

ЕРИЖОКОВ Р.А.,

ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия,
e-mail: ErizhokovRA@zdrav.mos.ru, ORCID: 0009-0007-3636-2889

САХАРОВА П.А.,

ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия,
e-mail: SakharovaPA@zdrav.mos.ru, ORCID: 0009-0008-1323-9308

КИРСАНОВ В.В.,

ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия,
e-mail: KirsanovVV2@zdrav.mos.ru, ORCID: 0009-0004-7467-0573

ХОРУЖАЯ А.Н.,

ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия,
e-mail: KhoruzhayaAN@zdrav.mos.ru, ORCID: 0000-0003-4857-5404

ВАРЮХИНА М.Д.,

к.м.н., ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия,
e-mail: VaryukhinaMD@zdrav.mos.ru, ORCID: 0000-0001-8870-7649

АРЗАМАСОВ К.М.,

д.м.н., ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия,
e-mail: ArzamasovKM@zdrav.mos.ru, ORCID: 0000-0001-7786-0349

ВЛАДЗИМИРСКИЙ А.В.,

д.м.н., ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия,
e-mail: VladzimirskijAV@zdrav.mos.ru, ORCID: 0000-0002-2990-7736

ВАСИЛЬЕВ Ю.А.,

д.м.н., ГБУЗ «НПКЦ ДиТ ДЗМ», г. Москва, Россия,
e-mail: VasilevYA1@zdrav.mos.ru, ORCID: 0000-0002-5283-5961

ОЦЕНКА КЛИНИЧЕСКОЙ ПРИМЕНИМОСТИ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ВЫЯВЛЕНИЯ ИНФИЛЬТРАТИВНЫХ ИЗМЕНЕНИЙ ПО РЕНТГЕНОГРАММАМ ОРГАНОВ ГРУДНОЙ КЛЕТКИ В УСЛОВИЯХ СТАЦИОНАРА

DOI: 10.25881/18110193_2026_2_36

Аннотация. Цель исследования: изучить возможности систем искусственного интеллекта (СИИ) в стационарных условиях, в частности в выявлении признаков инфильтрации по рентгенограммам органов грудной клетки (РГ ОГК).

Материалы и методы. Исходными данными являлись РГ ОГК в прямой проекции 1537 пациентов стационарных отделений в возрасте от 18 до 102 лет, полученные из Единой Радиологической Информационной Системы Единого Медицинского Информационно-Аналитического Сервиса (ЕРИС ЕМИАС). Факт наличия инфильтрации на РГ ОГК определялся врачом-рентгенологом (Рентгенолог 1) и рентгенологом-экспертом (Рентгенолог 2) со стажем работы больше 5 лет. Рентгенолог 2 размечал исследования в два этапа: на первом этапе были предоставлены только РГ ОГК, на втором этапе в дополнение к рентгенологическим данным был предоставлен доступ к информации о состоянии пациента.

Результаты. Показатели точности для пациентов с переходным состоянием составили 0,755 для СИИ, 0,756 для Рентгенолога 1 и 0,856 для Рентгенолога 2 соответственно. Статистически значимых различий точности между СИИ и Рентгенологом 1 не было выявлено ($p > 0,05$). При рассмотрении сценария использования СИИ как второго мнения точность для пациентов с переходным состоянием составила 0,833, что статистически значимо превышает точность Рентгенолога 1 и СИИ ($p < 0,05$).

Выводы. СИИ могут использоваться в качестве второго мнения при обнаружении признаков инфильтрации по данным РГ ОГК в условиях стационара. Однако в связи с тем, что диагностическая точность СИИ пока недостаточна для практического применения, требуется продолжение дальнейших исследований, направленных на обучение алгоритмов с использованием как визуальных данных, так и данных электронной медицинской карты.

Ключевые слова: рентгенография органов грудной клетки, инфильтрация, система искусственного интеллекта, клиническая применимость.

Для цитирования: Ерижиков Р.А., Сахарова П.А., Кирсанов В.В., Хоружая А.Н., Варюхина М.Д., Арзамасов К.М., Владимирский А.В., Васильев Ю.А. Оценка клинической применимости систем искусственного интеллекта для выявления инфильтративных изменений по рентгенограммам органов грудной клетки в условиях стационара. Врач и информационные технологии. 2026; 2: 36-49. DOI: 10.25881/18110193_2026_2_36.

ERIZHOKOV R.A.,

Moscow Center for Diagnostics and Telemedicine, Moscow, Russia,
e-mail: ErizhokovRA@zdrav.mos.ru, ORCID: 0009-0007-3636-2889

SAKHAROVA P.A.,

Moscow Center for Diagnostics and Telemedicine, Moscow, Russia,
e-mail: SakharovaPA@zdrav.mos.ru, ORCID: 0009-0008-1323-9308

KIRSANOV V.V.,

Moscow Center for Diagnostics and Telemedicine, Moscow, Russia,
e-mail: KirsanovVV2@zdrav.mos.ru, ORCID: 0009-0004-7467-0573

KHORUZHAYA A.N.,

Moscow Center for Diagnostics and Telemedicine, Moscow, Russia,
e-mail: KhoruzhayaAN@zdrav.mos.ru, ORCID: 0000-0003-4857-5404

VARYUKHINA M.D.,

PhD, Moscow Center for Diagnostics and Telemedicine, Moscow, Russia,
e-mail: VaryukhinaMD@zdrav.mos.ru, ORCID: 0000-0001-8870-7649

ARZAMASOV K.M.,

DSc, Moscow Center for Diagnostics and Telemedicine, Moscow, Russia,
e-mail: ArzamasovKM@zdrav.mos.ru, ORCID: 0000-0001-7786-0349

