

DOI: <https://doi.org/10.18027/2224-5057-2025-064>

Применение технологий искусственного интеллекта для прогнозирования риска рецидива при раке поджелудочной железы. Систематический обзор литературы и мета-анализ

М. Ш. Манукян¹, В. И. Павлова^{1,2,3}, Р. Ш. Абдулаева¹, Т. Г. Геворкян¹, С. С. Гордеев^{1,3,4}

¹ ФГБУ «Национальный медицинский исследовательский центр онкологии им. Н. Н. Блохина» Минздрава России; Россия, 115478 Москва, Каширское шоссе, 23;

² ГАУЗ ТО «Многопрофильный клинический медицинский центр «Медицинский город»; Россия, 625041 Тюмень, ул. Барнаульская, 32;

³ ФГБОУ ВО «Тюменский государственный медицинский университет» Минздрава России; Россия, 625023 Тюмень, ул. Одесская, 54;

⁴ ФГАОУ ВО «Первый Московский государственный медицинский университет им. И. М. Сеченова» Минздрава России (Сеченовский Университет); Россия, 119991, г. Москва, ул. Трубецкая, д. 8, стр. 2

Контакты: Мариам Ширавовна Манукян manukyanyamariam@gmail.com

Проведен систематический обзор и мета-анализ 10 исследований (2019–2024 гг.), оценивающих диагностическую точность алгоритмов искусственного интеллекта (ИИ) для прогнозирования рецидивов рака поджелудочной железы (РПЖ). Объединенные оценки чувствительности и специфичности составили 0,77 [95% ДИ: 0,58–0,95] и 0,79 [95% ДИ: 0,57–1,00] соответственно. Ключевыми ограничениями работы являлась высокая гетерогенность ($I^2 > 98\%$), которая может быть связана с малым числом включенных исследований, и недостаточная стандартизация методов валидации.

Введение: Использование технологий искусственного интеллекта открывает новые возможности в прогнозировании течения рака поджелудочной железы.

Цель: Проведение мета-анализа диагностической точности алгоритмов ИИ (чувствительности и специфичности) для прогнозирования рецидивов РПЖ и сравнительный анализ эффективности различных типов алгоритмов.

Методы: Был проведен систематический поиск литературы в ведущих научных базах данных, охватывающий публикации за период с 2019 по 2024 годы. В обзор включены исследования, в которых применялись методологии искусственного интеллекта для прогнозирования риска рецидива рака поджелудочной железы. Поиск и анализ данных осуществлялись в три этапа: первичный поиск исследований по ключевым словам и критериям включения; скрининг заголовков и аннотаций для отбора релевантных работ; детальная оценка полных текстов отобранных статей.

Синтез данных включал анализ производительности моделей ИИ, типов используемых данных (клинические, геномные, радиологические и др.), а также стратегий валидации и тестирования предложенных алгоритмов.

Для мета-анализа чувствительности и специфичности использована модель случайных эффектов с расчетом объединенных оценок, 95% доверительных интервалов и показателей гетерогенности (I^2 , τ^2). Дополнительно выполнена мета-регрессия для оценки влияния типа алгоритма на чувствительность. Статистический анализ проведен в R (пакет *metafor*) с визуализацией лесных диаграмм.

Результаты: Данный систематический обзор включил 10 исследований, из которых 5 были отобраны для мета-анализа. Результаты демонстрируют объединенную чувствительность 0,77 [95% ДИ: 0,58–0,95] и специфичность 0,79 [95% ДИ: 0,57–1,00] алгоритмов ИИ для прогнозирования рецидивов РПЖ. При анализе отдельных типов алгоритмов искусственные нейронные сети (ANN) показали объединенную чувствительность 0,87 [0,73–1,01], а метод опорных векторов (SVM) имел отрицательный коэффициент влияния на чувствительность (–0,45 [–0,69 – –0,21]). Мета-анализ выявил высокую гетерогенность между исследованиями ($I^2 = 98,84\%$ для чувствительности и $I^2 = 99,42\%$ для специфичности), что требует осторожности при интерпретации результатов.

Заключение: ИИ-модели демонстрируют потенциал для прогнозирования рецидивов РПЖ, но требуют стандартизации данных и проспективной валидации в клинической практике.

Ключевые слова: рак поджелудочной железы, ранний рецидив, искусственный интеллект, машинное и глубокое обучение, прогностические факторы

Для цитирования: Манукян М.Ш., Павлова В.И., Абдулаева Р.Ш. и соавт. Применение технологий искусственного интеллекта для прогнозирования риска рецидива при раке поджелудочной железы. Систематический обзор литературы и мета-анализ. Злокачественные опухоли 2025;15(4):56–64. DOI: <https://doi.org/10.18027/2224-5057-2025-064>

Using artificial intelligence tools to predict recurrence risk in pancreatic cancer. A systematic literature review and meta-analysis

M. Sh. Manukyan¹, V. I. Pavlova^{1, 2, 3}, R. Sh. Abdulaeva¹, T. G. Gevorkyan¹, S. S. Gordeev^{1, 3, 4}

¹ N. N. Blokhin National Medical Research Center of Oncology, Ministry of Health of Russia; 23 Kashirskoe Shosse, Moscow 115478, Russia

² Multidisciplinary Clinical Medical Center "Medical City"; 32 Barnaulskaya St., Tyumen 625041, Russia;

³ Tyumen State Medical University, Ministry of Health of Russia; 54 Odesskaya St., Tyumen 625023, Russia;

⁴ I. M. Sechenov First Moscow State Medical University, Ministry of Health of Russia (Sechenov University); Build. 2, 8 Trubetskaya St., Moscow 119991, Russia

Contacts: Mariam Shirakovna Manukyan manukyanmariam6@gmail.com

A systematic review and meta-analysis of 10 studies (2019–2024) evaluating the diagnostic accuracy of artificial intelligence (AI) algorithms for predicting pancreatic cancer (PCa) recurrence was conducted. The pooled sensitivity and specificity estimates were 0.77 [95% CI: 0.58–0.95] and 0.79 [95% CI: 0.57–1.00], respectively. Key limitations of the study included high heterogeneity ($I^2 > 98\%$), which could be related to the small number of included studies and insufficient standardization of the validation methods.

