

DOI: <https://doi.org/10.17816/DD321971>

Методология тестирования и мониторинга программного обеспечения на основе технологий искусственного интеллекта для медицинской диагностики

Ю.А. Васильев, А.В. Владзимирский, О.В. Омелянская, К.М. Арзамасов,
С.Ф. Четвериков, Д.А. Румянцев, М.А. Зеленова

Научно-практический клинический центр диагностики и телемедицинских технологий, Москва, Российская Федерация

АННОТАЦИЯ

Обоснование. Мировая сумма инвестиций в компании по разработке программного обеспечения на основе технологий искусственного интеллекта для медицинской диагностики составила 80 млн долларов в 2016 году, 152 млн долларов — в 2017 и, ожидаемо, продолжает расти. Активная деятельность компаний-производителей программного обеспечения должна соответствовать существующим клиническим, биоэтическим, правовым и методологическим основам и стандартам. Как на национальном, так и на международном уровне не существует единых стандартов и протоколов проведения испытаний и мониторинга программного обеспечения на основе технологий искусственного интеллекта для медицинской диагностики.

Цель — разработать универсальную методологию тестирования и мониторинга программного обеспечения на основе технологий искусственного интеллекта для медицинской диагностики, направленную на повышение его качества и внедрение в практическое здравоохранение.

Материалы и методы. В ходе аналитического этапа был проведён обзор литературы по базам данных PubMed и eLIBRARY. Практический этап включал апробацию разработанной методологии в рамках Эксперимента по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы.

Результаты. Разработана методология тестирования и мониторинга программного обеспечения на основе технологий искусственного интеллекта для медицинской диагностики, направленная на повышение качества данного программного обеспечения и его внедрение в практическое здравоохранение. Методология состоит из 7 этапов: самотестирование, функциональное тестирование, калибровочное тестирование, технологический мониторинг, клинический мониторинг, обратная связь и доработка.

Заключение. Отличительными особенностями методологии являются цикличность этапов тестирования, мониторинга и доработки программного обеспечения, приводящие к постоянному повышению его качества, наличие подробных требований к результатам его работы, участие врачей в его оценке. Методология позволит разработчикам программного обеспечения достичь высоких результатов и продемонстрировать достижения в различных направлениях, а пользователям — сделать осознанный и уверенный выбор среди программ, прошедших независимую и всестороннюю проверку качества.

Ключевые слова: программное обеспечение; искусственный интеллект; рентгенология; диагностическая визуализация; методология; контроль качества.

Как цитировать:

Васильев Ю.А., Владзимирский А.В., Омелянская О.В., Арзамасов К.М., Четвериков С.Ф., Румянцев Д.А., Зеленова М.А. Методология тестирования и мониторинга программного обеспечения на основе технологий искусственного интеллекта для медицинской диагностики // *Digital Diagnostics*. 2023. Т. 4, № 3. С. 252–267. DOI: <https://doi.org/10.17816/DD321971>

DOI: <https://doi.org/10.17816/DD321971>

Methodology for testing and monitoring artificial intelligence-based software for medical diagnostics

Yuriy A. Vasilev, Anton V. Vladzimirskyy, Olga V. Omelyanskaya, Kirill M. Arzamasov, Sergey F. Chetverikov, Denis A. Rumyantsev, Maria A. Zelenova

Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

ABSTRACT

BACKGROUND: The global amount of investment in companies developing artificial intelligence (AI)-based software technologies for medical diagnostics reached \$80 million in 2016, rose to \$152 million in 2017, and is expected to continue growing. While software manufacturing companies should comply with existing clinical, bioethical, legal, and methodological frameworks and standards, there is a lack of uniform national and international standards and protocols for testing and monitoring AI-based software.

AIM: This objective of this study is to develop a universal methodology for testing and monitoring AI-based software for medical diagnostics, with the aim of improving its quality and implementing its integration into practical healthcare.

MATERIALS AND METHODS: The research process involved an analytical phase in which a literature review was conducted on the PubMed and eLibrary databases. The practical stage included the approbation of the developed methodology within the framework of an experiment focused on the use of innovative technologies in the field of computer vision to analyze medical images and further application in the health care system of the city of Moscow.

RESULTS: A methodology for testing and monitoring AI-based software for medical diagnostics has been developed, aimed at improving its quality and introducing it into practical healthcare. The methodology consists of seven stages: self-testing, functional testing, calibration testing, technological monitoring, clinical monitoring, feedback, and refinement.

CONCLUSION: Distinctive features of the methodology include its cyclical stages of monitoring and software development, leading to continuous improvement of its quality, the presence of detailed requirements for the results of the software work, and the participation of doctors in software evaluation. The methodology will allow software developers to achieve significant outcomes and demonstrate achievements across various areas. It also empowers users to make informed and confident choices among software options that have passed an independent and comprehensive quality check.

Keywords: software; artificial intelligence; radiology; diagnostic imaging; methodology; quality control.

To cite this article:

Vasilev YuA, Vladzimirskyy AV, Omelyanskaya OV, Arzamasov KM, Chetverikov SF, Rumyantsev DA, Zelenova MA. Methodology for testing and monitoring artificial intelligence-based software for medical diagnostics. *Digital Diagnostics*. 2023;4(3):252–267. DOI: <https://doi.org/10.17816/DD321971>

DOI: <https://doi.org/10.17816/DD321971>

基于人工智能技术的医疗诊断软件测试和监测方法学

Yuriy A. Vasilev, Anton V. Vladzimirskyy, Olga V. Omelyanskaya, Kirill M. Arzamasov, Sergey F. Chetverikov, Denis A. Rumyantsev, Maria A. Zelenova

Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

简评

论证。2016年，全球对基于人工智能技术开发医疗诊断软件的公司投资额为8000万美元，2017年为1.52亿美元，并预料还将继续增长。软件公司的积极活动必须符合现有的临床、生物伦理、法律和方法学原理和标准。在国家与国际范围，基于人工智能技术的软件还没有统一的测试和监测标准和协议。

该研究的目的是开发一种通用方法，用于测试和监测基于人工智能技术的医疗诊断软件，以提高其质量和在实际医疗中的应用。

材料和方法。在分析阶段，对PubMed和eLIBRARY数据库进行了文献综述。实用阶段包括在《使用创新计算机视觉技术进行医学图像分析并进一步应用于莫斯科市医疗系统的实验》框架内批准所开发的方法学，并将其进一步应用于莫斯科的医疗保健系统。

结果。我们开发了一套基于人工智能技术的医疗诊断软件测试和监测方法学，旨在提高该软件的质量，并将其应用于实际医疗保健中。该方法学包括7个阶段：自我测试、功能测试、校准测试、技术监测、临床监测、反馈和改进。

结论。该方法学的显著特点是对软件进行周期性的监测和改进，从而不断提高其质量；对软件性能结果并医生参与软件评估提出详细要求。该方法学可使软件开发人员在各个领域取得优异成绩并展示成就，也可使用户在通过独立、全面质量控制的程序中做出明智、自信的选择。

关键词：软件；人工智能；放射学；诊断成像；方法学；质量控制。

引用本文：

Vasilev YuA, Vladzimirskyy AV, Omelyanskaya OV, Arzamasov KM, Chetverikov SF, Rumyantsev DA, Zelenova MA. 基于人工智能技术的医疗诊断软件测试和监测方法学. *Digital Diagnostics*. 2023;4(3):252-267. DOI: <https://doi.org/10.17816/DD321971>

收到: 06.04.2023

接受: 15.06.2023

发布日期: 30.08.2023

BACKGROUND

Global investment in developing software based on artificial intelligence (AI) technologies for medical diagnostics was \$80 million in 2016 and \$152 million in 2017; it is likely to grow continually [1]. In 2019, the Moscow government decided to conduct a large-scale scientific study (which is still ongoing in 2023) to evaluate the use of innovative computer vision technologies for analyzing medical images and further application in the Moscow healthcare system (hereinafter referred to as the Experiment).¹

Software manufacturers must comply with current clinical, bioethical, legal, and methodological principles and standards [1]. According to Russian legislation, before using AI-based software in clinical practice, it must be legally approved as a medical device, which requires the software to receive a marketing authorization (MA) from the Federal Service for Surveillance in Healthcare (Roszdravnadzor).²

Before submission, the software should be assessed in technical and clinical studies to ensure that the specified functions are met.³ However, due to particular aspects of AI-based software, such as a lack of user-friendly information regarding its operating process and decision-making principles, there are no uniform standards and test protocols for this purpose at the national and international levels [2]. The Food and Drug Administration (FDA) in the United States also establishes explicit criteria for evaluating and regulating AI-based software [1]. The inability to reliably confirm software compliance has negative consequences, including user distrust in the software, slower implementation in clinical practice, missing positive socioeconomic impacts from software, and slower overall development of the healthcare system [3].

