

МЕТОДОЛОГИЯ ПРОВЕДЕНИЯ ПОСТРЕГИСТРАЦИОННОГО КЛИНИЧЕСКОГО МОНИТОРИНГА ДЛЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ С ПРИМЕНЕНИЕМ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

DOI: 10.17691/stm2022.14.5.02

УДК 616:004.891.3

Поступила 27.06.2022 г.

- © **В.В. Зинченко**, начальник сектора клинических и технических испытаний отдела инновационных технологий;
К.М. Арзамасов, к.м.н., руководитель отдела медицинской информатики, радиомики и радиогеномики;
С.Ф. Четвериков, к.т.н., инженер сектора разработки систем внедрения медицинских интеллектуальных технологий отдела медицинской информатики, радиомики и радиогеномики;
А.В. Мальцев, к.т.н., старший научный сотрудник сектора клинических и технических испытаний отдела инновационных технологий;
В.П. Новик, научный сотрудник сектора разработки систем внедрения медицинских интеллектуальных технологий отдела медицинской информатики, радиомики и радиогеномики;
Е.С. Ахмад, научный сотрудник сектора клинических и технических испытаний отдела инновационных технологий;
Д.Е. Шарова, исполняющий обязанности руководителя отдела инновационных технологий;
А.Е. Андрейченко, к.ф.-м.н., ведущий научный сотрудник отдела медицинской информатики, радиомики и радиогеномики;
А.В. Владимирский, д.м.н., зам. директора по научной работе;
С.П. Морозов, д.м.н., профессор, врач

Научно-практический клинический центр диагностики и телемедицинских технологий
Департамента здравоохранения города Москвы, ул. Петровка, 24, стр. 1, Москва, 127051

Цель исследования — разработать методологию проведения пострегистрационного клинического мониторинга программного обеспечения на основе технологий искусственного интеллекта (ПО с ИИ), являющегося медицинским изделием.

Материалы и методы. В основу методологии пострегистрационного клинического мониторинга положены требования нормативных правовых актов Коллегии Евразийской экономической комиссии. Согласно данным требованиям, в рамках мониторинга должны проводиться анализ сообщений о неблагоприятных событиях, анализ периодических отчетов производителей о безопасности и эффективности ПО с ИИ, а также оценка системы сбора и анализа данных производителя о безопасности и эффективности медицинского изделия на пострегистрационном этапе. При разработке методологии также учитывали рекомендации International Medical Device Regulators Forum, документы Food and Drug Administration (США). Разработанная методология апробирована на ПО с ИИ, предназначенном для применения в области лучевой диагностики.

Результаты. Выделено три основных этапа пострегистрационного мониторинга ПО с ИИ: сбор обратной связи от пользователей, технологический мониторинг и клиническая валидация. Технологический мониторинг заключается в регулярной проверке качества выходных данных ПО с ИИ для своевременного выявления и устранения дефектов, а также в оценке стабильности работы продукта. Важным результатом служат упорядочивание перечня технологических дефектов ПО с ИИ и их классификация на примере лучевой диагностики. Использование данной методологии привело к постепенному снижению числа исследований с дефектами за счет своевременных доработок алгоритмов искусственного интеллекта: при последующих проверках отмечалось снижение числа дефектов до 5% по различным направлениям. Клиническая валидация подтвердила способность ПО с ИИ выдавать клинически значимые выходные данные, связанные с его целевым использованием в рамках установленного производителем функционального назначения. В процессе апробации определен порядок проведения данного испытания и основные требования к нему.

Заключение. Разработанная методология позволит обеспечить безопасность и эффективность применения ПО с ИИ с учетом специфики данных продуктов как нематериальных медицинских изделий. Приведенная в данной статье методология может использоваться разработчиками ПО с ИИ в процессе составления плана и при выполнении пострегистрационного клинического мониторинга.

Ключевые слова: искусственный интеллект; медицинское программное обеспечение; пострегистрационный клинический мониторинг.

Для контактов: Зинченко Виктория Валерьевна, e-mail: v.zinchenko@nrcmr.ru

Как цитировать: Zinchenko V.V., Arzamasov K.M., Chetverikov S.F., Maltsev A.V., Novik V.P., Akhmad E.S., Sharova D.E., Andreychenko A.E., Vladzmyrskyy A.V., Morozov S.P. Methodology for conducting post-marketing surveillance of software as a medical device based on artificial intelligence technologies. *Sovremennye tehnologii v medicine* 2022; 14(5): 15, <https://doi.org/10.17691/stm2022.14.5.02>

English

Methodology for Conducting Post-Marketing Surveillance of Software as a Medical Device Based on Artificial Intelligence Technologies

V.V. Zinchenko, Head of the Clinical and Technical Trials Sector, Department of Innovative Technologies;
K.M. Arzamasov, MD, PhD, Head of the Department of Medical Informatics, Radiomics and Radiogenomics;
S.F. Chetverikov, PhD, Engineer, Sector of the Development of Systems for the Implementation of Medical Intelligent Technologies, Department of Medical Informatics, Radiomics and Radiogenomics;
A.V. Maltsev, PhD, Senior Researcher, Sector of Clinical and Technical Trials, Department of Innovative Technologies;
V.P. Novik, Researcher, Sector of the Development of Systems for the Implementation of Medical Intelligent Technologies, Department of Medical Informatics, Radiomics and Radiogenomics;
E.S. Akhmad, Researcher, Sector of Clinical and Technical Trials, Department of Innovative Technologies;
D.E. Sharova, Deputy Head of the Department of Innovative Technologies;
A.E. Andreychenko, PhD, Leading Researcher, Department of Medical Informatics, Radiomics and Radiogenomics;
A.V. Vladzmyrskyy, MD, DSc, Deputy Director for Science;
S.P. Morozov, MD, DSc, Professor, Radiologist

Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department, Bldg 1, 24 Petrovka St., Moscow, 127051, Russia

The aim of the study was to develop a methodology for conducting post-registration clinical monitoring of software as a medical device based on artificial intelligence technologies (SaMD-AI).

Materials and Methods. The methodology of post-registration clinical monitoring is based on the requirements of regulatory legal acts issued by the Board of the Eurasian Economic Commission. To comply with these requirements, the monitoring involves submission of the review of adverse events reports, the review of developers' routine reports on the safety and efficiency of SaMD-AI, and the assessment of the system for collecting and analyzing developers' post-registration data on the safety and efficiency of medical devices. The methodology was developed with regard to the recommendations of the International Medical Device Regulators Forum and the documents issued by the Food and Drug Administration (USA). Field-testing of this methodology was carried out using SaMD-AI designed for diagnostic imaging.

Results. The post-registration monitoring of SaMD-AI consists of three key stages: collecting user feedback, technical monitoring and clinical validation. Technical monitoring involves routine evaluation of SaMD-AI output data quality to detect and remove flaws in a timely manner, and to secure the product stability. Major outcomes include an ordered list of technical flaws in SaMD-AI and their classification using evidence from diagnostic imaging studies. The application of this methodology resulted in a gradual reduction in the number of studies with flaws due to timely improvements in artificial intelligence algorithms: the number of flaws decreased to 5% in various aspects during subsequent testing. Clinical validation confirmed that SaMD-AI is capable of producing clinically meaningful outputs related to its intended use within the functionality determined by the developer. The testing procedure and the baseline testing framework were established during the field testing.

