

ГЛУШКОВ В.С.,

к.м.н., ФГБОУ Тюменский ГМУ Минздрава России, г. Тюмень, Россия,
e-mail: glushkovvs@tyumsmu.ru

ВДОВИН Е.П.,

д.ф.-м.н., доцент, ФГАОУ ВО «Тюменский государственный университет», г. Тюмень, Россия,
e-mail: e.p.vdovin@utmn.ru

ЕРМАКОВ Н.В.,

ООО "Ин Нова", г. Тюмень, Россия, e-mail: n.ermakov@innovalab.ru

БАКАНОВСКАЯ Л.Н.,

к.т.н., доцент, ФГАОУ ВО «Тюменский государственный университет», г. Тюмень, Россия,
e-mail: l.n.bakanovskaya@utmn.ru

ЧЕРНЫШЕВА Т.Ю.,

к.т.н., доцент, ФГАОУ ВО «Тюменский государственный университет», г. Тюмень, Россия,
e-mail: t.y.chernysheva@utmn.ru

КРАВЕЦ В.Д.,

ФГАОУ ВО «Тюменский государственный университет», г. Тюмень, Россия, e-mail: vkravets03@gmail.com

СОБОЛЕВ И.С.,

ФГАОУ ВО «Тюменский государственный университет», г. Тюмень, Россия,
e-mail: ilia.sobolev2014@yandex.ru

ВОЛКОВ Д.Е.,

ФГАОУ ВО «Тюменский государственный университет», г. Тюмень, Россия, e-mail: vbnzwolf@gmail.com

МИЛЯЕВ М.В.,

ФГАОУ ВО «Тюменский индустриальный университет», г. Тюмень, Россия, e-mail: epev.nd@mail.ru

ПОДХОД К РАЗРАБОТКЕ МОДУЛЬНОЙ АРХИТЕКТУРЫ БАЗЫ ДАННЫХ В ОБЛАСТИ ИНТЕНСИВНОЙ ТЕРАПИИ И РЕАНИМАЦИИ

DOI: 10.25881/18110193_2025_2_54

Аннотация. В статье представлено проектирование базы данных, предназначенной для оптимизации хранения и обработки медицинских данных, с акцентом на поддержку принятия решений в области интенсивной терапии и реанимации. Целью исследования является разработка логической модели базы данных на основе передовых принципов и методов, используемых в международных проектах открытых баз данных, способной минимизировать ошибки, связанные с человеческим фактором, и улучшить точность прогноза состояния пациентов в реальном времени.

Методология работы основана на сравнительном анализе существующих международных медицинских баз данных, таких как MIMIC-IV и eICU. Для проектирования новой базы данных применен инновационный модульный подход, который обеспечивает гибкость и масштабируемость системы.

Основные результаты работы заключаются в создании логической модели базы данных, которая может быть эффективно использована в российской системе здравоохранения, в том числе в удаленных и малоресурсных регионах. Логическая модель разработана с учётом специфики медицинских данных, включая модули для хранения информации о госпитализациях, показателях состояния пациентов, лабораторных исследованиях, медикаментозных назначениях и других аспектах клинической практики. Важной частью исследования является интеграция базы данных с российскими медицинскими информационными системами и адаптация к национальным стандартам и нормативным требованиям.

Созданная архитектура логической модели минимизирует влияние человеческого фактора, автоматизирует анализ данных и может использоваться в разработке систем поддержки принятия врачебных решений. Практическая значимость заключается в повышении качества медицинской помощи и снижении нагрузки на персонал. Система применима в российских учреждениях, включая удаленные регионы, и способствует цифровизации здравоохранения.

Ключевые слова: медицинская база данных, модульная архитектура базы данных, интенсивная терапия, реанимация, машинное обучение, обработка временных рядов, российская система здравоохранения.

Для цитирования: Глушков В.С., Вдовин Е.П., Ермаков Н.В., Бакановская Л.Н., Чернышева Т.Ю., Кравец В.Д., Соболев И.С., Волков Д.Е., Миляев М.В. Подход к разработке модульной архитектуры базы данных в области интенсивной терапии и реанимации. *Врач и информационные технологии*. 2025; 2: 54-69. doi: 10.25881/18110193_2025_2_54.

GLUSHKOV V.S.,

PhD, Federal State Budgetary Educational Institution of Higher Education "Tyumen State Medical University" of the Ministry of Health of the Russian Federation, Tyumen, Russia, e-mail: glushkovs@tyumsmu.ru

VDOVIN E.P.,

DSc, Associate Professor, University of Tyumen, Tyumen, Russia, e-mail: e.p.vdovin@utmn.ru

ERMAKOV N.V.,

"In Nova" LLC, Tyumen, Russia, e-mail: n.ermakov@innovalab.ru

BAKANOVSKAYA L.N.,

PhD, Associate Professor, University of Tyumen, Tyumen, Russia, e-mail: l.n.bakanovskaya@utmn.ru

CHERNYSHEVA T.YU.,

PhD, Associate Professor, University of Tyumen, Tyumen, Russia, e-mail: t.y.chernysheva@utmn.ru

KRAVETS V.D.,

University of Tyumen, Tyumen, Tyumen, Russia, e-mail: vkravets03@gmail.com

SOBOLEV I.S.,

University of Tyumen, Russia, Tyumen, Russia, e-mail: ilia.sobolev2014@yandex.ru

VOLKOV D.E.,

University of Tyumen, Tyumen, Russia, e-mail: vbnzwolf@gmail.com

MILYAEV M.V.,

Industrial University of Tyumen, Tyumen, Russia, e-mail: epev.nd@mail.ru

AN APPROACH FOR MODULAR DATABASE ARCHITECTURE DESIGN IN THE INTENSIVE CARE UNIT

DOI: 10.25881/18110193_2025_2_54

Abstract. This article presents the design of a database intended to optimize the storage and processing of medical data, with a focus on decision support in intensive care and resuscitation. The aim of the study is to develop a logical database model based on advanced principles and methods used in international open database projects, capable of minimizing human error and enhancing the accuracy of real-time patient prognosis.

The methodology is founded on a comparative analysis of existing international medical databases, such as MIMIC-IV and eICU. An innovative modular approach was applied in designing the new database, ensuring system flexibility and scalability. The primary outcome is the creation of a logical database model that can be effectively utilized within the Russian healthcare system, including remote and low-resource regions. The logical model was developed taking into account the specifics of medical data, including modules for storing information on hospitalizations, patient condition indicators, laboratory tests,

medication prescriptions and other aspects of clinical practice. An important part of the study is the integration of the database with Russian medical information systems and adaptation to national standards and regulatory requirements. The developed architecture of the logical model minimizes the influence of the human factor, automates data analysis and can be used in the development of medical decision support systems. The practical significance lies in improving the quality of medical care and reducing the burden on the staff. The system is applicable in Russian institutions, including remote regions, and contributes to the digitalization of healthcare.

