

Published: August 31, 2023

Citation: Sridhar GR and Lakshmi G, 2023. Ethical Issues of Artificial Intelligence in Diabetes Mellitus, Medical Research Archives, [online] 11(8).

<https://doi.org/10.18103/mra.v11i8.4287>

Copyright: © 2023 European Society of Medicine. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI

<https://doi.org/10.18103/mra.v11i8.4287>

ISSN: 2375-1924

RESEARCH ARTICLE

Ethical Issues of Artificial Intelligence in Diabetes Mellitus

Gumpeny Ramachandra Sridhar, MD DM FACE FRCP*¹;
Gumpeny Lakshmi MD²

¹ Endocrine and Diabetes Centre, Visakhapatnam, India

² Associate Professor of Medicine, Gayatri Vidya Parishad Institute of Healthcare & Medical Technology, Visakhapatnam, Andhra Pradesh, India

*Corresponding author: Email: grsridhar@hotmail.com

ABSTRACT

Artificial intelligence (AI) has permeated various branches of clinical medicine, with promising applications in predictive, diagnostic and therapeutic areas. Digital innovations are increasingly useful in the management of non communicable diseases in the form of tracking applications, data collection systems (EMRs) and wearable sensors. Globally, diabetes mellitus is the most common and serious non communicable disease. There is a mismatch between the people with diabetes and the number of healthcare professionals needed to manage them. Therefore, artificial intelligence has the potential to play a significant role in addressing the unmet need. Majority of AI applications have been developed for the diabetes population. Ethical issues arising from the application can be carried over to its application in other areas of clinical medicine. Among diabetic population, artificial intelligence has been prominently employed in screening for diabetic retinopathy. Continuous glucose monitoring and insulin pumps are other areas of application. Data collection and sharing through AI media can ease the burden of poor doctor-patient ratio, and improve efficacy of treatment. Despite its advantages, and the fact that citizen juries have been found to be favourable towards the use of AI in research and treatment, certain drawbacks continue to exist. With the threat of data theft and breach of privacy, due diligence must be given to ethical and legal aspects to protect the patient. It is acknowledged that AI can facilitate the decision making process but not entirely replace a physician's role. With able governing laws, systems to protect safety, minimize bias and improve transparency, AI and precision medicine could help control the burden of disease.

Keywords: Privacy, security, reliability, safety, fairness, transparency, accountability, retinopathy

Introduction

Artificial intelligence (AI) is a phenomenon whose time has come in clinical medicine, and particularly non-communicable diseases. Burgeoning data from multiple sources and enhancements in computing power leads to innovative paths where their confluence is rapidly impacting clinical care. AI promises to aid prediction, diagnosis, and management of diseases (1). Such is the scope of the phenomenon to answer questions beyond the capacity of humans (2) excitement at the pace that it took a while to recognize the unintended consequences (3) and thereby the ethical implications of the new technologies (4). Innovations are necessary to tackle the problems of non-communicable diseases (NCDs) and their complications. Digitization is a promising path which makes artificial intelligence (AI) possible. Digital data is captured from a variety of sources including electronic medical records (EMRs), and wearable sensors [5]. This has led to questions regarding ethics and legality in the use of AI in healthcare sector. Among non-communicable diseases, diabetes mellitus is predominant on a global scale. Many applications of AI have focused on diabetes; the first FDA approval was given for use of AI based digital retinal imaging to detect diabetic retinopathy. The principles of ethical issues in AI in diabetes can be broadly applied to other areas of clinical medicine.

AI has been defined as the 'use of computer-processing capabilities of symbols to find generic methods for automatic perceptual, cognitive and manipulating activities via algorithms'[5]. Due to the burgeoning availability of data which cannot be meaningfully interpreted by individuals or even by traditional statistical methods, there is need for more complex algorithmic computer analysis [6].