Conclusion. The developed methodology will ensure the safety and efficiency of SaMD-AI taking into account its specifics as intangible medical devices. The methodology presented in this paper can be used by SaMD-AI developers to plan and carry out the post-registration clinical monitoring.

Key words: artificial intelligence; medical software; a post-marketing surveillance.

Введение

Мониторингу безопасности медицинских изделий (МИ), который является одним из элементов государственного контроля за обращением МИ, отводится важная роль как во всем мире [1–3], так и в России

[4, 5]. Данный вид мониторинга нацелен на обеспечение безопасности и эффективности применения МИ в реальной практике. При использовании МИ высокого (третьего) класса риска к мониторингу предъявляются повышенные требования. Так, в решении Коллегии Евразийской экономической комиссии [6] определено,

что для МИ третьего класса риска мониторинг необходимо проводить ежегодно в течение трех лет после получения регистрационного удостоверения даже в случае невозникновения неблагоприятных событий или рисков со стороны продукта. Данный вид мониторинга носит название пострегистрационного клинического мониторинга (ПКМ).

В программном обеспечении с применением технологий искусственного интеллекта (ПО с ИИ) используются технологические решения, позволяющие имитировать когнитивные функции человека и получать результаты, сопоставимые как минимум с результатами интеллектуальной деятельности человека [7, 8]. ПО с ИИ требует особого контроля в процессе эксплуатации, так как в нем отсутствует интерпретируемость данных, возможно смещение результатов при использовании на популяции, отличной от применяемой при обучении ИИ, и т.д. [9, 10]. Кроме того, в отличие от программного обеспечения, являющегося МИ без технологий искусственного интеллекта, оно может в том числе включать глубокие нейронные сети с непрерывным (само)обучением [11]. В связи с этим ПО с ИИ отнесено к высокому — третьему — классу риска [12], а выполнение ПКМ является обязательным после регистрации МИ ежегодно в течение трех лет. Данная процедура должна повысить доверие к ПО с ИИ со стороны пользователей, а также обеспечить контроль уровня его безопасности и эффективности на протяжении всего периода эксплуатации (total product life cycle) [13–15].

Однако требования и рекомендации по проведению ПКМ в опубликованных документах являются общими для МИ и не учитывают специфику ПО с ИИ, требующего определенного подхода к данному виду контроля. Такой контроль должен обеспечивать как эффективность, так и безопасность готового продукта в реальной клинической практике.

Известны подходы к установлению требований к производителям ПО с ИИ по обеспечению должного уровня эффективности и безопасности продукта (например, Good Machine Learning Practice [16]). С целью определения возможных изменений ПО с ИИ его разработчики должны создавать документ «План управления изменениями» (configuration and change management), определенный ГОСТ Р МЭК 62304—2013, или Predetermined Change Control Plan из документации Food Drug Administration (FDA, США) [16].

В рамках ПКМ важно подтвердить безопасность, эффективность использования ПО с ИИ в клинической практике, а также получить и проанализировать обратную связь от пользователей. И хотя для МИ методология ПКМ уже разработана и отлажена, включая механизмы контроля и проверок, получения обратной связи от пользователей, для ПО с ИИ требуется разработка собственной методологии.

Целью данной работы является создание методологии пострегистрационного клинического мониторинга работы программного обеспечения с применением

технологий искусственного интеллекта, а также ее апробация.

Материалы и методы

В основу создания методологии ПКМ работы ПО с ИИ в рамках национального нормативного регулирования положены требования, утвержденные решением Коллегии Евразийской экономической комиссии [6]. Так, в данном решении указано, что мониторинг должен включать анализ сообщений о неблагоприятных событиях, анализ периодических отчетов производителей о безопасности и эффективности ПО с ИИ, а также оценку системы сбора и анализа данных производителя о безопасности и эффективности МИ на пострегистрационном этапе. ПКМ проводится в соответствии с планом, составленным изготовителем ПО с ИИ на этапе разработки МИ. Данный план должен включать цели и задачи, а также схему ПКМ с подробным обоснованием применяемых методов, с характеристикой популяции, критериями включения/исключения и др.

Кроме того, проанализированы рекомендации зарубежных профессиональных сообществ и регулирующих органов. Так, были рассмотрены рекомендации International Medical Device Regulators Forum (IMDRF), разработанные для ПО в качестве МИ [14, 17], документы FDA [11, 18], а также дополнительные литературные источники [2, 13, 15, 19, 20].

С учетом рекомендаций из приведенных источников выделены несколько этапов ПКМ.

1. Обратная связь. С целью анализа сообщений о неблагоприятных событиях должна быть организована система сбора и анализа обратной связи.

2. Технологический мониторинг. Для составления периодических отчетов о безопасности ПО с ИИ производители должны выполнять контроль продукта на отсутствие дефектов, которые могут привести к возникновению неблагоприятных событий.

3. Клиническая валидация. При проверке эффективности ПО с ИИ в процессе эксплуатации необходимо подтвердить метрики точности продукта по итогам тестирования на наборе данных.

Предлагаемая методология ПКМ работы ПО с ИИ была апробирована в рамках эксперимента по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы (далее — Эксперимент) [21–24]. Некоторые результаты апробации методологии ПКМ, полученные в ходе Эксперимента, приведены в данной статье в качестве иллюстрации.

Результаты

Предлагаемая в данной статье методология ПКМ для оценки работы ПО с ИИ показана на рис. 1. Блок «Обратная связь» включает сбор данных о

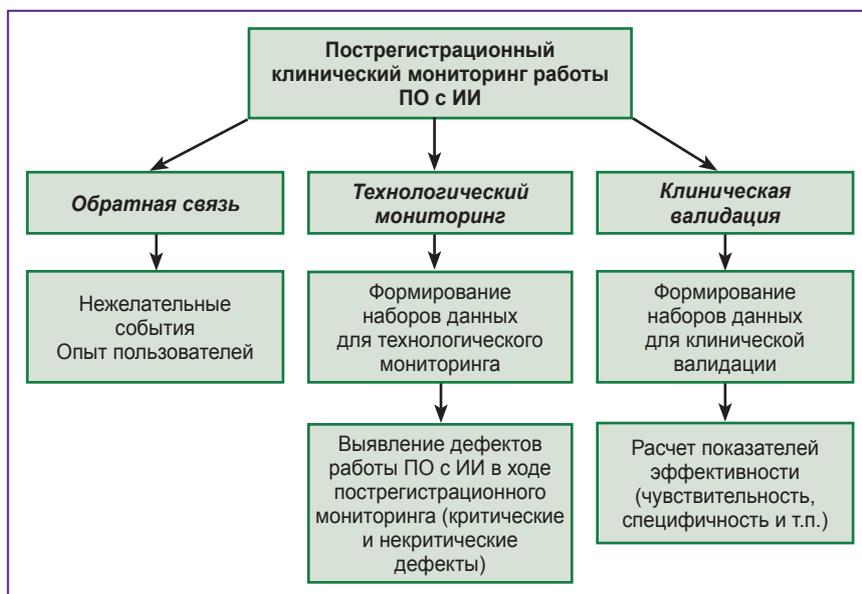


Рис. 1. Ключевые элементы пострегистрационного клинического мониторинга ПО с ИИ

неблагоприятных событиях, обратную связь от пользователей, которая в том числе может собираться посредством их анкетирования [21]. Данный этап является универсальным для всех видов МИ, и его обсуждение в связи с этим выходит за рамки данной публикации.

