

DOI: <https://doi.org/10.17816/DD321423>

Двойной просмотр результатов маммографии с применением технологий искусственного интеллекта: новая модель организации массовых профилактических исследований

Ю.А. Васильев¹, И.А. Тыров², А.В. Владзимирский¹, К.М. Арзамасов¹,
И.М. Шулькин¹, Д.Д. Кожихина¹, Л.Д. Пестренин¹

¹ Научно-практический клинический центр диагностики и телемедицинских технологий, Москва, Российская Федерация

² Департамент здравоохранения города Москвы, Москва, Российская Федерация

АННОТАЦИЯ

Обоснование. Доступность наборов медицинских данных и технологий разработки программного обеспечения на основе искусственного интеллекта в последние годы привела к увеличению количества решений для медицинской диагностики и маммографии в частности. Это программное обеспечение, зарегистрированное как медицинское изделие, может быть использовано для описания цифровых маммографий, что позволит в значительной мере сэкономить временные, материальные и кадровые ресурсы в здравоохранении при гарантированном сохранении качества профилактических исследований молочных желёз.

Цель — обосновать возможность и эффективность применения программного обеспечения на основе технологий искусственного интеллекта для первой интерпретации цифровых маммограмм при сохранении практики второго описания рентгеновских изображений врачом-рентгенологом.

Материалы и методы. Набор данных из 100 цифровых маммографических исследований, из них 50 — «Отсутствие целевой патологии», 50 — «Присутствие целевой патологии» (с признаками злокачественных новообразований), был обработан программным обеспечением на основе технологий искусственного интеллекта, зарегистрированным в Российской Федерации как медицинское изделие. Выполнен ROC-анализ. Ограничения исследования: значения метрик диагностической точности получены для версий программного обеспечения на основе технологий искусственного интеллекта, актуальных на конец 2022 года.

Результаты. При настройке на 80,0% чувствительность специфичность искусственного интеллекта составила 90,0% (95% ДИ 81,7–98,3), точность — 85,0% (95% ДИ 78,0–92,0). При настройке на 100% специфичность искусственный интеллект показал чувствительность 56,0% (95% ДИ 42,2–69,8), точность — 78,0% (95% ДИ 69,9–86,1). При настройке на 100% чувствительность специфичность искусственного интеллекта составила 54,0% (95% ДИ 40,2–67,8), точность — 77,0% (95% ДИ 68,8–85,2). Предложены два подхода, предусматривающие автономную первую интерпретацию цифровой маммографии посредством искусственного интеллекта. Первый подход заключается в оценке рентгеновского изображения с помощью искусственного интеллекта с более высокой чувствительностью, чем у двойного описания маммографии врачами-рентгенологами, при сопоставимом уровне специфичности. Второй подход подразумевает, что программное обеспечение на основе технологий искусственного интеллекта будет определять категорию маммографии («Отсутствие целевой патологии» или «Присутствие целевой патологии») с указанием степени своей «уверенности» в полученном результате в зависимости от «коридора», в который попадает предсказанное значение.

Заключение. Оба предложенных сценария использования программного обеспечения на основе технологий искусственного интеллекта с целью автономного первого описания цифровых маммограмм способны обеспечить качество диагностики, не уступающее двойному описанию снимков врачами-рентгенологами и даже превышающее его. Экономическая выгода от практической реализации данного подхода в масштабах страны может составлять от 0,6 до 5,5 млрд рублей ежегодно.

Ключевые слова: профилактические исследования; маммография; искусственный интеллект; диагностическая точность.

Как цитировать

Васильев Ю.А., Тыров И.А., Владзимирский А.В., Арзамасов К.М., Шулькин И.М., Кожихина Д.Д., Пестренин Л.Д. Двойной просмотр результатов маммографии с применением технологий искусственного интеллекта: новая модель организации массовых профилактических исследований // *Digital Diagnostics*. 2023. Т. 4, № 2. С. 93–104. DOI: <https://doi.org/10.17816/DD321423>

Рукопись получена: 17.03.2023

Рукопись одобрена: 25.04.2023

Опубликована: 25.05.2023



DOI: <https://doi.org/10.17816/DD321423>

Double-reading mammograms using artificial intelligence technologies: A new model of mass preventive examination organization

Yuriy A. Vasilev¹, Ilya A. Tyrov², Anton V. Vladzmyrskyy¹, Kirill M. Arzamasov¹, Igor M. Shulkin¹, Daria D. Kozhikhina¹, Lev D. Pestrenin¹

¹ Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

² Moscow Health Care Department, Moscow, Russian Federation

ABSTRACT

BACKGROUND: In recent years, the availability of medical datasets and technologies for software development based on artificial intelligence technology has resulted in a growth in the number of solutions for medical diagnostics, particularly mammography. Registered as a medical device, this program can interpret digital mammography, significantly saving time, material, and human resources in healthcare while ensuring the quality of mammary gland preventive studies.

AIM: This study aims to justify the possibility and effectiveness of artificial intelligence-based software for the first interpretation of digital mammograms while maintaining the practice of a radiologist's second description of X-ray images.

MATERIALS AND METHODS: A dataset of 100 digital mammography studies (50 — “absence of target pathology” and 50 — “presence of target pathology,” with signs of malignant neoplasms) was processed by software based on artificial intelligence technology that was registered as a medical device in the Russian Federation. Receiver operating characteristic analysis was performed. Limitations of the study include the values of diagnostic accuracy metrics obtained for software based on artificial intelligence technology versions, relevant at the end of 2022.

RESULTS: When set to 80.0% sensitivity, artificial intelligence specificity was 90.0% (95% CI, 81.7–98.3), and accuracy was 85.0% (95% CI, 78.0–92.0). When set to 100% specificity, artificial intelligence demonstrated 56.0% sensitivity (95% CI, 42.2–69.8) and 78.0% accuracy (95% CI, 69.9–86.1). When the sensitivity was set to 100%, the artificial intelligence specificity was 54.0% (95% CI, 40.2–67.8), and the accuracy was 77.0% (95% CI, 68.8–85.2). Two approaches have been proposed, providing an autonomous first interpretation of digital mammography using artificial intelligence. The first approach is to evaluate the X-ray image using artificial intelligence with a higher sensitivity than that of the double-reading mammogram by radiologists, with a comparable level of specificity. The second approach implies that artificial intelligence-based software will determine the mammogram category (“absence of target pathology” or “presence of target pathology”), indicating the degree of “confidence” in the obtained result, depending on the corridor into which the predicted value falls.

CONCLUSIONS: Both proposed approaches for using artificial intelligence-based software for the autonomous first interpretation of digital mammograms can provide diagnostic quality comparable to, if not superior to, double-image reading by radiologists. The economic benefit from the practical implementation of this approach nationwide can range from 0.6 to 5.5 billion rubles annually.

Keywords: artificial intelligence; diagnostic accuracy; mammography; preventive medicine.

