

<https://doi.org/10.29001/2073-8552-2026-41-3-214-224>  
УДК 616.5-006.81-073.97:004.89

# Пилотная проспективная оценка согласованности решений Melanoscope AI с экспертной дерматоскопической оценкой

Козачок Е.С.<sup>1</sup>, Карась С.И.<sup>2</sup>, Серегин С.С.<sup>3</sup>

<sup>1</sup> ФГБУН «Институт системного программирования имени В.П. Иванникова» Российской академии наук (ИСП РАН), 109004, Российская Федерация, Москва, ул. Александра Солженицына, 25

<sup>2</sup> Научно-исследовательский институт кардиологии, Томский национальный исследовательский медицинский центр Российской академии наук (НИИ кардиологии Томского НИМЦ), 634012, Российская Федерация, Томск, ул. Киевская, 111а

<sup>3</sup> БУЗ Орловской области «Орловский онкологический диспансер», 302020, Российская Федерация, Орел, Ипподромный пер., 2

## Аннотация

**Введение.** Ранняя диагностика злокачественных новообразований (ЗНО) кожи критически важна для прогноза, однако дефицит онкодерматологов в российских регионах ограничивает охват скрининга. Системы поддержки принятия врачебных решений (СППВР) на основе мобильной дерматоскопии представляют перспективный инструмент, при этом доказательство диагностической эффективности и оснащение первичного звена остаются первичными условиями, а на этапе клинической интеграции критичными становятся интерпретируемость вывода и стандартизация маршрутизации.

**Цель исследования:** оценить согласованность заключений СППВР Melanoscope AI (далее – система) с независимой экспертной дерматоскопической оценкой и изменение заключений врача общей практики после предъявления результата системы в условиях скрининговых сессий.

**Материал и методы.** Проведено одноцентровое проспективное пилотное исследование в четырех скрининговых сессиях (июнь 2025 – апрель 2026 гг.). Материалом служили дерматоскопические снимки – по одному наиболее подозрительному образованию на каждого из 176 обследованных, отобранному по результатам визуального осмотра. Оценивали согласованность заключений системы (злокачественное/доброкачественное новообразование) с заключением эксперта, сформированным до ознакомления с выводом системы. Образования, расцененные экспертом как злокачественные, подтверждались гистологически; доброкачественные случаи морфологически не верифицировались (частичная дифференциальная верификация). Изменение согласованности заключений врача общей практики (ОП) с экспертным заключением по одним и тем же случаям до и после предъявления результата системы оценивали точным критерием Мак-Немара.

**Результаты.** Совпадение бинарной классификации системы с независимым экспертным заключением по 176 образованиям составило 88,6% (95% ДИ 83,0–92,9), специфичность – 88,3% (82,5–92,7), прогностичность положительного результата – 20,0% (6,8–40,7). Все пять гистологически подтвержденных ЗНО (3 меланомы и 2 базальноклеточные карциномы) отнесены системой к красной зоне маршрутизации; ложноотрицательные решения при примененном референсном стандарте выявлены быть не могли, поскольку доброкачественные случаи гистологически не верифицировались. За шестью пациентами с диспластическими невусами установлено динамическое наблюдение. Согласованность заключений врача ОП с экспертным заключением возросла с 71,0 (63,9–77,2) до 82,4% (76,1–87,3); прирост – 11,4 п. п. (6,7–16,3), точный критерий Мак-Немара  $p = 1,9 \cdot 10^{-6}$ .

**Заключение.** Все гистологически подтвержденные ЗНО были отнесены системой к красной зоне маршрутизации, однако чувствительность и безопасность исключения злокачественного процесса настоящим пилотным исследованием не устанавливаются вследствие малого числа событий и отсутствия морфологической верификации отрицательных случаев. Согласованность заключений врача ОП с экспертным заключением после предъявления решения системы возросла. Необходима многоцентровая проверка согласованности решений системы на значительном контингенте обследуемых в условиях скрининга популяции, а также централизованный динамический мониторинг случаев, отнесенных системой к желтой зоне маршрутизации.

## Ключевые слова:

дерматоскопия; новообразования кожи; меланома; ранняя диагностика рака; поддержка принятия врачебных решений; компьютерная диагностика; глубокое обучение; скрининг; сортировка пациентов.

## Финансирование:

исследование выполнено без финансовой поддержки грантов, общественных, некоммерческих, коммерческих организаций и структур.

Козачок Елена Сергеевна, e-mail: e.kozachok@ispras.ru.

© Козачок Е. С., Карась С. И., Серегин С. С., 2026

|                                      |                                                                                                                                                                                                                                                                                                                                                                      |
|--------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Ресурсное обеспечение:</b>        | для выполнения вычислительных задач использовались ресурсы облачной платформы ИСП РАН.                                                                                                                                                                                                                                                                               |
| <b>Соответствие принципам этики:</b> | исследование проведено в соответствии с Хельсинкской декларацией. Все участники подписали информированное согласие на обработку обезличенных клинических данных и дерматоскопических изображений. Исследование носило наблюдательный характер и не предполагало изменения стандартного маршрута оказания медицинской помощи.                                         |
| <b>Для цитирования:</b>              | Козачок Е.С., Карась С.И., Серегин С.С. Пилотная проспективная оценка согласованности решений Melanoscope AI с экспертной дерматоскопической оценкой. <i>Сибирский журнал клинической и экспериментальной медицины</i> . 2026;41(3):214–224. <a href="https://doi.org/10.29001/2073-8552-2026-41-3-214-224">https://doi.org/10.29001/2073-8552-2026-41-3-214-224</a> |

# Pilot prospective assessment of the concordance of Melanoscope AI decisions with expert dermoscopic evaluation

Kozachok E.S.<sup>1</sup>, Karas S.I.<sup>2</sup>, Seregin S.S.<sup>3</sup>

<sup>1</sup> Ivannikov Institute for System Programming of the Russian Academy of Sciences (ISP RAS), 25, Alexander Solzhenitsyn str., Moscow, 109004, Russian Federation

<sup>2</sup> Cardiology Research Institute, Tomsk National Research Medical Center, Russian Academy of Sciences (Cardiology Research Institute, Tomsk NRMС), 111a, Kievskaya str., Tomsk, 634012, Russian Federation

<sup>3</sup> Orel Regional Oncology Dispensary, 2, Ippodromny lane, Orel, 302020, Russian Federation

## Abstract

**Introduction.** Early diagnosis of malignant skin neoplasms (MSN) is critical for prognosis, but a shortage of oncodermatologists in Russian regions limits screening coverage. Clinical decision support systems (CDSS) based on mobile dermoscopy represent a promising tool; proof of diagnostic effectiveness and equipping of primary care remain primary preconditions, while at the stage of clinical integration the interpretability of system output and standardization of patient routing become critical.

**Aim:** To assess the concordance between the conclusions of the Melanoscope AI clinical decision support system (hereafter, the system) and independent expert dermoscopic evaluation, and the change in the conclusions of a general practitioner after exposure to the system output in a screening setting.

**Material and Methods.** A single-center prospective pilot study was performed across four screening sessions (June 2025–April 2026). The material comprised dermoscopic images – one per examinee, of the lesion judged most suspicious on visual examination – from 176 participants. Concordance was assessed between the system's conclusions (malignant/benign) and the expert conclusion formed before exposure to the system's output. Lesions assessed by the expert as malignant were confirmed histologically; benign cases were not verified morphologically (partial differential verification). The change in concordance between a general practitioner's (GP) conclusions and the expert conclusion on the same cases, before and after exposure to the system's output, was assessed using the exact McNemar test.

**Results.** Agreement between the binary classification of the system and the independent expert conclusion across 176 lesions was 88.6% (95% CI 83.0–92.9), specificity 88.3% (82.5–92.7), positive predictive value 20.0% (6.8–40.7). All five histologically confirmed malignant neoplasms (3 melanomas and 2 basal cell carcinomas) were assigned by the system to the red routing zone; false-negative decisions could not be identified under the reference standard applied, since benign cases were not verified histologically. Six patients with dysplastic nevi were assigned to dynamic follow-up. Concordance between the GP's conclusions and the expert conclusion increased from 71.0% (63.9–77.2) to 82.4% (76.1–87.3); the gain was 11.4 percentage points (6.7–16.3), exact McNemar test  $p = 1.9 \cdot 10^{-6}$ .