After receiving an MA, post-marketing clinical monitoring should be performed to ensure the safety of using this software in clinical practice.⁴ However, the present criteria apply to various medical devices and do not consider special aspects of AI-based software for medical diagnostics [4]. According to the Decision of the Board of the Eurasian Economic Commission, medical devices of the third risk class (including AI-based software) must be monitored annually for 3 year after acquiring an MA.⁵ However, more frequent monitoring is required because of the high variability of medical data and the difficulty of predicting changes in

environmental conditions, such as the epidemiological situation [5]. Monitoring enables the identification of critical remarks on the results of software operations that require software improvement, and, when the software is finalized, repeated testing and monitoring should be performed.

A retrospective cohort study is the most appropriate for evaluating AI-based medical diagnostics software [1]. This software evaluation method has several disadvantages, the most significant of which is a difference in the actual results of software operation in ideal and practical settings [1]. A common example is the negative experience of introducing the first computer-aided diagnostic system for mammography screening. Large-scale multicenter studies found that using this software increased breast cancer diagnosis by 2%–10% [6]. In 1998, the FDA approved the software for use in clinical practice. However, in real-world settings, this software did not achieve positive results. When interpreting mammography results, it even leads to a decrease in detection rate and an increase in false positive results [6]. The study suggested that radiologists with varying degrees of expertise used the new technology in different ways. More experienced specialists did not pay attention to it, whereas less experienced ones made mistakes due to a false sense of security. The second explanation is that the software was ineffective in detecting certain forms of cancer, which were not found in previous examinations [1].

Therefore, although ethical and legal problems are the most common with AI-based software, there is also an important methodological problem, which can be defined as a lack of universal and comprehensive methodology for testing and monitoring AI-based software for medical diagnostics to improve its quality and further implement it in clinical practice [7]. According to the above, it is important to develop such a methodology. The methodology will not replace the existing legal methods for assessing the safety and effectiveness of software but will exist independently and contribute to the likelihood of successful Roszdravnadzor approval of software. After receiving an MA, this methodology will help further assess and improve software for its effective implementation in clinical practice.

This study aims to develop a universal methodology for testing and monitoring AI-based software for medical diagnostics to improve its quality and implement it in clinical practice.

¹ Decree No. 1543-PP of the Moscow Government dated November 21, 2019 on conducting an experiment on the use of innovative technologies in computer vision for analyzing medical images and further application in the Moscow healthcare system. Link: <https://docs.cntd.ru/document/563879961>.

² Decree No. 1906 of the Government of the Russian Federation dated November 24, 2020 on amendments to the Rules for state registration of medical devices. Link: <http://publication.pravo.gov.ru/Document/View/0001202011270010>.

³ Federal Law No. 323-FZ dated November 21, 2011. Basics of Health Protection of the Citizens in the Russian Federation. Article 38. Medical devices. Link: https://www.consultant.ru/document/cons_doc_LAW_121895/ddcfddbdbb49e64f085b65473218611b4bb6cd65/.

⁴ Order No. 980n of the Ministry of Health of Russia dated September 15, 2020 on approval of the procedure for monitoring the safety of medical devices. Link: <https://docs.cntd.ru/document/566006416>.

⁵ Decision No. 174 of the Board of the Eurasian Economic Commission dated December 22, 2015 on approval of the rules for monitoring the safety, quality, and effectiveness of medical devices. Link: <https://www.alt.ru/tamdoc/15kr0174/>.

MATERIALS AND METHODS

Study design

The presented methodology was developed by analyzing literature and personal experience.

Development of Methodology

The methodology was developed in two stages: analytical and practical.

To study existing methodologies, literature published in PubMed and eLIBRARY scientific libraries from 2018 to 2023 (the last 5 years) was reviewed using the search terms “methodology for evaluation AI in radiology” and “methodology for assessing AI in radiology.” After assessing their relevance, papers were included in the analysis by reading the title and abstract. There were 22 papers [1–22] and five legal acts examined.⁶

Based on the Unified Radiological Information Network (ERIS) of the Unified Medical Information and Analytical System of Moscow (EMIAS), the methodology was tested during the Experiment on using innovative computer vision technologies for analyzing medical images and further application in the Moscow healthcare system. Some testing results are presented in this article as an illustration.

Statistical justification of sample sizes

The number of studies included in the sample was determined at different stages.

1. At the self-testing stage, the size of a data set is not regulated and varies depending on the clinical problem solved by the software.⁷ The data sets used at the stages of self-testing, functional, and calibration testing are based on expert consensus data, with histological conclusions used in some cases (e.g., when assessing malignant neoplasms). The process of preparing data sets is described in detail in the corresponding regulations [19].
2. At the functional testing stage, the data set included five studies (based on GOST R 8.736-2011, multiple measurements require at least four measurements).⁸ An expert’s opinion is considered a true value. An expert is

a healthcare professional who has been working as a specialist for more than 5 years and has been trained in AI-based software to describe examinations in the related field (a specific modality and target abnormality). This stage requires at least one technical specialist and one medical expert.

3. At the stage of calibration testing, the data set includes 100 studies with a 50/50 balance (50% of examinations with target abnormality and 50% without it) [20, 21].⁹ At this stage, one technical specialist and one medical expert are required.
4. At the stage of technological monitoring, all examinations for the reporting period should be assessed by software for the presence of defects “a” and “b” (based on automated defect detection), with a sample of 80 examinations for defects “c” to “e” [20, 21].¹⁰ At least one technical specialist is required at this stage.
5. At the stage of clinical monitoring, the data set includes 80 examinations, and an expert’s opinion is considered the true value [20, 21].¹¹ At this stage, one expert is required.

Ethical review

This study was conducted as part of another study that had previously been approved by the local ethics committee (No. NCT04489992), “Experiment on the use of innovative technologies in computer vision for analyzing medical images and further application in the Moscow healthcare system” (Moscow experiment).

RESULTS

Based on the literature review, papers were found to describe individual stages of evaluating AI-based software for medical diagnostics, such as validation [1, 5, 8, 9], monitoring [10], implementation [7, 11–13], and regulation [14, 15]. However, there is no unified methodology for testing and monitoring AI-based software for medical diagnostics. There have been papers on the life cycle of AI-based software [16], but they are mainly related to nonmedical software and do not consider special aspects of AI-based software

⁶ Decree No. 1543-PP of the Moscow Government of the Russian Federation dated November 21, 2019. Link: <https://docs.cntd.ru/document/563879961>; Decree No. 1906 of the Government of the Russian Federation dated November 24, 2020. Link: <http://publication.pravo.gov.ru/Document/View/0001202011270010>; Article 38 of Federal Law No. 323-FZ dated November 21, 2011. Link: https://www.consultant.ru/document/cons_doc_LAW_121895/; Order No. 980n of the Ministry of Health of Russia dated September 15, 2020. Link: <https://docs.cntd.ru/document/566006416>; Order No. 134 of the Moscow Department of Health dated February 16, 2023 Link: https://mosmed.ai/documents/227/order_DZM__134_d_02/16/2023.pdf.

⁷ Center for Diagnostics and Telemedicine. Official website. Data sets. Link: <https://mosmed.ai/datasets/>.

⁸ GOST R 8.736-2011. National standard of the Russian Federation. State system for ensuring the uniformity of measurements. Multiple direct measurements. Methods for processing measurement results. Basic provisions. Link: <https://docs.cntd.ru/document/1200089016>.

