

Опыт использования технологий искусственного интеллекта в лабораторной медицине, их эффективность и сценарии применения: систематический обзор

Ю.А. Васильев¹, О.Г. Нанова¹, А.В. Владзимирский¹, А.С. Гольдберг², И.А. Блохин¹, Р.В. Решетников¹

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

² Российская медицинская академия непрерывного профессионального образования, Москва, Россия

АННОТАЦИЯ

Обоснование. Область лабораторной медицины в связи с нарастающим потоком данных нуждается в автоматизации и стандартизации рутинных процессов для разгрузки медицинских работников и высвобождения их времени на решение более специализированных задач. Модели машинного обучения и искусственные нейронные сети помогают распознавать изображения и анализировать большие массивы данных, что потенциально позволяет внедрить их в работу лабораторий для решения рутинных задач.

Цель — проанализировать мировую литературу в области применения технологий искусственного интеллекта в лабораторной медицине, оценить их возможности в отношении решения существующих задач, а также выявить возможные проблемы, затрудняющие внедрение искусственного интеллекта в лабораторные процессы.

Материалы и методы. Поиск работ проводили в поисковой системе PubMed, на сайтах производителей готовых лабораторных решений и в списках литературы других обзоров. Кроме того, использовали программу для управления библиографической информацией Mendeley. Временной интервал — 2019–2024 гг. Из найденных публикаций извлекали библиометрические данные, область исследований, основные методические характеристики, значения диагностической эффективности искусственного интеллекта и медицинских работников, число и опыт задействованных медицинских специалистов, подтверждённые результаты его внедрения. Качество исследований оценивали с помощью модифицированного опросника QUADAS-CAD.

Результаты. Всего в обзор включили 23 публикации, в которых представлены исследования на преаналитическом, аналитическом и постаналитическом этапах лабораторного анализа — 1, 19 и 3 соответственно. Большинство исследований проведено в области цитологии и микробиологии — 48 и 35% соответственно. Искусственный интеллект демонстрирует высокую эффективность в отношении решения задач на всех этапах лабораторного процесса. Кроме того, его диагностическая точность сопоставима с уровнем медицинских работников, а скорость принятия решений значительно выше. Тем не менее во всех работах наблюдали риск систематической ошибки, что связано с несбалансированностью выборок, отсутствием внешней валидации данных, а также точного их описания и методов анализа.

Заключение. Искусственный интеллект обладает высоким потенциалом в отношении диагностической точности и скорости работы, что делает его перспективным инструментом для внедрения в лабораторную практику и автоматизации рутинных процессов. Однако для этого необходимо стандартизировать методики исследования искусственного интеллекта для снижения риска систематических ошибок, установить референсные значения для лабораторий с целью обеспечения воспроизводимости и обобщаемости результатов, повысить осведомлённость медицинских работников и пациентов о принципах его работы для преодоления предубеждений, а также разработать надёжные механизмы защиты персональных данных при использовании искусственного интеллекта.

Ключевые слова: искусственный интеллект; машинное обучение; компьютерное зрение; лабораторная медицина; патоморфология.

КАК ЦИТИРОВАТЬ:

Васильев Ю.А., Нанова О.Г., Владзимирский А.В., Гольдберг А.С., Блохин И.А., Решетников Р.В.
Опыт использования технологий искусственного интеллекта в лабораторной медицине, их
эффективность и сценарии применения: систематический обзор // Digital Diagnostics. 2025. Т. 6, № 2.
С. XX–XX. DOI: 10.17816/DD635349 EDN: BXDWFO

Рукопись получена: 23.08.2024

Рукопись одобрена: 21.11.2024

Опубликована online: 30.05.2025

Статья доступна по лицензии CC BY-NC-ND 4.0 International

© Эко-Вектор, 2025

Experience with the use of artificial intelligence technologies in laboratory medicine, their efficiency and application scenarios: a systematic review

Yuriy A. Vasilev¹, Olga G. Nanova¹, Anton V. Vladzimirskyy¹, Arcadiy S. Goldberg², Ivan A. Blokhin¹,
Roman V. Reshetnikov¹

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

² The Russian Medical Academy of Continuous Professional Education, Moscow, Russia

ABSTRACT

BACKGROUND: Laboratory medicine due to the increasing dataflow needs in automatization and standardization of routine processes for medical workers relieving and freeing up their time for solving more specialized tasks. Rapidly developing artificial intelligence (AI) methods (that includes both machine learning methods and neural networks) designed for image recognition and big data analysis could be included in laboratory workflow for routine tasks solution.

AIM: to outline the scope of ai applications and assess the ability of ai to solve tasks in laboratory medicine. also, we aimed to reveal challenges that prevent the ai implementation in the laboratory workflow.

MATERIALS AND METHODS: PubMed and Mendeley databases, websites of commercial AI solutions manufacturers, and references of others reviews was searched for studies. The time interval was 2019-2024. The extracted data included bibliometric data, research area, main methodological characteristics, AI and medical workers diagnostic accuracy estimates, number and experience of the involved medical workers, and the confirmed results of AI implementation. The quality of the studies was assessed using the QUADAS-CAD questionnaire that was adapted for AI.

RESULTS: Twenty-tree studies were included in the review. Among them, one study was performed at pre-analytical phase, 19 studies were performed at analytical phase, and three – at post-analytical phase. Nearly half (48%) studies of analytical phase were performed in cytology. Next, 35% of studies were performed in microbiology. AI displays high efficiency for all phases of the laboratory workflow. The diagnostic accuracy of AI is comparable to diagnostic accuracy of medical workers. The speed of AI required for the task completing significantly exceeds the speed of medical workers. All studies had a risk of bias due to unbalanced samples, lack of external validation of the outcomes, and inaccurate description of data and methods.

CONCLUSIONS: AI has a good potential in terms of diagnostic accuracy and speed of work for implementation in laboratory practice and automation of routine laboratory processes. Set of issues must be resolved for AI implementation in laboratory practice. First, standardization of AI research methods to reduce the risk of bias. Second, standardization of laboratory reference values to generalize AI results. Third, increasing awareness among medical workers and patients about the principles of AI to overcome various kinds of prejudices, as well as developing systems to protect patients' personal data when using AI.

Keywords: artificial intelligence; machine learning; computer vision; laboratory medicine; pathology.

TO CITE THIS ARTICLE:

Vasilev YuA, Nanova OG, Vladzomyrskyy AV, Goldberg AS, Blokhin IA, Reshetnikov RV. Experience with the use of artificial intelligence technologies in laboratory medicine, their efficiency and application scenarios: systematic review *Digital Diagnostics*. 2025;6(2):XX–XX. DOI: 10.17816/DD635349
EDN: BXDWFO

Submitted: 23.08.2024

Accepted: 21.11.2024

Published online: 30.05.2025

The article can be used under the CC BY-NC-ND 4.0 International License
© Eco-Vector, 2025

ОБОСНОВАНИЕ

Лабораторная медицина — крайне нагруженная непрерывным потоком анализов и данных область. Классические стандартные протоколы лабораторной диагностики требуют от медицинских специалистов больших затрат времени и постоянной концентрации внимания [1, 2]. Именно в этой области остро стоит вопрос автоматизации процессов с целью снятия с медицинских работников нагрузки, связанной с рутинными процедурами, и переключения их внимания и усилий на более сложные специализированные задачи [3].

Технологии искусственного интеллекта (ИИ), включающие в себя в широком понимании как относительно простые методы машинного обучения, так и искусственные нейронные сети, быстро развиваются в последнее десятилетие и имеют потенциал стать удачным решением в отношении автоматизации рутинных процессов лабораторной медицины.

ЦЕЛЬ

Проанализировать мировую литературу в области применения технологий искусственного интеллекта в лабораторной медицине, оценить их возможности в отношении решения существующих задач, а также выявить возможные проблемы, затрудняющие внедрение искусственного интеллекта в лабораторные процессы.

МАТЕРИАЛЫ И МЕТОДЫ

СТРАТЕГИЯ ПОИСКА

Поиск публикаций выполняли с помощью поисковой системе PubMed [4], на сайтах производителей программ на основе технологий ИИ для лабораторной медицины, а также использовали метод «снежного кома» — поиск исследовательских работ в списках литературы публикаций.

Поисковая система PubMed. Временной интервал: с 2019 года по 01.06.2024. Поисковый запрос выглядел следующим образом: (((((((((((artificial intelligence) OR (deep learning)) OR (machine learning) OR (computer vision)) AND (clinical laboratory)) OR (laboratory medicine)) OR (clinical deployment)) AND (pathomorphology)) OR (digital pathology))) AND (computer-aided diagnosis)) OR (diagnostics)) AND (pathological image analysis)) NOT (radiomics).

В разделе фильтров выбраны следующие опции:

- для «Text Availability» с целью отбора статей с доступным полным текстом: Abstract, Full Text;
- для «Article Attribute» с целью отбора статей, которые содержат ссылки на ассоциированные клинические исследования либо наборы данных, подтверждающие достоверность полученных результатов: Associated Data;
- для «Article Type» с целью отбора наиболее убедительных доказательств: Clinical Trial, Randomized Controlled Trial.

Программа для управления библиографической информацией Mendeley. Временной интервал: 2019–2024 гг. Поиск выполняли с использованием ключевого слова: «artificial intelligence laboratory medicine».

Для поиска исследовательских работ с доказанным применением технологий ИИ в лабораторной практике изучены сайты производителей, упомянутые в отобранных публикациях, в частности разделы с научной литературой по использованию выпускаемого оборудования:

- Visiopharm¹;
- CyPath Lung²;
- EasyCell³;
- Copan⁴.

Временной интервал: 2023–2024 гг. Искли публикации, подтверждающие внедрение технологий ИИ в работу лабораторий.

Литературные обзоры, демонстрирующие опыт применения технологий ИИ в лабораторной медицине в период с 2023 по 2024 год: 2 публикации, отобранные из 12 найденных [2, 5].

КРИТЕРИИ ВКЛЮЧЕНИЯ

- публикации, имеющие как минимум английское резюме;
- статьи, опубликованные в рецензируемых научных журналах;
- препринты;
- статьи, опубликованные в конференционных сборниках.

КРИТЕРИИ ИСКЛЮЧЕНИЯ

- публикации, не связанные с лабораторной медициной и компьютерным зрением;
- публикации, не охватывающие вопросы медицины человека;
- обзоры литературы;
- конференционные тезисы.

Поисковая стратегия включала два этапа:

- первый — сначала анализировали названия и резюме всех найденных по поисковым запросам работ, затем отбирали соответствующие нашим задачам исследования;
- второй — анализировали полные тексты и их доступность из отобранного списка работ и составляли выборку для основного анализа обзора.

Публикации отбирал один эксперт, а финальный список включённых работ оценивали два эксперта. В качестве экспертов выступали научные сотрудники с опытом работы в медицинской информатике более 10 лет.

Наш обзор включает публикации, демонстрирующие применение технологий ИИ на всех трёх основных этапах лабораторного анализа:

- преаналитический;
- аналитический;
- постаналитический.

Поскольку их задачи и методы различны, каждый из этих этапов в систематическом обзоре мы рассмотрим отдельно.

ИЗВЛЕЧЕНИЕ ИНФОРМАЦИИ И ОЦЕНКА КАЧЕСТВА СТАТЬИ

Из полных текстов отобранных статей извлекали следующую информацию:

- библиометрические данные — имя первого автора, название статьи, год выхода, DOI (Digital Object Identifier — цифровой идентификатор объекта), название журнала, импакт-фактор журнала, страну выполнения исследования;
- направление исследований и их основные характеристики (объём выборки, дизайн исследования, наличие валидации на внешних данных, использованные лабораторные методы и модели ИИ);
- показатели диагностической эффективности ИИ [чувствительность, специфичность, площадь под кривой (AUC), точность, а также некоторые другие критерии эффективности, традиционные для лабораторной медицины];
- сравнения диагностической эффективности ИИ и медицинских специалистов;
- количество медицинских специалистов и уровень их квалификации;

¹ Visiopharm [Internet]. Denmark: Visiopharm®. 2001–2024. Режим доступа: <https://visiopharm.com/> Дата обращения: 12.10.2024.

² CyPath Lung [Internet]. San Antonio: CyPath® Lung. 2021–2024. Режим доступа: <https://www.cypathlung.com/> Дата обращения: 12.10.2024.

³ EasyCell [Internet]. Anyang-si: EasyCell Co., Ltd. 2020–2024. Режим доступа: <https://www.easycell.co/> Дата обращения: 12.10.2024.

⁴ Copan [Internet]. Murrieta: Copan Diagnostics Inc. 1999–2024. Режим доступа: <https://www.copanusa.com/> Дата обращения: 12.10.2024.

- оценка времени работы моделей машинного обучения и систем на основе ИИ (в том числе в сравнении с медицинскими специалистами);
- оценка экономической эффективности потенциального внедрения технологий ИИ;
- подтвержденные результаты внедрения ИИ.

Мы рассчитали средние показатели диагностической эффективности с использованием всех работ, где они обнаружены: медиану и 95% доверительный интервал (ДИ). Провели оценку качества отобранных публикаций с помощью модифицированного опросника QUADAS-CAD (Quality Assessment of Diagnostic Accuracy Studies Computer-Aided Detection) [6], разработанного для исследований с использованием ИИ.

Извлечение информации и оценка качества работ проведены одним экспертом. Результаты оценивали два эксперта с опытом работы в медицинской информатике более 10 лет.

РЕЗУЛЬТАТЫ

ПОИСК ЛИТЕРАТУРЫ И ОТБОР РАБОТ

На первом этапе найдено 2036 публикаций:

- поисковая система PubMed — 551;
- программа для управления библиографической информацией Mendeley — 1335;
- сайты — 17;
- литературные обзоры — 133.

На втором этапе отобрано 58 публикаций и исключено 1978. В основной анализ включили 23 публикации (приложение 1). Из систематического обзора исключили 35 публикаций (приложение 2). Среди основных причин исключения выделили следующие:

- отсутствие доступа к полному тексту;
- отсутствие факта применения технологий ИИ;
- техническая разработка метода без анализа медицинских данных.

Из включённых работ одна посвящена преаналитическому этапу лабораторного анализа, 19 — аналитическому, 3 — постаналитическому.

Основные характеристики исследований, представленных в публикациях, включённых в систематический обзор, продемонстрированы в приложении 3.

Характеристики выборок и используемые модели машинного обучения или готовые коммерческие решения на основе ИИ, описанные в исследованиях, продемонстрированы в приложении 4.

ПРЕАНАЛИТИЧЕСКИЙ ЭТАП

В исследовании, посвящённом преаналитическому этапу лабораторного анализа, рассматривают возможности применения технологий ИИ для выявления случаев неправильной маркировки пробирок [7]. Их эффективность в контроле качества сравнивают с результатами, достигаемыми медицинским персоналом лаборатории. Для такой проверки стандартно используют Δ-проверку (Delta-Check Methods) — сравнение последовательных во времени лабораторных результатов от одного и того же пациента и выявление ошибок при сильных их расхождениях. Рассматриваемое исследование выполнено ретроспективно, без внешней валидации, а также без тестирования в реальных лабораторных условиях. Ошибку маркировки симулировали на 50% уровне. Авторы разработали, обучили и протестировали в среде R 8 моделей машинного обучения (приложение 5). Для сравнения использовали 50 медицинских работников с разным опытом (приложение 6). Следует отметить, что не обнаружено достоверной связи ($p > 0,1$) между точностью контроля качества медицинских работников и их стажем. Все 8 моделей (0,865–0,921) превзошли медицинский персонал (0,778) по уровню точности решения задачи выявления неправильной маркировки (см. приложение 6). Модель нейронной сети была наиболее эффективной (0,921), модель простого дерева решений — наименее (0,865).

АНАЛИТИЧЕСКИЙ ЭТАП

Основной объём найденных по нашим запросам публикаций (19) демонстрирует исследования, выполненные на аналитическом этапе лабораторного анализа. Они выполнены в области цитологии, микробиологии, гистопатологии, паразитологии либо на пересечении этих областей.

В сфере цитологии задачи распределены следующим образом:

- диагностика опухолевых заболеваний — 2 исследования (анализ крови; анализ мокроты) [8, 9];

- диагностика гематологической патологии — 6 исследований (анализ крови — 5; мазок костного мозга — 1) [10–15].

В области микробиологии задачи работ распределены следующим образом:

- оценка устойчивости кишечной палочки (*Escherichia coli*) к 19 видам антибиотиков — одно исследование (общий анализ крови и мочи) [16];
- детекция стрептококков группы А (*Streptococcus*) — одно исследование (посев слизи из ротоглотки на агаре и кровяном агаре) [17];
- сегрегация патоген-положительных и -отрицательных посевов мочи с помощью подсчёта колоний без определения морфологии клеток — одно исследование (посев мочи на кровяном агаре и агаре МакКонки) [18];
- идентификация микобактерий в тканях человека с использованием окрашивания по Цилю–Нильсену — два исследования [19, 20];
- диагностика вагинитов с использованием вагинальных мазков — два исследования [21, 22].

