

ВАСИЛЬЕВ Ю.А.,

к.м.н., ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия; ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России,
г. Москва, Россия, e-mail: npcmr@zdrav.mos.ru

ПАМОВА А.П.,

к.м.н., ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия,
e-mail: PamovaAP@zdrav.mos.ru

АРЗАМАСОВ К.М.,

к.м.н., ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия; МИРЭА, г. Москва, Россия,
e-mail: ArzamasovKM@zdrav.mos.ru

ВЛАДИМИРСКИЙ А.В.,

д.м.н., ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия; ПМГМУ им. И.М. Сеченова,
г. Москва, Россия, e-mail: VladzimirskijAV@zdrav.mos.ru

ЗАЮНЧОВСКИЙ С.Ю.,

ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия,
e-mail: ZayunchkovskijSY@zdrav.mos.ru

ЗИНЧЕНКО В.В.,

ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия, e-mail: ZinchenkoVW1@zdrav.mos.ru

ПРЕДСТАВЛЕНИЕ МЕТРИК ДИАГНОСТИЧЕСКОЙ ТОЧНОСТИ В ЗАВИСИМОСТИ ОТ КЛАССИФИКАЦИИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ НА ОСНОВЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ОБЛАСТИ ЛУЧЕВОЙ ДИАГНОСТИКИ

DOI: 10.25881/18110193_2025_1_58

Аннотация. Цель: настоящее исследование направлено на разработку клинической классификации программного обеспечения (ПО) на основе искусственного интеллекта (ИИ) в области лучевой диагностики.

Материалы и методы: для проведения исследования был проведён всесторонний анализ доступной информации о ПО на основе ИИ в сфере лучевой диагностики с использованием отечественных и зарубежных баз данных. В процессе анализа были выявлены ключевые аспекты, включая клиническую применимость ПО на основе ИИ, диагностическую точность медицинских изделий с использованием ИИ в лучевой диагностике.

Результаты: была разработана клиническая классификация ПО на основе ИИ в области лучевой диагностики. Кроме того, было выявлено важное замечание относительно представления метрик диагностической точности ПО на основе ИИ. В результате этого предложенная классификация была расширена и дополнена определением уровня представления метрик диагностической точности в зависимости от клинической классификации.

Заключение: на основе проведенного исследования разработана клиническая классификация ПО на основе ИИ, что обеспечивает единый подход к представлению данных о диагностической точности со стороны разработчиков. Данный подход позволяет повысить прозрачность и сравнимость информации о различных ПО на основе

ИИ в медицинской практике, что способствует повышению эффективности и безопасности использования ПО на основе ИИ в медицинской практике. Результаты настоящего исследования имеют потенциал для масштабирования на другие области применения ИИ и могут быть использованы для совершенствования системы регулирования качества медицинских изделий с применением ИИ.

Ключевые слова: информационные технологии, искусственный интеллект, классификация, лучевая-диагностика, технические характеристики, диагностическая точность.

Для цитирования: Васильев Ю.А., Памова А.П., Арзамасов К.М., Владимирский А.В., Заюнчковский С.Ю., Зинченко В.В. Представление метрик диагностической точности в зависимости от классификации программного обеспечения на основе искусственного интеллекта в области лучевой диагностики. Врач и информационные технологии. 2025; 1: 58-69. doi: 10.25881/18110193_2025_1_58.

VASILEV YU.A.,

PhD, Moscow Center for Diagnostics & Telemedicine, Moscow, Russia; Pirogov National Medical and Surgical Center, Moscow, Russia, e-mail: npcmr@zdrav.mos.ru

PAMOVA A.P.,

PhD, Moscow Center for Diagnostics & Telemedicine, Moscow, Russia, e-mail: PamovaAP@zdrav.mos.ru

ARZAMASOV K.M.,

PhD, Moscow Center for Diagnostics & Telemedicine, Moscow, Russia; MIREA – Russian Technological University, Moscow, Russia, e-mail: ArzamasovKM@zdrav.mos.ru

VLADZYMYRSKIY A.V.,

DSc, Moscow Center for Diagnostics & Telemedicine, Moscow, Russia; I.M. Sechenov First Moscow State Medical University of the Ministry of Health of the Russian Federation (Sechenov University), Moscow, Russia, e-mail: VladzimirskijAV@zdrav.mos.ru, ORCID ID: 0000-0002-2990-7736