With increasing prevalence of NCDs, and insufficient healthcare members, AI could provide be a potential solution [7]. Advantages of digital data sharing are well known: health research is made possible across geographic boundaries to improve patient care [8]. Secondary use of data from EMRs and its relation to financial aspects of AI raise ethical and legal issues [9,10].

However, concerns about ethical malfeasance must not obscure the potential benefits of AI [11]. It can provide economic benefits by improving efficiency and productivity of healthcare. Thereby, AI can be more consistent than human beings and improve quality of work. 'Good' AI straddles shared ethical values such as 'benevolence, security, achievement and self-direction' [11]. Other benefits of AI include enhancing medical knowledge to improve care,

making medical expertise accessible to non-specialists, automating repetitive tasks and finally as an aid in equitable distribution of scarce resources [12].

AI also allows democratization of medical expertise so that generalists can access some of the skills of specialists. The most advanced example is in the recognition of actionable diabetic retinopathy by machine learning tools. A large number of people can be screened, so that only a small subset is referred for specialist ophthalmological care. This lessens the drudgery of repetitive work by specialists. AI pre-screening helps to triage routine tasks, leaving the specialist to deal with only those which needs their specific expertise.

As enticing as the potential of AI in diabetes clinical care is, the industrialization of AI requires that societal issues involving ethics and legality are addressed [6]: privacy, anonymity, fairness, explainability and interoperability are paramount.

These are similar to ethical aspects of genetic and genomic research involving biobanks[13], and of secondary usage of health data [8,9]. Recognized tensions exist in incentives and benefit, harm to groups, power structures and engagement by the researcher along with sharing of responsibility [8].

In view of the potential impact of AI applications in health care, a scoping review was recently carried out to understand what aspects of AI in health care are studied, from publications in MEDLINE, Scopus, Web of Science, CINAHL and PsycINFO databases [14]. Among a potential set of 9218 records, less than 0.5% (45/9218; 0.49%) met the inclusion criteria. The results were revealing: most were from high-income nations (33/45;73%) and were directed at care providers (25/45; 56%). Most studied clinical care aspects, involving support in decision-making by individual doctors. They were aimed at technical and computational aspects and in establishing effectiveness of AI interventions. Not much attention has been devoted to issues of building trust, addressing transparency ethics and developing explainability [14].

Ethical and legal aspects of artificial intelligence in diabetes health care

PRIVACY AND SECURITY

Privacy and security of data usage are important both ethically and legally. Issues may arise when there is an interaction between public sector, which has the data in electronic format and the private sector which operates through computing and

telecommunication channels [6]. Whereas anonymity and privacy are critical in healthcare, data availability in EMRs puts both in jeopardy. Adequate legislation is necessary to balance privacy and anonymity, and to ensure there are no breaches. One way of working around is to produce decentralized analytical solutions, which may be compromised with the involvement of corporate information technology in medical field.

Aspects to be considered include (a) data transference, transparency and openness (b) transfer of data must be proportional to the medical task (c) governance, not merely legislation must be created or strengthened where necessary [6]. Progress could be made in 'differentiated data protection,' which refers to blurred set of anonymized data for development rather than individual data with attendant risks of being traced back to a specific person [15].

Informed consent for devices or procedures employing AI is a formidable challenge. This could be a bottleneck in integrating AI with clinical practice [10]. Non-AI informed consent requires the clinician to educate the patient about the procedure used, the risks and benefits so that the patient in turn can make an informed choice. AI is far more complex, posing difficulties for the clinician to understand the concept in the first place. The opaque variables comprise the machine learning method employed, the data on which ML is trained as well as the potential biases in data. Should the clinician even notify the patient that AI is being used? All these are questions still awaiting answers [10].

In addition, health apps and chatbots are being released for a variety of day to day applications including diet advice, adherence and wearable sensors. Informed consent in these applications is often obtained online, without a face-to-face interaction between the patient and the clinician. In addition, software updates make informed consent even more complicated.

The borders of medical AI, in terms of what it is and what it can do are not clearly defined. The role of AI in medical care must therefore be defined and circumscribed [12], which requires proper framework of governance to protect human subjects from harm [7].