В двух публикациях продемонстрировано совместное применение методов цитологии и микробиологии — анализировали осадок мочи для выявления потенциальных патогенов и диагностики инфекций мочевыводящих путей, при этом использовали детекцию и подсчёт элементов осадка мочи [23, 24].

В.А. Mathison и соавт. [25] выполняли детекцию кишечных простейших (*Protozoa*) в окрашенных трихромом мазках кала человека. В качестве объектов распознавания выступали:

- кишечная лямблия (*Giardia duodenalis*), её цисты и трофозоиты;
- кишечные амёбы (*Entamoeba hartmanni*, *Entamoeba sp. non-hartmanni*, или крупная *Entamoeba sp.*) и их трофозоиты;
- *Dientamoeba fragilis*;
- бластоцисты (*Blastocystis spp.*);
- хиломастикс (*Chilomastix mesnili*) и его трофозоиты;
- карликовые амёбы (*Endolimax nana*) и их трофозоиты;
- *Iodamoeba buetschlii* и её трофозоиты;
- эритроциты;
- лейкоциты.

Кроме того, модели обучили также для идентификации дрожжей как антикласса с целью предотвращения их ошибочной классификации. Для микроорганизмов *Entamoeba spp.*, *C. mesnili*, *E. nana* и *I. buetschlii* модель обучили распознавать только активную стадию трофозоитов. Тем не менее она не идентифицировала их цисты в связи с недостаточным числом обучающих образцов и слабой выраженностью их морфологии в используемом методе окрашивания.

Из 19 включённых в обзор исследований, проведённых на аналитическом этапе лабораторного анализа, были:

- многоцентровыми (использовали данные из нескольких лабораторий) — 8 (42%) [9, 12, 13, 15, 18, 21];
- одноцентровыми (использовали данные только из одной лаборатории) — 11 (55%) [7, 8, 10, 11, 14, 16, 17, 19, 22, 23, 25];
- ретроспективными — 17 (90%) [8–11, 13–21, 23–26];
- проспективными — 1 (5%) [22];
- ретроспективными с включённым проспективным тестом — 1 (5%) [12].

Внешнюю валидацию модели использовали в двух (10%) исследованиях [12, 13].

Для описания объёма использованной выборки в исследованиях используют разные показатели (см. приложение 3).

- число пациентов;
- число образцов (мазков, анализов);
- число изображений и их областей.

Объём выборки может сильно варьировать между разными исследованиями:

- число пациентов — 103–8021;
- число образцов — 167–212 554;
- число изображений — 510–695 030;
- количество областей изображения — от 260 тыс. до 7 млн.

Если в исследовании отмечено наличие нескольких показателей объёма выборки, то число образцов всегда значительно превышает количество пациентов, а число проанализированных изображений — количество образцов.

Информация о возрасте пациентов найдена в 8 публикациях из 19, при этом возрастные интервалы широко варьируют как внутри исследований, так и между работами. Информация о половом составе выборок найдена в 10 работах из 19. Относительно равное соотношение полов в выборках выявлено в двух исследованиях [8, 12]. Однако следует отметить, что анемию чаще диагностируют у женщин [12]. Исследования, посвящённые диагностике инфекций мочевыводящих путей, характеризуются преобладанием женщин [16, 23], поскольку эта патология чаще представлена у них. В ряде работ наблюдают неравномерное распределения пола в выборке по неясным причинам [9, 15, 20, 26]. Расовый и этнический состав пациентов приведён в одном исследовании [9]:

- «белые» пациенты с отсутствием злокачественных новообразований — 110 (90,2%);
- «белые» пациенты с подтверждённым злокачественным новообразованием — 25 (89,3%);
- «небелые» пациенты с отсутствием злокачественных новообразований — 12 (9,8%);
- «небелые» пациенты с подтверждённым злокачественным новообразованием — 3 (10,7%);
- латиноамериканцы с отсутствием злокачественных новообразований — 15 (12,3%);
- латиноамериканцы с подтверждённым злокачественным новообразованием — 8 (28,6%);
- не латиноамериканцы с отсутствием злокачественных новообразований — 104 (85,2%);
- не латиноамериканцы с подтверждённым злокачественным новообразованием — 18 (64,3%);
- пациенты с отсутствием злокачественных новообразований и без данных о расовой и этнической принадлежности — 3 (2,5%);
- пациенты с подтверждённым злокачественным новообразованием и без данных о расовой и этнической принадлежности — 2 (7,1%).

В 12 исследованиях авторы использовали собственные модели, разработанные с помощью различных алгоритмов машинного обучения. В 7 исследованиях применяли модели в виде готовых коммерческих решений, в которых:

- сообщали об отсутствии конфликта интересов — 2 [11, 14];
- отсутствовала информация об отсутствии конфликта интересов — 2 [8, 26];
- присутствует конфликт интересов, когда производитель оборудования финансирует исследование (например, предоставляет оборудование и материалы для проведения исследования, либо авторы работы одновременно являются или являлись в прошлом сотрудниками фирмы-продавца оборудования) — 3 [9, 17, 18].

Сравнительный анализ моделей, полученных с помощью различных алгоритмов, проводили в 6 исследованиях [9, 12, 15, 16, 23, 24], их диагностической эффективности в сравнении с медицинскими специалистами — 9 [8, 11, 12, 17, 18, 20–22, 25].

Диагностическая эффективность искусственного интеллекта на аналитическом этапе

Данные о диагностической эффективности ИИ представлены в приложении 5.

Средние обобщённые показатели эффективности моделей машинного обучения достаточно высоки:

- чувствительность — 0,923 (95% ДИ 0,921–0,924, $n=34$);
- специфичность — 0,940 (95% ДИ 0,939–0,942), $n=34$;
- значение AUC — 0,915 (95% ДИ 0,914–0,916), $n=14$;
- точность — 0,929 (95% ДИ 0,928–0,930), $n=37$.

Показатели, характеризующие эффективность, могут сильно варьировать в разных областях лабораторной медицины в соответствии с решаемыми задачами.

Так, модели машинного обучения на основе результатов анализа крови для диагностики анемии [12], а также опухолевых заболеваний и анемии [10] обладают следующими показателями эффективности (минимальное и максимальное значение):

- чувствительность — 0,930–0,980;
- специфичность — 0,920–1,000;
- значение AUC — 0,900–0,990.

Модели машинного обучения, включающие данные результатов мазков мокроты, для диагностики опухолевых заболеваний [9] также имеют высокие показатели (минимальное и максимальное значение):

- чувствительность — 0,820–0,920;
- специфичность — 0,770–0,880;

- значение AUC — 0,850–0,940).

Модели машинного обучения для диагностики опухолевых заболеваний, полученные с использованием результатов мазков костного мозга [13], обладают следующими показателями эффективности (минимальное и максимальное значение):

- чувствительность — 0,857–0,992;
- специфичность — 0,917–0,933;
- значение AUC — 0,970–0,990;
- точность — 0,914–0,929.

Качество идентификации и подсчёта клеток крови сильно варьирует в зависимости от типа анализируемых клеток. S. Yoon и соавт. [11] приводят метрики диагностической точности для классификации клеток с помощью цифрового морфологического анализатора, на финальном этапе полученные результаты врач-гематолог проверил и доработал. Однако подробный алгоритм их проверки экспертом в статье не представлен, также мы не обнаружили его в описании цифрового морфологического анализатора Vision Pro® (West Medica, Австрия) на сайте производителя⁵. Чувствительность анализатора была высокой в отношении нормальных лейкоцитов и ядросодержащих эритроидных клеток крови (0,801–0,980), и относительно низкой — для бластов, миелоцитов и метамиелоцитов (0,765, 0,480 и 0,505, соответственно). В то же время отмечена высокая специфичность для всех типов клеток (0,981–1,000).

Е.А. Елагина и соавт. [15] провели сравнительный анализ различных моделей машинного обучения для распознавания клеток крови. Следует отметить, что классификационные модели свёрточной нейронной сети и опорных векторов обладали наибольшей эффективностью в отношении диагностической точности. При этом модель опорных векторов страдала от переобучения и требовала значительных вычислительных затрат. Классификационная модель на основе метода k-ближайших соседей обладала меньшей диагностической точностью в сравнении с моделями свёрточной нейронной сети и опорных векторов.

Н. Ayıldız и соавт. [16] оценивали устойчивость *E. coli* к антибиотикам с помощью методов машинного обучения. Точность моделей варьировала в зависимости от используемого метода машинного обучения для их создания и типа антибиотика: 0,680–0,980. Кроме того, с помощью технологий ИИ с высокой диагностической точностью возможно определять микобактерии в тканях человека [19, 20] (чувствительность — 0,957–0,987; специфичность — 0,987–1,000; значение AUC — 0,980; точность — 0,983–0,988). Показатели эффективности моделей машинного обучения для диагностики бактериальных вагинитов могут варьировать в зависимости от их задач: чувствительность — 0,841–0,957; специфичность — 0,659–0,994 [21, 22].

С помощью модели машинного обучения возможно эффективно идентифицировать стрептококки в посевах на агаре, её чувствительность и специфичность составляет 0,906 и 0,940 соответственно. Модель для детекции колоний бактерий в посевах мочи демонстрирует высокий показатель чувствительности (0,998) и средний специфичности (0,720) [17].

R.J. Burton и соавт. [23] продемонстрировали эффективность использования технологий ИИ с целью снижения нагрузки на лабораторию посредством уменьшения числа посевов. Различные модели машинного обучения использовали для анализа осадка мочи (подсчёт собственных клеток и бактерий) с целью выявления необходимости дальнейшего посева. Авторы выявили, что модель экстремального градиентного бустинга обладала наибольшей эффективностью. Так, её применение по сравнению со стандартной автоматизированной микроскопией с помощью эвристической модели приведёт к тому, что результат исследования одного из четырёх пациентов будет переведён из ложноположительного в истинно отрицательный, и одного из 11 — из ложноотрицательного в истинно положительный. Авторы утверждают, что результаты исследования беременных и пациентов детского возраста необходимо анализировать в качестве отдельных выборок. D. Avci и соавт. [24] разработали модель для детекции разных элементов осадка мочи, созданную на основе свёрточной нейронной сети, которая продемонстрировала высокую точность — 0,962–0,986.

M.B. Wallace и соавт. [26] изучали возможное снижение количества ложноотрицательных результатов обнаружения кишечных неоплазий при использовании технологий ИИ. Для этого пациенты проходили две последовательные колоноскопии в один день. Пациентам 1-й группы сначала проводили колоноскопию с использованием технологий ИИ, а затем — без них. Во 2-й

⁵ Digital microscopy and AI: clinical and research applications [Internet]. Перхтольдсдорф: West Medica. 2021–2024. Режим доступа: <https://wm-vision.com/en/product/hema> Дата обращения: 12.10.2024.

группе, напротив, первоначально выполняли колоноскопию без применения технологий ИИ, после чего проводили повторное исследование с их использованием. Авторы рассчитывали показатель пропуска аденомы (AMR): количество гистологически подтверждённых поражений, обнаруженных при второй колоноскопии, делённое на общее количество поражений, обнаруженных при первой и второй колоноскопии, выполненной в течение одного дня. Рассчитывали также среднее количество поражений, обнаруженных при второй колоноскопии, и долю ложноотрицательных случаев (отсутствие поражений при первой колоноскопии и как минимум одно при второй). Показатель AMR в 1-й и 2-й группах составил 0,155 (38 из 246) и 0,324 (80 из 247) соответственно. Следует отметить, что он был ниже в 1-й группе при поражениях 5 мм и менее (0,159 против 0,358) и их неполипоидном типе (0,168 против 0,458). Кроме того, показатель AMR был ниже как в проксимальном (0,183 против 0,325), так и дистальном отделе толстого кишечника (0,108 против 0,321). Среднее количество аденом при второй колоноскопии — меньше в 1-й группе по сравнению со 2-й ($0,330 \pm 0,630$ против $0,700 \pm 0,970$, $p < 0,001$). Частота ложноотрицательных результатов в 1-й и 2-й группах составила 0,068 (3 из 44 пациентов) и 0,296 (13 из 44) соответственно.

Сравнение диагностической точности искусственного интеллекта и медицинских работников на аналитическом этапе лабораторного анализа

Сравнительный анализ диагностической точности ИИ и медицинских работников на аналитическом этапе лабораторного анализа представлен в приложении 6.

В исследованиях с проведённым анализом диагностической эффективности ИИ и медицинских работников показано, что она либо превосходит людей [12, 17, 20], либо сопоставима с ними [8, 11, 17, 18, 21, 22]. Согласно результатам двух исследований, при работе с одним пациентом ИИ затрачивает меньше времени, чем человек, что свидетельствует о его более высокой скорости обработки данных [12, 20]. В некоторых случаях, например при подсчёте разных клеток крови, результат сильно варьирует в зависимости от типа клеток: для одних — согласованность между моделями машинного обучения и медицинским работником высока, а для других — низкая [11]. Рассмотрим более подробно отдельные случаи.

Модель машинного обучения, позволяющая предсказывать низкую концентрацию ферритина в крови на основании общего клинического анализа крови и содержания С-реактивного белка, обладала более высокой диагностической эффективностью (чувствительность — 0,930–0,980; специфичность — 0,920) по сравнению с врачами клинической лабораторной диагностики (чувствительность — 0,830–0,880; специфичность — 0,910–0,920). Кроме того, применение ИИ позволяет значительно сократить время на принятие решения: менее 1 с на пациента против 19–20 с у врача и 13–16 с при использовании ИИ в качестве вспомогательного инструмента. Авторы полагают, что низкое содержание ферритина у пациентов с анемией возможно точно предсказать с помощью модели машинного обучения на основании результатов рутинных лабораторных исследований [12].

В случае диагностики и стадирования злокачественных новообразований с помощью индекса пролиферативной активности опухолевой клетки (Ki-67) оценивают согласованность результатов, полученных при анализе ИИ гистологических исследований, и стандартного протокола. Для количественного показателя Ki-67 использовали коэффициент внутриклассовой корреляции, который составил 0,960 (95% ДИ 0,940–0,980), а для оценки стадии заболевания применяли критерий квадратичного взвешенного κ — 0,860 (95% ДИ 0,810–0,910). Полученные результаты демонстрируют высокую степень согласованности между методами. Авторы полагают, что использование ИИ при определении значения Ki67 обеспечивает сопоставимую диагностическую точность с ручной оценкой, при этом выступая в роли эффективного вспомогательного инструмента, способного сэкономить время при диагностике патологий [8].

T.T. Van и соавт. [17] оценивали способность программного обеспечения PhenoMATRIX® (Copan Diagnostics Inc., Соединённые Штаты Америки) с хромогенным модулем обнаружения на основе ИИ автоматически идентифицировать колонии стрептококка группы А, выращенные на хромогенном агаре Colorex Strep A agar® (CHROMagar, Франция). Результаты, полученные с помощью программного обеспечения, сравнивали с результатами ручного считывания медицинскими лаборантами, обладающими опытом интерпретации хромогенных сред. Программное обеспечение на основе ИИ продемонстрировало более высокую чувствительность в сравнении с лаборантами при сопоставимой специфичности. Авторы считают, что его использование для обнаружения колоний стрептококков группы А на хромогенной питательной среде может оптимизировать рабочий процесс за счёт повышения скорости диагностики. Отметим, что в Российской Федерации интерпретацию

результатов лабораторных исследований может проводить только врач, что снижает ценность полученных результатов.

S. Yoon и соавт. [11] в своём исследовании при подсчёте лейкоцитов вычисляли разницу между их количеством, определённым тремя разными методами:

- с помощью ручного подсчёта, который проводили два врача-гематолога с использованием светового микроскопа на 200х увеличении;
- с помощью цифрового морфологического анализатора Vision Pro® (West Medica, Австрия);
- с помощью данного цифрового морфологического анализатора с конечной реклассификацией результата экспертом (в соответствии с нераскрытой методикой разработчика программного обеспечения).

Для разных типов клеток наблюдают высокую степень вариабельности согласованности между тремя способами подсчёта.

M.L. Faron и соавт. [18] применяли программное обеспечение на основе ИИ WASPLab® (Copan, Италия) для автоматического анализа культур мочи с использованием кровяного агара и агара МакКонни. Результаты ручного подсчёта выступали в качестве референс-теста. Подсчёт колоний в посевах мочи с помощью программного обеспечения характеризовался следующими показателями эффективности: чувствительности и специфичности — 0,998 и 0,720 соответственно. Авторы полагают, что его использование для анализа изображений является высокочувствительным, поэтому возможно внедрение в лаборатории для пакетного анализа отрицательных культур с целью улучшения рабочего процесса. Несогласованность между мануальным и автоматическим подсчётом связано с наличием микроколоний. Важной задачей для автоматизации подсчёта культур остаётся стандартизация пороговых значений выявления колоний. Лаборатории, представившие свои результаты для исследования, использовали разные критерии для определения положительных и отрицательных результатов, что затрудняет использование технологий ИИ и снижает финальную эффективность инструмента.