⁹ Order No. 134 of the Moscow Healthcare Department dated February 16, 2023 on approval of the procedure and conditions for conducting an experiment on the use of innovative technologies in computer vision for analyzing medical images and further Use in the Moscow Healthcare System. Link: https://mosmed.ai/documents/227/order_DZM__134_d_16.02.2023.pdf.

¹⁰ Ibid.

¹¹ Ibid.

for medical diagnostics. Furthermore, there are guidelines for conducting research and writing scientific publications on AI-based software, but they do not assist with testing and monitoring software [17, 18]. It should be noted that no publications on software modification after testing and monitoring were found. However, software improvement is necessary to improve its quality and effective implementation in clinical practice.

As a result, the authors developed a methodology for testing and monitoring AI-based software for medical diagnostics to improve its quality and use in clinical practice. The methodology consists of seven stages, as shown in Figure 1. The purpose, primary actions, and results are described below for each stage.

Self-testing

The self-testing stage is intended to assess the technical compatibility of software with input data. Software developers (or suppliers) are provided access to an open data set containing files in the Digital Imaging and Communications in Medicine (DICOM) format with anonymized examples of diagnostic examinations.¹² The data set has the following parameters: modality, type of diagnostic procedure, manufacturer, and model of the diagnostic device [19].

Software compatibility with data enables software integration into a healthcare institution's radiology information network and continues with further evaluation, starting with the functional testing stage.¹³

Functional testing

Functional testing is a stage wherein software functions specified by a supplier are checked for availability and functionality. This testing is performed at the technical and clinical levels. On a technical level, the software is assessed based on the following criteria: prioritization of research (triage), availability of an additional series of images from the software, presence of the other series' name, presence of a graphical designation of software on the images of the different series, presence of a warning label "For research purposes only" on images and in DICOM SR, possibility of series synchronization, displaying the probability of abnormality, indication of the category of abnormality, and availability of complete DICOM SR protocol structure (Figures 2 and 3).

This part of functional testing should be performed by technical specialists in accordance with the basic functional requirements developed by the Moscow State Budgetary Institution "Scientific and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Department of Health" (Center for Diagnostics and Telemedicine).¹⁴ The medical assessment of software functions should be performed by medical experts in accordance with basic diagnostic requirements developed by the Center for Diagnostics and Telemedicine.¹⁵ Basic diagnostic requirements include criteria, such as mandatory and optional content of software response, format, and form of the submitted response. Basic functional and diagnostic requirements contain common requirements for all software

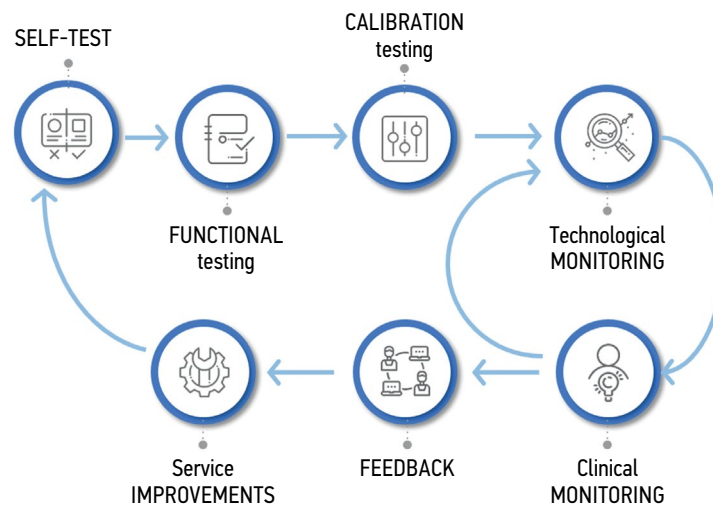


Fig. 1. Methodology for testing and monitoring artificial intelligence-based software for medical diagnostics.

¹² Center for Diagnostics and Telemedicine. Official website. Data sets. Link: <https://mosmed.ai/datasets/>.

¹³ Order No. 134 of the Moscow City Health Department dated February 16, 2023 on approval of the procedure and conditions for conducting an experiment on the use of innovative technologies in computer vision for analyzing medical images and further use in the Moscow healthcare system. Link: https://mosmed.ai/documents/227/order_DZM__134_d_16.02.2023.pdf.

¹⁴ Basic functional requirements for AI service results. Link: https://mosmed.ai/documents/218/Basic_functional_requirements_29.11.2022.pdf.

¹⁵ Basic diagnostic requirements for AI service results. Link: https://mosmed.ai/documents/226/Basic_diagnostic_requirements_22_02_2023.pdf.

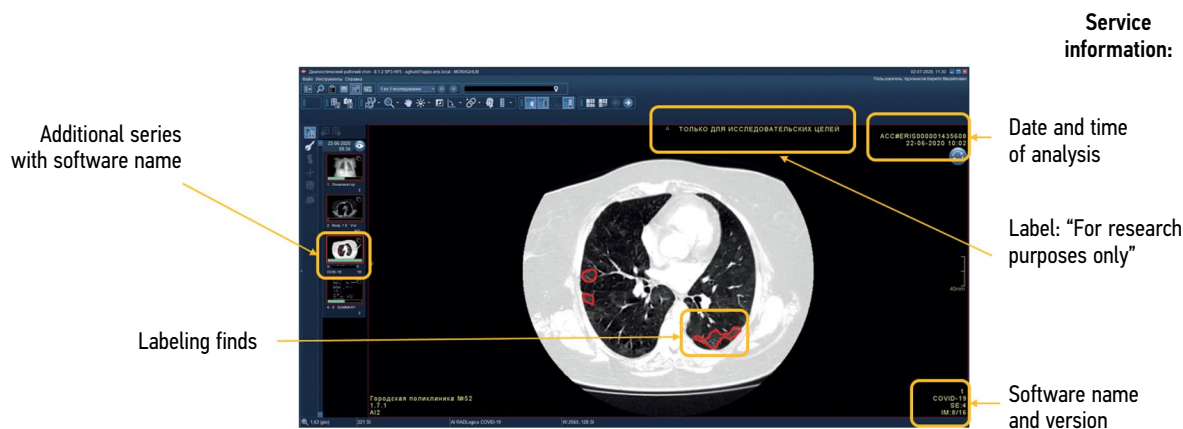


Fig. 2. Main components of the result of using artificial intelligence–based software with images: A reference example.

DICOM SR 1: Service information about research and service	<p>Inventory No. ERS 00001990912 Research date: August 12, 2020, 0:47 Date of conclusion: August 12, 2020, 1:08 Status Service name: COVID-19 Warning: "For research purposes only" Software version 1.8.0 Date and time of analysis: August 12, 2020, 01:08</p> <p>Purpose of the service Modality: CT Anatomical region: Chest Patients: Adults Purpose: Searching for signs and assessing the extent of lung damage due to COVID-19 pneumonia</p> <p>Quick Start Guide with lesions marked with a red outline</p> <p>Conclusion Result of AI processing COVID-19: In the left upper lobe, 0% of the lung parenchyma is affected (score 0) In the left lower lobe, 16.1% of the lung parenchyma are affected (score 2) In the right upper lobe, 0% of the lung parenchyma are affected (score 0) In the right middle lobe, 0% of the lung parenchyma is affected (score 0) In the right lower lobe, 0% of the lung parenchyma is affected (score 0) Total for both lungs: 8%. Severity: CT1</p> <p>Sign of COVID-19 Number of slices with abnormalities detected: 114 Total number of lung sections: 269 Percentage ratio of positive to total slices: 42%</p>	
DICOM SR 2: Information about the service and its purpose		
DICOM SR 3: Information on how to work with the service (user manual)		
DICOM SR 4: Conclusion of AI-service based on research results.		
DICOM SR 5: Detailing of findings		

Fig. 3. Main components of the result of using artificial intelligence–based software with DICOM SR: A reference example.

and specific requirements based on the clinical task for which the software is designed.

If critical nonconformities are identified, software testing is stopped until the supplier eliminates their causes. Inconsistencies with basic functional requirements are critical because they negatively affect the HCP work processes and, directly or indirectly, the patient’s life and health (Figures 4 and 5).