**Conclusion.** All histologically confirmed malignant neoplasms were assigned by the system to the red routing zone; however, sensitivity and the safety of ruling out malignancy are not established by this pilot study, owing to the small number of events and the absence of morphological verification of negative cases. Concordance between the GP's conclusions and the expert conclusion increased after exposure to the system's output. Multicenter verification of the system's decision concordance is needed on a larger cohort of examinees under population screening conditions.

|                  |                                                                                                                                                                           |
|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Keywords:</b> | dermoscopy; skin neoplasms; melanoma; early detection of cancer; decision support systems, clinical; diagnosis, computer-assisted; deep learning; mass screening; triage. |
|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                           |                                                                                                                                                                                                                                                                                                                                                                     |
|-------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Funding:</b>                           | the study was conducted without financial support from grants, public, non-profit, or commercial organizations and entities.                                                                                                                                                                                                                                        |
| <b>Resources:</b>                         | computational tasks were performed using the cloud platform resources of ISP RAS.                                                                                                                                                                                                                                                                                   |
| <b>Compliance with ethical standards:</b> | the study was conducted in accordance with the Declaration of Helsinki. All participants signed informed consent for the processing of anonymized clinical data and dermoscopic images. The study was observational in design and did not involve any change to the standard care pathway.                                                                          |
| <b>For citation:</b>                      | Kozachok E.S., Karas S.I., Seregin S.S. Pilot prospective assessment of the concordance of Melanoscope AI decisions with expert dermoscopic evaluation. <i>Siberian Journal of Clinical and Experimental Medicine</i> . 2026;41(3):214–224. <a href="https://doi.org/10.29001/2073-8552-2026-41-3-214-224">https://doi.org/10.29001/2073-8552-2026-41-3-214-224</a> |

## Введение

Меланома составляет менее 5% всех злокачественных новообразований (ЗНО) кожи, но дает около 80% смертельных исходов в этой группе. Пятилетняя выживаемость при выявлении в I стадии достигает 98%, в IV стадии – около 23%<sup>1</sup> [1]. В Российской Федерации меланома кожи остается одной из значимых причин онкологической летальности при сохраняющейся доле случаев, выявляемых на поздних стадиях [2, 3]. Ранняя визуальная диагностика обеспечивает критический выигрыш во времени до начала лечения, однако точность распознавания меланомы существенно зависит от квалификации специалиста и метода осмотра. По данным систематического обзора и метаанализа 100 исследований, чувствительность и специфичность диагностики новообразований кожи врачами первичного звена при использовании дерматоскопии составляют 81,4 и 80,1% соответственно, тогда как у опытных дерматологов применение дерматоскопии повышает шансы правильного распознавания меланомы в 5,7 раза по сравнению с клиническим осмотром [4]. Дерматоскопия повышает чувствительность выявления меланомы по сравнению с осмотром невооруженным глазом, однако данные о ее применении именно в первичном звене остаются ограниченными [5]. В условиях российской первичной медицинской сети, где плотность онкодерматологов в регионах существенно ниже, чем в крупных городах, задача инструментальной поддержки диагностического решения врача общей практики (ОП) приобретает прямое прикладное значение. Действующие клинические рекомендации включают врачей ОП в число специалистов, осуществляющих раннее выявление новообразований кожи [6].

Системы поддержки принятия врачебных решений (СППВР) на основе глубокого обучения способны достигать диагностической точности уровня онкодерматолога [7–9]. Коммерческие решения – SkinVision, Google Dermatology Assist, Botkin.AI, ПроРодинки – прошли апробацию в различных клинических условиях [10, 11]. Вместе с тем в открытой литературе фиксируется три устойчивых ограничения, препятствующих внедрению таких систем в российскую клиническую практику.

Первое ограничение – непрозрачность автоматического решения. Большинство коммерческих СППВР возвращают врачу только вероятностное предсказание или категориальный диагноз без визуализации признаков, повлиявших на классификацию. Это блокирует эксперт-

ный контроль и формирует у врача недоверие к системе как к инструменту клинического выбора [12]. В контексте медицинской системы искусственного интеллекта интерпретируемость рассматривается не как опциональная функция, а как обязательное условие безопасного клинического применения [13]. Анализ причин ошибочной классификации изображений опухолей кожи программами на основе искусственного интеллекта подтверждает практическую значимость интерпретируемого вывода [14].

Второе ограничение – отсутствие стандартизированной маршрутизации. Результат автоматической классификации обычно представлен либо непрерывной вероятностью, либо категориальным классом, но не определяет конкретного клинического действия в зависимости от уровня риска. Без формализованной схемы маршрутизации СППВР не может быть встроена в рабочий процесс учреждения с ограниченными ресурсами, поскольку алгоритм действий персонала остается на усмотрение каждого специалиста и теряет воспроизводимость.

Третье ограничение – отсутствие независимых апробаций на российской популяции. Преобладающая часть валидационных публикаций по ИИ-СППВР в дерматоскопии выполнена на западных популяциях с фототипами кожи II–III по Фицпатрику [15]. Результаты апробаций на I–IV фототипах, характерных для населения России, в открытой литературе практически не представлены [16].

Цель работы: оценить согласованность заключений СППВР Melanoscope AI с независимой экспертной дерматоскопической оценкой и изменение заключений врача ОП после предъявления результата системы в условиях скрининговых сессий.

## Материал и методы

### Система поддержки принятия врачебных решений Melanoscope AI

СППВР Melanoscope AI (далее – система) разработана с учетом требований ГОСТ Р 71671-2024 к системам поддержки принятия врачебных решений с применением искусственного интеллекта (свидетельство о государственной регистрации программы для ЭВМ № 2026664695, ИСП РАН). Мобильное приложение зарегистрировано как самостоятельная программа для ЭВМ (свидетельство № 2026665028, ИСП РАН). Государственная регистрация программы для ЭВМ обеспечивает правовую охрану про-

<sup>1</sup> SEER Cancer Stat Facts: Melanoma of the Skin. National Cancer Institute. Bethesda, MD. URL: <https://seer.cancer.gov/statfacts/html/melan.html> (29.05.2026).

граммного обеспечения и не эквивалентна регистрации медицинского изделия: регистрационное удостоверение медицинского изделия на момент проведения исследования отсутствовало, подготовка к регистрации ведется. В настоящем исследовании система применялась как вспомогательный инструмент, не заменяющий клиническое решение врача.

Система реализует каскадную двухэтапную классификацию дерматоскопических изображений и включает три функциональных компонента: мобильное приложение получения снимков оптическим дерматоскопом, серверную подсистему с расчетом вероятности ЗНО, модуль визуализации (рис. 1). Первый компонент передает обезличенную запись на сервер, второй – рассчитывает вероятности ЗНО/нозологического класса; третий – визуализирует зоны изображения, наиболее повлиявшие на решение модели (Grad-CAM/attention, тепловая карта внимания).

Классификация дерматоскопического изображения организована в два этапа:

– Этап 1: бинарное разделение ЗНО и доброкачественных образований; выход – вероятность злокачественного процесса  $P(\text{ЗНО})$ .

– Этап 2: дифференциация нозологического класса при условии  $P(\text{ЗНО}) \geq 0,50$ : меланома (MEL), плоскоклеточный рак (SCC), базальноклеточная карцинома (BCC).

Сравнительный анализ архитектур обоих этапов, стратегия доменной адаптации и метрики обучения являются предметом отдельного исследования и в настоящей статье не приводятся.

В клинической апробации использовалась зафиксированная редакция модели бинарного этапа каскада (ConvNeXt-S). Веса модели зафиксированы до проведения первой сессии (07.06.2025 г.) и не изменялись на протяжении всего периода апробации (июнь 2025–апрель 2026 гг.); все четыре сессии выполнены на одной и той же версии. Изображения апробации и изображения тех же пациентов не использовались при обучении, настройке пороговых значений и калибровке модели. Пересечение апробационной выборки с наборами, использованными при обосновании пороговых значений, отсутствует.

Врач получает три компонента результата: вероятность злокачественного процесса  $P(\text{ЗНО})$  на выходе первого этапа каскада; уточняющий нозологический класс на

выходе второго этапа (при достаточной величине вероятности отнесения к классу); тепловую карту внимания с подсветкой зон изображения, наиболее повлиявших на решение модели.

Следует отметить, что ранее опубликованные работы [17] описывают иной программный комплекс, правообладателем которого является ООО «Открытые инновации». Система Melanoscope AI, представляемая в настоящей статье, является самостоятельной новой разработкой ФГБУН «Институт системного программирования им. В.П. Иванникова» Российской академии наук (Москва). Настоящая работа использует редакцию системы с формализованным трехзонным алгоритмом маршрутизации и интерпретируемыми картами внимания.