Diabetic retinopathy (DR) is the most prominent area where AI is commercially employed, since 2018. Obtaining ethical consent in screening for DR is of practical importance. Ursin et al developed a checklist of items to be informed in the diagnosis of

DR using AI in primary care [12]. There were few records which met the inclusion criteria among articles from PubMed and Web of Science database. Requirements for the general practitioner were summed up as the need for their being informed about ethics of new technologies to communicate to patients. Contrariwise, patient's blind belief or fears must be allayed by informing them of risks, drawbacks and potential benefits [16]. These allow the better implementation of AI for the common good.

A citizens' juries were conducted to extend the knowledge about the controls citizens would seek for using EHRs in research, at baseline and after a deliberative process [9]. Among 34 jurors, at the end of the process, 33 jurors supported secondary use of data for research. Twenty-four desired that individuals have the choice of opting out, six the choice to opt in and three for use of all records without need for further consent. In terms of who gets access to data, public benefit was key in the jurors' opinion [9]. When informed of the benefits and risks associated with secondary sharing of EMRs, it was considered that citizen's privacy rights must not hinder research that can lead to societal benefit.

Surveillance is another contentious topic in AI. As of March 10th, 2022 a review identified 3556 scholarly publications obtained by using as keywords AI and surveillance in the Scopus database. The aim was to map scholarly studies on the subjects carried out by social sciences and humanities scholars [13]. Among the seven scholarly subjects identified, public health surveillance in the context of privacy and contact tracing apps during the pandemics was one. However, there were porous borders between dichotomous forms of surveillance. The author concluded that future research is necessary to compare the risk benefit consequences of AI surveillance.

Cybersecurity is another important aspect of AI in healthcare. Widespread use of internet of things (IoT) will increase the potential for breaches in cybersecurity. Particular attention should be paid to servers in hospitals, diagnostic centers and wearable devices. Trojan viruses may result in incorrect or even harmful treatment. Cybersecurity Act in European Union seeks to achieve robust cyber resilience, cybersecurity and trust. In the US, Cybersecurity and Infrastructure Security Act of 2018 (2018) was signed into law in November 2018 [10].

Despite these, oversight and regulation are likely to fall behind technologies they are designed to

regulate. There is need for patient consent and the use of sophisticated methods to anonymize and protect data [18].

Reliability and safety

Reliability and safety form the core ethical issues in AI technology. Yet there are few guidelines for AI in health; they must be developed in collaboration among government, industry and academia. Fundamentally, the data used to train AI systems must be reliable.

There is no clarity about legal responsibility for their actions between the developers of the technology and the users (viz doctors). A conflict exists if technologists must be held accountable if AI system in healthcare directly affects the patient. When clinicians can't clearly (understand) and explain the output from AI to the patient, they are not justified to use the data for actions [7]. AI devices help in the decision-making process about treatment, and do not replace doctors entirely. They must be validated and established via testing, measurements for dependability, performance, ethical compliance and safety. However, physicians cannot take cover against accountability by blaming the AI systems [7].

Ownership of the data is another contentious issue that must be addressed. Public may not be willing to share their health data to commercial establishments who look at profit. Those using patient data must demonstrate that they add value to the health of the patients whose data they are using [10]. In addition, privacy of patients must be ensured against deleterious actions by insurance, job options or even personal relations. Novel problems may arise such as whether patient data can be shared with family members. Finally, it is necessary to define under what situations patients can withdraw their data [10].

As part of the solutions to ensure AI is safe and effective, datasets must be obligated to be reliable and valid; software updates must be regularly performed and must be transparent. Governance is necessary for safe and effective employment of AI [6].

Fairness and inclusivity

A core principle is to treat patient data fairly and in a balanced manner, but also treat people in other groups in a similar way, viz, to be fair. This involves elimination of bias in research and clinical practice. AI models are created and used by and in humans, with inherent social biases. One must recognize it is the data that is given to the algorithm that is

responsible for the bias. Therefore, AI systems must be ethical and free from biases; in other words, 'responsible AI systems' must be 'transparent, explainable and accountable' [7].