Functional testing should be repeated after the supplier has eliminated the causes of critical nonconformities. This stage may be repeated no more than twice by the applicant. There are no time limits for the initial retesting after receiving the protocol with unsatisfactory test results. The second retesting should be performed no earlier than 3 months after

receiving the last protocol with unsatisfactory test results. If the second retest fails, the applicant may be provided an alternative scientific and practical cooperation option.¹⁶ If no critical inconsistencies are found, the software moves to the calibration testing stage.¹⁷

Calibration testing

Calibration testing is a stage wherein the diagnostic accuracy of software is determined. The main parameter is the area under the ROC curve (AUC). The optimal value of the activation threshold is determined by examining the ROC curve using Youden’s J statistic and maximizing the negative and positive predictive value. Other metrics include sensitivity, specificity, accuracy, and positive and negative

¹⁶ Order No. 134 of the Moscow City Health Department dated February 16, 2023 on approval of the procedure and conditions for conducting an experiment on the use of innovative technologies in computer vision for analyzing medical images and further use in the Moscow healthcare system. Link: https://mosmed.ai/documents/227/order_DZM__134_d_16.02.2023.pdf.

¹⁷ Ibid.

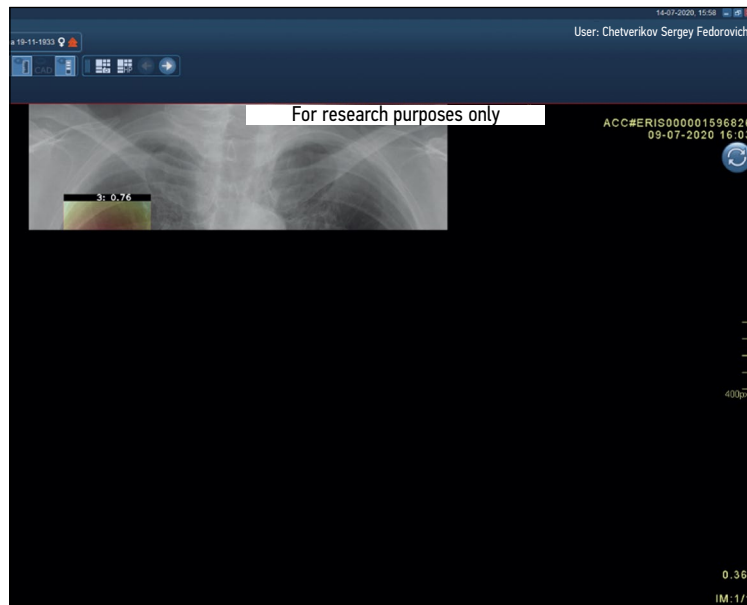


Fig. 4. Image clipping of additional series of artificial intelligence–based software: Critical noncompliance with basic functional requirements.

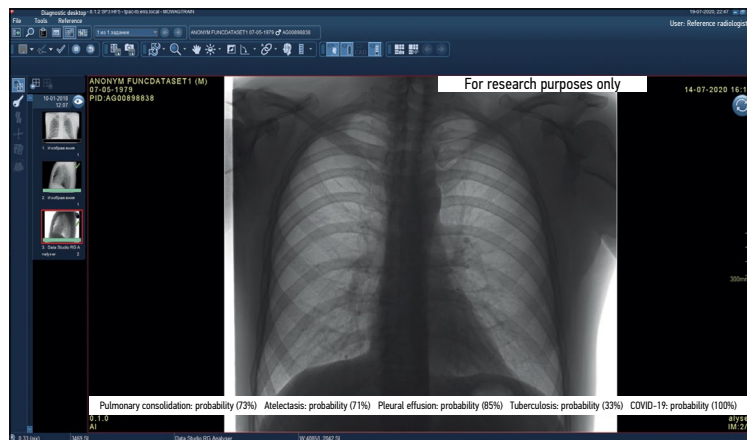


Fig. 5. Overlaying caption texts on images: Critical noncompliance with basic functional requirements.

predictive values. The minimum, average, and maximum time required to analyze one examination are also determined, and numbers of true positive, false negative, false positive, and true negative results are presented as a four-field table. Threshold values for some parameters are as follows: AUC ≥ 0.81 or 0.91 (depending on the clinical task); time spent on acceptance, processing of the study, and transmission of analysis results ≤ 6.5 min; and percentage of successfully processed examinations $\geq 90\%$ [21].¹⁸

Calibration testing results in a calibration protocol (Figure 6), which may contain critical and noncritical inconsistencies. Noncompliance with the above threshold values and significant deviations from methodological recommendations are considered crucial [21]. If they

are identified, software testing is stopped until they are eliminated. In their absence, the software may proceed to a prospective examination analysis as part of the periodic monitoring stage, which includes technological and clinical monitoring.¹⁹

Technological monitoring

Technological monitoring is a stage involving a periodic technical check of software results. This stage is required for rapid defect identification, timely quality control, and the prevention of functional software errors in radiology practice. Defects that can be identified at this stage are divided into the following groups:

(a) the processing time for one study exceeds 6.5 min,

¹⁸ Order No. 134 of the Moscow City Health Department dated February 16, 2023 on approval of the procedure and conditions for conducting an experiment on the use of innovative technologies in computer vision for analyzing medical images and further use in the Moscow healthcare system. Link: https://mosmed.ai/documents/227/order_DZM__134_d_16.02.2023.pdf.

¹⁹ Ibid.

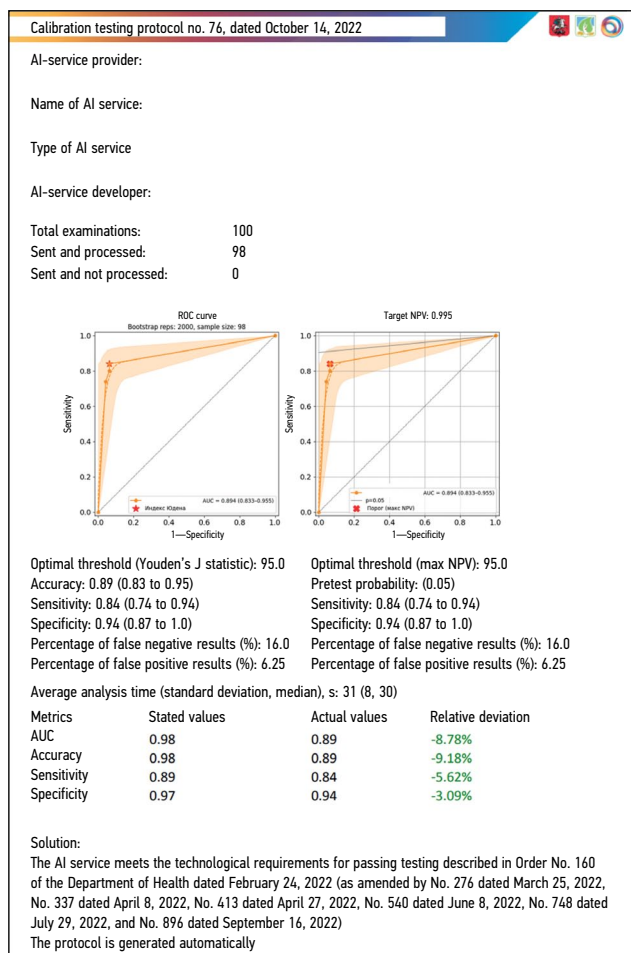


Fig. 6. Example of a calibration test protocol.

- (b) a lack of results from the examinations reviewed,
- (c) incorrect operation of the declared software functions, complicating the work of a radiologist or making it impossible to perform it with proper quality,
- (d) defects associated with the display of the image area, and
- (e) other violations of the integrity and contents of files containing research results, limiting diagnostic interpretation.

Defects “a” and “b” are monitored automatically for all examinations reviewed by software during the reporting period. For defects “c” and “d,” semi-automated monitoring is used with a sample of 80 examinations. An internal report form for monitoring software operation with instructions for monitoring technological defects has been developed for accurate defect assessment (Figure 7). Figure 8 shows graphical information on the average number of technological defects for the “chest radiography” area, with a tendency for the number of defects to decrease.