## Организация клинической апробации

### Дизайн и участники исследования

Клиническая апробация проведена в рамках серии профилактических мероприятий «День меланомы» из четырех сессий на базе Beauty Clinic (г. Орел).

Дизайн исследования – проспективная одноцентровая пилотная серия клинических наблюдений с независимой параллельной экспертной оценкой. Расчет размера выборки не проводился; объем выборки определялся числом участников профилактических сессий. Подготовка рукописи выполнена в соответствии с рекомендациями STARD-AI для исследований диагностической точности с применением искусственного интеллекта [18]. Экспертное заключение формировалось до ознакомления врача с результатом автоматической классификации.

Врач ОП формировал заключение по каждому случаю дважды в рамках одной сессии: первично – самостоятельно, в рамках стандартного протокола приема, без доступа к решению системы; повторно – после завершения первичного осмотра всех пациентов сессии, при одновременном предъявлении вероятности  $P(\text{ЗНО})$ , нозологического класса второго этапа и тепловой карты внимания. Порядок предъявления случаев при повторном чтении совпадал с порядком первичного осмотра; доступа к собственным первичным заключениям врач ОП не имел. На обоих чтениях врач ОП был ослеплен к экспертному заключению.

Оценку снимков проводили: канд. мед. наук

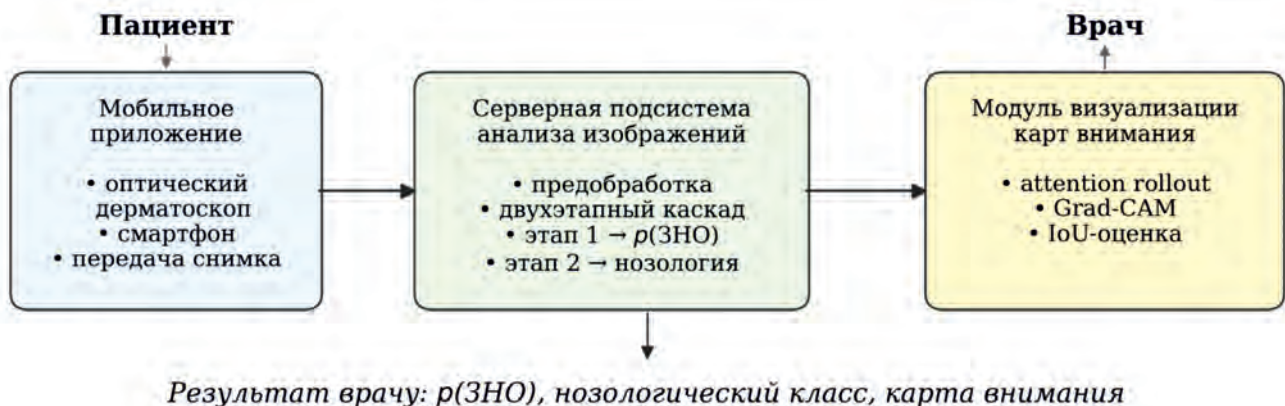


Рис. 1. Архитектура системы поддержки принятия врачебных решений Melanoscope AI: мобильное приложение получения снимков, серверная подсистема каскадного инференса и модуль визуализации карт внимания. Стрелки показывают потоки данных при стандартном цикле обследования

Fig. 1. Architecture of the Melanoscope AI CDSS: mobile image-acquisition application, server-side cascade inference subsystem and attention-map visualisation module. Arrows indicate data flow within a standard examination workflow

Серегин С.С. – главный онколог Орловской области, основной эксперт-референт; Козачок Е.С., врач-дерматолог – второй эксперт; врач общей практики Beauty Clinic.

Протокол участия пациента в апробации включал последовательность процедур:

- первичный визуальный осмотр кожного покрова с идентификацией объекта дерматоскопии;
- получение снимка объекта оптическим дерматоскопом, сопряженным со смартфоном, в соответствии с требованиями стандартной операционной процедуры (поляризованный режим, центрирование объекта, контроль качества изображения) [16];
- формирование независимого экспертного заключения без доступа к решению системы;
- классификация системой Melanoscope AI с фиксацией вероятности ЗНО, нозологического класса второго этапа и зоны карты внимания;
- при попадании в красную зону алгоритма маршрутизации – направление пациента на биопсию и гистологическое исследование с фиксацией результата в реестре направлений.

Применена дифференцированная верификация. Гистологическое исследование биопсийного материала выполнено в пяти случаях; направление на биопсию во всех случаях принималось экспертом по клиническим показаниям независимо от результата автоматической классификации, незавершенных биопсий и потерь на этапе верификации не было. Остальные образования, включая 20 образований красной зоны, не подтвержденных как злокачественные, и все 151 системно-отрицательное образование, классифицированы на основании экспертного клиничко-дерматоскопического заключения без морфологического подтверждения; дополнительного наблюдения системно-отрицательных случаев не проводилось. Референтом служило заключение основного эксперта; второй эксперт оценивал все 176 изображений независимо и был ослеплен как к заключению основного эксперта, так и к результату системы, расхождения разрешались в пользу заключения основного эксперта. Индивидуальные покейсовые записи второго эксперта не сохранялись, поэтому межэкспертное согласие количественно не оценивалось. Такой дизайн соответствует частичной дифференциальной верификации (partial verification bias), принятой в скрининговых исследованиях с ограниченными возможностями биопсии всех случаев, и является ограничением исследования: он может приводить к систематическому завышению специфичности, а ложноотрицательные решения системы при нем выявлены быть не могут.

Критерии включения обследуемых: возраст 18 лет и старше, наличие кожного новообразования, доступного для дерматоскопии, информированное согласие на участие в исследовании. Из обработки исключались изображения с  $image\_quality\_score \leq 2$  и пациенты, получавшие лечение оцениваемого новообразования менее чем за 3 мес. до сессии.

На четырех сессиях обследовано 176 человек; включены все обратившиеся, случаев исключения по критериям невключения не зарегистрировано. Отбраковки изображений по показателю  $image\_quality\_score \leq 2$  не потребовалось: все 176 снимков соответствовали требованиям стандартной операционной процедуры. У каждого пациента проводился визуальный осмотр кожного покрова с оценкой нескольких образований; для дерматоскопии и последующей автоматической классификации от-

биралось одно образование, расцененное как наиболее подозрительное по результатам визуального осмотра и с учетом жалоб пациента. Таким образом, анализируемая выборка представляет собой предварительно отобранные образования, а не сплошную выборку всех новообразований кожи обследованных.

В апробационную выборку включены данные 4 сессий скрининга – 176 первичных дерматоскопических снимков 176 пациентов. Распределение по сессиям приведено в таблице 1.

Таблица 1. Сессии клинической апробации

Table 1. Clinical validation sessions

| Сессия | Дата                          | Площадка               | Снимков |
|--------|-------------------------------|------------------------|---------|
| 1      | 07.06.2025                    | Beauty Clinic, г. Орел | 40      |
| 2      | 17.10.2025                    | Beauty Clinic, г. Орел | 43      |
| 3      | 26.12.2025                    | Beauty Clinic, г. Орел | 50      |
| 4      | 24.04.2026                    | Beauty Clinic, г. Орел | 43      |
| Итого  | июнь 2025–<br>апрель 2026 гг. | –                      | 176     |

Демографические характеристики выборки: возраст пациентов – от 19 до 84 лет, медиана – 42 года; 113 женщин (64,2%) и 63 мужчины (35,8%); преобладающие фототипы кожи по Фицпатрику – II (49,4%) и III (37,5%), что соответствует типичной популяции центральной части Российской Федерации.

### Этические аспекты исследования

Исследование проводилось в соответствии с Хельсинкской декларацией. Все пациенты подписали информированное согласие на участие в апробации, обработку обезличенных клинических данных и дерматоскопических изображений. Исследование носило характер клинической апробации (наблюдательная серия) и не предполагало изменения стандартного маршрута оказания медицинской помощи. Результаты системы использовались как вспомогательный инструмент и не заменяли клиническое решение врача. Направление на биопсию принималось врачом по клиническим показаниям. Протокол исследования и план статистического анализа предварительно не регистрировались.

