

Современные подходы машинного обучения в прогнозе сахарного диабета второго типа

**Буткова Т.В.¹, Куликова Л.И.^{*2}, Мальсагова К.А.¹, Петровский Д.В.¹,
Руднев В.Р.³, Рыскина Е.А.⁴, Куликов А.В.³, Кайшева А.Л.¹**

¹*Научно-исследовательский институт биомедицинской химии имени В.Н. Ореховича,
Москва, Россия*

²*Институт математических проблем биологии РАН — филиал Института
прикладной математики им. М.В. Келдыша РАН, Пуцино, Московская область, Россия*

³*Институт теоретической и экспериментальной биофизики РАН, Пуцино,
Московская область, Россия*

⁴*Национальный исследовательский университет «Высшая школа экономики», Москва,
Россия*

Аннотация. С 1990 по 2022 годы число людей с диабетом выросло с 200 до 830 миллионов. Большая часть приходится на инсулиннезависимый диабет 2 типа. Болезнь часто приводит к тяжёлым осложнениям: слепоте, почечной недостаточности, инфаркту, инсульту, ампутациям конечностей. По оценкам ВОЗ, в 2021 году от диабета и заболеваний почек умерло более двух миллионов человек. Глобальные расходы на здравоохранение, связанные с диабетом в целом, оценивались в 966 миллиардов долларов США в 2021 году и, по прогнозам, достигнут 1.05 миллиарда долларов США к 2045 году. Ранняя диагностика и индивидуальный подход к лечению помогают замедлить прогрессирование болезни и снизить риск сосудистых осложнений. На ранних стадиях изменение образа жизни может привести к длительной нормогликемии без применения лекарств. Одной из ключевых проблем остаётся профилактика гипогликемии, особенно ночной, которая может приводить к опасным осложнениям. Использование данных непрерывного мониторинга и алгоритмов машинного обучения позволяет прогнозировать риск гипогликемии с учётом индивидуальных особенностей пациента. Современные модели искусственного интеллекта строятся на анализе больших массивов данных. Решение задачи несбалансированности выборки повышает надёжность прогнозов. Для дальнейшего прогресса необходимо развитие больших библиотек данных и безопасных систем обмена информацией между устройствами мониторинга и приложениями. Применение таких технологий повышает эффективность терапии и качество жизни пациентов. Настоящий обзор освещает перспективы применения алгоритмов искусственного интеллекта в мониторинге и прогнозировании диабета.

Ключевые слова: сахарный диабет второго типа, ранняя диагностика, искусственный интеллект, модели машинного обучения

ВВЕДЕНИЕ

С 1990 по 2022 годы количество людей в мире, живущих с диабетом, большая часть которого приходится на инсулиннезависимый диабет 2 типа, выросло с 200 миллионов до 830 миллионов человек. Диабет может привести к серьёзным осложнениям, таким как слепота, почечная недостаточность, инфаркт миокарда, инсульт и ампутация

^{*}likulikova@mail.ru

нижних конечностей. По оценкам ВОЗ, в 2021 году в мире от диабета и связанных с ним заболеваний почек умерло более 2 миллионов человек. Кроме того, около 11 % всех смертей от сердечно-сосудистых заболеваний были связаны с повышенным уровнем глюкозы в крови [1].

Диабет повышает риск тяжелых осложнений, таких как ампутация, потеря зрения и почечная недостаточность, и связан с сердечно-сосудистыми заболеваниями, деменцией, некоторыми видами рака и инфекциями [2, 3].

Диабет (особенно 2 типа) поддается контролю: сбалансированная диета, физические нагрузки, медикаментозная терапия, а также регулярный мониторинг и своевременное лечение осложнений помогают предотвратить или замедлить прогрессирование болезни [4].

Вследствие высокого риска инвалидизирующих и смертельно опасных последствий диабета меры по его профилактике, а также отсрочка развития осложнений за счёт оптимизации питания и терапии представляют собой ключевые направления для улучшения показателей общественного здоровья.

Глобальные расходы на здравоохранение, связанные с диабетом, оценивались в 966 миллиардов долларов США в 2021 году и, по прогнозам, достигнут 1.05 миллиарда долларов США к 2045 году [5]. Нагрузка на систему здравоохранения, обусловленная ростом распространенности диабета, усугубляется увеличением числа осложнений: инсульт, ишемическая болезнь сердца, периферическая невропатия, ретинопатия, заболевания почек и заболевания периферических сосудов, злокачественные новообразования, инфекции стопы и почек, неалкогольная жировая болезнь печени, функциональные нарушения, когнитивные расстройства и аффективные расстройства [6, 7].

Раннее выявление сахарного диабета 2 (СД2) типа имеет принципиальное значение для замедления прогрессирования заболевания и предупреждения развития тяжёлых сосудистых осложнений. На ранних стадиях, когда ещё отсутствует выраженная гипергликемия, возможно достижение стойкой компенсации углеводного обмена за счёт изменения образа жизни: коррекции питания, повышение физической активности и нормализации массы тела. При раннем выявлении и активной коррекции факторов риска у части пациентов возможно достижение стойкой ремиссии без применения сахароснижающих препаратов [8].

Таким образом, несмотря на то что термин «излечение» не корректен применительно к хроническому заболеванию, ряду пациентов удаётся достичь нормогликемии и сохранить её в течение длительного времени при условии строгого соблюдения рекомендаций по питанию и физической нагрузке. Это подчёркивает ключевую роль своевременной диагностики и профилактики СД2 на доклинической стадии [8].

В настоящее время половина (50.1 %) людей с диабетом не знают, что у них диабет [9]. При выборе тактики лечения приоритетным является индивидуальный подход к каждому пациенту. Также на любом этапе лечения необходимо оценивать индивидуальный риск развития или наличие атеросклеротического сердечно-сосудистого заболевания (АССЗ), хронической сердечной недостаточности (ХСН), хронической болезни почек (ХБП), гипогликемий, а также влияние массы тела.

Прогнозирование и профилактика гипогликемии остаются одной из ключевых задач современной диабетологии, особенно в отношении эпизодов, происходящих во время ночного сна. У здоровых людей снижение уровня глюкозы во сне обычно сопровождается пробуждением, что предотвращает развитие тяжёлых последствий. У пациентов с сахарным диабетом 1 типа (СД1), получающих инсулинотерапию, этот защитный механизм значительно ослаблен, что объясняет высокую долю случаев тяжёлой гипогликемии именно в ночное время.

В последние годы многочисленные исследования демонстрируют перспективность использования данных непрерывного мониторинга уровня глюкозы в сочетании с методами искусственного интеллекта для прогнозирования и своевременного предупреждения гипогликемических состояний. Алгоритмы машинного обучения позволяют обрабатывать большие массивы данных, выявлять скрытые закономерности и строить прогностические модели.

Машинное обучение (МО) (от англ. «machine learning», ML), входящее в состав методов искусственного интеллекта (ИИ), основывается на способности вычислительных систем выполнять когнитивные функции, характерные для человека, такие как анализ информации, выявление закономерностей, прогнозирование и принятие решений. Ключевую роль в этом процессе играют статистические методы, теория вероятностей, байесовский подход, марковские цепи и искусственные нейронные сети.

Одной из распространённых проблем при построении прогностических моделей является несбалансированность выборки: гипогликемия развивается относительно редко и нерегулярно. Из-за этого записи с гипогликемическими эпизодами составляют незначительную долю среди всех наблюдений, что может приводить к переобучению модели и снижению точности предсказаний при добавлении или исключении данных. Для решения данной задачи рекомендуется использовать методы искусственного увеличения выборки редкого класса или, напротив, уменьшения доли преобладающего класса.

Таким образом, анализ литературных источников показывает, что интеграция методов искусственного интеллекта и непрерывного мониторинга глюкозы способствует развитию надёжных систем прогнозирования и предотвращения гипогликемических состояний, что играет важную роль в повышении безопасности пациентов и улучшении их качества жизни.

Важным условием для дальнейшего совершенствования этих технологий является создание обширных и структурированных библиотек данных. На сегодняшний день алгоритмы машинного обучения могут использовать разнообразные типы информации: гликемия, данные непрерывного мониторинга, показатели частоты сердечных сокращений, физической активности, данные о приёме пищи и лекарствах, а также генетические и метаболические маркёры. Интеграция этих данных позволяет строить многокомпонентные модели, способные не только прогнозировать риск гипогликемии, но и адаптироваться к индивидуальным особенностям пациента.

Кроме того, перспективным направлением является развитие предиктивных систем с обратной связью и адаптивных алгоритмов, которые в реальном времени анализируют поступающие данные и могут корректировать терапию или предупреждать пациента о повышенном риске. Внедрение таких интеллектуальных систем требует не только большого объема данных, но и эффективных протоколов обмена между устройствами мониторинга и введения инсулина и мобильными приложениями.

