

ДИАГНОСТИЧЕСКАЯ ТОЧНОСТЬ И СОГЛАСОВАННОСТЬ ИИ-СЕРВИСОВ МЕЖДУ СОБОЙ И С ВРАЧОМ-ЭКСПЕРТОМ ПРИ ОЦЕНКЕ МАММОГРАФИЧЕСКИХ ИССЛЕДОВАНИЙ ПО ШКАЛЕ BI-RADS: РЕТРОСПЕКТИВНОЕ АНАЛИТИЧЕСКОЕ ИССЛЕДОВАНИЕ

^{1,2}А. С. Азарян[✉], ¹М. Ю. Хрустачева[✉], ^{1,3}Л. Д. Пестренин[✉], ¹Ю. А. Васильев[✉], ^{1,4}А. В. Владзимирский[✉],
^{1,3}О. В. Омелянская[✉], ¹А. С. Доможирова[✉], ^{1,3,5}К. М. Арзамасов[✉], ¹А. С. Гацук[✉]

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

²Центральная государственная медицинская академия, Москва, Россия

³МИРЭА — Российский технологический университет, Москва, Россия

⁴Первый Московский государственный медицинский университет имени И. М. Сеченова, Москва, Россия

⁵Самарский государственный медицинский университет, Самара, Россия

ВВЕДЕНИЕ: По данным литературы, диагностическая точность маммографических ИИ-сервисов при определении категории BI-RADS может колебаться в достаточно широком диапазоне. При этом отсутствуют исследования, посвященные оценке согласованности между ИИ-сервисами и врачом-рентгенологом при определении категории BI-RADS.

ЦЕЛЬ: Оценить диагностическую точность и согласованность трех ИИ-сервисов между собой и с врачом-экспертом при оценке маммографических исследований по шкале BI-RADS.

МАТЕРИАЛЫ И МЕТОДЫ: Выполнено смешанное исследование, включавшее ретроспективный диагностический и аналитический компоненты. В анализ вошли 99 анонимизированных маммографических исследований. Исследования оценивались врачом-экспертом с определением категории BI-RADS 1–5 отдельно для правой и левой молочной железы; полученные оценки использовались в качестве референсного стандарта. Набор данных был обработан тремя коммерческими ИИ-сервисами, также определявшими категории BI-RADS 1–5.

Статистика: Статистический анализ проводился на уровне отдельных молочных желёз ($n=198$). Оценивали ROC AUC, чувствительность, специфичность и точность с 95% доверительными интервалами для двух бинарных шкал BI-RADS: BI-RADS 1–3 против 4–5 и BI-RADS 1–2 против 3–5. Согласованность между ИИ-сервисами и врачом-экспертом оценивали с помощью внутриклассовой корреляции и каппы Коэна. Для сравнения диагностических показателей применяли тест Мак-Немара, для оценки различий согласованности — bootstrap-анализ с 5000 повторений.

РЕЗУЛЬТАТЫ: Оценки диагностической точности представлены в виде диапазонов значений, полученных для отдельных ИИ-сервисов, оценки согласованности включают значения, полученные для отдельных пар ИИ-сервисов или пар «ИИ-сервис + врач». Диагностическая точность ИИ-сервисов по бинарной шкале BI-RADS № 1: ROC AUC — 0,644–0,876, точность — 0,788–0,909, специфичность — 0,816–0,977, чувствительность — 0,333–0,883. По бинарной шкале BI-RADS № 2: ROC AUC — 0,795–0,915, точность — 0,808–0,939, специфичность — 0,818–0,961, чувствительность — 0,739–0,870. Согласованность между парами ИИ-сервисов по бинарной шкале BI-RADS № 1 находилась в диапазоне от 0,232 до 0,504, по бинарной шкале BI-RADS № 2 — от 0,416 до 0,663, по полной шкале BI-RADS — от 0,387 до 0,564. Согласованность между ИИ-сервисами и врачом-экспертом по бинарной шкале BI-RADS № 1 находилась в диапазоне от 0,288 до 0,641, по бинарной шкале BI-RADS № 2 — от 0,518 до 0,832, по полной шкале BI-RADS — от 0,570 до 0,701.

ОБСУЖДЕНИЕ: Диагностическая точность ИИ-сервисов по бинарной шкале BI-RADS № 1 и бинарной шкале BI-RADS № 2 соответствует диапазонам в аналогичных исследованиях. По сравнению с результатами нашего предыдущего исследования согласованность по бинарной шкале BI-RADS № 1, бинарной шкале BI-RADS № 2 и полной шкале оказалась в подавляющем большинстве случаев ниже, чем согласованность между врачами-рентгенологами.

ЗАКЛЮЧЕНИЕ: Полученные результаты подтверждают данные литературы, свидетельствующие о широких диапазонах колебаний диагностической точности ИИ-сервисов. С учетом различных шкал BI-RADS согласованность ИИ-сервисов и врача-эксперта оказалась в подавляющем большинстве случаев ниже, чем согласованность между врачами-рентгенологами, которую мы оценивали в предыдущем исследовании.

КЛЮЧЕВЫЕ СЛОВА: маммография; искусственный интеллект; BI-RADS; точность; согласованность; рак молочной железы

* Для корреспонденции: Азарян Авет Сергеевич, e-mail: Dr.Azaryan@yandex.ru

Для цитирования: Азарян А.С., Хрустачева М.Ю., Пестренин Л.Д., Васильев Ю.А., Владзимирский А.В., Омелянская О.В., Доможирова А.С., Арзамасов К.М., Гацук А.С. Диагностическая точность и согласованность ИИ-сервисов между собой и с врачом-экспертом

при оценке маммографических исследований по шкале BI-RADS: ретроспективное аналитическое исследование // Лучевая диагностика и терапия. 2026. Т. 17, № 2. С. 86–97, doi: <http://doi.org/10.22328/2079-5343-2026-17-2-86-97>.

DIAGNOSTIC ACCURACY AND AGREEMENT OF AI SERVICES AMONG THEMSELVES AND WITH MEDICAL EXPERT IN EVALUATING MAMMOGRAPHY USING THE BI-RADS SCALE: A RETROSPECTIVE ANALYTICAL STUDY

^{1,2}Avet S. Azaryan[✉], ¹Margarita Yu. Khrustacheva[✉], ^{1,3}Lev D. Pestrenin[✉], ¹Yuriy A. Vasilev[✉],
^{1,4}Anton V. Vladzimirsky[✉], ¹Olga V. Omelyanskaya[✉], ¹Alla S. Domozhireva[✉], ^{1,3,5}Kirill M. Arzamasov[✉],
¹Anderei S. Gatsuk[✉]

¹Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Moscow, Russia

²Central State Medical Academy of Department of Presidential Affairs, Moscow, Russia

³MIREA — Russian Technological University, Moscow, Russia

⁴I. M. Sechenov First Moscow State Medical University, Moscow, Russian

⁵Samara State Medical University, Samara, Russia

INTRODUCTION: According to the literature, the diagnostic accuracy of mammographic AI services in determining the BI-RADS category can vary quite widely. However, there are no studies that evaluate agreement between AI systems and radiologists in determining the BI-RADS category.

OBJECTIVE: Evaluation of the diagnostic accuracy and agreement of three AI systems with each other and with an expert reader in mammographic examinations through BI-RADS.

MATERIALS AND METHODS: A mixed study was performed, which included retrospective diagnostic and analytical components. The analysis included 99 anonymized mammographic studies. The studies were evaluated by an expert reader with the definition of BI-RADS 1–5 categories separately for the right and left breast; the scores obtained were used as a reference standard. The data set was processed by three commercial AI services, which also identified BI-RADS 1–5 categories.

Statistics: Statistical analysis was performed at the level of individual breasts (n=198). ROC AUC, sensitivity, specificity, and accuracy were evaluated with 95% confidence intervals for two binary BI-RADS scales: BI-RADS 1–3 vs. 4–5 and BI-RADS 1–2 vs. 3–5. Agreement was calculated using the Pearson and Cohen intraclass correlation method for similar scales and the full BI-RADS scale. The McNemar test was used to compare diagnostic parameters, and a bootstrap analysis with 5,000 repetitions was used to assess consistency differences.