Анализ безопасности и эффективности ПО с ИИ в ходе мониторинга включает тестирование при исполь-

выявленной патологией («с патологией»). Такое распределение обусловлено тем, что при наличии патологии подключается дополнительный функционал — разметка патологических областей. Отобранные исследования с результатами ПО с ИИ просматриваются врачами-экспертами на наличие технологических дефектов. Исследование оценивается как «с патологией» в случае превышения вероятности

завании наборов данных для проведения технологического мониторинга и клинической валидации.

Технологический мониторинг

Технологический мониторинг представляет собой регулярную проверку (с периодичностью 1 раз в неделю или 1 раз в месяц), целью которой является контроль качества результатов работы ПО с ИИ для своевременного выявления и устранения технологических дефектов, а также оценки стабильности работы продукта.

Для выполнения такой проверки используется псевдослучайно отобранная выборка исследований с учетом следующей пропорции: 25% исследований, в которых ПО с ИИ не обнаружило патологию («без патологии»), и 75% — с

Перечень технологических дефектов для проверки в рамках пострегистрационного клинического мониторинга ПО с ИИ для лучевой диагностики

Тип дефекта	Подтип дефекта (критический/некритический)
1. Искажение изображений в результатах работы ПО с ИИ	1.1. Изображения обрезаны в диагностической области (критический)
	1.2. Изменена яркость/контрастность изображения, при этом она не подлежит коррекции (критический)
	1.3. Разметка за пределами целевого органа (критический)
2. Ошибки в результатах анализа/работы ПО с ИИ	2.1. Отсутствие результатов от ПО с ИИ (критический)
	2.2. Проанализированы не все необходимые изображения; проанализирована некорректная проекция или серия (критический)
	2.3. Отсутствие текстового заключения (критический)
	2.4. Наличие двух противоречащих текстовых заключений с результатами работы от одного ПО с ИИ (критический)
	2.5. Отсутствие предупреждающей надписи: «Исследование обработано ПО с ИИ. Требуется подтверждение врачом» (критический)
	2.6. Противоречие информации в текстовом заключении и в изображении от ПО с ИИ после обработки (критический)
	2.7. Отсутствие названия или сведений о версии ПО с ИИ (некритический)
	2.8. Отсутствие при выводе результатов от ПО с ИИ оригинальной серии изображений совместно с обработанными ПО с ИИ (если это было предусмотрено технической документацией на ПО с ИИ) (некритический)
3. Превышение среднего времени, затрачиваемого на анализ	3.1. Время обработки изображения больше времени, затрачиваемого врачом (критический)

наличия патологии оптимального порогового значения [24]. В противном случае исследование считается «без патологии».

По результатам предварительного анализа технологических дефектов нескольких ПО с ИИ, задействованных в Эксперименте, выделены основные категории и подкатегории возникающих ошибок (см. таблицу). Данные категории являются основными, поскольку влияют на безопасность и эффективность ПО с ИИ при эксплуатации пользователем: искажение изображений в результатах работы ПО с ИИ, ошибки работы ПО с ИИ, а также превышение времени обработки данных. Перечень был дополнен характеристиками дефектов, такими как критический/некритический. При этом критическим считали дефект, при наличии которого использование продукции по назначению практически невозможно или недопустимо (ГОСТ 15467—79).

В условиях научно-практического Эксперимента установлено, что доля исследований с технологическими дефектами не должна превышать 10% от набора данных, используемого в рамках проводимого тестирования. Такое требование также соответствует положениям п. 5.2 ГОСТ Р ИСО 2859-1—2007.

За время апробации процедуры технологического мониторинга участниками Эксперимента проведено более 550 проверок ПО с ИИ, из которых 60% приходится на компьютерную томографию, 28% — на рентгенографию/флюорографию и 12% — на маммографию.

Использование данной методологии привело к постепенному снижению числа исследований с дефектами за счет своевременных доработок алгоритмов искусственного интеллекта, и, как следствие, при последующих проверках отмечалось снижение числа дефектов до 5% для различных ПО с ИИ.

Рассмотрим технологический мониторинг ПО с ИИ на примере использования его в рентгенографии органов грудной клетки. Результаты для четырех изображений представлены на рис. 2–5, где отражены встречающиеся технологические дефекты при проверке ПО с ИИ, приведенные в таблице.

На рис. 2 представлены дефект 1.2 (изменена яркость/контрастность изображения, при этом она не подлежит коррекции) и дефект 2.8 (отсутствие при выводе результатов от ПО с ИИ оригинальной серии изображений совместно с обработанными ПО с ИИ). ПО с ИИ изменило оригинальную версию и контрастность исходного изображения таким образом, что исследуемая область не визуализируется и становится невозможным идентифицировать результаты работы ПО с ИИ.

На рис. 3 представлен технологический дефект 2.2 (проанализированы не все необходимые изображения, проанализирована некорректная проекция или серия). ПО с ИИ обработало и наложило тепловую карту на боковую проекцию рентгенографических снимков. Однако, согласно функциональному назна-

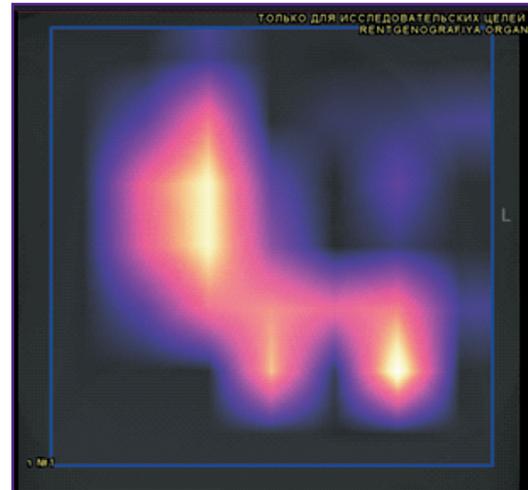


Рис. 2. Пример технологических дефектов (1.2 и 2.8, см. таблицу) работы ПО с ИИ в рентгенографии органов грудной клетки

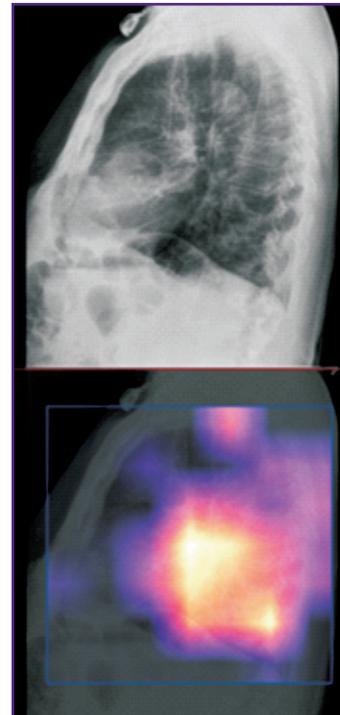


Рис. 3. Пример технологического дефекта (2.2, см. таблицу) работы ПО с ИИ в рентгенографии органов грудной клетки

чению ПО с ИИ, корректная обработка возможна только для прямой проекции рентгенограмм.