Keywords: *medical database, modular database architecture, intensive care, resuscitation, machine learning, time series analysis, Russian healthcare system.*

For citation: *Glushkov V.S., Vdovin E.P., Ermakov N.V., Bakanovskaya L.N., Chernysheva T.YU., Kravets V.D., Sobolev I.S., Volkov D.E., Milyaev M.V. An approach for modular database architecture design in the intensive care unit. Medical doctor and information technology. 2025; 2: 54-69. doi: 10.25881/18110193_2025_2_54.*

ВВЕДЕНИЕ

В Указе Президента Российской Федерации от 7 мая 2024 года № 309 [1] «О национальных целях развития Российской Федерации на период до 2030 года и на перспективу до 2036 года» Президент В.В. Путин установил новые целевые показатели и задачи, направленные на достижение национальной цели «Сохранение населения, укрепление здоровья и повышение благополучия людей, поддержка семьи». В Едином плане Правительства РФ [2] определены стратегические приоритеты и соответствующие показатели на ближайшие 12 лет, включая задачу создания и ввода в эксплуатацию к 2030 году цифровой платформы, направленной на формирование, поддержание и сохранение здоровья человека на протяжении всей его жизни, основанной на принципах управления на основе данных. Очевидно, что создание цифровой платформы способно существенно повысить качество предоставляемой медицинской помощи, а следовательно уровень здоровья населения, и в перспективе решить долгосрочные задачи, в том числе выхода на устойчивую траекторию экономического роста, что также установлено указом Президента.

Ещё одной из наиболее актуальных проблем для Российской Федерации является недостаток квалифицированных специалистов в удаленных регионах, так кадровый дефицит анестезиологов-реаниматологов в стране по данным на 2023 г. составил 85,6%, а коэффициент совместительства — 1,6. При этом обеспеченность кадрами многих регионов Южного, Северо-Западного, Центрального федеральных округов отстает от среднероссийских [3–5]. Эти данные свидетельствуют о существенной нагрузке на медицинский персонал отделений реанимации и интенсивной терапии (ОРИТ).

Таким образом, важной задачей системы здравоохранения РФ является активный переход к цифровым технологиям, особенно в области интенсивной терапии и реанимации, где требуется быстрое и точное принятие решений [6–11]. Вместо разработки медицинской информационной системы, охватывающей отдельные учреждения, регионы или группы организаций, необходимо создать единую федеральную систему.

Традиционные системы мониторинга в ОРИТ не всегда способны интегрировать информацию

из различных источников, что приводит к фрагментации данных и снижает точность прогнозов. Врачи по-прежнему тратят значительное время на ручной ввод данных в медицинские информационные системы, что снижает их продуктивность. Сложность одновременной интерпретации большого количества медицинских данных вручную снижает эффективность работы врачей. Однако широкое внедрение информационных систем в учреждения здравоохранения ожидаемо приводит к многолетнему накоплению данных о миллионах пациентов. В результате врачи и медицинский персонал сталкиваются с нагрузкой, обусловленной растущим объемом данных, которые необходимо анализировать и на основе которых необходимо оперативно принимать решения. Это осложняется ограниченным временем, часто недостатком инструментов и опыта для эффективной обработки и использования этих данных в повседневной практике [12, 13].

Особенно остро данная проблема проявляется в ОРИТ, где требуется круглосуточное наблюдение за состоянием пациентов. В результате наблюдений **формируются временные ряды данных в реальном времени, позволяющие при детальном изучении совершенствовать клиническую практику**. Временные ряды в медицинских данных содержат ключевую информацию о состоянии пациентов, включая динамику изменения жизненно важных показателей, реакции на терапию и риск осложнений. Анализ этих данных требует применения специальных методов для работы с нерегулярными интервалами, шумом и аномалиями.

Существующее ежегодное увеличение числа пациентов на 5–7% приводит к дефициту времени у медицинских работников, снижает их способность обрабатывать большие объемы медицинской информации и оперативно принимать обоснованные решения. В этих условиях возрастает вероятность ошибок, обусловленных человеческим фактором, что может привести к ухудшению состояния пациентов и увеличению смертности.

Другой важной проблемой является дефицит лекарственных препаратов в региональных больницах, что усложняет подбор эффективной схемы лечения [14] для пациентов в критическом состоянии.

Недостаток механизмов автоматического отслеживания состояния пациентов и оповещения родственников о критических изменениях также создает дополнительные риски и увеличивает нагрузку на медицинский персонал.

Описание выше ситуация в здравоохранении Российской Федерации и в Тюменской области, в частности, в последние годы вызывает необходимость масштабной цифровизации, которая может изменить систему здравоохранения, повышая доступность и качество медицинской помощи, особенно в регионах с ограниченными медицинскими ресурсами [15–18]. Решение данной задачи привлекло внимание к методам машинного обучения (МО) в медицинской сфере, которые эффективно используют большие объёмы доступных данных для извлечения прикладных знаний, тем самым предсказывая медицинские результаты и улучшая процесс принятия медицинских решений, что позволит автоматизировать процедуры:

- а) мониторинга и управления состоянием пациентов с акцентом на ОРПТ,
- б) обоснованного принятия решений.

Несмотря на уже имеющийся успех применения традиционного МО в области медицины, разработка эффективных прогнозных моделей остается сложной задачей. Из-за многомерной природы медицинских данных, как правило, для каждой новой задачи прогнозирования выбирается только ограниченный набор подходящих признаков из тысячи кандидатов, что требует трудоемкого и длительного процесса. Это часто нуждается в привлечении медицинских экспертов для извлечения, предварительной обработки и очистки данных из разных источников. Таким образом, перед ведущими учёными Тюмени и Тюменской области особо остро на сегодняшний день стоит задача решения таких вопросов, как:

- а) отсутствие поддержки временных рядов, т.к. критически важны временные зависимости между показателями состояния пациента, например, артериальное давление, пульс, газы крови;
- б) отсутствие механизмов интерпретации результатов, полученных с использованием искусственного интеллекта, основываясь на прогнозных моделях, а также объяснимых метрик, которые могут быть представлены

- врачам, и оперативное принятие решений;
- в) недостаточная интеграция данных внешних систем, возможная интеграция с амбулаторными данными или источниками из других отделений;
- г) отсутствие полноценных механизмов оповещения.

Целью настоящего исследования является разработка логической модели базы данных с модульной архитектурой, основанной на передовых принципах международных проектов Medical Information Mart for Intensive Care Database-IV (MIMIC-IV) и eICU Collaborative Research Database (eICU-CRD), для создания интегрированной медицинской информационной системы.