RESULTS: Diagnostic accuracy estimates are presented as ranges of values obtained for individual AI services, and consistency estimates include values obtained for individual pairs of AI services or pairs of «AI service + reader». Diagnostic accuracy of AI services on the BI-RADS binary scale No. 1: ROC AUC — 0.644–0.876, accuracy — 0.788–0.909, specificity — 0.816–0.977, sensitivity — 0.333–0.883. According to the BI-RADS No. 2 binary scale: ROC AUC — 0.795–0.915, accuracy — 0.808–0.939, specificity — 0.818–0.961, sensitivity — 0.739–0.870. Agreement between pairs of AI services on the BI-RADS binary scale No. 1 ranged from 0.232 to 0.504, on the BI-RADS binary scale No. 2 — from 0.416 to 0.663, on the full BI-RADS scale — from 0.387 to 0.564. Agreement between AI services and an expert reader on the binary scale BI-RADS No. 1 ranged from 0.288 to 0.641, on the binary scale BI-RADS No. 2 — from 0.518 to 0.832, on the full BI-RADS scale — from 0.570 to 0.701.

DISCUSSION: The diagnostic accuracy of AI services on the BI-RADS No. 1 binary scale and BI-RADS No. 2 binary scale corresponded to the ranges in similar studies. Compared to the results of our previous study, the agreement on the BI-RADS No. 1 binary scale, BI-RADS No. 2 binary scale, and the full scale was in the vast majority of cases was lower than the agreement between radiologists.

CONCLUSION: In most cases, there were no statistically significant differences between the diagnostic accuracy parameters of the AI services calculated for the BI-RADS binary scale No. 1 and No. 2. There were also no statistically significant differences in the agreement between pairs of AI services and between AI services and an expert doctor, depending on the type of BI-RADS scale. Considering the various BI-RADS scales, the agreement between AI services and an expert reader was in the vast majority of cases lower than the agreement between radiologists that we evaluated in a previous study.

KEYWORDS: mammography; artificial intelligence; BI-RADS; accuracy; agreement; breast cancer

* For correspondence: Avet S. Azaryan, e-mail: Dr.Azaryan@yandex.ru

For citation: Azaryan A.S., Khrustacheva M.Yu., Pestrenin L.D., Vasilev Yu.A., Vladzimirsky A.V., Omelyanskaya O.V., Domozhireva A.S., Arzamasov K.M., Gatsuk A.S. Diagnostic accuracy and agreement of AI services among themselves and with medical expert in evaluating mammography using the BI-RADS scale: a retrospective analytical study // *Diagnostic radiology and radiotherapy*. 2026. Vol. 17, No. 2. P. 86–97, <http://doi.org/10.22328/2079-5343-2026-17-2-86-97>.

Введение. Доля впервые выявленных случаев рака молочной железы увеличивается с каждым годом, что делает злокачественные новообразования (ЗНО) молочной железы самым распространенным типом злокачественных новообразований среди женщин [1]. При этом в России наблюдается рост активного выявления ЗНО молочных желез с 37,2% до 44% [2]. Рак молочной железы остается ведущей онкологической патологией у женского населения Российской Федерации, его вклад в структуру смертности женщин от злокачественных новообразований составляет около 15,9% [1, 3].

Маммографический скрининг рака молочной железы представляет собой эффективную меру для выявления ЗНО на ранней стадии и, таким образом, способствует снижению смертности [4]. Все маммографические исследования обязательно оцениваются по шкале BI-RADS. Оценки BI-RADS 1 и 2 соответствуют отрицательному результату в отношении ЗНО молочной железы. При BI-RADS 3 вероятность наличия ЗНО составляет более 0 и до 2%, при таком результате пациентам назначается наблюдение с 6-месячными интервалами и динамической маммографией. BI-RADS 4 и 5 означают высокую вероятность ЗНО (от 2% до 95% и более 95% соответственно) и направляются к врачу-онкологу, который организует выполнение биопсии патологического образования [5].

В связи с дефицитом кадров в лучевой диагностике [6] активно разрабатываются базы данных с наличием и отсутствием ЗНО [7], и внедряются технологии искусственного интеллекта в том числе и для анализа маммографических исследований [8]. Однако, по данным научной литературы, диагностическая точность таких ИИ-сервисов может колебаться в достаточно широком диапазоне. Так, например, согласно ряду публикаций, ROC AUC может составлять от 0,613 до 0,994 [9–11].

Помимо этого, необходимо отметить, что в научной литературе обычно изучается согласованность ИИ-сервисов и врачей-рентгенологов только при оценке плотности молочных желез по шкале ACR [12–14], тогда как при оценке вероятности ЗНО по шкале BI-RADS согласованность до сих пор не изучалась.

В связи с вышеизложенным, представляются актуальными исследования, направленные на оценку диагностической точности и согласованности ИИ-сервисов при оценке маммографических исследований по шкале BI-RADS.

Цель. Оценить диагностическую точность и согласованность трех ИИ-сервисов между собой и с врачом-экспертом при оценке маммографических исследований по шкале BI-RADS.

Материалы и методы. Выполнено смешанное исследование (количественный компонент — ретроспективное диагностическое исследование, качественный компонент — аналитическое исследование).

Данное исследование основано на результатах Эксперимента по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы, утвержденного этическим комитетом (выписка из протокола № 2 НЭК МРО РОПР от 20 февраля 2020 года), также зарегистрированного на ClinicalTrials (NCT04489992). Информированное согласие получено от каждого пациента.

Набор данных. Из Единого радиологического информационного сервиса Единой медицинской информационно-аналитической системы Москвы (ЕРИС ЕМИАС) случайным образом было отобрано и выгружено 99 маммографических исследований, выполненных за период с 13.11.2020 по 04.10.2021.

Критерии включения: исследования выполнены пациентам женского пола в возрасте старше 18 лет.

Критерии исключения: маммограммы с артефактами на изображениях, маммограммы без категорий BI-RADS; маммограммы с BI-RADS 0 (поскольку это неполная оценка, требующая дополнительной визуализации); маммограммы с BI-RADS 6 (данная категория подразумевает наличие информации о гистологически подтвержденном ЗНО молочной железы, тогда как ИИ-сервисы анализируют только изображения и соответственно не определяют данную категорию).

Перед выгрузкой все исследования были анонимизированы. Далее для каждого исследования врач-эксперт (рентгенолог с опытом работы не менее 5 лет и ученой степенью/званием по специальности «Рентгенология») определил категорию по шкале BI-RADS 1–5 для правой и левой молочной железы.

Размер выборки обусловлен имеющимися в наличии исследованиями, полученными при подготовке наборов данных. Согласно методу определения объема выборки, предложенному Т.М. Бобровской и соавт., для оценки диагностической точности ИИ-сервисов в данном исследовании минимальный объем выборки должен составлять 80 исследований [15].

ИИ-сервисы. Подготовленный нами набор данных был обработан тремя ИИ-сервисами российских коммерческих компаний (АО «МЕДИЦИНСКИЕ ТЕХНОЛОГИИ Лтд» (АО «МТЛ») — TrioDM (MTL), ООО «Медицинские скрининг системы» — Цельс ММГ, ООО «Платформа Третье Мнение» — Третье мнение ММГ), которые далее по тексту будут представлены в анонимизированном виде: ИИ-сервис «А», ИИ Сервис «В», ИИ Сервис «С».

Выбор ИИ-сервисов основывался на данных матрицы зрелости, которая используется в Московском эксперименте по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения этих технологий в системе здравоохранения (далее — московский эксперимент). Были

выбраны три ИИ-сервиса, которые продемонстрировали наибольшие значения клинической оценки.

С помощью данных ИИ-сервисов были определены категории BI-RADS 1–5 для каждой молочной железы.

Статистический анализ. Для оценки диагностической точности ИИ-сервисов были рассчитаны следующие параметры с 95% доверительными интервалами (ДИ): ROC AUC, чувствительность, специфичность, точность. Статистический анализ проводился на уровне отдельных молочных желёз (n=198). В качестве истинных значений (Ground Truth) использовались оценки врача-эксперта. Расчет параметров был выполнен с помощью веб-инструмента, разработанного ГБУЗ НПКЦ ДиТ ДЗМ (<https://roc-analysis.mosmed.ai>). Параметры диагностической точности рассчитывались для двух бинарных шкал BI-RADS.

внутриклассовой корреляции интерпретировались согласно шкале Landis и Koch: 0,81–1,00 — практически полная согласованность; 0,61–0,80 — существенная согласованность; 0,41–0,60 — умеренная согласованность, 0,21–0,40 — слабая согласованность, 0,00–0,20 — очень слабая согласованность.