На рис. 4 представлен технологический дефект 1.3 (разметка за пределами целевого органа). По результатам работы ПО с ИИ выявило находки вне области интереса — вне органа грудной клетки.

На рис. 5 представлены технологический дефект 2.5 (отсутствие предупреждающей надписи: «Исследование обработано ПО с ИИ. Требуется

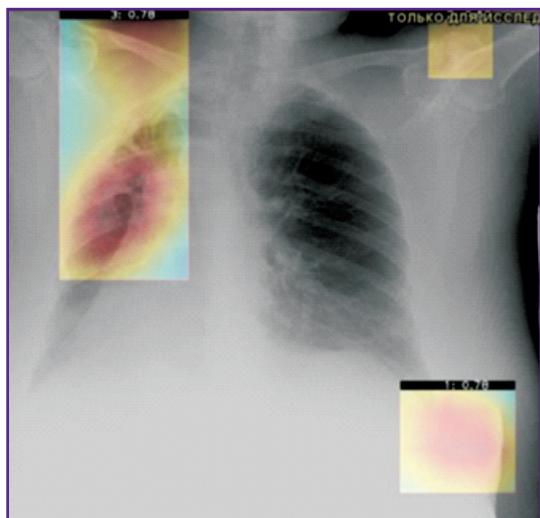


Рис. 4. Пример технологического дефекта (1.3, см. таблицу) работы ПО с ИИ в рентгенографии органов грудной клетки

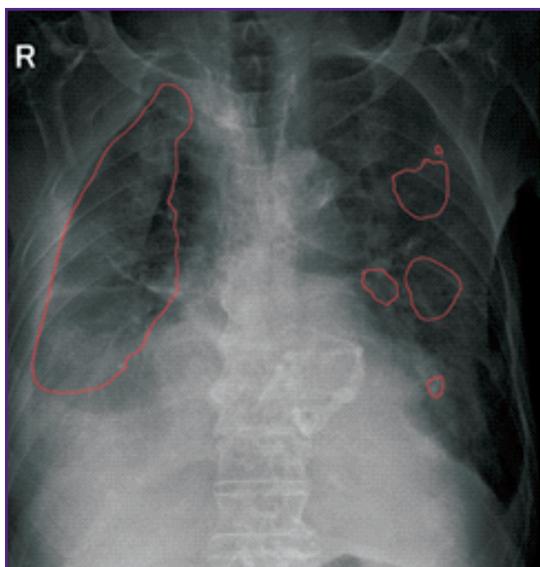


Рис. 5. Пример технологических дефектов (2.5 и 2.7, см. таблицу) работы ПО с ИИ в рентгенографии органов грудной клетки

подтверждение врачом») и дефект 2.7 (отсутствие названия или сведений о версии ПО с ИИ). Последствия таких дефектов заключаются в следующем: пациент может просмотреть данный снимок без уточняющей записи, что это результаты работы ПО с ИИ, а врач еще с ними не ознакомился и не подтвердил или опровергнул их, вследствие чего может возникнуть путаница или излишнее беспокойство со стороны пациента.

При анализе результатов апробации технологического мониторинга в Эксперименте за период с сентября по ноябрь 2021 г. было выявлено среднее число данных дефектов 13% (стандартное отклонение — 4,2%). При этом преобладали дефекты «отсутствие тексто-

го заключения» (28%) и «разметка за пределами целевого органа» (26%), которые относятся к критическим.

Клиническая валидация

Клиническая валидация подтверждает способность ПО с ИИ выдавать клинически значимые выходные данные, связанные с его целевым использованием в рамках установленного производителем функционального назначения (рис. 6) [17].

Для каждого ПО с ИИ должны быть заданы параметры эффективности (чувствительность, специфичность, точность и др.) и для проверки в рамках ПКМ переданы экспертам, проводящим тестирование. Набор показателей, определяемых в рамках клинической валидации, может изменяться в зависимости от назначения и функционала ПО с ИИ [23].

Важно отметить, что при тестировании ПО с ИИ выдается чаще всего не окончательное решение по наличию патологии на изображении, а количественный параметр, определяющий вероятность обнаружения патологии на каждом конкретном снимке, $p \in [0, 1]$. Наличие патологии на изображении признается в том случае, если величина p превышает заданное пороговое значение T . Чувствительность, специфичность и другие параметры эффективности тестируемого ПО с ИИ зависят от выбранного значения T .

При выборе оптимального порогового значения следует руководствоваться клинической задачей, на решение которой направлена работа данного ПО

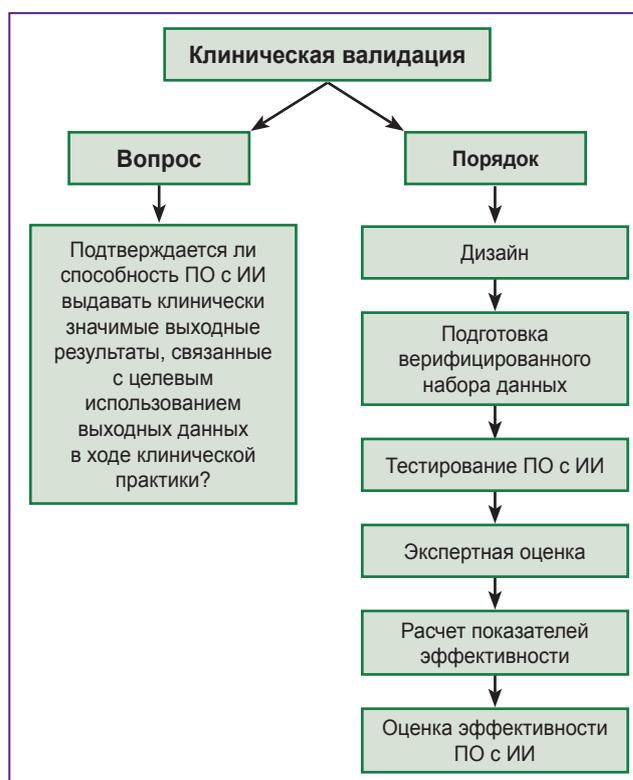


Рис. 6. Порядок выполнения клинической валидации ПО с ИИ

с ИИ. Так, например, для задач скрининга необходимо повышать чувствительность ПО с ИИ, поскольку клиническая задача заключается в минимизации пропуска патологии. Для этих целей может быть рекомендовано рассчитывать пороговое значение методом оценки предсказательной ценности для отрицательных результатов диагностического теста ($\max NPV$). Но данное ПО с ИИ с выставленным таким образом пороговым значением может быть бесполезно для врачей при работе в качестве экспертной системы поддержки принятия врачебных решений, когда необходим баланс между чувствительностью и специфичностью. В этом случае рекомендуется использовать пороговое значение, определенное методом максимизации индекса Юдена [23, 25]. Для узкоспециализированных задач выявления патологии может быть использован метод оценки предсказательной ценности для положительных результатов диагностического теста ($\max PPV$). Возможно также рассчитать порог для задания определенной чувствительности или специфичности ПО с ИИ в зависимости от конкретных условий. Выбор оптимального порога выполняют путем анализа характеристической кривой (ROC-кривая). Анализируя разные точки на кривой, выбирают значение порога, соответствующее наибольшему значению индекса Юдена: $Y = \text{sen} + \text{spe} - 1$, где sen — чувствительность, spe — специфичность. После определения оптимального порога производят расчет метрик эффективности ПО с ИИ [26]. С целью стандартизации получаемых результатов используются метрики диагностической точности, рассчитанные для метода Юдена.