A technological monitoring report is the deliverable of technological monitoring (Figure 9). If the percentage of detected defects exceeds 10%, then testing this software is suspended until the causes of the defects are eliminated. If the percentage of detected defects does not exceed 10%, then the operation of the software and its periodic monitoring continue.²⁰

Clinical monitoring

During periodic monitoring, a clinical assessment of software results is also performed by radiologists. Two

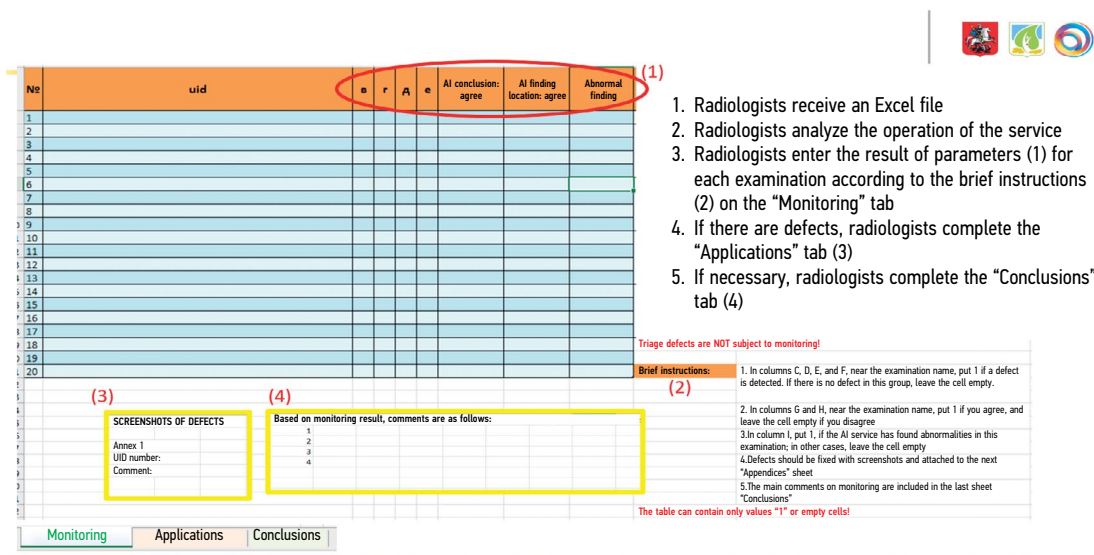


Fig. 7. Form of an internal report on monitoring the operation of artificial intelligence–based software.

²⁰ Order No. 134 of the Moscow City Health Department dated February 16, 2023 on approval of the procedure and conditions for conducting an experiment on the use of innovative technologies in computer vision for analyzing medical images and further use in the Moscow healthcare system.” Link: https://mosmed.ai/documents/227/order_DZM__134_d_16.02.2023.pdf.

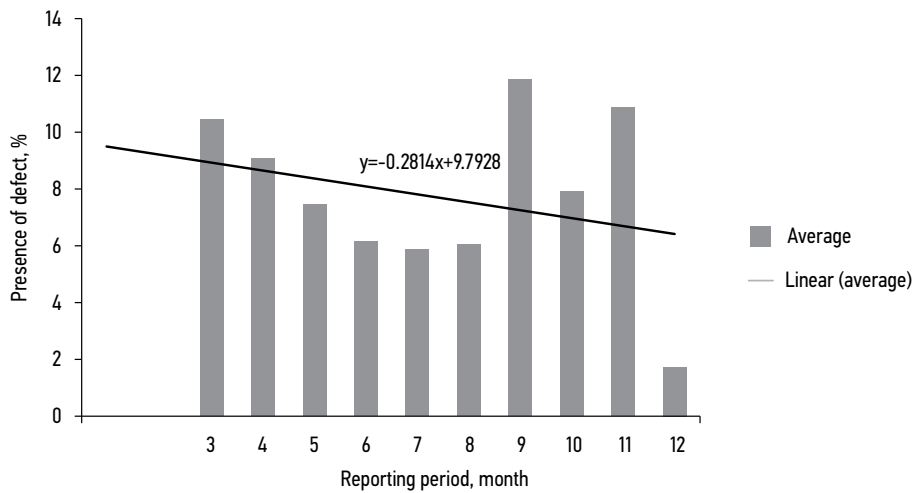


Fig. 8. Changes of technological software defects for "chest radiography" modality.

REPORT ON MONITORING OF TECHNOLOGICAL PARAMETERS OF THE SERVICE OPERATION.	
1. AI-service provider: Name of AI service: AI-service manufacturing company: AI-service ID in the experiment:	
2. Reporting period:	
3. Type of research:	
4. Clinical goal:	
5. Total number of examinations: 5.1. Sent for service analysis for the reporting period according to data extracted from ERIS EMIAS, No. 16108 5.2. Of them, unique according to data extracted from ERIS EMIAS, No. 16108	
6. Number of examinations that passed control, No. 16108 6.1. Passed manual control, No. 20	
7. Number of examinations with defects: 7.1. With technological defect "a", No., Appendix 1 32 7.2. With technological defect "b", No., Appendix 2 808 7.3. With technological defects "c" to "e", No., Appendix 3 1	
8. Relative share of examinations: 8.1. With technological defect "a" for 15,300 examinations, % 0 8.2. With technological defect "b" for 16,108 examinations, % 5 8.3. With technological defects "c" to "e" for two examinations, % 5	
9. Number of examinations without defects, pc 15267	
10. Solution: Ongoing AI-service participation in the experiment.	
11. Notes:	
Report date:	
Full name of the responsible person:	The report is generated automatically.
*0 non-unique examination extracted from ERIS EMIAS for the reporting period.	

Fig. 9. Example of a technology monitoring report.

main evaluation criteria include interpretation (conclusion) and localization (labeling) of an abnormal finding. During the assessment, the response options that clinicians can choose from include full compliance, incorrect assessment, false positive result, and false negative result. For example, the wording "Interpretation: Full compliance" is selected when a specialist fully agrees with the software conclusion, and the wording "Interpretation: Incorrect assessment" is

selected when the doctor partially agrees with a software conclusion (e.g., the specialist agrees with the presence of abnormal findings but disagrees with its details, or vice versa, they agree with details but disagree with the general conclusion about the possibility or severity of abnormal findings). If the specialist completely disagrees with the software conclusion, the wordings "Interpretation: False positive result" and "Interpretation: False negative result" are used (Figure 10).

The clinical assessment results are entered into the abovementioned internal monitoring report and imported into the monitoring software module, from which a final monitoring report is automatically generated.

Based on periodic monitoring, one of the following conclusions is adopted: "The participation of the software in the Experiment continues," "The participant in the Experiment needs to make changes to the operation of the software," and "The participation of the software in the Experiment is suspended until changes are made to the operation of the software."²¹

Feedback

The stage of radiologist feedback is required to assess the software's practical relevance. The feedback form is in the program window on the radiologist's automated workstation (Figure 11). The software's result may be agreed upon or disagreed upon by a radiologist. In case of disagreement, they select a reason. The primary causes include technological defects and diagnostic inaccuracy. It is necessary to obtain specialist feedback on 5% of all examinations assessed by software. In addition, feedback is collected through a survey of specialists to determine their satisfaction with the software.²²

²¹ Order No. 134 of the Moscow City Health Department dated February 16, 2023 on approval of the procedure and conditions for conducting an experiment on the use of innovative technologies in computer vision for analyzing medical images and further use in the Moscow healthcare system. Link: https://mosmed.ai/documents/227/order_DZM__134_d_16.02.2023.pdf.

²² Ibid.