Background: Artificial intelligence (AI) tools provide new possibilities in predicting the course of pancreatic cancer.

Purpose: To conduct a meta-analysis of the diagnostic accuracy of AI algorithms (sensitivity and specificity) for predicting PCa recurrence and to compare the effectiveness of different types of algorithms. **Methods:** A systematic literature search was conducted in leading scientific databases, covering publications from 2019 to 2024. The review included studies that applied artificial intelligence tools to predict the risk of pancreatic cancer recurrence. The data search and analysis were conducted in three stages: a primary search of studies using keywords and inclusion criteria; screening of the titles and abstracts to select relevant studies; and a detailed assessment of the full texts of the selected articles.

The data synthesis included an analysis of the performance of the AI models, the types of data used (clinical, genomic, radiological, etc.), and the validation and testing strategies for the proposed algorithms.

A random-effects model was used for the sensitivity and specificity meta-analysis, with the calculation of pooled estimates, 95% confidence intervals, and heterogeneity indices (I^2 , τ^2). A meta-regression was also performed to assess the impact of the algorithm type on sensitivity. Statistical analysis was carried out in R (metafor package) with forest plot visualization.

Results: This systematic review included 10 studies, of which 5 were selected for the meta-analysis. The results demonstrate a pooled sensitivity of 0.77 [95% CI: 0.58–0.95] and specificity of 0.79 [95% CI: 0.57–1.00] for AI algorithms in predicting PCa recurrence. In the analysis of individual algorithm types, artificial neural networks (ANNs) showed a pooled sensitivity of 0.87 [0.73–1.01], while support vector machines (SVMs) had a negative impact on sensitivity (coefficient -0.45 [-0.69 to -0.21]). The meta-analysis revealed high heterogeneity of the studies ($I^2 = 98.84\%$ for sensitivity and $I^2 = 99.42\%$ for specificity), requiring cautious interpretation of the results.

Conclusion: AI models are promising tools for predicting pancreatic cancer recurrence, but require data standardization and prospective validation in clinical practice.

Keywords: pancreatic cancer, early recurrence, artificial intelligence, machine and deep learning, prognostic factors

For citation: Manukyan M.Sh., Pavlova V.I., Abdulaeva R.Sh., et al. Using artificial intelligence tools to predict recurrence risk in pancreatic cancer. A systematic literature review and meta-analysis. *Zlokachestvennyye opuholi = Malignant Tumors* 2025;15(4):56–64 (In Russ.). DOI: <https://doi.org/10.18027/2224-5057-2025-064>

ВВЕДЕНИЕ

Выявление пациентов с высоким риском прогрессирования рака поджелудочной железы (РПЖ) позволяет рассмотреть возможность изменения тактики лечения, включая эскалацию режимов терапии, проведение неоадъювантной терапии или оптимизацию подходов к динамическому наблюдению после радикального лечения. Согласно данным литературы, своевременное выявление прогрессирования РПЖ и раннее начало лечения положи-

тельно влияют на показатели общей выживаемости (ОВ) [1]. Однако в настоящее время в клинической практике отсутствуют надежные инструменты для прогнозирования прогрессирования, что затрудняет выделение групп пациентов, требующих более интенсивного наблюдения.

В 2024 году был проведен мета-анализ 64 исследований, посвященных изучению различных прогностических факторов раннего рецидива РПЖ [1]. В анализируемых работах было выделено несколько статистически значимых параметров: уровень онкомаркера СА19-9, отсутствие

адьювантной химиотерапии, поражение лимфатических узлов, степень дифференцировки опухоли и ее размеры. На основе этих данных были разработаны номограммы, демонстрирующие C-индекс в диапазоне от 0,656 до 0,734. Несмотря на определенную прогностическую способность, эти шкалы не нашли широкого применения в клинической практике из-за их ограниченной эффективности [2–4].

Исследователи сталкиваются с рядом ограничений, включая гетерогенность пациентов, необходимость формирования больших выборок и наличие неочевидных факторов, влияющих на прогноз, которые не учитываются в классических дизайнах исследований. Эти ограничения подчеркивают необходимость разработки новых подходов к анализу данных и прогнозированию.

Быстрый прогресс в области вычислительных технологий и расширение технических возможностей делают искусственный интеллект (ИИ) многообещающим инструментом для решения указанных проблем. Однако остается неясным, насколько эти подходы превосходят традиционные номограммы, особенно с учетом значительных интеллектуальных и финансовых затрат, связанных с внедрением ИИ [5–7].

Несмотря на растущее количество исследований, посвященных применению ИИ в прогнозировании РПЖ, остается неясным, какие типы данных и методологии ИИ обеспечивают наиболее надежные результаты. В данном обзоре мы систематизируем результаты различных исследований, посвященных применению ИИ для прогнозирования рецидивов и прогрессирования РПЖ на основе клинических параметров, а также оценим потенциальное влияние этих подходов на клиническую практику.

