacement of humans by machines, injustice in the distribution of costs and benefits and data control, as well as the consequences. 3. Metaphysical questions about the nature of reality and humanity: these concern what machines should be allowed to decide autonomously. Economic consequences, employment, justice, freedom, human contact, individual autonomy, inequality, integrity, property, military use, power asymmetry, responsibility and sustainability fall into this category.	There is currently no consensus between the various approaches to governance and information security. Human rights legislation can resolve many social and ethical issues. Complexity of contexts and scenarios are factors that pluralize the ways of approaching them and allow scholars or professionals to keep an overview. Attention in the current use of AI and ML and, to some degree, in broader socio-technical systems.

Discussion

Although new technologies that use AI hold great promise for healthcare, ethics and human rights must be at the center of their design, deployment and use [8]. The incorporation of bioethical principles in the development of these technologies can help protect patient rights, minimize risks, establish responsibilities, and institute robust metrics to study their effectiveness and benefits [3].

From this point of view, some studies selected in this review established a direct relation between their results and the principles of bioethics: beneficence, nonmaleficence, autonomy and justice [14,20]. Another principle considered in the ethics of AI is explainability, which deals with the understanding of tools that use this technology [13,14,18]. The elaboration of these systems must consider the principles of beneficence, maintenance of values, responsibility and transparency—which comprises explainability [2].

Explainability is fundamental because the necessary and lasting trust can be built from it to promote the acceptance of AI with future users [1]. In one study, trust in the healthcare system and technology were the strongest and most consistent correlates of openness, concern, and perceived benefit [15]. In another survey, participants felt uncomfortable relying solely on recommendations made by an AI tool, without being able to directly assess its rationale [7]. Although systems are important in the decision-making process, experts must be able to understand and redo this process [1].

Even experts in AI development cannot determine how inputs are transformed into outputs, that is, how a personal profile generates a decision [21]. Unpredictability in ML thus becomes the key to the discussion, as it is a method for automating data analysis using algorithms that iteratively identify patterns in data and learn from them [4]. The consequences are impossible to measure when the machine is programmed to learn by itself [21].

What should machines be allowed to decide autonomously and who is responsible for the decision? There are no answers to these questions yet. For now, physicians must participate and understand AI, and then decide autonomously, though based on the AI recommendation [21].

Like unpredictability, prediction is another relevant point discussed in AI. A tool capable of predicting the risk and prioritizing patients for the treatment of chronic kidney disease helped in the rapid identification of individuals in need of renal replacement [17]. Another study evaluated the reaction and decision of nurses about the practice of euthanasia in terminal patients using eye-tracking technology, bringing a division between rational and emotional decisions, demonstrating the effectiveness of the prediction model [11]. Even when AI indicates treatments, health professionals and patients can make moral judgments that the program is uncapable of producing [13].

Despite any indications made by the AI, the final decision regarding health must be made by a professional [13]. A study conducted on patients who had been in a primary care unit found that they believe that their physicians should protect them from harm resulting from AI errors, so that the final decision should rest with the physician and the health insurance company [7]. In this regard and considering malpractice, insurance companies must be clear about coverage when decisions are made by AI systems, even if partially [6].

AI in the healthcare context is likely to give physicians more time for other tasks, such as establishing direct contact with their patients [6]. However, AI may cause conflicts about the role of physicians, who must provide care while feeding the system with patient data for support and clinical research purposes [13]. As AI is implemented for specific clinical tasks, the roles of health professionals will continue to evolve since various AI modules will be incorporated into care [6].

Professional integrity is one of the three main ethical values for big data in decision making since it encompasses responsibility for the patient, meaning medical training becomes central [13]. Physicians will need to adapt to their new roles as integrators and interpreters of information and as supporters of patients, and the medical education system will have to provide them with the necessary tools and methods [6]. AI tools are useful, but physicians have not been trained to understand them and, at the same time, their active participation in its development is vital [18].

To face the challenges, AI researchers and clinicians need to work together to prioritize and develop applications that meet crucial clinical needs [6]. Given the potential risk of causing harm, involving bioethicists in the design of these technologies is necessary [3]. Multidisciplinary and multisectoral collaborations will be needed to facilitate the development and deployment of AI applications in the medical field [6].

