FRAMEWORK OF RISK EVALUATION OF MEDICAL AI SYSTEMS

Volskaya EA[1] ✉, Rogova IV[2]

[1] NA Semashko National Research Institute of Public Health, Moscow, Russia

[2] Russian University of Medicine, Moscow, Russia

Medical technologies using artificial intelligence (AI) systems hold a firm place in real clinical practice as the main providers of important information for making medical decisions in diagnosis and treatment via assisting and auxiliary tools in the process of medical care provision. To obtain valid evidence of quality, effectiveness, and safety, AI software developers conduct clinical trials of these systems in accordance with current regulatory requirements [1], guided by the recommendations of recognized experts in the field of clinical research [2]. Ethical committees have a task to conduct a high-quality ethical review of the planned research, taking into account the specifics of AI technologies used in medicine and risks associated with their use.

Keywords: medical AI-system, ethical evaluation, risk evaluation, ethical postulates

Author contribution: the authors have made an equal contribution to the research and writing of the article.

✉ **Correspondence should be addressed**: Elena Alekseevna Volskaya
Chernyakhovsky St., 4/a, apt. 52, Moscow, 125319, Russia; vols-elena@yandex.ru

ПОДХОДЫ К ОЦЕНКЕ РИСКОВ МЕДИЦИНСКИХ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Е. А. Вольская[1] ✉, И. В. Рогова[2]

[1] Национальный НИИ общественного здоровья имени Н. А. Семашко, Москва, Россия

[2] Российский университет медицины, Москва, Россия

Медицинские технологии с использованием систем искусственного интеллекта (ИИ) занимают прочное место в реальной клинической практике в качестве основных поставщиков важной информации для принятия врачебных решений при диагностике и лечении, в формате ассистирующих и вспомогательных инструментов в процессе оказания медицинской помощи. Для получения валидных доказательств качества, эффективности и безопасности разработчики программного обеспечения с использованием ИИ (ПО с ИИ) проводят клинические исследования этих систем в соответствии с действующими нормативными требованиями [1], руководствуясь рекомендациями признанных экспертов в сфере клинических исследований [2]. Перед этическими комитетами стоит задача качественно провести этическую экспертизу планируемых исследований, учитывая специфику технологий ИИ, применяемых в медицине, и риски, связанные с их применением.

Ключевые слова: медицинские системы искусственного интеллекта, этическая экспертиза, оценка рисков, этические постулаты

Вклад авторов: авторы внесли равный вклад в проведение исследования и написание статьи.

✉ **Для корреспонденции**: Елена Алексеевна Вольская
ул. Черняховского, д. 4/а, кв. 52, г. Москва, 125319, Россия; vols-elena@yandex.ru

Artificial intelligence (AI) systems have rapidly entered all spheres of society, including healthcare and medicine. In accordance with current regulatory requirements, AI medical systems are subject to state registration as medical devices (software with artificial intelligence technology or software with AI) [3], which is carried out by Roszdravnadzor. At the beginning of October 2024, 37 medical devices using AI technologies were registered [4].

It should be noted that major clinical trials with AI-based software and, moreover, AI medical systems are associated with many difficulties. Therefore, they are not often conducted in our country and abroad. Thus, in an analysis performed using the US FDA database, only 20% of approved medical AI systems had passed pre-registration clinical trials by 2023, and no randomized trials were recorded among them [5]. And this is despite the fact that the FDA imposes clear requirements on the registration dossier in terms of information about the studies:

– demonstration of the desired medical benefit at set values of certain quality indicators;
– comparison of the evaluated product with classical clinical diagnostic or therapeutic procedures (reference standard);
– demonstration of technical/analytical capabilities;
– a modern prospective randomized multi-center study;
– demonstration of clinical efficacy, etc. [6].

In our country, universities and research institutes are conducting proactive research in the field of using artificial intelligence systems to provide medical care to patients along with major developments by serious manufacturers of medical AI systems, which are submitted to Roszdravnadzor for registration and implementation in medical practice. Such projects, especially if they are carried out as part of dissertation, are usually subject to examination by independent ethics committees (IECs).

Currently, the (IECs) have gained their first experience in ethical evaluation of independent research of medical AI systems. Most often, we are talking about navigation systems that use augmented reality for surgery, software for automatic image analysis for diagnostic purposes, medical decision support systems, etc. Not all systems are original; they include adaptation projects for using a medical device in a new field. Developers consider these studies, including within thesis works, as pilot projects. In case of positive results, they are planning to continue development.