МАТЕРИАЛЫ И МЕТОДЫ

Для достижения поставленной цели был проведен систематический поиск научной литературы, охватывающий статьи, опубликованные в период с 2019 по 2024 годы. Поиск осуществлялся в четырех ведущих онлайн-базах данных: PubMed, ScienceDirect, NATURE, MedRxiv, BioRxiv и Google Scholar. Основным поисковым запросом был сформулирован следующим образом: «("Artificial intelligence" OR "Machine learning" OR "Deep learning" OR "supervised learning" OR "unsupervised learning" OR "reinforcement learning") AND ("Pancreatic Cancer" OR "Pancreatic adenocarcinoma") AND (diagnose OR detect OR predict* OR screen*)». Этот запрос использовался для поиска в базах данных PubMed и Google Scholar. Однако из-за ограничений на количество символов в других базах данных (ScienceDirect, NATURE, MedRxiv и BioRxiv) был применен сокращенный вариант запроса: *«(Artificial intelligence OR Machine learning OR Deep learning) AND (Pancreatic Cancer) AND (predict) AND (Pancreas)».

В обзор были включены только те исследования, которые соответствовали определенным критериям. Рассматривались работы, посвященные применению методов искусственного интеллекта (ИИ) для прогнозирования

риска прогрессирования рака поджелудочной железы, с акцентом на клинические параметры. Включены статьи из рецензируемых журналов, при этом клинические случаи, материалы конференций, обзоры, диссертации, аннотации и редакционные статьи были исключены. Исследования, использующие не-ИИ методы для диагностики рака поджелудочной железы, а также работы, предоставляющие исключительно теоретическую основу для моделей ИИ без практического применения, также не рассматривались. Ограничений по методам, дизайну или географическому расположению исследований не применялось, однако учитывались только публикации на английском языке.

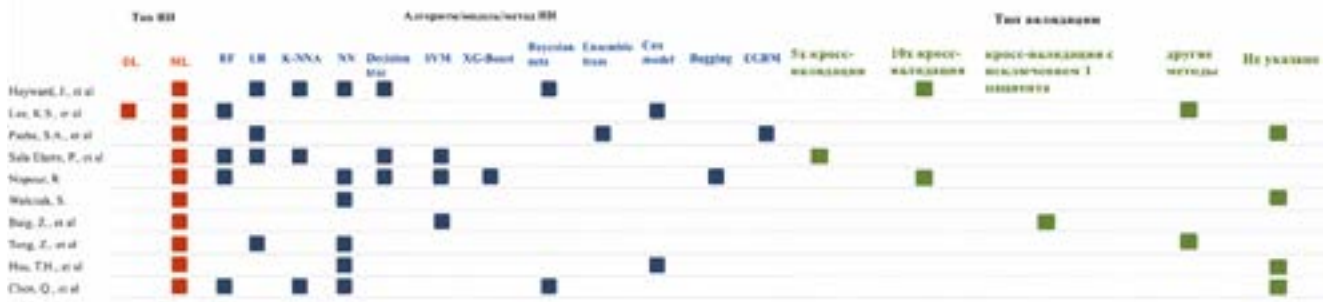
Процесс отбора исследований состоял из трех этапов. На первом этапе был проведен поиск литературы в указанных базах данных, после чего для удаления дубликатов среди найденных исследований использовался инструмент Rayyan. На втором этапе два независимых рецензента проанализировали заголовки и аннотации всех найденных статей, исключив исследования, не соответствующие теме обзора. На заключительном этапе рецензенты независимо рассмотрели полные тексты статей, прошедших предыдущий этап, а все несоответствия между рецензентами разрешались в ходе обсуждения. Для оценки уровня согласия между рецензентами был рассчитан коэффициент Каппа Козна, значение которого составило 0,93 для анализа полных текстов, что свидетельствует о высоком уровне согласия.

После извлечения данных из исследований был проведен нарративный синтез. Основной фокус — методы ИИ (SVM, нейронные сети, случайные леса), их цели, характеристики, источники данных и языки программирования. Также анализировались метрики: точность, специфичность, чувствительность, прецизионность. Данные систематизировали в Microsoft Excel.

Далее мы провели мета-анализ исследований для оценки специфичности и чувствительности различных алгоритмов искусственного интеллекта, а также мета-регрессию для оценки влияния различных алгоритмов искусственного интеллекта (ИИ) на показатель чувствительности. В связи с ожидаемой гетерогенностью между исследованиями при мета-анализе использована модель случайных эффектов (Random-Effects Model). Проведена оценка общего объединенного эффекта (с 95% доверительным интервалом), степень гетерогенности (τ^2 , I^2 , Q-тест) отдельно для специфичности и чувствительности. Анализ выполнен в программном обеспечении R (пакет metafor) с визуализацией лесных диаграмм.

РЕЗУЛЬТАТЫ ПОИСКА

Сначала мы идентифицировали 23539 статей, используя 6 баз данных: PubMed (n = 417), Science Direct (n = 3587), NATURE (n = 537), Google Scholar (n = 18100), BioRxiv (n = 737) и MedRxiv (n = 161). Все статьи из PubMed были проанализированы на основе заданного запроса. Из-за большого объема статей из Science Direct, BioRxiv и Google Scholar



DL (Deep Learning) — Глубокое обучение;
 ML (Machine Learning) — Машинное обучение;
 RF (Random Forest) — Случайный лес;
 LR (Logistic Regression) — Логистическая регрессия;
 K-NNA (K-Nearest Neighbors) — Метод k-ближайших соседей;
 NN (Neural Networks) — Нейронные сети;
 Decision tree — дерево решений;
 SVM (Support Vector Machine) — Метод опорных векторов;