### Статистический анализ

Долевые показатели приведены с 95% доверительными интервалами по методу Уилсона; для показателей, рассчитанных на малом числе событий, использован точный метод Клоппера – Пирсона. Согласованность заключений врача ОП с экспертным заключением без доступа и с доступом к результату системы сравнивалась на связанных выборках (одни и те же 176 образований) критерием Мак-Немара в точной биномиальной формулировке. Точный вариант выбран ввиду малого числа дискордантных пар и нулевого значения одной из них, при которых асимптотическое приближение и поправка Йейтса неприменимы. Отношение шансов дискордантных пар при отсутствии обратных переходов не определено, поэтому величина эффекта охарактеризована долей дискордантных пар в пользу решения с доступом к системе с точным биномиальным доверительным интервалом и разностью связанных долей с доверительным интервалом по методу Ньюкомба. Расчеты выполнены в Python 3.11 (statsmodels 0.14, scipy 1.14).

## Результаты клинической апробации

### Оценка согласованности решений

Совпадение автоматической классификации с независимым экспертным заключением составило 88,6% (156 совпадений из 176 снимков). Матрица ошибок для бинарной задачи «злокачественное/доброкачественное» приведена в таблице 2.

Таблица 2. Матрица ошибок автоматической классификации ( $n = 176$ )

Table 2. Confusion matrix of the automatic classification ( $n = 176$ )

| Решения            | Эксперт: злокач. | Эксперт: доброкач. | Итого |
|--------------------|------------------|--------------------|-------|
| Система: злокач.   | TP = 5           | FP = 20            | 25    |
| Система: доброкач. | FN = 0           | TN = 151           | 151   |
| Итого              | 5                | 171                | 176   |

Примечание: значения TP и FP определены относительно экспертного клиничко-дерматоскопического заключения, гистологическое подтверждение получено только для пяти образований, отнесенных экспертом к злокачественным. Значение FN отражает отсутствие выявленных ложноотрицательных решений при примененном референсном стандарте и не может интерпретироваться как доказательство их отсутствия.

Операционные характеристики решений системы при бинарной классификации новообразований относительно экспертного заключения: общая точность (доля правильных классификаций) – 88,6% (95% ДИ 83,0–92,9); специфичность – 88,3% (82,5–92,7); прогностичность положительного результата (PPV) – 20,0% (6,8–40,7). Все пять образований, расцененных экспертом как злокачественные и подтвержденных гистологически, отнесены системой к красной зоне (5/5; точный 95% ДИ 47,8–100,0). Чувствительность и прогностичность отрицательного результата относительно морфологического диагноза не рассчитывались: доброкачественные по обоим методам образования гистологически не верифицировались, поэтому ложноотрицательные решения системы выявлены быть не могли.

Низкое значение PPV (20,0%) закономерно для скрининговой популяции с распространенностью ЗНО 2,8% (5/176) среди отобранных образований. При низкой априорной вероятности система даже с высокой специфичностью 88% дает значительную долю ложноположительных результатов среди направленных в красную зону. Это обстоятельство не снижает диагностической ценности системы в скрининговом контексте, однако означает, что в среднем 4 из 5 пациентов красной зоны имеют доброкачественные образования. Данный факт должен быть разъяснен персоналу и пациентам, а также учтен при оценке нагрузки на биопсийные службы.

Анализ ложноположительных решений системы, т. е. случаев доброкачественных новообразований (FP = 20) показал наличие среди них пигментированных себорейных кератозов с псевдосетчатым паттерном ( $n = 9$ ), диспластических невусов с выраженной атипией пигментной сети ( $n = 6$ ), гемангиом с темной пигментной коркой ( $n = 5$ ). Нозологическая принадлежность этих 20 образований установлена по дерматоскопическому заключению эксперта без гистологического подтверждения. Все ложноположительные случаи попадают в красную зону маршрутизации, что обеспечивает их направление к специалисту ценой избыточных консультаций онкодерматолога. Ложноотрицательные решения при примененном референсном стандарте выявлены быть не могли,

поскольку доброкачественные по обоим методам образования гистологически не верифицировались.

Суммарная распространенность онкологических и предраковых новообразований кожи составила  $11/176 = 6,3\%$  от числа обследованных в клинической апробации. Все 3 случая меланомы попали в красную зону по результатам первого этапа каскада. Два случая базальноклеточной карциномы отнесены системой к красной зоне с уточнением класса на втором этапе каскада и подтверждены гистологически. Шесть случаев диспластических невусов с атипичными признаками дерматоскопической картины включены в программу динамического наблюдения с повторной дерматоскопией через 6 мес.: четыре выявлены в первой сессии (07.06.2025), два – во второй (17.10.2025). Повторный осмотр выполнен для всех шести случаев; признаков злокачественной трансформации не выявлено.

### Изменение согласованности заключений врача общей практики после предъявления результата системы

Доля совпадений заключения врача ОП с экспертным заключением без доступа к результатам работы системы составила 71,0% (125 из 176; 95% ДИ 63,9–77,2). После предъявления результата автоматической классификации и тепловой карты внимания доля совпадений возросла до 82,4% (145 из 176; 76,1–87,3), прирост – 11,4 п. п. (95% ДИ 6,7–16,3). Структура парных переходов приведена в таблице 3: заключение врача ОП изменилось в 20 случаях, все изменения произошли в сторону согласия с экспертным заключением, обратных переходов не зафиксировано (20 из 20; точный 95% ДИ доли переходов в пользу согласия 83,2–100,0; точный критерий Мак-Немара,  $p = 1,9 \cdot 10^{-6}$ ). Отношение шансов дискордантных пар при отсутствии обратных переходов не определено. Все изменения решений пришлось на желтую (12 из 30 случаев зоны) и красную (8 из 25) зоны маршрутизации; в зеленой зоне (121 случай) заключение врача ОП не изменилось ни разу.

Таблица 3. Парное сравнение заключений врача общей практики с экспертным заключением без доступа и с доступом к результату системы ( $n = 176$ )

Table 3. Paired comparison of the general practitioner conclusions with the expert conclusion without and with access to the system output ( $n = 176$ )

| Без доступа к системе | Совпало с экспертом | Не совпало | Всего |
|-----------------------|---------------------|------------|-------|
| Совпало с экспертом   | 125                 | 0          | 125   |
| Не совпало            | 20                  | 31         | 51    |
| Всего                 | 145                 | 31         | 176   |

### Алгоритм маршрутизации пациента

Результат первого этапа каскада – непрерывная вероятность злокачественного процесса Р(ЗНО) – не может непосредственно определять клиническое действие без формальной схемы интерпретации. Общепринятой клинической практикой является разбивка непрерывной шкалы риска на дискретные зоны, каждой из которых соответствует стандартный алгоритм действий медицинского персонала [19].

Пороговые значения обоснованы раздельно и приняты в настоящем исследовании в готовом виде [20]. Значение  $p = 0,15$  соответствует оптимуму индекса Юдена

бинарного этапа отбора: на внутренней отложенной выборке оптимум достигается при  $p = 0,17$ , на клиническом наборе Сеченовского университета – при  $p = 0,15$ , на наборе Melanoscope – при  $p = 0,18$ , то есть положение оптимума устойчиво при переносе между наборами, несмотря на снижение ROC-AUC с 0,965 до 0,858 (рис. 2). Значение  $p=0,50$  соответствует нейтральной рабочей точке правила максимума апостериорной вероятности и принято как порог запуска нозологической дифференциации второго этапа, где решение влечет направление на биопсию. Результаты настоящей апробации согласуются с принятыми порогами: в зеленой зоне ( $p < 0,15$ ) ложно-отрицательных случаев не зафиксировано ни в одной из четырех сессий, а в красной зоне ( $p \geq 0,50$ ) выявлены все подтвержденные ЗНО (5/5). При примененном референсном стандарте эти наблюдения не могут рассматриваться как независимая валидация пороговых значений.

Алгоритм маршрутизации реализует трехзонную схему (рис. 3):

- зеленая зона ( $p < 0,15$ ) – фиксация в медицинской карте, информирование пациента, повторный осмотр через 6–12 мес.;
- желтая зона ( $0,15 \leq p < 0,50$ ) – направление к онкодерматологу, повторная дерматоскопия, при повторном попадании в желтую или красную зону – биопсия;
- красная зона ( $p \geq 0,50$ ) – безотлагательное направление к онкологу или онкодерматологу, внеочередной прием, биопсия или эксцизия.