Table 1. Gradation of potential risk levels for the use of medical AI systems proposed by the International Forum of Medical Device Regulators (IMDRF, 2014)

Clinical situation/ condition	The importance of information received from SaMD to take a medical decision		
	For treatment or diagnosis	For clinical management	For patient-management information
Critical	IV	III	II
Serious	III	II	I
Non-serious	II	I	I

Table 2. Risk assessment according to the Order of the Ministry of Health of the Russian Federation No. 686n dated July 7, 2020

The type of information provided by the software and its significance for making a medical decision	Categories of software application conditions		
	Category A for emergency cases, surgical intervention	Category B for emergency care without surgery	Category C for planned medical care
Does not require clarification, indicates the need for emergency action	3 — high risk (NB! + AI systems)	2b — increased risk	2a — medium risk
to be clarified	2b — increased risk	2a — medium risk	1 — low risk
3 — no need for immediate medical actions	2a — medium risk	1 — low risk	1 — low risk

The IEC needs to assess the risks of using an AI system in a clinical trial. Of course, IECs mainly follow relevant regulatory acts such as Helsinki Declaration of the WMA, Rules of Good Clinical Practice of the EAEU, and current GOST on clinical trials of medical devices [7], etc. The classification of risks of medical devices, which includes three risk classes (with two subclasses in class 2), should also be taken into account. Despite the fact that the safety degree of patients and subjects of research at the stage of medical device development is the main principle of risk ranking, nevertheless, the specific traits of AI software, including AI medical systems, take into account not only additional parameters [8], but also the entire existing regulatory framework for dealing with AI. Moreover, its detailed and systematic analysis is presented in scientific publications [9].

In 2024, the International Forum of Medical Device Regulators published the final document on the risk categories of software as a medical device (SaMD) [10]. These were the first recommendations on AI-specific software risk classification intended for use in medical technologies, including medical AI systems.

The document provides a matrix (Table 1), based on the clinical situation for which the AI medical system is intended, whereas the second parameter is the importance of medical decision support provided by AI software for a specific clinical situation. According to these criteria, four levels of risk are proposed. They ranged from the first, low level to the fourth, very high and critical risk level.

Three types of a clinical situations were considered:
- critical, when emergency (including surgical) medical care is needed for a patient with life-threatening conditions, including incurable conditions;
- clinical situations requiring serious therapeutic interventions, when a quick decision is required and time constraints can affect the ability of the decision-maker to correctly evaluate the information provided to him by the AI system;
- a clinical situation or patient's condition that does not require serious therapeutic interventions, when there is time to clarify the information received.

Another parameter that determines the risk level is the importance of the information provided by the AI system for making a medical (clinical) decision:
- information provided by SaMD should be used to make an immediate medical decision;
- information important for the diagnosis (detection) of a disease or condition, for clinical decisions on patient management, and for subsequent diagnosis and/or determination of a treatment plan;
- information important for determining the options for planned treatment, diagnosis, prevention, and alleviation of the disease symptoms.

In 2020, the Russian Ministry of Health issued Order No. 686h [11], which introduced very significant substantive changes to Order No. 4h 2012 'On Approval of the Nomenclature Classification of Medical Devices'. Section 'III. Classification of software that is a medical device' of Appendix No. 2 to the order appeared to be the most essential one.

In fact, this classification is based on a concept very close to that proposed in 2014 by the International Forum of Regulators. According to the Order, the structure of risk classes of software that is a medical device (including an AI medical system) fully and verbatim corresponds to that for medical devices. The only difference is that instead of the term 'medical devices' the phrase 'software' (software) is included: class 1 for low-risk software, class 2a for software with medium-risk, class 2b for higher-risk software, class 3 for high-risk software. It is noted that software is given a risk class regardless of the risk class of the medical device in combination with which it is used.

Two criteria are used to determine the level of risk: the type of information provided by the AI system and the clinical conditions of using the AI system.

There are three types of information provided by the AI system:
1) information that does not require clarification in order to make an informed clinical/medical decision and indicates the need for immediate actions;
2) information that needs to be clarified in order to make an informed clinical/medical decision;

3) information that does not show the need for immediate medical actions.