ZAYUNCHKOVSKIY S.YU.,

Moscow Center for Diagnostics & Telemedicine, Moscow, Russia, e-mail: ZayunchkovskijSY@zdrav.mos.ru

ZINCHENKO V.V.,

Moscow Center for Diagnostics & Telemedicine, Moscow, Russia, e-mail: ZinchenkoVV1@zdrav.mos.ru

PRESENTATION OF DIAGNOSTIC ACCURACY METRICS BASED ON CLASSIFICATION OF ARTIFICIAL INTELLIGENCE SOFTWARE IN RADIOLOGY

DOI: 10.25881/18110193_2025_1_58

Abstract. *The implementation of artificial intelligence in healthcare is a key direction for technology development in Russia, aimed at improving the quality of medical services and increasing diagnostic accuracy. However, the lack of standards for presenting metrics of diagnostic accuracy of artificial intelligence-based software (AI-based software) complicates comparative analysis and selection of the most suitable software for medical organizations. Therefore, developing a detailed classification of AI-based software is an important task for ensuring safety and quality of medical care, as well as determining the interchangeability of AI-based medical devices.*

Purpose: This study aims to develop a clinical classification of AI-based software in radiology.

Materials and Methods: To conduct the study, a comprehensive analysis of available information on AI-based software in radiology was conducted using domestic and foreign databases. In the process of analysis, key aspects were identified, including clinical applicability of AI-based software, diagnostic accuracy of medical devices using AI in radiology.

Results: a clinical classification of AI-based software in the field of radiology was developed. In addition, an important observation regarding the representation of diagnostic accuracy metrics of AI-based software was identified. As a result, the proposed classification was extended and supplemented by defining the level of representation of diagnostic accuracy metrics depending on the clinical classification.

Conclusion: based on the conducted research, a clinical classification of AI-based software has been developed, which provides a unified approach to the presentation of data on diagnostic accuracy by developers. This approach improves the

transparency and comparability of information about different AI-based software in medical practice, thereby improving the efficiency and safety of AI-based software use in medical practice. The results of this study have the potential to be scaled to other AI applications and can be used to improve the quality regulation system for AI-enabled medical devices.

Keywords: *information technology, artificial intelligence, classification, radiology, technical specifications, diagnostic accuracy.*

For citation: *Vasilev Yu.A., Pamova A.P., Arzamasov K.M., Vladzimirskyy A.V., Zayunchkovskiy S.Yu., Zinchenko V.V. Presentation of diagnostic accuracy metrics based on classification of artificial intelligence software in radiology. Medical doctor and information technology. 2025; 1: 58-69. doi: 10.25881/18110193_2025_1_58.*

ВВЕДЕНИЕ

Согласно указу Президента от 10.10.2019 № 490 внедрение искусственного интеллекта (ИИ) в сферу здравоохранения стало одним из ключевых направлений развития подобных технологий в Российской Федерации (РФ). Это связано со стремлением улучшить качество медицинских услуг, повысить точность диагностики и оптимизировать процессы лечения. Разработкой программного обеспечения (ПО) на основе ИИ занимаются организации, специализирующиеся на создании цифровых продуктов для медицинской сферы.

Внедрение технологий ИИ в здравоохранение происходит после успешного прохождения технических, клинических испытаний и дальнейшей государственной регистрации ПО на основе ИИ как медицинского изделия [1], либо в составе другого медицинского изделия. Во время клинических испытаний основное внимание уделяется анализу метрик диагностической точности ПО на основе ИИ, которые заявляются в регуляторные органы со стороны разработчика.

Настоящая публикация – это обобщение полученного опыта в процессах подготовки к обращению и собственно обращения медицинских изделий на основе ПО с ИИ в части корректности оценки технических характеристик таких изделий и представления результатов их измерений.

В сфере здравоохранения наиболее широкое применение технологий ИИ наблюдается в лучевой диагностике (ПНСТ 873-2023 «Системы искусственного интеллекта в лучевой диагностике. Основные положения»), а среди ключевых метрик диагностической точности ПО на основе ИИ (под которыми в данной публикации имеются в виду метрики, указанные ниже, относящиеся к **техническим характеристикам** ПО), активно применяющихся в радиологической практике, выделяются: чувствительность (от англ. «sensitivity»); специфичность (от англ. «specificity»); точность (диагностическая, от англ. «accuracy»); площадь под характеристической кривой (ROC-AUC от англ. «area under receiver operating characteristic curve»); индекс Дайса-Соренсена («Sørensen–Dice index»).