При апробации системы 176 пациентов распределились по зонам риска следующим образом (рис. 4): зеленая зона – 121 пациент (68,8%), желтая – 30 пациентов (17,0%), красная – 25 пациентов (14,2%). Все три подтвержденных меланомы и две базальноклеточные карциномы были отнесены к красной зоне уже на первом этапе каскада (бинарная оценка Р(ЗНО); второй этап лишь уточняет нозологический класс и на факт попадания в крас-

ную зону не влияет). Из желтой зоны 6 пациентов с диспластическими невусами направлены на динамическое наблюдение; ни один не подтвержден как злокачественный по результатам последующего осмотра. Применение алгоритма привело к направлению 25 пациентов (14,2%) в красную зону с рекомендацией внеочередной консультации и биопсии. Злокачественный процесс подтвержден у пяти из них, то есть доля направлений, не подтвержденных морфологически, составила 80%. Еще 30 пациентов (17,0%) отнесены к желтой зоне с рекомендацией повторной дерматоскопии. Эти показатели определяют ресурсную нагрузку на биопсийную и консультативную службы при внедрении алгоритма. Предусмотренный алгоритмом повторный осмотр пациентов зеленой зоны через 6–12 мес. является элементом маршрутизации; в рамках настоящего исследования наблюдение системно-отрицательных случаев не проводилось.

## Обсуждение

Сравнение с существующими СППВР выполнено по критериям, релевантным для оценки их готовности к клиническому применению в условиях российской медицинской практики. Ключевым среди них является интерпретируемость как условие формирования обоснованного доверия врача к системе (табл. 4) [21].

Рассмотренные системы методологически неоднородны, что ограничивает их прямое сопоставление по единственному показателю точности и требует комплексной оценки. SkinVision и Google Derm. Assist реализованы как потребительские мобильные приложения скрининга, рассчитанные на самостоятельное обследование пациентом без интеграции в маршрут оказания медицинской помощи. Их валидационные выборки сформированы преимущественно из европейской и североамериканской популяции, что не гарантирует сопоставимой точности на фототипах I–IV, преобладающих у российских пациентов.

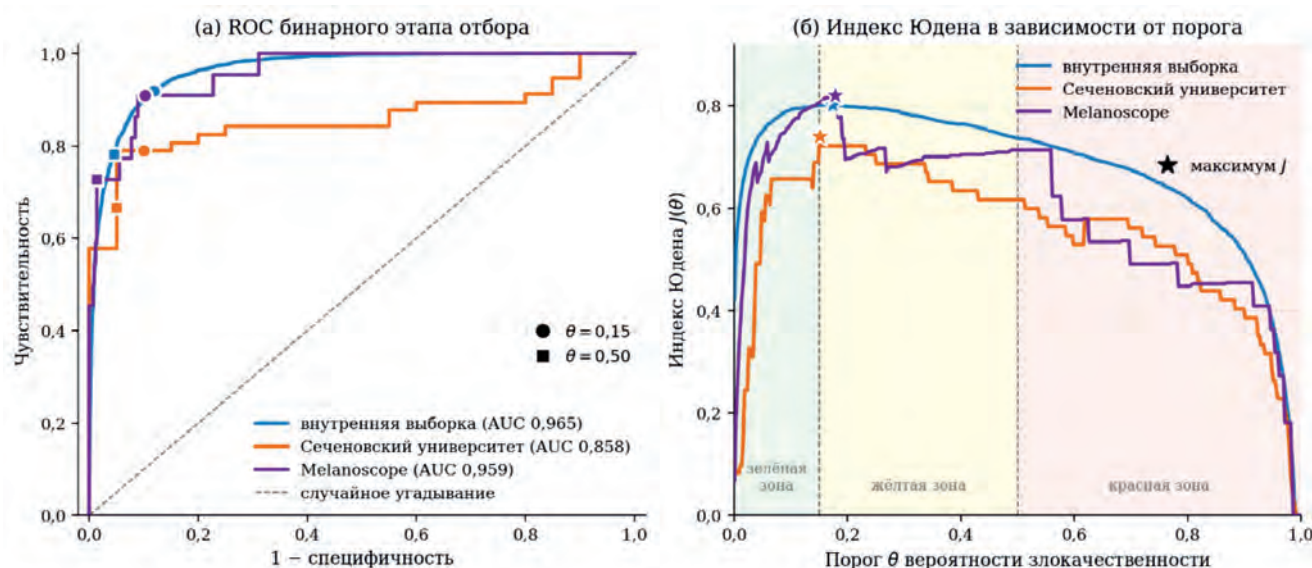


Рис. 2. Обоснование пороговых значений алгоритма маршрутизации. а – ROC-кривые бинарного этапа отбора каскада (ConvNeXt-S) на внутренней отложенной выборке и двух независимых клинических наборах; отмечены рабочие точки  $p = 0,15$  (круг) и  $p = 0,50$  (квадрат); б – индекс Юдена в зависимости от порога; максимумы (звезды) достигаются при  $p = 0,17$  (внутренняя выборка), 0,15 (Сеченовский университет) и 0,18 (Melanoscope), то есть на границе зеленой и желтой зон маршрутизации, устойчиво при переносе между наборами

Fig. 2. Justification of the routing algorithm thresholds. а – ROC curves of the cascade binary triage stage (ConvNeXt-S) on the internal held-out set and two independent clinical sets; operating points  $p = 0.15$  (circle) and  $p = 0.50$  (square) are marked. б – Youden index versus threshold; the maxima (stars) are attained at  $p = 0.17$ , 0.15 and 0.18 respectively, i.e. at the green/yellow routing zone boundary, consistently across sets

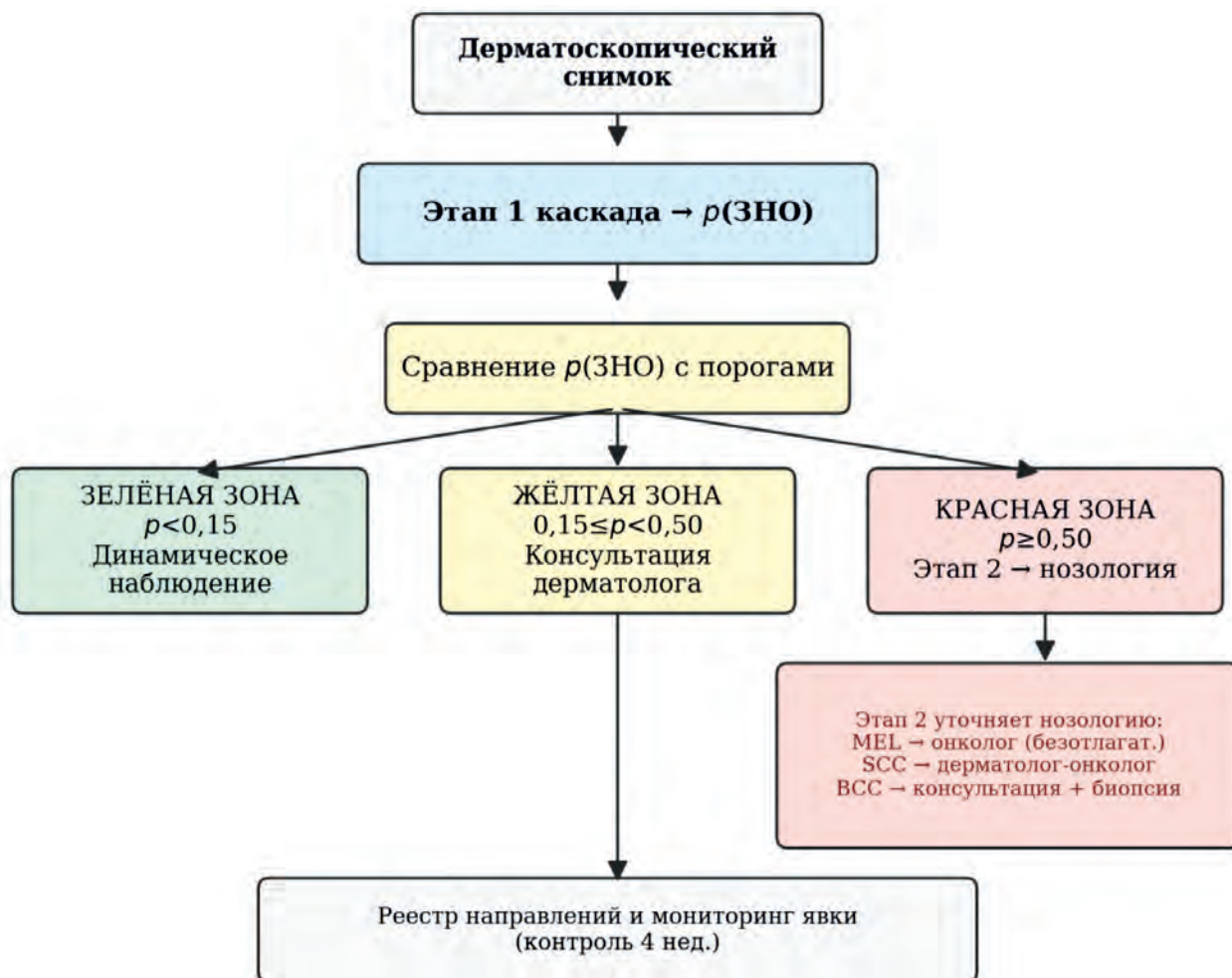


Рис. 3. Блок-схема трехзонного алгоритма маршрутизации пациента. Вычисляется  $P(\text{ЗНО})$ ; в красной зоне второй этап уточняет нозологический класс и определяет степень срочности направления