S. Zurac и соавт. [20] предложили автоматический метод идентификации *Mycobacterium tuberculosis* в образцах, окрашенных по методу Циля–Нильсена, и в тканях человека с использованием глубоких нейронных сетей. Данный метод продемонстрировал более высокие диагностические результаты (чувствительность — 0,957; специфичность — 1,000; точность — 0,983) в сравнении с патологоанатомами (чувствительность — 0,391–0,957; специфичность — 0,756–0,946; точность — 0,833). Среднее время, затрачиваемое патологоанатомами на исследование образца, варьировало от 5,48 до 17,06 мин: на анализ положительных слайдов уходило меньше времени, тогда как на отрицательные (истинно или ложноотрицательные) — больше. Для каждого специалиста наибольшая продолжительность обследования зафиксирована в отношении отрицательных случаев (истинно отрицательный для семи врачей и ложноотрицательный для одного), а наименьшая — истинно положительных. Время, затраченное патологоанатомами на исследование образца с помощью ИИ, варьировало от 9 с до 2 мин для положительных слайдов (в среднем 0,61 мин). Таким образом, автоматический метод идентификации на основе технологий ИИ, использованный в работе, экономит как минимум треть времени специалиста. Кроме того, он позволяет снижать вероятность человеческих ошибок, вызванных усталостью и потерей внимания.

В двух исследованиях сравнивали эффективность технологий ИИ и медицинских работников в диагностике вагинитов. В данных работах диагностическая точность ИИ сопоставима с результатами, показанными медицинскими работниками [21, 22]. В частности, в исследовании Z. Wang и соавт. [21] чувствительность модели свёрточной нейронной сети составила 0,914, а у медицинских специалистов (трёх лаборантов и двух врачей-акушеров-гинекологов) — 0,943. Она продемонстрировала более высокую специфичность (0,913 против 0,731) и точность (0,893 против 0,837). Авторы отмечают, что изменение качества изображений (например, цвета, яркости) влияет на точность работы модели свёрточной нейронной сети. Результаты данных исследований свидетельствуют о перспективности использования автоматизированной микроскопии для повышения качества первичной диагностики инфекционных и неинфекционных вагинитов.

В.А. Mathison и соавт. [25] выявили высокую степень согласованности между моделью свёрточной нейронной сети и медицинскими лаборантами при обнаружении и классификации кишечных простейших в образцах кала, окрашенных трихромом. Уровень положительной и отрицательной согласованности составил 0,989 (95% ДИ 0,938–1,000) и 0,981 (95% ДИ 0,934–0,998) соответственно. Модель показала высокую воспроизводимость при анализе слайдов, содержащих несколько классов, один класс или не содержащих паразитов. Авторы считают, что цифровое сканирование слайдов в

сочетании с протестированной моделью свёрточной нейронной сети — надёжный инструмент, дополняющий традиционные методы обнаружения кишечных простейших.

Подтверждённые результаты внедрения технологий искусственного интеллекта на аналитическом этапе лабораторного анализа

S. Kurstjens и соавт. [12] внедрили разработанную модель машинного обучения в работу лаборатории в тестовом режиме на 1 мес. Python скрипт реализован во внутренней лабораторной системе. В октябре 2021 г. проспективно проанализированы все результаты анализов взрослых пациентов первичного звена с анемией. Концентрацию ферритина измеряли у всех взрослых. В течение 21 дня при анализе данных 391 уникального пациента модель машинного обучения позволила выявить 18 новых случаев дефицита железа, не выявленных ранее. Авторы полагают, что технологии ИИ позволяют точнее и быстрее диагностировать низкое содержание железа в крови, что позволяет им быть полезным вспомогательным инструментом для врачей. Однако отмечены некоторые ограничения. Например, данная модель валидирована, поэтому её можно использовать в группе пациентов с определёнными характеристиками — взрослые пациенты с диагностированной анемией. Кроме того, следует учитывать, что в разных лабораториях используют различные референсные значения содержания ферритина в крови.

Ни в одном из включённых исследований не анализировали экономическую эффективность внедрения технологий ИИ в лабораторную практику.

ПОСТАНАЛИТИЧЕСКИЙ ЭТАП

Исследования, реализованные на постаналитическом этапе лабораторного анализа, выполнены в области контроля качества — направлены на поиск ошибок и аномальных значений в результатах анализов. Следует отметить, что работы с использованием технологий ИИ на постаналитическом этапе проводили преимущественно в следующих направлениях:

- анализ больших данных (Big Data);
- симуляции данных (например, искусственного внесения в данные ошибки);
- Δ -анализ — сравнительное исследование последовательных во времени анализов одного и того же пациента.

Во всех трёх исследованиях показано превосходство ИИ по сравнению с классическими статистическими методами PBRTQC (Patient-Based Real-Time Quality Control) контроля качества в контексте показателей эффективности.

Y. Liang и соавт. [27] изучали новый протокол стабильности данных путём объединения Δ -данных с методами машинного обучения для улучшения возможностей обнаружения событий контроля качества. Авторы сравнивали данные Δ -типа и данные одного типа, обработанные с использованием пределов усечения в PBRTQC на основе статистического метода. Сравнения проводили для семи параметров крови (см. приложение 3). В качестве клинического показателя использовали также количество затронутых пациентов от начала внесения систематической ошибки до её обнаружения, оптимальное значение которого должно быть минимальным. Результаты исследования показали, что диагностические параметры модели «случайного леса» значительно превосходят (см. приложение 5) стандартные статистические методы PBRTQC.

R. Zhou и соавт. [28] провели сравнительный анализ различных алгоритмов постаналитического контроля качества биохимических тестов, направленных на определение содержания общего простатоспецифического антигена. В исследовании рассмотрены четыре традиционные модели контроля качества (PBRTQC) — Moving Average; Moving Median; Moving Standard Deviation; Moving Sum of Number of Patient Results. Кроме того, авторы сравнивали эффективность трёх моделей машинного обучения — «случайного леса», опорных векторов и нейронной сети. Также протестирован алгоритм слияния информационной энтропии, объединяющий все три модели машинного обучения. При симуляции в данные вносили ошибки восьми разных уровней (0,01–0,20 мкг/л), а также использовали шесть вариантов блоков различного размера. Все модели машинного обучения и их комбинация превосходили методы стандартных алгоритмов PBRTQC по параметрам диагностической точности. Модель слияния превосходила по эффективности каждую из трёх моделей машинного обучения по отдельности. Модель «случайного леса» предрасположена к переобучению. Модель опорных векторов демонстрировала затруднения при множественной классификации, в то время как модель нейронной сети столкнулась с трудностями при принятии решений. Модель слияния по точности превосходила следующие модели:

- опорных векторов — на 8,7%;

- «случайного леса» — на 9,6%;
- нейронной сети — на 6,9%;
- стандартные PBRTQC — на 20%.

Кроме того, эффективность методов PBRTQC варьировала в зависимости от уровня внесённой ошибки, тогда как модели машинного обучения демонстрировали стабильную работу независимо от её величины.

Н. Wang и соавт. [29] в своём исследовании тестировали возможности нескольких моделей машинного обучения в отношении верификации данных биохимических тестов, включающих 52 биохимических параметра. Финальная обобщённая модель показала уровень прохождения и ложноотрицательных результатов 89,60 и 0,095% соответственно. Полученная модель позволила сократить количество недействительных отчётов примерно на 80% по сравнению с теми, которые оценивали с помощью стандартного алгоритма, повысив эффективность работы и снизив нагрузку на персонал биохимической лаборатории. Примечательно, что два обнаруженных ложноотрицательных отчёта, принадлежали пациентам с экстремальными для выборки значениями возраста: 4 мес. и 92 года. Кроме того, результаты (уровень прохождения) стандартного лабораторного алгоритма (50,20–65,10%) флуктуировали в течение рабочего процесса, тогда как для модели машинного обучения они были относительно стабильны (87,00–94,00%).

ОЦЕНКА КАЧЕСТВА МЕТОДОЛОГИИ ВСЕХ ИССЛЕДОВАНИЙ

Оценка качества методологии проанализированных исследований с использованием модифицированного опросника QUADAS-CAD представлена в **приложении 7**. Во всех проанализированных исследованиях ($n=23$, 100%) присутствует вероятность систематической ошибки из-за особенностей использования методик (**рис. 1**). Только в незначительной части исследований выборки были сбалансированы по уровню патологии ($n=3$, 13,0%) и демографическим характеристикам ($n=2$, 8,3%). Поскольку соответствующие вопросы являются сигнальными для домена D1 (Patient Selection), риск систематической ошибки в нём высок или вызывает опасения во всех проанализированных случаях. В некоторых исследованиях ($n=9$, 39,1%) из описания методики не было возможности определить, пересекались ли обучающая и тестовая выборка, что представляет собой ключевой вопрос в рамках домена D2 (Index Test). В ряде случаев некоторые вопросы домена D2 (например, «Если использовался порог патологии, был ли он установлен заранее?») и домена D3 (например, «Может ли референсный стандарт правильно классифицировать целевое состояние?» и «Были ли результаты референсных стандартов подготовлены или проверены с необходимым уровнем экспертизы?») были неприменимыми. Это связано с тем, что в некоторых исследованиях [11, 18, 19, 25] оценивали способность моделей подсчитывать разные типы клеток без диагностики заболеваний, либо использовали симуляцию данных [7, 27–29].

В большинстве случаев референсный стандарт позволял корректно классифицировать целевые состояния ($n=19$, 82,6%). Тем не менее в 9 исследованиях (39,1%) уровень экспертной подготовки и критерии оценки референсных стандартов недостаточно определены. Минимальный риск систематической ошибки обнаружен для домена D4, оценивающего прозрачность полученных результатов.

ОБСУЖДЕНИЕ

ОБЛАСТИ ПРИМЕНЕНИЯ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Выполненный систематический обзор продемонстрировал широкий спектр областей лабораторной медицины, в которых возможно применение технологий ИИ. Особое внимание следует уделить тому, что аналитический этап, с одной стороны, и преаналитический и постаналитический этапы — с другой, представляют собой две обширные сферы со своими специфическими задачами и методами их решения. На аналитическом этапе лабораторного анализа основным направлением применения технологий ИИ является распознавание объектов различной морфологии на лабораторных изображениях и их количественный анализ. Наблюдают большое разнообразие областей, где необходимо решать эту задачу:

- анализ образцов крови с целью диагностики различных гематологических заболеваний — лейкозов и анемий;
- анализ образцов осадка мочи для выявления инфекций мочевыводящих путей;

- выявление разного рода микроорганизмов (от бактерий до простейших) в пробах тканей, мазках и посевах;
- анализ образцов мокроты методами проточной цитометрии для выявления злокачественных новообразований лёгких;
- анализ результатов биопсии костного мозга для выявления онкологических заболеваний крови.

В большинстве проанализированных исследований авторы использовали собственные решения, созданные в среде Python или в среде R. Наиболее эффективный результат (в плане диагностической точности, скорости и отсутствия проблем переобучения) показывают модели свёрточных нейронных сетей. Тем не менее в некоторых исследованиях авторы использовали готовые коммерческие решения для анализа образцов крови, посевов мочи и слизи из ротоглотки, а также образцов мокроты с помощью проточной цитометрии.

На преаналитическом и постаналитическом этапах основной задачей является обеспечение контроля качества данных [2, 30], что включает выявление ошибочной маркировки пробирок, а также поиск выпадающих или ошибочных значений результатов анализов как для каждого пациента персонально, так и для выборки в целом. В этом контексте наиболее востребованы инструменты из области больших данных [31].

Неравномерное распределение числа работ по разным этапам лабораторного анализа (значительное преобладание работ на аналитическом этапе) отражает структуру нашего поискового запроса при отборе литературы. В дальнейшем каждая из этих областей заслуживает отдельного исследования.

Диагностическая эффективность искусственного интеллекта и его внедрение в лабораторный процесс

Во всех проанализированных работах продемонстрирована высокая диагностическая точность моделей машинного обучения, достаточная для их внедрения в лабораторную практику. Эффективность ИИ сопоставима с результатами работы медицинских специалистов высшей квалификации и превышает показатели специалистов начального уровня. Так, для аналитического этапа обобщённые оценки диагностической эффективности ИИ составляют:

- чувствительность — 0,923;
- специфичность — 0,940;
- точность — 0,929.

На преаналитическом этапе точность моделей машинного обучения в отношении выявления ошибочной маркировки пробирок составила 0,865–0,921. На постаналитическом этапе показатели, отражающие эффективность применения ИИ в контексте контроля качества данных (чувствительность, специфичность, точность), достигают 0,990. По скорости диагностики обученные модели машинного обучения значительно превосходят медицинских работников при анализе изображений и данных.

Внедрение технологий искусственного в лабораторный процесс и проблемы, препятствующие этому

Несмотря на высокую эффективность ИИ в экспериментах, случаи реального его внедрения в практику единичны и находятся на экспериментальном этапе. Мы обнаружили только одну публикацию, посвящённую опыту экспериментального внедрения технологий ИИ (в течение 1 мес.) в лабораторную практику. Следует отметить, что существует несколько групп проблем, препятствующих данному процессу [32, 33].

- Причины, связанные со структурой экспериментальных работ, где получены оценки высокой эффективности ИИ. Важной и наиболее частой проблемой здесь является несбалансированность выборок при тестировании моделей машинного обучения [34]. Преимущественно наблюдают несбалансированность в отношении исследуемых патологий — выборка представлена пациентами с патологией. Это, очевидно, можно объяснить тем, что исследования проводят в медицинских организациях, специализирующихся на конкретных заболеваниях, и учёным доступны именно такие выборки. Тем не менее это повышает вероятность систематической ошибки в рамках исследования и снижения эффективности обученной модели в условиях реальной практики. Это же относится и к несбалансированности выборок по демографическим характеристикам, в частности по полу и возрасту. Часто в экспериментальную выборку включают всех доступных пациентов. Однако обнаружено, что для пациентов младшей и старшей возрастных

групп, а также для беременных эффективность работы ИИ в некоторых случаях может иметь существенные отличия.

- Отсутствие внешней валидации результатов тестирования моделей машинного обучения в большинстве работ, что приводит к невозможности генерализовать полученные в эксперименте результаты. Например, показано, что качество изображений (например, цвет и яркость) могут влиять на результат работы ИИ.

В некоторых исследованиях с целью увеличения объёма выборки используют метод создания из одного изображения нескольких, нарезая его на области. Далее их используют как независимые экземпляры в общей выборке в контексте применения технологий ИИ. Перед проведением подобной процедуры необходима предварительная проверка, которая позволит убедиться, что анализируемые участки одного изображения независимы друг от друга. Её отсутствие может вызвать псевдорепликацию, что сопровождается завышением эффективности используемой модели.

Кроме того, выявлены исследования, в которых присутствует конфликт интересов, когда производитель оборудования прямо или опосредованно финансирует работу, что повышает риск систематической ошибки при оценке эффективности ИИ. Тем не менее мы не обнаружили ни одной работы, где показана его низкая эффективность. Это косвенно указывает на наличие публикационной предвзятости, при которой отрицательные результаты недоступны для научного сообщества.

Общая сложность медицинских задач также в некоторых случаях труднопреодолима. Например, на одном и том же изображении системы ИИ могут распознавать некоторые типы клеток с очень высокой эффективностью, тогда как другие — с недопустимо низкой.

Структура работы и потребности лаборатории часто сложнее экспериментальных условий. Более того, условия работы разных лабораторий даже в одной области могут варьировать (например, использование своих референсных значений). Потребность в использовании технологий ИИ обусловлена экономией времени и ресурсов. Однако существуют факторы, которые способны нивелировать потенциальные преимущества их применения в реальных условиях:

- необходимость большого количества аннотированных изображений;
- вариабельность чувствительности и специфичности (например, высокая эффективность при обнаружении истинно положительных случаев при одновременном её снижении в выявлении истинно отрицательных);
- необходимость обучения медицинских работников;
- обязательная верификация результатов классификации человеком;
- сложность при интерпретации выводов и вариабельность результатов (работа нейронной сети — «чёрный ящик» для пользователя).

Оценка временных и финансовых затрат, связанных с внедрением технологий ИИ в лабораторную практику, остаётся неопределённой.

Таким образом, модели машинного обучения имеют хороший потенциал в качестве вспомогательного инструмента для медицинских работников в области лабораторной медицины. Технологии ИИ способны автоматизировать рутинные лабораторные процессы, способствуют их стандартизации [35, 36], позволяют решать широкий спектр задач в патоморфологии. Возникает необходимость проведения экспериментальных исследований в области применения ИИ с учётом существующих методических проблем [37], а также оценки соотношения затрат и ожидаемой выгоды при его внедрении в лабораторную практику.