Currently, AI models are trained with data sourced from high-income settings, which results in a biased and discriminant system. One must be sensitive to biases that can creep at every stage of development, and avoid or at least minimize the risk [10]. Biased algorithms due to skewed training data sets led to injustice related to ethnicity, colour of skin or gender. Non-representative datasets lead to false diagnosis especially in phenotype-genotype associations.

Such biases can be minimized by collection of equitable data from all populations. Similarly, there are debates about the explainability versus the 'black box' model of AI development. Some argue that what is relevant is whether the algorithm is effective, rather than how it reaches the decision. Another issue is where the AI is used. Often experts in resource-rich settings have access, which leads to deprivation for the rest.

There is need for developing a clear path for ethical use of AI, although fairness is considered to be subjective. Nonetheless, the concept of fairness must be embedded in AI models [6]. Structural inequalities, biases and racism that are inherent in datasets can worsen social injustices [15]. Attention must be paid to avoid racism at the stages of both input and analysis. Ways to achieve this include awareness of the ways in which data science can perpetuate racism, seeking diverse and representative perspectives from the public and integrating them in the research plan, and emphasizing intersectional analysis of data [19]. Finally, the reporting of disaggregated data on ethnicity should become routine.

A recent analysis of AI use in different ethn racial groups showed that among publications specifically reporting race, about 69% were white, 17% black and 3.7% were Asian. Only two articles out of 10 reported inclusion of Native Americans [20]. It underscores the importance of addressing ethn racial inequities before they are embedded into the healthcare systems.

Finally, intellectual property (IP) aspects and commercial protection form an important subject for consideration. A number of variables come into play including long contracts, copyright and trade protection. IP often results in litigation about access to data and its analysis [10]. Problems arise because of conflicts among the nature, availability and

legalities of overlapping rights. Some solutions include flexible data exclusivity regimes which often requires doctrinal and normative challenges to the system.

Transparency and accountability

Transparency and accountability are the last but not the least important aspects in ethics of AI in healthcare. They involve human understanding of how decisions are made by AI in a manner that is transparent to researchers, healthcare providers as well as patients.

Difficulties arise because of the 'black box' nature of machine learning algorithms. There is much complexity involved and interpretability was reported to be considered as a 'human-computer' interaction problem [6]. Should machine-learning based decision support systems not be comprehensible to humans, the medical expert is left in an unenviable position of having to vouch for the trustworthiness of the system. A middle path may consider integrating medical expert's knowledge into the AI and ML models.

Explainability elicits debates which go beyond the technicalities of AI. It may be divided into (a) providing assessment of the need for explainability for clinical practice and (b) ethical evaluation of what the term means to be adopted into AI driven methods in clinical practice [21]. This requires that AI developers, healthcare workers and legislators comprehend the challenges of opaque algorithms. Explainability comprises of informed consent, certification and approval by regulatory agencies and liability based on Western legal viewpoint. Currently, regulatory agencies require explainability only rather vaguely; they would be defined in future. However, there are no clearcut answers as to what extent the patient must be informed that decision about treatment is based on AI. The conflict between regulation and innovation presently lies in the Court of law.

Explainability from a medical viewpoint consists of (a) understanding how the AI system arrives at a decision (b) identifying the important features for an individual prediction. At present, clinical validation is most widely discussed, while explainability is often considered as a 'second thought' [21]. Since AI systems cannot be perfectly accurate, explainability ensures that disagreement between AI system and human experts can be resolved.

In terms of the patients' relation to explainability, patient-centered care must be responsive and respectful of each person, taking them as equal

partners in decision making. The way forward seems to be a synthesis of available information to aid patients in understanding their risk and outcomes. A personalized conversation aid can help bridge the communication gap between the doctors and patients [21]. Four ethical principles operate in explainability, viz autonomy, beneficence, non-maleficence and justice.