To cite this article

Vasilev YuA, Tyrov IA, Vladzmyrskyy AV, Arzamasov KM, Shulkin IM, Kozhikhina DD, Pestrenin LD. Double-reading mammograms using artificial intelligence technologies: A new model of mass preventive examination organization. *Digital Diagnostics*. 2023;4(2):93–104. DOI: <https://doi.org/10.17816/DD321423>

Received: 17.03.2023

Accepted: 25.04.2023

Published: 25.05.2023

DOI: <https://doi.org/10.17816/DD321423>

使用人工智能对乳腺X线摄影结果进行两次查看：一种组织大规模预防性研究的新模式

Yuriy A. Vasilev¹, Ilya A. Tyrov², Anton V. Vladzimirsky¹, Kirill M. Arzamasov¹, Igor M. Shulkin¹, Daria D. Kozhikhina¹, Lev D. Pestrenin¹¹ Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation² Moscow Health Care Department, Moscow, Russian Federation

简评

论证。近年来，医疗数据集和人工智能软件技术开发的可达性，使得医疗诊断，特别是乳腺X线摄影的解决方案激增。这种登记为医疗设备的软件可被用于描述数乳腺X线摄影，这允许提供医疗服务时在很大程度上节省时间、物质和人力的资源，同时确保乳房预防性检查的质量。

该研究的目的是证明使用基于人工智能技术的软件对数字乳房X光照片进行第一次解读的可用性和有效性，同时保持放射科医生对X射线图像进行第二次描述的做法。

材料和方法。我们用在俄罗斯联邦登记为医疗设备的基于人工智能技术的软件处理了100张数字乳房X光照片的数据集，其中50张为“无目标病变”，50张为“存在目标病变”（有恶性肿瘤症状）。进行了ROC分析。研究局限性：诊断准确性度量值是基于人工智能技术的软件版本获得的，是在2022年底有效的。

结果。当设置为80.0%的灵敏度时，人工智能显示出90.0%的特异度（95% CI 81.7–98.3）和85.0%的准确性（95% CI 78.0–92.0）。当设置为100%的特异度时，人工智能显示出56.0%的灵敏度（95% CI 42.2–69.8）和78.0%的准确性（95% CI 69.9–86.1）。当设置为100%灵敏度时，人工智能的特异度为54.0%（95% CI 40.2–67.8），准确性为77.0%（95% CI 68.8–85.2）。提出了两种方法，涉及通过人工智能对数字乳腺X线摄影进行的第一次自主解读。第一种方法是利用人工智能评估X射线图像，其灵敏度高于放射科医生进行的双重乳腺X线摄影描述，特异度水平相当。第二种方法是基于人工智能技术的软件将对乳腺X线摄影进行分类（“无目标病变”或“存在目标病变”），表明其对结果的“信心”程度，取决于预测值所处的“走廊”。

结论。使用基于人工智能技术的软件对数字乳房X光照片进行第一次自主描述的两种提出方案都能提供与放射科医生对图像进行的双重描述相同甚至更高的诊断质量。在全国范围内在实践中实现这种方法的经济效益可能是每年6亿至55亿卢布。

关键词：预防性研究；乳腺X线摄影；人工智能；诊断准确性。

To cite this article

Vasilev YuA, Tyrov IA, Vladzimirsky AV, Arzamasov KM, Shulkin IM, Kozhikhina DD, Pestrenin LD. 使用人工智能对乳腺X线摄影结果进行两次查看：一种组织大规模预防性研究的新模式. *Digital Diagnostics*. 2023;4(2):93–104. DOI: <https://doi.org/10.17816/DD321423>

收到: 17.03.2023

接受: 25.04.2023

发布日期: 25.05.2023

ОБОСНОВАНИЕ

Злокачественные новообразования молочной железы представляют собой значимую медицинскую, социально-экономическую и демографическую проблему, будучи ведущей онкологической патологией и лидером в структуре смертности среди женского населения. С 2011 по 2019 год наблюдался устойчивый рост заболеваемости (с 45,24 до 53,34 на 100 тыс. населения соответственно); в 2020 году отмечалось резкое её снижение до 47,39 на 100 тыс. населения с возобновлением роста в 2021 году [1]. Такая динамика объясняется приостановкой программ массовых профилактических осмотров в период пандемии COVID-19. С одной стороны, ситуация полностью обоснованная, с другой — демонстрирует уязвимость системы здравоохранения. Вынужденное перераспределение ресурсов негативно сказалось на скрининге социально значимых заболеваний. Но и вне контекста пандемии своевременность выявления злокачественных новообразований молочных желёз остаётся неудовлетворительной. Крайне высок показатель запущенности: более 27,0% случаев злокачественных новообразований выявляется на поздних (III–IV) стадиях. Обращает на себя внимание положительная тенденция — устойчивый рост индекса накопления контингента больных злокачественными новообразованиями молочных желёз: в 2011 году соответствующий показатель составлял 9,5, в 2021 году — уже 11,9. Это свидетельствует о планомерном улучшении качества и эффективности методов лечения злокачественных новообразований молочных желёз [2].

Таким образом, требуется развитие массовых профилактических исследований как с точки зрения увеличения охвата и вовлечённости населения, производительности, качества, экономической эффективности, так и с позиций формирования устойчивости и бесперебойной доступности. Реализация указанных задач на фоне явного прогресса методов лечения позволит получить принципиально новые результаты в преодолении проблемы злокачественных новообразований молочных желёз.

В настоящее время самый распространённый способ скрининга злокачественных новообразований — рентгеновская маммография.

Действующими нормативно-правовыми актами предусматривается двойной просмотр результатов профилактической маммографии. Это означает, что изображения каждой обследованной пациентки должны быть независимо интерпретированы и описаны двумя врачами-рентгенологами. Целесообразность такого подхода неоднократно подтверждена как отечественными, так и зарубежными авторами. Объединённая частота выявления патологических изменений при двойном просмотре выше. Однократный просмотр по сравнению с двойным снижает чувствительность для всех категорий BI-RADS (Breast Imaging Reporting and Data System — система интерпретации и протоколирования визуализации молочной

железы). Более того, отказ от двойного просмотра несёт разнообразные негативные последствия для обследуемых женщин [3, 4]. Вместе с тем двойной просмотр чреват определёнными негативными аспектами (ресурсоёмкость, качество, несовершенная система финансирования).

Ресурсоёмкость. В условиях первичного звена здравоохранения необходимо задействовать двух врачей-рентгенологов для описания каждого профилактического исследования, при том что подавляющее большинство любых скрининговых исследований относится к категории «Отсутствие целевой патологии». Существует риск сокрытия кадрового дефицита за счёт приписок и иных манипуляций, что, в конечном счёте, негативно сказывается на здоровье женского населения. Вместе с тем реальная потребность в объёме профилактических маммографий и темпы наращивания парка оборудования принципиально превышают все возможности по увеличению количества кадров.

Качество. Интерпретация результатов маммографии требует специальных навыков и фактически является субспециализацией современной рентгенологии. Это только усугубляет проблему кадрового дефицита: «механическое» наращивание числа врачей-рентгенологов не решит проблему качественного и своевременного выявления злокачественных новообразований молочной железы.