ОГРАНИЧЕНИЯ СИСТЕМАТИЧЕСКОГО ОБЗОРА

Разнообразие задач лабораторной медицины крайне велико и затрагивает многие медицинские и технологические области. Для поиска литературы мы использовали поисковую систему PubMed и программу для управления библиографической информацией Mendeley, что позволило отобрать исследования по наиболее часто встречающимся темам. Однако широкий спектр задач, ограничения структуры запроса и доступа к некоторым исследованиям не позволяет рассмотреть все существующие здесь вопросы, кроме того, некоторые темы не затронуты. Например, к таким областям относят вопросы роботизации забора образцов (в частности, забор венозной крови) [38], оптимизации назначений анализов и прогнозирования. Также по причине большого разнообразия задач во включённых исследованиях и, как следствие, невозможности группировки полученных моделей, разработанных при использовании различных методов машинного обучения, мы рассчитали только усреднённые оценки диагностической точности, не проводя метаанализ для отдельных моделей в соответствии с рекомендациями Кокрейновского руководства [39].

Помимо собственно медицинских, технических и экономических аспектов, внедрение технологий ИИ в практику сопровождается гуманитарными проблемами [40]. Например, разного рода страхи и психологические сложности медицинских работников и пациентов, основанные преимущественно на недостаточной осведомлённости о них. Кроме того, важной проблемой при использовании технологий ИИ является защита персональных данных пациентов. Все эти вопросы требуют отдельного исследования и решения.

Важно отметить, что практическое применение моделей машинного обучения и систем на основе ИИ возможно только после получения статуса медицинского изделия с технологиями ИИ. Его присваивают централизованно в Российской Федерации (Росздравнадзор, регистрационное удостоверение) и Соединённых Штатах Америки (Food and Drug Administration, FDA), либо децентрализованно — аккредитованными частными уполномоченными органами в странах Европейского союза (маркировка CE) [41]. Публикации, которые мы проанализировали, включая работу с экспериментальным этапом внедрения технологий ИИ в лабораторный процесс [12], не содержали информации о регистрации статуса медицинских изделий, что ещё раз подтверждает начальные этапы их внедрения в исследуемую область медицины.

ЗАКЛЮЧЕНИЕ

Потенциал применения технологий ИИ в лабораторной медицине охватывает все этапы лабораторного процесса и остаётся особенно актуальным.

Анализ выявленных исследований показал их распределение по всем этапам лабораторного анализа — преаналитическому, аналитическому и постаналитическому. При этом на аналитическом этапе выполнено большинство исследований (83,6%). Основное внимание в них уделяли диагностике гематологических и онкологических заболеваний. Также представлены исследования, направленные на выявление патогенных бактерий в тканевых пробах, моче и мазках. Кроме того, по одной работе посвящено вопросам паразитологии и гистопатологии. Работы, выполненные на преаналитическом и постаналитическом этапах, направлены на разработку эффективных методов контроля качества лабораторных отчётов с помощью технологий ИИ. Текущая стадия их внедрения в лабораторную медицину характеризуется ранним этапом развития, что подтверждается преобладанием собственных разработок. Лишь в 30,4% включённых работ использовали готовые коммерческие решения.

Модели машинного обучения и системы на основе ИИ демонстрируют высокие показатели эффективности, сравнимые с показателями высококвалифицированных медицинских работников или превосходящие их. Однако анализ качества методологии включённых работ, который мы провели, продемонстрировал высокую вероятность систематической ошибки во всех оцениваемых доменах, за исключением прозрачности полученных результатов. Высокий риск систематических ошибок обусловлен несбалансированностью выборок по представленным патологиям и демографическим характеристикам, возможной псевдорепликацией данных, отсутствием внешней валидации результатов, что в совокупности затрудняет их обобщение.

Завышение оценок эффективности ИИ во включённых исследованиях косвенно подтверждают малочисленные попытки внедрения разработанных моделей в рутинную практику. Мы выявили единственный эксперимент, продемонстрировавший положительные результаты в выявлении новых случаев дефицита железа.

Таким образом, технологии ИИ обладают значительным потенциалом в повышении эффективности и скорости выполнения рутинных лабораторных процессов, способствуя их автоматизации и стандартизации, а также высвобождению времени медицинского персонала для решения более сложных задач. Тем не менее для полноценного их внедрения в лабораторную практику необходимо комплексное решение ряда вопросов, связанных с оценкой его надёжности, воспроизводимости и практического применения.

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Приложение 1. Список публикаций, включённых в систематический обзор, и их характеристики. DOI: 10.17816/DD635349-4334766

Приложение 2. Список публикаций, исключённых из систематического обзора. DOI: 10.17816/DD635349-4334769

Приложение 3. Основные характеристики исследований, представленных в публикациях, включённых в систематический обзор. DOI: 10.17816/DD635349-4334770

Приложение 4. Характеристики выборок, используемые модели машинного обучения или готовые коммерческие решения, представленные в исследованиях. DOI: 10.17816/DD635349-4334771

Приложение 5. Эффективность искусственного интеллекта в исследованиях. DOI: 10.17816/DD635349-4334772

Приложение 6. Сравнительный анализ диагностической эффективности искусственного интеллекта и медицинских специалистов. DOI: 10.17816/DD635349-4334773

Приложение 7. Оценка качества методологии исследований с использованием модифицированного опросника QUADAS-CAD. DOI: 10.17816/DD635349-4334767

Вклад авторов. Ю.А. Васильев, А.В. Владзимирский, А.С. Гольдберг — разработка концепции исследования; О.Г. Нанова, И.А. Блохин, Р.В. Решетников — сбор и анализ литературных данных, написание и редактирование текста рукописи. Все авторы одобрили рукопись (версию для публикации), а также согласились нести ответственность за все аспекты работы, гарантируя надлежащее рассмотрение и решение вопросов, связанных с точностью и добросовестностью любой её части.

Этическая экспертиза. Неприменимо.

Источники финансирования. Данная статья подготовлена авторским коллективом в рамках НИР «Научное обоснование методов лучевой диагностики опухолевых заболеваний с использованием радиомического анализа», (№ ЕГИСУ: № 123031500005-2) в соответствии с Приказом от 22.12.2023 г. № 1258 "Об утверждении государственных заданий, финансовое обеспечение которых осуществляется за счет средств бюджета города Москвы государственным бюджетным (автономным) учреждениям подведомственным Департаменту здравоохранения города Москвы, на 2024 год и плановый период 2025 и 2026 годов" Департамента здравоохранения города Москвы.

Раскрытие интересов. Авторы заявляют об отсутствии отношений, деятельности и интересов за последние три года, связанных с третьими лицами (коммерческими и некоммерческими), интересы которых могут быть затронуты содержанием статьи.

Оригинальность. При создании настоящей работы авторы не использовали ранее опубликованные сведения (текст, иллюстрации, данные).

Доступ к данным. Редакционная политика в отношении совместного использования данных к настоящей работе применима. Все данные, полученные в настоящем исследовании, доступны в статье и в приложении к ней. В частности, в Приложении 1-7.

Генеративный искусственный интеллект. При создании настоящей статьи технологии генеративного искусственного интеллекта не использовали.

Рассмотрение и рецензирование. Настоящая работа подана в журнал в инициативном порядке и рассмотрена по обычной процедуре. В рецензировании участвовали два члена редакционной коллегии и научный редактор издания.

ADDITIONAL INFORMATION

Supplement 1. List of publications included in the systematic review and their characteristics. DOI: 10.17816/DD635349-4334766

Supplement 2. List of publications excluded from the systematic review. DOI: 10.17816/DD635349-4334769

Supplement 3. Main characteristics of the studies presented in the publications included in the systematic review. DOI: 10.17816/DD635349-4334770

Supplement 4. Characteristics of samples and machine learning models used, or commercially available solutions presented in the studies. DOI: 10.17816/DD635349-4334771

Supplement 5. Effectiveness of artificial intelligence in studies. DOI: 10.17816/DD635349-4334772

Supplement 6. Comparative analysis of the diagnostic efficiency of artificial intelligence and medical workers. DOI: 10.17816/DD635349-4334773

Supplement 7. Assessment of the quality of research methodology using the modified QUADAS-CAD questionnaire. DOI: 10.17816/DD635349-4334768

Authors' contribution. Yu.A. Vasilev, A.V. Vladzymyrskyy, A.S. Goldberg: development of the research concept, improvement of the final version of the manuscript; O.G. Nanova, I.A. Blokhin, R.V. Reshetnikov: literature review, data analysis, writing the text of the manuscript, text editing. Thereby, all authors provided approval of the version to be published and agree to be accountable for all aspects of the work in ensuring

that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Ethics approval. Not applicable.

Funding sources. This paper was prepared by a team of authors as a part of the research project "Scientific substantiation of diagnostic radiology methods of tumors using radiomic analysis", (EGISU No. 123031500005-2) in accordance with Order No. 1258 dated December 22, 2023: "On approval of state assignments funded from the Moscow city budget to state budgetary (autonomous) institutions subordinated to the Moscow Healthcare Department for 2024 and the planning period of 2025 and 2026".

Disclosure of interests. The authors declare the absence of relationships, activities or interests over the past three years related to third parties (commercial and non-commercial) whose interests may be affected by the content of the article.

Statement of originality. In creating this work, the authors did not use previously published information (text, illustrations, data).

Data availability statement. The editorial policy on data sharing is applicable to this work. All data obtained in this study are available in the article and in its appendix. In particular, in Appendix 1-7.

Generative AI.

Provenance and peer-review. This article was submitted to the Journal on an unsolicited basis and reviewed according to the usual procedure. Two members of the editorial board and the science editor were involved in the review process.

СПИСОК ЛИТЕРАТУРЫ | REFERENCES

1. Bonert M, Zafar U, Maung R, et al. Pathologist workload, work distribution and significant absences or departures at a regional hospital laboratory. *PLOS ONE*. 2022;17(3):e0265905. doi: [10.1371/journal.pone.0265905](https://doi.org/10.1371/journal.pone.0265905) EDN: [UFNVFE](#)
2. Hou H, Zhang R, Li J. Artificial intelligence in the clinical laboratory. *Clinica Chimica Acta*. 2024;559:119724. doi: [10.1016/j.cca.2024.119724](https://doi.org/10.1016/j.cca.2024.119724) EDN: [PBDERB](#)
3. Munari E, Scarpa A, Cima L, et al. Cutting-edge technology and automation in the pathology laboratory. *Virchows Archiv*. 2023;484(4):555–566. doi: [10.1007/s00428-023-03637-z](https://doi.org/10.1007/s00428-023-03637-z) EDN: [OSGENI](#)
4. Vasilev YuA, Vladzimirskyy AV, Omelyanskaya OV, et al. *Guidelines for preparing a systematic review*. Moscow: State Budget-Funded Health Care Institution of the City of Moscow "Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of Moscow Health Care Department"; 2023. 34 p. (In Russ.) EDN: [XKXHDA](#)
5. Anjankar AP, Jha RK, Lambe S. Implementation of artificial intelligence in laboratory medicine. *Journal of Datta Meghe Institute of Medical Sciences University*. 2023;18(4):598–601. doi: [10.4103/jdmimsu.jdmimsu_486_22](https://doi.org/10.4103/jdmimsu.jdmimsu_486_22) EDN: [VBNWUF](#)
6. Kodenko MR, Vasilev YuA, Vladzimirskyy AV, et al. Diagnostic accuracy of ai for opportunistic screening of abdominal aortic aneurysm in ct: a systematic review and narrative synthesis. *Diagnostics*. 2022;12(12):3197. doi: [10.3390/diagnostics12123197](https://doi.org/10.3390/diagnostics12123197) EDN: [ERWYPX](#)
7. Farrell CJ. Identifying mislabelled samples: machine learning models exceed human performance. *Annals of Clinical Biochemistry: International Journal of Laboratory Medicine*. 2021;58(6):650–652. doi: [10.1177/00045632211032991](https://doi.org/10.1177/00045632211032991) EDN: [MQQLCW](#)
8. Lea D, Gudlaugsson EG, Skaland I, et al. Digital image analysis of the proliferation markers Ki67 and phosphohistone H3 in gastroenteropancreatic neuroendocrine neoplasms: accuracy of grading compared with routine manual hot spot evaluation of the Ki67 index. *Applied Immunohistochemistry & Molecular Morphology*. 2021;29(7):499–505. doi: [10.1097/pai.0000000000000934](https://doi.org/10.1097/pai.0000000000000934) EDN: [XIKRGL](#)
9. Lemieux ME, Reveles XT, Rebeles J, et al. Detection of early-stage lung cancer in sputum using automated flow cytometry and machine learning. *Respiratory Research*. 2023;24(1):23. doi: [10.1186/s12931-023-02327-3](https://doi.org/10.1186/s12931-023-02327-3) EDN: [HSQBUA](#)
10. Kimura K, Tabe Y, Ai T, et al. A novel automated image analysis system using deep convolutional neural networks can assist to differentiate MDS and AA. *Scientific Reports*. 2019;9(1):1–9. doi: [10.1038/s41598-019-49942-z](https://doi.org/10.1038/s41598-019-49942-z) EDN: [PXXHII](#)
11. Yoon S, Hur M, Park M, et al. Performance of digital morphology analyzer Vision Pro on white blood cell differentials. *Clinical Chemistry and Laboratory Medicine (CCLM)*. 2021;59(6):1099–1106. doi: [10.1515/cclm-2020-1701](https://doi.org/10.1515/cclm-2020-1701) EDN: [GVMONA](#)

EDN: BXDWFO

12. Kurstjens S, de Bel T, van der Horst A, et al. Automated prediction of low ferritin concentrations using a machine learning algorithm. *Clinical Chemistry and Laboratory Medicine (CCLM)*. 2022;60(12):1921–1928. doi: [10.1515/cclm-2021-1194](https://doi.org/10.1515/cclm-2021-1194) EDN: [HDJWKG](#)
13. Wang M, Dong C, Gao Y, et al. A deep learning model for the automatic recognition of aplastic anemia, myelodysplastic syndromes, and acute myeloid leukemia based on bone marrow smear. *Frontiers in Oncology*. 2022;12: 844978. doi: [10.3389/fonc.2022.844978](https://doi.org/10.3389/fonc.2022.844978) EDN: [BQFWSO](#)
14. Kim H, Lee GH, Yoon S, et al. Performance of digital morphology analyzer Medica EasyCell assistant. *Clinical Chemistry and Laboratory Medicine (CCLM)*. 2023;61(10):1858–1866. doi: [10.1515/cclm-2023-0100](https://doi.org/10.1515/cclm-2023-0100) EDN: [ZDXONI](#)
15. Elagina EA, Margun AA. Research of machine learning methods in the problem of identification of blood cells. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. 2021;21(6):903–911. doi: [10.17586/2226-1494-2021-21-6-903-911](https://doi.org/10.17586/2226-1494-2021-21-6-903-911) EDN: [ZVQLEV](#)
16. Ayyıldız H, Arslan Tuncer S. Is it possible to determine antibiotic resistance of *E. coli* by analyzing laboratory data with machine learning? *Turkish Journal of Biochemistry*. 2021;46(6):623–630. doi: [10.1515/tjb-2021-0040](https://doi.org/10.1515/tjb-2021-0040) EDN: [JTZHYJ](#)
17. Van TT, Mata K, Bard JD. Automated detection of *Streptococcus pyogenes* pharyngitis by use of Colorex Strep A CHROMagar and WASPLab artificial intelligence chromogenic detection module software. *Journal of Clinical Microbiology*. 2019;57(11):e00811-19. doi: [10.1128/JCM.00811-19](https://doi.org/10.1128/JCM.00811-19)
18. Faron ML, Buchan BW, Relich RF, et al. Evaluation of the WASPLab segregation software to automatically analyze urine cultures using routine blood and MacConkey agars. *Journal of Clinical Microbiology*. 2020;58(4):e01683-19. doi: [10.1128/jcm.01683-19](https://doi.org/10.1128/jcm.01683-19) EDN: [UDENAP](#)
19. Yang M, Nurzynska K, Walts AE, Gertych A. A CNN-based active learning framework to identify mycobacteria in digitized Ziehl–Neelsen stained human tissues. *Computerized Medical Imaging and Graphics*. 2020;84:101752. doi: [10.1016/j.compmedimag.2020.101752](https://doi.org/10.1016/j.compmedimag.2020.101752) EDN: [AYLPVY](#)
20. Zurac S, Mogodici C, Poncu T, et al. A new artificial intelligence-based method for identifying mycobacterium tuberculosis in Ziehl–Neelsen stain on tissue. *Diagnostics*. 2022;12(6):1484. doi: [10.3390/diagnostics12061484](https://doi.org/10.3390/diagnostics12061484) EDN: [IJUCYT](#)
21. Wang Z, Zhang L, Zhao M, et al. Deep neural networks offer morphologic classification and diagnosis of bacterial vaginosis. *Journal of Clinical Microbiology*. 2021;59(2):e02236-20. doi: [10.1128/JCM.02236-20](https://doi.org/10.1128/JCM.02236-20) EDN: [GBZITD](#)
22. Lev-Sagie A, Strauss D, Ben Chetrit A. Diagnostic performance of an automated microscopy and pH test for diagnosis of vaginitis. *NPJ Digital Medicine*. 2023;6(1):66. doi: [10.1038/s41746-023-00815-w](https://doi.org/10.1038/s41746-023-00815-w) EDN: [SVUVPI](#)
23. Burton RJ, Albur M, Eberl M, Cuff SM. Using artificial intelligence to reduce diagnostic workload without compromising detection of urinary tract infections. *BMC Medical Informatics and Decision Making*. 2019;19:171. doi: [10.1186/s12911-019-0878-9](https://doi.org/10.1186/s12911-019-0878-9)
24. Avci D, Sert E, Dogantekin E, et al. A new super resolution Faster R-CNN model based detection and classification of urine sediments. *Biocybernetics and Biomedical Engineering*. 2023;43(1):58–68. doi: [10.1016/j.bbe.2022.12.001](https://doi.org/10.1016/j.bbe.2022.12.001) EDN: [HJRRRR](#)
25. Mathison BA, Kohan JL, Walker JF, et al. Detection of intestinal protozoa in trichrome-stained stool specimens by use of a deep convolutional neural network. *Journal of Clinical Microbiology*. 2020;58(6):e02053-19. doi: [10.1128/jcm.02053-19](https://doi.org/10.1128/jcm.02053-19) EDN: [GWHHRT](#)
26. Wallace MB, Sharma P, Bhandari P, et al. Impact of artificial intelligence on miss rate of colorectal neoplasia. *Gastroenterology*. 2022;163(1):295–304.e5. doi: [10.1053/j.gastro.2022.03.007](https://doi.org/10.1053/j.gastro.2022.03.007) EDN: [CVAOAF](#)
27. Liang Y, Wang Z, Huang D, et al. A study on quality control using delta data with machine learning technique. *Heliyon*. 2022;8(8):e09935. doi: [10.1016/j.heliyon.2022.e09935](https://doi.org/10.1016/j.heliyon.2022.e09935) EDN: [XNSZKR](#)
28. Zhou R, Liang Y, Cheng H, et al. A multi-model fusion algorithm as a real-time quality control tool for small shift detection. *Computers in Biology and Medicine*. 2022;148:105866. doi: [10.1016/j.compbiomed.2022.105866](https://doi.org/10.1016/j.compbiomed.2022.105866) EDN: [OBKKZC](#)
29. Wang H, Wang H, Zhang J, et al. Using machine learning to develop an autoverification system in a clinical biochemistry laboratory. *Clinical Chemistry and Laboratory Medicine (CCLM)*. 2020;59(5):883–891. doi: [10.1515/cclm-2020-0716](https://doi.org/10.1515/cclm-2020-0716) EDN: [SVNLZY](#)
30. Lippi G, Mattiuzzi C, Favaloro E. Artificial intelligence in the pre-analytical phase: state-of-the art and future perspectives. *Journal of Medical Biochemistry*. 2024;43(1):1–10. doi: [10.5937/jomb0-45936](https://doi.org/10.5937/jomb0-45936) EDN: [PVAVYI](#)