Таким образом, развитие инфраструктуры хранения и обмена данными, а также разработка безопасных и конфиденциальных платформ обработки медицинской информации являются ключевыми факторами успешной интеграции технологий искусственного интеллекта в клиническую практику.

Настоящий обзор посвящен анализу литературных данных, связанных с развитием данного направления. Источниками информации являлись рецензируемые научные журналы и открытые информационные системы. Поиск литературы был проведён в базе данных PubMed с использованием ключевых слов «diabetes, type 2, artificial intelligence». На начальном этапе система выдала 1173 публикаций. При этом научные публикации охватили период с 1990-х годов до 2025 года. После применения фильтров «Humans», «Adults 19+ years» общее число статей составило 496. Были применены

дополнительные критерии отбора в соответствии с представленными разделами исследования.

РЕЗУЛЬТАТЫ

1. Тенденции развития методов машинного обучения

Перспективно применение моделей машинного обучения и искусственного интеллекта в создании персонализированных систем прогнозирования диабета [10]. В последнее десятилетие получили развития математические алгоритмы, основанные на методах машинного обучения для прогноза риска развития заболевания, выявления факторов диабета, анализа возможных осложнений, разработки лекарств и методов лечения диабета, изучения влияния генетических и экологических факторов на возникновение и прогрессирование диабета [11].

Начало использования методов (Первый этап, рисунок 1) искусственного интеллекта и машинного обучения для анализа факторов риска развития сахарного диабета 2 типа было положено в 1990-х годах благодаря применению простых статистических моделей, таких как логистическая регрессия и регрессия Кокса. В 1992 году была опубликована статья М.В. Кнуймана и его соавторов, посвященная анализу высокого уровня смертности среди участников с инсулиннезависимым сахарным диабетом, проведенному с использованием модели регрессии пропорциональных рисков Кокса [12]. В исследовании приняли участие более 800 жителей Западной Австралии с инсулиннезависимым сахарным диабетом в период 1978–1982 гг.

В 1993 году Дж. А. Маршалл и соавт. опубликовали исследование, посвященное изучению повышенной в 2-5 раз частоте заболеваемости СД2 у латиноамериканцев по отношению к представителям европейского происхождения, не относящихся к латиноамериканской этнической группе, в Соединенных Штатах с 1984 по 1986 год [13]. В исследовании использовали логистическую модель.

Также в 1993 году вышло исследование Лиу и соавт. [14], в котором исследовали связь уровня содержания гликированного гемоглобина, глюкозы плазмы и уровня диабетической ретинопатии среди индейцев пима ($n = 789$) с СД2 с применением логистических моделей.

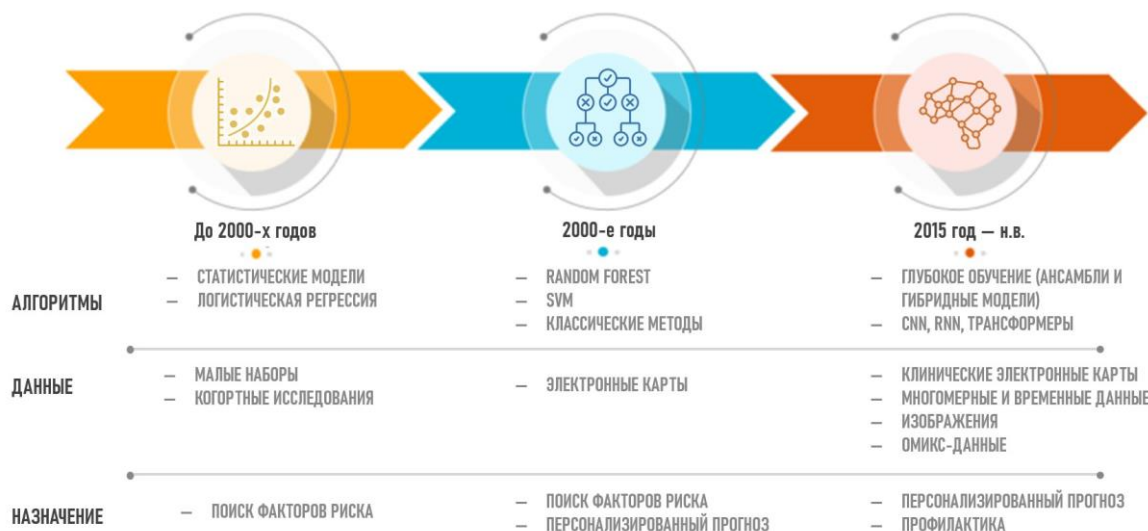


Рис. 1. Алгоритмы от ретроспективного анализа к персонализированной медицине.

На первом этапе использования методов искусственного интеллекта и машинного обучения для анализа факторов риска развития сахарного диабета 2 типа в качестве входных данных применялись клинические параметры: уровень гликированного

гемоглобина, глюкозы натощак, результаты перорального глюкозотолерантного теста, антропометрические данные: возраст, индекс массы тела и сведения из опросников, в том числе семейный анамнез.

Второй этап развития методов машинного обучения (рис. 1) характеризуется внедрением сложных алгоритмов для выявления нелинейных зависимостей. Широкое распространение получили деревья решений и их ансамбли: случайный лес (random forest), градиентный бустинг (gradient boosting, xgboost), подход ускоренного градиентного бустинга LightGBM (light gradient boosting machine). Особенности этих подходов по отношению к простым моделям являются хорошая интерпретируемость, устойчивость к шуму и высокая точность.

В 2017 году Аарон Баум и соавт. использовали деревья решений для задачи прогноза риска развития сердечно-сосудистых осложнений при диабете 2 типа в условиях снижения веса участников [15]. В исследовании участвовали 4901 пациент с СД2, наблюдения проводились в период с 2001 г. по 2004 г.

В 2020 году опубликовано исследование, посвященное прогнозу уровня глюкозы в крови у пациентов с СД2 с использованием алгоритма случайного леса, метода опорных векторов и нейронной сети с обратным распространением ошибки [16]. Наблюдения проводились в период с января 2016 года по декабрь 2017 года, в исследовании приняли участие 2787 пациентов из 27 медицинских центров.

В 2020 году группа итальянских исследователей под руководством Мишель Бернардини [17] разработала и протестировала модель предсказания сахарного диабета 2 типа с использованием метода опорных векторов. В качестве данных использовали записи электронных медицинских карт 2433 участников. Модель SB-SVM продемонстрировала высокое качество диагностики СД2 по критериям AUC и recall.

Третий этап (рис. 1) – это период развития моделей глубокого обучения и анализа больших, многомерных и неструктурированных данных. Глубокие нейронные сети DNN (deep neural network), рекуррентные нейронные сети RNN (recurrent neural network), сети долгой краткосрочной памяти LSTM (long short-term memory) и многослойные перцептроны MLP (multilayered perceptron), сверточные нейронные сети CNN (convolutional neural network) применяются для анализа молекулярно-генетических и клинических данных, включая медицинские изображения (офтальмологические снимки, изображения магнитно-резонансной ангиографии, КТ-ангиографии, кардиологические снимки), данные, полученные с носимых устройств (активность, сон), текстовые данные (врачебные записи).

В 2023 году опубликовано исследование, в котором применен алгоритм долгой краткосрочной памяти LSTM для мониторинга и прогноза уровня глюкозы в капиллярной крови у женщин с гестационным сахарным диабетом [18]. В исследовании участвовали 1110 женщин, ежедневно измерявших уровень глюкозы в крови. Учет клинических данных в виде временных рядов проводился посредством мобильного приложения GDM-Health. Авторы отмечают дешевизну и доступность такого мониторинга и прогноза уровня глюкозы в крови.

В 2024 году Сяобо Ци и соавт. предложили модель прогноза риска развития СД2, основанную на методе глубокого обучения, коэффициента корреляции Кендалла и механизма внимания [19]. В исследовании использованы наборы клинических данных из Лаборатории городской медицинской больницы (Ирак) для 1000 пациентов и набор данных из Национального института диабета, заболеваний пищеварительной системы и почек, который включал 768 пациентов племени пима (PIMA Indian dataset) [19].

В таблице 1 представлены некоторые исследования, посвященные созданию алгоритмов ИИ для прогноза развития СД2. Для сравнения эффективности методов используются значения точности (accuracy), прецизионности (precision), специфичности (specificity), чувствительности (sensitivity, recall) и F1-оценка (F1-score), мы приводим те, которые указаны в источниках.