The clinical conditions of the AI medical system are also divided into three categories:

– category A is assigned if the AI system is intended for use in emergency situations, during surgical interventions, as well as in providing care for diseases with a high risk to individual and public health;

– category B is given in case of emergency care or medical care without surgical intervention, with a moderate risk to public health;

– category C is provided in routine medical care, medical care using non-invasive methods, with low risks to public health.

If we structure the paragraphs of rather extensive section III of Appendix 2, we get a table reflecting a very logical risk rating system (Table 2). There is only one exception in the well-structured classification of risks, depending on two criteria such as importance of the information provided by AI and complexity of the clinical situation. It concerns software using artificial intelligence technologies: any AI systems are classified as the ones with the highest risk and belong to Class 3 (clause 15.1.1 of section III of Appendix 2).

Attributing all AI medical systems to the highest risk class without exception might seem excessively rough. Although when it comes to AI systems designed to assist an operator during a surgery, and when the accuracy of the information provided depends on success of the operation, health and life of the patient, such roughness is justified and appropriate. For example, if an AI system performs diagnostic image analysis at the time when a doctor decides on treatment strategy for a patient with an acute stroke, when rapid and accurate differentiation between ischemic and hemorrhagic strokes is crucial for the choice of therapy. However, many AI systems have been introduced into clinical practice and continue to be implemented, providing auxiliary information for medical decision-making in much milder conditions. In fact, they could be classified as 2b or even 2a risk classes.

However, along with risk classification, ethics committees, when examining planned software research for medical technologies, including AI systems, should take into account other risks that the system may be associated with both during the research and in the future. These risks include:

– breach of confidentiality: in the worst case, discrimination in a social environment with consequences for mental health;

– influencing a medical choice, for example, when teaching SaMD using archived data, some of which may be biased;

– loss of a personal contact between the patient and the doctor;

– misleading with low-quality information about the AI system used in the process of providing medical care;

– anxiety, stress, and hypochondria developed due to the constant and frequent use of SaMD;

– errors in interpreting the system's response (incorrect self-treatment);

– technical failures, AI system hacks, cyber-attacks, etc.

To prevent these and other risks, the possibility of which cannot be excluded during the use of medical AI systems, it is necessary not only to minimize their negative effects at the research stages, but also to promote the responsible attitude of developers, control the use of these systems in real clinical practice, and increase patient loyalty to them. Distrust of innovative medical systems by patients can reduce the effectiveness of their use [12]. Therefore, ethics committees should perhaps expand the perspective of prognostic assessment when examining the planned studies, including the likely humanitarian impact of the application of the developed AI system on patients in clinical practice.

To date, it is possible to identify the main ethical postulates that should be followed by both developers of medical AI systems when designing developments, and by ethical committees when evaluating the developments:

– final decision-making authority should always remain with the doctor, since he is responsible for the medical care provided;

– control and storage of confidential medical data should be guaranteed, and periodic independent audits of data protection of subjects should be facilitated;

– patients/consumers should be fully informed about the AI systems used in the applied technologies. Ethics committees should monitor not only information material intended for research subjects, but also information related to the use of the AI system independently or as part of other medical technologies and intended for patients in clinical practice.

CONCLUSION

The system regulating the field of medical AI technologies is being formed and developed both in our country and around the globe (WHO, UNESCO, IMDRF and other organizations). The main provisions, conceptual framework, classification features, etc. are being developed and introduced into the sphere of AI technologies, which lays the foundation for unified approaches to the development of medical AI systems. Thus, in early October 2024 Rosstandart approved two important documents in the field of medical AI technologies such as the main provisions on medical decision support systems [13] and the main provisions on predictive analytics systems based on artificial intelligence [14]. The National Standard of the Russian Federation "Artificial intelligence systems in clinical medicine. Part 1. Clinical assessment" [15] was established as well.

Standardization of the field of medical AI systems is carried out in a very timely manner, since the number of AI systems being introduced into medical practice is constantly increasing. This leads to an increased public interest in both the use of AI software in everyday clinical practice and ethical aspects of development and application of medical AI technologies, which is reflected in the growing number of publications related to this topic.

Ethical issues related to introduction of innovative cognitive technologies capable of imitating thought processes into society are becoming the subject of discussion at representative international forums [16]. They are in the focus of attention of large public associations, such as the Alliance in the Field of Artificial Intelligence, which developed the Code of Ethics of AI [17], and are of scientific interest to serious scientific research teams [18].