EDN: BXDWFO

31. Blatter TU, Witte H, Nakas CT, Leichtle AB. Big data in laboratory medicine-FAIR quality for AI? *Diagnostics*. 2022;12(8):1923. doi: [10.3390/diagnostics12081923](https://doi.org/10.3390/diagnostics12081923) EDN: [MCJCST](#)
32. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet Digital Health*. 2021;3(11):e745–e750. doi: [10.1016/s2589-7500\(21\)00208-9](https://doi.org/10.1016/s2589-7500(21)00208-9) EDN: [EHUNYG](#)
33. Paranjape K, Schinkel M, Hammer RD, et al. The value of artificial intelligence in laboratory medicine. *American Journal of Clinical Pathology*. 2020;155(6):823–831. doi: [10.1093/ajcp/aqaa170](https://doi.org/10.1093/ajcp/aqaa170) EDN: [KUADLL](#)
34. Ghosh K, Bellinger C, Corizzo R, et al. The class imbalance problem in deep learning. *Machine Learning*. 2022;113(7):4845–4901. doi: [10.1007/s10994-022-06268-8](https://doi.org/10.1007/s10994-022-06268-8) EDN: [AQXQUP](#)
35. Certificate of state registration of a computer program No. 2023665713/ 19.07.2023. Byul. No. 7. Vasilev YuA, Vladzymyrskyy AV, Omelyanskaya OV, et al. *Web platform for technological and clinical monitoring of the results of digital medical image analysis algorithms*. Available from: https://elibrary.ru/download/elibrary_54200632_17081735.PDF (In Russ.) EDN: [JIEPJK](#)
36. Zinchenko VV, Arzamasov KM, Kremneva EI, et al. Technological defects in software based on artificial intelligence. *Digital Diagnostics*. 2023;4(4):593–604. doi: [10.17816/DD501759](https://doi.org/10.17816/DD501759) EDN: [ORUFMM](#)
37. Sharova DE, Garbuk SV, Vasilyev YuA. Artificial intelligence systems in clinical medicine: the world's first series of national standards. *Standards and Quality*. 2023;(1):46–51. doi: [10.35400/0038-9692-2023-1-304-22](https://doi.org/10.35400/0038-9692-2023-1-304-22) EDN: [SNMGQA](#)
38. Laddi A, Goyal S, Savlania A. Vein segmentation and visualization of upper and lower extremities using convolution neural network. *Biomedical Engineering. Biomedizinische Technik*. 2024;69(5):455–464. doi: [10.1515/bmt-2023-0331](https://doi.org/10.1515/bmt-2023-0331) EDN: [PRAAZI](#)
39. Macaskill P, Takwoingi Y, Deeks JJ, Gatsonis C. Chapter 9: Understanding meta-analysis. In: Deeks JJ, Bossuyt PM, Leeflang MM, Takwoingi Y, editors. *Cochrane handbook for systematic reviews of diagnostic test accuracy. version 2.0* (updated July 2023). Cochrane; 2023 [cited 2024 Aug 17]. Available from: <https://training.cochrane.org/handbook-diagnostic-test-accuracy/current>
40. Pennestri F, Banfi G. Artificial intelligence in laboratory medicine: fundamental ethical issues and normative key-points. *Clinical Chemistry and Laboratory Medicine (CCLM)*. 2022;60(12):1867–1874. doi: [10.1515/cclm-2022-0096](https://doi.org/10.1515/cclm-2022-0096) EDN: [ZOALXU](#)
41. Muehlematter UJ, Daniore P, Vokinger KN. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. *The Lancet Digital Health*. 2021;3(3):e195–e203. doi: [10.1016/s2589-7500\(20\)30292-2](https://doi.org/10.1016/s2589-7500(20)30292-2) EDN: [UWEZGN](#)

ОБ АВТОРАХ/ AUTHORS' INFO

Автор, ответственный за переписку:	
* Нанова Ольга Геннадьевна , канд. биол. наук; адрес: Россия, 127051, Москва, ул. Петровка, д. 24, стр. 1; ORCID: 0000-0001-8886-3684; eLibrary SPIN: 6135-4872; e-mail: nanova@mail.ru	* Olga G. Nanova , Cand. Sci. (Biology); address: 24 Petrovka st, bldg 1, Moscow, Russia, 127051; ORCID: 0000-0001-8886-3684; eLibrary SPIN: 6135-4872; e-mail: nanova@mail.ru
Соавторы (должны быть приведены в порядке их перечисления в списке авторов рукописи):	
Васильев Юрий Александрович , канд. мед. наук; ORCID: 0000-0002-5283-5961; eLibrary SPIN: 4458-5608; e-mail: npcmr@zdrav.mos.ru	Yuriy A. Vasilev , MD, Cand. Sci. (Medicine); ORCID: 0000-0002-5283-5961; eLibrary SPIN: 4458-5608; e-mail: npcmr@zdrav.mos.ru
Владзимирский Антон Вячеславович , д-р мед. наук; ORCID: 0000-0002-2990-7736; eLibrary SPIN: 3602-7120; e-mail: VladzimirskijAV@zdrav.mos.ru	Anton V. Vladzymyrskyy , MD, Dr. Sci. (Medicine); ORCID: 0000-0002-2990-7736; eLibrary SPIN: 3602-7120; e-mail: VladzimirskijAV@zdrav.mos.ru
Гольдберг Аркадий Станиславович , канд. мед. наук;	Arcadiy S. Goldberg , MD, Cand. Sci. (Medicine);

EDN: BXDWFO

ORCID: 0000-0002-2787-4731; eLibrary SPIN: 8854-0469; e-mail: goldarcadiy@gmail.com	ORCID: 0000-0002-2787-4731; eLibrary SPIN: 8854-0469; e-mail: goldarcadiy@gmail.com
Блохин Иван Андреевич , канд. мед. наук; ORCID: 0000-0002-2681-9378; eLibrary SPIN: 3306-1387; e-mail: i.blokhin@npcmr.ru	Ivan A. Blokhin , MD, Cand. Sci. (Medicine); ORCID: 0000-0002-2681-9378; eLibrary SPIN: 3306-1387; e-mail: i.blokhin@npcmr.ru
Решетников Роман Владимирович , канд. физ.-мат. наук; ORCID: 0000-0002-9661-0254; eLibrary SPIN: 8592-0558; e-mail: ReshetnikovRV1@zdrav.mos.ru	Roman V. Reshetnikov , Cand. Sci. (Physics and Mathematics); ORCID: 0000-0002-9661-0254; eLibrary SPIN: 8592-0558; e-mail: ReshetnikovRV1@zdrav.mos.ru

Accepted for publication

ПРИЛОЖЕНИЯ

Приложение 1. Список публикаций, включённых в систематический обзор, и их характеристики.

Supplement 1. List of publications included in the systematic review and their characteristics.

Авторы	Год	Название работы	Название журнала	Импакт-фактор журнала	Этап лабораторного анализа	Задача исследования	Страна исследования
C.J. Farrell [7]	2021	Identifying mislabelled samples: machine learning models exceed human performance	<i>Annals of Clinical Biochemistry</i>	2,30	Преаналитический	Сравнение производительности 8 моделей машинного обучения и лабораторного персонала при выявлении смоделированных неправильно маркированных образцов	Австралия
S. Kurstjens и соавт. [12]	2022	Automated prediction of low ferritin concentrations using a machine learning algorithm	<i>Clinical Chemistry and Laboratory Medicine (CCLM)</i>	5,10	Аналитический	Диагностика анемии — автоматическое прогнозирование низкого содержания ферритина (дополнительно определения витамина B12 и фолиевой кислоты)	Нидерланды
R.J. Burton и соавт. [23]	2019	Using artificial intelligence to reduce diagnostic workload without compromising detection of urinary tract infections	<i>BMC Medical Informatics and Decision Making</i>	9,30	Аналитический	Детекция образцов мочи для выявления потребности в последующем посеве с целью выявления патогенов	Великобритания
Е.А. Елагина и соавт. [15]	2021	Исследование методов машинного обучения в задаче идентификации клеток крови	<i>Научно-технический вестник информационных технологий, механики и оптики</i>	РИНЦ 0,39	Аналитический	Исследовать применение методов машинного обучения к задаче идентификации и классификации клеток крови для увеличения скорости распознавания без ухудшения качества	Россия
H. Ayyıldız и соавт. [16]	2021	Is it possible to determine antibiotic resistance of <i>E. coli</i> by analyzing laboratory data with machine learning?	<i>Turkish Journal of Biochemistry</i>	0,70	Аналитический	Оценка устойчивости кишечной палочки к антибиотикам (19 видов антибиотиков) с помощью ИИ моделей машинного обучения	Турция
D. Lea и соавт. [8]	2021	Digital image analysis of the proliferation markers Ki67 and phosphohistone H3 in gastroenteropancreatic neuroendocrine neoplasms: accuracy of grading compared with routine manual hot spot evaluation of the Ki67 index	<i>Applied Immunohistochemistry & Molecular Morphology</i>	1,60	Аналитический	Диагностика и стадирование опухолевых заболеваний	Норвегия
M.B. Wallace и соавт. [26]	2022	Impact of artificial intelligence on miss rate of colorectal neoplasia	<i>Gastroenterology</i>	29,40	Аналитический	Изучить возможное снижение уровня ложноотрицательных результатов обнаружения кишечных неоплазий с помощью технологий ИИ	США, Великобритания, Италия, Ирландия

Digital Diagnostics | Digital Diagnostics
Систематические обзоры | Systematic reviews
DOI: <https://doi.org/10.17816/DD635349>

EDN: BXDWFO

Авторы	Год	Название работы	Название журнала	Импакт-фактор журнала	Этап лабораторного анализа	Задача исследования	Страна исследования
M.E. Lemieux и соавт. [9]	2023	Detection of early-stage lung cancer in sputum using automated flow cytometry and machine learning	<i>Respiratory Research</i>	5,90	Аналитический	Диагностика опухолевых заболеваний лёгких	Канада, США
H. Kim и соавт. [14]	2023	Performance of digital morphology analyzer Medica EasyCell assistant	<i>Clinical Chemistry and Laboratory Medicine (CCLM)</i>	5,10	Аналитический	Идентификация и подсчёт клеток крови для тестирования и сравнительного анализа разных методов	Южная Корея
T.T. Van и соавт. [17]	2019	Automated detection of <i>Streptococcus pyogenes</i> pharyngitis by use of colorex Strep A CHROMagar and WASPLab artificial intelligence chromogenic detection module software	<i>Journal of Clinical Microbiology</i>	4,96	Аналитический	Детекции стрептококков группы А	США
S. Yoon и соавт. [11]	2021	Performance of digital morphology analyzer Vision Pro on white blood cell differentials	<i>Clinical Chemistry and Laboratory Medicine (CCLM)</i>	5,10	Аналитический	Дифференциация лейкоцитов	Южная Корея
M. Wang и соавт. [13]	2022	A deep learning model for the automatic recognition of aplastic anemia, myelodysplastic syndromes, and acute myeloid leukemia based on bone marrow smear	<i>Frontiers in Oncology</i>	4,70	Аналитический	Диагностика гематологических заболеваний	Китай
K. Kimura и соавт. [10]	2019	A novel automated image analysis system using deep convolutional neural networks can assist to differentiate MDS and AA	<i>Scientific Reports</i>	4,60	Аналитический	Классификация клеток крови (17 типов) и диагностика гематологических заболеваний	Япония
M.L. Faron и соавт. [18]	2020	Evaluation of the WASPLab segregation software to automatically analyze urine cultures using routine blood and MacConkey agars	<i>Journal of Clinical Microbiology</i>	1,80	Аналитический	Сегрегация патоген-положительных и -отрицательных посевов мочи	США
M. Yang и соавт. [19]	2020	A CNN-based active learning framework to identify mycobacteria indigitized Ziehl–Neelsen stained human tissues	<i>Computerized Medical Imaging and Graphics</i>	2,40	Аналитический	Идентификация микобактерий в тканях человека	США, Польша
S. Zuras и соавт. [20]	2022	A new artificial intelligence-based method for identifying mycobacterium tuberculosis in Ziehl–Neelsen stain on tissue	<i>Diagnostics</i>	2,49	Аналитический	Идентификация микобактерий в тканях человека, диагностика туберкулёза	Румыния
B.A. Mathison и соавт. [25]	2020	Detection of intestinal protozoa in trichrome-stained stool specimens by use of a deep convolutional neural network	<i>Journal of Clinical Microbiology</i>	1,80	Аналитический	Детекция кишечных простейших	США
Z. Wang и соавт. [21]	2021	Deep neural networks offer morphologic classification and diagnosis of bacterial vaginosis	<i>Journal of Clinical Microbiology</i>	1,80	Аналитический	Диагностика вагинитов, детекция и классификация трёх категорий по шкале Ньюджента	США

EDN: BXDWFO

Авторы	Год	Название работы	Название журнала	Импакт-фактор журнала	Этап лабораторного анализа	Задача исследования	Страна исследования
A. Lev-Sagie и соавт. [22]	2023	Diagnostic performance of an automated microscopy and pH test for diagnosis of vaginitis	<i>NPJ Digital Medicine</i>	1,80	Аналитический	Диагностика вагинитов	Китай, США
D. Avci и соавт. [24]	2023	A new super resolution Faster R-CNN model based detection and classification of urine sediments	<i>Biocybernetics and Biomedical Engineering</i>	15,20	Аналитический	Диагностика инфекций мочевыводящих путей и заболевания почек с помощью микроскопии мочи	Израиль
Y. Liang и соавт. [27]	2022	A study on quality control using delta data with machine learning technique	<i>Heliyon</i>	4,00	Постаналитический	Повышение контроля качества, разработка и оценка эффективности нового протокола обеспечения стабильности данных путём совместного Δ-анализа и машинного обучения	Китай
R. Zhou и соавт. [28]	2022	A multi-model fusion algorithm as a real-time quality control tool for small shift detection	<i>Computers in Biology and Medicine</i>	7,7	Постаналитический	Разработка метода машинного обучения для автоверификации системы биохимических анализов	Италия, Китай
H. Wang и соавт. [29]	2021	Using machine learning to develop an autoverification system in a clinical biochemistry laboratory	<i>Clinical Chemistry and Laboratory Medicine (CCLM)</i>	5,1	Постаналитический	Разработка системы автоверификации на основе методов ИИ для биохимических отчетов	Китай

Примечание. РИНЦ — российский индекс научного цитирования; США — Соединённые Штаты Америки.