XG-Boost (Extreme Gradient Boosting) — Экстремальный градиентный бустинг;
 Bayesian nets — Байесовские сети;
 Ensemble trees — ансамблевые деревья;
 Cox model — модель пропорциональных рисков Кокса;
 Bagging — Бэггинг;
 EGBM (Extreme Gradient Boosting Machine) — Экстремальный градиентный бустинг;

Рисунок 1. Типы использованных методов искусственного интеллекта (n = 10 исследований)

Figure 1. Types of artificial intelligence methods used (n = 10 studies)

Таблица 1. Характеристики техник ИИ, использованных в каждом исследовании

Table 1. Characteristics of the AI techniques used in each study

Автор, год	Использованные алгоритмы	Измеряемый результат	Результаты	Количество пациентов
Hayward J. et al., 2010 [8]	Байесовские сети, деревья решений, k-ближайшие соседи, нейронные сети	ECOG PS 6 месяцев	Точность: 49% без ИИ, 65% с ИИ. Байесовская модель: 0,56 без ИИ, 0,79 с ИИ.	87
Walczak S. et al., 2017 [9]	Модель нейронной сети	7-месячная общая выживаемость (ОВ)	Чувствительность 91%, специфичность 38%	219
Elarre S.P. et al., 2019 [10]	Логистическая регрессия, дерево решений, случайный лес, метод опорных векторов (SVM), k-ближайшие соседи (KNN)	Рецидив через 2 года после операции	Чувствительность 0,70, специфичность 0,73, точность 0,71 (95% ДИ 0,56–0,84, p = 0.005), средний AUC0,75	45 (дополнительно внешняя когорта пациентов для валидации n = 11)
Tong Z. et al., 2020 [11]	Три модели нейронных сетей	8-месячная общая выживаемость (ОВ)	3 модели нейронных сетей против логистической регрессии AUC (0,811 против 0,680; 0,844 против 0,722; 0,921 против 0,849, все p < 0,05). Чувствительность 0,8241, специфичность 0,8961.	168 (в тренировочной когорте — 133; в когорте валидации — 35)
Lee K. et al., 2021 [7]	Случайный лес и модель пропорциональных рисков Кокса	Выживаемость без прогрессирования	C-индекс 0,6805 (RF) и 0,7738 (модель Кокса)	4846 (в тренировочной когорте — 3635; в когорте валидации — 1211)
Baig Z. et al., 2021 [12]	Нелинейный метод опорных векторов	2-летняя общая выживаемость (ОВ)	Точность 75%, чувствительность 41,9%, специфичность 98%	93 (в тренировочной когорте — 83; в когорте валидации — 10)
Hsu T. et al., 2021 [13]	Кривые Каплана-Мейера и регрессионный анализ Кокса	Общая выживаемость (ОВ)	Средняя выживаемость 15 ± 12 против 22 ± 12 (p < 0.05), скорректированный HR 1,58 (95% ДИ 1,03–3,33)	136
Chen Q. et al., 2024 [14]	k-ближайшие соседи, нейронные сети, наивный Байес, случайный лес	5-летняя выживаемость (ОВ)	0,7345-летняя выживаемость, 0,7955-летняя общая выживаемость AUC	24044 (в тренировочной когорте — 19235; в тестовой когорте — 4809)
Pasha S. et al., 2024 [15]	Логистическая регрессия с лассо-регуляризацией, модель градиентного бустинга	Завершение лечения	AUC0,67	208
Norpour R. et al., 2024 [16]	6 алгоритмов ML: случайный лес, бэггинг, XG-Boost (лучший результат)	Общая выживаемость (ОВ)	Чувствительность 94,96%, специфичность 93,62%, точность 94,55%, F-мера 96,00%, AUC0,933 (95% ДИ 0,906–0,958)	654 (дополнительно внешняя когорта пациентов для валидации n = 52)

были пересмотрены только первые 200 результатов (отсортированных по релевантности). В общей сложности для детального анализа было включено 1006 статей.

Из 1006 проанализированных статей 966 были исключены по следующим причинам: 317 статей не относились к искусственному интеллекту, 185 статей не были сосредоточены на раке поджелудочной железы, 191 статья являлась литературным обзором, и 273 статьи не имели отношения к риску рецидива. В конечном итоге было выделено 40 статей, из них 30 были исключены в связи с несоответствием изучаемых параметров (включались не только клинические данные, но и радиомика, геномика или патомика). В конечном итоге было включено в анализ 10 статей.

ХАРАКТЕРИСТИКИ ВКЛЮЧЕННЫХ СТАТЕЙ

Все включенные исследования были опубликованы в рецензируемых журналах (10/10, 100%). Количество участников в включенных исследованиях варьировало от 45 до 24044, медиана составила 188 пациентов.

ХАРАКТЕРИСТИКИ ИСПОЛЬЗУЕМЫХ ТЕХНИК ИИ

Типы используемых техник искусственного интеллекта представлены на рисунке 1.

Во всех 10 исследованиях (100%) использовали алгоритмы машинного обучения, в одной работе также применялся алгоритм глубокого обучения (10%). В исследованиях применялись различные алгоритмы ИИ, нейронные сети использовались в 60% работ, случайный лес — 40%, а дерево решений — в 30% случаев. Подробная информация касательного характеристик применяемых алгоритмов представлена в таблице 1.

Из методов валидации 10-кратная кросс-валидация применялась в 20% исследований. В 40% работ метод валидации не был указан.

Среди изученных алгоритмов XGBoost (экстремальный градиентный бустинг) и нейронные сети продемонстрировали следующие результаты.