However, issues of methodology for the ethical evaluation of clinical trials of medical technologies and systems using AI, as well as ethical aspects related to the introduction of these technologies into clinical practice, their perception by the patient community and variants of sociomental reactions, remain controversial. Obviously, experts in the field of research ethics still have to work together, in discussions and exchange of opinions, to develop criteria for ethical assessment and reference points for ethical committees.

References

1. Decision No.29 of the EEC Council dated February 12, 2016 «On Rules for clinical and laboratory tests (research) of medical products». Russian.
2. Ethical Review of Biomedical Trials. Guidelines for Ethics Committees. Ed. Hohlov AL. M., 2021; 792 p. Russian.
3. Letter of the Russian Federal Ministry of Health dated February 13, 2020 No. 02I-297/20 "On Software". Russian.
4. Gusev A. Registered medical devices using artificial intelligence technologies. Russian. Available from URL: https://webiomed.ru/blog/zaregistrirovannye-meditsinskie-izdeliia-ai/?ysclid=m2u6g82m51527476645 (accessed: 14.10.2024) Russian.
5. Medical AI Evaluation. Available from URL: https://ericwu09.github.io/medical-ai-evaluation/ (accessed: 14.10.2024)
6. Artificial Intelligence and Machine Learning in Software as a Medical Device. FDA. 2021. Available from URL: https://www.google.com/search?q=Artificial+Intelligence+and+Machine+Learning+in+Software+as+a+Medical+Device&rlz=1C1PRUC_en (accessed: 14.10.2024)
7. GOST R ISO 14155-2022 Clinical Research of medical devices involving Human Participants. Ressian. Available from URL: https://docs.cntd.ru/document/1200194060?ysclid=m2u8wxt2o368417633 (accessed: 14.10.2024) Russian.
8. Gusev AV, Kobyakova OS, Kovshirina YuV, Rebrova OYu. Ethical review of clinical research for AI systems. Guidelines. M.: RIO CNIIOIZ MZ RF, 2024; 32 p. DOI: 10.21045/978-5-94116-182-9-2024. Russian.
9. Hohlov AL, Belousov DYu. Ethical aspects of Sofrware with AI technology. Good clinical Prectice. 2021; 1: 70–84. Russian.
10. Software as a Medical Device: Possible Framework for Risk Categorization and Corresponding Considerations. IMDRF. 2014. URL: https://www.imdrf.org/documents/software-medical-device-possible-framework-risk-categorization-and-corresponding-considerations (accessed: 14.10.2024)
11. Order of the Ministry of Health of the Russian Federation of July 7, 2020 No. 686n "On Amendments to Annexes No. 1 and No. 2 to the Order of the Ministry of Health of the Russian Federation of June 6, 2012 No. 4n "On Approval of the Nomenclature Classification of Medical Devices". Russian.
12. Hohlov AL, Zarubina TV, Kotlovsky Myu, et al. Mechanisms for Introduction of Artificial Intelligence in Healthcare: new ethical Challenges. Russian. Medical Ethics. 2024; 3: 4–10.
13. GOST R 71671-2024 System for supporting medical decision-making using artificial intelligence. Main provisions. Available from URL: https://nd.gostinfo.ru/document/7582986.aspx (accessed: 14.10.2024) Russian.
14. GOST R 71672-2024 "Predictive analytics systems using artificial intelligence in medicine. Main provisions. Available from URL: https://nd.gostinfo.ru/document/7582951.aspx(accessed: 14.10.2024) Russian.
15. GOST R 59921.1-2022. Artificial intelligence systems in clinical medicine. P. 1. Clinical evaluation. Available from URL: https://files.stroyinf.ru/Data/782/78211.pdf (accessed: 14.10.2024) Russian.
16. IV AI Ethics Forum. Moscow. October 9, 2024. Available from URL: https://aiethic.ru/ (accessed: 14.10.2024) Russian.
17. AI Ethics Code. Available from URL: https://ethics.a-ai.ru/ (accessed: 14.10.2024) Russian.
18. Ethical review in the field of AI. Project of HSE AI Research Centre. Available from URL: https://cs.hse.ru/aicenter/ethical_review (accessed: 14.10.2024) Russian.