Приложение 2. Список публикаций, исключённых из систематического обзора

Supplement 2. List of publications excluded from the systematic review

Публикация	Причина исключения
Karsenti D, Tharsis G, Perrot B, et al. Effect of real-time computer-aided detection of colorectal adenoma in routine colonoscopy (COLO-GENIUS): a single-centre randomised controlled trial. <i>The Lancet Gastroenterology & Hepatology</i> . 2023;8(8):726–734. doi: 10.1016/s2468-1253(23)00104-8 EDN: ORVAVB	Нет открытого доступа к полному тексту статьи
Lin A, van Diemen PA, Motwani M, et al. Machine learning from quantitative coronary computed tomography angiography predicts fractional flow reserve-defined ischemia and impaired myocardial blood flow. <i>Circulation: Cardiovascular Imaging</i> . 2022;15(10):e014369. doi: 10.1161/circimaging.122.014369 EDN: KVMGRT	Не лабораторная медицина
Lui TKL, Hang DV, Tsao SKK, et al. Computer-assisted detection versus conventional colonoscopy for proximal colonic lesions: a multicenter, randomized, tandem-colonoscopy study. <i>Gastrointestinal Endoscopy</i> . 2023;97(2):325–334.e1. doi: 10.1016/j.gie.2022.09.020 EDN: DUZTNI	Нет открытого доступа к полному тексту статьи
Lipkin I, Telluri A, Kim Y, et al. Coronary CTA With AI-QCT interpretation: comparison with myocardial perfusion imaging for detection of obstructive stenosis using invasive angiography as reference standard. <i>American Journal of Roentgenology</i> . 2022;219(3):407–419. doi: 10.2214/AJR.21.27289 EDN: UDABGS	Не лабораторная медицина
Zad Z, Jiang VS, Wolf AT, et al. Predicting polycystic ovary syndrome with machine learning algorithms from electronic health records. <i>Frontiers in Endocrinology</i> . 2024;15:1298628. doi: 10.3389/fendo.2024.1298628 EDN: JSMNVT	Не лабораторная медицина
Laurie MA, Zhou SR, Islam MT, et al. Bladder cancer and artificial intelligence. <i>Urologic Clinics of North America</i> . 2024;51(1):63–75. doi: 10.1016/j.ucl.2023.07.002 EDN: LOOCOU	Обзор
Saputra DCE, Sunat K, Ratnaningsih T. A New Artificial intelligence approach using extreme learning machine as the potentially effective model to predict and analyze the diagnosis of anemia. <i>Healthcare</i> . 2023;11(5):697. doi: 10.3390/healthcare11050697	Не использовали компьютерное зрение
Ward PK, Roose S, Ayana M, et al. A comprehensive evaluation of an artificial intelligence based digital pathology to monitor large-scale deworming programs against soil-transmitted helminths: a study protocol. <i>PLoS ONE</i> . 2024;19(10):e0309816. DOI: 10.1371/journal.pone.0309816 EDN: RNWSWB	Протокол планируемого исследования
Wojtas N, Wiczorek M, Belkot Z. Malaria detection using custom semantic segmentation neural network architecture. <i>Medycyna Weterynaryjna</i> . 2023;79(8):406–412.	Ветеринарная медицина

Digital Diagnostics | Digital Diagnostics
Систематические обзоры | Systematic reviews
DOI: <https://doi.org/10.17816/DD635349>

EDN: BXDWFO

Публикация	Причина исключения
doi: 10.21521/mw.6804 EDN: CQNCPCG	
Eckardt JN, Hahn W, Röhlig C, et al. Mimicking clinical trials with synthetic acute myeloid leukemia patients using generative artificial intelligence. <i>Blood</i> . 2023;142(Suppl. 1):2268–2268. doi: 10.1182/blood-2023-179817 EDN: ZDJRHJ	Графическое резюме
Lorenzon AR, Jacobs C, Milanezi de Souza RC, et al. Non-invasive method for blastocyst ploidy prediction based on artificial intelligence, time-lapse images and patient parameters. <i>Fertility and Sterility</i> . 2021;116(3):e102. doi: 10.1016/j.fertnstert.2021.07.285	Резюме
El Emrani S, Abdoun O. Artificial neural network for thyroid disease diagnosis In: Ezziyanni M, Kacprzyk J, Balas VE, editors. <i>Proceedings of the International Conference on Advanced Intelligent Systems for Sustainable Development (AIISSD 2023). Lecture Notes in Networks and Systems</i> ; 2024. Cham: Springer; 2024. doi: 10.1007/978-3-031-52385-4_25	Нет открытого доступа к полному тексту материалов конференции
Barouqa MI, Gosnell HL, Gil MR. Preoperative determination of red blood cell transfusion for orthotopic liver transplantation in the age of artificial intelligence. <i>American Journal of Clinical Pathology</i> . 2023;160(Suppl. 1):S138. doi: 10.1093/ajcp/aqad150.298	Резюме
Montjean D, Godin-Pagé MH, Benkhalifa M, Miron P. P-312 Automated sperm selection software (SiD) is as good as human selection for ICSI. <i>Human Reproduction</i> . 2023;38(Suppl. 1):dead093.670. doi: 10.1093/humrep/dead093.670 EDN: ESAFHB	Резюме
Marzizarani SB, Zolfaghari M, Sajedi H. IoMT-based automated leukemia classification using CNN and higher order singular value decomposition. In: <i>Proceedings of the 9th International Conference on Web Research (ICWR)</i> . Tehran; 2023. P. 317–321. doi: 10.1109/ICWR57742.2023.10139301	Материалы конференции, недостаточно данных для включения в анализ
Monteiro ACB, Iano Y, França RP, Arthur R. A Metaheuristic algorithm for classification of white blood cells in healthcare informatics. In: Oliva D, Houssein EH, Hinojosa S, editors. <i>Metaheuristics in machine learning: theory and applications. Studies in computational intelligence</i> . Cham: Springer; 2021. P. 219–238. doi: 10.1007/978-3-030-70542-8_10	Нет открытого доступа к полному тексту главы книги
Lin G, Wang H, Sha M, et al. Design of a multi-data fusion intelligent venipuncture blood sampling robot. In: <i>Proceedings of the 7th International Conference on Control and Robotics Engineering (ICCRE)</i> . Beijing, 2022 Apr 15–17. Beijing: IEEE; 2022. P. 10–15. doi: 10.1109/ICCRE55123.2022.9770260	Нет открытого доступа к полному тексту материалов конференции
Tian Y, Zhao D, Wang T. An improved YOLO Nano model for dorsal hand vein detection system. <i>Medical & Biological Engineering & Computing</i> . 2022;60(5):1225–1237. doi: 10.1007/s11517-022-02551-x EDN: YAEQTN	Нет открытого доступа к полному тексту статьи
Xia W, Lu Z, Huang Y, et al. MAGIC: Manifold and graph integrative convolutional network for low-dose CT reconstruction. <i>IEEE Transactions on Medical Imaging</i> . 2021;40(12):3459–3472. doi: 10.1109/tmi.2021.3088344 EDN: LIKHWV	Разработка метода
Liu Q, Yu L, Luo L, et al. Semi-supervised medical image classification with relation-driven self-ensembling model. <i>IEEE Transactions on Medical Imaging</i> . 2020;39(11):3429–3440. doi: 10.1109/tmi.2020.2995518 EDN: HCUMAB	Разработка метода
Zhou R, Liang Y, Cheng HL, et al. A highly accurate delta check method using deep learning for detection of sample mix-up in the clinical laboratory. <i>Clinical Chemistry and Laboratory Medicine (CCLM)</i> . 2021;60(12):1984–1992. doi: 10.1515/cclm-2021-1171 EDN: GZTHYE	Разработка метода
Farrell CJ, Makuni C, Keenan A, et al. A machine learning model for the routine detection of “wrong blood in complete blood count tube” errors. <i>Clinical Chemistry</i> . 2023;69(9):1031–1037. doi: 10.1093/clinchem/hvad100 EDN: YKNRVI	Нет открытого доступа к полному тексту статьи
Jawahar M, H S, L JA, Gandomi AH. ALNett: a cluster layer deep convolutional neural network for acute lymphoblastic leukemia classification. <i>Computers in Biology and Medicine</i> . 2022;148:105894. doi: 10.1016/j.compbiomed.2022.105894 EDN: QPYXCEV	Нет открытого доступа к полному тексту статьи
Lyu H, Xu F, Jin T, et al. Automated detection of multi-class urinary sediment particles: an accurate deep learning approach. <i>Biocybernetics and Biomedical Engineering</i> . 2023;43(4):672–683. doi: 10.1016/j.bbe.2023.09.003 EDN: XGGDBL	Нет открытого доступа к полному тексту статьи
Li Q, Yu Z, Qi T, et al. Inspection of visible components in urine based on deep learning. <i>Medical Physics</i> . 2020;47(7):2937–2949. doi: 10.1002/mp.14118 EDN: AWIHBC	Обзор метода
Yildirim M, Bingol H, Cengil E, et al. Automatic classification of particles in the urine sediment test with the developed artificial intelligence-based hybrid model. <i>Diagnostics</i> . 2023;13(7):1299. doi: 10.3390/diagnostics13071299 EDN: YNHFGI	Исследование в области радиомикки
Liang Y, Padoan A, Wang Z, et al. Machine learning-based nonlinear regression-adjusted real-time quality control modeling: a multi-center study. <i>Clinical Chemistry and Laboratory Medicine (CCLM)</i> . 2023;62(4):635–645. doi: 10.1515/cclm-2023-0964 EDN: YOGOIK	Нет открытого доступа к полному тексту статьи
Hatami B, Asadi F, Bayani A, et al. Machine learning-based system for prediction of ascites grades in patients with liver cirrhosis using laboratory and clinical data: design and implementation study. <i>Clinical Chemistry and Laboratory Medicine (CCLM)</i> . 2022;60(12):1946–1954. doi: 10.1515/cclm-2022-0454 EDN: DIXDIG	Нет открытого доступа к полному тексту статьи
Lu XJ, Yang XJ, Sun JY, et al. FibroBox: a novel noninvasive tool for predicting significant liver fibrosis and cirrhosis in HBV infected patients. <i>Biomarker Research</i> . 2020;8(1):48. doi: 10.1186/s40364-020-00215-2 EDN: BDTCRW	Не лабораторная медицина

EDN: BXDWFO

Публикация	Причина исключения
Kate RJ, Pearce N, Mazumdar D, Nilakantan V. A continual prediction model for inpatient acute kidney injury. <i>Computers in Biology and Medicine</i> . 2020;116:103580. doi: 10.1016/j.combiomed.2019.103580 EDN: FCQYFN	Нет открытого доступа к полному тексту статьи
Steinbach D, Ahrens PC, Schmidt M, et al. Applying machine learning to blood count data predicts sepsis with ICU admission. <i>Clinical Chemistry</i> . 2024;70(3):506–515. doi: 10.1093/clinchem/hvae001 EDN: GLTLTL	Не компьютерное зрение
Ng DP, Simonson PD, Tarnok A, et al. Recommendations for using artificial intelligence in clinical flow cytometry. <i>Cytometry Part B: Clinical Cytometry</i> . 2024;106(4):228–238. doi: 10.1002/cyto.b.22166 EDN: VONANS	Обзор метода
Suster CJE, Pham D, Kok J, Sintchenko V. Emerging applications of artificial intelligence in pathogen genomics. <i>Frontiers in Bacteriology</i> . 2024;3:1326958. doi: 10.3389/fbri.2024.1326958 EDN: IUCEPH	Литературный обзор
Llamas-Velasco M, Ovejero-Merino E. Artificial intelligence in dermatopathologic diagnosis. <i>Piel</i> . 2024;39(8):512–517. doi: 10.1016/j.piel.2024.01.002 EDN: GLWBYS	Литературный обзор
Yu Z, Li J, Wen X, et al. Retracted article: AMLnet, a deep-learning pipeline for the differential diagnosis of acute myeloid leukemia from bone marrow smears. <i>Journal of Hematology & Oncology</i> . 2023;16(1):27. doi: 10.1186/s13045-023-01419-3 EDN: IKGJCK	Публикацию отозвал главный редактор журнала 01.08.2024, поскольку авторы исследования не предоставили документы, подтверждающие одобрение локально этическим комитетом

Приложение 3. Основные характеристики исследований, представленных в публикациях, включённых в систематический обзор

Supplement 3. Main characteristics of the studies presented in the publications included in the systematic review

Авторы	Область исследования	Используемые лабораторные методы	Время сбора исследуемого материала	Критерии включения	Критерии исключения	Дизайн исследования	Внешняя валидация
C.J. Farrell [7]	контроль качества — ошибка маркировки пробирок	• обезличенные анализы мочи; • симуляция ошибок маркировки на уровне 50%; • метод Δ-проверки	с 01.06.2019 по 26.10.2020	—	—	• одноцентровое; • ретроспективное	нет
S. Kurstjens и соавт. [12]	цитология	• ОАК — СРБ	2021 год	• взрослые с анемией	• недостаток данных; • возраст младше 18 лет	• многоцентровое; • ретроспективное с проспективным тестом в лаборатории (1 мес.)	да
R.J. Burton и соавт. [23]	цитология и микробиология	• микроскопия мочи — подсчёт лейкоцитов, эритроцитов, бактерий	2016–2017 гг.	—	• образцы, подвергшиеся ручной микроскопии; • образцы полученные от пациентов при катетеризации	• одноцентровое; • ретроспективное	нет
Е.А. Елагина и соавт. [15]	цитология	• анализ крови	—	• пациенты с различными типами анемии; • возраста — 15–41 лет	—	• многоцентровое; • ретроспективное	нет
Н. Ayıldız и соавт. [16]	микробиология	• ОАК; • общий анализ мочи	2019–2020 гг.	• пациенты с резистентными и/или рецидивирующими инфекциями мочевыводящих путей	—	• одноцентровое; • ретроспективное	нет
D. Lea и соавт. [8]	цитология	• анализ крови — подсчёт окрашенных клеток крови	2003–2013 гг.	• пациенты с диагностированными опухолевыми заболеваниями,	• наличие у пациентов смешанных нейроэндокринных или	• одноцентровое; • ретроспективное	нет

EDN: BXDWFO

Авторы	Область исследования	Используемые лабораторные методы	Время сбора исследуемого материала	Критерии включения	Критерии исключения	Дизайн исследования	Внешняя валидация
		(Ki-67)		проходившие лечение в больнице при университете Ставангера	• ненейроэндокринных новообразований; • недостаток опухолевого материала		
M.B. Wallace и соавт. [26]	гистопатология	• колоноскопия; • эндоскопия; • биопсия; • гистопатология	—	• пациенты в возрасте 45 лет и старше, проходящие скрининг на злокачественные новообразования кишечника	• беременные или женщины, планирующие беременность; • пациенты с воспалительными заболеваниями кишечника или после резекции толстой кишки; • пациенты с противопоказаниями к колоноскопии; • пациенты с симптомами, требующими выборочной биопсии толстой кишки	• многоцентровое; • ретроспективное	нет
M.E. Lemieux и соавт. [9]	цитология	• анализ мокроты методом проточной цитометрии	с апреля 2018 г. по ноябрь 2019 г.; с июля 2020 г. по ноябрь 2021 г.	• группа пациентов без опухолевых заболеваний — возраст 52–79 лет; ИКЧ <20; • группа пациентов с подтверждёнными опухолевыми заболеваниями лёгкого — диагноз подтверждали биопсией после сдачи образца мокроты, не было ограничений по возрасту или истории курения	—	• многоцентровое; • ретроспективное	нет
H. Kim и соавт. [14]	цитология	• ОАК — подсчёт лейкоцитов и тромбоцитов	—	• нормальные образцы — здоровые люди, выявленные при обычном физикальном осмотре; • патологические образцы — пациенты с патологическими результатами ОАК и/или гематологическими расстройствами разной степени выраженности	—	• одноцентровое; • ретроспективное	нет
T.T. Van и соавт. [17]	микробиология	• посев слизи из ротоглотки на агаре и кровяном агаре	с сентября 2017 г. по январь 2019 г.	• пациенты, поступившие в отделение неотложной помощи детской больницы Лос-Анджелеса с предположительным диагнозом бактериальный фарингит	—	• одноцентровое; • ретроспективное	нет
S. Yoon и соавт. [11]	цитология	• ОАК — подсчёт лейкоцитов	с августа 2017 г. по ноябрь 2018 г.	• нормальные образцы — здоровые люди, выявленные при обычном физикальном осмотре; • патологические образцы — пациенты с патологическими результатами ОАК и/или гематологическими расстройствами разной степени выраженности	—	• одноцентровое; • ретроспективное	нет