Формирование набора данных для проведения клинической валидации. В целях оценки эффективности работы ПО с ИИ в рамках клинической валидации выполняют тестирование с применением наборов данных, объем которых достаточен для получения результатов с заявленным уровнем точности. Это отличает данный этап проверки от тестирования на наличие технологических дефектов.

Объем набора данных (т.е. размер выборки) для оценки характеристик ПО с ИИ устанавливаются с помощью методик определения размера выборки для пропорций и с учетом правил статистики (например, с использованием метода установления требуемой точности для оцениваемых чувствительности и специфичности, которые заявляет производитель) [27, 28]. Подробно процесс формирования наборов данных представлен в работах [29–31].

1. Определение целей и задач. Должна быть определена цель формирования набора данных, только тогда можно оценить, является ли доступ к данным или другая деятельность по их обработке доступными:

- какие данные допустимо собирать;
- как их следует использовать (применительно к каким задачам);
- кому их следует раскрывать (доступ третьих лиц);

в течение какого времени они должны быть доступны.

Постановка задачи подготовки набора данных должна включать определение предметной области и выбор методов обработки.

2. Получение одобрения комитета по этике (при необходимости).

3. Организация доступа к набору данных. Процесс организации доступа должен быть задокументирован, должны быть обеспечены процессы защиты данных, в том числе персональной информации, согласно действующим нормативным правовым актам. Организация доступа должна обеспечивать скорость передачи данных, соответствующую целям и задачам такого доступа.

4. Сбор данных. Включает представление медицинских данных (феноменов, синдромов, заболеваний, исходов) согласно их частоте встречаемости и заболеваемости в популяции (если это определено целью испытаний). Объем выборки и частота встречаемости должны быть определены в ходе статистического расчета в соответствии с целью формирования набора данных.

5. Деидентификация. Любая персональная информация должна быть удалена как из метаданных, так и из исходных данных.

6. Структурирование набора данных. Может быть снижена размерность набора данных.

7. Фильтрация. Этап фильтрации набора данных позволяет снизить затраты на разметку данных за счет исключения не соответствующих заданным параметрам.

8. Разметка (аннотация) данных. Виды аннотаций представлены в работах [29, 30].

9. Организация хранения и доступа к верифицированному набору данных. Хранение данных может быть организовано на локальном сервере или с использованием облачного хранения (ГОСТ Р ИСО/МЭК 17826—2015).

Пример клинической валидации. В рамках работы Научно-практического клинического центра диагностики и телемедицинских технологий Сектором клинических и технических испытаний апробирована разработанная методология ПКМ, которая включает клиническую валидацию ПО с ИИ.

Первый этап клинической валидации — формирование верифицированного набора данных, т.е. набора с подтвержденной медицинской информацией.

С целью расчета соотношений изображений (норма/патология) в наборе данных (выборки) для проверки заявленной чувствительности подбираются изображения, содержащие патологические признаки, во время как для оценки специфичности — изображения, не содержащие указанных в документации производителя патологий (признаков).

С целью формирования первоначального набора данных использовали клинические данные из Единого радиологического информационного сервиса Единой

медицинской информационно-аналитической системы города Москвы (ЕРИС ЕМИАС), полученные путем отбора по следующим параметрам: наименование процедуры, вид диагностического устройства, вид медицинской организации, дата, возраст пациента. Использованы данные с диагностических аппаратов 18 различных моделей, 4 из которых относятся к категории флюорографов (670 исследований), остальные 14 — к категории рентгенодиагностических аппаратов (866 исследований). Для отобранных исследований из ЕРИС ЕМИАС также выгружены заключения врачей-рентгенологов. Сформированный предварительный набор данных был визуально проанализирован врачами-рентгенологами из группы исследователей. В ходе анализа исключены:

спорные изображения, не содержащие выраженных признаков патологий, перечисленных в документации производителя, и которые нельзя достоверно отнести к категории «норма»;

изображения с недостаточным техническим качеством (низкая контрастность и т.п.).

Верификацию сформированного набора данных проводили врачи-рентгенологи из группы исследователей, имеющие ученую степень и опыт работы по данной специализации. Кроме того, выполнены процедуры согласно приведенному ранее алгоритму, в том числе деидентификация.

Второй этап клинической валидации — оценка работы ПО с ИИ с верифицированным набором данных. Данный этап проводили с использованием программного обеспечения Python. Выполняли расчет параметров эффективности ПО с ИИ (чувствительность, специфичность), а также сопоставление рассчитанных значений с показателями точности, заявленными производителем ПО с ИИ.

Благодаря проведенному анализу в ходе клинической валидации удалось подтвердить заявленные производителем характеристики и сделать достоверные выводы об эффективности и безопасности ПО с ИИ.

Обсуждение

Предлагаемая в данной статье методология ПКМ обеспечивает мониторинг безопасности и эффективности ПО с ИИ согласно установленным требованиям нормативной документации, при этом учитывает специфику ПО с ИИ как инновационного продукта.

В существующих нормативных документах приведены общие требования к данному виду мониторинга МИ и обусловлена необходимость разработки конкретных требований, применимых к ПО с ИИ. Разработанная схема ПКМ работы ПО с ИИ отвечает общемировым требованиям по проверке и тестированию на соответствие МИ, но с учетом специфики ПО с ИИ. Так, в работе Y. Park с соавт. [32] говорится о том, что важно проводить не только тестирование на наличие технологических дефектов, но и оценку, сосредоточенную на клинической практике.

Приведенные в настоящей статье этапы гарантируют комплексный мониторинг ПО с ИИ, который удовлетворяет всем требованиям методологии, представленной в работе [32].

В статье S. Venjmens и соавт. [9] подняты актуальные вопросы быстрого обновления ПО с ИИ как нематериального МИ и, соответственно, увеличенной по времени повторной регистрации ПО с ИИ. По решению FDA для данного вида МИ необходима уникальная система регистрации, заключающаяся в рассмотрении полного жизненного цикла продукта. Методология должна учитывать возможность внесения изменений в процессе адаптации и эксплуатации ПО с ИИ, при этом обеспечивая его безопасность и эффективность в качестве МИ [11]. В статье C.J. Kelly с соавт. [10] рассматриваются шаги, необходимые для обеспечения безопасной, но в то же время недлительной регистрации ПО с ИИ и внедрения данных продуктов в клиническую практику. Авторы отмечают, что важно и нужно проводить оценку работы ПО с ИИ на реальных клинических данных и сравнение с результатами предыдущих оценок с целью исключения дрейфа характеристик ПО с ИИ. Разработчики ПО с ИИ должны проявлять бдительность в отношении потенциальных опасностей, включая проблемы применения для новых групп пациентов и непреднамеренные негативные последствия для показателей здоровья. Необходимо анализировать не только основные показатели эффективности работы (метрики) ПО с ИИ (площадь под характеристической кривой), но и прогнозные показатели, т.е. положительные и отрицательные прогностические значения.