Fig. 3. Flowchart of the three-zone patient routing algorithm.  $p(\text{malignancy})$  is computed; in the red zone the second stage refines the nosological class and defines the urgency of referral

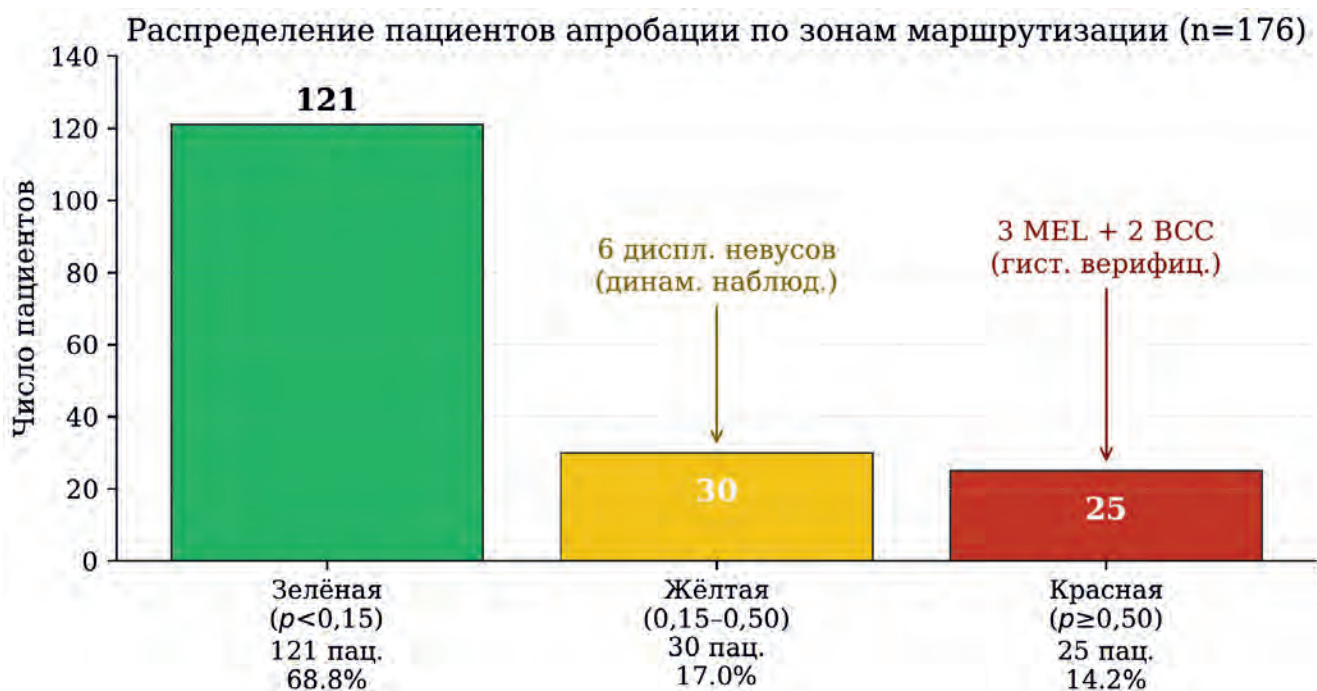


Рис. 4. Распределение 176 пациентов апробационной выборки по зонам алгоритма маршрутизации

Fig. 4. Distribution of 176 patients in validation sample across the routing zones

**Таблица 4.** Сравнение системы поддержки принятия врачебных решений по диагностике новообразований кожи**Table 4.** Comparison of CDSSs for skin-lesion diagnosis

| Критерий                                                | Melanoscope AI (наст. ред.) | SkinVision | Google Derm. Assist | Botkin.AI |
|---------------------------------------------------------|-----------------------------|------------|---------------------|-----------|
| Визуальное обоснование решения (тепловая карта)         | +                           | –          | –                   | –         |
| Стандарт. алгоритм маршрутизации                        | + (3 зоны)                  | –          | –                   | –         |
| Независимая апробация в РФ                              | + (4 сессии 2025 – 2026)    | –          | –                   | огранич.  |
| Учет фототипов I–IV РФ                                  | +                           | –          | –                   | –         |
| Мобильная дерматоскопия                                 | +                           | частично   | –                   | –         |
| Регистрационное удостоверение медицинского изделия в РФ | –                           | –          | –                   | +         |

Примечание: сведения о сравниваемых системах получены по публикациям [10, 11] и открытым материалам производителей, дата обращения к источникам – 08.08.2026. Сопоставление носит ориентировочный характер: системы различаются по целевому пользователю, контексту применения и составу функциональных модулей, а операционные характеристики получены на несопоставимых выборках и референсных стандартах, поэтому прямое сравнение показателей точности не приводится. Google Dermatology Assist на дату обращения не находится в разработке, в связи с чем не может рассматриваться как актуальный коммерческий аналог.

Botkin.AI – отечественная система поддержки принятия решений для врача, зарегистрированная как медицинское изделие, однако не содержащая алгоритма маршрутизации пациентов по степени риска и визуального обоснования решения, а также не имеющая в открытом доступе данных об операционных характеристиках по диагностике меланомы. Указанные различия в целевой аудитории, контексте применения и составе функциональных модулей систем определяют необходимость их сопоставления не по изолированной диагностической метрике, а по совокупности критериев, отражающих методологические основания и готовность к клиническому внедрению в условиях российского здравоохранения.

Для сравнения отобраны шесть критериев, отражающих не точечную диагностическую эффективность, а готовность СППВР к клиническому внедрению в первичном звене: наличие визуального обоснования решения (интерпретируемость), реализация алгоритма маршрутизации пациентов по степени риска, факт независимой апробации в РФ, учет фототипов кожи I–IV по Фитцпатрику, возможность мобильной дерматоскопии и наличие регистрационного удостоверения медицинского изделия в РФ.

Отличительной особенностью представленной редакции системы является интегрированная цепочка «каскадная классификация – карта внимания – алгоритм маршрутизации». По совокупности этих компонентов рассмотренные коммерческие аналоги, по доступным сведениям, сопоставимого решения не предоставляют. Число выявленных ЗНО в апробации мало (5 случаев, из них 3 меланомы), поэтому операционные характеристики системы по отдельным нозологиям настоящим исследованием не оцениваются; их получение требует расширения выборки в рамках многоцентрового исследования.

Метод количественной оценки согласованности карт внимания с экспертной аннотацией дерматоскопических структур по метрике IoU разработан авторами и применен к четырем архитектурам на отдельном размеченном наборе изображений. Для 176 клинических образований настоящего исследования экспертная аннотация структур не выполнялась, поэтому согласованность предъявлявшихся врачу ОП карт внимания количественно не оценивалась. Соответственно, изменение заключений врача ОП не может быть отнесено к вкладу визуального объяснения: вероятность, нозологический класс и тепловая карта предъявлялись одновременно, и раздельная оценка их влияния дизайном исследования не предусмотрена.

Данные внешней оценки бинарного этапа показывают

снижение ROC-AUC с 0,965 на внутренней отложенной выборке до 0,858 на независимом клиническом наборе при сохранении положения оптимального порога [20]. Это указывает на доменный сдвиг между наборами и подтверждает необходимость многоцентровой проверки до клинического внедрения.

Перспективы дальнейших исследований включают многоцентровое формирование выборки и проверку согласованности решений системы на значительном контингенте обследуемых в условиях скрининга популяции, значительный интервал времени между заключениями врача ОП, не имеющего и имеющего информацию о решении системы, оценка взаимосвязи решений врача ОП со стадией выявления ЗНО, а также централизованный динамический мониторинг случаев, отнесенных системой к желтой зоне маршрутизации.

### Ограничения исследования

Подробный разбор ограничений исследования необходим для повышения объективности оценки полученных результатов.