При оценке основных технических характеристик не обойтись без понятия «точность измерений», принятого в метрологии. «Точность измерений», «точность результата измерения»

– это близость измеренного значения к истинному значению измеряемой величины. Любая измеренная техническая характеристика отражает истинное значение измеряемой величины лишь с определённой погрешностью. Следовательно, возникает вопрос о представлении результатов измерений и их округлении до приемлемых значений. В контексте здравоохранения для ПО на основе ИИ это необходимо для следующих целей:

1. Обеспечения безопасности и качества медицинской помощи;
2. Обоснования клинической значимости применяемых технологий ИИ;
3. Для правильного расчета конечного результата, поставленного в той или иной задаче;
4. Для удобства представления результатов измерений (испытаний) потребителю и другим заинтересованным сторонам;
5. Для соблюдения требований стандартов (в частности ГОСТ Р 8.736-2011) и других нормативных документов;
6. Для единообразия форм оформления документации (технической, эксплуатационной и в целом – конструкторской), а также протоколов испытаний;
7. Для совершенствования и развития методов измерений и повышения компетентности персонала.

Необходимо отметить, что в соответствии со ст. 38 Федерального закона от 21.11.2011 № 323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации» в Государственный реестр медицинских изделий и организаций, осуществляющих производство и изготовление медицинских изделий, должны вноситься сведения о взаимозаменяемых медицинских изделиях. Критериям взаимозаменяемости медицинских изделий уделяется особое внимание контролирующими и надзорными органами. Взаимозаменяемость медицинских изделий в соответствии с указанной статьей Федерального закона определяется следующим образом: медицинские изделия признаются взаимозаменяемыми, если они «сравнимы по функциональному назначению, качественным и **техническим характеристикам** и способны заменить друг друга». Приведенные выше метрики являются одними из основных технических характеристик для ПО на основе ИИ. Таким образом, вынесение решения о взаимозаменяемости ПО на основе

ИИ, являющихся самостоятельными медицинскими изделиями, невозможно без анализа этих существенных технических характеристик. Ниже в публикации будет показана значимость корректности определения метрик (технических характеристик) диагностической точности ПО на основе ИИ в клинической практике для вынесения решения о взаимозаменяемости.

В свете активно развивающейся стандартизации искусственного интеллекта в здравоохранении сформировались вопросы, которые требуют внимания как со стороны разработчиков, так и со стороны медицинского сообщества [2]. На данный момент отсутствуют стандарты представления дробной части метрик диагностической точности, которые должны быть указаны разработчиками при регистрации ПО на основе ИИ. Существующий подход, при котором каждый разработчик самостоятельно определяет с каким уровнем точности он предоставит данные регуляторным органам, может способствовать допуску к использованию в здравоохранении ПО на основе ИИ низкого качества, что представляет потенциальную угрозу здоровью пациентов. Кроме этого, данный подход препятствует сравнительному анализу ПО на основе ИИ с идентичными функциями, что усложняет процесс выбора наиболее подходящего ПО для медицинской организации.

Таким образом, акцент в данной публикации делается на разработку детализированной классификации ПО с технологиями ИИ. Эта классификация призвана служить инструментом для определения уровня представления метрик диагностической точности, который должен учитываться разработчиками в конструкторской документации на стадиях проектирования и разработки, а также указан при представлении информации в регуляторные органы здравоохранения и для конечных пользователей цифрового продукта – медицинских работников [3].

Цель данной работы заключается в создании классификации ПО на основе ИИ в области лучевой диагностики, основанной на клинической применимости и диагностической точности подобных медицинских изделий. Этот подход предполагает комплексный анализ и систематизацию критериев, определяющих эффективность и надежность ПО на основе ИИ в контексте лучевых исследований.