Для оценки статистической значимости различий параметров диагностической точности использовался тест Мак-Немара, согласованности — метод отбора с возвратом (bootstrap) с количеством выборок равным 5000.

Результаты
1. Результаты оценки маммографических исследований врачом-экспертом. Средний возраст пациенток составил 56,0 лет. Категории по шкале BI-RADS 1–5 для правой и левой молочной железы, определенные врачом-экспертом, представлены в табл. 1.

Таблица 1

Распределение исследований по категориям полной шкалы BI-RADS в соответствии с оценками эксперта (абсолютные и относительные значения)

Distribution of studies by categories of the full BI-RADS scale according to expert assessments (absolute and relative values)					
Локализация	BI-RADS 1	BI-RADS 2	BI-RADS 3	BI-RADS 4	BI-RADS 5
Правая молочная железа, абс. (%)	20 (20,2)	56 (56,6)	10 (10,1)	9 (9,1)	4 (4)
Левая молочная железа, абс. (%)	15 (15,1)	62 (62,7)	10 (10,1)	10 (10,1)	2 (2)

Бинарная шкала BI-RADS № 1: 0 — нет признаков ЗНО (BI-RADS 1, 2, 3), 1 — есть признаки ЗНО (BI-RADS 4, 5).

Бинарная шкала BI-RADS № 2: 0 — нет признаков ЗНО (BI-RADS 1, 2), 1 — есть признаки ЗНО (BI-RADS 3, 4, 5).

Для обеспечения сопоставимости наших результатов с результатами, представленными в научной литературе, мы также перевели полную шкалу BI-RADS в два вида бинарных шкал (табл. 2 и 3).

2. Параметры диагностической точности ИИ-сервисов. Параметры диагностической точно-

Таблица 2

Распределение исследований по категориям бинарной шкалы BI-RADS № 1 в соответствии с оценками эксперта (абсолютные и относительные значения)

Distribution of studies by BI-RADS 1 binary scale categories according to expert assessments (absolute and relative values)		
Локализация	0 — нет признаков ЗНО (BI-RADS 1, 2, 3)	1 — есть признаки ЗНО (BI-RADS 4, 5)
Правая молочная железа, абс. (%)	86 (86,8)	13 (13,2)
Левая молочная железа	87 (87,8)	12 (12,2)

Для количественной оценки степени согласованности между различными ИИ-сервисами и врачом-экспертом при определении риска рака молочной железы по шкале BI-RADS был применен метод внутриклассовой корреляции (Intraclass Correlation Coefficient, ICC) Пирсона. В нашем исследовании использовалась двусторонняя модель ICC с доверительными интервалами 95%, что соответствует рекомендациям по оценке согласованности между различными методами измерений. Также для оценки согласованности рассчитывалась каппа Коэна.

Внутриклассовая корреляция рассчитывалась для полной шкалы BI-RADS и для двух бинарных шкал, описание которых представлено выше. Значения

сти анализируемых ИИ-сервисов представлены в табл. 4.

При оценке ИИ-сервисов по обоим бинарным шкалам BI-RADS ИИ-сервис «В» продемонстрировал статистически значимо более высокую специфичность по сравнению с ИИ-сервисами «А» и «С» при оценке правой молочной железы. При оценке левой молочной железы специфичность ИИ-сервиса «В» также была статистически значимо выше, чем у ИИ-сервиса «А».

Точность ИИ-сервиса «В» оказалась статистически значимо выше точности ИИ-сервиса «А» при оценке по бинарной шкале BI-RADS № 2 для правой и левой молочных желез.

Таблица 3

Распределение исследований по категориям бинарной шкалы BI-RADS № 2 в соответствии с оценками эксперта (абсолютные и относительные значения)

Table 3

Distribution of studies by BI-RADS 2 binary scale categories according to expert assessments (absolute and relative values)

Локализация	0 — нет признаков ЗНО (BI-RADS 1, 2)	1 — есть признаки ЗНО (BI-RADS 3, 4, 5)
Правая молочная железа, абс. (%)	76 (76,8)	23 (23,2)
Левая молочная железа, абс. (%)	77 (77,8)	22 (22,2)

ROC AUC ИИ-сервиса «С» оказался статистически значимо выше, чем у ИИ-сервисов «А» и «В» только по бинарной шкале BI-RADS № 1 для левой молочной железы, тогда как в остальных случаях статистически значимые различия отсутствовали.

чески значимые различия наблюдались только для левой молочной железы (рис. 1).

Для значений каппы Коэна наблюдалась аналогичная картина. В целом значения каппы Коэна практически полностью соответствовали значениям

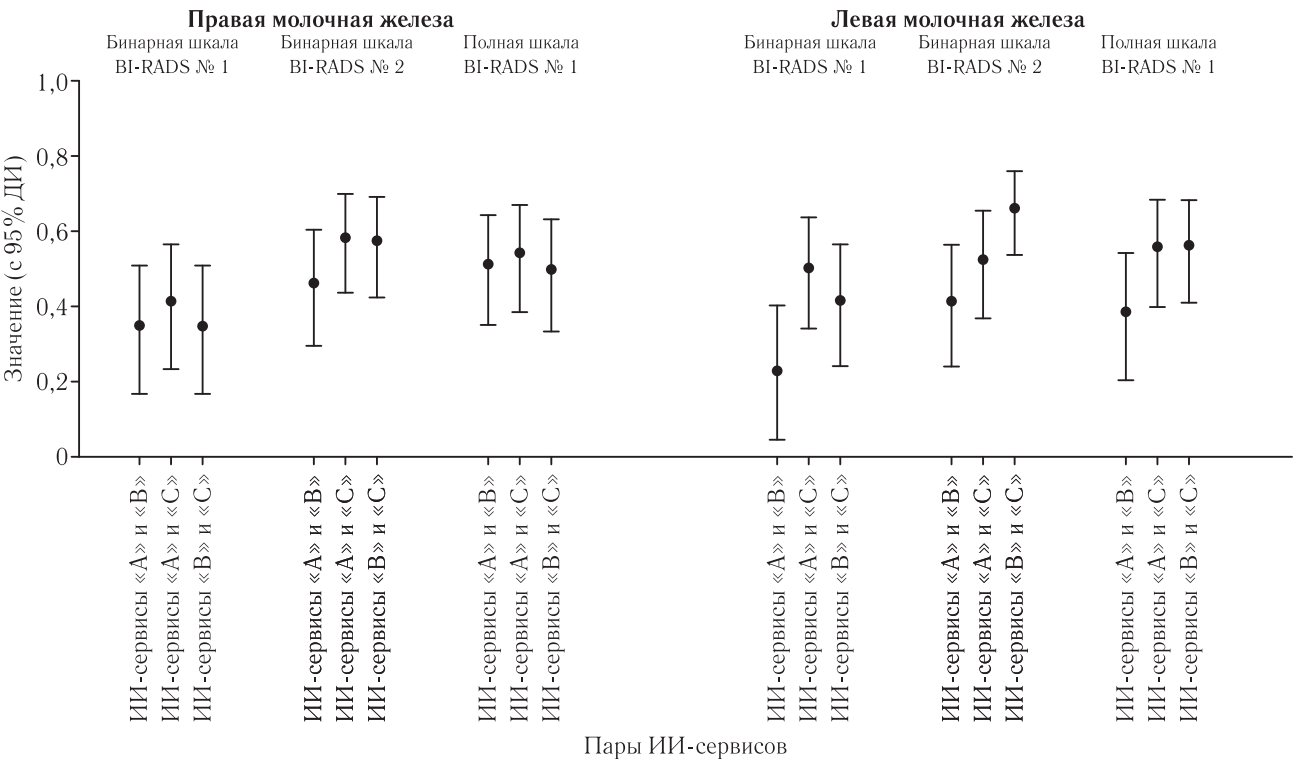


Рис. 1. Согласованность между ИИ-сервисами при оценке маммографических исследований по двум бинарным шкалам и полной шкале BI-RADS, рассчитанная методом внутриклассовой корреляции

Fig. 1. Agreement between AI services in assessing mammographic examinations using two binary scales and the full BI-RADS scale, calculated using the intraclass correlation method

В большинстве случаев чувствительность ИИ-сервисов статистически значимо не отличалась.

3. Согласованность ИИ-сервисов. Результаты оценки согласованности ИИ-сервисов между собой представлены в табл. 5.