Литература

1. Решение Совета Евразийской экономической комиссии от 12.02.2016 № 29 «О правилах проведения клинических и клинико-лабораторных испытаний (исследований) медицинских изделий».
2. Этическая экспертиза биомедицинских исследований. Руководство для комитетов по этике. Под общей ред. А. Л. Хохлова. 3-е изд., перераб. и доп. М. ОКИ, 2021; 792 с.
3. Письмо Федеральной службы по надзору в сфере здравоохранения от 13 февраля 2020 г. № 02И-297/20 «О программном обеспечении».
4. Гусев А. Зарегистрированные медицинские изделия, использующие технологии искусственного интеллекта. Режим доступа: [Электронный ресурс] URL: https://webiomed.ru/blog/zaregistrirovannye-meditsinskie-izdeliia-ai/?ysclid=m2u6g82m51527476645 (дата обращения 14.10.2024).
5. Medical AI Evaluation. Available from URL: https://ericwu09.github.io/medical-ai-evaluation/ (accessed: 14.10.2024).
6. Artificial Intelligence and Machine Learning in Software as a Medical Device. FDA. 2021. Available from URL: https://www.google.com/search?q=Artificial+Intelligence+and+Machine+Learning+in+Software+as+a+Medical+Device&rlz=1C1PRUC_en (accessed: 14.10.2024).
7. ГОСТ Р ИСО 14155-2022. Клинические исследования медицинских изделий, проводимые с участием человека в качестве субъекта. Надлежащая клиническая практика. Режим доступа: [Электронный ресурс] URL: https://docs.cntd.ru/document/1200194060?ysclid=m2u8wxt2o368417633 (дата обращения 14.10.2024).
8. Гусев А. В., Кобякова О. С., Ковширина Ю. В., Реброва О. Ю. Этическая экспертиза клинических испытаний систем искусственного интеллекта. Методические рекомендации. М.: РИО ЦНИИОИЗ МЗ РФ, 2024; 32 с. DOI: 10.21045/978-5-94116-182-9-2024.
9. Хохлов А. Л., Белоусов Д. Ю. Этические аспекты применения программного обеспечения с технологией искусственного интеллекта. Качественная клиническая практика. 2021; 1: 70–84.
10. Software as a Medical Device: Possible Framework for Risk Categorization and Corresponding Considerations. IMDRF. 2014. Available from URL: https://www.imdrf.org/documents/software-medical-device-possible-framework-risk-categorization-and-corresponding-considerations (accessed: 14.10.2024).
11. Приказ Министерства здравоохранения Российской Федерации от 7 июля 2020 г. № 686н «О внесении изменений в приложения № 1 и № 2 к приказу Министерства здравоохранения Российской Федерации от 6 июня 2012 г. № 4н "Об утверждении номенклатурной классификации медицинских изделий"».
12. Хохлов А. Л., Зарубина Т. В., Котловский М. Ю. и др. Механизмы внедрения технологий искусственного интеллекта в здравоохранение: новые этические вызовы. Медицинская этика. 2024; 3: 4–10.
13. ГОСТ Р 71671-2024 «Система поддержки принятия врачебных решений с применением искусственного интеллекта. Основные положения. Режим доступа: [Электронный ресурс] URL: https://nd.gostinfo.ru/document/7582986.aspx (дата обращения 14.10.2024).
14. ГОСТ Р 71672-2024 «Системы прогнозной аналитики на основе искусственного интеллекта в клинической медицине. Основные положения». Режим доступа: [Электронный ресурс] URL: https://nd.gostinfo.ru/document/7582951.aspx (дата обращения 14.10.2024).
15. ГОСТ Р 59921.1-2022. Системы искусственного интеллекта в клинической медицине. Ч. 1. Клиническая оценка. Режим доступа: [Электронный ресурс] URL: https://files.stroyinf.ru/Data/782/78211.pdf (дата обращения 14.10.2024).
16. IV Форум этики в сфере искусственного интеллекта. Москва. 9 октября 2024. Режим доступа: [Электронный ресурс] URL: https://aiethic.ru/ (дата обращения 14.10.2024).
17. Кодекс этики в сфере ИИ. Режим доступа: [Электронный ресурс] URL: https://ethics.a-ai.ru/ (дата обращения 14.10.2024).
18. «Этическая экспертиза в сфере искусственного интеллекта». Проект Центра ИИ НИУ ВШЭ. Режим доступа: [Электронный ресурс] URL: https://cs.hse.ru/aicenter/ethical_review (дата обращения 14.10.2024).