VLADZYMYRSKYY A.V.,

DSc, Moscow Center for Diagnostics and Telemedicine, Moscow, Russia,
e-mail: VladzimirskijAV@zdrav.mos.ru, ORCID: 0000-0002-2990-7736

VASILEV Y.A.,

DSc, Moscow Center for Diagnostics and Telemedicine, Moscow, Russia,
e-mail: VasilevYA1@zdrav.mos.ru, ORCID: 0000-0002-5283-5961

EVALUATION OF THE CLINICAL APPLICABILITY OF ARTIFICIAL INTELLIGENCE SYSTEMS FOR DETECTING INFILTRATIVE CHANGES ON CHEST X-RAYS IN INPATIENT CARE

DOI: 10.25881/18110193_2026_2_36

Abstract. Aim: to investigate the capabilities of artificial intelligence systems (AIS) in an inpatient setting, specifically in identifying signs of infiltration in chest radiographs (RG OGK)

Materials and methods. The initial data were chest X-rays (CXRs) in a frontal projection of 1,537 inpatients aged 18 to 102 years, obtained from the Unified Radiological Information System of the Unified Medical Information and Analytical Service (ERIS EMIAS). The presence of infiltration on the chest X-ray was determined by a radiologist (Radiologist 1) and an expert radiologist (Radiologist 2) with more than 5 years of experience. Radiologist 2 annotated the studies in two stages: at the first stage, only the chest X-rays were provided; at the second stage, in addition to the radiological data, access to information about the patient's condition was provided.

Results. The accuracy rates for patients with a shifting state were 0.755 for AIS, 0.756 for Radiologist 1, and 0.856 for Radiologist 2, respectively. No statistically significant differences in accuracy were found between AIS and Radiologist 1 ($p > 0.05$). When considering the scenario of using the AIS as a second opinion the accuracy for patients with a shifting state was 0.833, which is statistically significantly higher than the accuracy of Radiologist 1 and the AIS ($p < 0.05$).

Conclusion. AIS can be used as a second opinion when detecting signs of infiltration based on RG OGK data in a hospital setting. However, since the diagnostic accuracy of the AIS is still insufficient for practical use, further research is needed to train algorithms using both visual data and electronic medical record data.

Keywords: chest x-ray, infiltration, artificial intelligence system, clinical applicability.

For citation: Erizhokov R.A., Sakharova P.A., Kirsanov V.V., Khoruzhaya A.N., Varyukhina M.D., Arzamasov K.M., Vladzymyrskyy A.V., Vasilev Y.A. Evaluation of the clinical applicability of artificial intelligence systems for detecting infiltrative changes on chest X-rays in inpatient care. Medical doctor and information technology. 2026; 2: 36-49. DOI: 10.25881/18110193_2026_2_36.

ВВЕДЕНИЕ

Лидирующую позицию среди медицинских исследований, анализ которых автоматизируется, занимают рентгенография органов грудной клетки (РГ ОГК) и флюорография (ФЛГ): из всех лучевых исследований, обработанных медицинскими изделиями с искусственным интеллектом в России, на конец 2023 года 45% составили РГ ОГК и ФЛГ [1].

В настоящее время самым перспективным направлением является разработка систем искусственного интеллекта (СИИ), которые обладают высокой чувствительностью и позволяют не пропустить целевую патологию. Так, например, в исследовании представлены результаты модели, выполняющей автоматическую сортировку РГ ОГК [2]. В 99,95% случаев исследования были правильно отнесены СИИ к категории «норма».

Использование системы поддержки принятия врачебных решений (СППВР) позволяет врачу-рентгенологу снизить число пропущенных патологий на медицинских изображениях, однако так же приводит к росту числа ложноположительных результатов. Hwang E.J. и соавт. тестировали СИИ на выборке из 4208 РГ ОГК, изначально оцененных врачами как нормальные, с целью поиска пропущенных признаков патологии [3]. СИИ обнаружила патологию на 694 РГ, при этом из них только 103 действительно были с патологией.

Важно отметить, что задача для СИИ в диагностике патологий на РГ ОГК осложнена по нескольким причинам. В частности, образования разного происхождения на РГ ОГК могут иметь схожие визуальные характеристики (форму или интенсивность), что заставляет врачей интерпретировать рентгенологическую картину, опираясь на историю болезни и результаты клинических анализов, однако СИИ, как правило, обучаются только на данных лучевых исследований. При этом сужение фокуса внимания рентгенолога на основной патологии, которая послужила причиной направления на исследование, может приводить к пропуску другой патологии. Кроме того, ещё один фактор возможных ошибок — склонность повторять ошибки предыдущего рентгенолога, т.е. полагать, что, если коллега описал предыдущее исследование как нормальное, значит, так оно и есть [4].

Также особенности проектирования СИИ не позволяют им учитывать клиническую информацию. Поэтому использование СИИ для оценки состояния пациента в динамике может быть осложнено тем, что СИИ воспринимает остаточные проявления перенесенного заболевания как признаки вновь выявленной патологии, в связи с чем возрастает число ложноположительных результатов [5].

С другой стороны, недостаток клинической информации может привести к пропуску патологии. Так, Hofmeister и соавт. протестировали СИИ на двух когортах госпитализированных пациентов с подозрением на пневмонию [6]. В одной из когорт у врача с данными компьютерной томографии (КТ) AUROC выявления пневмонии составила 0,92 по сравнению с СИИ (0,74) и врачом-рентгенологом (0,72). Также авторы отмечают, что эффективность СИИ в диагностике заболеваний ОГК может быть переоценена, особенно если оценивать работу СИИ на реальных данных. Точность и чувствительность СИИ составили в среднем 0,72 и 0,73, соответственно, по обоим когортам. Авторы объясняют более низкую, чем ожидалось, точность тем, что около 60% РГ были выполнены пациентам в лежачем положении.