МАТЕРИАЛЫ И МЕТОДЫ

Данная публикация подготовлена авторским коллективом в рамках научно-исследовательской работы «Научно-методические основы цифровой трансформации службы лучевой диагностики», (№ ЕГИСУ: № 123031400118-0) в соответствии с Приказом от 22.12.2023 г. № 1258 «Об утверждении государственных заданий, финансовое обеспечение которых осуществляется за счет средств бюджета города Москвы государственным бюджетным (автономным) учреждениям подведомственным Департаменту здравоохранения города Москвы, на 2024 год и плановый период 2025 и 2026 годов» Департамента здравоохранения города Москвы.

Были проанализированы материалы, связанные с процедурами подготовки к регистрации ПО на основе ИИ (например, заявления о государственной регистрации, если они опубликованы разработчиком ПО), применяемого в области лучевой диагностики. Дополнительно были проанализированы публикации, в которых некоторые разработчики указывали метрики диагностической точности своих изделий, а также публикации, которые содержали какую-либо классификацию ПО на основе ИИ (в период с 15.05.2024 по 30.05.2024). Для ПО, зарегистрированного на территории РФ, был использован официальный ресурс Росздравнадзора, раздел «Медицинские изделия», подраздел «Регистрация медицинских изделий», «Государственная регистрация медицинских изделий». Для анализа ПО на основе ИИ, зарегистрированного за пределами РФ, был выбран веб-сайт Food and Drug Administration (FDA) Соединенных Штатов Америки. Исследование проводилось в рамках раздела «Medical Devices», подраздела «Digital Health Center of Excellence», категории «Software as a Medical Device (SaMD)», и последующего подраздела «Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices». Основное внимание было уделено ПО на основе ИИ в лучевой диагностике. Информация на сайте FDA охватывает продукты из разных стран, что позволяет избежать географического смещения выборки и обеспечивает объективность анализа.

Для расширения спектра исследуемого ПО на основе ИИ за последние 5 лет, его классификации и углубления понимания его применения в медицинской практике, был осуществлен

дополнительный поиск в научных базах данных, включая Российский индекс научного цитирования (РИНЦ), PubMed и Google Scholar. Поиск проводился с использованием ключевых слов, отражающих основные аспекты интересующей области: «искусственный интеллект», «artificial intelligence», «системы», «медицинские изделия», «medical devices», «машинное обучение», «machine learning», «медицина», «medicine», «классификация», «classification», «taxonomy», «specification», «спецификация», «metrics».

Кроме того, для сбора данных о метриках диагностической точности исследуемых ПО на основе ИИ были изучены официальные веб-ресурсы разработчиков. Этот подход позволил получить непосредственную информацию о технических характеристиках и возможностях ИИ, а также о методологиях, применяемых для оценки их эффективности в клинической практике.

Таким образом, найдено и проанализировано более 50 публикаций и веб-ресурсов. В ходе проведенного исследования были собраны данные о метриках диагностической точности как отечественного, так и зарубежного ПО на основе ИИ, используемого в лучевой диагностике. Полученная информация стала основой для выполнения сравнительного анализа, направленного на выявление различий и сходств в представлении метрик диагностической точности между различными медицинскими ПО на основе ИИ.

Критерии исключения публикаций и веб-ресурсов из последующего анализа: нет данных о метриках диагностической точности (не указаны).

РЕЗУЛЬТАТЫ

В процессе исследования обнаружена значимая проблема: разработчики программного обеспечения, основанного на технологиях ИИ, часто не раскрывают метрики диагностической точности своих продуктов. На веб-платформах, специализирующихся на медицинском оборудовании, доступная информация зачастую не отражает реальные показатели диагностической точности и носит преимущественно маркетинговый характер. Более того, разработчики ПО на основе ИИ предпочитают не публиковать научные статьи, связанные с их разработками, и чаще всего не предоставляют данные, используемые при подаче заявок на регистрацию продуктов в регуляторные органы, такие как FDA или Росздравнадзор. Существующие классификации ПО на основе ИИ, разработанные другими авторами, сосредоточены на технических характеристиках систем, в то время как клиническая применимость остается вне их внимания [4]. Это обстоятельство существенно затрудняет комплексную оценку диагностической точности различного ПО на основе ИИ и его взаимозаменяемости.

В таблице 1 представлены примеры того, как разработчики ИИ из области компьютерного

Таблица 1 — Примеры указания метрик диагностической точности со стороны разработчиков ПО на основе ИИ, зарегистрированных как медицинские изделия в Росздравнадзоре и FDA.

