

## Российская база данных реанимационных пациентов — RICD

А. В. Гречко, М. Я. Ядгаров\*, А. А. Яковлев, Л. Б. Берикашвили,  
А. Н. Кузовлев, П. А. Поляков, И. В. Кузнецов, В. В. Лихванцев

Федеральный научно-клинический центр реаниматологии и реабилитологии,  
Россия, 107031, г. Москва, ул. Петровка, д. 25, стр. 2

**Для цитирования:** А. В. Гречко, М. Я. Ядгаров, А. А. Яковлев, Л. Б. Берикашвили, А. Н. Кузовлев, П. А. Поляков, И. В. Кузнецов, В. В. Лихванцев. Российская база данных реанимационных пациентов — RICD. *Общая реаниматология*. 2024; 20 (3): 22–31. <https://doi.org/10.15360/1813-9779-2024-3-22-31> [На русск. и англ.]

\*Адрес для корреспонденции: Михаил Яковлевич Ядгаров, [myadgarov@fnkcr.ru](mailto:myadgarov@fnkcr.ru)

### Резюме

В эпоху цифровизации здравоохранения научное сообщество сталкивается с потребностью в структурированных и доступных базах данных для научных исследований и технологических проектов в области искусственного интеллекта, связанных с разработкой новых методов диагностики и лечения.

**Цель проекта:** разработать базу данных, содержащую анонимизированные медицинские данные всех пациентов, проходивших лечение в Федеральном научно-клиническом центре реаниматологии и реабилитологии (ФНКЦ РР), и предоставить врачам и научным сотрудникам ФНКЦ РР и других центров возможность доступа к структурированным данным пациентов для последующего анализа и проведения исследований.

**Материалы и методы.** Для сбора и представления данных использовали медицинскую информационную систему ФНКЦ РР и инструменты «Асклепиус», PL/SQL, Microsoft Office Excel, Power Query M, Microsoft PowerBI, Open data editor и Python. С целью предоставления возможности открытого доступа к базе данных и защиты персональных данных пациентов их анонимизировали.

**Результаты.** Представили проект RICD (Russian Intensive Care Dataset, <https://fnkcr-database.ru/>) — первую в Российской Федерации базу данных реанимационных пациентов, разработанную в ФНКЦ РР на основе передовых принципов и методов, используемых в международных проектах открытых баз данных — «eICU Program» от Philips Healthcare, «MIMIC-IV» и «MIMIC-III». Созданная база данных представляет информацию о 7730 госпитализациях 5115 пациентов (с учетом повторных госпитализаций), включая данные о 3291 госпитализации в отделения реанимации и интенсивной терапии. Общее число записей в RICD превышает 14 млн. RICD представляет медико-антропометрические данные, информацию о движении пациентов внутри учреждения, диагнозы, сведения о проводимой терапии, результаты лабораторных исследований, оценки по шкалам и исходы госпитализации. RICD также содержит данные ряда витальных параметров, собираемых с прикроватных мониторов и иного оборудования отделений реанимации и интенсивной терапии, с частотой до 10 оценок в час.

**Заключение.** RICD позволяет проводить глубокий анализ и исследования клинических практик в сфере анестезиологии-реаниматологии и интенсивной терапии, а также разрабатывать инструменты поддержки принятия клинических решений и применять методы машинного обучения для решения задач диагностики и улучшения исходов лечения пациентов. Благодаря доступности и детальной структуризации данных, база станет полезным инструментом решения как научных, так и практических задач в области интенсивной терапии.

**Ключевые слова:** база данных; реанимационные пациенты; интенсивная терапия; искусственный интеллект; машинное обучение; системы поддержки принятия клинических решений; <https://fnkcr-database.ru/>

**Конфликт интересов.** Авторы заявляют о потенциальном конфликте интересов, в случае коммерциализации представленной в данной статье базы данных.

### RICD: Russian Intensive Care Dataset

Andrey V. Grechko, Mikhail Y. Yadgarov\*, Alexey A. Yakovlev, Levan B. Berikashvili,  
Artem N. Kuzovlev, Petr A. Polyakov, Ivan V. Kuznetsov, Valery V. Likhvantsev

Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology,  
25 Petrovka Str., Bldg. 2, 107031 Moscow, Russia

### Summary

In the era of healthcare digital transformation, the scientific community faces the need for structured and available datasets for research and technological projects in the field of artificial intelligence, related to the development of new diagnostic and treatment methods.