Таблица 1. Обзор алгоритмов прогноза основных факторов развития СД2

Год	Алгоритм	Параметры проверки	Тип данных	Выборка, чел.	Ссылка
2015	Логистическая регрессия, линейный дискриминантный анализ, наивный байесовский анализ и метод опорных векторов	точность: 0.97–0.99	генетические данные (17 473 гена) микрофлоры кишечника	489	[20]
2016	Логистическая регрессия, машина опорных векторов и искусственная нейронная сеть	точность: 0.85 чувствительность: 0.85 F1-оценка: 0.85	электрохимические измерения показателей в слюне	175	[21]
2017	Сигмоидальная функция — глубокая нейронная сеть с кросс-энтропией в качестве функции потерь	точность: 0.921 прецизионность: 0.915 чувствительность: 0.867 F1-оценка: 0.823	клинические данные всего 62 параметра: физическое обследование, анализ крови, анализ функции	110300 (8 групп)	[22]
2018	Нейронная сеть и логистическая регрессия	точность: 0.78 F1-оценка: 0.84	клинические данные: число беременностей, глюкозотолерантный тест, АД, толщина кожной складки, 2 уровень инсулина, ИМТ, возраст, предрасположенность к СД	768 беременных женщин	[23]
2019	Глубокая нейронная сеть с тремя встроенными и двумя скрытыми слоями	точность: 0.84 чувствительность: 0.96	Клинические данные всего 1312 параметра: базовые признаки (возраст, пол, ИМТ, АД), диагностические признаки на основе кодов МКБ-9, медикаментов и лабораторных тестов	9948, из которых 1904 с СД2	[24]
2020	Свёрточная нейронная сеть и двунаправленная долговременная кратковременная память	точность: 0.928 прецизионность: 0.923 чувствительность: 0.905 F1-оценка: 0.913	Клинические данные всего 17 блоков: диагноз, основная жалоба, важные показатели, общее состояние, аллергический анамнез, статус питания и др.	18625, из них 5645 с СД2	[25]
2023	Рекуррентные модели нейронных сетей с долговременной краткосрочной памятью	точность: 0.96	Прогнозный мониторинг уровня глюкозы в крови по клиническим данным: уровень глюкозы в крови четыре раза в день	1282 с ГСД	[18]

Год	Алгоритм	Параметры проверки	Тип данных	Выборка, чел.	Ссылка
2024	Глубокая нейронная сеть с применением коэффициента корреляции Кендалла и механизма внимания (KCCAM DNN)	точность: 0.991 чувствительность: 0.973 F1-оценка: 0.974	Клинические данные всего 8 параметров: беременности, уровень глюкозы, АД, толщина кожной складки, инсулин, ИМТ, наследственность, возраст	1768	[19]

ИМТ – индекс массы тела; АД – артериальное давление; ГСД – гестационный сахарный диабет

2. Источники данных для машинного обучения

В области машинного обучения для анализа сахарного диабета используются разнообразные источники данных, включая клинические, молекулярно-генетические, изображения и поведенческие данные. В последние годы клинические данные о пациенте ведутся в электронной медицинской карте (ЭМК), которая стала стандартом во многих странах, постепенно заменяя бумажные носители. ЭМК является цифровым аналогом бумажных медицинских карт пациента, где вся информация о здоровье хранится в электронном виде (записи о приемах у специалистов, результаты анализов и диагностических исследований, сведения о больничных листах, справках, вакцинации, рецепты на получение лекарственных препаратов и др.). И предлагают ряд преимуществ, таких как удобство доступа и интеграции данных, снижение вероятности потери или повреждения информации, а также упрощение обмена и анализа данными между медицинскими учреждениями [26].

Приказ Минздрава РФ от 7 сентября 2020 года № 947н «Об утверждении Порядка организации системы документооборота в сфере охраны здоровья в части ведения медицинской документации в форме электронных документов» регламентирует ведение медицинской документации в электронном виде. Внедрение ЭМК является частью общей стратегии цифровизации здравоохранения в России, и переход на электронные медкарты стал обязательным для всех медицинских организаций с 2024 года (Распоряжение Правительства Российской Федерации от 17.04.2024 № 959-р).

Клинические и лабораторные данные

Электронные медицинские карты включают в себя историю болезни участников исследования, результаты проведения биохимических анализов, а также схемы назначения лекарств и динамику состояния. Наиболее часто в литературе упоминаются следующие лабораторные и клинические показатели в контексте создания инструментов машинного обучения: уровень глюкозы в крови и тест на толерантность к глюкозе (более 30 исследований), уровень глюкозы в моче (2 исследования), триглицериды (более 20 исследований), липопротеины низкой, очень низкой плотности (более 20 исследований), ферменты АЛТ и АСТ (около 10 исследований), уровень мочевины в крови (6 исследований), гликированный гемоглобин (около 20 исследований); инсулин (около 30 исследований), тиреотропный гормон; катионы в крови (4 исследования), содержание С-реактивного белка, холестерина (около 30 исследований) [10].

Для создания новых инструментов МО доступны обширные банки аннотированных данных для различных нозологий. Биобанки предоставляют обширные массивы

генетической и клинической информации, которые используются для изучения механизмов прогрессирования заболевания, факторов риска и разработки потенциальных стратегий терапии. Биобанки UK Biobank [27], FinnGen [28], Biobank Japan [29] и All of Us Research Program (США) [30] позволяют исследователям выявлять подгруппы в популяциях с преддиабетом и разрабатывать персонализированные подходы для профилактики.

Использование биобанков способствует трансформации подходов к лечению. Анализ генетических данных пациентов позволяет выявить генетические факторы, влияющие на течение заболевания и эффективность терапии. На основе этих данных формируются персонализированные рекомендации по образу жизни и медикаментозному лечению: с учётом генетического профиля пациенту могут быть предложены конкретные диеты, режимы физической активности и лекарственные препараты, направленные на оптимальный контроль уровня глюкозы в крови и снижение риска осложнений [31,32].

В России развитие биобанков осуществляется, в частности, в рамках проекта Научного центра мирового уровня «Цифровой биодизайн и персонализированное здравоохранение». На базе Института биомедицинской химии (ИБМХ) создан и развивается виртуальный биобанк «БиоВитрина» [33] — инновационная цифровая платформа для каталогизации и управления биообразцами. «БиоВитрина» представляет собой веб-репозиторий, позволяющий исследователям подбирать образцы для планирования экспериментов, а также хранить и систематизировать результаты аналитических исследований. Платформа способствует расширению научного сотрудничества между учреждениями и развитию персонализированной медицины. В настоящее время в системе представлены образцы по ряду ключевых направлений, охватывающих социально значимые области медицины. Работа по дальнейшему расширению и интеграции биобанка активно продолжается. [31,34].

Также доступны для исследователей специализированные датасеты. Например, «Pima Indians Diabetes Dataset», включающий сведения о 768 женщинах с данными о беременности, ИМТ, уровне глюкозы и др.[35]. В России разработана веб-доступная пополняемая база данных клинико-эпидемиологического мониторинга сахарного диабета (РУ №2020622447). Проект реализуется ГНЦ РФ ФГБУ «НМИЦ эндокринологии» Минздрава России совместно с АО «Астон Консалтинг».

Медицинская визуализация

Офтальмологические снимки, в том числе фотоизображения глазного дна и оптическая когерентная томография, применяются для выявления заболеваний сетчатки и скрининга диабетической ретинопатии с использованием искусственного интеллекта. Поскольку люди с диабетом регулярно проходят процедуру визуализации сетчатки для оценки своего статуса диабетической ретинопатии, применение методов искусственного интеллекта может в перспективе использоваться в качестве доступного комплексного скрининга множественных системных микро- и макрососудистых осложнений диабета [36]. На сегодня доступны исследования с указанием обширных наборов данных: 131 238 изображений от 66 500 участников исследования UK Biobank (Великобритания) [37]; 79 511 пациентов из SiDRP (Сингапур) [38]; SDRN-NDS 24 012 и 202 843 человека с СД1 и СД2 (Шотландия); 798 866 снимков глазного дна от BRAVE (Китай) [39] и другие исследования.

Молекулярно-генетические данные

СД2 относится к группе метаболических заболеваний, в связи с этим важным представляется сбор омикс-данных, то есть данных об особенностях молекулярного состава (генетические полиморфизмы, белки и протеоформы, транскрипты, метаболиты) биологических образцов. Применение методов машинного обучения к

таким данным позволяет уточнить патогенез СД2 на молекулярном уровне, прогнозировать риск развития заболевания, выявлять новые кандидатные биомаркеры [40]. Помимо этого, методы машинного обучения позволяют интегрировать из нескольких источников биологическую и молекулярную информацию.

На сегодня мировое медицинское и биологическое сообщество накопило большое количество молекулярной информации для различных состояний здоровья человека, в том числе для СД2. Эта информация организована в репозитории и банки данных, которые, как правило, доступны для исследователей.