Таким образом, целью нашего исследования является оценка диагностической точности СИИ в выявлении признаков инфильтрации по данным РГ ОГК среди пациентов стационарных отделений, наблюдаемых в динамике.

МАТЕРИАЛЫ И МЕТОДЫ

Расчет и обоснование размера выборки

Материалом для ретроспективного исследования послужили доступные в Единой Радиологической Информационной Системе Единого Медицинского Информационно-Аналитического Сервиса (ЕРИС ЕМИАС) г. Москвы результаты РГ ОГК 1537 пациентов в возрасте от 18 до 102 лет, находившихся на стационарном лечении в лечебных учреждениях г. Москвы в 2024 году.

Исследования были выбраны в ЕРИС ЕМИАС случайным образом и были предварительно анонимизированы.

Критериями включения пациентов в исследование были:

- возраст пациентов не менее 18 лет;
- наличие трех последовательных РГ ОГК, выполненных в пределах одного месяца

(разница между первым и последним исследованием должна составлять не более 30 дней);

- наличие результатов работы одного из трех СИИ, предназначенных для выявления нескольких патологий РГ ОГК: «Третье Мнение», «Цельс» и «АрхиМед Aivory Chest X».

Таким образом, исходная выборка составила 4611 исследований с описаниями и заключениями Рентгенолога 1 (врач-рентгенолог без требований к опыту работы, работающий в условиях круглосуточного стационара клинической больницы) и СИИ. Все исследования были выполнены пациентам только в прямой проекции.

Все сервисы участвовали в Московском Эксперименте по применению компьютерного зрения в лучевой диагностике [7, 8]. Поскольку условием участия СИИ в данном эксперименте было соответствие единым базовым функциональным и диагностическим требованиям к СИИ для выявления признаков патологии на РГ ОГК [9], в рамках данной задачи было принято решение считать три сервиса взаимозаменяемыми.

Согласно разметке Рентгенолога 1, соотношение численности классов «патология» и «норма» составило 851:3760. В качестве ожидаемого значения AUROC рассматривалось значение 0,74, полученное Hofmeister и соавт., поскольку данная работа также содержит оценку СИИ в стационаре. В качестве значения AUROC нулевой гипотезы рассматривалось значение 0,81 согласно требованиям к СИИ, участвующим в Московском Эксперименте по применению компьютерного зрения в лучевой диагностике. Расчет размера выборки производился с помощью программного обеспечения MedCalc при установленном значении ошибки первого рода 0,05 и мощности теста 80% [10]. Согласно расчету, минимальное число случаев нормы и патологии, которое необходимо включить в выборку, составило 562 и 127, соответственно, то есть выборка отвечала требованиям минимального размера для получения статистически значимых результатов.

Методы исследования

Исследования пациентов были разбиты на три группы:

- исследования пациентов, у которых по данным всех трех РГ Рентгенологом 1 не было выявлено признаков инфильтрации

легочной ткани (только «норма») (67,1% исследований);

- исследования пациентов, у которых по данным всех трех РГ Рентгенологом 1 были выявлены признаки инфильтрации легочной ткани (только «патология») (6,3% исследований);
- остальные исследования (пациенты с переходным состоянием) (26,6% исследований).

Схема отбора исследований показана на рисунке 1.

Разметка исследований

Разметка исследований подразумевала отнесение каждого исследования к категории «норма» или «патология». Первая разметка исследований была выполнена Рентгенологом 1 в процессе описания исследований в стационарах г. Москвы.

Для следующего этапа разметки привлекался Рентгенолог 2 (врач-рентгенолог с опытом работы более 5 лет), который оценивал только данные РГ ОГК. У Рентгенолога 2 не было доступа к ЭМК, а также к предыдущим и последующим исследованиям пациента. На оценку к Рентгенологу 2 были отправлены исследования только тех пациентов, которые были отнесены к группе с переходным состоянием и у которых хотя бы одно из трех исследований было интерпретировано по-разному СИИ и Рентгенологом 1. В остальных случаях, когда ответ СИИ и Рентгенолога 1 не противоречили друг другу, за истинное значение был принят консенсусный результат.

На третьем этапе разметки Рентгенологу 2 в дополнение к РГ предоставлялся доступ к ЭМК пациента. Поскольку при данном типе разметки были учтены не только визуальные характеристики РГ, но и результаты предыдущих инструментальных исследований и клинические данные, в дальнейшем при исследовании групп пациентов с переходным состоянием она использовалась в качестве эталонной разметки.

На рисунке 2 показана схема исследования.

Дополнительно в рамках исследования была проведена оценка возможностей СИИ при её использовании в качестве второго мнения для Рентгенолога 1. Для этого применялся метод консенсусной оценки: заключение о наличии патологии выносилось в том случае, если хотя бы один из участников (Рентгенолог 1 или СИИ) выявлял патологию. Данный подход позволил