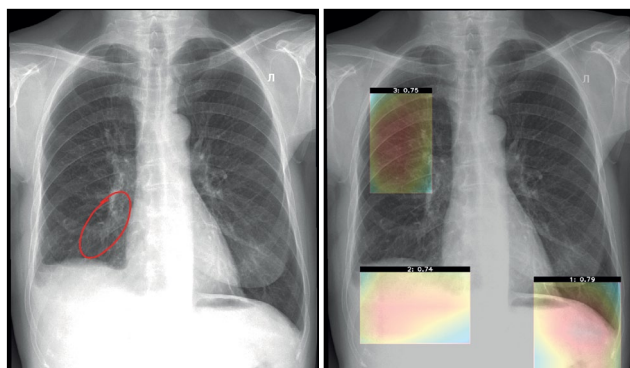


Fig. 10. False negative (the subsegmental atelectasis is not detected in the lower lobe of the right lung): Noncritical noncompliance with basic diagnostic requirements.

Finalization

Suppose a critical comment regarding the software operation is identified at functional, calibration testing, and periodic monitoring stages. In that case, software testing is suspended until the causes of the comment are eliminated. Software finalization is performed by the supplier, which serves as a “secret box” for the healthcare organization. Suppose the modifications required do not involve changes in the initially declared functions or technical architecture and do not affect the diagnostic accuracy of the software. In that case, the applicant can proceed to the next stage of the methodology immediately after making modifications.

If the applicant makes modifications that affect the initially declared functions, technical architecture and diagnostic accuracy of the software, functional and calibration testing should be repeated regardless of what stage of the software methodology it was at.²³

DISCUSSION

This paper presents a methodology for testing and monitoring the results of AI-based software for medical diagnostics to improve its quality and implement it in clinical practice. The key reasons for its development include the lack of specific requirements for testing and monitoring AI-based software for medical diagnostics in existing regulatory documentation and the lack of regulated principles for software selection by a healthcare organization among various software programs on the market. This methodology does not conflict with legal requirements but considers special characteristics of AI-based software for medical diagnostics. The methodology includes seven unique, clearly organized, scientifically validated stages [1–4, 19–21]; it is supported by legislative documents.²⁴

The presence of developed basic functional and diagnostic requirements used at the functional testing stage is a key element of the methodology.²⁵ The defect and requirement systematization is unique (their detailed descriptions are not provided in the reviewed sources). It is especially worth noting the differentiation between critical and noncritical noncompliance, which is useful for software developers and users. Documents from the Institute of Data Sciences of the American College of Radiologists, which describe the clinical tasks solved using the software and the expected input and output data, are well known on a global scale.²⁶

Another important advantage of the methodology is the mandatory software calibration using local data (calibration testing stage) and subsequent validation using real-world data (periodic monitoring stage). According to a foreign systematic review [22], only 6% of AI-based software passed

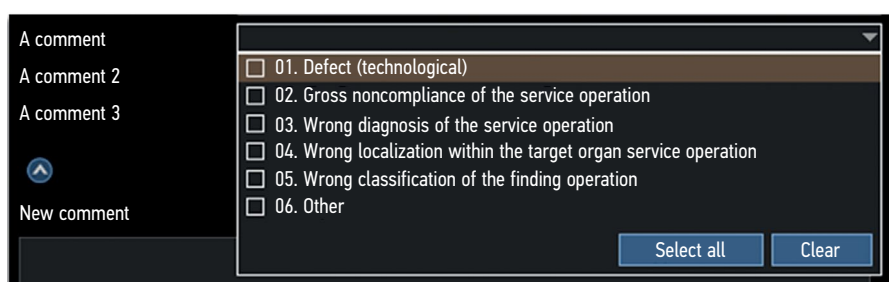


Fig. 11. A feedback window in the user interface.

²³ Order No. 134 of the Moscow City Health Department dated February 16, 2023 on approval of the procedure and conditions for conducting an experiment on the use of innovative technologies in computer vision for analyzing medical images and further use in the Moscow healthcare system. Link: https://mosmed.ai/documents/227/order_DZM__134_d_16.02.2023.pdf.

²⁴ Decree No. 1543-PP of the Moscow Government dated November 21, 2019 Link: <https://docs.cntd.ru/document/563879961>; Order No. 134 of the Moscow Department of Health dated February 16, 2023. Link: https://mosmed.ai/documents/227/order_DZM__134_d_02/16/2023.pdf.

²⁵ Basic functional requirements for AI service results Link: https://mosmed.ai/documents/218/Basic_functional_requirements_29.11.2022.pdf; Basic diagnostic requirements for AI service results Link: https://mosmed.ai/documents/226/Basic_diagnostic_requirements_22_02_2023.pdf.

²⁶ ACR Data Science Institute Releases Landmark Artificial Intelligence Use Cases. 2018. Link: <https://www.acr.org/Media-Center/ACR-News-Releases/2018/ACR-Data-Science-Institute-Releases-Landmark-Artificial-Intelligence-Use-Cases>.

the external validation stages. Validation can be “broad” and “narrow” [8]. The purpose of “narrow” validation is to determine the “correctness” of the product, that is, to what extent the results of its use correspond to the purposes of its use. This may include clinical validation and usability assessment. “Broad” validation encompasses “narrow” validation and is also associated with quality control, which ensures that software was developed following best practices and methods. This includes algorithm analysis, software testing, and documentation research. In this case, the internal structure of the software is assessed, and it is designated as a “white box” [8].

Moreover, it is important to mention the stage of software finalization after identifying critical inconsistencies. Software finalization provides a gradual decrease in the number of technological defects and an increase in software diagnostic accuracy. Therefore, the methodology will enable developers of AI-based software for medical diagnostics to achieve excellent results in various areas. Users will be able to make an informed and confident choice among software products that have passed an independent quality check, leading to the implementation of software in clinical practice, reducing the workload of radiologists, and increasing the efficiency of diagnostic examination interpretation. As a result, the initial goal of AI-based process automation will be achieved.

This methodology does not replace established medical device registration procedures. Moreover, the entire method or its stages and approaches may be used by regulatory authorities to assess the safety and effectiveness of AI-based software, and it may also be part of a manufacturer’s quality management system. The methodology can be used by software developers to prepare a post-registration clinical monitoring plan (which must be submitted as part of a set of documents when registering medical devices) and by healthcare organizations to select the most suitable software for specific conditions and purposes [4]. The methodology is indefinitely applicable, and it complies with the requirements of the Eurasian Economic Commission for 3 year and FDA recommendations for monitoring throughout the entire period of product operation.

Having MA for AI-based software does not eliminate the need to perform all stages of testing in accordance with the presented methodology. Such an approach is justified for at least two reasons. First, a MA may have been obtained by testing specific diagnostic equipment, and the results of the software may change when performed on other equipment. Second, a MA could be obtained to solve a specific clinical task; software developers could add functionalities in the future.

Our paper presented cases from radiologist practice, but the methodology may be adapted to AI-based software used in other areas of clinical medicine. In this case, adjusting certain forms, such as a list of technological defects and a clinical assessment, will be necessary.

Limitations of the study

A limitation of the methodology is the separation of a manufacturer and an assessor. In several methods, software is developed and assessed by one company (concept-to-implementation methodology) [16]. In the case of the presented method, the software is assessed by a third party closer to implementation. Errors a developer makes early in development may still be identified, but correcting them may be more challenging.

The software assesses several examinations at the periodic monitoring stage (>1,000). Due to limited resources, a small number of medical experts, and their high workload, it is impossible to provide quality control for all examinations. Despite the automated generation of a representative pseudo-random sample of examinations, systematic sampling errors may cause errors to be undetected during the periodic monitoring stage.

Research prospects

1. Publication of software evaluation results using the presented methodology (hypothesis: software evaluation based on the presented methodology improves diagnostic accuracy and practical relevance of AI-based software in medical diagnostics).
2. Comparison of software that received and did not receive Roszdravnadzor MAs using the presented methodology.
3. Developing a testing stage as part of the methodology to evaluate software processing results of “unsatisfactory” examinations (with unsuitable anatomical regions, modality, artifacts, improper patient positioning, implants, and other unsuitable medical equipment for this software).