Достижение поставленной цели определяет решение следующих *научно-исследовательских задач*:

1. Анализ и систематизация подходов к проектированию и разработке открытых медицинских баз данных:
 - изучение структуры и архитектуры существующих проектов MIMIC-III и MIMIC-IV;
 - анализ основных принципов и методов проектирования баз данных, применяемых в этих проектах;
 - сравнение и выбор наиболее подходящих решений для разработки новой базы данных.
2. Проектирование и разработка прототипа базы данных:
 - определение требований к структуре и функциональным возможностям разрабатываемой базы данных;
 - разработка логической модели данных, включающей необходимые таблицы, связи между ними и механизмы обеспечения целостности данных.

Таким образом, разработанная база данных позволит реализовать следующие функции в медицинской информационной системе:

1. автоматизировать анализ медицинских данных и прогнозирования критических ситуаций, минимизировав ошибки, связанные с человеческим фактором;
2. поддерживать выбор оптимальной схемы лечения с учетом наличия лекарственных препаратов и персональных данных пациента;
3. отслеживать состояние пациентов в режиме

реального времени и предоставлять медицинскому персоналу рекомендации по лечению на основе индивидуальных параметров пациента;

4. автоматизировать процесс оповещения родственников о состоянии пациента и критических изменениях в его состоянии.

Объектом исследования является процесс разработки и внедрения базы данных на основе передовых принципов и методов, используемых в международных проектах открытых баз данных MIMIC-III и MIMIC-IV.

Предметом исследования являются методы и технологии проектирования, организации и управления данными в медицинских информационных системах, а также их интеграция в единую базу данных для интенсивной терапии.

МАТЕРИАЛЫ И МЕТОДЫ

Для проведения исследования были проанализированы структуры существующих медицинских наборов данных MIMIC-IV и eICU, которые зарекомендовали себя в качестве высокоценного ресурса для мировых научных исследований и разработки новых методов диагностики и лечения, основанных на данных, собранных в ОРИТ и в отделении неотложной помощи. Разработчиками [19] утверждается, что на 2024 г. опубликовано свыше 1500 научных работ, в которых использованы данные проектов MIMIC-III, MIMIC-IV и eICU. Это говорит об их высокой научной ценности, валидации данных и возможности использовать для развития предиктивных моделей, подтверждения точности и применимости цифрового двойника в условиях, приближенных к реальной практике, так как качество и объем данных влияют на точность прогнозов и снижают число ошибок, связанных с человеческим фактором. В ходе аналитического изучения

MIMIC-IV и eICU был проведен систематический анализ и критический обзор научных статей и диссертаций с использованием международных научных баз данных с высоким уровнем цитирования Springer, IEEE, MDPI, Lancet, Scopus. Основное внимание уделено публикациям, освещающим убедительные доказательства эффективности использования данных MIMIC-IV:

- а) прогнозирование состояния пациентов в ОРИТ на основе временных рядов данных [20–31];
- б) интеграция данных с внешних источников, таких как амбулаторные записи и данные из других отделений. [32–38];
- в) реализация эффективных механизмов поддержки принятия решений [39–49].

Так как MIMIC-IV — это новая версия датасета, которая стала более структурированной, то за основу данного исследования был взят именно этот проект и версия MIMIC-IV v3.1 on BigQuery [50]. **MIMIC-IV** — это крупный медицинский датасет, структурированный на основе модульного подхода, является уникальной базой данных для исследований в области медицины и искусственного интеллекта по нескольким причинам.

Основные преимущества MIMIC-IV для проведения исследований (рисунок 1):

Объем данных включает в себя 546 028 уникальных госпитализаций, что соответствует 223 452 уникальным пациентам. В дополнение к этому имеется 94 458 записей о пребывании в ОРИТ для 65 366 уникальных записей. Эти данные содержат записи витальных показателей, лабораторных исследований и медицинских процедур, что делает MIMIC-IV одним из самых обширных наборов данных в области критической медицины.

Разнообразие данных: данные представлены в различных форматах, включая как

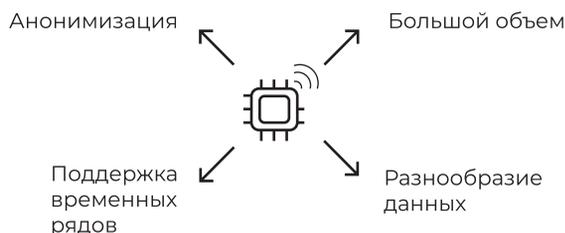


Рисунок 1 — Характеристики данных MIMIC-IV.

структурированные, так и неструктурированные данные. Структурированные данные, такие как результаты лабораторных анализов, легко поддаются количественному анализу. В то же время неструктурированные данные, например, текстовые записи врачей, предоставляют ценную информацию, которая может быть использована для более глубокого понимания клинических случаев и принятия решений.

MIMIC-IV содержит **временные ряды данных**, что открывает возможности для применения алгоритмов МО и статистического анализа. Эти временные ряды позволяют исследователям прогнозировать изменения состояния пациентов во времени, что может быть критически важным для улучшения качества медицинского обслуживания и принятия решений в экстренных ситуациях.

Одним из ключевых аспектов MIMIC-IV является **анонимизация данных**, что обеспечивает защиту конфиденциальности пациентов. Все записи были обработаны таким образом, чтобы исключить возможность идентификации индивидуальных лиц, что позволяет исследователям безопасно использовать данные для анализа и разработки новых методов лечения, не нарушая этических норм.

База данных MIMIC-IV структурирована в два основных модуля: **hosp** и **icu**, что обеспечивает синхронизацию данных. Модуль **hosp** содержит информацию, полученную из общей электронной медицинской записи больницы (EHR), в то время как модуль **icu** включает данные, собранные из клинической информационной системы ОРИТ (MetaVision). MIMIC-IV можно связать с другими проектами MIMIC, опубликованными на PhysioNet (рисунок 2) [51]:

Каждый модуль содержит таблицы, хранящие данные по определённой области. Например, данные госпитализации пациентов, лабораторные анализы и назначения лекарств разделены на независимые модули.

Уникальные единые идентификаторы используются для связи всех модулей:

subject_id — уникальный идентификатор таблицы **patients** модуля MIMIC-IV-Core однозначно определяет отдельного пациента. Любые строки, связанные с одним **subject_id**, относятся к одному и тому же пациенту. Как первичный ключ для таблицы, **subject_id** уникален для каждой строки (записи).

hadm_id — уникальный идентификатор госпитализации таблицы **admissions** модуля MIMIC-IV-Core, который хранит запись о

MIMIC-IV is separated into "modules" to reflect the provenance of the data.