**Objective:** to develop a dataset containing anonymized medical data of all patients treated at the Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology (FRCCR), and provide access for doctors and scientists of FRCCR and other centers to structured patient data for subsequent analysis and research.

**Materials and Methods.** The FRCCR medical information system and the tools «Asclepius», PL/SQL, Microsoft Office Excel, Power Query M, Microsoft PowerBI, Open data editor, and Python were used for data collection and representation. To provide open access to the dataset and protect the personal data of patients, the information was anonymized.

**Results.** We introduce the RICD (Russian Intensive Care Dataset, <https://fnkrr-database.ru/>) — the first dataset of intensive care patients in the Russian Federation, developed at FRCCR based on advanced principles and methods used in international open database projects — «eICU Program» from Philips Healthcare, «MIMIC-IV», and «MIMIC-III». The developed dataset contains information on 7,730 hospitalizations of 5,115 patients (including readmissions), covering data from 3,291 hospitalizations in the intensive care units (ICUs). The total number of records in the RICD exceeds 14 million. The RICD presents medical-anthropometric data, patient movement within the institution, diagnoses, information on therapy provided, results of laboratory tests, scale assessments, and outcomes of hospitalization. RICD also contains data on several vital parameters collected from bedside monitors and other equipment of ICUs, with up to 10 evaluations per hour.

**Conclusion.** The RICD allows for in-depth analysis and research of clinical practices in intensive care, enabling the development of clinical decision support tools and the application of machine learning methods to enhance diagnostic tools and improve patient outcomes. With its accessibility and detailed data structure, the dataset serves as a valuable tool for both scientific research and practical applications in intensive care.

**Keywords:** *dataset; critically ill patients; intensive care; artificial intelligence; machine learning; clinical decision support systems; <https://fnkrr-database.ru/>*

**Conflict of interest.** The authors declare a potential conflict of interest in case of commercialization of the dataset presented in this article.

Read the full-text English version at [www.reanimatology.com](http://www.reanimatology.com)

## Введение

В современной медицине наблюдается бурное развитие электронных систем здравоохранения, благодаря которым данные, собранные в ходе рутинной клинической практики, систематизируются и хранятся в медицинских учреждениях по всему миру. Особое внимание развитию подобных систем уделяется в области интенсивной терапии в связи с необходимостью непрерывного контроля витальных функций пациентов отделений реанимации и интенсивной терапии (ОРИТ). Такой подход приводит к формированию значительных объемов данных, позволяющих при детальном изучении совершенствовать клиническую практику. Проекты «MIMIC-III» [1] (база данных, представляющая данные более 40000 пациентов, поступивших в ОРИТ медицинского центра Beth Israel Deaconess, США за 2001–2012 гг.) и «MIMIC-IV» (данные госпитализаций 69653 пациентов из медицинского центра Beth Israel Deaconess, США за 2008–2019 гг.) [2], а также «eICU» (многоцентровая база данных, включающая деперсонализированные медицинские данные более 200000 случаев поступления в ОРИТ в десятках медицинских центров США в период 2014–2015 гг.) [3] демонстрируют, как ценная информация, собранная в ОРИТ, может быть использована для научных исследований и разработки новых методов диагностики и лечения. На 2024 г. опубликовано свыше 1500 научных работ, в которых использованы данные проектов «MIMIC-III», «MIMIC-IV» и «eICU», в т. ч. в таких журналах, как *Critical Care* [4], *Nature* [5], *Lancet* [6], *JAMA* [7], *BMC Anesthesiology* [8] и др.

Важный вклад в развитие клинических исследований внесли и другие международные

проекты, такие как «HiRID» [9] (36098 госпитализаций в ОРИТ, Швейцария), «AmsterdamUMCdb» [10] (23106 госпитализаций в ОРИТ, Нидерланды), а также база данных «The Children's Hospital at Zhejiang University School of Medicine» [11] (13941 госпитализация в детское ОРИТ, Китай). Все эти репозитории представляют уникальные данные о пациентах, их лечении и клинических исходах, однако следует отметить, что ввиду значительного отличия систем здравоохранения разных стран, использование зарубежных проектов открытых баз данных в отечественной практике ограничено. В 2023 г. ФГБОУ ВО СибГМУ Минздрава России представил первый репозиторий клинических данных в Российской Федерации, «SibMED Data Clinical Repository», включающий анонимизированные данные более чем 20000 госпитализаций в 10 многопрофильных клиниках СибГМУ [12]. К настоящему моменту (2024 г.) в данном репозитории не представлены мониторируемые данные реанимационных пациентов, что не позволяет проводить глубокие исследования в области интенсивной терапии.