При XGBoost чувствительность составила 94,96%, специфичность — 93,62%, AUC0,933 (95% ДИ 0,906–0,958). Результаты применения нейронных сетей варьировались в исследованиях, чувствительность данного алгоритма составила 82–91% и специфичность — 38–89%.

Отдельное внимание стоит уделить параметрам, анализируемым в исследованиях. Большинство из них имело схожий спектр оцениваемых характеристик, но лабораторные показатели и проведение нео- или адьювантной терапии отражены в 30% работ, а сопутствующие заболевания оценивались в 40% исследований. Подробнее оцениваемые параметры отражены в таблице 2.

МЕТА-АНАЛИЗ

Результаты чувствительности и специфичности метода были представлены в 5 исследованиях, на основе которых проведен мета-анализ.

Чувствительность

В анализ вошли данные пяти независимых исследований ($n = 45–654$), в которых сравнивались следующие модели:

- Искусственные нейронные сети (ANN),
- Логистическая регрессия, метод опорных векторов (SVM), случайный лес, k-ближайших соседей (KNN),
- Нелинейный SVM,
- XGBoost.

Таблица 2. Основные клинические параметры, используемые в исследованиях

Table 2. Main clinical parameters used in the studies

Категория параметров	Параметры	Ссылки на исследования	% статей
Демографические данные	Возраст, пол, национальность, ИМТ, курение, алкоголизм, наследственность	1–10	100%
Опухолевые характеристики	Размер опухоли, стадия (T, N), степень дифференцировки, гистологический тип, локализация первичной опухоли	1–10	100%
Лабораторные показатели	Уровень СА19–9, РЭА, билирубин, альбумин, АЛТ, АСТ, креатинин	3,6,8	30%
Хирургические параметры	Вид операции, края резекции, вовлеченность сосудов, перинеуральная/лимфо-васкулярная инвазия Резекция портальной вены Послеоперационные осложнения	1,2,4,5,7,8,10	70%
Лечение	Неoadьювантная химиотерапия, адьювантная химиотерапия, лучевая терапия	4,6,7	30%
Сопутствующие заболевания	Сахарный диабет 2 типа, сердечная недостаточность, вирусный гепатит В	3,6,7,8	40%
Иные признаки	ECOG, саркопения, механическая желтуха и проводилась ли декомпрессия желчных протоков	1,4,9	30%

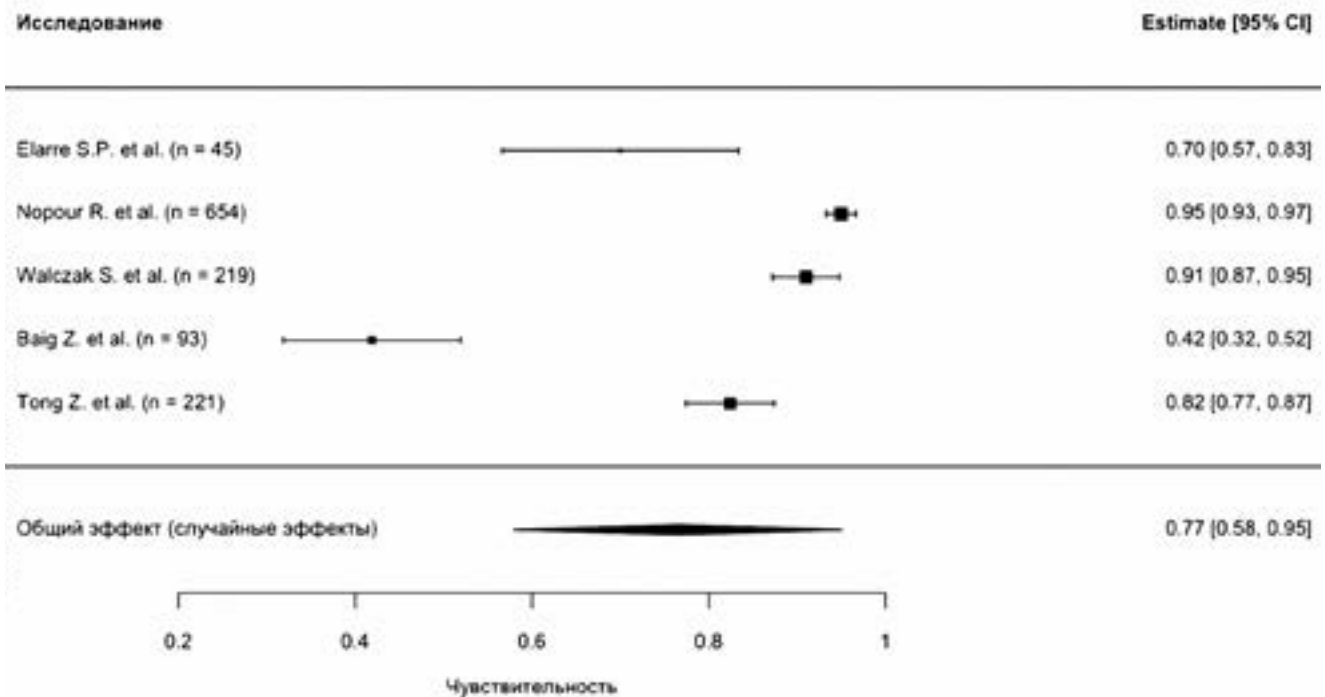


Рисунок 2. График forest plot объединенных оценок чувствительности по включенным исследованиям