Несовершенная система финансирования. Использование двойного тарифа, т.е. последовательная оплата двух услуг применима только в случаях наличия отдельного тарифа на описание результатов маммографии; соответственно, одна услуга включает выполнение и описание маммографии, вторая — только описание. В противоположном случае осуществляется некорректное финансирование (оплачиваются две услуги, включающие как выполнение, так и описание маммографии). Более распространена ситуация недофинансирования медицинских организаций: услуга выполнения и описания оплачивается однократно, второе описание не оплачивается.

Потенциально указанные проблемы могут быть решены за счёт автоматизации — применения технологий искусственного интеллекта (ТИИ) для интерпретации результатов маммографии.

В научной литературе имеются доказательные данные о высоком качестве соответствующих технических решений. Фактически ряд продуктов на основе ТИИ по своей диагностической точности в выявлении определённых рентгенологических признаков идентичен среднестатистическому врачу-рентгенологу. Так, объединённая чувствительность, специфичность и площадь под характеристической кривой составляют для автоматизированного анализа 75,4; 90,6 и 0,89%, для врача-рентгенолога — 73,0; 88,6 и 0,85% (статистически значимые различия отсутствуют) [5]. По данным метаанализа [6], объединённая чувствительность, специфичность и площадь

под характеристической кривой для автоматизированного анализа маммографии составляет 91,4; 91,6 и 94,5% соответственно. Коэффициенты сходства IoU-точности (intersection-over-union) локализации патологических находок, отсегментированных нейросетевой моделью и врачом-рентгенологом, составляют 0,86 и 0,96 соответственно [7]. Однако, на наш взгляд, в контексте скрининга злокачественных новообразований молочной железы ТИИ должны внедряться не просто в качестве дискретных систем поддержки принятия врачебных решений, но в виде автоматизированных процессов.

Цель исследования — обосновать модель организации массовых профилактических исследований методом маммографии с применением технологий искусственного интеллекта.

МАТЕРИАЛЫ И МЕТОДЫ

Исследование выполнено в рамках Эксперимента по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы (далее — Московский эксперимент), проводимого с 2020 года при поддержке Правительства Москвы (mosmed.ai). Достигнутый уровень точности программного обеспечения (ПО) на основе ТИИ (AUC, площадь под кривой >0,81), установленный на основе анализа 61 497 маммографий в течение первого года Московского эксперимента, позволил провести данное исследование [8].

Дизайн исследования

Смешанный (количественный компонент — ретроспективное диагностическое исследование, качественный компонент — аналитическое исследование).

Технологии искусственного интеллекта

Для применения в клинической практике ПО на основе ТИИ должно быть зарегистрировано как медицинское изделие. В Российской Федерации статусом медицинских изделий обладают программные продукты на основе технологий искусственного интеллекта, предназначенные для автоматизированного анализа результатов маммографии, от компаний ООО «Медицинские скрининг-системы» (РЗН 2021/14449), и ООО «Платформа Третье Мнение» (РЗН2022/16534).

В исследовании использовалось медицинское изделие, представляющее собой ПО на основе ТИИ — продукт одной из указанных выше компаний. В рамках научного Московского эксперимента ПО на основе ТИИ было интегрировано в Единый радиологический информационный сервис Единой медицинской информационно-аналитической системы г. Москвы (ЕРИС ЕМИАС). ПО на основе ТИИ принимало на вход цифровую маммографию в формате DICOM, анализировало её и возвращало в качестве ответа

текстовое описание (DICOM SR), изображения с разметкой патологических областей (DICOM SC), вероятности наличия патологии в исследовании в целом. Результаты анализа, выполненного ПО на основе ТИИ, становились доступны в ЕРИС ЕМИАС наравне с исходными результатами исследования. Для исследования использовали значения вероятности патологии (вероятность наличия злокачественного процесса) в исследовании в целом; оценка корректности выставления BI-RADS, а также точности локализации патологических находок в настоящем исследовании не проводилась.

Набор данных

В проведённое нами исследование включено 100 цифровых маммограмм, выполненных в рамках скринингового исследования, из них с патологией — 50. Средний возраст обследованных пациентов составил 63 ± 6 года.

Присвоение исследованиям категорий «Отсутствие целевой патологии» или «Присутствие целевой патологии» происходило на основании консенсуса двух экспертов (врачей-рентгенологов со стажем работы по направлению «Маммография» более 5 лет). Для группы «Присутствие целевой патологии» основным критерием включения в исследование была гистологическая верификация. Для группы «Отсутствие целевой патологии» отличающиеся друг от друга мнения врачей-рентгенологов были критерием исключения. Критериями исключения для обеих групп являлись возраст пациентов менее 18 лет и некачественные исследования (I по шкале PGMI), которые определялись на этапе проведения экспертом разметки при подготовке набора данных.

В качестве целевой патологии рассматривались патологические признаки на маммографии, соответствующие 3, 4 и 5-й категориям BI-RADS — «Присутствие целевой патологии». При классификации исследования «Отсутствие целевой патологии» подразумевалось отсутствие подозрения на злокачественные новообразования молочной железы, т.е. наличие категории BI-RADS 1 или 2.

Распределение исследований по типам в соответствии с классификацией ACR (American College of Radiology) было следующим: в группе «Отсутствие целевой патологии»: А — 26, В — 16, С — 5, D — 3; в группе «Присутствие целевой патологии»: А — 15, В — 24, С — 11, D — 0.

Набор данных сформирован из исследований, выполненных на маммографах FUJIFILM Corporation (Япония). В процессе подготовки набора данных участвовали следующие медицинские организации: Городская поликлиника (ГП) № 22, филиал № 1; Диагностический клинический центр № 1; ГП № 8; ГП № 36; ГП № 22; ГП № 209; Диагностический центр № 3, филиал № 4; Консультативно-диагностическая поликлиника № 121; Клинико-диагностический центр № 4; Городская клиническая больница имени М.П. Кончаловского Департамента здравоохранения города Москвы, Поликлиническое отделение № 3.

Этическая экспертиза

Данное исследование основано на результатах Эксперимента по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения г. Москвы, утверждённого этическим комитетом (выписка из протокола № 2 НЭК МРО РОПР от 20.02.2020), также зарегистрированного на ClinicalTrials (NCT04489992).

Статистический анализ

Для обработки полученных данных использовали метод построения и анализа характеристической кривой (ROC-анализ) посредством специально разработанного web-инструмента (<https://roc-analysis.mosmed.ai/>) [9, 10]. Истинные значения были представлены бинарной разметкой исследований (0 — отсутствие патологии, 1 — наличие целевой патологии). Результатом являлась вероятность наличия патологии, определённая ПО на основе ТИИ. Полученные данные в CSV-формате загружали в электронную форму, после чего WEB-инструмент выполнял построение характеристической кривой. В интерактивном режиме определяли пороговые значения, соответствующие самой крайней левой точке с чувствительностью 100% и самой крайней правой экспериментальной точке со специфичностью 100%. В дальнейшем проводился анализ остальных метрик диагностической точности для установленного порогового значения. Для анализа использовали термины классической четырёхпольной таблицы. Истинно положительным считалось верное отнесение маммографии к группе «Присутствие целевой патологии», истинно отрицательным — верное отнесение маммографии к группе «Отсутствие целевой патологии». Ложноположительным считалось ошибочное отнесение маммографии из группы «Отсутствие целевой патологии» к группе «Присутствие целевой патологии», а ложноотрицательным — результат отнесения маммографии из группы «Присутствие целевой патологии» к группе «Отсутствие целевой патологии».