ПО на основе ИИ в лучевой диагностике	Коэффициент Дайса-Серенсена	Точность (Accuracy)	Коэффициент внутриклассовой корреляции	Площадь под кривой (AUC)	Чувствительность	Специфичность
iCAS [6]	0,86±0,17	–	–	–	–	–
AVIEW CAC [7]	0,58±0,08	99,2%	0,958	–	0,83	0,83
AI-Rad Companion (Pulmonary) [8]	–	0,88	–	90%	–	–
inHEART Models [9]	от 79,8% до 87,3%	–	–	–	–	–
ПроРодинки [10]	–	92%	–	–	–	–
WEBIOMED [11]	–	78%	–	0,77	–	–
Botkin.AI [12]	–	–	–	92,74	0,902 (далее 100%)	0,937 (93,7%)
Sciberia Lungs [13]	–	99%	–	–	–	–

Таблица 1 — Примеры указания метрик диагностической точности со стороны разработчиков ПО на основе ИИ, зарегистрированных как медицинские изделия в Росздравнадзоре и FDA (продолжение)

ПО на основе ИИ в лучевой диагностике	Коэффициент Дайса-Серенсёна	Точность (Accuracy)	Коэффициент внутрикласовой корреляции	Площадь под кривой (AUC)	Чувствительность	Специфичность
Платформа третье мнение [14]	–	–	–	0,910 (далее 91%)	81,7%	94,5%
ЦЕЛЬС [15]	–	–	–	91%	85%	93%
ТУ 62.01.29-001-96876180-2019 (Фтизисбиомед) [16]	–	0,92	–	0,95	0,92	0,92

зрения в лучевой диагностике раскрывают метрики диагностической точности. Следует отметить, что большая часть ключевых метрик обычно указывается с точностью до двух десятичных знаков после запятой [5].

ОБСУЖДЕНИЕ

Разработчик ПО на основе ИИ в своей публикации в разделе «проведение пилотных проектов с применением платформы ИИ» округляет значение чувствительности до ближайшего целого числа (100%), указывая эту метрику в общедоступных источниках именно в таком виде. Однако в табличных данных этой же публикации указано среднее значение чувствительности на тестовых данных в другом формате и составляет 0,902 [12]. Таким образом, 100% чувствительности (или равной 1,0) на тестовых данных достигнуто не было. Другой разработчик наоборот, указывает значения чувствительности и специфичности с очень высокой точностью, до третьего знака после запятой [14]. Если обратить внимание на индекс Дайса-Серенсёна, можно увидеть, что для ПО на основе ИИ inHEART Models представлено три знака после запятой, а для iCAS – BR/LV/LG и AVIEW CAC всего два знака.

Исходя из полученных результатов, можно говорить об определённой разобщённости в указании метрик диагностической точности различными разработчиками ПО на основе ИИ. Рассмотрим следующую ситуацию: медицинская организация (МО) внедряет ПО на основе ИИ и заявляет, что для них важно, чтобы значение AUC было $\geq 0,81$. Предположим, что при тестировании фактически достигнутое значение AUC составило 0,807, что при округлении до двух знаков после

запятой позволяет разработчику и его ПО быть допущенными к использованию в МО. Если не округлять текущее значение метрики AUC, то ПО на основе ИИ не смогло бы соответствовать указанному выше требованию. Таким образом, на основе результатов проведённого анализа была предложена классификация ПО на основе ИИ в лучевой диагностике, которая призвана уточнить требования к правилам округления и указания дробной части метрик диагностической точности. Эта классификация предназначена для использования разработчиками ПО с технологиями ИИ при регистрации программных продуктов в качестве медицинских изделий. Она направлена на обеспечение единого подхода к представлению данных о диагностической точности, что в свою очередь способствует повышению прозрачности и сравнимости информации о различных ПО на основе ИИ, используемых в медицинской практике, и, таким образом, взаимозаменяемости. Кроме того, предложенная классификация масштабируема не только на лучевую диагностику, но и на другие области здравоохранения, например, системы мониторинга, системы поддержки принятия врачебных решений.