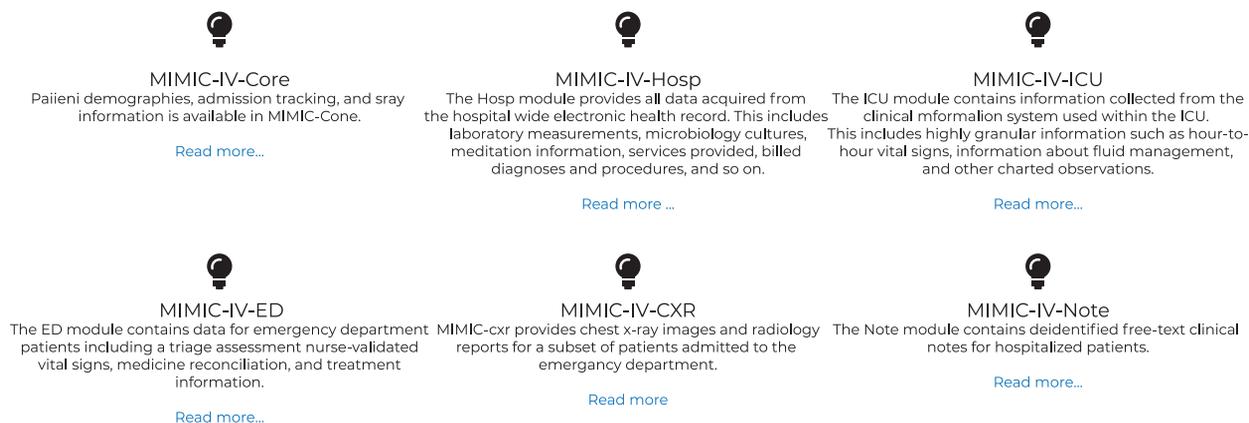


Рисунок 2 — Основные структурные компоненты MIMIC-IV v3.1 [51].

поступлении одного пациента в больницу. Таблица admissions имеет внешний ключ (ссылку) на таблицу patients.

stay_id — идентификатор, который однозначно идентифицирует отдельное пребывание одного пациента в отделении неотложной помощи таблицы main модуля MIMIC-IV-ED. В данной таблице также хранится subject_id, который определяет уникального пациента. Любые строки, связанные с одним subject_id, относятся к одному и тому же пребыванию пациента.

Таблица diagnoses_icd модуля MIMIC-IV-Hosp имеет subject_id (ссылку) на таблицу patients. Эта таблица содержит запись всех диагнозов с использованием онтологий МКБ-9 и МКБ-10, за которые пациенту был выставлен счет во время пребывания в больнице.

Как видим из примеров выше, уникальные идентификаторы позволяют связывать данные из разных модулей и в то же время сохраняют возможность каждому модулю быть автономным.

Модульный подход в организации баз данных (рисунок 3) предполагает разделение данных на независимые логические блоки (модули), каждый из которых отвечает за конкретную

функцию или область данных. Разделение таблиц по модулям оптимизирует работу с данными и анализ, минимизируя объём данных, которые необходимо обрабатывать для конкретных задач, и обеспечивает масштабируемость: легко добавлять новые модули, не влияя на существующую систему. Например, можно добавить модуль Imaging для хранения данных медицинских изображений (КТ, МРТ).

Каждый модуль может обновляться или изменяться, не затрагивая другие, что повышает гибкость системы и устойчивость к изменениям. Модульная структура позволяет разрабатывать и тестировать модули независимо. Это снижает риски ошибок и упрощает процесс обновления системы.

Можно выделить следующие преимущества модульного подхода. Улучшает интеграцию: упрощается взаимодействие с внешними источниками данных и импорт данных из внешних источников. Например, данные лабораторий можно подключить к модулю Diagnostics, а устройства мониторинга пациентов — к модулю Monitoring и т.п. Объединение данных для МО: слияние данных из модулей Diagnostics и Imaging для анализа. Повышенная производительность

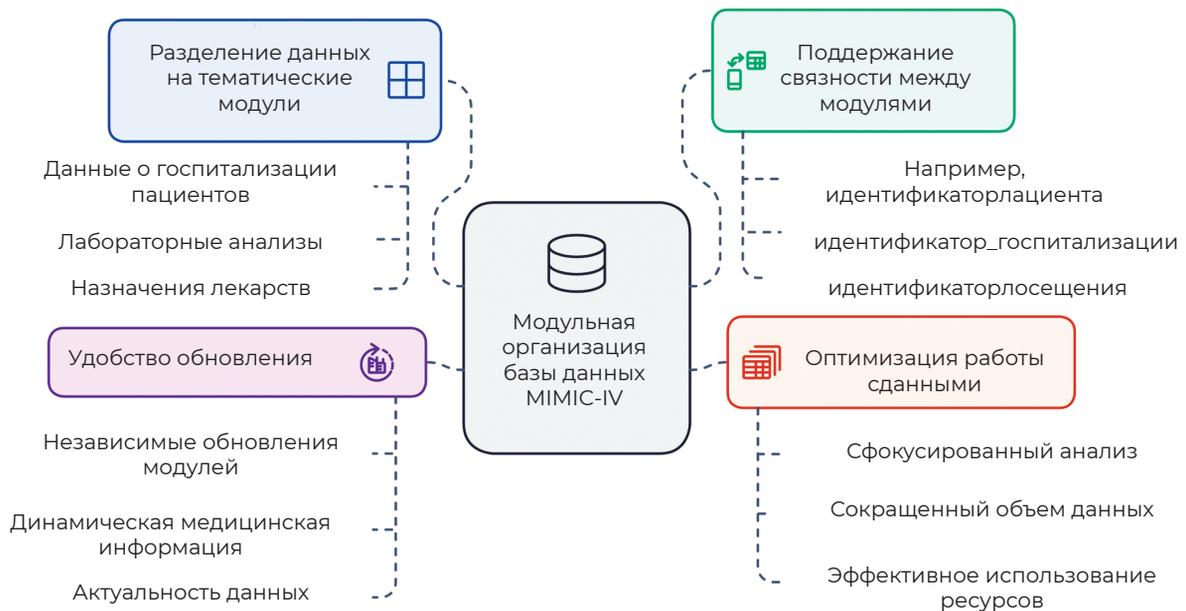


Рисунок 3 — Преимущества модульного подхода MIMIC-IV.

Таблица 1 — Структура модулей базы данных MIMIC-IV

Модуль	Описание	Ключевые таблицы
MIMIC-IV-Core	Общая информация о госпитализациях и пациентах	admissions patients services stays transfers
MIMIC-IV-Hosp	Все данные, полученные из электронной медицинской карты больницы. Информация включает в себя информацию о пациенте и поступлении, лабораторные измерения, микробиологические исследования, некоторые таблицы включают также данные из внешних учреждений (например, амбулаторные лабораторные анализы), прием лекарств и выставленные счета за диагнозы.	d_hcpcs d_icd_diagnoses d_icd_procedures d_labitems d_micro diagnoses_icd drgcodes emar emar_detail hcpcsevents labevents microbiologyevents procedures_icd
MIMIC-IV-ICU	Данные о нахождении пациентов в ОРИТ, собранные из клинической информационной системы, используемой в ОРИТ. Включает в себя высокодетализированную информацию, такую как почасовые показатели жизнедеятельности, информацию о регулировании жидкости и другие картированные наблюдения.	icustays d_items chartevents datetimeevents inputevents outputevents procedureevents
MIMIC-IV-CXR	Содержит связующую таблицу между файлом DICOM, соответствующему одной рентгенограмме грудной клетки и пациентом, с клиническими данными из других модулей MIMIC-IV	cxr

и ускорение анализа данных: запросы обращаются только к необходимому модулю, что сокращает нагрузку на базу данных.