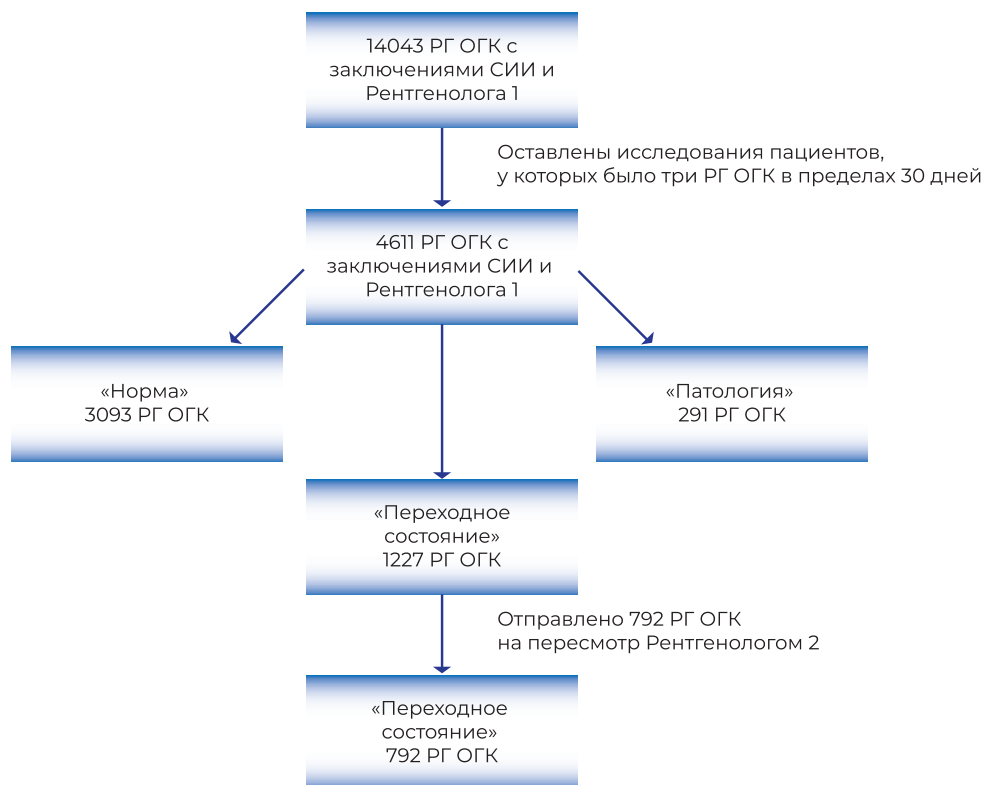


Рисунок 1 – Схема отбора исследований РГ ОГК.

нам оценить максимально достижимую чувствительность при совместном анализе РГ ОГК.

Методы оценки разметки

Поскольку не для всех исследований от СИИ были представлены вероятности наличия патологии (инфильтрации), было принято решение перевести предсказания в бинарный вид «патология/норма» с порогом $\geq 0,5$.

В качестве метрик оценки эффективности СИИ использовались точность (Acc), чувствительность (Se) и специфичность (Sp).

Для определения степени согласованности между разметками рассчитывался коэффициент Каппа Козна. Его значения интерпретировались следующим образом: незначимое согласие (до 0,4), умеренное согласие (от 0,41 до 0,6), существенное согласие (от 0,61 до 0,8) и полное согласие (0,81 и выше) разметок [11].

Статистические методы

Статистическая значимость различий значений метрик классификации между разметками

исследовалась с помощью теста МакНемара [12]. Для сравнения долей признака и распределения категориальных переменных в двух выборках использовался тест хи-квадрат.

Доверительные интервалы метрик рассчитывались с использованием бутстрапа, где за нижнюю и верхнюю границу были взяты 2,5 и 97,5 процентиля бутстрапированных выборок соответственно.

Статистические гипотезы проверялись при уровне доверительной вероятности (р-значение) 0,05.

РЕЗУЛЬТАТЫ

Характеристика исходной выборки

Характеристика выборки приведена в таблице 1. Возраст представлен как среднее (стандартное отклонение).

Сравнение диагностической точности СИИ с оценкой Рентгенолога 1

На основании разметки РГ ОГК Рентгенологом 1 были рассчитаны метрики классификации

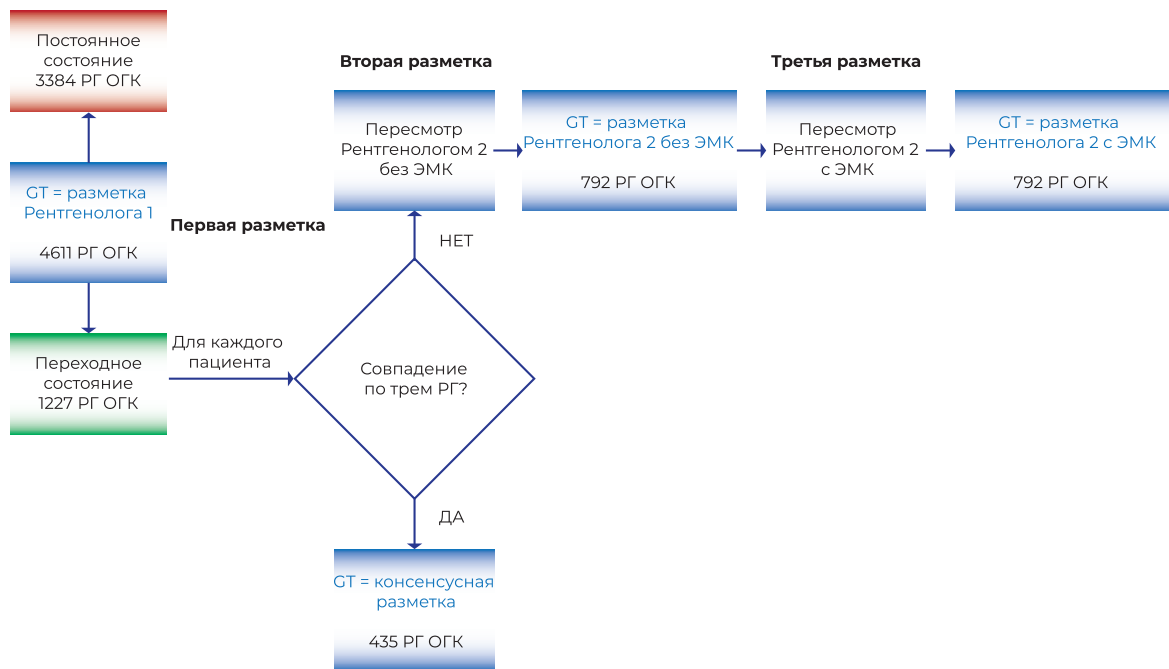


Рисунок 2 – Схема исследования.