Необходимо уточнить, что высокая степень детализации метрик не является обязательной для всего ПО на основе ИИ. Чтобы определить, какому ПО на основе ИИ действительно требуется указание максимально высокой степени точности с тремя знаками после запятой, а какому нет, мы предлагаем ориентироваться на нашу классификацию (Таблица 2) [6, 7, 9, 18]. Следует подчеркнуть, что в случае, если ПО на основе ИИ работает не автономно от врача, не используется в экстренном порядке, а также не

применяется в области поиска злокачественных новообразований, травм, нарушений кровообращения, болезней органов дыхания (в том числе COVID-19), нарушений гемостаза, заболеваний почек, достаточно указывать два знака после запятой (0.XX). В иных случаях необходимо указывать три знака (0.XXX), что более наглядно отражено в таблице 2.

Для демонстрации того факта, что в некоторых случаях важна высокая детализация указания диагностических метрик, приведём несколько примеров расчёта. По данным Всемирной организации здравоохранения (ВОЗ), одним из высоко летальных заболеваний является

туберкулёз (в том числе – лёгких) (www.who.int, дата обращения к сайту: 24.09.2024). По данным «Федеральной службы государственной статистики» – «здравоохранение» – «заболеваемость населения социально значимыми болезнями» впервые в жизни активный туберкулёз (<https://rosstat.gov.ru/folder/13721>) в 2022 году у взрослого населения России был выявлен у 45,6 тысяч человек. Пересчитаем количество больных, которых ПО на основе ИИ признает ошибочно здоровыми в случае указания диагностической метрики (чувствительность) до двух десятых, вместо трёх десятых (Таблица 3), в случае массовых диагностических исследований.

Таблица 2 — Классификация ПО на основе ИИ и представление дробной части метрик диагностической точности в зависимости от предназначения ПО



Таблица 3 — Расчёт количества людей (на 1000 человек), которых ПО на основе ИИ признает ошибочно здоровыми в зависимости от уровня чувствительности, когда на самом деле у них туберкулёз

Значение чувствительности	Количество больных туберкулёзом, которых ПО на основе ИИ признает ошибочно здоровыми на 1000 человек
0,8	200
0,84	160
0,836	164

Таблица 4 — Расчет количества пациентов (на 100 человек), которым будет ошибочно введён препарат при заданном уровне специфичности

Значение специфичности	Препарат будет введён по ошибке этому количеству человек
0,8	20
0,84	16
0,836	16,4

Таким образом, если разработчик округляет значение метрики с реальных 0,836 до чувствительности равной 0,84 – это приведёт к недооценке реального количества больных туберкулёзом. При чувствительности равной 0,836 – на самом деле будет пропущено на 4 человека с туберкулёзом больше, чем заявляет разработчик с чувствительностью 0,84. Может возникнуть вопрос – тогда почему не указывать значение метрик диагностической точности с ещё большей точностью? Данный вопрос является дискуссионным. В текущей публикации мы предлагаем остановиться на максимальных трех знаках после запятой. Однако, мы предполагаем, что всё зависит от того, в какой конкретной области применяется ПО на основе ИИ, на экономических затратах, связанных с недообследованными пациентами и другими факторами.

Однозначно, высокая точность указания диагностических метрик важна для ПО на основе ИИ, используемых в целях оказания экстренной (неотложной) помощи. Например, ИИ-система предназначена для решения о введении норадреналина в случае шокового состояния пациента без наблюдения врача реаниматолога (автономная система). Более высокий уровень специфичности даст для конечного пользователя (врач-реаниматолог) понимание того, насколько можно доверять такой ИИ-системе (Таблица 4) [18].

Таким образом, при указании со стороны разработчика всего двух знаков после запятой снова наблюдается та же ситуация, что и с чувствительностью. Одновременно с этим, указание всего одного знака после запятой сильно занижает реальную пользу ПО на основе ИИ.

Почему тогда не указывать для любого вида ПО на основе ИИ всегда три знака после запятой? Рассмотрим этот вопрос на примере ПО для морфометрических исследований. Для ПО на основе ИИ для морфометрических исследований (измерений линейных и других размеров анатомических и патологических структур) в лучевой диагностике указание двух знаков после запятой в индексе Дайса-Серенсена является достаточным и предпочтительным. Это решение основано на анализе ошибок, которые допускают сами врачи-рентгенологи при проведении исследований, а также на основании погрешности измерений, которую даёт медицинское оборудование. Двух знаков после запятой вполне достаточно, чтобы отразить реальную способность ПО на основе ИИ измерять органы и ткани без необходимости абсолютной точности. Учитывая вариативность ошибок измерения даже среди профессиональных медицинских работников, подобное представление метрик диагностической точности обеспечивает баланс между техническими характеристиками ПО на основе ИИ и его практической применимостью [19].