1. Выборка в 176 снимков достаточна для демонстрации работоспособности системы и первичной оценки прироста согласованности врача ОП с экспертным заключением. Для устойчивых оценок чувствительности и специфичности по редким классам ЗНО полученные значения можно использовать только как ориентировочные. Малая подвыборка злокачественных случаев (5 человек) обусловлена скрининговым дизайном исследования (добровольное участие, распространенность 2,8%), что ограничивает применение этих оценок на популяциях с более высокой априорной вероятностью, не позволяет оценить точность по разным фототипам, полу и возрасту.

2. Анализировались предварительно отобранные образования – по одному наиболее подозрительному на пациента, а не все новообразования кожи обследованных. Распространенность ЗНО 2,8% характеризует отобранные образования, а не обследованную популяцию; полученные значения специфичности и прогностической ценности не переносимы на тотальный осмотр кожного покрова.

3. Гистологическая верификация выполнена в пяти случаях; остальные образования классифицированы по экспертному клинко-дерматоскопическому заключению без морфологического подтверждения и без последующего наблюдения. При таком референсном стандарте ложноотрицательные решения системы не выявляемы, а специфичность может быть завышена; чувствительность и прогностичность отрицательного результата

относительно морфологического диагноза настоящим исследованием не устанавливаются. Индивидуальные покейсовые заключения второго эксперта не сохранялись, поэтому межэкспертное согласие количественно не оценивалось.

4. Ограничением также является одноцентровый дизайн исследования: для масштабирования результатов в иных условиях, на ином оборудовании и демографических группах требуется многоцентровая проверка.

5. Оценка прироста согласованности получена в дизайне с повторным чтением одних и тех же изображений одним специалистом в рамках одной сессии, при котором второй проход выполнялся с доступом к результату системы и в том же порядке предъявления. Такой дизайн не исключает эффекта согласия с подсказкой и эффекта запоминания изображений, поэтому наблюдаемый прирост следует трактовать как оценку сверху. Полное отсутствие обратных переходов (0 из 176 пар) согласуется как с гипотезой о пользе системы, так и с гипотезой о преимущественном согласии врача с ее выводом; настоящий дизайн не позволяет их разделить. При нуле наблюдаемых случаев расхождения верхняя граница 95% доверительного интервала их доли составляет 1,7%, что не исключает клинически значимого автоматизационного смещения в более крупной выборке. Подтверждение требует протокола с ослепленным повторным чтением, разнесением проходов во времени и участием нескольких врачей сходной квалификации.

6. Отсутствие независимого от разработчиков и экспертов внешнего аудита данных в данной работе будет устранено в многоцентровом исследовании.

7. Оценка точности нозологической дифференциации второго этапа каскада в задачи настоящего исследования не входила; матрица ошибок по нозологическим классам не приводилась. Приоритетом являлась оценка бинарного этапа, определяющего попадание образования в зоны маршрутизации. Протокол исследования и план статистического анализа предварительно не регистрировались.

## Заключение и выводы

В работе представлены результаты проспективного одноцентрового пилотного исследования согласованности решений СППВР Melanoscope AI с независимой экспертной оценкой и гистологической верификацией выявленных ЗНО в условиях скрининговой программы. Нами получены следующие основные результаты:

1. Общая точность автоматической классификации относительно экспертного заключения составила 88,6% (95% ДИ 83,0–92,9), специфичность – 88,3% (82,5–92,7). Все пять гистологически подтвержденных ЗНО (3 меланомы, 2 базальноклеточные карциномы) отнесены системой к красной зоне маршрутизации; вследствие малого числа событий и отсутствия морфологической верификации отрицательных случаев чувствительность и безопасность исключения злокачественного процесса настоящим исследованием не устанавливаются и требуют независимой многоцентровой проверки.

2. Согласованность решений врача ОП с экспертным заключением возросла с 71,0 (95% ДИ 63,9–77,2) до 82,4% (76,1–87,3) после предъявления результата системы; прирост – 11,4 п. п. (95% ДИ 6,7–16,3), точный критерий Мак-Немара  $p = 1,9 \cdot 10^{-6}$ . Причинный вклад отдельных компонентов вывода системы и воспроизводимость

эффекта требуют подтверждения в контролируемом многоэкспертном дизайне.

## Благодарности

Авторы выражают благодарность коллективу Beauty Clinic (г. Орел) за организационную поддержку проведения профилактических сессий «День меланомы». Работа выполнена в рамках плана научно-исследовательских работ ИСП РАН.

**Декларация ИИ.** В процессе подготовки рукописи авторы не использовали генеративные технологии искусственного интеллекта.

## Литература / References

- Sung H., Ferlay J., Siegel R.L., et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J. Clin.* 2021;71(3):209–249. DOI: [10.3322/caac.21660](https://doi.org/10.3322/caac.21660)
- Каприн А.Д., Старинский В.В., Шахзадова А.О. Злокачественные новообразования в России в 2023 году (заболеваемость и смертность). М.: МНИОИ им. П.А. Герцена – филиал ФГБУ «НМИЦ радиологии» Минздрава России, 2024:276. ISBN 978-5-85502-298-8.
- Kaprin A.D., Starinskiy V.V., Shakhzadova A.O. Malignant neoplasms in Russia in 2023 (morbidity and mortality). Moscow: P.A. Herzen MORI – branch of the National Medical Research Radiological Centre, 2024:276. (In Russ.) ISBN 978-5-85502-298-8.
- Шахзадова А.О., Старинский В.В., Лисичникова И.В. Состояние онкологической помощи населению России в 2022 году. *Сибирский онкологический журнал*. 2023;22(5):5–13. DOI: [10.21294/1814-4861-2023-22-5-5-13](https://doi.org/10.21294/1814-4861-2023-22-5-5-13) EDN: [PESHHL](https://www.edn.ru/PESHHL)
- Shakhzadova A.O., Starinskiy V.V., Lisichnikova I.V. State of cancer care for the population of Russia in 2022. *Siberian Journal of Oncology*. 2023;22(5):5–13. (In Russ.) DOI: [10.21294/1814-4861-2023-22-5-5-13](https://doi.org/10.21294/1814-4861-2023-22-5-5-13) EDN: [PESHHL](https://www.edn.ru/PESHHL)
- Chen J.Y., Fernandez K., Fadadu R.P., et al. Skin cancer diagnosis by lesion, physician, and examination type: a systematic review and meta-analysis. *JAMA Dermatology*. 2025;161(2):135–146. DOI: [10.1001/jamadermatol.2024.4382](https://doi.org/10.1001/jamadermatol.2024.4382)
- Dinnes J., Deeks J.J., Chuchu N., et al. Dermoscopy, with and without visual inspection, for diagnosing melanoma in adults. *Cochrane Database of Systematic Reviews*. 2018;12(12):CD011902. DOI: [10.1002/14651858.CD011902.pub2](https://doi.org/10.1002/14651858.CD011902.pub2)
- Esteva A., Kuprel B., Novoa R.A., et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115–118. DOI: [10.1038/nature21056](https://doi.org/10.1038/nature21056)
- Brinker T.J., Hekler A., Enk A.H., et al. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. *Eur. J. Cancer*. 2019;113:47–54. DOI: [10.1016/j.ejca.2019.04.001](https://doi.org/10.1016/j.ejca.2019.04.001)
- Tschandl P., Rinner C., Apalla Z., et al. Human-computer collaboration for skin cancer recognition. *Nature Medicine*. 2020;26(8):1229–1234. DOI: [10.1038/s41591-020-0942-0](https://doi.org/10.1038/s41591-020-0942-0)
- Udrea A., Mitra G.D., Costea D., et al. Accuracy of a smartphone application for triage of skin lesions based on machine learning algorithms. *J. Eur. Acad. Dermatol. Venereol.* 2020;34(3):648–655. DOI: [10.1111/jdv.15935](https://doi.org/10.1111/jdv.15935)
- Liu Y., Jain A., Eng C., et al. A deep learning system for differential diagnosis of skin diseases. *Nature Medicine*. 2020;26(6):900–908. DOI: [10.1038/s41591-020-0842-7](https://doi.org/10.1038/s41591-020-0842-7)
- Wysocki O., Davies J.K., Vigo M., et al. Assessing the communication gap between AI models and healthcare professionals: explainability, utility and trust in AI-driven clinical decision-making. *Artificial Intelligence*. 2023;319:103839. DOI: [10.1016/j.artint.2022.103839](https://doi.org/10.1016/j.artint.2022.103839)
- Reyes M., Meier R., Pereira S., et al. On the interpretability of artificial intelligence in radiology: challenges and opportunities. *Radiology: Artificial Intelligence*. 2020;2(3):e190043. DOI: [10.1148/ryai.2020190043](https://doi.org/10.1148/ryai.2020190043)
- Корабельников Д.И., Ламоткин А.И. Анализ причин ошибочной классификации изображений немеланоцитарных опухолей кожи программами на основе технологий искусственного интеллекта. *Сибирский журнал клинической и экспериментальной медицины*. 2026;41(1):221–231. DOI: [10.29001/2073-8552-2026-2881](https://doi.org/10.29001/2073-8552-2026-2881) EDN: [DLMLOC](https://www.edn.ru/DLMLOC)
- Korabelnikov D.I., Lamotkin A.I. Analysis of the causes of misclassification of non-melanocytic skin tumor images by artificial-intelligence-based