Таблица 1 – Характеристика пациентов

Состояние	Пол	Число пациентов (n, %)	Возраст (лет)
Переходное	Мужской	214 (52,3%)	60,3 (16,9)
	Женский	195 (47,7%)	74,2 (15,1)
	Всего	409 (100%)	66,9 (17,5)
Норма	Мужской	518 (50,2%)	61,4 (16,2)
	Женский	513 (49,8%)	70,7 (16,0)
	Всего	1031 (100%)	66,0 (16,8)
Патология	Мужской	60 (61,9%)	61,1 (19,8)
	Женский	37 (38,1%)	72,2 (15,3)
	Всего	97 (100%)	65,4 (18,9)

Таблица 2 – Метрики СИИ в группах пациентов с нормой, патологией и переходным состоянием: Ground Truth (GT) = разметка Рентгенолога 1

Состояние	Acc	Se	Sp
Норма	0,808 (0,794-0,822)	–	0,808 (0,794-0,822)
Патология	0,773 (0,725-0,818)	0,772 (0,725-0,821)	–
Переходное	0,707 (0,682-0,732)	0,708 (0,674-0,744)	0,708 (0,672-0,742)

СИИ для всех трех групп. Результаты представлены в таблице 2 значениями метрик и доверительными интервалами.

Различия в точности диагностики СИИ патологии у пациентов с постоянным (нормой и патологией) и переходным состоянием были

Таблица 3 – Оценка статистической значимости доли правильных результатов классификации в зависимости от состояния пациента с помощью теста хи-квадрат: GT = разметка Рентгенолога 1

Состояние	Acc	p-значение
Постоянное	0,805	< 0,0001
Переходное	0,707	

изучены с помощью теста хи-квадрат. Результаты представлены в таблице 3.

Сравнение диагностической точности СИИ с оценкой Рентгенолога 2 без доступа к ЭМК

Чтобы сравнить диагностическую точность врача и СИИ в сходных условиях (у Рентгенолога 2 и СИИ есть доступ только к РГ пациента), мы решили пересмотреть РГ пациентов с переходным состоянием без обращения к другим исследованиям пациента и сравнить разметки, полученные СИИ и Рентгенологом 2.

В таблице 4 приведены метрики СИИ и Рентгенолога 1, рассчитанные в группе с переходным состоянием относительно разметки Рентгенолога 2 без ЭМК.

Для установления статистически значимых различий между разметкой СИИ и разметкой Рентгенолога 1 относительно разметки Рентгенолога 2 без ЭМК был проведен тест МакНемара, результат которого представлен таблицей сопряженности, р-значением и значением отношения шансов (OR) с 95%-ым доверительным интервалом.

Сравнение диагностической точности СИИ с оценкой Рентгенолога 2 с ЭМК

Следующим этапом Рентгенологу 2 было предложено повторно разметить те же РГ, но уже имея возможность использовать ЭМК пациента для получения клинической информации. Данная разметка в дальнейшем была использована

Таблица 4 – Метрики СИИ и Рентгенолога 1 в группе пациентов с переходным состоянием после второй разметки: GT = разметка Рентгенолога 2 без ЭМК

Разметка	Acc	Se	Sp
СИИ	0,817 (0,796-0,838)	0,750 (0,720-0,783)	0,917 (0,892-0,940)
Рентгенолог 1	0,758 (0,735-0,781)	0,678 (0,645-0,712)	0,874 (0,844-0,901)

Таблица 5 – Результаты теста МакНемара для разметок СИИ и Рентгенолога 1: GT = разметка Рентгенолога 2 без ЭМК

СИИ	Рентгенолог 1		Всего	p-значение	OR
	True	False			
True	787	216	1003	0,0001	1,510 (1,217-1,879)
False	143	81	224		
Всего	930	297	1227		

Таблица 6 – Метрики разметок СИИ, Рентгенолога 1 и Рентгенолога 2 без ЭМК в группе пациентов с переходным состоянием: GT = разметка Рентгенолога 2 с ЭМК

Разметка	Acc	Se	Sp
Рентгенолог 1	0,756 (0,733-0,780)	0,674 (0,638-0,707)	0,881 (0,852-0,909)
Рентгенолог 2 без ЭМК	0,856 (0,836-0,874)	0,872 (0,848-0,897)	0,830 (0,794-0,864)
СИИ	0,755 (0,731-0,778)	0,694 (0,660-0,730)	0,849 (0,817-0,878)

в качестве референсной. В таблице 6 приведены метрики диагностической точности для всех разметок относительно разметки Рентгенолога 2 с ЭМК.

На рисунке 3 показана матрица коэффициентов Каппа Коэна между разметками исследований пациентов из группы с переходным состоянием.

С помощью теста хи-квадрат были рассчитаны р-значения. Все коэффициенты Каппа Коэна являются статистически значимыми ($p < 0,0001$).

В таблицах 7–9 представлены результаты теста МакНемара, выраженные р-значениями.