Although people may be uneasy with a 'black-box' approach to AI in health care, one finds a trade-off between explainability and accuracy. Often, explainable AI systems are less accurate, and the reverse is also true [22]. It is unclear whether decisions based on AI need to be explained. To gain insight about explainability, two citizens juries were organized to determine whether AI systems provide explanation, even if that leads to less accurate decisions [23]. The results showed that in healthcare scenarios, citizens may prefer system accuracy more than explainability. Therefore, the public should be involved in the development of explainability in AI [19]. Rather than make categorical rules, a more nuanced domain specific considerations must be employed in such situations. Further work is needed in this area.

A practical adverse outcome was reported when IBM Watson for Oncology was used to suggest recommendations for cancer treatments because of using a few synthetic cancer cases [24]. These bring to the fore requirements for datasets which are reliable and valid, as well as transparency. Ideally, all data and algorithms must be available for the public to assess, although this could be a counsel of perfection. Nonetheless, developers must be transparent about the data that is used and potential biases in software [10].

Kawamleh argued against requirements for explainability for AI in health care, in relation to informed consent [25]. Based on legal requirements for informed consent in US, it is posited that it is doubtful if humans possess accurate insight into their own diagnostic criteria, much less explain such forms of reasoning. It is concluded that an epistemically opaque AI algorithms need not be considered a violation of patients right to informed consent [25].

Other ethical and legal aspects in AI for healthcare

Other unexplored issues are linked to fundamental aspects of reality, or the so-called 'metaphysical issues' of ethics. They are related to concepts of symbolic and logical representation of the external world. Although not urgent, they delve into the nature of humanity [11].

Legal issues regulate human societies; developments in AI remained isolated from legality until they occupied the social space and could impact people [6], specifically in relation to privacy and anonymity, liability and accuracy [7].

Digital therapeutics, not to be confused with digital medicine is based on software programs, or rather, 'software as a medical device (SaMD) [26]. They are often coupled with principles of AI in providing treatments particularly for behavioural and psychological aspects. They must be subjected to clinical trials, with attendant ethical and logistical challenges.

Varying aspects of ethics must be considered by AI developers across the span of the AI lifecycle including use of data and its management, development of models, their use and monitoring [27].

Guidelines are being developed for regulation of AI in healthcare arena. New regulations from the European Union were published to manage ethical

and legal hurdles within the legal framework, while drafting fresh regulations [28]. They covered the broad areas of oversight, transparency, diversity, accountability, and societal and environmental well-being. The goal of putting in place responsible AI systems is to ensure transparency, explainability and accountability [7].

Future of AI and healthcare

As of now, efforts of AI in healthcare are focused on developing cloistered models that can identify and predict narrow clinical conditions (eg identification of diabetic retinopathy, predicting the risk of hypoglycemia with the use of insulin pump and continuous glucose monitors). The next step in the development entails embedding AI in a broader framework where there is a confluence of personalized care, clinical decision support, early detection of disease as well as tracking the progression of disease. A synergy between AI and precision medicine could eventually lower the burden of disease in the population. Newer issues of ethics and legality are inevitable and await solutions [29].