Several datasets must be manipulated to develop AI tools, meaning this technology can only advance with the use of big data [5]. To develop and train algorithms, data from health records need to be accessed, and patients are not always aware that such information is being shared [16]. This dilemma requires an in-depth study of these emerging technologies, and the evolution of bioethical principles of patient privacy and confidentiality [3].

Discussions around the ethics of electronic health records and AI have primarily focused on privacy, confidentiality, data security, informed consent and data ownership; however, the relevance of each varies depending on differences in culture, literacy, relationships. patient-provider relationship, available infrastructure and the regulations of each country [4].

Despite the many discussions about the ethics of big data in a variety of contexts, little guidance is available on what values are at stake and how decisions should be made in an increasingly complex healthcare and research environment [5]. Anonymity and patient privacy are essential to create trust, in addition to the need to list permission levels in the connections between health institution, data owner (patient) and related for-profit institutions (health plans, providers and treatment/exam partners) [16].

Several complex ethical issues arise when considering the use of big data and, therefore, an ethical framework to address them and guide actions is important [5]. Likewise, such ethical planning must be applied in other projects that involve AI in health, such as a structure to be used in all stages of obtaining drones in public health, involving ethical principles, human values, design standards and requirements [14].

The role of big data and AI companies, a profit-oriented segment, is questioned given ethical implications in the organization of health [18]. The sustainability of health systems is currently undermined by how health innovations are designed and brought to market undermines [12]. The financial aspects and the power concentrated in the owners of data and technologies seem to be more relevant than those related to equity, prejudice and inequalities [18].

Who will control or profit from the application of AI remains to be determined, so, the priority must be the balance between regulatory safeguards and market forces to ensure that patients benefit [6]. The creation of regulations is a fundamental point in the evolution of this topic [18].

Despite holding immense potential to correct human errors and improve care delivery, ML-based AI applications can also aggravate biases [3]. Patients are concerned that AI tools may reinforce existing biases, which can occur if a learning dataset is biased or developers unintentionally incorporate their own bias into an algorithm [7], which is already a possibility with CDSS AI algorithms [13].

AI can be transformative in social aspects for public health in countries with few resources, because as these locations become more connected and create higher quality data, the ability of AI tools to address health challenges is likely to increase [4].

Portability is another important aspect when discussing AI availability in remote locations, for example, more affordable and portable MRI scanners offer opportunities to address unmet health needs and inequities in remote and resource-limited settings [19]. The infrastructure required in these environments demands substantial investments to be implemented, thus not always having the necessary access to upload large datasets to cloud systems [4]. In this case, data privacy and security regulations (local and international) and how to provide support and clinical referral in these communities located in remote locations become relevant [19].

In the study of AI in computer-assisted diagnosis, claims for compensation for mineworkers with occupational lung cancer were accelerated [20]. Even with low demand for public and private investments and a high risk of disqualification of the tool in legal spheres, this approach provides a proven social benefit when ethical protocols and monitoring are part of the process from the beginning.

Two predictive AI technologies were better evaluated among older adults and individuals with lower digital literacy, demonstrating that relief for serious diseases such as heart attack and cancer, cases involving urgency, vulnerability and risk, can be accepted under a reliability premise [15].

While patients are generally excited about the possibility of AI improving their care, they are also concerned about the safety and oversight possibilities [7,15]. Concerns that AI tools can increase healthcare costs and that such costs may be passed on to patients are also raised [7]. However, since medical decisions are made by automated systems, cost-effectiveness and expense rationing can be prioritized [3].

The use of AI in clinical decision making is probably inevitable given the rising costs of healthcare [13]. Prospective studies will be able to better identify the weaknesses of AI models in heterogeneous clinical environments and indicate ways to integrate them into current clinical workflows [6]. However, efforts to overcome technical challenges in AI deployment must be followed from the beginning by ethical aspects [20].

Corroborating the literature, evident benefits and risks in the use of AI in health exist. In order to maximize the benefits to the public interest and limit the threats, the World Health Organization proposes six principles for the use of AI and that focus on the topics addressed in this discussion, namely: 1) protecting human autonomy; 2) promoting well-being, security and the public interest; 3) ensuring transparency, explainability and intelligibility; 4) promoting responsibility and accountability; 5) ensuring inclusion and equity; and 6) promoting technologies that are responsive and sustainable [8].