Все статистические показатели, представленные в результатах, рассчитаны с 95% доверительным интервалом (ДИ) методом случайных выборок с возвратом (так называемый бутстреппинг, от англ. bootstrapping) с количеством итераций 1000.

РЕЗУЛЬТАТЫ

Нами предлагается следующая модель. Исследование «Профилактическая рентгеновская маммография» назначает лечащий врач в соответствии с действующими нормативно-правовыми актами, клиническими

рекомендациями. Непосредственно выполняет исследование рентгенолаборант. Результаты исследования в цифровом виде поступают в архив медицинских изображений в составе медицинской информационной системы медицинской организации и/или информационной системы в сфере здравоохранения субъекта Российской Федерации. Первый просмотр выполняется ПО (медицинским изделием на основе ТИИ, допущенным к обращению в установленном порядке). В результате первого просмотра в информационной системе формируется электронная медицинская запись¹, полученная в автоматическом режиме и содержащая (а) отдельную серию изображений с графической маркировкой и/или температурной картой областей с патологическими изменениями при их наличии; (б) структурированный отчёт, краткое руководство пользователя, заключение, детализацию и вероятность наличия патологических изменений. Второй просмотр выполняется врачом-рентгенологом. В результате второго просмотра в информационной системе формируется электронный медицинский документ — протокол описания и заключение.

Выполнено исследование возможности обеспечения ПО на основе ТИИ требуемого уровня диагностической точности.

Характеристическая кривая, отражающая работу исследуемого ПО на основе ТИИ, представлена на рис. 1.

Распределение маммографий по группам (истинно положительные, истинно отрицательные, ложноотрицательные, ложноположительные) в зависимости от выбранного порогового значения представлено в табл. 1.

При выставлении порогового значения, соответствующего 100,0% специфичности (пороговое значение 0,93), искусственный интеллект корректно определил маммографии из группы «Присутствие целевой патологии», т.е. ложноположительные срабатывания отсутствовали. Из 50 маммографий, которые были в группе «Присутствие целевой патологии», искусственный интеллект при указанной настройке порогов корректно определил 28 (56,0%) исследований. Для таких настроек ПО на основе ТИИ получаем следующие значения метрик диагностической точности: чувствительность 56,0% (95% ДИ 42,2–69,8), специфичность 100,0% (95% ДИ 100,0–100,0), точность 78,0% (95% ДИ 69,9–86,1).

При выставлении порогового значения, соответствующего 100% чувствительности (пороговое значение 0,25), ложноотрицательные результаты отсутствовали, а количество истинно отрицательных составило 27 (54,0% всех исследований группы «Отсутствие целевой патологии»). При такой настройке ПО на основе ТИИ получаем чувствительность 100,0% (95% ДИ

¹ Медицинские записи, полученные в автоматическом режиме в результате применения медицинских изделий, допущенных к обращению, не требуют подписания электронной подписью медицинского работника (в соответствии с Приказом Министерства здравоохранения Российской Федерации от 07.09.2020 N 947н «Об утверждении Порядка организации системы документооборота в сфере охраны здоровья в части ведения медицинской документации в форме электронных документов»).

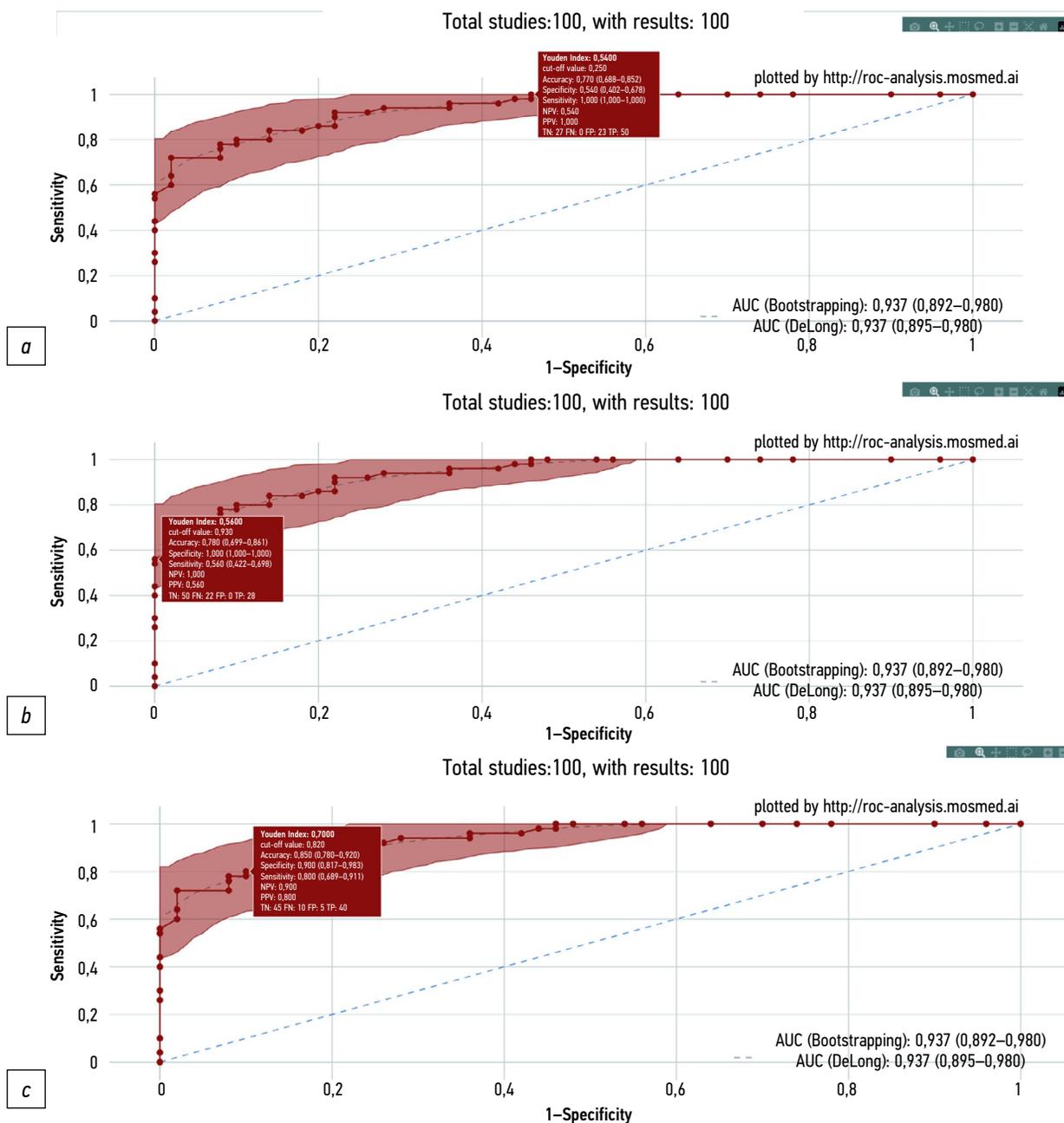


Рис. 1. Характеристическая кривая программного обеспечения на основе технологий искусственного интеллекта. Заливкой показан 95% доверительный интервал. Отдельно выделены экспериментальные значения, соответствующие 100,0% чувствительности (а), 100,0% специфичности (b) и 80,0% чувствительности (с). Для каждой из этих экспериментальных точек в прямоугольнике указаны значения метрик диагностической точности при соответствующем пороговом значении.