References

1. Hasanzad M, Aghaei Meybodi HR, Sarhangi N, Larijani B. Artificial intelligence perspective in the future of endocrine diseases. *J Diabetes Metab Disord.* 2022;21(1):971-978. *J Diabetes Metab Disord.* 2022;21(1):971-978.
2. Gubbi S, Hamet P, Tremblay J, Koch CA, Hannah-Shmouni F. Artificial Intelligence and Machine Learning in Endocrinology and Metabolism: The Dawn of a New Era. *Front Endocrinol (Lausanne).* 2019;10:185. doi: 10.3389/fendo.2019.00185. PMID: 30984108; PMCID: PMC6448412.
3. London AJ. Artificial intelligence in medicine: Overcoming or recapitulating structural challenges to improving patient care? *Cell Rep Med.* 2022;3(5):100622. doi: 10.1016/j.xcrm.2022.100622. Epub 2022 Apr 27. PMID: 35584620; PMCID: PMC9133460.
4. Haug CJ, Drazen JM. Artificial Intelligence and Machine Learning in Clinical Medicine, 2023. *N Engl J Med.* 2023;388(13):1201-1208. doi: 10.1056/NEJMr2302038. PMID: 36988595.
5. Sridhar GR, Lakshmi G. Artificial Intelligence in Medicine: Diabetes as a Model. In: Srinivasa K.G., G. M. S., Sekhar S.R.M. (eds) *Artificial Intelligence for Information Management: A Healthcare Perspective.* Studies in Big Data, vol 88. Singapore, Springer 2021: 88. https://doi.org/10.1007/978-981-16-0415-7_14
6. Vellido A. Societal Issues Concerning the Application of Artificial Intelligence in Medicine. *Kidney Dis (Basel).* 2019;5:11-17. doi: 10.1159/000492428. Epub 2018 Sep 3. PMID: 30815459; PMCID: PMC6388581.
7. Naik N, Hameed BMZ, Shetty DK, Swain D, Shah M, Paul R, Aggarwal K, Ibrahim S, Patil V, Smriti K, Shetty S, Rai BP, Chlostia P, Somani BK. Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility? *Front Surg.* 2022;9:862322. doi: 10.3389/fsurg.2022.862322. PMID: 35360424; PMCID: PMC8963864.
8. Sabatello M, Martschenko DO, Cho MK, Brothers KB. Data sharing and community-engaged research. *Science.* 2022;378(6616):141-143. doi: 10.1126/science.abq6851. Epub 2022 Oct 13. PMID: 36227983.
9. Tully MP, Bozentko K, Clement S, Hunn A, Hassan L, Norris R, Oswald M, Peek N. Investigating the Extent to Which Patients Should Control Access to Patient Records for Research: A Deliberative Process Using Citizens' Juries. *J Med Internet Res.* 2018;20:e112. doi: 10.2196/jmir.7763. PMID: 29592847; PMCID: PMC5895919.
10. Sara Gerke, Timo Minssen, Glenn Cohen. Ethical and legal challenges of artificial intelligence-driven healthcare, In (Ed) Adam Bohr, Kaveh Memarzadeh. *Artificial Intelligence in Healthcare,* Academic Press,2020,295-336,<https://doi.org/10.1016/B978-0-12-818438-7.00012-5>.
11. Stahl, B.C. (2021). Ethical Issues of AI. In: *Artificial Intelligence for a Better Future. SpringerBriefs in Research and Innovation Governance.* 2021; Springer, Cham. https://doi.org/10.1007/978-3-030-69978-9_4
12. Price II WN. Artificial intelligence in the medical system: four roles for potential transformation. *Yale J Law Technol* 2019;212:122-32. <https://balkin.blogspot.com/2018/10/four-roles-for-artificial-intelligence.html>
13. Koonrungsomboon N, Hirayama K. Editorial: Ethical and regulatory challenges in genetic and genomic research involving stored biological specimens. *Front Genet.* 2022;13:1062188. doi: 10.3389/fgene.2022.1062188. PMID: 36338956; PMCID: PMC9634577.
14. Sharma M, Savage C, Nair M, Larsson I, Svedberg P, Nygren JM. Artificial Intelligence Applications in Health Care Practice: Scoping Review. *J Med Internet Res.* 2022;24:e40238. doi: 10.2196/40238. PMID: 36197712; PMCID: PMC9582911.
15. Chikhaoui E, Alajmi A, Larabi-Marie-Sainte S. Artificial Intelligence Applications in Healthcare Sector: Ethical and Legal Challenges. *Emerging Science J* 2022;6:717-37 Doi: 10.28991/ESJ-2022-06-04-05
16. Ursin F, Timmermann C, Orzechowski M, Steger F. Diagnosing Diabetic Retinopathy With Artificial Intelligence: What Information Should Be Included to Ensure Ethical Informed Consent? *Front Med (Lausanne).* 2021;8:695217. doi: 10.3389/fmed.2021.695217. PMID: 34368192; PMCID: PMC8333706.
17. Saheb T. "Ethically contentious aspects of artificial intelligence surveillance: a social science perspective". *AI Ethics.* 2022:1-11. doi: 10.1007/s43681-022-00196-y. Epub ahead of print. PMID: 35874304; PMCID: PMC9294797.
18. Murdoch B. Privacy and artificial intelligence: challenges for protecting health information in a new era. *BMC Med Ethics.* 2021;22:122. doi: 10.1186/s12910-021-00687-3. PMID: 34525993; PMCID: PMC8442400.