Кроме указанных рекомендаций FDA относительно мониторинга ПО с ИИ недавно выпущено руководство Управления медицинских наук Сингапура [33]. Согласно данному документу, также требуется проводить мониторинг ПО с ИИ, включающий сбор и анализ данных реальной клинической практики, результаты которого должны периодически направляться в регулирующий орган. В Европейском союзе также утвержден документ, который включает аналогичные требования, предъявляемые к мониторингу любого медицинского программного обеспечения [2]. Необходимым является периодический сбор и анализ данных для оценки безопасности и эффективности продукта в реальной практике, а также анализ обратной связи. Таким образом, прослеживается тенденция следовать одним и тем же требованиям, предъявляемым к ПКМ как с использованием медицинского программного обеспечения в общем (что отмечено в документе Европейского союза), так и ПО с ИИ в частности, как это указано в FDA и требованиях Управления медицинских наук Сингапура. Наша работа также включает классификацию возможных ошибок ПО с ИИ, которая используется в рамках технологического мониторинга, что представляет собой более расширенную версию предложенных научных подходов.

Важность этапа технологического мониторинга ПО с ИИ, относящегося к программному обеспечению про-

дуктов, которые являются МИ, отмечается в нескольких публикациях FDA. В программное обеспечение вносят изменения после его регистрации в качестве МИ [34]. По статистике FDA, 79% выявленных ошибок программного обеспечения на пострегистрационном этапе связаны как раз с внесением изменений в него [18]. Отмечается также, что большая часть ошибок, связанных с программным обеспечением, наблюдается у МИ со средним риском [35]. Представленная в данной статье методология систематизации дефектов ПО с ИИ является уникальной, так как в рассмотренных источниках не приведены детальные описания возможных дефектов, которые могут возникать в процессе работы ПО с ИИ. Особенно ценным представляется выполненное в данной работе разделение дефектов ПО с ИИ на критические и некритические. Это может быть использовано как разработчиками ПО с ИИ, так и пользователями при оценке наличия неблагоприятных событий, на которые стоит обратить внимание в процессе формирования отчета о ПКМ, а также в процессе составления и оценки обратной связи от пользователей.

Предложенные подходы ПКМ могут быть реализованы для мониторинга безопасности и эффективности ПО с ИИ со стороны регулирующих органов, а также как часть системы менеджмента качества производителей. ISO/TR 20416:2020 [36] рекомендует обеспечивать связь плана ПКМ со всеми процессами в рамках контроля качества ПО с ИИ (менеджмент риска, клиническая оценка и др.). Для того, чтобы обеспечить прослеживаемость результатов ПО с ИИ, также важно проводить анализ и сравнение результатов текущего ПКМ с предыдущими данными [37], периодически оценивать тренды параметров эффективности и обратной связи, а также технологического мониторинга для своевременного внесения корректирующих действий и обеспечения безопасности и эффективности ПО с ИИ.

Приведенная в данной статье методология может применяться разработчиками ПО с ИИ в процессе составления и выполнения плана ПКМ, который должен быть представлен в комплекте документов на регистрацию МИ. Хотя Евразийская экономическая комиссия устанавливает необходимость проведения такого мониторинга и предоставления отчетов производителями ПО с ИИ в регулирующие органы в течение 3 лет, FDA рекомендует проводить такой мониторинг на всем протяжении периода эксплуатации продукта.

Приведенные в данной работе примеры относятся к области лучевой диагностики, однако описанная методология ПКМ может быть применима для ПО с ИИ, используемого в любых областях клинической медицины, в которой применяются медицинские данные [29, 31]. Изменения в основном потребуются в формировании списка технологических дефектов и их классификации, так как данная информация будет специфичной для каждой области медицины.

Необходимые этапы ПКМ работы ПО с ИИ из пока-

занной в статье методологии представлены в серии национальных стандартов, посвященных клинической медицине [30, 38–41]: в этих документах приведен план ПКМ, цели и задачи, адаптированные именно для ПО с ИИ.

Заключение

Выполнение пострегистрационного клинического мониторинга регламентировано действующими нормативными правовыми актами, однако требования, как правило, установлены для медицинских изделий в целом. Предложенная методология пострегистрационного клинического мониторинга работы программного обеспечения на основе технологий искусственного интеллекта на примере лучевой диагностики включает оценку обратной связи, технологический мониторинг на наличие дефектов в работе программного обеспечения на основе технологий искусственного интеллекта, а также оценку эффективности применения такого программного обеспечения путем выполнения клинической валидации. Осуществление такого мониторинга позволит обеспечить безопасность и эффективность применения программного обеспечения на основе технологий искусственного интеллекта с учетом специфики данных продуктов как нематериальных медицинских изделий.

Информация об источниках финансирования.

Данная статья подготовлена авторским коллективом в рамках научно-исследовательской работы (№ЕГИСУ: АААА-А21-121012290079-2) в соответствии с программой Департамента здравоохранения города Москвы «Научное обеспечение столичного здравоохранения» на 2020–2022 годы.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Литература/References

1. *Guidance for post-market surveillance and market surveillance of medical devices, including in-vitro-diagnostics*. URL: https://www.who.int/docs/default-source/essential-medicines/in-vitro-diagnostics/draft-public-pmsdevices.pdf?sfvrsn=f803f68a_2.
2. *Guidance on Clinical Evaluation (MDR) / Performance Evaluation (IVDR) of medical device software*. URL: https://ec.europa.eu/health/sites/default/files/md_sector/docs/md_mdcg_2020_1_guidance_clinic_eva_md_software_en.pdf.
3. *Postmarket Surveillance Under Section 522 of the Federal Food, Drug, and Cosmetic Act. Guidance for Industry and Food and Drug Administration Staff*. URL: <https://www.fda.gov/media/81015/download>.
4. *Федеральный закон от 21.11.2011 №323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации» (с изм. и доп., вступ. в силу с 01.10.2021)*.
Federal'nyy zakon ot 21.11.2011 No.323-FZ "Ob osnovakh okhrany zdorov'ya grazhdan v Rossiyskoy Federatsii" (s izm. i dop., vstup. v silu s 01.10.2021) [Federal Law No.323-FZ of November 21, 2011 "On the fundamentals of protecting the

health of citizens in the Russian Federation” (as amended and supplemented, effective from October 1, 2021)].

5. Приказ Минздрава России от 14.09.2012 №175н «Об утверждении Порядка осуществления мониторинга безопасности медицинских изделий» (зарегистрирован Минюстом России 25.12.2012 №26356).

Prikaz Minzdrava Rossii ot 14.09.2012 No.175n "Ob utverzhdenii Poryadka osushchestvleniya monitoringa bezopasnosti meditsinskikh izdeliy" (zaregistririvan Minyustom Rossii 25.12.2012 No.26356) [Order of the Ministry of Health of Russia dated September 14, 2012 No.175n "On approval of the procedure for monitoring the safety of medical devices" (registered by the Ministry of Justice of Russia on December 25, 2012 No.26356)].