Генетические банки данных могут быть общими для разных состояний организма в норме и патологии и специализированными для СД2. Наиболее распространенными являются несколько. «GWAS Catalog» предоставляет доступную для поиска базу данных ассоциаций SNP-признаков (www.ebi.ac.uk/gwas/studies) [41]. Данный ресурс содержит сведения о 7653 SNP-признаках, которые были выявлены в 254 исследованиях, и, вероятно, ассоциированы с СД2. Среди специализированных репозиториев стоит отметить «Type 2 Diabetes KP genetic association datasets» (t2d.hugeamp.org), в котором представлены 6557 наборов генетических данных [42].

Банки данных протеомных исследований включают аннотированные наборы масс-спектрометрических данных, полученных для биологических образцов. Так, в одном из крупнейших банков ProteomeXchange (www.proteomexchange.org) [43] содержатся 72 набора протеомных данных для СД2. PRIDE (www.ebi.ac.uk/pride) является основным членом консорциума ProteomeXchange содержит сведения 246 наборах протеомных данных для СД2 [44]. База знаний MassIVE (massive.ucsd.edu) также является одним из крупнейших сервисов обмена протеомными данными и содержит 99 наборов данных, связанных с сахарным диабетом.

Развиты также репозитории данных метаболомных экспериментов. Так, в репозитории MetaboLights (www.ebi.ac.uk/metabolights/index) содержится 66 наборов данных для сахарного диабета [45]. Банк данных HMDB (The human metabolome databank, hmdb.ca) [46] содержит данные о 762 метаболитах, ассоциированных с развитием сахарного диабета

Данные носимых устройств

Телереабилитация и подходы реабилитации на основе носимых устройств признаются за их положительное влияние на управление сахарным диабетом 2 типа. Эти подходы позволяют людям контролировать физическую активность и лечебные мероприятия. Телереабилитация подразумевает удаленное наблюдение с помощью видеозвонков, мобильных приложений или онлайн-платформ [47,48].

Сегодня в этом отношении доступно несколько проектов, направленных на повышение качества жизни и мотивацию пациентов к заботе о своем здоровье. Разработанные в рамках этих проектов приложения или устройства являются коммерчески доступными или представленными в научной литературе.

Few Touch Application (FTA) является исследовательской платформой, разработанной как инструмент для помощи в самоконтроле и управлении состоянием организма при СД (персональное медицинское приложение) [49]. В рамках этого проекта было проведено исследование, отслеживающее требования, потребности, поведение и опыт пользователей (пациентов), которые использовали мобильное приложение FTA. Приложение может автоматически передавать данные об уровне глюкозы, обмениваться важной информацией через SMS, вести цифровой дневник СД 2 типа, интегрировать дневник с медицинскими учреждениями, вести дневник фотографий еды, отслеживать физическую активность и отображать информацию о питании.

MyNetDiary – мобильное и веб-приложение. Основной компонент этой системы сосредоточен на ведении цифрового дневника диеты, персонализированном

управлении состоянием организма и поддержании здорового образа жизни. Он позволяет отмечать и отслеживать уровень глюкозы в крови, показатели гликированного гемоглобина, артериальное давление, диету и прием лекарств. Это коммерческий продукт, к преимуществам которого относится возможность подключения к различным глюкометрам, имеется большая база данных продуктов с указанием их калорийности (www.mynetdiary.com).

Приложение Glooko используется для дистанционного взаимодействия медицинских работников с пациентами с целью обеспечения удаленного мониторинга, тем самым снижая расходы на лечение и неотложную медицинскую помощь. Приложение совместимо с широким спектром глюкометров, а также позволяет отслеживать рацион питания и физическую активность. Медицинский персонал имеет возможность проводить мониторинг параметров состояния пациентов (уровень глюкозы, потребление пищи, лекарств и т.д.), а также устанавливать метки для уведомления о конкретных пациентах с факторами риска (glooko.com).

В рамках реализации проекта diabICT (diabict.fer.hr) была разработана платформа, состоящая из трех частей: сенсорного браслета, программной платформы и устройства, позволяющего передавать данные с глюкометра на смартфон. Система позволяет отслеживать уровень глюкозы, активность и диету, а также предлагать упражнения для повышения ежедневной активности. Еще раз подчеркивается мотивация самоконтроля у пользователей и принятия активной роли в управлении диабетом.

Как можно заметить, доступно множество вариантов для управления, отслеживания, помощи в диагностике и лечения диабета с новыми решениями, которые появляются все чаще.

Однако наиболее часто для мониторинга состояния пациенты используют глюкометры, поскольку они являются наиболее привычными в повседневной жизни. На рынке представлен широкий спектр глюкометров таких как Dexcom, FreeStyle Libre и др.

Непрерывная система мониторинга уровня глюкозы Dexcom G6 не требует уколов пальцев или анализа крови. Представляет собой небольшой прибор в виде пластыря, содержащий сенсор, который непрерывно проводит измерение уровня сахара в интерстициальной жидкости и каждые пять минут передает данные об уровне сахара в приложение на смартфоне пользователя.

Система MiniMed 670G, первая гибридная замкнутая система, постоянно отслеживает уровень сахара в крови и автоматически выдает инсулин, но пациентам все еще приходится самостоятельно вводить дозу болюсного инсулина (инсулина, обычно принимаемого во время еды) в зависимости от того, что вы едите, а также «сообщать» ему, когда начинают заниматься физическими упражнениями, которые снижают уровень сахара.

Abbott выпустила беспроводной монитор FreeStyle Libre, новый класс устройств для мониторинга уровня сахара. Пользователь закрепляет сенсор на верхней части руки, который измеряет уровень сахара в интерстициальной жидкости.

NovoPen 6 – интеллектуальные инсулиновые ручки с возможностью интеграции со смартфоном, автоматически записывают информацию о дозировке инсулина при каждой инъекции, что обеспечивает полноценный гликемический контроль.

Компания Insulet предлагает систему менеджмента инсулина OmniPod (OmniPod Insulin Management System), которая состоит из одноразового инфузионного аппарата (инсулиновой помпы) и портативного беспроводного персонального контрольного устройства, которое подключается к помпе посредством Bluetooth, и представляет собой альтернативу традиционным методам доставки инсулина. Это гибридная система с замкнутой обратной связью со встроенным алгоритмом модели предиктивного анализа для расчета дозировки для каждого конкретного пациента. Она может использоваться вместе с базальным (предназначен для обеспечения нормального

функционирования организма) и болюсным инсулином (для компенсации употребляемой еды). Благодаря своей простой и удобной конструкции одноразовый блок обеспечивает до трех дней непрерывной доставки инсулина.

Другие аналогичные устройства представлены системой DBLG1 (www.diabeloop.com), BEAT Technology Platform (nemaumedical.com/technology) и др.

Анкетирование

Анкетирование играет важную роль в комплексной оценке состояния пациентов с сахарным диабетом 2 типа. Опросники позволяют собрать объективные данные о принимаемой терапии, психологическом состоянии [50] и поведенческих аспектах, которые невозможно получить только с помощью лабораторных или инструментальных методов исследования. Эти данные важны для персонализированной терапии, оценки рисков осложнений и повышения эффективности назначенного лечения.

Одним из наиболее часто используемых опросников является Diabetes Self-Management Questionnaire (DSMQ) [51], направленный на оценку эффективности самоконтроля у пациентов. Он охватывает такие аспекты, как соблюдение диеты, регулярность физической активности, контроль уровня сахара и взаимодействие с медицинскими специалистами. Высокие баллы по DSMQ коррелируют с лучшими показателями гликемического контроля, что делает его полезным инструментом в рутинной практике.

Для оценки влияния СД2 на качество жизни широко применяется Diabetes Quality of Life Measure (DQOL) [52]. Этот опросник учитывает физическое и эмоциональное состояние, восприятие ограничений, связанных с болезнью, и уровень удовлетворённости лечением. Анализ ответов помогает выявить скрытые проблемы, требующие немедленного вмешательства – например, психологическую помощь, связанную с основным заболеванием.

Также в клинической практике применяются инструменты для оценки психологического дистресса, например, Problem Areas in Diabetes Scale (PAID) [53] и Diabetes Distress Scale (DDS) [54]. Они позволяют выявить эмоциональное выгорание, тревогу и депрессию у пациентов, что особенно важно, учитывая их влияние на приверженность терапии и общие исходы лечения.

В России в клинических рекомендациях по лечению сахарного диабета 2 типа [50] помимо перечисленных шкал используется Финская шкала риска диабета (FINDRISC) – это простой инструмент из восьми вопросов, используемый для оценки риска развития диабета 2 типа у человека. Он не требует никаких анализов крови и опирается на легкодоступную информацию, такую как возраст, индекс массы тела, окружность талии, физическая активность, диета и семейный анамнез диабета. FINDRISC был проверен в различных группах населения и также используется для выявления лиц с риском метаболического синдрома, непереносимости глюкозы [55].