По бинарной шкале BI-RADS № 1 согласованность ИИ-сервисов между собой, оценивавшаяся методом внутриклассовой корреляции, находилась в диапазоне от 0,232 до 0,504. По бинарной шкале BI-RADS № 2 согласованность находилась в диапазоне от 0,416 до 0,663. По полной шкале BI-RADS согласованность находилась в диапазоне от 0,387 до 0,564.

При оценке согласованности между ИИ-сервисами методом внутриклассовой корреляции статисти-

чески значимые различия наблюдались только для левой молочной железы (рис. 2).

Результаты оценки согласованности ИИ-сервисов с врачом-рентгенологом представлены в табл. 6.

По бинарной шкале BI-RADS № 1 согласованность ИИ-сервисов с врачом-экспертом, оценивавшаяся методом внутриклассовой корреляции, находилась в диапазоне от 0,288 до 0,641. По бинарной шкале BI-RADS № 2 согласованность находилась в диапазоне от 0,518 до 0,832. По полной шкале BI-RADS согласованность находилась в диапазоне от 0,570 до 0,701. Был выявлен ряд статистически значимых различий между уровнями согласованности. Так, например, по бинарной шкале BI-RADS № 2 согласованность ИИ-сервиса «В» с врачом

Таблица 4

Table 4

Диагностическая точность ИИ-сервисов для бинарных шкал BI-RADS № 1 и BI-RADS № 2

Diagnostic accuracy of AI services for binary scales BI-RADS No. 1 and BI-RADS No. 2

Молочная железа	Параметр	Бинарная шкала BI-RADS № 1						Бинарная шкала BI-RADS № 2					
		ИИ Сервис «А»	ИИ Сервис «В»	ИИ Сервис «С»	р для «А» и «В»	р для «А» и «С»	р для «В» и «С»	ИИ Сервис «А»	ИИ Сервис «В»	ИИ Сервис «С»	р для «А» и «В»	р для «А» и «С»	р для «В» и «С»
Правая	ROC AUC	0,744 (0,607–0,880)	0,758 (0,615–0,904)	0,744 (0,603–0,886)	0,878	1,000	0,896	0,812 (0,719–0,910)	0,915 (0,841–0,988)	0,804 (0,710–0,899)	0,121	0,850	0,064
	Точность	0,838 (0,766–0,911)	0,91 (0,866–0,973)	0,838 (0,766–0,911)	0,076	1,000	0,076	0,828 (0,754–0,903)	0,939 (0,892–0,986)	0,838 (0,766–0,911)	0,026	1,000	0,021
	Специфичность	0,872 (0,802–0,943)	0,977 (0,945–1,000)	0,872 (0,802–0,943)	0,011	1,000	0,003	0,842 (0,760–0,903)	0,961 (0,917–1,000)	0,868 (0,792–0,944)	0,022	0,790	0,039
	Чувствительность	0,615 (0,351–0,880)	0,538 (0,267–0,809)	0,615 (0,351–0,880)	1,000	1,000	1,000	0,783 (0,614–0,951)	0,870 (0,732–1,000)	0,739 (0,560–0,919)	0,726	1,000	0,453
Левая	ROC AUC	0,700 (0,544–0,852)	0,644 (0,505–0,779)	0,870 (0,757–0,993)	0,544	0,009	0,002	0,795 (0,699–0,896)	0,883 (0,800–0,969)	0,870 (0,783–0,957)	0,132	0,115	0,795
	Точность	0,788 (0,707–0,868)	0,879 (0,814–0,943)	0,909 (0,852–0,966)	0,078	0,004	0,581	0,808 (0,731–0,886)	0,919 (0,866–0,973)	0,899 (0,840–0,958)	0,034	0,063	0,774
	Специфичность	0,816 (0,735–0,897)	0,954 (0,910–0,998)	0,920 (0,862–0,977)	0,004	0,022	0,453	0,818 (0,732–0,904)	0,948 (0,898–0,998)	0,922 (0,862–0,982)	0,030	0,076	0,726
	Чувствительность	0,583 (0,304–0,862)	0,333 (0,067–0,600)	0,883 (0,662–1,000)	0,375	0,250	0,031	0,773 (0,598–0,948)	0,818 (0,657–0,979)	0,818 (0,657–0,979)	1,000	1,000	1,000

Примечание: в круглых скобках представлены значения 95% ДИ. Полуширным выделены значения $p < 0,05$.
Note: 95% CI values are shown in parentheses. Values in bold indicate $p < 0,05$.

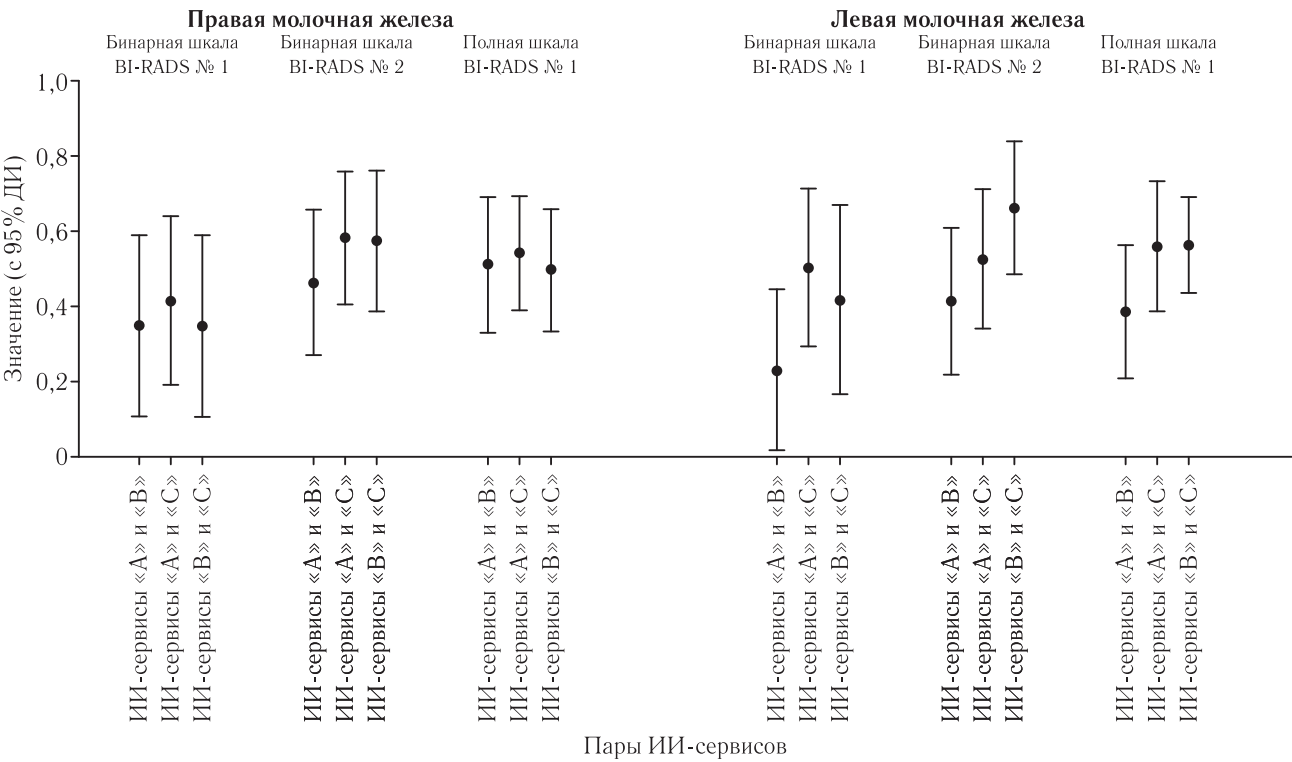


Рис. 2. Согласованность между ИИ-сервисами при оценке маммографических исследований по двум бинарным шкалам и полной шкале BI-RADS, рассчитанная методом каппы Коэна

Fig. 2. Agreement between AI services in assessing mammographic examinations using two binary scales and the full BI-RADS scale, calculated using Cohen's kappa method

была выше, чем согласованность ИИ-сервисов «А» и «С» с врачом для правой молочной железы, и выше, чем согласованность ИИ-сервиса «А» и «С» с врачом для левой молочной железы (рис. 3).