Final considerations

Understanding AI in healthcare requires understanding its steps. Data feeding, development of algorithms, and decisions come out; in this process, the four principles of bioethics, added to those of explainability and human rights legislation, are the guarantor constructs of ethics in the process.

The guarantee of anonymity and consent for the use of data, without commercial purposes holding greater relevance in AI implementation, remain diverse and legally regionalized in research, diagnosis and treatment procedures. Transparency is, therefore, the main foundation in AI in healthcare and the questions are broader than the answers.

Understanding how decisions were programmed, taking them out of the black box of incomprehensible technology, is the central aspect. When the machine learns by itself, can the physician be held accountable for the decisions? Regulations need to evolve, both in defining responsibilities and in understanding and trusting the autonomy of these decisions generated by AI. Permissions for machines to decide on their own, based on their own learning, is an evolving concept, given that not even specialists in this development can foresee its consequences, both in terms of assertiveness and reliability, as well as in maintaining the initially defined use.

What can be said is that we are at a path of no return where all the issues raised in this article will need to be regulated considering access, privacy, social issues and justice, still being touched by interests of control, power and profitability.

This study did not receive any specific funding from public, commercial or non-lucrative agencies.

Pesquisa

References

1. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. Wiley Interdiscip Rev Data Min Knowl Discov [Internet]. 2019 [acesso 5 nov 2021];9(4):e1312. DOI: 10.1002/widm.1312
2. Tai MCT. The impact of artificial intelligence on human society and bioethics. Tzu Chi Med J [Internet]. 2020 [acesso 5 nov 2021];32(4):339-43. DOI: 10.4103/tcmj.tcmj_71_20
3. Nabi J. How bioethics can shape artificial intelligence and machine learning. Hastings Cent Rep [Internet]. 2018 [acesso 5 nov 2021];48(5):10-3. DOI: 10.1002/hast.895

4. Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? BMJ Glob Heal [Internet]. 2018 [acesso 5 out 2021];3(4):e000798. DOI: 10.1136/bmjgh-2018-000798
5. Xafis V, Schaefer GO, Labude MK, Brassington I, Ballantyne A, Lim HY *et al*. An ethics framework for big data in health and research. Asian Bioeth Rev [Internet]. 2019 [acesso 5 nov 2021];11(3):227-254. DOI: 10.1007/s41649-019-00099-x
6. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng [Internet]. 2018 [acesso 5 nov 2021];2(10):719-731. DOI: 10.1038/s41551-018-0305-z
7. Richardson JP, Smith C, Curtis S, et al. Patient apprehensions about the use of artificial intelligence in healthcare. NPJ Digit Med [Internet]. 2021 [acesso 5 nov 2021];4(1):140. DOI: 10.1038/s41746-021-00509-1
8. World Health Organization. Ethics and governance of artificial intelligence for health [Internet]. Geneva: WHO Guidance; 2021 [acesso 4 nov 2021]. Disponível: https://bit.ly/3Ry4M8P
9. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ [Internet]. 2009 [acesso 5 set 2021];339(7716):332-6. DOI: 10.1136/bmj.b2535
10. Santos CMDC, Pimenta CADM, Nobre MRC. A estratégia PICO para a construção da pergunta de pesquisa e busca de evidências. Rev Latinam Enferm [Internet]. 2007 [acesso 5 set 2021];15(3):508-11. DOI: 10.1590/S0104-11692007000300023
11. Fernandes DL, Siqueira-Batista R, Gomes AP, Souza CR, Costa IT, Cardoso FSL *et al*. Investigation of the visual attention role in clinical bioethics decision-making using machine learning algorithms. Procedia Comput Sci [Internet]. 2017 [acesso 11 out 2022];108:1165-74. DOI: 10.1016/j.procs.2017.05.032
12. Silva HP, Lehoux P, Hagemeister N. Developing a tool to assess responsibility in health innovation: results from an international delphi study. Heal Policy Technol [Internet]. 2018 [acesso 11 out 2022];7(4):388-96. DOI: 10.1016/j.hlpt.2018.10.007
13. Lysaght T, Lim HY, Xafis V, Ngiam KY. AI-assisted decision-making in healthcare. Asian Bioeth Rev [Internet]. 2019 [acesso 11 out 2022];11(3):299-314. DOI: 10.1007/s41649-019-00096-0
14. Cawthorne D, Robbins-van Wynsberghe A. An ethical framework for the design, development, implementation, and assessment of drones used in public healthcare. Sci Eng Ethics [Internet]. 2020 [acesso 11 out 2022];26(5):2867-91. DOI: 10.1007/s11948-020-00233-1
15. Antes AL, Burrous S, Sisk BA, Schuelke MJ, Keune JD, DuBois JM. Exploring perceptions of healthcare technologies enabled by artificial intelligence: an online, scenario-based survey. BMC Med Inform Decis Mak [Internet]. 2021 acesso 11 out 2022];21(1):221. DOI: 10.1186/s12911-021-01586-8
16. Batlle JC, Dreyer K, Allen B, Cook T, Roth CJ, Kitts AB *et al*. Data sharing of imaging in an evolving health care world: report of the ACR Data Sharing Workgroup, part 1: data ethics of privacy, consent, and anonymization. J Am Coll Radiol [Internet]. 2021 [acesso 11 out 2022];18(12):1646-54. DOI: 10.1016/j.jacr.2021.07.014
17. Green JA, Ephraim PL, Hill-Briggs F, Browne T, Strigo TS, Hauer CL *et al*. Integrated digital health system tools to support decision making and treatment preparation in CKD: the PREPARE NOW Study. Kidney Med [Internet]. 2021 [acesso 11 out 2022];3(4):565-75.e1. DOI: 10.1016/j.xkme.2021.03.009
18. Martinho A, Kroesen M, Chorus C. A healthy debate: exploring the views of medical doctors on the ethics of artificial intelligence. Artif Intell Med [Internet]. 2021 [acesso 11 out 2022];121(October):102190. DOI: 10.1016/j.artmed.2021.102190
19. Shen FX, Wolf SM, Bhavnani S, Deoni S, Elison JT, Fair D *et al*. Emerging ethical issues raised by highly portable MRI research in remote and resource-limited international settings. Neuroimage [Internet]. 2021 [acesso 11 out 2022];238:118210. DOI: 10.1016/j.neuroimage.2021.118210
20. Spiegel JM, Ehrlich R, Yassi A, Riera F, Wilkinson J, Lockhart K *et al*. Using artificial intelligence for high-volume identification of silicosis and tuberculosis a bio-ethics approach. Ann Glob Heal [Internet]. 2021 [acesso 11 out 2022];87(1):58. DOI: 10.5334/aogh.3206
21. Stahl BC, Andreou A, Brey P, Hatzakis T, Kirichenko A, Macnish K *et al*. Artificial intelligence for human flourishing: beyond principles for machine learning. J Bus Res [Internet]. 2021 [acesso 11 out 2022];124:374-88. DOI: 10.1016/j.jbusres.2020.11.030

Mariele Abadia Elias – Graduate – marieleabadia@uniplaclages.edu.br
0000-0001-5724-8400

Luciana Arruda Faversani – Graduate – luarrudaf@uniplaclages.edu.br
0000-0002-5495-1122

Josiane Aparecida Vieira Moreira – Graduate – josiane@uniplaclages.edu.br
0000-0003-2683-5172

Anelise Viapiana Masiero – PhD – anemasiero@uniplaclages.edu.br
0000-0003-3203-5876

Natalia Veronez da Cunha – PhD – nat_cunha@uniplaclages.edu.br
0000-0002-8522-5836

Correspondence
Mariele Abadia Elias – Av. Marechal Castelo Branco, 170, bloco 2, Universitário CEP 88509-900. Lages/SC, Brasil.

Participation of the authors
Mariele Abadia Elias, Luciana Arruda Faversani and Josiane Aparecida Vieira Moreira were responsible for the conceptualization, method, research, writing and preparation of the original article. Anelise Viapiana Masiero and Natalia Veronez da Cunha were responsible for preparation, revising and editing of the article.

Received: 4.7.2023
Revised: 7.6.2023
Approved: 10.12.2023