ЗАКЛЮЧЕНИЕ

В ходе проведения исследования отмечено, что разработчики ПО стараются не раскрывать метрики диагностической точности своих продуктов в открытом доступе, что затрудняет объективную оценку реальной значимости подобного рода медицинских изделий.

Была разработана клиническая классификация для ПО на основе ИИ в лучевой диагностике, опираясь на которую можно понять, насколько важна точность передачи значения метрик диагностической точности от разработчика ПО на основе ИИ в регуляторные органы в области здравоохранения и для конечного пользователя. Потенциально, данная классификация может быть масштабируема на другие виды медицинских изделий, включающих в себя применение ПО на основе ИИ.

Выводы, полученные в ходе исследования, могут быть использованы для разработки и

внедрения ПО на основе ИИ в области лучевой диагностики и в смежных областях, а также для совершенствования системы регулирования и контроля качества подобного рода медицинских изделий.

Финансирование. Данная статья подготовлена авторским коллективом в рамках НИР «Научно-методические основы цифровой трансформации службы лучевой диагностики», (№ ЕГИСУ: № 123031400118-0) в соответствии с Приказом от 22.12.2023 г. № 1258 «Об утверждении государственных заданий, финансовое обеспечение которых осуществляется за счет средств бюджета города Москвы государственным бюджетным (автономным) учреждениям подведомственным Департаменту здравоохранения города Москвы, на 2024 год и плановый период 2025 и 2026 годов» Департамента здравоохранения города Москвы.

ЛИТЕРАТУРА/REFERENCES

1. Базовые рекомендации к работе сервисов искусственного интеллекта для лучевой диагностики: методические рекомендации / сост. С. П. Морозов, Л. Р. Абуладзе, А. Е. Андрейченко и др. // Серия «Лучшие практики лучевой и инструментальной диагностики». – Вып. 119. – М.: ГБУЗ «НПКЦ ДиТ ДЗМ», 2022. – 68 с. [Bazovye rekomendacii k rabote servisov iskusstvennogo intellekta dlja luchevoj diagnostiki: metodicheskie rekomendacii. S. P. Morozov, L. R. Abuladze, A. E. Andrejchenko, et al. Serija «Luchshie praktiki luchevoj i instrumental'noj diagnostiki». M.: GBUZ «NPKC DiT DZM», 2022; 119: 68. (In Russ.)]
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. [Morozov SP, Zinchenko VV, Khoruzhaya AN, et al. Standardization of artificial intelligence in healthcare: Russia becomes the leader. Medical doctor and information technology. 2021; 2: 12-19. (In Russ.)] doi: 1025881/18110193_2021_2_12.
3. Васильев Ю.А., Владимирский А.В., Омелянская О.В. и др. Методология тестирования и мониторинга программного обеспечения на основе технологий искусственного интеллекта для медицинской диагностики // Digital Diagnostics. – 2023. – Т.4. – №3. – С.252-267. [Vasilev YuA, Vladymyrskyy AV, Omelyanskaya OV, et al. Methodology for testing and monitoring artificial intelligence-based software for medical diagnostics. Digital Diagnostics. 2023; 4(3): 252-267. (In Russ.)] doi: 10.17816/DD321971.
4. Кукшев В.И. Классификация систем искусственного интеллекта // Экономические стратегии. – 2020. – Т.6. – С.58-67. [Kukshev VI. Classification of artificial intelligence systems. Digital Society. 2020; 6: 58-67. (In Russ.)] doi: 10.33917/es-6.172.2020.58-67.
5. Jungmann F, Müller L, Hahn F, et al. Commercial AI solutions in detecting COVID-19 pneumonia in chest CT: not yet ready for clinical implementation? European Radiology. Springer Berlin Heidelberg, 2022; 32(5): 3152-3160. doi: 10.1007/s00330-021-08409-4.
6. High-rad.com [Internet]. iCAS. Available at: <https://www.high-rad.com>. Accessed 21.05.2024.
7. Corelinesoft.com [Internet]. AVIEW CAC. Available at: <https://www.corelinesoft.com>. Accessed 21.05.2024.
8. Ai-rad-companion [Internet]. AI-Rad Companion (Pulmonary) Available at: <https://www.siemens-healthineers.com/en-iq/digital-health-solutions/ai-rad-companion>. Accessed 21.05.2024.