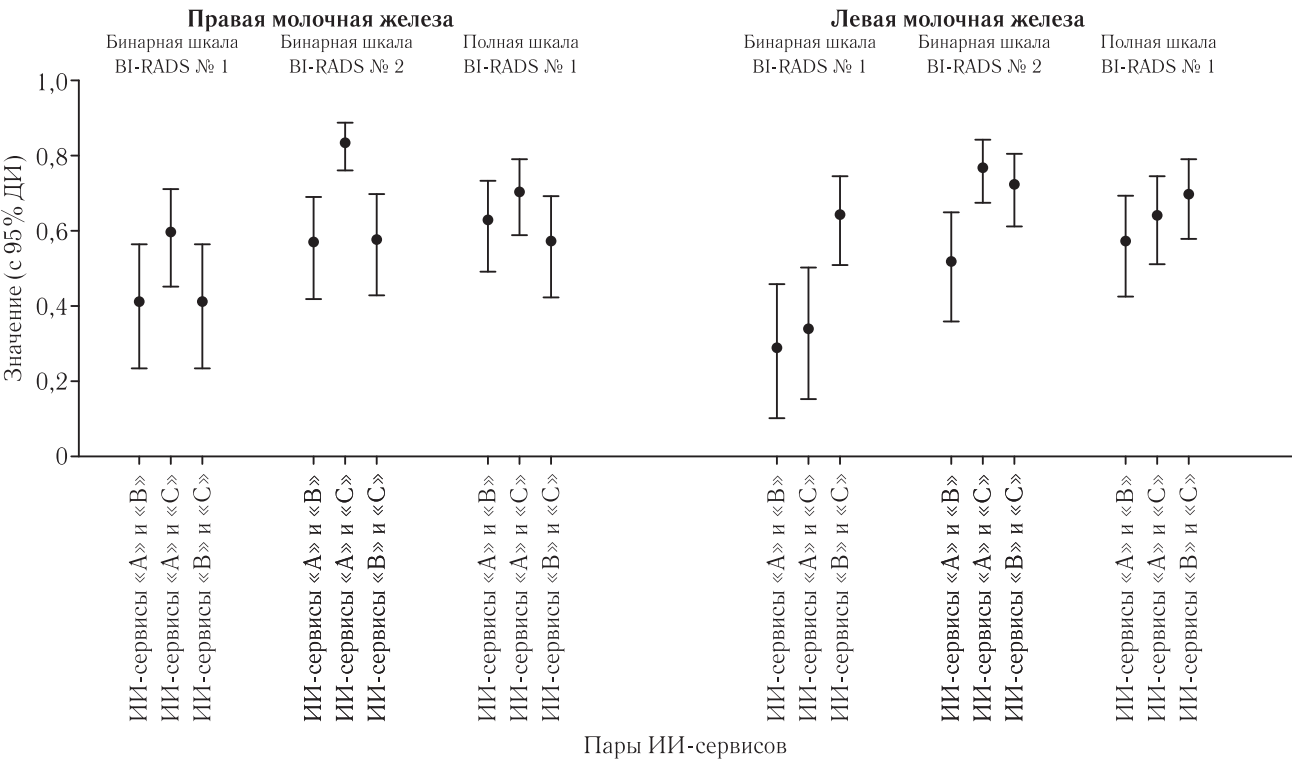


Рис. 3. Согласованность ИИ-сервисов с врачом-рентгенологом при оценке маммографических исследований по двум бинарным шкалам и полной шкале BI-RADS, рассчитанная методом внутриклассовой корреляции

Fig. 3. Agreement between AI services and a radiologist in assessing mammographic examinations using two binary scales and the full BI-RADS scale, calculated using the intraclass correlation method

Для значений каппы Коэна наблюдалась аналогичная картина. В целом значения каппы Коэна практически полностью соответствовали значениям внутрикласовой корреляции, но имели более широкие доверительные интервалы (рис. 4).

соответствует бинарной шкале BI-RADS № 2 в нашем исследовании. В таких работах ROC AUC находился в диапазоне от 0,620 до 0,940, точность — 0,720–0,877, специфичность — 0,370–0,970, чувствительность — от 0,706, до 0,870 [19–

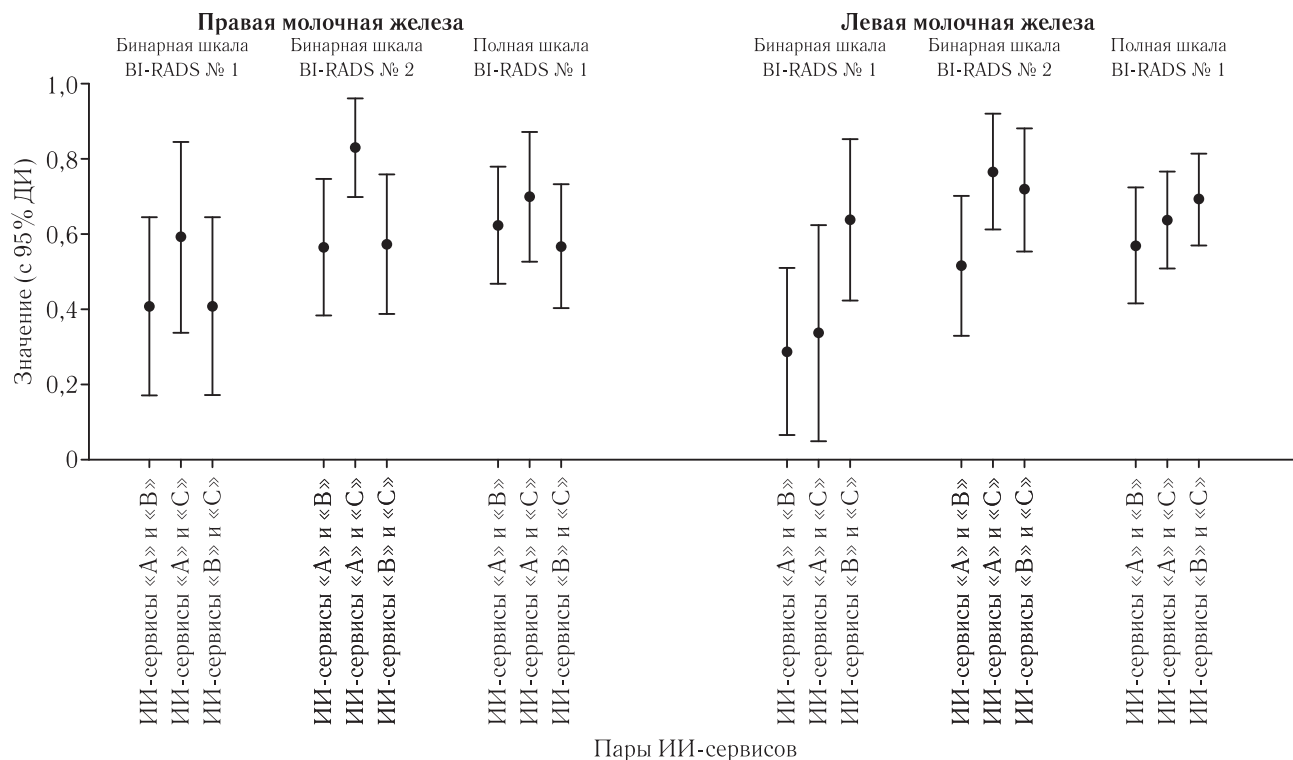


Рис. 4. Согласованность ИИ-сервисов с врачом-рентгенологом при оценке маммографических исследований по двум бинарным шкалам и полной шкале BI-RADS, рассчитанная методом каппы Коэна

Fig. 4. Agreement between AI services and a radiologist in assessing mammographic examinations using two binary scales and the full BI-RADS scale, calculated using Cohen's kappa method

Обсуждение

1. Диагностическая точность ИИ-сервисов.

В рамках нашего исследования было показано, что диапазоны параметров диагностической точности ИИ-сервисов по бинарной шкале BI-RADS № 1 были следующими: ROC AUC — 0,644–0,876, точность — 0,788–0,909, специфичность — 0,816–0,977, чувствительность — 0,333–0,883. При оценке по бинарной шкале BI-RADS № 2: ROC AUC — 0,795–0,915, точность — 0,808–0,939, специфичность — 0,818–0,961, чувствительность — 0,739–0,870.