Таблица 1. Четырёхполюсная таблица для разных пороговых значений

Группа	Число исследований	Результат	100% чувствительность	100% специфичность	Сбалансированная чувствительность и специфичность
«Присутствие целевой патологии»	50	Истинно положительный	50	28	80,0%**
		Ложноположительный	23	0	
«Отсутствие целевой патологии»	50	Истинно отрицательный	27	50	90,0%*
		Ложноотрицательный	0	22	

Примечание. * Процент истинно отрицательных результатов рассчитан относительно объёма группы «Отсутствие целевой патологии»; ** процент истинно положительных результатов рассчитан относительно объёма группы «Присутствие целевой патологии».

100,0–100,0), специфичность 54,0% (95% ДИ 40,2–67,8), точность 77,0% (95% ДИ 68,8–85,2).

При выставлении порогового значения, руководствуясь методом максимизации индекса Юдена (пороговое значение 0,82), значение чувствительности составило 80,0%, количество истинно отрицательных результатов — 45 (90,0% всех исследований группы «Отсутствие целевой патологии»), а количество истинно положительных — 40 (80,0% всех исследований группы «Присутствие целевой патологии»). При такой настройке ПО на основе ТИИ получаем чувствительность 80,0% (95% ДИ 68,9–91,1), специфичность 90,0% (95% ДИ 81,7–98,3), точность 85,0% (95% ДИ 78,0–92,0).

ОБСУЖДЕНИЕ

Резюме основного результата исследования

Для ПО на основе ТИИ, используемого при реализации услуги «Описание и интерпретация данных маммографического исследования с использованием искусственного интеллекта», выставление разных пороговых значений позволяет установить значения метрик чувствительности, специфичности и точности, соответствующие или превышающие таковые для двойного чтения маммографии врачами-рентгенологами.

Обсуждение основного результата исследования

В Российской Федерации ежегодно выполняется около 8,2 млн профилактических маммографий², двойная оценка которых требует значительного количества временных, кадровых и финансовых ресурсов. Использование ПО на основе ТИИ для первого просмотра маммографий позволит уменьшить вышеуказанные затраты с одновременным сохранением или даже улучшением качества диагностики. Предлагается рассмотреть два подхода к настройке ПО на основе ТИИ для первого просмотра маммографий.

Первый подход предполагает использование ПО на основе ТИИ со сбалансированными значениями чувствительности и специфичности. В нашем случае чувствительность ПО на основе ТИИ составила 80,0%, что превышает чувствительность двойного просмотра маммографий врачами-рентгенологами, определённую в обзорных публикациях (72,0–73,0%) [5, 6]. При этом специфичность работы искусственного интеллекта (90,0%), полученная в ходе нашего исследования, не уступает специфичности работы двух врачей-рентгенологов (88,0–98,0%) [5, 6]. Ожидается также, что ПО на основе ТИИ в сочетании с оценкой одного рентгенолога будет обладать более

высокой общей точностью интерпретации маммографии по сравнению с интерпретацией только одним рентгенологом, что подтверждается рядом научных публикаций [11–13]. Результатом работы ПО на основе ТИИ при данном подходе будет электронная медицинская запись, содержащая заключение о категории изображения («Отсутствие целевой патологии») или «Присутствие целевой патологии»).

Второй подход подразумевает, что ПО на основе ТИИ будет определять категорию изображения («Отсутствие целевой патологии») или «Присутствие целевой патологии») с указанием степени своей «уверенности» в полученном результате. Общая концепция метода представлена на рис. 2.

Как упоминалось в разделе «Результаты», при построении ROC-кривой для ПО на основе ТИИ были определены пороговые значения (cut-off value) для предсказанных значений при 100% чувствительности и 100% специфичности (0,25 и 0,93 соответственно). На основе этих данных предлагается относить предсказанное значение к одному из трёх «коридоров», которые соответствуют разным результатам классификации и разной степени «уверенности» ПО на основе ТИИ в них:

- 1) «зелёный коридор»: предсказанные значения находятся в диапазоне от 0 до 0,25 и соответствуют категории «Отсутствие целевой патологии» с уверенностью 100%;
- 2) «красный коридор»: предсказанные значения находятся в диапазоне от 0,93 до 1,0 и соответствуют категории «Присутствие целевой патологии» с уверенностью 100%;
- 3) «жёлтый коридор»: предсказанные значения находятся в диапазоне от 0,25 до 0,93 включительно и соответствуют категории «Отсутствие целевой патологии» или «Присутствие целевой патологии», но вероятность правильности классификации менее 100%.

Предсказанное значение и цвет «коридора», в который оно попадает, предлагается добавить в описание маммографии, выполненное ПО на основе ТИИ. При наличии этой информации врач-рентгенолог, который описывает изображение с результатами работы ПО на основе ТИИ, будет знать, насколько можно доверять результатам его работы. С одной стороны, это поможет врачу сохранять повышенную бдительность при описании маммографии из «жёлтого коридора», с другой — подобный подход в долгосрочной перспективе может повысить доверие врачей-рентгенологов к результатам работы ПО на основе ТИИ. Это связано с тем, что искусственный интеллект даже при текущем высоком уровне точности пока не способен однозначно правильно классифицировать 100% анализируемых изображений.

² Тюрин И.Е. Отчет за 2020 г. главного внештатного специалиста Минздрава России по лучевой и инструментальной диагностике Тюрина И.Е. [электронный ресурс]. Режим доступа: https://static-0.minzdrav.gov.ru/system/attachments/attaches/000/056/620/original/Отчет_за_2020_год_Тюрин.pdf?1624967722.