Оценка симптомов диабетической нейропатии (DNS) – это основанная на симптомах система оценки, используемая для выявления и диагностики дистальной полинейропатии у лиц с диабетом. Она фокусируется на наличии специфических невропатических симптомов, таких как боль, онемение, покалывание и атаксия (неустойчивость при ходьбе). Данная шкала является ценным инструментом для врачей, позволяющим быстро выявлять диабетическую нейропатию на основе симптомов, сообщаемых пациентами, обеспечивая баланс между простотой, скоростью и клинической достоверностью [56,57].

Использование валидированных опросников в сочетании с цифровыми технологиями – мобильными приложениями, системами дистанционного мониторинга и ИИ-алгоритмами – расширяет возможности анкетирования. Оно становится частью персонализированного подхода в управлении СД2, позволяя врачу учитывать как

объективные клинические данные, так и субъективный опыт пациента, а также более точно подбирать методы лечения и поддержки.

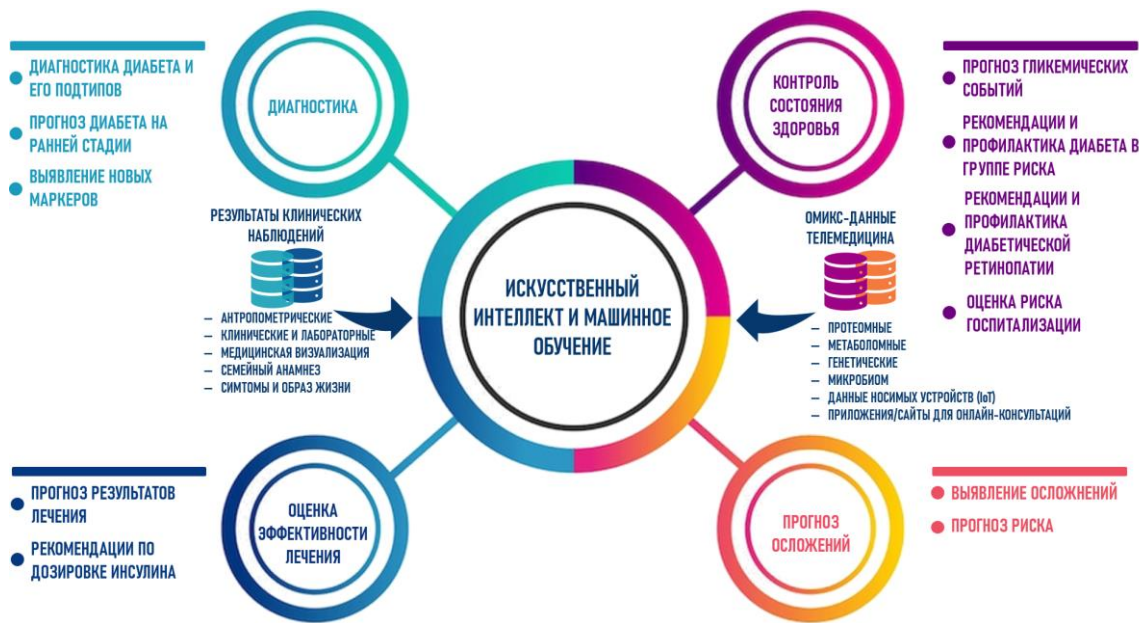


Рис. 2. Практическое применение методов машинного обучения в прогнозе сахарного диабета второго типа [58].

Таким образом, возможный путь решения проблем в лечении диабета и метаболических заболеваний – это анализ обширных наборов данных и прогностическое моделирование. На сегодня разработаны алгоритмы искусственного интеллекта, которые могут автоматизировать скрининг клинических и молекулярно-генетических данных для прогноза, лечения и мониторинга состояния пациента (рис. 2).

Мобильные приложения и носимые устройства также могут способствовать повышению эффективности мониторинга уровня глюкозы. Применение носимых устройств и технологий интернета вещей (IoT) позволяет перейти от эпизодического контроля к непрерывному мониторингу клинических показателей в режиме реального времени. Носимые устройства для контроля гликемии с использованием технологий искусственного интеллекта представляют собой перспективное направление в компенсации сахарного диабета. Такие устройства обеспечивают автоматизированную подачу инсулина за счёт интеграции инсулиновых помп с системами непрерывного мониторинга уровня глюкозы (CGM, Continuous Glucose Monitoring). Алгоритмы ИИ в составе этих устройств анализируют данные в реальном времени и адаптируют дозировку инсулина в зависимости от текущих и прогнозируемых показателей гликемии. Эти системы также известны как «искусственная поджелудочная железа» и направлены на снижение колебаний глюкозы, минимизацию гипо- и гипергликемий и улучшение общего контроля заболевания без необходимости постоянного вмешательства со стороны пациента.

3. Будущие направления развития области

В перспективе дальнейшего развития методов машинного обучения в прогнозе, диагностике и лечении сахарного диабета 2 типа можно выделить ключевые направления. Первое направление связано с повышением точности ранней диагностики и стратификации риска. Машинное обучение уже доказало свою эффективность в ранней идентификации лиц с риском развития сахарного диабета 2 типа и его осложнений, таких как ретинопатия, нефропатия и диабетическая стопа. Например,

модели, использующие «деревья решений» и «случайный лес», могут с высокой точностью классифицировать пациентов по группам риска на основе клинико-лабораторных и демографических данных. Это обеспечивает более целенаправленный подход к профилактике и мониторингу заболевания [59,60].

Второе важное направление – персонализированная терапия, включая индивидуальный подбор лекарственных средств и дозировок. Методы машинного обучения позволяют учитывать генетические особенности, стиль жизни, данные сенсоров с приборов и истории болезни для точной настройки лечения. Это особенно важно при подборе режима инсулинотерапии, где внедрение систем «model-informed precision dosing» (MIPD) демонстрирует клиническую эффективность [61]. Современные интеллектуальные системы способны рекомендовать оптимальные комбинации препаратов, предсказывать ответ на терапию и минимизировать вероятность развития осложнений, в том числе сводя к минимуму риск возникновения гипо- или гипергликемии. Это особенно важно, когда лечение должно применяться у особых групп пациентов, таких как дети, пожилые люди, полиморбидные пациенты.

Третье направление связано с развитием систем, обеспечивающих автоматизированную подачу инсулина. Такие системы интегрированы с устройствами непрерывного мониторинга глюкозы и инсулиновыми помпами. Они обеспечивают автоматическую адаптацию доз инсулина в реальном времени, улучшая контроль гликемии и снижая нагрузку на пациента. Это направление активно развивается в клинических исследованиях и уже внедряется в повседневную практику для больных с СД1 с перспективой адаптации для СД2 [62].

Четвёртая перспектива развития методов машинного обучения в лечении сахарного диабета представляет собой мониторинг осложнений с помощью компьютерных алгоритмов и анализа биомаркеров. Системы на базе искусственного интеллекта уже используются для анализа изображений глазного дна при диагностике диабетической ретинопатии, а также для оценки функции почек по данным биохимического анализа крови и мочи. Это позволяет осуществлять раннюю диагностику и предупреждение прогрессирования хронических осложнений диабета [63].

В дополнение к этим направлениям, следует отметить, что в ближайшей перспективе инструменты позволят создать универсальные системы поддержки принятия клинических решений, интегрированные в ЭМК. Такие системы смогут автоматизировать диагностику метаболических нарушений, в том числе преддиабета и сахарного диабета 2 типа, оптимизировать дозирование лекарственных препаратов и обеспечивать непрерывный мониторинг клинического состояния пациента на всех этапах лечения. Однако для внедрения в клиническую практику необходимы стандарты качества, интерпретируемость моделей, защита данных и этическое регулирование.

ЗАКЛЮЧЕНИЕ

Сахарный диабет 2 типа остается важной проблемой здравоохранения, что обусловлено ростом заболеваемости, высоким уровнем осложнений и инвалидизации, существенными экономическими затратами на лечение. В этой связи внедрение методов машинного обучения открывает перспективы для ранней диагностики, прогноза течения заболевания и персонализированного подхода управления терапией.

Проведенный анализ показал, что современные алгоритмы машинного обучения, включая ансамблевые методы и машинные методы глубокого обучения развиваются в направлении прогноза риска развития СД2, используя клинико-лабораторные данные и анкетирование, повышения эффективности контроля гликемии за счет анализа данных непрерывного мониторинга и прогнозирования гипогликемий, выявления осложнений (ретинопатия, нефропатия) посредством обработки медицинских изображений и электронных медицинских карт.

Преимуществами технологий искусственного интеллекта и машинного обучения является возможность анализа сложных многомерных данных, включая генетические, протеомные, метаболомные и клинические. Развитие биобанков и аннотированных данных, стандартизированных электронных медицинских карт, внедрение методов объяснимого искусственного интеллекта обеспечат дальнейшее повышение точности алгоритмов в прогнозе и лечения СД2.