Модульная структура MIMIC-IV поддерживает международные стандарты кодирования информации, такие как HL7 FHIR. Например, модуль MIMIC-IV-Hosp можно построить на основе ресурса FHIR Observation. Модуль MIMIC-IV-ICU может использовать ресурс FHIR MedicationRequest.

Таким образом, анализируемый датасет состоит из нескольких модулей, каждый из которых содержит определенный тип информации о пациентах и таблиц (таблица 1):

Как видно из логики структуры, каждый модуль MIMIC-IV фокусируется на одной области или задаче и может быть частью системы поддержки принятия решений (СППР) для оценки рисков и определения приоритетности лечения пациентов в ОРИТ. Такой подход позволяет масштабировать систему, добавляя новые модули (например, для анализа изображений).

Проведённый анализ медицинских проектов, таких как MIMIC-III, MIMIC-IV и eICU, показал сложную модульную архитектуру базы данных, включающую множество таблиц, временных рядов, текстов и изображений. Выбор подходящей структуры данных и её организация существенно повлияют на: скорость извлечения данных для обучения моделей, точность моделей за счёт устранения избыточности или недостатка данных, гибкость и расширяемость системы при изменении данных или требований, что необходимо учесть в разрабатываемой базе данных.

РЕЗУЛЬТАТЫ

Логическая модель (рисунок 4) разрабатываемой базы данных была построена на основе опросов заказчиков и функций будущей базы данных, основываясь на аналогичной структуре MIMIC-IV, и структурирована на следующие модули:

Модуль госпитализации — содержит сущности, относящиеся к госпитализации пациента: пациент, госпитализация, обстоятельство посещения.

Модуль госпитализации в ОРИТ — содержит сущности, относящиеся к госпитализации пациента в ОРИТ: госпитализация, госпитализация в ОРИТ и регистрация показателей и иной информации о состоянии пациента.

Модуль показателей — содержит сущности, относящиеся к регистрации показателей состояния пациента при его нахождении в ОРИТ: показатель пациента, регистрация показателя пациента, регистрация основных показателей пациента, единица измерения.

Модуль препаратов — содержит сущности, относящиеся к введению препаратов в организм пациента при его нахождении в ОРИТ: препарат, группа препарата, назначение препарата, регистрация введения препарата, лекарственное средство.

Модуль дренажей — содержит сущности, относящиеся к дренажам пациента: дренажное оборудование, регистрация дренажей пациента.

Модуль лабораторных исследований — содержит сущности, относящиеся к лабораторным исследованиям пациента: профиль лабораторного исследования, лабораторный тест, лабораторное исследование.

Модуль диагнозов — содержит сущности, относящиеся к диагнозам пациента за текущую госпитализацию: диагноз, код МКБ.

Модуль истории болезни — содержит сущности, относящиеся к истории болезни пациента, включая назначения и результаты фармакотерапии, и сопутствующие заболевания пациента: сопутствующее заболевание, регистрация результата осмотра в истории болезни, результат фармакотерапии.

Сущности объединены связями в общей схеме, но для удобства восприятия поделены на логические части. Сущности были разделены на семантические модули — одна сущность может присутствовать в нескольких модулях. Для каждой таблицы определены уникальные идентификаторы (например, `patient_id`, `admission_id`), первичные ключи для каждой таблицы, внешние ключи для связи таблиц между модулями.

ОБСУЖДЕНИЕ

Достоверный анализ медицинских данных требует высокой точности и полноты исходной информации. Однако базы данных, такие как MIMIC-IV, содержащие данные пациентов из ОРИТ, часто сталкиваются с проблемой пропусков и нерегулярности временных рядов. Эти особенности обусловлены как особенностями сбора данных (например, пропуск регистраций), так и техническими ограничениями оборудования. Пропуски данных могут значительно повлиять на точность предсказаний, интерпретацию результатов и принятие клинических решений.

В разрабатываемой базе данных реализуются современные методы обработки пропусков и временных рядов, которые могут быть интегрированы в модульную архитектуру базы данных, что будет являться преимуществом и научной новизной системы. Описанный подход к архитектуре разрабатываемой базы данных позволяет повысить качество данных и точность аналитических выводов, что особенно важно для поддержки принятия решений в клинической практике и использования алгоритмов МО.

Разрабатываемая модульная структура впервые адаптирована для поставленных задач:

1. *Реализуется интеграция базы данных с российскими медицинскими информационными системами* (например, «1С:Медицина», «МедИнфо»). Это позволяет обеспечить совместимость с уже существующими платформами для управления медицинскими данными в больницах.
2. *Локализация форматов данных: модуль госпитализации*, который отвечает за данные из общей больничной системы EHR, адаптирован под российские стандарты медицинской документации, такие как унифицированные формы записи медицинской помощи (например, Форма № 027/y).
3. *Поддержка нормативных требований*: включение модулей в систему было разработано с учетом российских требований по защите персональных данных (152-ФЗ) и стандартов хранения медицинской информации. Это обеспечивает правовую основу для использования таких систем в РФ.
4. *Расширение модуля госпитализации в ОРИТ*: модуль дополнен поддержкой данных из российских клинических протоколов (например, рекомендации Минздрава России для ОРИТ).

Это помогает учитывать локальные особенности ведения пациентов.

5. Включение русского языка в обработку текстовых данных: реализована возможность анализа текстовых данных врачебных записей на русском языке с использованием методов NLP, что позволяет обрабатывать электронные истории болезни и рекомендации на национальном языке.
6. Адаптация к удаленным регионам: разработана инфраструктура для работы с распределенными данными в условиях ограниченной цифровой инфраструктуры, что важно для внедрения системы в удаленных регионах РФ.

Научная и практическая значимость адаптации

Эти изменения делают разрабатываемую базу данных не просто исследовательским инструментом, но и практически применимой платформой для анализа и оптимизации работы российских медицинских учреждений.

Проведенный анализ и разработанная логическая схема базы данных с модульной структурой, аналогичной MIMIC-IV, обладает значительным потенциалом для:

1. выявления скрытых закономерностей между различными показателями здоровья.
2. персонализации лечения, учитывая индивидуальные особенности пациента.
3. прогнозирования реакции пациента на терапевтические методы.