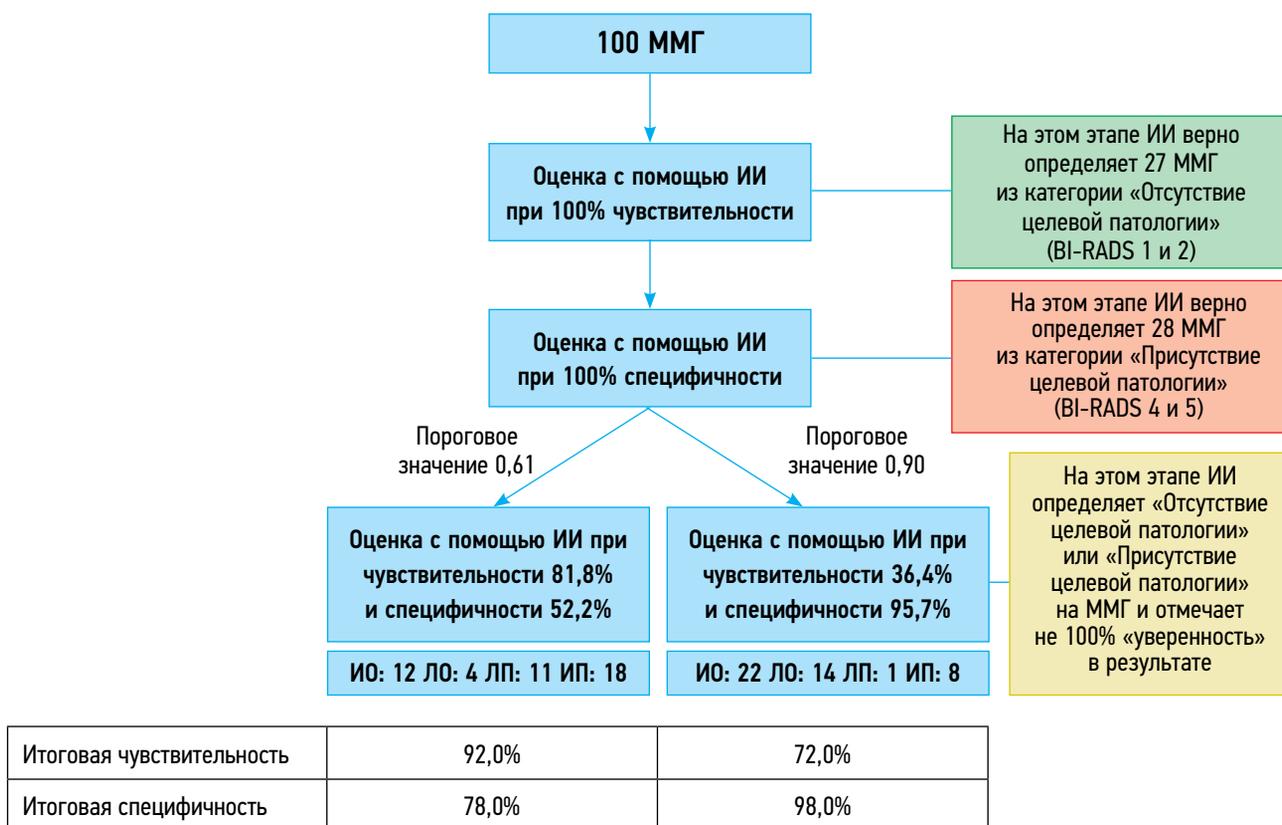


Рис. 2. Концепция подхода к первому описанию маммографии с помощью искусственного интеллекта, в основе которой лежит бинарная классификация изображений с указанием степени «уверенности» программного обеспечения на основе технологий искусственного интеллекта в полученных результатах.

Примечание. ММГ — маммография; ИИ — искусственный интеллект; ИО — истинно отрицательный; ЛО — ложноотрицательный; ЛП — ложноположительный; ИП — истинно положительный.

Преимуществом второго подхода является не только способность искусственного интеллекта отнести часть маммографий к категории «Отсутствие целевой патологии» или «Присутствие целевой патологии» со 100% степенью «уверенности» (зелёный и красный «коридоры» для предсказанных значений), но и возможность изменения порогового значения с целью балансировки чувствительности и специфичности ПО на основе ТИИ для анализа маммографий, попадающих в «жёлтый коридор». В зависимости от клинической задачи можно

установить более высокую чувствительность, которая обеспечит лучшую выявляемость патологии при наименьшем количестве ложноотрицательных результатов или более высокую специфичность для уменьшения количества ложноположительных результатов (табл. 2) [5, 6].

Результаты проведённого исследования демонстрируют возможность автономного применения ПО на основе ТИИ для первого прочтения маммографии, тем не менее в дальнейшем необходимо оптимизировать программное обеспечение для более эффективного выявления «Отсутствия

Таблица 2. Чувствительность и специфичность разных подходов к анализу маммограмм

Анализ результатов профилактической маммографии	Чувствительность, %	Специфичность, %
Двойной просмотр врачами*	72,0–73,0	88,0–98,0
Первый подход с использованием искусственного интеллекта для первого прочтения маммографии (бинарная классификация)	80,0	90,0
Второй подход с использованием искусственного интеллекта для первого прочтения маммографии (бинарная классификация со степенью «уверенности») при пороговом значении 0,61	92,0	78,0
Второй подход с использованием искусственного интеллекта для первого прочтения маммографии (бинарная классификация со степенью «уверенности») при пороговом значении 0,90	72,0	98,0

Примечание. * По литературным данным [5, 6].

целевой патологии» и «Присутствия целевой патологии». Необходимо также отметить, что для настоящего исследования изначально отбирались только качественно выполненные маммографии. Однако современное ПО на основе ТИИ имеет функцию автономного контроля качества маммографии. При практическом внедрении ПО на основе ТИИ может осуществлять кроме клинической оценки результатов маммографии также техническую оценку качества [8].

Экономическое обоснование двойного просмотра результатов маммографии с применением технологий искусственного интеллекта

В рамках исследования проанализированы тарифные соглашения на оплату медицинской помощи, оказываемой по территориальной программе обязательного медицинского страхования, принятые в регионах Российской Федерации на 2023 год.

В 19 субъектах Российской Федерации (22,4% от количества всех субъектов Российской Федерации) отдельно выделена стоимость медицинской услуги «А06.20.004 Маммография», которая проводится в рамках диспансеризации. В 4 субъектах из 19 также определена стоимость медицинской услуги «А06.30.002 Описание и интерпретация рентгенографических изображений (при повторном описании маммографии)». Во всех остальных субъектах Российской Федерации первый этап диспансеризации взрослого населения учитывается комплексной услугой без уточнения стоимости входящей в него маммографии.

Стоимость медицинской услуги по описанию маммографии варьирует от 114,97 до 1034,93 рублей. По состоянию на 01 марта 2023 года медицинская услуга «Описание и интерпретация данных маммографического исследования с использованием искусственного интеллекта» применяется только в рамках Московского эксперимента [8]. Согласно тарифному соглашению на оплату медицинской помощи, оказываемой по территориальной программе обязательного медицинского страхования г. Москвы, стоимость данной медицинской услуги составляет 239,00 рублей³.