6. Решение Коллегии Евразийской экономической комиссии от 22 декабря 2015 г. №174 «Об утверждении Правил проведения мониторинга безопасности, качества и эффективности медицинских изделий».

Reshenie Kollegii Evraziyskoy ekonomicheskoy komissii ot 22 dekabrya 2015 g. No.174 "Ob utverzhdenii Pravil provedeniya monitoringa bezopasnosti, kachestva i effektivnosti meditsinskikh izdeliy" [Decision of the Board of the Eurasian Economic Commission dated December 22, 2015 No.174 "On approval of the Rules for monitoring the safety, quality and effectiveness of medical devices"].

7. Карпов О.Э., Клименко Г.С., Лебедев Г.С. Применение интеллектуальных систем в здравоохранении. *Современные наукоемкие технологии* 2016; 7–1: 38–43.

Karpov O.E., Klimenko G.S., Lebedev G.S. Application of intelligent systems in health care. *Sovremennyye naukoemkie tehnologii* 2016; 7–1: 38–43.

8. Гусев А.В., Добридюк С.Л. Искусственный интеллект в медицине и здравоохранении. *Информационное общество* 2017; 4–5: 78–93.

Gusev A.V., Dobridnyuk S.L. Artificial intelligence in medicine and healthcare. *Informacionnoe obsestvo* 2017; 4–5: 78–93.

9. Benjamens S., Dhunoo P., Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digit Med* 2020; 3(1): 118, <https://doi.org/10.1038/s41746-020-00324-0>.

10. Kelly C.J., Karthikesalingam A., Suleyman M., Corrado G., King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med* 2019; 17(1): 195, <https://doi.org/10.1186/s12916-019-1426-2>.

11. *Proposed Regulatory Framework for Modifications to Artificial Intelligence Machine Learning (AI ML)-Based Software as a Medical Device (SaMD). Discussion paper and request for feedback.* URL: <https://www.fda.gov/media/122535/download>.

12. Приказ МЗ РФ от 6 июня 2012 г. №4н «Об утверждении номенклатурной классификации медицинских изделий».

Prikaz MZ RF ot 6 iyunya 2012 g. No.4n "Ob utverzhdenii nomenklaturnoy klassifikatsii meditsinskikh izdeliy" [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"].

13. Harvey H.B., Gowda V. How the FDA regulates AI. *Acad Radiol* 2020; 27(1): 58–61, <https://doi.org/10.1016/j.acra.2019.09.017>.

14. Shuren J.; IMDRF Chair. "Software as a Medical Device": possible framework for risk categorization and corresponding considerations. 2014. URL: <https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf>.

15. Higgins D.C. OnRAMP for regulating artificial intelligence in medical products. *Adv Intell Syst* 2021; 3(11): 2100042, <https://doi.org/10.1002/aisy.202100042>.

16. *Software as a Medical Device (SaMD) Action Plan.* URL: <https://www.fda.gov/media/145022/download>.

17. *Software as a Medical Device (SaMD): Clinical Evaluation.* URL: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/software-medical-device-samd-clinical-evaluation>.

18. *General Principles of Software Validation; Final Guidance for Industry and FDA Staff.* URL: <https://www.fda.gov/files/medical%20devices/published/General-Principles-of-Software-Validation---Final-Guidance-for-Industry-and-FDA-Staff.pdf>.

19. Higgins D., Madai V.I. From bit to bedside: a practical framework for artificial intelligence product development in healthcare. *Adv Intell Syst* 2020; 2: 2000052, <https://doi.org/10.1002/aisy.202000052>.

20. Sounderajah V., Ashrafian H., Golub R.M., Shetty S., De Fauw J., Hooft L., Moons K., Collins G., Moher D., Bossuyt P.M., Darzi A., Karthikesalingam A., Denniston A.K., Mateen B.A., Ting D., Treanor D., King D., Greaves F., Godwin J., Pearson-Stuttard J., Harling L., McInnes M., Rifai N., Tomasev N., Normahani P., Whiting P., Aggarwal R., Vollmer S., Markar S.R., Panch T., Liu X.; STARD-AI Steering Committee. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. *BMJ Open* 2021; 11(6): e047709, <https://doi.org/10.1136/bmjopen-2020-047709>.

21. Морозов С.П., Владимировский А.В., Ледихова Н.В., Андрейченко А.Е., Арзамасов К.М., Баланюк Э.А., Гомболевский В.А., Ермолаев С.О., Живоденко В.С., Идрисов И.М., Кирпичев Ю.С., Логунова Т.А., Нуждина В.А., Омелянская О.В., Раковчен В.Г., Слепушкина А.В. Московский эксперимент по применению компьютерного зрения в лучевой диагностике: вовлеченность врачей-рентгенологов. *Врач и информационные технологии* 2020; 4: 14–23, <https://doi.org/10.37690/1811-0193-2020-4-14-23>.

Morozov S.P., Vladymyrskyu A.V., Ledikhova N.V., Andrejchenko A.E., Arzamasov K.M., Balanjuk E.A., Gombolevskij V.A., Ermolaev S.O., Zhivodenko V.S., Idrisov I.M., Kirpichev Ju.S., Logunova T.A., Nuzhdina V.A., Omeljanskaja O.V., Rakovchen V.G., Slepushkina A.V. Moscow experiment on computer vision in radiology: involvement and participation of radiologists. *Vrac i informacionnye tehnologii* 2020; 4: 14–23, <https://doi.org/10.37690/1811-0193-2020-4-14-23>.

22. *Эксперимент по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы.* URL: <https://mosmed.ai/ai/>.

Eksperiment po ispol'zovaniyu innovatsionnykh tekhnologii v oblasti komp'yuternogo zreniya dlya analiza meditsinskikh izobrazheniy i dal'neyshego primeneniya v sisteme zdravookhraneniya goroda Moskvu [Experiment on the use of innovative computer vision technologies for medical image analysis and subsequent applicability in the healthcare system of Moscow]. URL: <https://mosmed.ai/ai/>.

23. Морозов С.П., Владимировский А.В., Кляшторный В.Г., Андрейченко А.Е., Кульберг Н.С., Гомболевский В.А., Сергунова К.А. Клинические испытания программного обеспечения на основе интеллектуальных технологий (лучевая диагностика). В кн.: *Серия «Лучшие практики лучевой и инструментальной диагностики»*. Выпуск 57. М; 2019; 51 с.

Morozov S.P., Vladzimirskyy A.V., Klyashtornyy V.G., Andreychenko A.E., Kulberg N.S., Gombolevsky V.A., Sergunova K.A. Klinicheskie ispytaniya programmnoy obespecheniya na osnove intellektual'nykh tekhnologiy (luhevaya diagnostika). V kn.: *Seriya «Luchshie praktiki luhevoy i instrumental'noy diagnostiki»*. Vypusk 57 [Clinical acceptance of software based on artificial intelligence technologies (radiology). In: Series “Best practices in medical imaging”. Issue 57]. Moscow; 2019; 51 p.