Для Б.Т.В., М.К.А., П.Д.В., Р.Е.А. и К.А.Л. работа выполнена в рамках Программы фундаментальных научных исследований в Российской Федерации на долгосрочный период (2021–2030 годы); К.Л.И. работала по госзаданию ИПМ им. М.В. Келдыша РАН (FFMN-2025-0024), Р.В.Р., К.А.В. – по госзаданию ИТЭБ РАН (FFRS-2024-0015).

СПИСОК ЛИТЕРАТУРЫ

1. Zhou B., Rayner A.W., Gregg E.W., Sheffer K.E., Carrillo-Larco R.M., Bennett J.E., Shaw J.E., Paciorek C.J., Singleton R.K., Barradas Pires A., et al. Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants. *The Lancet*. 2024. V. 404, No. 10467. P. 2077–2093.
2. American Diabetes Association Professional Practice Committee. 9. Pharmacologic Approaches to Glycemic Treatment: Standards of Care in Diabetes-2024. *Diabetes Care*. 2024. V. 47. No. Suppl.1. P. S158–S178. doi: [10.2337/dc24-S009](https://doi.org/10.2337/dc24-S009)
3. Adler A.I., Coleman R.L., Leal J., Whiteley W.N., Clarke P., Holman R.R. Post-trial monitoring of a randomised controlled trial of intensive glycaemic control in type 2 diabetes extended from 10 years to 24 years (UKPDS 91). *Lancet Lond. Engl.* 2024. V. 404. No. 10448. P. 145–155. doi: [10.1016/S0140-6736\(24\)00537-3](https://doi.org/10.1016/S0140-6736(24)00537-3)
4. Davies M.J., Aroda V.R., Collins B.S., Gabbay R.A., Green J., Maruthur N.M., Rosas S.E., Del Prato S., Mathieu C., Mingrone G. et al. Management of Hyperglycemia in Type 2 Diabetes, 2022. A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes Care*. 2022. V. 45. No. 11. P. 2753–2786. doi: [10.2337/dci22-0034](https://doi.org/10.2337/dci22-0034)
5. Sun H., Saeedi P., Karuranga S., Pinkepank M., Ogurtsova K., Duncan B.B., Stein C., Basit A., Chan J.C.N., Mbanya J.C. et al. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res. Clin. Pract.* 2022. V. 183. Article No. 109119. doi: [10.1016/j.diabres.2021.109119](https://doi.org/10.1016/j.diabres.2021.109119)
6. Rosenfeld R.M., Grega M.L., Karlsen M.C., Abu Dabrh A.M., Aurora R.N., Bonnet J.P., Donnell L., Fitzpatrick S.L., Frates B., Joy E.A. et al. Lifestyle Interventions for Treatment and Remission of Type 2 Diabetes and Prediabetes in Adults: A Clinical Practice Guideline from the American College of Lifestyle Medicine. *Am. J. Lifestyle Med.* 2025. V. 19. P. 10S–131S. doi: [10.1177/15598276251325488](https://doi.org/10.1177/15598276251325488)
7. Tomic D., Shaw J.E., Magliano D.J. The burden and risks of emerging complications of diabetes mellitus. *Nat. Rev. Endocrinol.* 2022. V. 18. No. 9. P. 525–539. doi: [10.1038/s41574-022-00690-7](https://doi.org/10.1038/s41574-022-00690-7)
8. Дедов И.И., Шестакова М.В., Майоров А.Ю., Мокрышева Н.Г., Андреева Е.Н., Безлепкина О.Б., Петеркова В.А., Артемова Е.В., Бардюгов П.С., Бешлиева Д.Д. и др. Алгоритмы специализированной медицинской помощи больным сахарным диабетом. *Сахарный диабет*. 2023. Т. 26. № 2S. doi: [10.14341/DM13042](https://doi.org/10.14341/DM13042)
9. Saeedi P., Petersohn I., Salpea P., Malanda B., Karuranga S., Unwin N., Colagiuri S., Guariguata L., Motala A.A., Ogurtsova K. et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the

- International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res. Clin. Pract.* 2019. V. 157. doi: [10.1016/j.diabres.2019.107843](https://doi.org/10.1016/j.diabres.2019.107843)
10. Kiran M., Xie Y., Anjum N., Ball G., Pierscionek B., Russell D. Machine learning and artificial intelligence in type 2 diabetes prediction: a comprehensive 33-year bibliometric and literature analysis. *Front. Digit. Health.* 2025. V. 7. Article No. 1557467. doi: [10.3389/fdgth.2025.1557467](https://doi.org/10.3389/fdgth.2025.1557467)
11. Kavakiotis I., Tsave O., Salifoglou A., Maglaveras N., Vlahavas I., Chouvarda I. Machine Learning and Data Mining Methods in Diabetes Research. *Comput. Struct. Biotechnol. J.* 2017. V. 15. P. 104–116. doi: [10.1016/j.csbj.2016.12.005](https://doi.org/10.1016/j.csbj.2016.12.005)
12. Knuiman M.W., Welborn T.A., Whittall D.E. An analysis of excess mortality rates for persons with non-insulin-dependent diabetes mellitus in Western Australia using the Cox proportional hazards regression model. *Am. J. Epidemiol.* 1992. V. 135. No. 6. P. 638–648. doi: [10.1093/oxfordjournals.aje.a116343](https://doi.org/10.1093/oxfordjournals.aje.a116343)
13. Marshall J.A. Hamman R.F., Baxter J., Mayer E.J., Fulton D.L., Orleans M., Rewers M., Jones R.H. Ethnic differences in risk factors associated with the prevalence of non-insulin-dependent diabetes mellitus. The San Luis Valley Diabetes Study. *Am. J. Epidemiol.* 1993. V. 137. No. 7. P. 706–718. doi: [10.1093/oxfordjournals.aje.a116731](https://doi.org/10.1093/oxfordjournals.aje.a116731)
14. Liu Q.Z., Pettitt D.J., Hanson R.L., Charles M.A., Klein R., Bennett P.H., Knowler W.C. Glycated haemoglobin, plasma glucose and diabetic retinopathy: cross-sectional and prospective analyses. *Diabetologia.* 1993. V. 36. No. 5. P. 428–432. doi: [10.1007/BF00402279](https://doi.org/10.1007/BF00402279)
15. Baum A., Scarpa J., Bruzelius E., Tamler R., Basu S., Faghmous J. Targeting Weight Loss Interventions to Reduce Cardiovascular Complications of Type 2 Diabetes: A Machine Learning-Based Post-Hoc Analysis of Heterogeneous Treatment Effects in the Look AHEAD Trial. *Lancet Diabetes Endocrinol.* 2017. V. 5. P. 808–815. doi: [10.1016/S2213-8587\(17\)30176-6](https://doi.org/10.1016/S2213-8587(17)30176-6)
16. Wang J., Wang M.-Y., Wang H., Liu H.-W., Lu R., Duan T.-Q., Li C.-P., Cui Z., Liu Y.-Y., Lyu Y.-J. et al. Status of Glycosylated Hemoglobin and Prediction of Glycemic Control among Patients with Insulin-Treated Type 2 Diabetes in North China: A Multicenter Observational Study. *Chin. Med. J. (Engl.)*. 2020. V. 133. P. 17–24, doi: [10.1097/CM9.0000000000000585](https://doi.org/10.1097/CM9.0000000000000585)
17. Bernardini M., Romeo L., Misericordia P., Frontoni E. Discovering the Type 2 Diabetes in Electronic Health Records Using the Sparse Balanced Support Vector Machine. *IEEE J. Biomed Health Inform.* 2020. V. 24. P. 235–246. doi: [10.1109/JBHI.2019.2899218](https://doi.org/10.1109/JBHI.2019.2899218)
18. Lu H.Y., Lu P., Hirst J.E., Mackillop L., Clifton D.A. A Stacked Long Short-Term Memory Approach for Predictive Blood Glucose Monitoring in Women with Gestational Diabetes Mellitus. *Sensors (Basel)*. 2023. V. 23. Article No. 7990. doi: [10.3390/s23187990](https://doi.org/10.3390/s23187990)
19. Qi X., Lu Y., Shi Y., Qi H., Ren L. A deep neural network prediction method for diabetes based on Kendall's correlation coefficient and attention mechanism. *PLoS One*. 2024. V. 19. Article No. e0306090. doi: [10.1371/journal.pone.0306090](https://doi.org/10.1371/journal.pone.0306090)
20. Cai L., Wu H., Li D., Zhou K., Zou F. Type 2 Diabetes Biomarkers of Human Gut Microbiota Selected via Iterative Sure Independent Screening Method. *PLoS One*. 2015. V. 10. Article No. e0140827. doi: [10.1371/journal.pone.0140827](https://doi.org/10.1371/journal.pone.0140827)
21. Malik S., Khadgawat R., Anand S., Gupta S. Non-invasive detection of fasting blood glucose level via electrochemical measurement of saliva. *SpringerPlus*. 2016. V. 5. Article No. 701. doi: [10.1186/s40064-016-2339-6](https://doi.org/10.1186/s40064-016-2339-6)
22. Maxwell A., Li R., Yang B., Weng H., Ou A., Hong H., Zhou Z., Gong P., Zhang C. Deep learning architectures for multi-label classification of intelligent health risk