В рамках настоящего исследования проанализированы два подхода к определению необходимого размера финансирования скрининговых маммографий в пределах территории Российской Федерации. Первый подход заключался в выполнении расчётов на основе стоимости медицинских услуг в Москве на 2023 год. Стоимость одного описания маммографии врачом-рентгенологом составляет 178,00 рублей. Следовательно, двойное описание каждой маммографии врачами-рентгенологами будет стоить 356,00 рублей. В свою очередь стоимость описания маммографии искусственным интеллектом и врачом-рентгенологом, как было упомянуто выше, составляет

239,00 рублей. Таким образом, при среднем количестве ежегодных маммографий в России 8,2 млн стоимость двойного их описания врачами-рентгенологами составит 2,9 млрд рублей, а стоимость описания искусственным интеллектом и врачом-рентгенологом — 1,9 млрд рублей. Потенциальная экономия средств за счёт использования ПО на основе ТИИ может составить 1,0 млрд рублей ежегодно. Второй подход заключался в выполнении расчётов на основе стоимости медицинских услуг в субъектах Российской Федерации на 2023 год. Учитывался также процент экономии средств при интерпретации маммографии искусственным интеллектом и врачом-рентгенологом по сравнению с двойной интерпретацией врачами-рентгенологами в Москве, который составил

$$\frac{178,00 \times 2 - 239,00}{178,00 \times 2} \times 100\% = 32,8\%.$$

При стоимости медицинской услуги по описанию маммографии врачом-рентгенологом в субъектах Российской Федерации от 114,97 до 1034,93 рублей двойные описания маммографии врачами-рентгенологами стоят от 229,94 до 2069,86 рублей. Предполагая, что интерпретация маммографии искусственным интеллектом и врачом-рентгенологом в субъектах Российской Федерации будет так же, как и в Москве, дешевле двойного описания врачами-рентгенологами на 32,8%, получаем стоимость описания маммографии искусственным интеллектом и врачом-рентгенологом в пределах от 154,51 до 1390,94 рублей. Таким образом, при среднем количестве ежегодных маммографий в России 8,2 млн стоимость двойного описания врачами-рентгенологами составит от 1,8 до 16,9 млрд рублей, а стоимость описания искусственным интеллектом и врачом-рентгенологом — от 1,2 до 11,4 млрд рублей. Потенциальная экономия средств за счёт использования ПО на основе ТИИ на общегосударственном уровне может составить от 0,6 до 5,5 млрд рублей ежегодно.

Ограничения исследования

Значения метрик диагностической точности получены для версий ПО на основе ТИИ, актуальных на конец 2022 года.

Для пациентов группы «Отсутствие целевой патологии» оценка динамики класса по шкале BI-RADS не проводилась, что можно расценивать как ограничение исследования.

ЗАКЛЮЧЕНИЕ

Показана возможность и доказана перспективность применения ТИИ для первого описания маммографических изображений. Медицинское изделие — программное обеспечение на основе ТИИ — обладает чувствительностью

³ Тарифное соглашение на оплату медицинской помощи, оказываемой по Территориальной программе обязательного медицинского страхования города Москвы на 2023 год (Заключено в г. Москве 30.12.2022). Режим доступа: <https://www.consultant.ru/law/review/208575247.html>.

и специфичностью, равными чувствительности и специфичности работы двух врачей-рентгенологов или превосходящими их. Модель применения ПО на основе ТИИ в качестве первого чтения в сочетании со вторым прочтением врачом-рентгенологом позволяет в масштабе всей страны получить экономическую выгоду от реализации в объёме от 0,6 до 5,5 млрд рублей ежегодно.

ДОПОЛНИТЕЛЬНО

Источник финансирования. Исследование выполнено в рамках научно-исследовательской работы «Научные методологии устойчивого развития технологий искусственного интеллекта в медицинской диагностике», регистрационный номер ЕГИСУ: 123031500004-5.

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Вклад авторов. Все авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, проведение исследования и подготовку статьи, прочли и одобрили финальную версию перед публикацией). Наибольший вклад распределён следующим образом: Ю.А. Васильев, И.А. Тыров — концепция исследования, финальная вычитка текста; А.В. Владзимирский — концепция и дизайн исследования, написание и редактирование текста; К.М. Арзамасов — концепция

и дизайн исследования, анализ полученных данных, написание и редактирование текста; И.М. Шулькин — дизайн исследования, сбор и обработка материалов, финальная вычитка текста; Д.Д. Кожихина — дизайн исследования, сбор и обработка материалов; Л.Д. Пестренин — анализ полученных данных, написание текста.

ADDITIONAL INFORMATION

Funding source. This article was prepared by a group of authors as a part of the research and development effort titled “Evidence-based methodologies for sustainable development of artificial intelligence in medical imaging”, registration number to EGISU: 123031500004-5.

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. Vasilev, I.A. Tyrov — research concept, final proofreading of the text; A.V. Vladzimirsky — research concept and design, writing and editing of the text; K.M. Arzamasov — research concept and design, analysis of the data obtained, writing and editing of the text; I.M. Shulkin — research design, collection and processing of materials, final proofreading the text; D.D. Kozhikhina — research design, collection and processing of materials; L.D. Pestrenin — analysis of the data obtained, writing the text.

СПИСОК ЛИТЕРАТУРЫ

1. Злокачественные новообразования в России в 2021 году (заболеваемость и смертность) / под ред. А.Д. Каприна, В.В. Старинского, А.О. Шахзадовой. Москва, 2022. 252 с.
2. Состояние онкологической помощи населению России в 2021 году / под ред. А.Д. Каприна, В.В. Старинского, А.О. Шахзадовой. Москва, 2022. 239 с.
3. Chen Y., James J.J., Michalopoulou E., et al. Performance of radiologists and radiographers in double reading mammograms: The UK national health service breast screening program // *Radiology*. 2023. Vol. 306, N 1. P. 102–109. doi: 10.1148/radiol.212951
4. Euler-Chelpin M.V., Lillholm M., Napolitano G., et al. Screening mammography: Benefit of double reading by breast density // *Breast Cancer Res Treat*. 2018. Vol. 171, N 3. P. 767–776. doi: 10.1007/s10549-018-4864-1
5. Hickman S.E., Woitek R., Le E.P., et al. Machine learning for workflow applications in screening mammography: Systematic review and meta-analysis // *Radiology*. 2022. Vol. 302, N 1. P. 88–104. doi: 10.1148/radiol.2021210391
6. Liu J., Lei J., Ou Y., et al. Mammography diagnosis of breast cancer screening through machine learning: A systematic review and meta-analysis // *Clin Exp Med*. 2022. doi: 10.1007/s10238-022-00895-0
7. Рожкова Н.И., Ройтберг П.Г., Варфоломеева А.А., и др. Сегментационная модель скрининга рака молочной железы на основе нейросетевого анализа рентгеновских изображений // *Сеченовский вестник*. 2020. Т. 11, № 3. С. 4–14. doi: 10.47093/2218-7332.2020.11.3.4-14
8. Компьютерное зрение в лучевой диагностике: первый этап Московского эксперимента. Монография / под ред. Ю.А. Васильева, А.В. Владзимирского. Москва: Издательские решения, 2022. 388 с.
9. Патент РФ на изобретение № 2022617324/05.04.2022. Бюл. № 4. Морозов С.П., Андрейченко А.Е., Четвериков С.Ф., и др. Веб-инструмент для выполнения ROC-анализа результатов диагностических тестов. Режим доступа: <https://www.elibrary.ru/item.asp?id=48373757>. Дата обращения: 10.03.2023.
10. Morozov S.P., Vladzimirsky A.V., Klyashtorny V.G., et al. Clinical acceptance of software based on artificial intelligence technologies (radiology). Moscow, 2019. 45 p. (Сер. Best practices in medical imaging).
11. Schaffter T., Buist D.S., Lee C.I., et al. Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms // *JAMA Netw Open*. 2020. Vol. 3, N 3. P. e200265. doi: 10.1001/jamanetworkopen.2020.0265
12. Wan Y., Tong Y., Liu Y., et al. Evaluation of the combination of artificial intelligence and radiologist assessments to interpret malignant architectural distortion on mammography // *Front Oncol*. 2022. N 12. P. 880150. doi: 10.3389/fonc.2022.880150
13. Leibig C., Brehmer M., Bunk S., et al. Combining the strengths of radiologists and AI for breast cancer screening: A retrospective analysis // *Lancet Digit Health*. 2022. Vol. 4, N 7. P. e507–e519. doi: 10.1016/S2589-7500(22)00070-X