19. Knight HE, Deeny SR, Dreyer K, Engmann J, Mackintosh M, Raza S, Stafford M, Tesfaye R, Steventon A. Challenging racism in the use of health data. *Lancet Digit Health*. 2021;3:e144-e146. doi: 10.1016/S2589-7500(21)00019-4. Epub 2021 Feb 3. PMID: 33549513.
20. Pham Q, Gamble A, Hearn J, Cafazzo JA. The Need for Ethnoracial Equity in Artificial Intelligence for Diabetes Management: Review and Recommendations. *J Med Internet Res*. 2021;23:e22320. doi: 10.2196/22320. PMID: 33565982; PMCID: PMC7904401.
21. Amann J, Blasimme A, Vayena E, Frey D, Madai VI; Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak*. 2020;20:310. doi: 10.1186/s12911-020-01332-6. PMID: 33256715; PMCID: PMC7706019.
22. Gunning D, Stefik M, Choi J, Miller T, Stumpf S, Yang GZ. XAI-Explainable artificial intelligence. *Sci Robot*. 2019 ;4):eaay7120. doi: 10.1126/scirobotics.aay7120. PMID: 33137719.
23. van der Veer SN, Riste L, Cheraghi-Sohi S, Phipps DL, Tully MP, Bozentko K, Atwood S, Hubbard A, Wiper C, Oswald M, Peek N. Trading off accuracy and explainability in AI decision-making: findings from 2 citizens' juries. *J Am Med Inform Assoc*. 2021;28:2128-2138. doi: 10.1093/jamia/ocab127. PMID: 34333646; PMCID: PMC8522832.
24. Brown J. IBM Watson reportedly recommended cancer treatments that were 'unsafe and incorrect'. Gizmodo, <https://gizmodo.com/ibm-watson-reportedly-recommended-cancer-treatments-tha-1827868882>; 2018
25. Kawamleh, Suzanne. (2022). Against explainability requirements for ethical artificial intelligence in health care. *AI and Ethics*. 2022; 1-16. 10.1007/s43681-022-00212-1.
26. Refolo P, Sacchini D, Raimondi C, Spagnolo AG. Ethics of digital therapeutics (DTx). *Eur Rev Med Pharmacol Sci*. 2022;26:6418-6423. doi: 10.26355/eurrev_202209_29741. PMID: 36196692.
27. Solanki, P., Grundy, J. & Hussain, W. Operationalising ethics in artificial intelligence for healthcare: a framework for AI developers. *AI Ethics*; 2022. <https://doi.org/10.1007/s43681-022-00195-z>
28. Meszaros Janos, Minari Jusaku, Huys Isabelle. The future regulation of artificial intelligence systems in healthcare services and medical research in the European Union. *Frontiers in Genetics*.13:2022. URL=<https://www.frontiersin.org/articles/10.3389/fgene.2022.927721>.DOI=10.3389/fgene.2022.927721
29. Shandhi MMH, Dunn JP. AI in medicine: Where are we now and where are we going? *Cell Rep Med*. 2022;3:100861. doi: 10.1016/j.xcrm.2022.100861. PMID: 36543109.