9. Florkow MC, Nguyen CH, Sakkars RJB, et al. Magnetic resonance imaging-based bone imaging of the lower limb: Strategies for generating high-resolution synthetic computed tomography. *Journal of Orthopaedic Research*. 2024; 42(4): 843-854. doi: 10.1002/jor.25707.
10. ПроРодинки [интернет]. Доступно по: <https://prorodinki.ru>. Ссылка активна на 21.05.2024. [Prorodinki.ru. Available at: <https://prorodinki.ru>. Accessed 21.05.2024. (In Russ.)]
11. Webiomed.ru [интернет]. Доступно по: <https://webiomed.ru/products/webiomed-dhra/>. Ссылка активна на 21.05.2024. [Webiomed Available at: <https://webiomed.ru/products/webiomed-dhra/>. Accessed 21.05.2024. (In Russ.)]
12. Дрокин И.С., Еричева Е.В., Бухвалов О.Л. и др. Опыт разработки и внедрения системы поиска онкологических образований с помощью искусственного интеллекта на примере рентгеновской компьютерной томографии лёгких // *Врач и информационные технологии*. – 2019. – №3. – С.48-57. [Drokin IS, Elicheva EV, Bukhvalov OL, et al. Experience in developing and implementing an oncological formations searching system using artificial intelligence with the example of X ray computed tomography of the lungs. *Medical doctor and information technology*. 2019; 3: 48-57. (In Russ.)]
13. Sciberia Lungs [Electronic resource]. Available at: <https://sciberia.ru/?product=lungs>. Accessed 23.05.2024.
14. Программный модуль для анализа маммограмм ТУ 58.29.31-003-21494354-2021. М.: Платформа третье мнение, 2021. – С.40. [Programmnyj modul' dlja analiza mammogramm TU 58.29.31-003-21494354-2021. M.: Platforma tret'e mnenie. 2021: 1-40. (In Russ.)]
15. Матвиенко А.В. О требованиях к внедрению искусственного интеллекта в здравоохранение. ПО Цельс [интернет]. Доступно по: https://portal.egisz.rosminzdrav.ru/files/26.05.2023_Каспий_Матвиенко.pdf. Ссылка активна на 21.05.2024. [Matvienko A.V. O trebovaniyah k vnedreniyu iskusstvennogo intellekta v zdravooohranenie. PO Cel's. Available at: https://portal.egisz.rosminzdrav.ru/files/26.05.2023_Каспий_Матвиенко. Accessed 21.05.2024. (In Russ.)]
16. Классен В. Фтизисбиомед. Медицинский сервис фтизисбиомед для автоматизированного анализа рентгенограмм органов грудной клетки/флюорограмм (искусственный медицинский интеллект) [интернет]. 2021. Доступно по ссылке: https://files.sk.ru/navigator/company_files/1121134/1640275876_FBMTTM-RU.pdf. Ссылка активна на 24.05.2024. [Klassen V. Ftizisbiomed. Medicinskij servis ftizisbiomed dlya avtomatizirovannogo analiza rentgenogramm organov grudnoj kletki/flyuorogramm (iskusstvennyj medicinskij intellekt). Available at: https://files.sk.ru/navigator/company_files/1121134/1640275876_FBMTTM-RU.pdf. Accessed 24.05.2024. (In Russ.)]
17. Dice LR. Measures of the Amount of Ecologic Association Between Species. *Ecology*. 1945; 26(3): 297-302.
18. Erickson BJ, Kitamura F. Magician's Corner: 9. Performance Metrics for Machine Learning Models. *Radiology: Artificial Intelligence*. 2021; 3(3): e200126. doi: 10.1148/ryai.2021200126.
19. de Maissin A, Vallée R, Flamant M, et al. Multi-expert annotation of Crohn's disease images of the small bowel for automatic detection using a convolutional recurrent attention neural network. *Endoscopy International Open*. 2021; 9(7): e1136-e1144. doi: 10.1055/a-1468-3964.