EDN: BXDWFO

Авторы	Область исследования	Используемые лабораторные методы	Время сбора исследуемого материала	Критерии включения	Критерии исключения	Дизайн исследования	Внешняя валидация
M. Wang и соавт. [13]	цитология	• мазок костного мозга — идентификации клеток	—	• пациенты с одним из подтверждённых диагнозов — миелодиспластический синдром, миелонный лейкоз, апластическая анемия	—	• многоцентровое; • ретроспективное	да
K. Kimura и соавт. [10]	цитология	• анализ крови — определение дисморфных клеток в мазках периферической крови	2017–2018 гг.	—	—	• одноцентровое; • ретроспективное	нет
M.L. Faron и соавт. [18]	микробиология	• посев мочи на агаре и агаре МакКонки — подсчёт колоний без определения морфологии клеток	—	—	—	• многоцентровое; • ретроспективное	нет
M. Yang и соавт. [19]	микробиология	• анализ тканей человека — окрашивание по Цилю–Нильсену	—	—	—	• одноцентровое; • ретроспективное	нет
S. Zurac и соавт. [20]	микробиология	• анализ тканей человека — окрашивание по Цилю–Нильсену	—	—	—	• ретроспективное	нет
B.A. Mathison и соавт. [25]	паразитология	• окрашенные трихромом мазки кала	—	• образцы, содержащие паразитов; • сканы с площадью в фокусе 80% и более	• некачественные сканы с площадью не в фокусе более 20%	• одноцентровое; • ретроспективное	нет
Z. Wang и соавт. [21]	микробиология	• вагинальные мазки — оценка по шкале Ньюджента	с января 2018 г. по сентябрь 2019 г.	—	—	• многоцентровое; • ретроспективное	нет
A. Lev-Sagie и соавт. [22]	микробиология	• микроскопия	—	• женщины с вагинальными жалобами — выделения, неприятный запах, зуд, жжение, боль или сухость; • возраст — 18 лет и старше	• пациенты без информированного согласия; • неинтерпретируемые образцы	• одноцентровое; • проспективное	нет
D. Avcı и соавт. [24]	цитология и микробиология	• детекция и подсчёт элементов в осадке мочи (10 элементов разного размера): эритроциты, лейкоциты, клетки эпителия, кристаллы, бактерии, грибы, паразитические организмы, сперматозоиды, посторонние загрязнения	—	—	—	• ретроспективное	нет
Y. Liang и соавт. [27]	контроль качества — цитология	• анализ крови — подсчёт клеток крови; • симуляция данных	2019 год	—	• пациенты, имеющие менее двух последовательных результатов	• многоцентровое; • ретроспективное	нет
R. Zhou и соавт. [28]	контроль качества — биохимия	• содержание простатоспецифического антигена; • симуляция данных	с января 2016 г. по июнь 2021 г.	—	—	• одноцентровое; • ретроспективное	нет

EDN: BXDWFO

Авторы	Область исследования	Используемые лабораторные методы	Время сбора исследуемого материала	Критерии включения	Критерии исключения	Дизайн исследования	Внешняя валидация
Н. Wang и соавт. [29]	контроль качества — биохимия	• симуляция данных	с июня 2018 г. по август 2019 г.	—	—	• одноцентровое; • ретроспективное	нет

Примечание. СРБ — С-реактивный белок; Ki-67 — индекс пролиферативной активности опухолевой клетки; ИКЧ — индекс курящего человека; ОАК — общий анализ крови.

Приложение 4. Характеристики выборок, используемые модели машинного обучения или готовые коммерческие решения, представленные в исследованиях

Supplement 4. Characteristics of samples and machine learning models used, or commercially available solutions presented in the studies

Авторы	Объём выборки	Возраст, лет	Пол, мужчины и женщины	Представленность патологии в выборке	Используемые модели машинного обучения или готовые коммерческие решения	Сравнение моделей между собой	Сравнение моделей с врачами
C.J. Farrell [7]	141 396 образцов	—	—	—	8 моделей в R-среде: • модель нейронной сети; • модель экстремального градиентного бустинга; • модель опорных векторов; • модель «случайного леса»; • модель логистической регрессии; • модель k-ближайших соседей; • модели дерева принятия решений (2 модели)	да	да
S. Kurstjens и соавт. [12]	отчёты трёх лабораторий: • n=3797; • n=8021; • n=191	• медиана — 74 (среднее — 68); • медиана — 73 (среднее — 74) • медиан — 73 (среднее — 74)	• 48 и 52%; • 49 и 51%; • 57 и 43%	• анемия — 100%	Phython скрипт для машинного обучения с использованием метода «случайного леса»: • модель JBN-S; • модель Medlon-R	да	да
R.J. Burton и соавт. [23]	212 554 образца мочи	—	54 070 и 158 422 — 1:3	—	три модели, разработанные в Phython: • модель «случайного леса»; • модель однослойной нейронной сети; • модель экстремального градиентного бустинга	да	нет
Е.А. Елагина и соавт. [15]	изображения: • тромбоцитов — 1000; • лейкоцитов — 1000; • моноцитов — 1000	15–41	25 и 165	• анемия — 100%	• модель опорных векторов; • модель k-ближайших соседей; • модель свёрточной нейронной сети; • модель нейронной сети с прямым распространением	да	нет
Н. Ayyıldız и соавт. [16]	103 пациента	1–93	33 и 70	• инфекции мочевыводящих путей — 100%	• модель k-ближайших соседей; • модель нейронной сети; • модель опорных векторов; • модель дерева принятия решений	да	нет

EDN: BXDWFO

Авторы	Объём выборки	Возраст, лет	Пол, мужчины и женщины	Представленность патологии в выборке	Используемые модели машинного обучения или готовые коммерческие решения	Сравнение моделей между собой	Сравнение моделей с врачами
D. Lea и соавт. [8]	159 пациентов: • 500–2000 раковых клеток в точках анализа при микроскопии	61,8 (ДИ 12,5–94,2)	91 (57,2%) и 68 (42,8%)	—	готовое коммерческое решение — Image Analysis Software® (Visiopharm, Дания)	нет	да
M.B. Wallace и соавт. [26]	230 пациентов	старше 45	157 (68,26%) и 73 (31,74%)	—	готовый модуль GI-Genius® (Medtronic, США), созданный с помощью свёрточной нейронной сети	нет	нет
M.E. Lemieux и соавт. [9]	182 пациента	• отсутствие опухолевых заболеваний — 65 (53–75); • опухолевые заболевания — 73 (54–79)	• отсутствие опухолевых заболеваний — 57 (46,7%) и 65 (53,3%); • опухолевые заболевания — 21 (75,0%) и 7 (25,0%)	—	готовое коммерческое решение для раннего выявления рака лёгких CyPath® Lung (bioAffinity Technologies, Inc., США)	да	нет
H. Kim и соавт. [14]	225 образцов	—	—	• норма — 100; • лимфома — 10; • ОМЛ — 18; • МПН — 6; • ПМ — 7; • лейкопения разной выраженности — 45	готовое коммерческое решение — гематологическая система EasyCell® assistant (Medica Corporation, США)	нет	нет
T.T. Van и соавт. [17]	480 пациентов	—	—	• бактериальный фарингит — 192	готовое коммерческое решение на основе искусственного интеллекта — программное обеспечение с модулем хромогенного обнаружения PhenoMATRIX® (Copan, США)	нет	да
S. Yoop и соавт. [11]	200 образцов	• здоровые — 49 (18–78); • пациенты с патологическими образцами — 61 (15–88)	—	• норма — 100; • ОМЛ — 21; • МДС — 13; • МПН — 11; • ПМ — 15; • лейкопения разной степени — 36; • ОЛЛ — <5; • ОНЛ — <5; • ОЛСФ — <5; • ХЛЛ — <5; • ХМЛ — <5; • ЛПЛ — <5; • АЦ — <5; • ЗЛ — <5	готовое коммерческое решение — цифровой морфологический анализатор Vision Pro® (West Medica, Австрия)	нет	да

Авторы	Объём выборки	Возраст, лет	Пол, мужчины и женщины	Представленность патологии в выборке	Используемые модели машинного обучения или готовые коммерческие решения	Сравнение моделей между собой	Сравнение моделей с врачами
M. Wang и соавт. [13]	547 образцов	—	—	• МДС — 246; • АА — 141; • ОМЛ — 160	модель свёрточной нейронной сети Resnet 50	нет	нет
K. Kimura и соавт. [10]	3261 образцов — 695 030 изображений	—	—	• МДС — 94; • МПН — 127; • ОМЛ — 38; • ОЛЛ — 27; • лимфомы — 324; • ММ — 82; • АА — 42	система распознавания изображений клеток с модулем на основе свёрточной нейронной сети и модулем классификации	нет	нет
M.L. Faron и соавт. [18]	12 931 образцов	—	—	• инфекции мочевыводящих путей — 43,9%	готовое коммерческое решение — программное обеспечение на основе искусственного интеллекта WASPLab® (Cora, Италия)	нет	да
M. Yang и соавт. [19]	167 образцов	—	—	• туберкулёз — 51 (30,5%)	• две модели свёрточной нейронной сети; • модель логистической регрессии	нет	нет
S. Zurac и соавт. [20]	510 изображений — >260 тыс. положительных участков и >7 млн отрицательных	медиана 42,25 (от 1 до 80)	37 (61,67%) и 23 (38,33%)	• туберкулёз — 23 (38,33%); • воспаления другой этиологии — 37 (61,67%)	модель глубокой нейронной сети для обнаружения микобактерий	нет	да
B.A. Mathison и соавт. [25]	1394–23 566 экземпляров с различными патогенами	—	—	• протозойные кишечные инфекции: <i>G. duodenalis</i> и цисты — 23; <i>G. duodenalis</i> и трофозоиты — 21; <i>Blastocystis sp.</i> — 61; <i>D. fragilis</i> — 29; <i>E. hartmanni</i> — 10; <i>Entamoeba spp.</i>, non-hartmanni — 34); <i>E. coli</i> — 21; <i>E. histolytica/E. dispar</i> — 10; <i>E. polecki</i> — 1; <i>Entamoeba sp.</i> — 3; <i>C. mesnili</i> — 15; <i>E. nana/I. buetschlii</i> — 36	модель свёрточной нейронной сети	нет	нет
Z. Wang и соавт. [21]	29 095 изображений	—	женщины — 100%	• нормальная флора влагалища — 16 490;	модель свёрточной нейронной сети	нет	да

Авторы	Объём выборки	Возраст, лет	Пол, мужчины и женщины	Представленность патологии в выборке	Используемые модели машинного обучения или готовые коммерческие решения	Сравнение моделей между собой	Сравнение моделей с врачами
				• изменённая флора влагалища — 4660; • бактериальный вагинит — 2130			
A. Lev-Sagie и соавт. [22]	226 пациентов	—	женщины — 100%	• кандидозный вагинит (<i>Candida albicans</i>) — (32%); • бактериальный вагинит — 16%; • вагинальная атрофия — 12%; • отсутствие диагноза — 8%; • другие состояния — 13%; аэробные вагиниты — 12%; • цитолитический вагинит — 7%	модель свёрточной нейронной сети для классификации изображений с множественными метками	нет	да
D. Avci и соавт. [24]	25 771 изображений	—	—	—	модель свёрточной нейронной сети	да	нет
Y. Liang и соавт. [27]	423 290 лабораторных результатов	от 14 до 60	—	—	модель «случайного леса»	да	нет
R. Zhou и соавт. [28]	84 241 лабораторных результатов	—	—	—	модели, реализованные в Python 3.7.3: • «случайного леса»; • опорных векторов; • нейронной сети	да	нет
H. Wang и соавт. [29]	196 591 образцов и 3 756 239 отчётов	—	—	—	• модель наивного байесовского классификатора; • модель k-ближайших соседей; • модель «случайного леса»; • модель экстремального градиентного бустинга	нет	нет

Примечание. JBN-S — модель, разработанная на основе данных лабораторных отчётов, полученных с помощью химического анализатора Siemens® (Siemens Healthineers, Германия); Medlon-R — модель, разработанная на основе данных лабораторных отчётов, полученных с помощью химического анализатора Roche® (Roche Diagnostics, Германия); ДИ — доверительный интервал; США — Соединённые Штаты Америки; СГА — стрептококки группы А; ОМЛ — острый миелоидный лейкоз; МПН — миелопролиферативное новообразование; ПМ — плазмноклеточная миелома; МДС — миелодиспластический синдром; ОЛЛ — острый лимфобластный лейкоз; ОНЛ — острый недифференцированный лейкоз; ОЛСФ — острый лейкоз со смешанным фенотипом; ХЛЛ — хронический лимфоцитарный лейкоз; ХМЛ — хронический миеломоноцитарный лейкоз; ЛПЛ — лимфоплазмочитарная лимфома; АЦ — аутоиммунная цитопения; ЗЛ — злокачественная лимфома; АА — апластическая анемия; ММ — множественная миелома.

Приложение 5. Эффективность искусственного интеллекта в исследованиях

Supplement 5. Effectiveness of artificial intelligence in studies

Авторы	Модель или готовое коммерческое решение	Чувствительность	Специфичность	Значение AUC	Точность
C.J. Farrell [7]	Модель нейронной сети	0,916	0,925	0,977 (95% ДИ 0,975–0,979)	0,921
	Модель экстремального градиентного бустинга	0,916	0,915	0,975 (95% ДИ 0,973–0,978)	0,915
	Модель опорных векторов	0,906	0,924	0,973 (95% ДИ 0,970–0,975)	0,915
	Модель «случайного леса»	0,905	0,916	0,970 (95% ДИ 0,968–0,972)	0,910
	Модель сложного дерева принятия решений	0,878	0,908	0,932 (95% ДИ 0,928–0,937)	0,894
	Модель логистической регрессии	0,862	0,917	0,956 (95% ДИ 0,953–0,959)	0,890
	Модель k-ближайших соседей	0,819	0,916	0,947 (95% ДИ 0,944–0,951)	0,869
S. Kurstjens и соавт. [12]	Модель простого дерева принятия решений	0,886	0,844	0,901 (95% ДИ 0,896–0,907)	0,865
	Модель JBN-S	0,930	0,920	0,920	—
R.J. Burton и соавт. [23]	Модель Medlon-R	0,980	0,920	0,900	—
	Эвристическая модель	0,957	0,521	—	0,639
	Модель «случайного леса»	0,960	0,634	0,908	—
	Модель нейронной сети	0,740	0,891	0,906	—
	Модель нейронной сети с ресэмплированием	0,906	0,752	0,904	—
Е.А. Елагина и соавт. [15]	Модель экстремального градиентного бустинга	0,967	0,541	0,910	—
	Модель опорных векторов (тромбоциты)	—	—	—	0,984
	Модель опорных векторов (моноциты)	—	—	—	0,882
	Модель опорных векторов (лейкоциты)	—	—	—	0,859
	Модель глубокой нейронной сети (тромбоциты)	—	—	—	0,995
	Модель глубокой нейронной сети (моноциты)	—	—	—	0,999
	Модель глубокой нейронной сети (лейкоциты)	—	—	—	0,998
	Модель k-ближайших соседей (тромбоциты)	—	—	—	0,990
	Модель k-ближайших соседей (моноциты)	—	—	—	0,710
	Модель k-ближайших соседей (лейкоциты)	—	—	—	0,800
	Модель нейронной сети прямого распространения	—	—	—	0,960
	Модель свёрточной нейронной сети	—	—	—	0,980
	Модель опорных векторов	—	—	—	0,980
Н. Аyyıldız и соавт. [16]	Модель k-ближайших соседей	—	—	—	0,920
	Модель нейронной сети (амикацин)	—	—	—	0,960
	Модель нейронной сети (ампициллин)	—	—	—	0,770
	Модель нейронной сети (цефтазидим)	—	—	—	0,620
	Модель нейронной сети (цефиксим)	—	—	—	0,630
	Модель нейронной сети (цефотаксим)	—	—	—	0,680
	Модель нейронной сети (колистиметат натрия)	—	—	—	0,950
	Модель нейронной сети (ципрофлоксацин)	—	—	—	0,760
	Модель нейронной сети (цефепим)	—	—	—	0,700
	Модель нейронной сети (эртапенем)	—	—	—	0,960
	Модель нейронной сети (нитрофурантоин)	—	—	—	0,900
	Модель нейронной сети (фосфомидин)	—	—	—	0,980