В ряде научных публикаций авторы представили результаты оценки маммографических ИИ-сервисов по шкале, аналогичной бинарной шкале BI-RADS № 1, использовавшейся в нашем исследовании. В таких работах ROC AUC находился в диапазоне от 0,644 до 0,988, точность — 0,732–0,968, специфичность — 0,588–0,963, чувствительность — от 0,740, до 0,923 [16–18]. Наши результаты в целом соответствуют представленным диапазонам, однако один из ИИ-сервисов продемонстрировал аномально низкую чувствительность 0,333.

Однако в некоторых публикациях авторы при сведении полной шкалы BI-RADS к бинарной относят категорию 3 к группе «Есть признаки ЗНО», что

21]. Наши результаты соответствуют представленным диапазонам.

2. Согласованность ИИ-сервисов. По данным нашего исследования, согласованность ИИ-сервисов по бинарной шкале BI-RADS № 1 находилась в диапазоне от 0,232 до 0,504. По бинарной шкале BI-RADS № 2 согласованность ИИ-сервисов находилась в диапазоне от 0,416 до 0,663. По полной шкале BI-RADS согласованность ИИ-сервисов находилась в диапазоне от 0,387 до 0,564.

При оценке согласованности между ИИ-сервисами и врачом-экспертом по бинарной шкале BI-RADS № 1 согласованность находилась в диапазоне от 0,288 до 0,641. По бинарной шкале BI-RADS № 2 согласованность ИИ-сервисов находилась в диапазоне от 0,518 до 0,832. По полной шкале BI-RADS согласованность ИИ-сервисов находилась в диапазоне от 0,570 до 0,701.

По сравнению с результатами нашего предыдущего исследования, посвященного оценке согласованности между врачами-рентгенологами [22], согласованность ИИ-сервисов и врача-эксперта (0,288 до 0,641) по бинарной шкале BI-RADS № 1 оказалась во всех случаях ниже, чем согласованность между врачами-рентгенологами (0,810–0,941).

По бинарной шкале BI-RADS № 2 согласованность ИИ-сервисов и врача-эксперта (0,518 до 0,832) в большинстве случаев оказалась ниже согласованности между врачами-рентгенологами (0,788–0,911).

По полной шкале BI-RADS согласованность ИИ-сервисов и врача-эксперта (0,570 до 0,701) во всех случаях была ниже, чем согласованность между врачами-рентгенологами (0,836–0,963).

Интересным также представляется тот факт, что согласованность ИИ-сервисов и врача-эксперта в нашем исследовании находилась на уровне согласованности опытных рентгенологов (0,682 (95% ДИ 0,660–0,703)), полученном в исследовании Н. Тап и соавт. [16].

Оценивая уровни согласованности ИИ-сервисов и врача-эксперта, полученные в нашем исследовании, в целом, можно предположить, что более низкая согласованность по сравнению с врачами-рентгенологами, которую мы получили в предыдущем исследовании, может быть обусловлена тем, что ИИ-сервисы работают исключительно с маммографическими изображениями, тогда как врачи-рентгенологи обязательно имеют доступ к электронным медицинским картам пациентов и архиву исследований (при наличии).

Согласованность между ИИ-сервисами от умеренной до существенной может быть обусловлена несколькими факторами: размер и особенности разметки использовавшегося для обучения набора данных [23], особенности сетевой архитектуры ИИ-алгоритма [23, 24], пороговые значения [25] и др. Помимо этого, в масштабном исследовании, посвященном оценке 1537 ИИ-алгоритмов для анализа маммографических исследований, было показано, что при сопоставимой диагностической точности ИИ-алгоритмы могут быть более чувствительными к разным видам ЗНО молочных желез [26].

Ограничениями исследования можно считать использование заключения врача-эксперта в качестве референсного стандарта без гистологической верификации, относительно небольшой размер выборки для оценки согласованности (особенно

каппы Козна), а также невозможность детального анализа архитектурных особенностей коммерческих ИИ-сервисов.

Выводы

1. По бинарной шкале BI-RADS № 1 ROC AUC ИИ-сервисов находился в диапазоне 0,644–0,876, точность — 0,788–0,909, специфичность — 0,816–0,977, чувствительность — 0,333–0,883. По бинарной шкале BI-RADS № 2 ROC AUC ИИ-сервисов находился в диапазоне 0,795–0,915, точность — 0,808–0,939, специфичность — 0,818–0,961, чувствительность — 0,739–0,870. Полученные результаты подтверждают данные литературы, свидетельствующие о широких диапазонах колебаний диагностической точности ИИ-сервисов.

2. Согласованность между парами ИИ-сервисов находилась в диапазоне от слабой (0,232) до существенной (0,663). Согласованность между ИИ-сервисами и врачом-экспертом находилась в диапазоне от слабой (0,288) до практически полной (0,832).

3. С учетом различных шкал BI-RADS согласованность ИИ-сервисов и врача-эксперта оказалась в подавляющем большинстве случаев ниже, чем согласованность между врачами-рентгенологами, которую мы оценивали в предыдущем исследовании. Это может быть обусловлено тем, что ИИ-сервисы работают только с изображениями одного исследования, тогда как у врачей всегда есть доступ к архиву исследований и электронной медицинской карте пациентов.

Заключение. ИИ-сервисы продемонстрировали вариабельную диагностическую точность при оценке изменений молочных желез по бинарным шкалам BI-RADS, что согласуется с данными литературы. Согласованность между ИИ-сервисами, а также между ИИ-сервисами и врачом-экспертом варьировала от слабой до практически полной. При этом в большинстве случаев она была ниже ранее установленной межэкспертной согласованности врачей-рентгенологов, что может быть связано с ограничением ИИ-сервисов анализом только изображений одного исследования без учета архивных данных и информации электронной медицинской карты.

Сведения об авторах:

Азарян Авет Сергеевич — аспирант государственного бюджетного учреждения здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы»; 127051, Москва, ул. Петровка, д. 24; ассистент кафедры рентгенологии и ультразвуковой диагностики федерального государственного бюджетного учреждения дополнительного профессионального образования «Центральная государственная медицинская академия» Управления делами Президента Российской Федерации, 121359, Москва, ул. Маршала Тимошенко д. 19, с. 1А; e-mail: Dr.Azaryan@yandex.ru; ORCID 0009–0007–8975–0017;

Хрустачева Маргарита Юрьевна — аспирант государственного бюджетного учреждения здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы»; 127051, Москва, ул. Петровка, д. 24; e-mail: Myukhrustacheva@mail.ru; ORCID 0009–0001–1381–0809;

Пестренкин Лев Дмитриевич — младший научный сотрудник отдела медицинской информатики, радиомикрии и радиогеномики государственного бюджетного учреждения здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы»; 127051, Москва, ул. Петровка, д. 24; e-mail: PestreninLD@zdrav.mos.ru; ORCID 0000–0002–1786–4329;

Васильев Юрий Александрович — доктор медицинских наук, директор государственного бюджетного учреждения здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы»; 127051, Москва, ул. Петровка, д. 24; e-mail: nrcmr@zdrav.mos.ru; ORCID 0000–0002–5283–5961;

Владимирский Антон Вячеславович — доктор медицинских наук, заместитель директора по научной работе государственного бюджетного учреждения здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы»; 127051, Москва, ул. Петровка, д. 24, e-mail: VladzimirskijAV@zdrav.mos.ru; ORCID 0000–0002–2990–7736;

Омелянская Ольга Васильевна — руководитель по управлению подразделениями Дирекции наука государственного бюджетного учреждения здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы»; 127051, Москва, ул. Петровка, д. 24, e-mail: OmelyanskayaOV@zdrav.mos.ru; ORCID 0000–0002–0245–4431;

Доможирова Алла Сергеевна — доктор медицинских наук, председатель научной проблемной комиссии, ученый секретарь государственного бюджетного учреждения здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы»; 127051, Москва, ул. Петровка, д. 24, e-mail: 2356563@mail.ru; ORCID 0000–0003–0806–3164;

Арзамасов Кирилл Михайлович — доктор медицинских наук, заведующий отдела медицинской информатики, радиомики и радиогеномики государственного бюджетного учреждения здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы»; 127051, Москва, ул. Петровка, д. 24; e-mail: ArzamasovKM@zdrav.mos.ru; ORCID 0000–0001–7786–0349;

Гацук Андрей Сергеевич — младший научный сотрудник отдела медицинской информатики, радиомики и радиогеномики государственного бюджетного учреждения здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы»; 127051, Москва, ул. Петровка, д. 24; e-mail: GatsukAS@zdrav.mos.ru.