- prediction. *BMC Bioinformatics*. 2017. V. 18. Article No. 523. doi: [10.1186/s12859-017-1898-z](https://doi.org/10.1186/s12859-017-1898-z)
23. Dwivedi A.K. Analysis of computational intelligence techniques for diabetes mellitus prediction. *Neural Comput. & Applic.* 2018. V. 30. P. 3837–3845. doi: [10.1007/s00521-017-2969-9](https://doi.org/10.1007/s00521-017-2969-9)
24. Nguyen B.P., Pham H.N., Tran H., Nghiem N., Nguyen Q.H., Do T.T.T., Tran C.T., Simpson C.R. Predicting the onset of type 2 diabetes using wide and deep learning with electronic health records. *Comput. Methods Programs Biomed.* 2019. V. 182. Article No. 105055. doi: [10.1016/j.cmpb.2019.105055](https://doi.org/10.1016/j.cmpb.2019.105055)
25. Wang T., Xuan P., Liu Z., Zhang T. Assistant diagnosis with Chinese electronic medical records based on CNN and BiLSTM with phrase-level and word-level attentions. *BMC Bioinformatics*. 2020. V. 21. Article No. 230. doi: [10.1186/s12859-020-03554-x](https://doi.org/10.1186/s12859-020-03554-x)
26. Гусев А.В., Зингерман Б.В., Тюфлин Д.С., Зинченко В.В. Электронные медицинские карты как источник реальных клинических данных. *Реальные данные и доказательства*. 2022. Т. 2. № 2. С. 8–20. doi: [10.37489/2782-3784-myrd-13](https://doi.org/10.37489/2782-3784-myrd-13)
27. UK Biobank. URL: <https://www.ukbiobank.ac.uk> (accessed: 30.06.2025).
28. FinnGen: an expedition into genomics and medicine. URL: <https://www.finnngen.fi/en> (accessed: 30.06.2025).
29. BioBank Japan. URL: <https://biobankjp.org/en/#gsc.tab=0> (accessed: 30.06.2025).
30. All of Us Research Hub. URL: <https://allofus.nih.gov/> (accessed: 30.06.2025).
31. Bukreeva A.S., Malsagova K.A., Petrovskiy D.V., Butkova T.V., Nakhod V.I., Rudnev V.R., Izotov A.A., Kaysheva A.L. Biobank Digitalization: From Data Acquisition to Efficient Use. *Biology (Basel)*. 2024. V. 13. Article No. 957. doi: [10.3390/biology13120957](https://doi.org/10.3390/biology13120957)
32. Xu W., Liang X., Chen L., Hong W., Hu X. Biobanks in chronic disease management: A comprehensive review of strategies, challenges, and future directions. *Heliyon*. 2024. V. 10. Article No. e32063. doi: [10.1016/j.heliyon.2024.e32063](https://doi.org/10.1016/j.heliyon.2024.e32063)
33. БиоВитрина. URL: <https://bb.ibmc.msk.ru/Main/Catalog> (дата обращения: 30.06.2025).
34. Букреева А.С., Кайшева А.Л., Мальсагова К.А., Пономаренко Е.А. Веб-репозиторий биологических образцов «БиоВитрина» версия 2.0: патент №2023685684 от 29.11.2023. URL: <https://www.fips.ru/publication-web/publications/document?type=doc&tab=PrEVM&id=1301E5EE-C372-4096-B338-228F6203A209> (дата обращения: 30.06.2025).
35. Smith J.W., Everhart J.E., Dickson W.C., Knowler W.C., Johannes R.S. Using the ADAP Learning Algorithm to Forecast the Onset of Diabetes Mellitus. *Proc. Annu. Symp. Comput. Appl. Med. Care*. 1988. P. 261–265.
36. Yang Q., Bee Y.M., Lim C.C., Sabanayagam C., Yim-Lui Cheung C., Wong T.Y., Ting D.S.W., Lim L.L., Li H., He M. et al. Use of artificial intelligence with retinal imaging in screening for diabetes-associated complications: systematic review. *EClinicalMedicine*. 2025. V. 81. Article No. 103089. doi: [10.1016/j.eclinm.2025.103089](https://doi.org/10.1016/j.eclinm.2025.103089)
37. Zhu Z., Shi D., Guankai P., Tan Z., Shang X., Hu W., Liao H., Zhang X., Huang Y., Yu H. et al. Retinal age gap as a predictive biomarker for mortality risk. *Br. J. Ophthalmol.* 2023. V. 107. P. 547–554. doi: [10.1136/bjophthalmol-2021-319807](https://doi.org/10.1136/bjophthalmol-2021-319807)
38. Betzler B.K., Chee E.Y.L., He F., Lim C.C., Ho J., Hamzah H., Tan N.C., Liew G., McKay G.J., Hogg R.E. et al. Deep learning algorithms to detect diabetic kidney disease

- from retinal photographs in multiethnic populations with diabetes. *J. Am. Med. Inform. Assoc.* 2023. V. 30. No. 12. P. 1904–1914. doi: [10.1093/jamia/ocad179](https://doi.org/10.1093/jamia/ocad179)
39. Ma Y., Xiong J., Zhu Y., Ge Z., Hua R., Fu M., Li C., Wang B., Dong L., Zhao X. et al. Deep learning algorithm using fundus photographs for 10-year risk assessment of ischemic cardiovascular diseases in China. *Sci. Bull.* 2022. V. 67. No. 1. P. 17–20. doi: [0.1016/j.scib.2021.08.016](https://doi.org/10.1016/j.scib.2021.08.016)
40. Rönn T., Perfilyev A., Oskolkov N., Ling C. Predicting type 2 diabetes via machine learning integration of multiple omics from human pancreatic islets. *Sci. Rep.* 2024. V. 14. No. 1. Article No. 14637. doi: [10.1038/s41598-024-64846-3](https://doi.org/10.1038/s41598-024-64846-3)
41. Cerezo M., Sollis E., Ji Y., Lewis E., Abid A., Bircan K.O., Hall P., Hayhurst J., John S., Mosaku A. et al. The NHGRI-EBI GWAS Catalog: standards for reusability, sustainability and diversity. *Nucleic Acids Res.* 2025. V. 53. P. D998–D1005. doi: [0.1093/nar/gkae1070](https://doi.org/10.1093/nar/gkae1070)
42. Costanzo M.C., von Grotthuss M., Massung J., Jang D., Caulkins L., Koesterer R., Gilbert C., Welch R.P., Kudtarkar P., Hoang Q. et al. The Type 2 Diabetes Knowledge Portal: An open access genetic resource dedicated to type 2 diabetes and related traits. *Cell Metab.* 2023. V. 35. No. 4. P. 695–710.e6. doi: [10.1016/j.cmet.2023.03.001](https://doi.org/10.1016/j.cmet.2023.03.001)
43. Deutsch E.W., Bandeira N., Perez-Riverol Y., Sharma V., Carver J.J., Mendoza L., Kundu D.J., Wang S., Bandla C., Kamatchinathan S. et al. The ProteomeXchange consortium at 10 years: 2023 update. *Nucleic Acids Res.* 2023. V. 51. P. D1539–D1548. doi: [10.1093/nar/gkac1040](https://doi.org/10.1093/nar/gkac1040)
44. Perez-Riverol Y., Bandla C., Kundu D.J., Kamatchinathan S., Bai J., Hewapathirana S., John N.S., Prakash A., Walzer M., Wang S. et al. The PRIDE database at 20 years: 2025 update. *Nucleic Acids Res.* 2025. V. 53. P. D543–D553. doi: [10.1093/nar/gkae1011](https://doi.org/10.1093/nar/gkae1011)
45. Yurekten O., Payne T., Tejera N., Amaladoss F.X., Martin C., Williams M., O'Donovan C. MetaboLights: open data repository for metabolomics. *Nucleic Acids Res.* 2023. V. 52. P. D640–D646. doi: [10.1093/nar/gkad1045](https://doi.org/10.1093/nar/gkad1045)
46. Wishart D.S., Guo A., Oler E., Wang F., Anjum A., Peters H., Dizon R., Sayeeda Z., Tian S., Lee B.L. et al. HMDB 5.0: the Human Metabolome Database for 2022. *Nucleic Acids Res.* 2022. V. 50. P. D622–D631. doi: [10.1093/nar/gkab1062](https://doi.org/10.1093/nar/gkab1062)
47. Khan H., Ali K., Aslam A., Singla D., Aman I. Progress in the management of type 2 diabetes mellitus: a narrative review of telerehabilitation and wearable devices. *Ann. Pediatr. Endocrinol. Metab.* 2025. V. 30. No. 2. P. 69–76. doi: [10.6065/apem.2448162.081](https://doi.org/10.6065/apem.2448162.081)
48. Casas L.A., Alarcón J., Urbano A., Peña-Zárate E.E., Sangiovanni S., Libreros-Peña L., Escobar M.F. Telemedicine for the management of diabetic patients in a high-complexity Latin American hospital. *BMC Health Serv. Res.* 2023. V. 23. No. 1. Article No. 314. doi: [10.1186/s12913-023-09267-0](https://doi.org/10.1186/s12913-023-09267-0)
49. Årsand E., Frøisland D.H., Skrøvseth S.O., Chomutare T., Tatara N., Hartvigsen, G., Tufano J.T. Mobile Health Applications to Assist Patients with Diabetes: Lessons Learned and Design Implications. *J. Diabetes Sci. Technol.* 2012. V. 6. P. 1197–1206. doi: [10.1177/193229681200600525](https://doi.org/10.1177/193229681200600525)
50. Старостина Е.Г., Володина М.Н., Старостин И.В., Бобров А.Е. Депрессия и сахарный диабет как коморбидные заболевания. *Русский медицинский журнал.* 2017. Т. 22. С. 1613–1620.
51. Schmitt A., Gahr A., Hermanns N., Kulzer B., Huber J., Haak T. The Diabetes Self-Management Questionnaire (DSMQ): development and evaluation of an instrument to assess diabetes self-care activities associated with glycaemic control. *Health Qual. Life Outcomes.* 2013. V. 11. Article No. 138. doi: [10.1186/1477-7525-11-138](https://doi.org/10.1186/1477-7525-11-138)