CONCLUSION

A methodology for testing and monitoring AI-based software for medical diagnostics has been developed to improve its quality and implement it in clinical practice. The method consists of seven stages: self-testing, functional testing, calibration testing, process monitoring, clinical monitoring, feedback, and finalization. The methodology is characterized by the presence of cyclical stages of testing, monitoring, and software finalization, which results in continuous improvement in software quality, the availability of explicit requirements for software results, and the involvement of HCPs in software evaluation. The methodology will enable software developers to achieve excellent results and demonstrate achievements in various areas. Users will be able to make an informed and confident decision among software products that have passed an independent and comprehensive quality check.

ADDITIONAL INFORMATION

Funding source. This article was prepared by a group of authors as a part of the research and development effort titled

“Development of a platform for improving the quality of AI services for clinical diagnostics,” No. 123031400006-0 in accordance with the Order No. 1196 dated December 21, 2022 “On approval of state assignments funded by means of allocations from the budget of the city of Moscow to the state budgetary (autonomous) institutions subordinate to the Moscow Health Care Department, for 2023 and the planned period of 2024 and 2025” issued by the Moscow Health Care Department.

Competing interests. The authors declare that they have no competing interests.

Authors' contribution. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation

of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work. Yu.A. Vasiliev — development of the concept, approval of the final version of the manuscript; A.V. Vladzimirsky — development of the concept, approval of the final version of the manuscript; O.V. Omelyanskaya — development of methodology, approval of the final version of the manuscript; K.M. Arzamasov — concept development, research, editing and approval of the final version of the manuscript; S.F. Chetverikov — development of methodology, research; D.A. Rummyantsev — literature review, writing and editing the text of the article; M.A. Zelenova — editing the text of the article.

REFERENCES

- Oakden-Rayner L, Palme LJ. Artificial intelligence in medicine: Validation and study design. In: Ranschart E, Morozov S, Algra P, eds. Artificial intelligence in medical imaging. Cham: Springer; 2019. P. 83–104.
- Morozov SP, Zinchenko VV, Khoruzhaya AN, et al. Standardization of artificial intelligence in healthcare: Russia is becoming a leader. *Doctor Inform Technol.* 2021;(2):12–19. (In Russ). doi: 10.25881/18110193_2021_2_12
- Mello AA, Utkin LV, Trofimova TN. Artificial intelligence in medicine: The current state and main directions of development of intellectual diagnostics. *Radiation Diagnost Therapy.* 2020;(1):9–17. (In Russ). doi: 10.22328/2079-5343-2020-11-1-9-17
- Zinchenko VV, Arzamasov KM, Chetverikov SF, et al. Methodology of post-registration clinical monitoring for software using artificial intelligence technologies. *Modern Technol Med.* 2022;14(5):15–25. (In Russ). doi: 10.17691/stm2022.14.5.02
- Tanguay W, Acar P, Fine B, et al. Assessment of radiology artificial intelligence software: A validation and evaluation framework. *Can Assoc Radiol J.* 2023;74(2):326–333. doi: 10.1177/08465371221135760
- Kohli A, Jha S. Why CAD failed in mammography. *J Am Coll Radiol.* 2018;15(3 Pt B):535–537. doi: 10.1016/j.jacr.2017.12.029
- Recht MP, Dewey M, Dreyer K, et al. Integrating artificial intelligence into the clinical practice of radiology: Challenges and recommendations. *Eur Radiol.* 2020;30(6):3576–3584. doi: 10.1007/s00330-020-06672-5
- Higgins DC, Johnher C. Validation of artificial intelligence containing products across the regulated healthcare industries. *Ther Innov Regul Sci.* 2023;57(4):797–809. doi: 10.1007/s43441-023-00530-4
- Rudolph J, Schachtner B, Fink N, et al. Clinically focused multi-cohort benchmarking as a tool for external validation of artificial intelligence algorithm performance in basic chest radiography analysis. *Sci Rep.* 2022;12(1):12764. doi: 10.1038/s41598-022-16514-7
- Allen B, Dreyer K, Stibolt R, et al. Evaluation and real-world performance monitoring of artificial intelligence models in clinical practice: Try it, buy it, check it. *J Am Coll Radiol.* 2021;18(11):1489–1496. doi: 10.1016/j.jacr.2021.08.022
- Strohman L, Hehakaya C, Ranschaert ER, et al. Implementation of artificial intelligence (AI) applications in radiology: Hindering and facilitating factors. *Eur Radiol.* 2020;30(10):5525–5532. doi: 10.1007/s00330-020-06946-y
- Sohn JH, Chillakuru YR, Lee S, et al. An open-source, vendor agnostic hardware and software pipeline for integration of artificial intelligence in radiology workflow. *J Digit Imaging.* 2020;33(4):1041–1046. doi: 10.1007/s10278-020-00348-8
- Wichmann JL, Willeminck MJ, De Cecco CN. Artificial intelligence and machine learning in radiology: Current state and considerations for routine clinical implementation. *Invest Radiol.* 2020;55(9):619–627. doi: 10.1097/RLI.0000000000000673
- Larson DB, Harvey H, Rubin DL, et al. Regulatory frameworks for development and evaluation of artificial intelligence-based diagnostic imaging algorithms: Summary and recommendations. *J Am Coll Radiol.* 2021;18(3 Pt A):413–424. doi: 10.1016/j.jacr.2020.09.060
- Milam ME, Koo CW. The current status and future of FDA-approved artificial intelligence tools in chest radiology in the United States. *Clin Radiol.* 2023;78(2):115–122. doi: 10.1016/j.crad.2022.08.135
- De Silva D, Alahakoon D. An artificial intelligence life cycle: From conception to production. *Patterns (NY).* 2022;3(6):100489. doi: 10.1016/j.patter.2022.100489
- Cerdá-Alberich L, Solana J, Mallol P, et al. MAIC-10 brief quality checklist for publications using artificial intelligence and medical images. *Insights Imaging.* 2023;14(1):11. doi: 10.1186/s13244-022-01355-9
- Vasey B, Novak A, Ather S, et al. DECIDE-AI: A new reporting guideline and its relevance to artificial intelligence studies in radiology. *Clin Radiol.* 2023;78(2):130–136. doi: 10.1016/j.crad.2022.09.131
- Regulations for the preparation of data sets with a description of approaches to the formation of a representative sample of data. Moscow: Scientific and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Department of Health of the City of Moscow; 2022. 40 p. (Best practices in radiological and instrumental diagnostics; Part 1). (In Russ).
- Chetverikov S, Arzamasov KM, Andreichenko AE, et al. Approaches to sampling for quality control of artificial intelligence systems in biomedical research. *Modern Technol Med.* 2023;15(2):19–27. (In Russ). doi: 10.17691/stm2023.15.2.02

21. Morozov SP, Vladzimirsky AV, Klyashtorny VG, et al. Clinical trials of software based on intelligent technologies (radiation diagnostics). Moscow: Scientific and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Department of Health of the City of Moscow; 2019. 33 p. (In Russ).

22. Kim DW, Jang HY, Kim KW, et al. Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: Results from recently published papers. *Korean J Radiol.* 2019;20(3):405–410. doi: 10.3348/kjr.2019.0025

СПИСОК ЛИТЕРАТУРЫ

1. Oakden-Rayner L., Palme L.J. Artificial intelligence in medicine: Validation and study design. In: Ranschart E., Morozov S., Algra P., eds. *Artificial intelligence in medical imaging.* Cham: Springer, 2019. P. 83–104.