Information about the authors:

Avet S. Azaryan — postgraduate student, Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department, 24 Petrovka St., Moscow, 127051, Russia; assistant Professor at the Department of the Central State Medical Academy, Marshal Timoshenko St., 19, p. 1A, Moscow, 121359, Russia; e-mail: Dr.Azaryan@yandex.ru; ORCID 0009–0007–8975–0017;

Leo D. Pestrenin — junior Research Fellow, Department of Medical Informatics, Radiomics and Radiogenomics, Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department, 24 Petrovka St., Moscow, 127051, Russia; e-mail: PestreninLD@zdrav.mos.ru; ORCID 0000–0002–1786–4329;

Margarita S. Khrustacheva — postgraduate student, Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department, 24 Petrovka St., Moscow, 127051, Russia; e-mail: Myukhrustacheva@mail.ru; ORCID 0009–0001–1381–0809;

Yuriy A. Vasilev — Dr. of Sci. (Med.), Medical Director, Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department, 24 Petrovka St., Moscow, 127051, Russia; e-mail: nrcmr@zdrav.mos.ru; ORCID 0000–0002–5283–5961;

Anton V. Vladzimirskiy — Dr. of Sci. (Med.), Deputy Director for R&D, Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department, 24 Petrovka St., Moscow 127051, Russia; e-mail: VladzimirskijAV@zdrav.mos.ru; ORCID 0000–0002–2990–7736;

Olga V. Omelyanskaya — Chief Administrative Officer of R&D/CAO of R&D State Budget-Funded Health Care Institution of the City of Moscow «Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department»; e-mail: OmelyanskayaOV@zdrav.mos.ru; ORCID 0000–0002–0245–4431;

Alla S. Domozhirova — Dr. of Sci. (Med.), Chair of the Scientific Problem Commission, Scientific Secretary State Budget-Funded Health Care Institution of the City of Moscow «Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department»; ORCID 0000–0003–0806–3164;

Kirill M. Arzamasov — Dr. of Sci. (Med.), Head of Department of Medical Informatics, Radiomics, and Radiogenomics, Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department, 24 Petrovka St., Moscow, 127051, Russia; e-mail: ArzamasovKM@zdrav.mos.ru; ORCID 0000–0001–7786–0349;

Anderei S. Gatsuk — junior Research Fellow, Department of Medical Informatics, Radiomics and Radiogenomics, Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department, 24 Petrovka St., Moscow, 127051, Russia; e-mail: GatsukAS@zdrav.mos.ru.

Вклад авторов. Все авторы подтверждают соответствие своего авторства, согласно 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 распределен следующим образом: концепция и план исследования — Ю. А. Васильев, А. В. Владимирский, О. В. Омелянская, А. С. Доможирова; сбор и анализ данных — А. С. Азарян, М. Ю. Хрустачева, Л. Д. Пестренин; подготовка рукописи — А. С. Азарян, Л. Д. Пестренин, К. М. Арзамасов, А. С. Гацук.

Authors' contributions. All authors confirm their authorship according to the international ICMJE criteria (all authors made substantial contributions to the conception, research, and preparation of the article, and read and approved the final version before publication). Special contribution: YuAV, AVV, OVO, ASD aided in the concept and plan of the study; ASA, MYuKh, LDP provided collection and analysis of data; ASA, LDP, KMA, ASG preparation of the manuscript.

Потенциальный конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Disclosure: the author declares no conflict of interest.

Соответствие принципам этики: Исследование основано на результатах Эксперимента по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы, утвержденного этическим комитетом (выписка из протокола № 2 НЭК МРО РОПР от 20 февраля 2020 года), также зарегистрированного на ClinicalTrials (NCT04489992). Информированное согласие получено от каждого пациента.

Adherence to ethical standards: The study is based on the results of an Experiment on the use of innovative technologies in the field of computer vision for the analysis of medical images and further application in the healthcare system of the city of Moscow, approved by the Ethics Committee (extract from Protocol No. 2 of NEC MRO RORR dated February 20, 2020), also registered on ClinicalTrials (NCT04489992). Informed consent was obtained from each patient.

Финансирование: Данная статья подготовлена авторским коллективом в рамках НИР «Создание системы контроля качества автономного искусственного интеллекта», в соответствии с приказом от 27.11.2025 № 1213 Департамента здравоохранения города Москвы.

Funding: This paper was prepared by a group of authors as a part of the research and development effort titled «Development of a quality control system for autonomous artificial intelligence», in accordance with the Order No. 1213 dated November 27, 2025 issued by the Moscow Health Care Department.