Figure 2. Forest plot of pooled sensitivity estimates from included studies

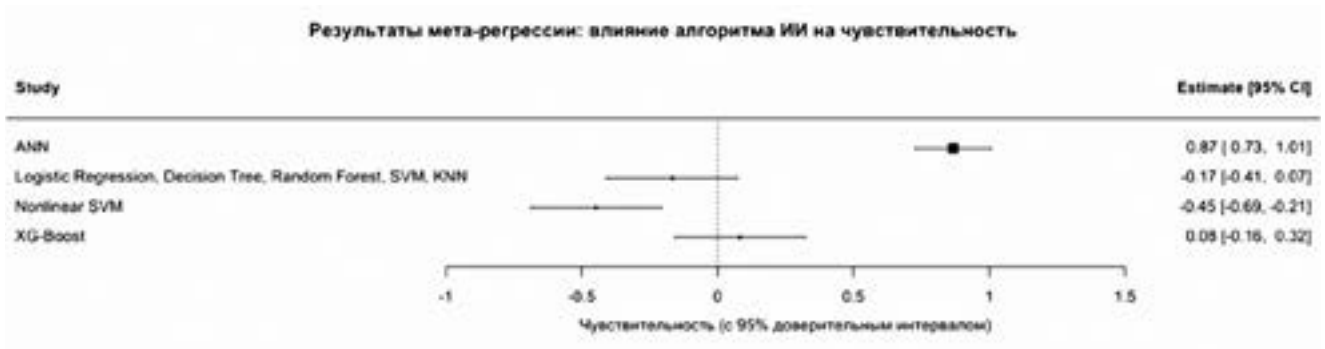


Рисунок 3. График forest plot для оценки влияния алгоритма ИИ на чувствительность

Figure 3. Forest plot for assessing the impact of the AI algorithm on sensitivity

Учитывая малое число включенных исследований и ожидаемую высокую гетерогенность, применена модель случайных эффектов (Random-Effects Model).

Объединенная оценка чувствительности составила 0,77 [95% ДИ: 0,58–0,95], что указывает на умеренно высокую диагностическую точность (рис. 2). Однако была получена высокая гетерогенность между исследованиями $I^2 = 98,84\%$ ($p < 0,0001$), при $\tau^2 = 0,0436$ ($SE = 0,0319$). Высокий I^2 при относительно низком τ^2 может объясняться малым числом исследований ($k = 5$). Диапазон чувствительности составил от 0,42 [0,32–0,52] [12] до 0,95 [0,93–0,97] [16].

Анализ влияния типа алгоритма на показатели чувствительности выявил следующее.

Алгоритмы ANN показали коэффициент 0,87 [95% ДИ: 0,73–1,01], алгоритм SVM продемонстрировал коэффициент $-0,45$ [95% ДИ: $-0,69$ — $-0,21$] (рис. 3).

Специфичность

На основе тех же 5 исследований проведен мета-анализ для оценки специфичности различных алгоритмов искусственного интеллекта. Также применена модель случайных эффектов (Random-Effects Model).

Объединенная оценка специфичности составила 0,79 [95% ДИ: 0,57–1,00], что указывает на высокую диагностическую точность. Однако обнаружена высо-

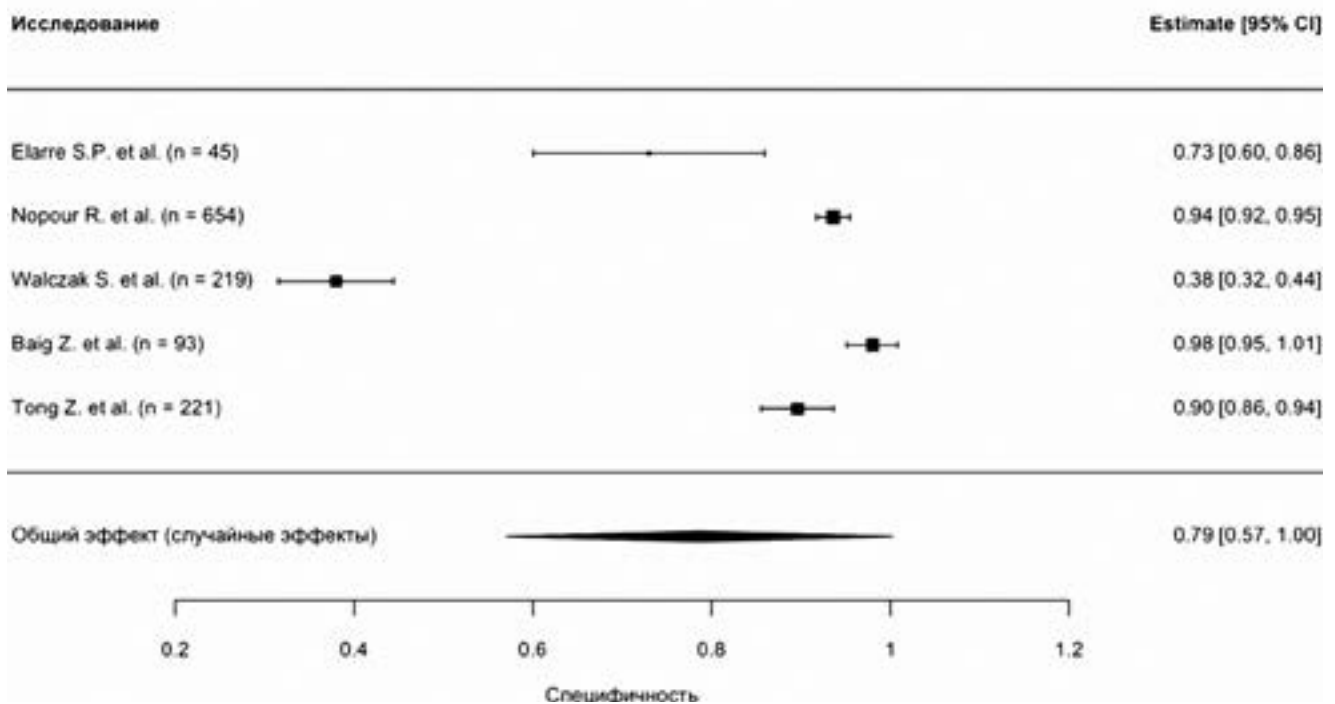


Рисунок 4. График forest plot объединенных оценок специфичности по включенным исследованиям