52. Bujang M.A., Adnan T.H., Mohd Hatta N.K.B., Ismail M., Lim C.J. A Revised Version of Diabetes Quality of Life Instrument Maintaining Domains for Satisfaction, Impact, and Worry. *J. Diabetes Res.* 2018. V. 2018. Article No. 5804687. doi: [10.1155/2018/5804687](https://doi.org/10.1155/2018/5804687)
53. *Problem Areas In Diabetes Scale (PAID)*. URL: https://professional.diabetes.org/sites/default/files/media/ada_mental_health_toolkit_questionnaires.pdf (accessed: 30.06.2025).
54. *Diabetes Distress Scale (DDS)*. URL: <https://www.mdcalc.com/calc/10523/diabetes-distress-scale-dds17> (accessed: 30.06.2025).
55. Bernabe-Ortiz A., Perel P., Miranda J.J., Smeeth L. Diagnostic accuracy of the Finnish Diabetes Risk Score (FINDRISC) for undiagnosed T2DM in Peruvian population. *Prim. Care Diabetes*. 2018. V. 12. No. 6. P. 517–525. doi: [10.1016/j.pcd.2018.07.015](https://doi.org/10.1016/j.pcd.2018.07.015)
56. Meijer J.W.G., Smit A.J., Sonderen E.V., Groothoff J.W., Eisma W.H., Links T.P. Symptom scoring systems to diagnose distal polyneuropathy in diabetes: the Diabetic Neuropathy Symptom score. *Diabet. Med.* 2002. V. 19. No. 11. P. 962–965. doi: [10.1046/j.1464-5491.2002.00819.x](https://doi.org/10.1046/j.1464-5491.2002.00819.x)
57. Yang Z., Chen R., Zhang Y., Huang Y., Hong T., Sun F., Ji L., Zhan S. Scoring systems to screen for diabetic peripheral neuropathy. *Cochrane Database Syst. Rev.* 2018. V. 2018. No. 7. Article No. CD010974. doi: [10.1002/14651858.CD010974](https://doi.org/10.1002/14651858.CD010974)
58. Iftikhar M., Saqib M., Qayyum S.N., Asmat R., Mumtaz H., Rehan M., Ullah I., Ud-Din I., Noori S., Khan M. et al. Artificial intelligence-driven transformations in diabetes care: a comprehensive literature review. *Ann. Med. Surg.* 2024. V. 86. No. 9. P. 5334–5342. doi: [10.1097/MS9.0000000000002369](https://doi.org/10.1097/MS9.0000000000002369)
59. Oikonomou E.K., Khera R. Machine learning in precision diabetes care and cardiovascular risk prediction. *Cardiovascular Diabetology*. 2023 V. 22. Article No. 259 doi: [10.1186/s12933-023-01985-3](https://doi.org/10.1186/s12933-023-01985-3)
60. Corrao S., Janić M., Maggio V., Rizzo M. Machine learning and deep learning in diabetology: revolutionizing diabetes care. *Front. Clin. Diabetes Healthc.* 2025. V. 6. Article No. 1547689. doi: [10.3389/fcdhc.2025.1547689](https://doi.org/10.3389/fcdhc.2025.1547689)
61. Darwich A.S., Polasek T.M., Aronson J.K., Ogungbenro K., Wright D.F.B., Achour B., Reny J.L., Daali Y., Eiermann B., Cook J. et al. Model-Informed Precision Dosing: Background, Requirements, Validation, Implementation, and Forward Trajectory of Individualizing Drug Therapy. *Annu. Rev. Pharmacol. Toxicol.* 2021. V. 61. P. 225–245. doi: [10.1146/annurev-pharmtox-033020-113257](https://doi.org/10.1146/annurev-pharmtox-033020-113257)
62. Emerson H., Guy M., McConville R. Offline reinforcement learning for safer blood glucose control in people with type 1 diabetes. *J. Biomed. Inform.* 2023. V. 142. Article No. 104376. doi: [10.1016/j.jbi.2023.104376](https://doi.org/10.1016/j.jbi.2023.104376)
63. Alsadoun L., Ali H., Mushtaq M., Mushtaq M., Burhanuddin M., Anwar R., Liaqat M., Bokhari S.F.H., Hasan A.H., Ahmed F. et al. Artificial Intelligence (AI)-Enhanced Detection of Diabetic Retinopathy from Fundus Images: The Current Landscape and Future Directions. *Cureus*. V. 16. No. 8. Article No. e67844. doi: [10.7759/cureus.67844](https://doi.org/10.7759/cureus.67844)

Рукопись поступила в редакцию 11.05.2025, переработанный вариант поступил 27.06.2025.

Дата опубликования 05.08.2025.

Modern Machine Learning Approaches for Predicting Type 2 Diabetes Mellitus

Butkova T.V.¹, Kulikova L.I.², Malsagova K.A.¹, Petrovskiy D.V.¹,
Rudnev V.R.³, Ryskina E.A.⁴, Kulikov A.V.³, Kaysheva A.L.¹

¹*Institute of Biomedical Chemistry, Moscow, Russia*

²*Institute of Mathematical Problems of Biology RAS, Keldysh Institute of Applied Mathematics RAS, Pushchino, Russia*

³*Institute of Theoretical and Experimental Biophysics RAS, Pushchino, Russia*

⁴*HSE University, Moscow, Russia*

Abstract. Between 1990 and 2022, the number of people with diabetes increased from 200 million to 830 million, with the majority of cases attributed to non-insulin-dependent type 2 diabetes. This condition often leads to severe complications such as blindness, kidney failure, myocardial infarction, stroke, and limb amputations. According to WHO estimates, more than 2 million people died in 2021 from diabetes and kidney-related diseases. Global healthcare expenditures related to diabetes were estimated at 966 billion USD in 2021 and are projected to reach 1.054 trillion USD by 2045. Early diagnosis and a personalized approach to treatment can slow disease progression and reduce the risk of vascular complications. At early stages, lifestyle modification alone can achieve long-term normoglycemia without the need for medication. However, nearly half of all people with diabetes worldwide remain undiagnosed. One of the key challenges remains the prevention of hypoglycemia, particularly nocturnal hypoglycemia, which can lead to dangerous complications. Today, the use of continuous glucose monitoring data combined with machine learning algorithms enables the prediction of hypoglycemia risk, taking into account individual patient characteristics. Modern artificial intelligence models rely on large-scale, multimodal data analysis and can adapt in real time. Addressing the issue of data imbalance improves the reliability of predictions. Further progress requires the development of large data libraries and secure systems for information exchange between monitoring devices and clinical applications. The implementation of such technologies enhances treatment effectiveness and improves patients' quality of life. This review highlights the prospects for integrating artificial intelligence into diabetes monitoring and prediction.

Key words: *type 2 diabetes mellitus, early diagnosis, artificial intelligence, machine learning models*