Поступила/Received: 20.01.2026

Принята к печати/Accepted: 29.05.2026

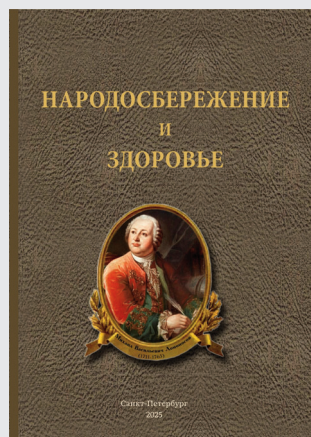
Опубликована/Published: 29.06.2026

ЛИТЕРАТУРА / REFERENCES

1. Bray F., Laversanne M., Sung H. et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries // *CA: A Cancer Journal for Clinicians*. 2024. Vol. 74, No. 3. P. 229–263. doi: 10.3322/caac.21834.
2. Владимирский А.В., Васильев Ю.А., Шахов А.В. и др. Влияние технологий искусственного интеллекта на активную выявляемость рака молочной железы: эффективность и масштабирование // *Здравоохранение стран СНГ*. 2026. Т. 2, № 1. С. 11–19. [Vladymyrsky A.V., Vasilev Yu.A., Shakhov A.V. et al. The impact of artificial intelligence technologies on active breast cancer detection: efficiency and scalability. *The CIS Healthcare*, 2026, Vol. 2, No. 1, pp. 11–19 (In Russ.)]. doi: 10.21045/3033-6341-2026-2-1-11-19.
3. Мерабишвили В.М. Состояние онкологической помощи в России: рак молочной железы среди женского населения. Заболеваемость, смертность, достоверность учета, детальная локализационная и гистологическая структура. Популяционное исследование на уровне федерального округа // *Вопросы онкологии*. 2022. Т. 68, № 3. С. 286–293. [Merabishvili V.M. The state of oncological care in Russia: breast cancer among the female population. Morbidity, mortality, reliability of accounting, detailed localization and histological structure. Population-based research at the federal district level. *Issues of oncology*, 2022, Vol. 68, No. 3, pp. 286–293 (In Russ.)]. doi: 10.37469/0507-3758-2022-68-3-286-293.
4. Ren W., Chen M., Qiao Y., Zhao F. Global guidelines for breast cancer screening: A systematic review // *Breast. Churchill Livingstone*. 2022. Vol. 64. P. 85–99. doi: 10.1016/j.breast.2022.04.003.
5. Spak D.A., Plaxco J.S., Santiago L., Dryden M.J., Dogan B.E. BI-RADS® fifth edition: A summary of changes // *Diagnostic and Interventional Imaging*. 2017. Vol. 98, No. 3. P. 179–190. doi: 10.1016/j.diii.2017.01.001.
6. Васильев Ю.А., Сон И.М., Пестренин Л.Д., Владимирский А.В. Результативность профилактической маммографии в Российской Федерации: сравнение итогов первого этапа диспансеризации в 2019 и 2022 гг. // *Менеджер здравоохранения*. 2024. Т. 11. С. 63–76. [Vasiliev Yu.A., Son I.M., Pestrenin L.D., Vladimírsky A.V. Effectiveness of preventive mammography in the Russian Federation: comparison of the results of the first stage of medical examination in 2019 and 2022. *Healthcare Manager*, 2024, Vol. 11, pp. 63–76 (In Russ.)]. doi: 10.21045/1811-0185-2024-11-63-76.
7. Cai H., Wang J., Dan T. et al. An Online Mammography Database with Biopsy Confirmed Types // *Scientific Data*. 2023. Vol. 10. P. 123. doi: 10.1038/s41597-023-02025-1.
8. Васильев Ю.А., Тыров И.А., Владимирский А.В. и др. Двойной просмотр результатов маммографии с применением технологий искусственного интеллекта: новая модель организации массовых профилактических исследований // *Digital Diagnostics*. 2023. Т. 4, № 2. С. 93–104. [Vasiliev Yu.A., Tyrov I.A., Vladimírsky A.V. et al. Double viewing of mammography results using artificial intelligence technologies: a new model for organizing mass preventive research. *Digital Diagnostics*, 2023, Vol. 4, No. 2, pp. 93–104 (In Russ.)]. doi: 10.17816/DD321423.
9. Арзамасов К.М., Васильев Ю.А., Владимирский А.В. и др. Применение компьютерного зрения для профилактических исследований на примере маммографии // *Профилактическая медицина*. 2023. Т. 26, № 6. С. 117–123. [Arzamasov K.M., Vasiliev Yu.A., Vladimírsky A.V. et al. Application of computer vision for preventive examinations using mammography as an example. *Preventive Medicine*, 2023, Vol. 26, No. 6, pp. 117–123 (In Russ.)]. doi: 10.17116/profmed20232606117.
10. Dang L.A., Chazard E., Poncelet E. et al. Impact of artificial intelligence in breast cancer screening with mammography // *Breast Cancer*. 2022. Vol. 29, No. 6. P. 967–977. doi: 10.1007/s12282-022-01375-9.
11. 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. Vol. 23, No. 6. P. 2341–2356. doi: 10.1007/s10238-022-00895-0.
12. Le Boulc'h M., Bekhouche A., Kermarrec E. et al. Comparison of breast density assessment between human eye and automated software on digital and synthetic mammography: Impact on breast cancer risk // *Diagn. Interv. Imaging*. 2020. Vol. 101, No. 12. P. 811–819. doi: 10.1016/j.diii.2020.07.004.
13. Ji H., Jang Mjin, Chang J.M. Variability in Breast Density Estimation and Its Impact on Breast Cancer Risk Assessment // *J. Breast Cancer*. 2024. Vol. 27, No. 5. P. 334–342. doi: 10.4048/jbc.2024.0101.
14. Rigaud B., Weaver O.O., Dennison J.B. et al. Deep Learning Models for Automated Assessment of Breast Density Using Multiple Mammographic Image Types // *Cancers (Basel)*. 2022. Vol. 14, No. 20, P. 5003. doi: 10.3390/cancers14205003.
15. Бобровская Т.М., Васильев Ю.А., Никитина Н.Ю. и др. Объем выборки для оценки диагностической точности программного обеспечения на основе технологий искусственного интеллекта в лучевой диагностике // *Сибирский журнал клинической и экспериментальной медицины*. 2024. Т. 39, № 3. С. 188–198. [Bobrovskaya T.M., Vasiliev Yu.A., Nikitina N.Yu. et al. Sample size for assessing the diagnostic accuracy of software based on artificial intelligence technologies in radiation diagnostics. *Siberian Journal of Clinical and Experimental Medicine*, 2024, Vol. 39, No. 3, pp. 188–198 (In Russ.)]. doi: 10.29001/2073-8552-2024-39-3-188-198.
16. Tan H., Wu Q., Wu Y. et al. Mammography-based artificial intelligence for breast cancer detection, diagnosis, and BI-RADS categorization using multi-view and multi-level convolutional neural networks // *Insights Imaging*. 2025. Vol. 16, No. 1. P. 109. doi: 10.1186/s13244-025-01983-x.

17. Васильев Ю.А., Владимирский А.В., Арзамасов К.М. и др. Первые 10 000 маммографических исследований, выполненных в рамках услуги «описание и интерпретация данных маммографического исследования с использованием искусственного интеллекта» // *Менеджер здравоохранения*. 2023. № 8. С. 54–67. [Vasiliev Yu.A., Vladimirovsky A.V., Arzamasov K.M. et al. The first 10,000 mammographic examinations performed within the framework of the service «description and interpretation of mammographic examination data using artificial intelligence». *Healthcare Manager*, 2023, No. 8, pp. 54–67 (In Russ.)]. doi: 10.21045/1811-0185-2023-8-54-67.
18. Sasaki M., Tozaki M., Rodriguez-Ruiz A. et al. Artificial intelligence for breast cancer detection in mammography: experience of use of the ScreenPoint Medical Transpara system in 310 Japanese women // *Breast Cancer*. 2020. Vol. 27, No. 4. P. 642–651. doi: 10.1007/S12282-020-01061-8.
19. Arzamasov K., Vasilev Y., Vladymyrsky A. et al. An International Non-Inferiority Study for the Benchmarking of AI for Routine Radiology Cases: Chest X-ray, Fluorography and Mammography // *Healthcare*. 2023. Vol. 11, No. 12. P. 1684. doi: 10.3390/healthcare11121684.
20. Li H.Y., Lin Y., Hong Y.T., Chou C.P. Comparative Study of Artificial Intelligence-based System Alone in Synthesized Mammography Versus Radiologist's Interpretation of Digital Breast Tomosynthesis in Screening Women // *Journal of Radiological Science*. 2024. Vol. 49. No. 1. P. 59–65. doi: 10.4103/jradiolsci.JRA-DIOLSCI-D-23-00027.
21. Caldas F.A.A., Caldas H.C., Henrique T. et al. Evaluating the performance of artificial intelligence and radiologists accuracy in breast cancer detection in screening mammography across breast densities // *European Journal of Radiology Artificial Intelligence*. 2025. Vol. 2. P. 100013. doi: 10.1016/j.ejrai.2025.100013.
22. Азарян А.С., Пестренин Л.Д., Васильев Ю.А., Ахмад Е.С., Арзамасов К.М. Согласованность между рентгенологами города Москвы при интерпретации маммографических исследований по шкале BI-RADS // *Альманах клинической медицины*. 2024. Т. 52, № 7. С. 377–384. [Azaryan A.S., Pestrenin L.D., Vasiliev Yu.A., Akhmad E.S., Arzamasov K.M. Agreement between Moscow radiologists in the interpretation of mammographic studies using the BI-RADS scale. *Almanac of Clinical Medicine*, 2024, Vol. 52, No. 7, pp. 377–384 (In Russ.)]. doi: 10.18786/2072-0505-2024-52-035.
23. Taya M. Comparative Performance of Artificial Intelligence Algorithms for Screening Mammography. *Radiol. Imaging Cancer*. 2020. Vol. 2, No. 6. P. e209034. doi: 10.1148/rycan.2020209034.
24. Renard X., Laugel T., Detyniecki M. Understanding Prediction Discrepancies in classification // *Machine Learning*. 2024. Vol. 113, No. 3. P. 7997–8026. doi: 10.1007/s10994-024-06557-4.
25. Han T., Yun H., Sur Y.K., Park H. Optimizing Artificial Intelligence Thresholds for Mammographic Lesion Detection: A Retrospective Study on Diagnostic Performance and Radiologist-Artificial Intelligence Discordance // *Diagnostics*. 2025. Vol. 15, No. 11. P. 1368. doi: 10.3390/diagnostics15111368
26. Chen Y., Partridge G.J.W., Vazirabad M. et al. Performance of Algorithms Submitted in the 2023 RSNA Screening Mammography Breast Cancer Detection AI Challenge // *Radiology*. 2025. Vol. 316, No. 2. P. e241447. doi: 10.1148/radiol.241447.

Уважаемые коллеги!



В монографии, подготовленной в виде научных очерков, представлены наиболее значимые события становления в России направления народосбережения и охраны здоровья. Исторические события изложены в сочетании следующих важных факторов — демографии и народонаселения, прироста и потерь населения, длительности жизни, заболеваний, влияющих на численность населения, и других факторов. Рассмотрены основные критерии оценки народосбережения: количественные (рождаемость, численность, заболеваемость, потери и др.) и качественные (удовлетворенность и качество жизни, коэффициент счастья, географические и материальные влияния и др.). Описаны психологические и социальные причины воздействия на рождаемость и смертность от разных причин и др. В подготовке книги участвовали ведущие специалисты — представители различных отраслей науки: врачи, эпидемиологи, психологи, биологи и демографы и др. Книга по содержанию и изложению носит междисциплинарный характер, что расширяет круг заинтересованных читателей. Изложенные материалы могут быть интересны и познавательны для специалистов, работающих в сфере народосбережения, ученых, преподавателей, студентов университетов и иных читателей.

Подробная информация о книге
на сайте издательства <https://www.bmosc-spb.ru>.