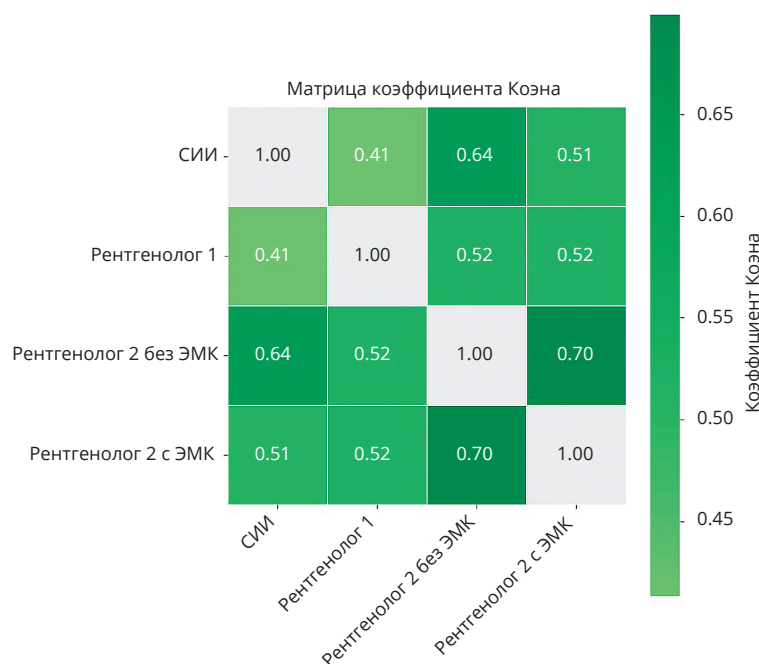


Рисунок 3 – Матрица коэффициентов Каппа Коэна между разметками.

Таблица 7 – Результаты теста МакНемара для разметок СИИ и Рентгенолога 1: GT = разметка Рентгенолога 2 с ЭМК

СИИ	Рентгенолог 1		Всего	р-значение	OR
	True	False			
True	747	179	926	1	0,994 (0,804-1,230)
False	180	121	301		
Всего	927	300	1227		

Таблица 8 – Результаты теста МакНемара для разметок СИИ и Рентгенолога 2 без ЭМК: GT = разметка Рентгенолога 2 с ЭМК

СИИ	Рентгенолог 2 без ЭМК		Всего	р-значение	OR
	True	False			
True	876	50	926	<0,0001	3,480 (2,528-4,867)
False	174	127	301		
Всего	1050	177	1227		

**Таблица 9 – Результаты теста МакНемара для разметок СИИ и Рентгенолога 1:
GT = разметка Рентгенолога 2 без ЭМК**

Рентгенолог 1	Рентгенолог 2 без ЭМК		Всего	p-значение	OR
	True	False			
True	840	87	927	0,0001	2,414 (1,871-3,136)
False	210	90	300		
Всего	1050	177	1227		

**Таблица 10 – Метрики совместной разметки и доля ЛО в группе пациентов
с переходным состоянием: GT = разметка Рентгенолога 2 с ЭМК**

Разметка	Acc	Se	Sp	ЛО (%)
Совместная разметка	0,833 (0,812-0,855)	0,868 (0,844-0,892)	0,778 (0,741-0,815)	8,0
СИИ	0,755 (0,731-0,778)	0,694 (0,660-0,730)	0,849 (0,817-0,878)	18,6
Рентгенолог 1	0,756 (0,733-0,780)	0,674 (0,638-0,707)	0,881 (0,852-0,909)	19,8

**Таблица 11 – Результаты теста МакНемара для разметки СИИ и совместной разметки:
GT = разметка Рентгенолога 2 с ЭМК**

СИИ	Совместная разметка		Всего	p-значение	OR
	True	False			
True	892	34	926	<0,0001	3,824 (2,604-5,754)
False	130	171	301		
Всего	1022	177	1227		

**Таблица 12 – Результаты теста МакНемара для разметки Рентгенолога 1
и совместной разметки: GT = разметка Рентгенолога 2 с ЭМК**

СИИ	Совместная разметка		Всего	p-значение	OR
	True	False			
True	877	50	927	<0,0001	2,900 (2,089-4,084)
False	145	155	300		
Всего	1022	205	1227		

**Таблица 13 – Результаты теста МакНемара для разметки Рентгенолога 2 без ЭМК
и совместной разметки: GT = разметка Рентгенолога 2 с ЭМК**

СИИ	Совместная разметка		Всего	p-значение	OR
	True	False			
True	957	93	1050	0,031	0,699 (0,501-0,970)
False	65	112	177		
Всего	1022	205	1227		

С целью изучения возможности уменьшить число пропущенных случаев инфильтрации был рассмотрен сценарий использования СИИ в качестве второго мнения для Рентгенолога 1. Метрики классификации и доля

ложноотрицательных результатов (ЛО) совместной разметки, СИИ и Рентгенолога 1 представлены в таблице 10.

В таблицах 11–13 представлены результаты теста МакНемара, выраженные p-значениями.

ОБСУЖДЕНИЕ

Как правило, ошибки работы СИИ обусловлены рядом причин. Так, в случае оценки РГ, выполненных только в прямой проекции, даже врач-рентгенолог может пропустить патологию, т.к. в ряде случаев вывод о наличии признаков инфильтрации можно будет сделать только при наличии боковой проекции. Исследования нашей выборки были выполнены только в прямой проекции. Ценную помощь в выявлении инфильтрации могут также оказать результаты дополнительного обследования: в частности, КТ, однако в нашем исследовании в большинстве случаев она не была выполнена пациентам. Таким образом,

данные обстоятельства обуславливают большое количество пропусков патологии как в случае оценки Рентгенологом 2 без доступа к клинической информации из ЭМК, так и СИИ.

На рисунке 4а приведена РГ, на которой Рентгенолог 1 и СИИ не выявили патологию из-за низкой информативности РГ в прямой проекции. При этом на рисунке 4б приведены данные КТ ОГК пациента, по которым можно однозначно установить наличие признаков инфильтрации.