Этот потенциал обеспечивается несколькими ключевыми характеристиками системы:

1. Выявление скрытых закономерностей
 - *Объединение разнородных данных:* Модульная структура позволяет хранить различные типы данных (временные ряды, лабораторные результаты, демографические данные) в одной системе, связывая их с уникальным идентификатором пациента.
 - *Гибкость и масштабируемость:* Новые модули можно добавлять без нарушения существующей структуры. Например, можно интегрировать геномные данные или данные носимых устройств.

Механизм:

2. Анализ временных рядов:
 - Сопоставление изменений физиологических параметров (например, частота

сердечных сокращений, артериальное давление) с клиническими событиями (диагнозы, процедуры) для выявления паттернов ухудшения состояния пациента.

- Применение алгоритмов временных рядов (LSTM, Prophet) для анализа тенденций.
3. Корреляция многомерных данных:
 - Сопоставление лабораторных анализов с демографическими данными и медицинскими процедурами.
 - Использование методов корреляционного анализа или кластеризации для обнаружения закономерностей (например, связь уровня сахара и давления у разных групп пациентов).
 4. Применение методов МО:
 - Алгоритмы МО (градиентный бустинг, случайный лес, нейронные сети) могут обучаться на данных из разных модулей для поиска сложных нелинейных взаимосвязей. Например, данные о пациентах, поступающих с подозрением на сепсис, могут быть использованы для выявления предикторов, предсказывающих развитие критического состояния, в частности на основе комбинации количества лейкоцитов, биохимических показателей и параметров гемодинамики. Другим примером может служить предсказание риска развития сердечно-сосудистых заболеваний на основе анамнеза и текущих физиологических показателей.

ЗАКЛЮЧЕНИЕ

Использование базы данных MIMIC-IV (данные госпитализаций 69653 пациентов из медицинского центра Beth Israel Deaconess, США за 2008–2019 гг.), которая является одной из крупнейших и наиболее полных баз данных электронных медицинских записей, открывает новые возможности для создания интеллектуальных систем поддержки принятия решений в медицине. Внедрение таких систем в российскую медицинскую практику позволит не только повысить качество оказываемой помощи, но и снизить нагрузку на медицинский персонал, уменьшить количество ошибок, связанных с человеческим фактором, и улучшить прогнозы для пациентов.

Поставленная цель работы достигнута. Разработанная модель базы данных не только соответствует современным требованиям к медицинским информационным системам, но и предлагает инновационные решения для повышения качества медицинской помощи. Практическая значимость работы заключается в возможности её внедрения в российскую систему здравоохранения, что особенно актуально для удаленных и малоресурсных регионов.

Исследование выполняется *при поддержке Министерства науки и высшего образования Российской Федерации* в рамках проекта "Фундаментальные проблемы методики разработки и связанного с ней правового и этического регулирования в сфере применения систем и моделей искусственного интеллекта" (FEWZ-2024-0016).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

ЛИТЕРАТУРА/REFERENCES

1. Указ Президента Российской Федерации от 07.05.2024 г. №309 «О национальных целях развития Российской Федерации на период до 2030 года и на перспективу до 2036 года». [Ukaz Prezidenta Rossijskoj Federacii ot 07.05.2024 g. №309 «O nacional'nyh celyah razvitiya Rossijskoj Federacii na period do 2030 goda i na perspektivu do 2036 goda». (In Russ.)]
2. Единый план по достижению национальных целей развития Российской Федерации до 2030 года и на перспективу до 2036 года (утв. Правительством РФ). [Edinyj plan po dostizheniyu nacional'nyh celej razvitiya Rossijskoj Federacii do 2030 goda i na perspektivu do 2036 goda (utv. Pravitel'stvom RF). (In Russ.)]
3. Покида А.Н., Зыбуновская Н.В. Здоровье в восприятии россиян и реальные медицинские практики // Здоровье населения и среда обитания. 2021. — Т.29. — №7. — С.19–27. [Pokida AN, Zybunovskaya NV. Health in the perception of Russians and real medical practices. Zdorov'e Naseleniya i Sreda Obitaniya. 2021; 29(7): 19-27. (In Russ.)] doi: 10.35627/2219-5238/2021-29-7-19-27.
4. Сычев Е.В., Есауленко И.Э., Петрова Т.Н., Петров И.С. Особенности кадровой политики сельского здравоохранения и пути повышения ее эффективности // Наука молодых (Eruditio Juvenium). — 2023. — Т.11. — №4. — С.535-544. [Sychev EV, Esaulenko IE, Petrova TN, Petrov IS. Personnel Policy of Rural Healthcare and Ways to Improve its Effectiveness. Science of the young (Eruditio Juvenium). 2023; 11(4): 535-544. (In Russ.)] doi: 10.23888/HMJ2023114535-544.
5. Романюк Т.И., Поздняков Д.Ю., Мушенок Ф.Б. Использование возможностей машинного обучения и искусственного интеллекта в отделениях анестезиологии и реанимации // Врач и информационные технологии. — 2021. — №2. — С.60-71. [Ramaniuk TI, Pozdnyakov DYU, Mushenok FB. Use of the opportunities of machine learning and artificial intelligence in the intensive care unit. Medical doctor and information technology. 2021; 2: 60-71. (In Russ.)] doi: 1025881/18110193_2021_2_60.
6. Романюк Т.И., Поздняков Д.Ю., Мушенок Ф.Б. Использование возможностей машинного обучения и искусственного интеллекта в отделениях анестезиологии и реанимации // Врач и информационные технологии. — 2021. — №2. — С.60-71. [Ramaniuk TI, Pozdnyakov DYU, Mushenok FB. Use of the opportunities of machine learning and artificial intelligence in the intensive care unit. Medical doctor and information technology. 2021; 2: 60-71. (In Russ.)] doi: 1025881/18110193_2021_2_60.
7. Горбань В.И., Щеголев А.В., Проценко Д.Н. и др. Цифровизация службы анестезиологии и реаниматологии: многоцентровое анкетное исследование // Вестник интенсивной терапии им. А.И.Салтанова. — 2024. — №2. — С.43-53. [Gorban VI, Shchegolev AV, Protsenko DN, et al. Digitalization of anesthesiology and resuscitation services: multicenter questionnaire study. Annals of Critical Care. 2024; 2: 43-53. (In Russ.)] doi: 10.21320/1818-474X-2024-2-43-53.
8. Гусаров В.Г., Замятин М.Н., Гороховатский Ю.И. и др. Безопасность пациента как основа стратегии развития службы анестезиологии и реаниматологии Пироговского центра // Вестник Национального медико-хирургического Центра им. Н. И. Пирогова. — 2022. — №4. — С.4-12. [Gusarov VG, Zamyatin MN, Gorohovatskij YUI, et al. Patient safety as the foundation for the development strategy of the Department of Anesthesiology and Intensive Care of the Pirogov Center. Bulletin of Pirogov National Medical & Surgical Center. 2022; 17(4-2): 4-12. (In Russ.)] doi: 10.25881/20728255_2022_17_4_2_4.
9. Карпов О.Э., Гусаров В.Г., Замятин М.Н. и др. Интеграция цифровых решений в работу службы анестезиологии и реаниматологии многопрофильной клиники // Вестник НМХЦ им. Н.И. Пирогова. — 2020. — №15(3-2). — С.106-113. [Karpov OE, Gusarov VG, Zamyatin MN, et al. Digital Solutions Integration into the Anesthesiology Service of a Multidisciplinary Clinic. Bulletin of Pirogov National Medical & Surgical Center. 2020; 15(3-2): 106-113. (In Russ.)] doi: 10.25881/BPNMSC.2020.33.66.020.
10. Клыкков А.И., Фролов Н.С. Особенности приоритетного проекта «Совершенствование процессов организации медицинской помощи на основе внедрения информационных технологий» // Смоленский