Figure 4. Forest plot of pooled specificity estimates from included studies

Таблица 3. Основные показатели гетерогенности в мета-анализах

Table 3. Main indicators of heterogeneity in meta-analyses

Параметр	Чувствительность (Мета-анализ)	Специфичность (Мета-анализ)	Интерпретация различий
Общий эффект [95% ДИ]	0,77 [0,58; 0,95]	0,79 [0,57; 1,00]	Схожая средняя точность, но шире ДИ у специфичности
τ^2 (гетерогенность)	0,0436 (SE = 0,0319) низкая/умеренная	0,0596 (SE = 0,0430) умеренная	Умеренно выше вариативность у специфичности
I^2 (%)	98,84%	99,42%	Крайне высокая гетерогенность в обоих случаях
H^2	86,18	171,95	В 2 раза больше вариаций между исследованиями для специфичности
Q-тест (p-value)	< 0,0001	< 0,0001	Статистически значимая гетерогенность

кая гетерогенность между исследованиями $I^2 = 99,42\%$ ($p < 0.0001$). Показатель τ^2 составил 0,0596 (SE = 0,0430), что соответствует умеренной гетерогенности. Такие различия вероятно связаны с малым число исследований.

Наблюдается разброс показателей специфичности: так, в работе Baig Z. et al. результат был равен 0,98 [0,95–1,01], а в исследовании Walczak S. et al. — 0,38 [0,32–0,44]. В трех оставшихся работах специфичность была выше 0,9. (рис. 4)

В таблице 3 представлены показатели гетерогенности между мета-анализами чувствительности и специфичности.

ОБСУЖДЕНИЕ

Анализ данных литературы показал, что методы машинного обучения (ML) и глубокого обучения (DL) активно применяются для прогнозирования отдаленных результатов лечения РПЖ.

Проведенный мета-анализ продемонстрировал высокую прогностическую эффективность алгоритмов ИИ в предсказании рецидивов рака поджелудочной железы. Наиболее стабильным и эффективным алгоритмом для оценки чувствительности модели был ANN (0,87 [0,73–1,01]), но полученные результаты требуют осторожной интерпретации. Основной проблемой стала высокая гетерогенность ($I^2 > 98\%$), обусловленная разнородностью используемых алгоритмов, различиями в раз-

мерах выборок и методологических подходах. Например, только 30% исследований учитывали уровень онкомаркера СА19-9, а методы валидации в 40% работ не были указаны. Малый объем включенных исследований ($k = 5$) привел к широким доверительным интервалам, что снижает надежность выводов. Кроме того, потенциальное смещение публикаций, связанное с включением только работ, сообщающих положительные результаты, могло исказить общие оценки. Для преодоления этих ограничений необходима стандартизация протоколов исследований, включая обязательное описание методов валидации и использование единых критериев оценки клинических параметров. Помимо этого, есть еще несколько трудностей для внедрения ИИ в клиническую практику. Для обучения ИИ-моделей требуются большие объемы данных, что может быть ограничено в условиях реальной клиниче-

ской практики. Также внедрение ИИ-технологий требует значительных финансовых и интеллектуальных затрат, что может стать барьером для их широкого применения. Добавление геномных и радиомических параметров в модели ИИ могут позволить повысить их точность, но в тоже время интеграция изображений и молекулярных сигнатур в модели приведет к еще большим сложностям к их интеграции в клиническую практику. Поэтому вопрос об оптимизации учитываемых клинических параметров остается важным для улучшения прогностической ценности модели.

Таким образом, несмотря на положительные результаты, дальнейшие исследования должны быть направлены на устранение методологических недостатков и обеспечение воспроизводимости моделей в реальной клинической среде.