EDN: BXDWFO

Авторы	Модель или готовое коммерческое решение	Чувствительность	Специфичность	Значение AUC	Точность
	Модель нейронной сети (гентамицин)	—	—	—	0,840
	Модель нейронной сети (левофлоксацин)	—	—	—	0,980
	Модель нейронной сети (пиперацillin+[тазобактам])	—	—	—	0,920
	Модель нейронной сети (Ко-тримоксазол [сульфаметоксазол+триметоприм])	—	—	—	0,790
M.E. Lemieux и соавт. [9]	CyPath® Lung (bioAffinity Technologies, Inc., США)	0,820	0,880	0,890 (95% ДИ 0,830–0,890)	—
	CyPath® Lung (bioAffinity Technologies, Inc., США), размер узелков <20 мм	0,920	0,870	0,940 (95% ДИ 0,890–0,990)	—
	Navios EX® (Beckman Coulter, США), проточная цитометрия	0,830	0,770	0,850 (95% ДИ 0,710–0,980)	—
H. Kim и соавт. [14]	Гематологическая система EasyCell® assistant (Medica Corporation, США), первичная классификация лейкоцитов	0,850–1,000	0,908–1,000	—	—
	Гематологическая система EasyCell® assistant (Medica Corporation, США), подсчёт других типов клеток в патологических образцах	0,986	0,074 (до верификации); 0,68 (после верификации пользователем)	—	—
	Гематологическая система EasyCell® assistant (Medica Corporation, США), подсчёт ядросодержащих эритроцитов в патологических образцах	0,963	0,141 (до верификации); 0,840 (после верификации пользователем)	—	—
T.T. Van и соавт. [17]	Программное обеспечение с модулем хромогенного обнаружения PhenoMATRIX® (Copan, США)	0,906	0,940	—	—
S. Yoon и соавт. [11]	Цифровой морфологический анализатор Vision Pro® (West Medica, Австрия), сегментоядерные нейтрофилы	0,980	0,987	—	—
	Цифровой морфологический анализатор Vision Pro® (West Medica, Австрия), палочкоядерные нейтрофилы	0,892	0,991	—	—
	Цифровой морфологический анализатор Vision Pro® (West Medica, Австрия), лимфоциты	0,971	0,981	—	—
	Цифровой морфологический анализатор Vision Pro® (West Medica, Австрия), моноциты	0,841	0,995	—	—
	Цифровой морфологический анализатор Vision Pro® (West Medica, Австрия), эозинофилы	0,903	0,998	—	—
	Цифровой морфологический анализатор Vision Pro® (West Medica, Австрия), базофилы	0,801	1,000	—	—
	Цифровой морфологический анализатор Vision Pro® (West Medica, Австрия), метамиелоциты	0,505	0,999	—	—
	Цифровой морфологический анализатор Vision Pro® (West Medica, Австрия), миелоциты	0,480	0,997	—	—
	Цифровой морфологический анализатор Vision Pro® (West Medica, Австрия), бласты	0,765	0,995	—	—
	Цифровой морфологический анализатор Vision Pro® (West Medica, Австрия), ядросодержащие эритроциты	0,951	0,997	—	—
M. Wang и соавт. [13]	Модель свёрточной нейронной сети, бинарная классификация (МДС)	0,992 (95% ДИ 0,980–1,000)	0,867–0,967 (95% ДИ 0,836–0,983)	0,990 (95% ДИ 0,980–0,990)	0,914 (95% ДИ 0,895–0,934)
	Модель свёрточной нейронной сети, классификация с тремя исходами (МДС, ОМЛ, АА)	0,857 (95% ДИ 0,828–0,886)	0,880–0,986 (95% ДИ 0,868–0,990)	0,970 (95% ДИ 0,960–0,980)	0,929 (95% ДИ 0,916–0,941)

Авторы	Модель или готовое коммерческое решение	Чувствительность	Специфичность	Значение AUC	Точность
K. Kimura и соавт. [10]	Система распознавания изображений клеток с модулем на основе свёрточной нейронной сети и модулем классификации (17 типов клеток крови и 97 морфологических признаков)	>0,935	>0,960	—	—
	Модель для диагностики МДС и АА	0,962	1,000	0,990	—
M.L. Faron и соавт. [18]	Программное обеспечение на основе искусственного интеллекта WASPLab® (Cоран, Италия)	0,998 (0,997–0,999)	0,720 (0,487–0,840)	—	—
M. Yang и соавт. [19]	Модель свёрточной нейронной сети	0,987–0,988	0,981–0,994	—	0,985–0,990
S. Zurac и соавт. [20]	Модель глубокой нейронной сети для обнаружения микобактерий	0,957	1,000	0,977	0,983
Z. Wang и соавт. [21]	Модель свёрточной нейронной сети	0,914	0,913	0,975	0,893
A. Lev-Sagie и соавт. [22]	Модель свёрточной нейронной сети (<i>Candida albicans</i>)	0,841 (95% ДИ 0,737–0,909)	0,659 (95% ДИ 0,571–0,736)	—	—
	Модель свёрточной нейронной сети (бактериальные вагиниты)	0,909 (95% ДИ 0,764–0,969)	0,994 (95% ДИ 0,969–0,999)	—	—
D. Avcı и соавт. [24]	Модель свёрточной нейронной сети (AlexNet)	—	—	—	0,986
	Модель свёрточной нейронной сети (VGG-19)	—	—	—	0,964
	Модель свёрточной нейронной сети (VGG-16)	—	—	—	0,962
Y. Liang и соавт. [27]	Модель «случайного леса»	0,990	0,990	0,990	0,990
	Методы PBRTQC (Single MA–MovSD), лимфоциты, тромбоциты, объём гемоглобина	0,071–0,388	0,960–1,000	0,529–0,674	—
	Методы PBRTQC (Δ MA–MovSD), лимфоциты, тромбоциты, объём гемоглобина	0,067–0,975	0,961–1,000	0,527–0,968	—
R. Zhou и соавт. [28]	Модель слияния (модель «случайного леса»+модель опорных векторов+модель глубокой нейронной сети)	0,933–0,934	0,998–0,999	—	0,927–0,930
	Методы PBRTQC	0,666–0,976	0,538–0,792	—	0,712–0,964
H. Wang и соавт. [29]	Система автоверификации на основе искусственного интеллекта	0,999	0,981	0,998	—
	Модель экстремального градиентного бустинга (1 этап)	0,925	0,837	0,972	—
	Модель «случайного леса» (1 этап)	0,873	0,952	0,957	—
	Модель наивного байесовского классификатора (1 этап)	0,944	0,713	0,955	—
	Модель k-ближайших соседей (1 этап)	0,962	0,512	0,922	—
	Модель экстремального градиентного бустинга (2 этап)	0,924–0,990	0,876–0,903	0,9780–1,000	—
	Модель «случайного леса» (2 этап)	0,921–0,990	0,791–0,884	0,942–0,958	—
	Модель наивного байесовского классификатора (2 этап)	0,689–0,737	0,931–0,943	0,932–0,943	—
	Модель k-ближайших соседей (2 этап)	0,829–0,860	0,873–0,893	0,878–0,910	—

Примечание. AUC (Area Under Curve) — площадь под кривой; JBN-S — модель, разработанная на основе данных лабораторных отчётов, полученных с помощью химического анализатора Siemens® (Siemens Healthineers, Германия); Medlon-R — модель, разработанная на основе данных лабораторных отчётов, полученных с помощью химического анализатора Roche® (Roche Diagnostics, Германия); ДИ — доверительный интервал; США — Соединённые Штаты Америки; МДС — миелодиспластический синдром; ОМЛ — острый миелоидный лейкоз; АА — апластическая анемия; PBRTQC (Probability-Based Real-Time Quality Control) — метод контроля качества в реальном времени; Single MA (Moving Average) — одиночное скользящее среднее; MovSD (Moving Standard Deviation) — скользящее стандартное отклонение; Δ MA (Delta Moving Average) — разность скользящего среднего.

Приложение 6. Сравнительный анализ диагностической эффективности искусственного интеллекта и медицинских специалистов

Supplement 6. Comparative analysis of the diagnostic efficiency of artificial intelligence and medical workers

Авторы	Субъекты диагностики	Чувствительность	Специфичность	Точность	Число медицинских работников и их опыт	Другие критерии эффективности
C.J. Farrell [7]	Модель нейронной сети	—	—	0,921	—	—
	Модель простого дерева решений	—	—	0,865	—	—
	Медицинский работник	—	—	0,778	лаборанты — 28; научные сотрудники — 15; старшие научные сотрудники — 5; другие специалисты — 2. Медицинский стаж — медиана 4,5 года валидации биохимических результатов (интервал 0–40 лет)	—
S. Kurstjens и соавт. [12]	Модель JBN-S	0,930	0,920	—	—	время на пациента — <1 с
	Медицинский работник	0,830	0,920	—	врачи клинической лабораторной диагностики — 6	время на пациента — 19 с
	Медицинский работник+Модель JBN-S	0,870	0,900	—	—	время на пациента — 13 с
	Модель Medlon-R	0,980	0,920	—	—	время на пациента — <1 с
	Медицинский работник	0,880	0,910	—	врачи клинической лабораторной диагностики — 6	время на пациента — 20 с
	Медицинский работник+Модель Medlon-R	0,840	0,930	—	—	время на пациента — 16 с
D. Lea и соавт. [8]	—	—	—	—	—	• согласованность идентификации и стадирования: внутрикласовый коэффициент корреляции — 0,960 (95% ДИ 0,940–0,980); • κ — 0,860 (95% ДИ 0,810–0,910)
T.T. Van и соавт. [17]	Программное обеспечение с модулем хромогенного обнаружения PhenoMATRIX® (Coran, США)	0,906	0,940	—	—	—
	Медицинский работник	0,875	0,977	—	лаборанты, обученные считать хромогенные среды и посевы на наличие гемолитических колоний	—
S. Yoon и соавт. [11]	—	—	—	—	врачи-гематологи — 2 (предварительная классификация); врач-гематолог — 1 (конечная реклассификация)	• усреднённый коэффициент корреляции между ручным и автоматическим анализом — 0,360 (предварительная классификация) и 0,900 (конечной реклассификации); • высокие коэффициенты корреляции между ручным и автоматическим анализом для патологических образцов при подсчёте нейтрофилов ($r=0,880$) и лимфоцитов ($r=0,840$); • средний коэффициент корреляции для моноцитов ($r=0,520$) и эозинофилов ($r=0,650$);

EDN: BXDWFO

Авторы	Субъекты диагностики	Чувствительность	Специфичность	Точность	Число медицинских работников и их опыт	Другие критерии эффективности
						низкий коэффициент корреляции для базофилов ($p=0,340$)
M.L. Faron и соавт. [18]	—	—	—	—	—	- истинно положительные образцы — 5678; - ложноотрицательные образцы — 9; - истинно отрицательные образцы — 5598; - ложноположительные образцы — 2180
S. Zurac и соавт. [20]	Модель глубокой нейронной сети для обнаружения микобактерий	0,957	1,000	0,983	—	—
	Медицинский работник	0,391–0,957	0,756–0,946	0,830	врачи-патологоанатомы с различным медицинским стажем — 8 лаборанты	—
B.A. Mathison и соавт. [25]	—	—	—	—	—	- положительное согласие — 0,989 (95% ДИ 0,938–1,000); - отрицательное согласие — 0,981 (95% ДИ 0,934–0,998)
Z. Wang и соавт. [21]	Модель свёрточной нейронной сети	0,914	0,913	0,893	—	—
	Медицинский работник	0,943	0,731	0,837	лаборанты — 3; врачи-акушеры-гинекологи — 2	—
A. Lev-Sagie и соавт. [22]	Модель свёрточной нейронной сети (<i>Candida albicans</i>)	0,841	0,659	—	—	—
	Модель свёрточной нейронной сети (бактериальные вагиниты)	0,909	0,994	—	—	—
	Медицинский работник	0,868	0,975	—	врач-терапевт — 1	—

Примечание. JBN-S — модель, разработанная на основе данных лабораторных отчётов, полученных с помощью химического анализатора Siemens® (Siemens Healthineers, Германия); Medlon-R — модель, разработанная на основе данных лабораторных отчётов, полученных с помощью химического анализатора Roche® (Roche Diagnostics, Германия); ДИ — доверительный интервал; США — Соединённые Штаты Америки.

Appendix 7. Assessment of the quality of research methodology using the modified QUADAS-CAD questionnaire

Домены	C.J. Farrell [7]	S. Kurstjens и соавт. [12]	R.J. Burton и соавт. [23]	Е.А. Елагина и соавт. [15]	Н. Ayyildiz и соавт. [16]	D. Lea и соавт. [8]	M.B. Walla се и соавт. [26]	M.E. Lemieu х и соавт. [9]	H. Kim и соавт. [14]	T.T. Van и соавт. [17]	S. Yoon и соавт. [11]	M. Wang и соавт. [13]
Отбор Пациентов (D1)	Были ли данные (обучающие и тестовые наборы) сбалансированы по тяжести (включая отсутствие) целевой патологии?											
	не применимо	нет	неясно	нет	нет	нет	неясно	нет	да	нет	да	нет
	Были ли данные (обучающие и тестовые наборы) сбалансированы с точки зрения демографических факторов?											
	нет	да	нет	нет	нет	нет	нет	да	нет	нет	нет	нет
	Избежало ли исследование неуместных исключений?											
	неясно	да	да	да	неясно	нет	да	да	да	да	да	да
Индексный тест (D2)	Если использовалась нейронная сеть, были ли наборы данных обучения и тестирования не пересекающимися?											
	да	да	да	да	неясно	неясно	неясно	да	неясно	неясно	неясно	да
	Если использовалась нейронная сеть, был ли рационализирован размер каждого набора данных?											
	да	да	да	да	неясно	нет	нет	нет	неясно	неясно	неясно	да
	Если использовался порог патологии, был ли он установлен заранее?											
	не применимо	да	неясно	неясно	неясно	нет	неясно	нет	да	неясно	не применимо	неясно
Референсный стандарт (D3)	Если использовался порог принятия решения (для искусственного интеллекта), был ли он задан заранее?											
	неясно	да	неясно	неясно	неясно	нет	неясно	нет	да	да	да	да
	Может ли референсный стандарт правильно классифицировать целевое состояние?											
	да	да	да	да	да	да	неясно	да	да	да	да	да
Получение результатов (D4)	Были ли результаты референсных стандартов подготовлены или проверены с необходимым уровнем экспертизы?											
	да	да	неясно	неясно	неясно	неясно	неясно	неясно	неясно	да	да	да
	Была ли прозрачность в том, как были получены результаты?											
	да	да	да	да	да	да	нет	да	да	да	да	да
	Все ли данные пациентов имели один и тот же референсный стандарт?											
	да	да	да	да	да	да	да	да	да	да	да	да
	К. Kimura и соавт. [10]											
	M.L. Faron и соавт. [18]											
Отбор Пациентов (D1)	M. Yang и соавт. [19]											
	S. Zurac и соавт. [20]											
	B.A. Mathison и соавт. [25]											
	Z. Wang и соавт. [21]											
	A. Lev-Sagie и соавт. [22]											
Индексный тест (D2)	D. Avci и соавт. [24]											
	Y. Liang и соавт. [27]											
	R. Zhou и соавт. [28]											
	H. Wang и соавт. [29]											
	Были ли данные (обучающие и тестовые наборы) сбалансированы по тяжести (включая отсутствие) целевой патологии?											
	нет	да	нет	нет	нет	нет	нет	нет	нет	нет	нет	нет
	Были ли данные (обучающие и тестовые наборы) сбалансированы с точки зрения демографических факторов?											
	нет	нет	нет	нет	нет	нет	нет	нет	нет	нет	нет	нет
	Избежало ли исследование неуместных исключений?											
	неясно	неясно	неясно	неясно	да	неясно	да	неясно	да	неясно	неясно	неясно
	Если использовалась нейронная сеть, были ли наборы данных обучения и тестирования не пересекающимися?											
	неясно	неясно	да	да	да	да	неясно	да	да	да	да	да
	Если использовалась нейронная сеть, был ли рационализирован размер каждого набора данных?											
	неясно	неясно	да	да	да	да	да	да	да	да	да	да
	Если использовался порог патологии, был ли он установлен заранее?											
	да	не применимо	не применимо	неясно	не применимо	да	неясно	да	не применимо	не применимо	не применимо	не применимо
	Если использовался порог принятия решения (для ИИ), был ли он задан заранее?											

EDN: BXDWFO

	да	да	да	да	да	нет	да	да	нет	да	да
Референсный стандарт (D3)	Может ли референсный стандарт правильно классифицировать целевое состояние?										
	да	да	да	да	да	да	да	да	не применимо	не применимо	не применимо
	Были ли результаты референсных стандартов подготовлены или проверены с необходимым уровнем экспертизы?										
	да	неясно	да	да	да	да	да	да	неясно	х	х
Получение результатов (D4)	Была ли прозрачность в том, как были получены результаты?										
	да	да	да	да	да	да	да	да	да	да	да
	Все ли данные пациентов имели один и тот же референсный стандарт?										
	да	нет	да	да	да	да	да	да	да	да	да

Примечание. QUADAS-CAD (Quality Assessment of Diagnostic Accuracy Studies Computer-Aided Detection) — специализированный модифицированный опросник для оценки риска систематических ошибок и применимости исследований в области технологий искусственного интеллекта.

Рис. 1. Оценки риска систематической ошибки
Assessment of Diagnostic Accuracy S
для оценки риска систематических о

Рис. 1. Оценки риска систематической ошибки
Assessment of Diagnostic Accuracy S
для оценки риска систематических о