- медицинский альманах. — 2019. — №3. — С.78-81. [Klykov AI, Frolov NS. Features of priority project «Improvement of medical care through the introduction of information technologies». Smolensk Medical Almanac. 2019; 3: 78-81. (In Russ.)]
11. Корчагин Е.Е., Гордеева Н.В., Демко И.В. и др. Использование информационных систем в здравоохранении // Сибирское медицинское обозрение. — 2019. — №3. — С.106-111. [Korchagin EE, Gordeeva NV, Demko IV, et al. Use of information systems in healthcare. Siberian Medical Review. 2019; 3: 106-111. (In Russ.)] doi: 10.20333/2500136-2019-3-106-111.
 12. Клейменова Е.Б., Яшина Л.П. Ноль медицинских информационных технологий в обеспечении безопасности пациентов // Врач и информационные технологии. — 2020. — №3. — С.13-24. [Kleymenova E B, Yashina L P. The role of health information technology in promoting patient safety. Medical doctor and information technology. 2020; 3: 13-24. (In Russ.)] doi: 10.37690/1811-0193-2020-3-13-24.
 13. Михеев А.Е. Возможности, проблемы и перспективы информационных технологий в сфере клинической безопасности. // Менеджер здравоохранения. — 2023. — №1. — С.5-20. [Mikheev AE. Opportunities, problems and prospects of information technologies in the field of clinical safety. Manager Zdravoohranenia. 2023; 1: 5-20. (In Russ.)] doi: 10.21045/1811-0185-2023-S-5-20.
 14. Кошечкин К.А. Регулирование искусственного интеллекта в медицине // Пациентоориентированная медицина и фармация. — 2023. — №1(1). — С.32-40. [Koshechkin KA. Regulation of artificial intelligence in medicine. Patient-Oriented Medicine and Pharmacy. 2023; 1(1): 32-40. (In Russ.)] doi: 10.37489/2949-1924-0005.
 15. Mirskikh I, Mingaleva Z, Kuranov V, Matseeva S. Digitization of Medicine in Russia: Mainstream Development and Potential. In: Antipova, T. (eds) Integrated Science in Digital Age 2020. ICIS 2020. Lecture Notes in Networks and Systems. 2021; 136. Springer, Cham. doi: 10.1007/978-3-030-49264-9_30.
 16. Kononova O, Prokudin D, Timofeeva A, Matrosova E. The Digital Era of Healthcare in Russia: Case Study. Lecture Notes in Information Systems and Organization, in: Zaramenskikh E, Fedorova A. (ed.), Digital Transformation and New Challenges. 2021. P.265-286. Springer.
 17. Яновская О., Кулагина Н., Логачева Н., Цифровое неравенство российских регионов // Sustainable Development and Engineering Economics 1. — 2022. — № 5. — С.77-98. [Yanovskaya O, Kulagina N, Logacheva N. Digital inequality of Russian regions. Sustainable Development and Engineering Economics 1. 2022; 5: 77-98. (In Russ.)] doi: 10.48554/SDEE.2022.1.5.
 18. Кром И.Л., Еругина М.В., Еремина М.Г. и др. Оптимизация медицинской помощи в региональном здравоохранении: перспективы и барьеры // Социология медицины. — 2023. — Т.22. — №1. — С.19-27. [Krom IL, Erugina MV, Yeremina MG, et al. Optimization of medical care in regional health care: perspectives and barriers. Sociology of Medicine. 2023; 22(1): 19-27. (In Russ.)] doi: 10.17816/socm252064.
 19. Гречко А.В., Ядгаров М.Я., Яковлев А.А. и др. Российская база данных реанимационных пациентов — RICD // Общая реаниматология. — 2024. — №20(3). — С.22-31. [Grechko AV, Yadgarov MY, Yakovlev AA, et al. RICD: Russian Intensive Care Dataset. General Reanimatology. 2024; 20(3): 22-31. (In Russ.)] doi: 10.15360/1813-9779-2024-3-22-31.
 20. Chen C, Ma S, Liao L, Xiao Y, Dai H. Effects of mesenchymal stem cells on postresuscitation renal and intestinal injuries in a porcine cardiac arrest model. Shock. 2023; 59(5): 803-809. doi: 10.1097/SHK.0000000000002107.
 21. Kakadiaris A. Evaluating the Fairness of the MIMIC-IV Dataset and a Baseline Algorithm: Application to the ICU Length of Stay Prediction. License: arXiv.org perpetual non-exclusive license arXiv:2401.00902v1 [cs.LG] 31 Dec 2023. doi: 10.48550/arXiv.2401.00902.
 22. González-Nóvoa JA, Busto L, Rodríguez-Andina JJ, et al. Using Explainable Machine Learning to Improve Intensive Care Unit Alarm Systems. Sensors. 2021; 21: 7125. doi: 10.3390/s21217125.
 23. Alabdulhafith M, Saleh H, Elmannai H, et al. A Clinical Decision Support System for Edge/Cloud ICU Readmission Model Based on Particle Swarm Optimization, Ensemble Machine Learning, and Explainable Artificial Intelligence. IEEE Access. 2023; 11: 100604–100621. doi: 10.1109/ACCESS.2023.3312343.
 24. Amico B. Explainable Temporal Data Mining Techniques to Support the Prediction Task in Medicine. S.S.D. INF/01 — Informatics. 2023. P.165.
 25. Thorsen-Meyer HC, Nielsen AB, Nielsen AP, et al. Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records. Lancet Digit Health. 2020; 2(4): e179-e191. doi: 10.1016/S2589-7500(20)30018-2.
 26. Shishkin S, Sheiman I, Vlassov V, Potapchik E, Sazhina S. Structural changes in the Russian health care system: do they match European trends? Health Econ Rev. 2022; 12(1): 29. doi: 10.1186/s13561-022-00373-z.
 27. Ye Z, An S, Gao Y, et al. Association between the triglyceride glucose index and in-hospital and 1-year mortality in patients with chronic kidney disease and coronary artery disease in the intensive care unit. Cardiovasc Diabetol. 2022; 110(2023). doi: 10.1186/s12933-023-01843-2.
 28. Xu J, Cai H, Zheng, X. Timing of vasopressin initiation and mortality in patients with septic shock: analysis of the MIMIC-III and MIMIC-IV databases. BMC Infect Dis. 2023; 199(2023). doi: 10.1186/s12879-023-08147-6.