ЛИТЕРАТУРА/REFERENCES

1. SEER Cancer Statistics: Pancreas. U.S. Department of Health and Human Services, National Cancer Institute. Available at: <https://seer.cancer.gov/statfacts/html/pancreas.html>
2. Leonhardt C.S., Gustorff C., Klaiber U., et al. Prognostic Factors for Early Recurrence After Resection of Pancreatic Cancer: A Systematic Review and Meta-Analysis. *Gastroenterology* 2024;167(5):977–992. <https://doi.org/10.1053/j.gastro.2024.05.028>
3. Song W., Miao D.L., Chen L. Nomogram for predicting survival in patients with pancreatic cancer. *Onco Targets Ther* 2018;11:539–545. <https://doi.org/10.2147/OTT.S154599>
4. Goldstein D, Von Hoff D.D., Chiorean E.G., et al. Nomogram for estimating overall survival in patients with metastatic pancreatic cancer. *Pancreas* 2020;49(6):744–750. <https://doi.org/10.1097/MPA.0000000000001563>
5. Tran K.A., Kondrashova O., Bradley A., et al., Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med* 2021;13(1):152. <https://doi.org/10.1186/s13073-021-00968-x>
6. Palumbo D., Mori M., Prato F., et al. Prediction of early distant recurrence in upfront resectable pancreatic adenocarcinoma: a multidisciplinary, machine learning-based approach. *Cancers (Basel)* 2021;13(19):4938. <https://doi.org/10.3390/cancers13194938>
7. Lee, K.S., et al. Usefulness of artificial intelligence for predicting recurrence following surgery for pancreatic cancer: Retrospective cohort study. *Int J Surg*, 2021.93:p.106050.
8. Hayward, J., Alvarez S.A., Ruiz C., et al., Machine learning of clinical performance in a pancreatic cancer database. *Artif Intell Med* 2010;49(3):187–95. <https://doi.org/10.1016/j.artmed.2010.04.009>
9. Walczak S., Velanovich V. An evaluation of artificial neural networks in predicting pancreatic cancer survival. *J Gastrointest Surg* 2017;21(10):1606–1612. <https://doi.org/10.1007/s11605-017-3518-7>
10. Sala Elarre P., Oyaga-Iriarte E., Yu K.H., et al., Use of Machine-Learning Algorithms in Intensified Preoperative Therapy of Pancreatic Cancer to Predict Individual Risk of Relapse. *Cancers (Basel)* 2019;11(5):606. <https://doi.org/10.3390/cancers11050606>
11. Tong Z., Liu Y., Ma H., et al. Development, validation and comparison of artificial neural network models and logistic regression models predicting survival of unresectable pancreatic cancer. *Front Bioeng Biotechnol* 2020;8:196. <https://doi.org/10.3389/fbioe.2020.00196>
12. Baig Z., Abu-Omar N., Khan R., et al. Prognosticating outcome in pancreatic head cancer with the use of a machine learning algorithm. *Technol Cancer Res Treat* 2021;20:15330338211050767. <https://doi.org/10.1177/15330338211050767>
13. Hsu T.H., Schawkat K., Berkowitz S.J., et al., Artificial intelligence to assess body composition on routine abdominal CT scans and predict mortality in pancreatic cancer- A recipe for your local application. *Eur J Radiol* 2021;142:109834. <https://doi.org/10.1016/j.ejrad.2021.109834>
14. Chen, Q., Hu Y., Lin W., et al. Studying the impact of marital status on diagnosis and survival prediction in pancreatic ductal carcinoma using machine learning methods. *Sci Rep* 2024;14(1):5273. <https://doi.org/10.1038/s41598-024-53145-6>

15. Pasha S.A., Khalid A., Levy T., et al. Machine learning to predict completion of treatment for pancreatic cancer. *J Surg Oncol* 2024;130(8):1605–1610. <https://doi.org/10.1002/jso.27812>
16. Nopour R. Establishment of prediction model for mortality risk of pancreatic cancer: a retrospective study. *BMC Med Inform Decis Mak* 2024;24(1):181. <https://doi.org/10.1186/s12911-024-02590-4>

ВКЛАД АВТОРОВ

С. С. Гордеев, М. Ш. Манукян: концепция и дизайн исследования, статистическая обработка;
М. Ш. Манукян, Р. Ш. Абдулаева, Т. Г. Геворкян, В. И. Павлова: сбор и обработка материалов;
М. Ш. Манукян: написание текста;
С. С. Гордеев: редактирование текста.

Все авторы одобрили финальную версию статьи перед публикацией, выразили согласие нести ответственность за все аспекты работы, подразумевающую надлежащее изучение и решение вопросов, связанных с точностью или добросовестностью любой части работы.

ORCID АВТОРОВ

Манукян Мариам Ширавковна
<https://orcid.org/0000-0002-5084-4872>
Павлова Валерия Игоревна
<https://orcid.org/0000-0002-0899-0809>
Абдулаева Рукият Шамильевна
<https://orcid.org/0009-0004-6399-963X>
Геворкян Тигран Гагикович
<https://orcid.org/0009-0008-3486-302X>
Гордеев Сергей Сергеевич
<https://orcid.org/0000-0002-9303-8379>

Конфликт интересов. Авторы заявляют об отсутствии возможных конфликтов интересов.

Финансирование исследования. Исследование проведено при финансовой поддержке аналитического центра правительства Российской Федерации (Соглашение No. 70-2024-000121 dd 29.03.2024. IGK 000000D730324P540002)

Статья поступила в редакцию журнала 25.08.2025, прошла рецензирование 02.09.2025, принята к печати 01.10.2025.

AUTHORS' CONTRIBUTION

S. S. Gordeev, M. Sh. Manukyan: concept and design of the study, statistical processing;
M. Sh. Manukyan, R. Sh. Abdulaeva, T. G. Gevorkyan, V. I. Pavlova: collection and processing of material;
M. Sh. Manukyan: text writing;
S. S. Gordeev: text editing.

All authors have approved the final version of the article before publication, agreed to assume responsibility for all aspects of the work, implying proper review and resolution of issues related to the accuracy or integrity of any part of the work.

ORCID OF AUTHORS

Manukyan Mariam Shirakovna
<https://orcid.org/0000-0002-5084-4872>
Pavlova Valeria Igorevna
<https://orcid.org/0000-0002-0899-0809>
Abdulaeva Rukiyat Shamiliyevna
<https://orcid.org/0009-0004-6399-963X>
Gevorkyan Tigran Gagikovich
<https://orcid.org/0009-0008-3486-302X>
Gordeev Sergey Sergeevich
<https://orcid.org/0000-0002-9303-8379>

Conflict of interest. The authors declare that there are no possible conflicts of interest.

Research funding. The study was conducted with the financial support of the Analytical Center for the Government of the Russian Federation (Agreement No. 70-2024-000121 dd 29.03.2024. IGK 000000D730324P540002)

Received 25 August 2025.
Reviewed 02 September 2025.
Accepted for publication 01 October 2025.