- software. *Siberian Journal of Clinical and Experimental Medicine*. 2026;41(1):221–231. (In Russ.) DOI: [10.29001/2073-8552-2026-2881](https://doi.org/10.29001/2073-8552-2026-2881) EDN: [DLMLOC](https://doi.org/10.29001/2073-8552-2026-2881)
14. Combalia M., Codella N., Rotemberg V., et al. Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: the 2019 International Skin Imaging Collaboration grand challenge. *The Lancet Digital Health*. 2022;4(5):e330–e339. DOI: [10.1016/S2589-7500\(22\)00021-8](https://doi.org/10.1016/S2589-7500(22)00021-8)
  15. Kozachok E.S. Methodology for Creating a Clinically Verified Dermoscopic Image Dataset [Preprint]. 2026. arXiv:2605.25168 [cs.CV]. DOI: [10.48550/arXiv.2605.25168](https://doi.org/10.48550/arXiv.2605.25168)
  16. Козачок Е.С., Серегин С.С., Козачок А.В. и др. Интеллектуальная система поддержки принятия решений врача в диагностике новообразований кожи на основе анализа дерматоскопических изображений. *Известия Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение*. 2025;15(3):50–65. DOI: [10.21869/2223-1536-2025-15-3-50-65](https://doi.org/10.21869/2223-1536-2025-15-3-50-65)
  17. Kozachok E.S., Seregin S.S., Kozachok A.V., et al. An intelligent clinical decision support system for diagnosis of skin lesions based on dermoscopic image analysis. *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta. Seriya: Upravlenie, vychislitel'naya tekhnika, informatika. Meditsinskoe priborostroenie*. 2025;15(3):50–65. (In Russ.) DOI: [10.21869/2223-1536-2025-15-3-50-65](https://doi.org/10.21869/2223-1536-2025-15-3-50-65)
  18. Sounderajah V., Guni A., Liu X., et al.; STARD-AI Steering Committee. The STARD-AI reporting guideline for diagnostic accuracy studies using artificial intelligence. *Nature Medicine*. 2025;31(10):3283–3289. DOI: [10.1038/s41591-025-03953-8](https://doi.org/10.1038/s41591-025-03953-8)
  19. Argenziano G., Soyer H.P., Chimenti S., et al. Dermoscopy of pigmented skin lesions: results of a consensus net meeting via the Internet. *J. Am. Acad. Dermatol.* 2003;48(5):679–693. DOI: [10.1067/mjd.2003.281](https://doi.org/10.1067/mjd.2003.281)
  20. Козачок Е.С., Серегин С.С., Козачок А.В. и др. Каскадная классификация дерматоскопических изображений новообразований кожи с управляемой чувствительностью и внешней клинической валидацией [Препринт]. 2026. arXiv:2606.13135 [cs.CV]. DOI: [10.48550/arXiv.2606.13135](https://doi.org/10.48550/arXiv.2606.13135)
  21. Kozachok E.S., Seregin S.S., Kozachok A.V., et al. Cascade classification of dermoscopic images of skin neoplasms with controllable sensitivity and external clinical validation [Preprint]. 2026. arXiv:2606.13135 [cs.CV]. (In Russ.) DOI: [10.48550/arXiv.2606.13135](https://doi.org/10.48550/arXiv.2606.13135)
  22. Tun H.M., Abdul Rahman H., Naing L., Malik O.A. Trust in artificial intelligence-based clinical decision support systems among health care workers: systematic review. *J. Med. Internet Res.* 2025;27:e69678. DOI: [10.2196/69678](https://doi.org/10.2196/69678)

## Информация о вкладе авторов

Козачок Е.С. – концепция и дизайн исследования, разработка системы Melanoscope AI и метода IoU-оценки интерпретируемости, сбор и анализ данных, написание и редактирование рукописи; Карась С.И. – критический пересмотр рукописи на предмет значимого интеллектуального содержания, научное и методологическое редактирование, утверждение окончательного варианта рукописи; Серегин С.С. – клиническая апробация, независимая экспертная оценка дерматоскопических снимков, направление пациентов на биопсию, верификация гистологических результатов.

**Конфликт интересов:** авторы декларируют отсутствие финансовых конфликтов интересов, связанных с публикацией настоящей статьи. Авторы признают наличие структурного конфликта интересов: Козачок Е.С. является одновременно разработчиком системы Melanoscope AI, вторым экспертом-дерматологом в клинической апробации и аналитиком, ответственным за сбор и обработку данных. Независимая первичная экспертная оценка дерматоскопических снимков, направление пациентов на биопсию и верификация гистологических результатов выполнены Серегиним С.С. независимо; референсным служило заключение основного эксперта, не участвовавшего в разработке системы.

## Сведения об авторах

**Козачок Елена Сергеевна**, специалист, Институт системного программирования им. В.П. Иванникова РАН (ИСП РАН), Москва, Россия, e-mail: [e.kozachok@ispras.ru](mailto:e.kozachok@ispras.ru); <http://orcid.org/0009-0007-9432-1663>.

**Карась Сергей Иосифович**, д-р мед. наук, специалист, отдел координации научной и образовательной деятельности, НИИ кардиологии Томского НИМЦ, Томск, Россия, e-mail: [ksi@cardio-tomsk.ru](mailto:ksi@cardio-tomsk.ru); <http://orcid.org/0000-0001-6716-856X>.

**Серегин Сергей Сергеевич**, канд. мед. наук, врач-онколог, Орловский онкологический диспансер, Орел, Россия, e-mail: [serega\\_s2004@mail.ru](mailto:serega_s2004@mail.ru); <http://orcid.org/0000-0002-7248-402X>.

Поступила 01.06.2026;  
рецензия получена 28.08.2026;  
принята к публикации 02.09.2026.

## Information on author contributions

Kozachok E. S.– study concept and design, development of the Melanoscope AI system and the IoU interpretability assessment method, data collection and analysis, manuscript writing and editing; Karas S. I.– critical revision of the manuscript for important intellectual content, scientific and methodological editing, approval of the final version of the manuscript; Seregin S. S.– clinical validation, independent expert assessment of dermoscopic images, patient referral for biopsy, histological result verification.

**Conflict of interest:** the authors declare no financial conflict of interests related to the publication of this article. The authors acknowledge a structural conflict of interest: Kozachok E.S. is concurrently the developer of the Melanoscope AI system, the second dermatologist expert in the clinical validation, and the analyst responsible for data collection and processing. Independent primary expert assessment of dermoscopic images, patient referral for biopsy, and histological verification were performed independently by Seregin S.S.; the reference standard was the conclusion of the principal expert, who was not involved in the development of the system.

## Information about the authors

**Elena S. Kozachok**, Specialist, Ivannikov Institute for System Programming of the Russian Academy of Sciences (ISP RAS), Moscow, Russia, e-mail: [e.kozachok@ispras.ru](mailto:e.kozachok@ispras.ru); <http://orcid.org/0009-0007-9432-1663>.

**Sergey I. Karas**, Dr. Sci. (Med.), Specialist, Department of Coordination of Scientific and Educational Activities, Research Institute of Cardiology, Tomsk NRMC, Tomsk, Russia, e-mail: [ksi@cardio-tomsk.ru](mailto:ksi@cardio-tomsk.ru); <http://orcid.org/0000-0001-6716-856X>.

**Sergey S. Seregin**, Cand. Sci. (Med.), Oncologist, Orel Regional Oncology Dispensary, Orel, Russia, e-mail: [serega\\_s2004@mail.ru](mailto:serega_s2004@mail.ru); <http://orcid.org/0000-0002-7248-402X>.

Received 01.06.2026;  
review received 28.08.2026;  
accepted for publication 02.09.2026.