29. Di Martino F, Delmastro F. Explainable AI for clinical and remote health applications: a survey on tabular and time series data. *Artif Intell Rev.* 2023; 56: 5261-5315. doi: 10.1007/s10462-022-10304-3.
30. Morid AM, Sheng O, Dunbar J. Time series prediction using deep learning methods in healthcare. *ACM Trans Manage Info Syst.* 2023; 14: 1-29. doi: 10.1145/3531326.
31. Idowu EAA, Teo J, Salih S, Valverde J, Yeung JA. Streams, rivers and data lakes: an introduction to understanding modern electronic healthcare records. *Clin Med (Lond).* 2023; 23(4): 409. doi: 10.7861/clinmed.2022-0325.
32. Stephany ND, Nan K, Douglas C, et al. HL7 FHIR-based tools and initiatives to support clinical research: a scoping review, *Journal of the American Medical Informatics Association.* 2022; 29(9): 1642-1653. doi: 10.1093/jamia/ocac105.
33. Dhayne H, Haque R, Kilany R, Taher Y. In Search of Big Medical Data Integration Solutions - A Comprehensive Survey. In *IEEE Access.* 2019; 7: 91265-91290. doi: 10.1109/ACCESS.2019.2927491.
34. Parciak M, Suhr M, Schmidt C, et al. FAIRness through automation: development of an automated medical data integration infrastructure for FAIR health data in a maximum care university hospital. *BMC Med Inform Decis Mak* 23. 2023; 94. doi: 10.1186/s12911-023-02195-3.
35. Bernonille S, Nies J, Pedersen HG, et al. Three different cases of exploiting decision support services for adverse drug event prevention. *Stud Health Technol Inform.* 2011; 166: 180-8.
36. Imran S, Mahmood T, Morshed A, ad Sellis T. Big data analytics in healthcare – A systematic literature review and roadmap for practical implementation, in *IEEE/CAA Journal of Automatica Sinica.* 2021; 8(1): 1-22. doi: 10.1109/JAS.2020.1003384.
37. Ciampi M, Sicuranza M, Silvestri S. A Privacy-Preserving and Standard-Based Architecture for Secondary Use of Clinical Data. *Information* 2022; 13(2): 87. doi: 10.3390/info13020087.
38. Jiang J, Hewner S, Chandola V. Explainable Deep Learning for Readmission Prediction with Tree-GloVe Embedding, 2021 IEEE 9th International Conference on Healthcare Informatics (ICHI). 09-12 August 2021. P.138-147. doi: 10.1109/ICHI52183.2021.00031.
39. Blinov P, Avetisian M, Kokh V, Umerenkov D, Tuzhilin A. Predicting Clinical Diagnosis from Patients Electronic Health Records Using BERT-Based Neural Networks. In: Michalowski M, Moskovitch R. (eds) *Artificial Intelligence in Medicine. AIME 2020. Lecture Notes in Computer Science, 2020;* 12299. Springer. Cham. doi: 10.1007/978-3-030-59137-3_11.
40. Kontsevaya A, Bobrova N, Barbarash O, et al. The management of acute myocardial infarction in the Russian Federation: protocol for a study of patient pathways. *Wellcome Open Res.* 2017; 2: 89. doi: 10.12688/wellcomeopenres.12478.2.
41. Brankovic A, Hassanzadeh H, Good N, et al. Explainable machine learning for real-time deterioration alert prediction to guide pre-emptive treatment. *Sci Rep.* 2022; 12: 11734. doi: 10.1038/s41598-022-15877-1.
42. Pang K, Li L, Ouyang W, Liu X, Tang Y. Establishment of ICU Mortality Risk Prediction Models with Machine Learning Algorithm Using MIMIC-IV Database. *Diagnostics.* 2022; 12(5): 1068. doi: 10.3390/diagnostics12051068.
43. Sun Y, He Z, Ren, J, et al. Prediction model of in-hospital mortality in intensive care unit patients with cardiac arrest: a retrospective analysis of MIMIC-IV database based on machine learning. *BMC Anesthesiol.* 2023; 23: 178. doi: 10.1186/s12871-023-02138-5.
44. Liu W, Tao G, Zhang Y, et al. A Simple Weaning Model Based on Interpretable Machine Learning Algorithm for Patients With Sepsis: A Research of MIMIC-IV and eICU Databases. *Front. Med.* 2022; 8: 814566. doi: 10.3389/fmed.2021.814566.
45. Kovalchuk SV, Funkner AA, Metsker OG, Yakovlev AN. Simulation of patient flow in multiple healthcare units using process and data mining techniques for model identification. *Journal of biomedical informatics,* 2018; 82: 128-142. doi: 10.1016/j.jbi.2018.05.004.
46. Cao Y, Li Y, Wang M, et al. Interpretable machine learning for predicting risk of invasive fungal infection in critically ill patients in the intensive care unit: a retrospective cohort study based on MIMIC-IV database. *Shock.* 2024; 61(6): 817-827. doi: 10.1097/SHK.0000000000002312.
47. Sun Y, He Z, Ren J, et al. Prediction model of in-hospital mortality in intensive care unit patients with cardiac arrest: a retrospective analysis of MIMIC -IV database based on machine learning. *BMC Anesthesiol.* 2023; 23: 178. doi: 10.1186/s12871-023-02138-5.
48. Pang K, Liang L, Wen O, Xing L, Yongzhong T. Establishment of ICU Mortality Risk Prediction Models with Machine Learning Algorithm Using MIMIC-IV Database. *Diagnostics,* 2022; 12(5): 1068. doi: 10.3390/diagnostics12051068.
49. Shu T, Huang J, Deng J, et al. Development and assessment of scoring model for ICU stay and mortality prediction after emergency admissions in ischemic heart disease: a retrospective study of MIMIC-IV databases. *Intern Emerg Med.* 2023; 18: 487-497. doi: 10.1007/s11739-023-03199-7.
50. Johnson AEW, Bulgarelli L, Shen L, et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data.* 2023; 10(1). doi: 10.1038/s41597-022-01899-x.
51. Medical Information Mart for Intensive Care. MIMIC Online Documentation: <https://mimic.mit.edu>.