24. Andreychenko A.E., Logunova T.A., Gombolevskiy V.A., Nikolaev A.E., Vladzimirskyy A.V., Sinitsyn V.E., Morozov S.P. *A methodology for selection and quality control of the radiological computer vision deployment at the megalopolis scale*. medRxiv; 2022, <https://doi.org/10.1101/2022.02.12.22270663>.

25. Ying G.S., Maguire M.G., Glynn R.J., Rosner B. Calculating sensitivity, specificity, and predictive values for correlated eye data. *Investig Ophthalmol Vis Sci* 2020; 61(11): 29, <https://doi.org/10.1167/iovs.61.11.29>.

26. Arroyo-Palacios J., Rudz M., Fidler R., Smith W., Ko N., Park S., Bai Y., Hu X. Characterization of shape differences among ICP pulses predicts outcome of external ventricular drainage weaning trial. *Neurocrit Care* 2016; 25(3): 424–433, <https://doi.org/10.1007/s12028-016-0268-4>.

27. Наркевич А.Н., Виноградов К.А. Методы определения минимально необходимого объема выборки в медицинских исследованиях. *Социальные аспекты здоровья населения* 2019; 65(6): 10.

Narkevich A.N., Vinogradov K.A. Methods for determining the minimum required sample size in medical research. *Social'nye aspekty zdorov'a naseleniya* 2019; 65(6): 10.

28. Pepe M.S. Chapter 8: study design and hypothesis testing. In: *The statistical evaluation of medical tests for classification and prediction*. New York: Oxford University Press; 2003; p. 319.

29. Павлов Н.А., Андрейченко А.Е., Владимирский А.В., Ревазян А.А., Кирпичев Ю.С., Морозов С.П. Эталонные медицинские датасеты (MosMedData) для независимой внешней оценки алгоритмов на основе искусственного интеллекта в диагностике. *Digital Diagnostics* 2021; 2(1): 49–66.

Pavlov N.A., Andreychenko A.E., Vladzimirskyy A.V., Revazyan A.A., Kirpichev Y.S., Morozov S.P. Reference medical datasets (MosMedData) for independent external evaluation of algorithms based on artificial intelligence in diagnostics. *Digital Diagnostics* 2021; 2(1): 49–66.

30. Зинченко В.В., Хоружая А.Н., Шарова Д.Е., Ахмад Е.С., Мокиенко О.А., Владимирский А.В., Морозов С.П. Стандартизация в области регулирования технологий искусственного интеллекта в российском здравоохранении. *Казанский медицинский журнал* 2021; 102(6): 923–933.

Zinchenko V.V., Khoruzhaya A.N., Sharova D.E., Akhmad E.S., Mokienko O.A., Vladzimirskyy A.V., Morozov S.P. Standardization in regulating artificial intelligence systems in Russian healthcare. *Kazanskij medicinskij zurnal* 2021; 102(6): 923–933.

31. ГОСТ Р 59921.5—2022 «Системы искусственного интеллекта в клинической медицине. Часть 5. Требования к структуре и порядку применения набора данных для обучения и тестирования алгоритмов».

GOST R 59921.5—2022 “Sistemy iskusstvennogo intellekta v klinicheskoy meditsine. Chast' 5. Trebovaniya k strukture i poryadku primeneniya nabora dannykh dlya obucheniya i testirovaniya algoritmov” [GOST R 59921.5—2022 “Artificial

intelligence systems in clinical medicine. Part 5. Requirements for the structure and order of using a data set for training and testing algorithms”].

32. Park Y., Jackson G.P., Foreman M.A., Gruen D., Hu J., Das A.K. Evaluating artificial intelligence in medicine: phases of clinical research. *JAMIA Open* 2020; 3(3): 326–331, <https://doi.org/10.1093/jamiaopen/ooaa033>.

33. Health Sciences Authority Singapore. *Regulatory Guidelines for Software Medical Devices — a Life Cycle Approach. Revision 2.0*. URL: [https://www.hsa.gov.sg/docs/default-source/hprg-mdb/guidance-documents-for-medical-devices/regulatory-guidelines-for-software-medical-devices---a-life-cycle-approach_r2-\(2022-apr\)-pub.pdf](https://www.hsa.gov.sg/docs/default-source/hprg-mdb/guidance-documents-for-medical-devices/regulatory-guidelines-for-software-medical-devices---a-life-cycle-approach_r2-(2022-apr)-pub.pdf).

34. Zinchenko V., Chetverikov S., Akhmad E., Arzamasov K., Vladzimirskyy A., Andreychenko A., Morozov S. Changes in software as a medical device based on artificial intelligence technologies. *Int J Comput Assist Radiol Surg* 2022; 17(10): 1969–1977, <https://doi.org/10.1007/s11548-022-02669-1>.

35. Ronquillo J.G., Zuckerman D.M. Software-related recalls of health information technology and other medical devices: implications for FDA regulation of digital health. *Milbank Q* 2017; 95(3): 535–553, <https://doi.org/10.1111/1468-0009.12278>.

36. ISO/TR 20416:2020. *Medical devices — Post-market surveillance for manufacturers*. URL: <https://www.iso.org/standard/67942.html>.

37. Article 78 — *Post-market surveillance system of the manufacturer*. URL: https://lexpacency.org/eu/32017R0746/ART_78/.

38. ГОСТ Р 59921.2—2021 «Системы искусственного интеллекта в клинической медицине. Часть 2. Программа и методика технических испытаний».

GOST R 59921.2—2021 “Sistemy iskusstvennogo intellekta v klinicheskoy meditsine. Chast' 2. Programma i metodika tekhnicheskikh ispytaniy” [GOST R 59921.2—2021 “Artificial intelligence systems in clinical medicine. Part 2. Program and methodology of technical validation”].

39. ГОСТ Р 59921.3—2021 «Системы искусственного интеллекта в клинической медицине. Часть 3. Управление изменениями в системах искусственного интеллекта с непрерывным обучением».

GOST R 59921.3—2021 “Sistemy iskusstvennogo intellekta v klinicheskoy meditsine. Chast' 3. Upravlenie izmeneniyami v sistemakh iskusstvennogo intellekta s nepreryvnyim obucheniem” [GOST R 59921.3—2021 “Artificial intelligence systems in clinical medicine. Part 3. Change management in AI systems with lifelong learning capability”].

40. ГОСТ Р 59921.4—2021 «Системы искусственного интеллекта в клинической медицине. Часть 4. Оценка и контроль эксплуатационных параметров».

GOST R 59921.4—2021 “Sistemy iskusstvennogo intellekta v klinicheskoy meditsine. Chast' 4. Otsenka i kontrol' ekspluatatsionnykh parametrov” [GOST R 59921.4—2021 “Artificial intelligence systems in clinical medicine. Part 4. Evaluation and control of performance parameters”].

41. ГОСТ Р 59921.6—2021 «Системы искусственного интеллекта в клинической медицине. Часть 6. Общие требования к эксплуатации».

GOST R 59921.6—2021 “Sistemy iskusstvennogo intellekta v klinicheskoy meditsine. Chast' 6. Obshchie trebovaniya k ekspluatatsii” [GOST R 59921.6—2021 “Artificial intelligence systems in clinical medicine. Part 6. General performance requirements”].