2. Морозов С.П., Зинченко В.В., Хоружая А.Н., и др. Стандартизация искусственного интеллекта в здравоохранении: Россия выходит в лидеры // *Врач и информационные технологии.* 2021. № 2. С. 12–19. doi: 10.25881/18110193_2021_2_12

3. Мелдо А.А., Уткин Л.В., Трофимова Т.Н. Искусственный интеллект в медицине: современное состояние и основные направления развития интеллектуальной диагностики // *Лучевая диагностика и терапия.* 2020. № 1. С. 9–17. doi: 10.22328/2079-5343-2020-11-1-9-17

4. Зинченко В.В., Арзамасов К.М., Четвериков С.Ф., и др. Методология проведения пострегистрционного клинического мониторинга для программного обеспечения с применением технологий искусственного интеллекта // *Современные технологии в медицине.* 2022. Т. 14, № 5. С. 15–25. doi: 10.17691/stm2022.14.5.02

5. Tanguay W., Acar P., Fine B., et al. Assessment of radiology artificial intelligence software: A validation and evaluation framework // *Can Assoc Radiol J.* 2023. Vol. 74, N 2. P. 326–333. doi: 10.1177/08465371221135760

6. Kohli A., Jha S. Why CAD failed in mammography // *J Am Coll Radiol.* 2018. Vol. 15, N 3, Pt. B. P. 535–537. doi: 10.1016/j.jacr.2017.12.029

7. Recht M.P., Dewey M., Dreyer K., et al. Integrating artificial intelligence into the clinical practice of radiology: Challenges and recommendations // *Eur Radiol.* 2020. Vol. 30, N 6. P. 3576–3584. doi: 10.1007/s00330-020-06672-5

8. Higgins D.C., Johner C. Validation of artificial intelligence containing products across the regulated healthcare industries // *Ther Innov Regul Sci.* 2023. Vol. 57, N 4. P. 797–809. doi: 10.1007/s43441-023-00530-4

9. Rudolph J., Schachtner B., Fink N., et al. Clinically focused multi-cohort benchmarking as a tool for external validation of artificial intelligence algorithm performance in basic chest radiography analysis // *Sci Rep.* 2022. Vol. 12, N 1. P. 12764. doi: 10.1038/s41598-022-16514-7

10. Allen B., Dreyer K., Stibolt R., et al. Evaluation and real-world performance monitoring of artificial intelligence models in clinical practice: Try it, buy it, check it // *J Am Coll Radiol.* 2021. Vol. 18, N 11. P. 1489–1496. doi: 10.1016/j.jacr.2021.08.022

11. Strohm L., Hehakaya C., Ranschaert E.R., et al. Implementation of artificial intelligence (AI) applications in radiology: Hindering and facilitating factors // *Eur Radiol.* 2020. Vol. 30, N 10. P. 5525–5532. doi: 10.1007/s00330-020-06946-y

12. Sohn J.H., Chillakuru Y.R., Lee S., et al. An open-source, vendor agnostic hardware and software pipeline for integration of artificial

intelligence in radiology workflow // *J Digit Imaging.* 2020. Vol. 33, N 4. P. 1041–1046. doi: 10.1007/s10278-020-00348-8

13. Wichmann J.L., Willeminck M.J., De Cecco C.N. Artificial intelligence and machine learning in radiology: Current state and considerations for routine clinical implementation // *Invest Radiol.* 2020. Vol. 55, N 9. P. 619–627. doi: 10.1097/RLI.0000000000000673

14. Larson D.B., Harvey H., Rubin D.L., et al. Regulatory frameworks for development and evaluation of artificial intelligence-based diagnostic imaging algorithms: Summary and recommendations // *J Am Coll Radiol.* 2021. Vol. 18, N 3, Pt. A. P. 413–424. doi: 10.1016/j.jacr.2020.09.060

15. Milam M.E., Koo C.W. The current status and future of FDA-approved artificial intelligence tools in chest radiology in the United States // *Clin Radiol.* 2023. Vol. 78, N 2. P. 115–122. doi: 10.1016/j.crad.2022.08.135

16. De Silva D., Alahakoon D. An artificial intelligence life cycle: From conception to production // *Patterns (NY).* 2022. Vol. 3, N 6. P. 100489. doi: 10.1016/j.patter.2022.100489

17. Cerdá-Alberich L., Solana J., Mallol P., et al. MAIC-10 brief quality checklist for publications using artificial intelligence and medical images // *Insights Imaging.* 2023. Vol. 14, N 1. P. 11. doi: 10.1186/s13244-022-01355-9

18. Vasey B., Novak A., Ather S., et al. DECIDE-AI: A new reporting guideline and its relevance to artificial intelligence studies in radiology // *Clin Radiol.* 2023. Vol. 78, N 2. P. 130–136. doi: 10.1016/j.crad.2022.09.131

19. Регламент подготовки наборов данных с описанием подходов к формированию репрезентативной выборки данных. Москва: Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы, 2022. 40 с. (Лучшие практики лучевой и инструментальной диагностики; Часть 1).

20. Четвериков С.Ф., Арзамасов К.М., Андрейченко А.Е., и др. Подходы к формированию выборки для контроля качества работы систем искусственного интеллекта в медико-биологических исследованиях // *Современные технологии в медицине.* 2023. Т. 15, № 2. С. 19–27. doi: 10.17691/stm2023.15.2.02

21. Морозов С.П., Владзими́рский А.В., Кляшторный В.Г., и др. Клинические испытания программного обеспечения на основе интеллектуальных технологий (лучевая диагностика). Москва: Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы, 2019. 33 с.

22. Kim D.W., Jang H.Y., Kim K.W., et al. Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: Results from recently published papers // *Korean J Radiol.* 2019. Vol. 20, N 3. P. 405–410. doi: 10.3348/kjr.2019.0025

AUTHORS' INFO

*** Denis A. Rumyantsev;**

address: 24/1 Petrovka street, 127051 Moscow, Russia;
ORCID: 0000-0001-7670-7385;
eLibrary SPIN: 8734-2085;
e-mail: RumyantsevDA3@zdrav.mos.ru

Yuriy A. Vasilev, MD, Cand. Sci. (Med.);

ORCID: 0000-0002-0208-5218;
eLibrary SPIN: 4458-5608;
e-mail: npcmmr@zdrav.mos.ru

Anton V. Vladzimirskyy, MD, Dr. Sci. (Med.);

ORCID: 0000-0002-2990-7736;
eLibrary SPIN: 3602-7120;
e-mail: npcmmr@zdrav.mos.ru

Olga V. Omelyanskaya;

ORCID: 0000-0002-0245-4431;
eLibrary SPIN: 8948-6152;
e-mail: npcmmr@zdrav.mos.ru

Kirill M. Arzamasov, MD, Cand. Sci. (Med.);

ORCID: 0000-0001-7786-0349;
eLibrary SPIN: 3160-8062;
e-mail: ArzamasovKM@zdrav.mos.ru

Sergei F. Chetverikov, Cand. Sci. (Engin.);

ORCID: 0000-0002-3097-8881;
eLibrary SPIN: 3815-8870;
e-mail: ChetverikovSF@zdrav.mos.ru

Maria A. Zelenova;

ORCID: 0000-0001-7458-5396;
eLibrary SPIN: 3823-6872;
e-mail: ZelenovaMA@zdrav.mos.ru

ОБ АВТОРАХ

*** Румянцев Денис Андреевич;**

адрес: Россия, 127051, Москва, ул. Петровка, д. 24, стр. 1;
ORCID: 0000-0001-7670-7385;
eLibrary SPIN: 8734-2085;
e-mail: RumyantsevDA3@zdrav.mos.ru

Васильев Юрий Александрович, канд. мед. наук;

ORCID: 0000-0002-0208-5218;
eLibrary SPIN: 4458-5608;
e-mail: npcmmr@zdrav.mos.ru

Владимирский Антон Вячеславович, д-р мед. наук;

ORCID: 0000-0002-2990-7736;
eLibrary SPIN: 3602-7120;
e-mail: npcmmr@zdrav.mos.ru

Омелянская Ольга Васильевна;

ORCID: 0000-0002-0245-4431;
eLibrary SPIN: 8948-6152;
e-mail: npcmmr@zdrav.mos.ru

Арзамасов Кирилл Михайлович, канд. мед. наук;

ORCID: 0000-0001-7786-0349;
eLibrary SPIN: 3160-8062;
e-mail: ArzamasovKM@zdrav.mos.ru

Четвериков Сергей Федорович, канд. тех. наук;

ORCID: 0000-0002-3097-8881;
eLibrary SPIN: 3815-8870;
e-mail: ChetverikovSF@zdrav.mos.ru

Зеленова Мария Александровна;

ORCID: 0000-0001-7458-5396;
eLibrary SPIN: 3823-6872;
e-mail: ZelenovaMA@zdrav.mos.ru

* Corresponding author / Автор, ответственный за переписку