REFERENCES

1. Malignant neoplasms in Russia in 2021 (morbidity and mortality). Ed by A.D. Kaprin, V.V. Starinsky, A.O. Shahzadova. Moscow; 2022. 252 p. (In Russ).
2. The state of oncological assistance to the population of Russia in 2021. Ed by A.D. Kaprin, V.V. Starinsky, A.O. Shahzadova. Moscow; 2022. 239 p. (In Russ).
3. Chen Y, James JJ, Michalopoulou E, et al. Performance of radiologists and radiographers in double reading mammograms: The UK national health service breast screening program. *Radiology*. 2023;306(1):102–109. doi: 10.1148/radiol.212951
4. Euler-Chelpin MV, Lillholm M, Napolitano G, et al. Screening mammography: Benefit of double reading by breast density. *Breast Cancer Res Treat*. 2018;171(3):767–776. doi: 10.1007/s10549-018-4864-1
5. Hickman SE, Woitek R, Le EP, et al. Machine learning for workflow applications in screening mammography: Systematic review and meta-analysis. *Radiology*. 2022;302(1):88–104. doi: 10.1148/radiol.2021210391
6. Liu J, Lei J, Ou Y, et al. Mammography diagnosis of breast cancer screening through machine learning: A systematic review and meta-analysis. *Clin Exp Med*. 2022. doi: 10.1007/s10238-022-00895-0
7. Rozhkova NI, Rojtbjerg PG, Varfolomeeva AA, et al. Neural network-based segmentation model for breast cancer X-ray screening. *Sechenov medical journal*. 2020;11(3):4–14 (In Russ). doi: 10.47093/2218-7332.2020.11.3.4-14
8. Vasilev JA, Vladzimirskyy AV. Computer vision in radiology: The first stage of the Moscow experiment: Monograph. Moscow: Izdatel'skie resheniya; 2022. 388 p. (In Russ).
9. Patent RUS № 2022617324/05.04.2022. Byul. № 4. Morozov SP, Andreichenko AE, Chetverikov SF, et al. A web-based tool for performing ROC analysis of diagnostic test results. Available from: <https://www.elibrary.ru/item.asp?id=48373757>. Accessed: 10.03.2023. (In Russ).
10. Morozov SP, Vladzimirskyy AV, Klyashtornyy VG, et al. Clinical acceptance of software based on artificial intelligence technologies (radiology). Moscow; 2019. 45 p. (Ser. Best practices in medical imaging).
11. Schaffter T, Buist DS, Lee CI, et al. Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms. *JAMA Netw Open*. 2020;3(3):e200265. doi: 10.1001/jamanetworkopen.2020.0265
12. Wan Y, Tong Y, Liu Y, et al. Evaluation of the combination of artificial intelligence and radiologist assessments to interpret malignant architectural distortion on mammography. *Front Oncol*. 2022;(12):880150. doi: 10.3389/fonc.2022.880150
13. Leibig C, Brehmer M, Bunk S, et al. Combining the strengths of radiologists and AI for breast cancer screening: A retrospective analysis. *Lancet Digit Health*. 2022;4(7):e507–e519. doi: 10.1016/S2589-7500(22)00070-X

ОБ АВТОРАХ

* **Пестренин Лев Дмитриевич**, мл. науч. сотр.;
адрес: Россия, 127051, Москва, ул. Петровка, д. 24, стр. 1;
ORCID: <http://orcid.org/0000-0002-1786-4329>;
eLibrary SPIN: 7193-7706; e-mail: PestreninLD@zdrav.mos.ru

Васильев Юрий Александрович, канд. мед. наук;
ORCID: <http://orcid.org/0000-0002-0208-5218>;
eLibrary SPIN: 4458-5608; e-mail: npcmmr@zdrav.mos.ru

Тыров Илья Александрович;
ORCID: <http://orcid.org/0000-0001-9337-624X>;
eLibrary SPIN: 8625-3458; e-mail: npcmmr@zdrav.mos.ru

Владзими́рский Антон Вячеславович, д-р мед. наук,
профессор; ORCID: <http://orcid.org/0000-0002-2990-7736>;
eLibrary SPIN: 3602-7120; e-mail: npcmmr@zdrav.mos.ru

Арзамасов Кирилл Михайлович, канд. мед. наук;
ORCID: <http://orcid.org/0000-0001-7786-0349>;
eLibrary SPIN: 3160-8062; e-mail: ArzamasovKM@zdrav.mos.ru

Шулькин Игорь Михайлович;
ORCID: <http://orcid.org/0000-0002-7613-5273>;
eLibrary SPIN: 5266-0618; e-mail: npcmmr@zdrav.mos.ru

Кожихина Дарья Дмитриевна;
ORCID: <http://orcid.org/0000-0001-7690-8427>;
eLibrary SPIN: 5869-3854; e-mail: npcmmr@zdrav.mos.ru

AUTHORS' INFO

* **Lev D. Pestrenin**, Junior Research Associate;
address: 24/1 Petrovka street, 127051 Moscow, Russia;
ORCID: <http://orcid.org/0000-0002-1786-4329>;
eLibrary SPIN: 7193-7706; e-mail: PestreninLD@zdrav.mos.ru

Yuriy A. Vasilev, MD, Cand. Sci. (Med);
ORCID: <http://orcid.org/0000-0002-0208-5218>;
eLibrary SPIN: 4458-5608; e-mail: npcmmr@zdrav.mos.ru

Ilya A. Tyrov;
ORCID: <http://orcid.org/0000-0001-9337-624X>;
eLibrary SPIN: 8625-3458; e-mail: npcmmr@zdrav.mos.ru

Anton V. Vladzimirskyy, MD, Dr. Sci. (Med), Professor;
ORCID: <http://orcid.org/0000-0002-2990-7736>;
eLibrary SPIN: 3602-7120; e-mail: npcmmr@zdrav.mos.ru

Kirill M. Arzamasov, MD, Cand. Sci. (Med);
ORCID: <http://orcid.org/0000-0001-7786-0349>;
eLibrary SPIN: 3160-8062; e-mail: ArzamasovKM@zdrav.mos.ru

Igor M. Shulkin;
ORCID: <http://orcid.org/0000-0002-7613-5273>;
eLibrary SPIN: 5266-0618; e-mail: npcmmr@zdrav.mos.ru

Daria D. Kozhikhina;
ORCID: <http://orcid.org/0000-0001-7690-8427>;
eLibrary SPIN: 5869-3854; e-mail: npcmmr@zdrav.mos.ru

* Автор, ответственный за переписку / Corresponding author