Небольшая часть ложноотрицательных предсказаний СИИ была связана с тем, что они локализовывали область инфильтрации и гидроторакса, но выявляли только гидроторакс. На рисунке 5

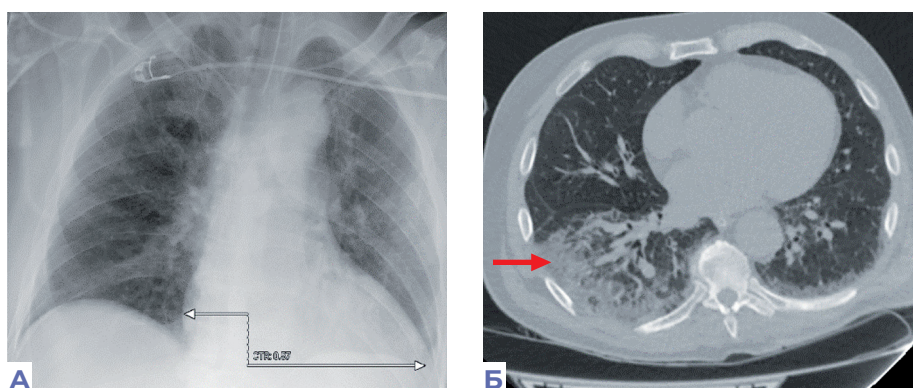


Рисунок 4 – Инфильтрация, пропущенная СИИ и Рентгенологом 1.
На рисунке а) показана РГ после обработки СИИ. На рисунке б) приведен результат КТ-исследования, с помощью которого Рентгенологом 2 была выявлена патология (зона инфильтрации указана стрелкой).

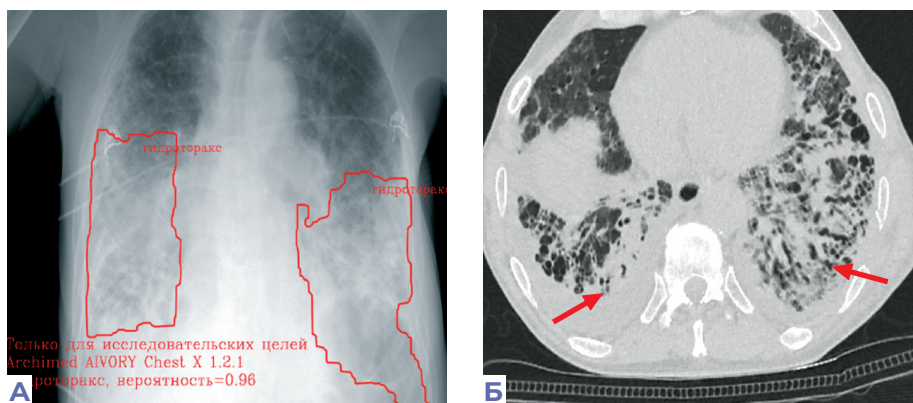


Рисунок 5 – Ложноотрицательный случай обработки РГ.
На рисунке а) СИИ верно локализовал патологию и верно выявил гидроторакс, но пропустил инфильтрацию. По данным КТ на рисунке б), зоны соответствуют инфильтрации (указаны стрелкой).

приведена РГ, на которой СИИ верно выявил гидроторакс, но пропустил инфильтрацию.

Настоящее исследование показало, что применение СИИ как самостоятельной системы выявления признаков инфильтрации в стационарных отделениях в настоящее время нецелесообразно из-за большой вероятности пропуска целевой патологии. Согласно таблице 11, среди пациентов с переходным состоянием чувствительность СИИ составила 0,678, что является низким показателем для клинического применения и демонстрирует необходимость доработки СИИ для данного сценария использования.

Разметка Рентгенолога 1 наиболее часто согласуется с разметкой СИИ, если состояние пациента не изменяется. Согласно таблице 3, точность СИИ для пациентов с постоянным и переходным состоянием относительно разметки Рентгенолога 1 составила 0,805 и 0,707, соответственно. Согласно результатам теста хи-квадрат, приведенным в таблице 3, это различие является статистически значимым ($p < 0,0001$).

Наибольшую согласованность сервис показал с Рентгенологом 2, который выносит заключение о состоянии пациента непосредственно по РГ ОГК, не имея при этом анамнеза заболевания и предыдущих инструментальных исследований. Это объясняется тем, что в данном случае как врач, так и СИИ, опираются только на данные РГ. Согласно метрикам таблицы 4 и результатам теста МакНемара, приведенным в таблице 5, точность сервиса (0,817) статистически значимо выше, чем точность Рентгенолога 1 (0,758) относительно разметки Рентгенолога 2 без ЭМК. Согласно рисунку 3, коэффициент Каппа Козна разметки СИИ с разметкой Рентгенолога 2 без ЭМК составил 0,64, что является показателем существенного согласия.

Понимание истории возникновения и развития заболевания, а также доступ к результатам дополнительных исследований, оказывает значительное влияние на итоговый результат. При пересмотре РГ с возможностью доступа к ЭМК Рентгенолог 2 изменил свое мнение в 22,4% случаев.

Согласно таблице 7, относительно эталонной разметки статистически значимых различий между диагностической точностью разметок Рентгенолога 1 и СИИ не выявлено ($p = 1$).

Согласно таблицам 8 и 9, разметка Рентгенолога 2 без ЭМК статистически значимо превосходит разметку Рентгенолога 1 и СИИ по точности ($p < 0,0001$).

Кроме того, нами был протестирован сценарий использования СИИ в качестве СППВР. Согласно таблицам 10–13, точность совместной разметки составила 0,833, что статистически значимо превышает точность Рентгенолога 1 (0,756) и СИИ (0,755), работающих самостоятельно. Совместное принятие решения о наличии патологии помогло снизить долю ЛО с 19,8% у Рентгенолога 1 и 18,6% у СИИ до 8,0%.

Наше исследование имеет ряд ограничений. В частности, результаты дополнительных КТ-исследований были доступны лишь в 80% случаев, что могло повлиять на качество референсной разметки. Ещё одно ограничение состоит в том, что у нас нет информации о положении, в котором пациентам были выполнены исследования (стоя или лежа). Также, ввиду большого количества РГ для пересмотра, Рентгенологом 2 были отобраны только исследования пациентов с переходным состоянием, что сократило исходную выборку и не позволило обобщить результаты для пациентов, состояние которых не менялось.

ЗАКЛЮЧЕНИЕ

Согласно проведенному исследованию, мнения врача-рентгенолога и СИИ наиболее часто расходятся при обработке РГ пациентов, у которых при оценке инфильтрации легочной ткани в динамике отмечалось появление или исчезновение признаков данной патологии на РГ. Таким образом, клиническое применение СИИ, в основе которых лежат алгоритмы, анализирующие РГ без учета клинического контекста или информации о результатах дополнительных обследований, для оценки патологических процессов в динамике в настоящее время нецелесообразно ввиду низкой диагностической точности.

Наибольшую согласованность с врачом СИИ демонстрируют в случаях, когда врач и сервис имеют в распоряжении одинаковые исходные данные, то есть исключительно РГ, описывающие текущее состояние пациента. При добавлении клинического контекста мнение врача часто меняется, однако сервис не может воспользоваться новой информацией, что приводит к снижению качества его работы.

В настоящем исследовании показано, что использование СИИ в качестве СППВР может существенно снизить долю ЛО и повысить диагностическую точность при оценке инфильтрации легочной ткани на РГ ОГК. Требуется продолжение исследования возможностей СИИ, в частности с использованием иного подхода к

проектированию СИИ, позволяющего им учитывать дополнительную клиническую информацию.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

ЛИТЕРАТУРА/REFERENCES

1. Gusev AV, Artemova OR, Vasiliev YuA, et al. Integration of AI-based software as a medical device into Russian healthcare system: results of 2023. National Health Care (Russia). 2024; 5: 17-24. doi: 10.47093/2713-069X.2024.5.2.17-24.
2. Vasilev YA, Sychev DA, Bazhin AV, et al. Autonomous artificial intelligence for sorting results of preventive radiological examinations of chest organs: medical and economic efficiency. Digital Diagnostics. 2025; 6: 5-22. doi: 10.17816/DD641703.
3. Hwang EJ, Park J, Hong W, et al. Artificial intelligence system for identification of false-negative interpretations in chest radiographs. European Radiology. 2022; 32: 4468-4478. doi: 10.1007/s00330-022-08593-x.
4. De Groot PM, Carter BW, Abbott GF, et al. Pitfalls in Chest Radiographic Interpretation: Blind Spots. Seminars in Roentgenology. 2015; 50: 197-209. doi: 10.1053/j.ro.2015.01.008.
5. Пилус П.С., Дрокин И.С., Баженова Д.А., и др. Оценка перспектив использования технологий искусственного интеллекта для анализа КТ-изображений органов грудной клетки с целью выявления признаков злокачественных новообразований в легких. Медицинская визуализация. 2023; 27(2): 138-146. [Pilyus PS, Drokin IS, Bazhenova DA, et al. Otsenka perspektiv ispol'zovaniya tekhnologiy iskusstvennogo intellekta dlya analiza KT-izobrazheniy organov grudnoy kletki s tsel'yu vyyavleniya priznakov zlokachestvennykh novoobrazovaniy v legkikh. Meditsinskaya vizualizatsiya. 2023; 27(2): 138-146. (In Russ.)]. doi: 10.24835/1607-0763-2023-2-138-146.
6. Hofmeister J, Garin N, Montet X, et al. Validating the accuracy of deep learning for the diagnosis of pneumonia on chest x-ray against a robust multimodal reference diagnosis: a post hoc analysis of two prospective studies. European Radiology Experimental. 2024; 8: 20. doi: 10.1186/s41747-023-00416-y.
7. Сервисы искусственного интеллекта в лучевой диагностике.. Доступно по: <https://mosmed.ai/>. Ссылка активна на 05.09.2025. [Servisy iskusstvennogo intellekta v luchevoi diagnostike. Available at: <https://mosmed.ai/>. Accessed: 05.09.2025. (In Russ.)]
8. Искусственный интеллект в лучевой диагностике: Per Aspera Ad Astra / Под ред. Ю.А. Васильева и А.В. Владзимирского. М.: Издательские решения; 2025. 491 с. [Vasil'ev YuA, Vladzimyrskiy AV, editors. Iskustvennyy intellekt v luchevoi diagnostike: Per Aspera Ad Astra. Moscow: Izdatel'skie resheniya; 2025: 491. (In Russ.)]
9. Морозов С.П., Абуладзе Л.Р., Андрейченко А.Е. и др. Базовые рекомендации к работе сервисов искусственного интеллекта для лучевой диагностики: методические рекомендации. Серия «Лучшие практики лучевой и инструментальной диагностики». Вып. 119. — М.: ГБУЗ «НПКЦ ДиТ ДЗМ», 2022. 68 с. [Morozov SP, Abuladze LR, Andreychenko AE, et al. Bazovye rekomendatsii k rabote servisov iskusstvennogo intellekta dlya luchevoi diagnostiki: metodicheskiye rekomendatsii. Seriya «Luchshiye praktiki luchevoy i instrumental'noy diagnostiki». Vyp. 119. Moscow: GBUZ «NPKTs DiT DZM»; 2022: 68. (In Russ.)]
10. Schoonjans F. MedCalc's Sample size for Area under ROC curve. Available at: <https://www.medcalc.org/calculator/sample-size-area-under-roc-curve.php>. Accessed April 05.09.2025.
11. Andropova PL, Gavrilov PV, Kazantseva IP, et al. Interexpert Agreement Between Emergency Neuroradiologists with Different Levels of Experience in the Rating of ASPECTS. Radiology — practice. 2022; 5: 10-25. doi: 10.52560/2713-0118-2022-5-10-25.
12. Wetcher-Hendricks D. An Expanded McNemar Approach for Evaluation of Longitudinal Data 2025. Preprints 2025. doi: 10.20944/preprints202503.2111.v1.