В настоящее время интеграция технологий искусственного интеллекта (ИИ) в медицинскую практику открывает новые возможности для диагностики, лечения и мониторинга состояния пациентов, что особенно актуально для пациентов реанимационного профиля. Технологии ИИ охватывают широкий спектр методов и подходов, включая нейронные сети, методы машинного обучения и экспертные системы. Эти инструменты обладают уникальной способностью анализировать сложные и объемные

наборы данных, включая медицинские изображения, текстовые документы (такие как дневниковые записи, выписные и этапные эпикризы, протоколы и заключения) и базы данных медицинских параметров [13]. Одним из наиболее перспективных направлений использования ИИ в области интенсивной терапии является мониторинг и анализ состояния пациентов в реальном времени. С использованием технологий ИИ возможно прогнозирование развития осложнений и неблагоприятных исходов. Системы, использующие технологии ИИ, могут выявлять опасные состояния, требующие внимания медицинского персонала, тем самым являясь важнейшими элементами систем поддержки принятия клинических решений, позволяющих уменьшить число врачебных ошибок и повысить качество оказываемой медицинской помощи [14, 15]. Для развития подобных систем требуется разработка крупных структурированных и доступных баз данных.

В настоящее время в Федеральном научно-клиническом центре реаниматологии и реабилитологии (ФНКЦ РР) реализована концепция «цифровой» реанимации, заключающаяся в том, что все данные реанимационных пациентов, включая непрерывно мониторируемые параметры, собираются, хранятся на серверах организации и являются доступными для анализа.

Цель проекта — разработать базу данных, содержащую анонимизированные медицинские данные всех пациентов, проходивших лечение в ФНКЦ РР, и предоставить врачам и научным сотрудникам ФНКЦ РР и других центров возможность доступа к структурированным данным пациентов для последующего анализа и проведения исследований.

### Материал и методы

**Получение данных.** Для получения данных использовали стандартные инструменты медицинской информационной системы (МИС) ФНКЦ РР, построенной на базе программного обеспечения и реляционной системы управления базами данных (СУБД) «Асклепиус» (Oracle 11g). Основной массив данных выгружали из различных модулей МИС (регистратуры, клинического, лабораторного, аптечного и др.) с использованием виртуального модуля «Конструктор запросов». Дополнительные запросы выполняли на языке PL/SQL непосредственно из СУБД. Период выгрузки данных: декабрь 2017 г. — июль 2023 г. Часть данных по результатам лабораторных исследований получили путем создания отчета в лабораторной информационной системе «Алиса». Для структурирования данных создавали макросы на языке «VBA» в Microsoft Office Excel 2021 и использовали язык программирования «Power Query M». Для демонстрации аналитических возможностей базы данных применяли BI систему «Microsoft PowerBI», язык программирования «Python» в среде «PyCharm 2023.2.4», и IBM SPSS Statistics 27.0. Для стандартизации создавали справочники, кото-

рые позволяли унифицировать и стандартизовать данные из различных источников. Итоговые файлы базы данных создавали с использованием технологии «Frictionless data», реализованной в программе «Open data editor» [16].

**Анонимизация и псевдонимизация.** Все данные деперсонализировали (анонимизировали) с целью предоставления открытого доступа. Методика анонимизации данных разработана на основе Национального стандарта Российской Федерации «Информатизация здоровья. Псевдонимизация» ГОСТ Р 55036-2012/ISO/TS 25237:2008 (утв. приказом Федерального агентства по техническому регулированию и метрологии от 29 октября 2012 г. N 585-ст) и международных проектов открытых баз данных. Методика анонимизации была утверждена локальным Этическим комитетом ФНКЦ РР (№4/23/2 от 20.12.2023 г.). Удалили прямые (явные) и косвенные идентификаторы пациентов. Для сохранения информативности базы данных в каждой таблице создали новые параметры, в которых отображали информацию о времени исследования/осмотра или оценки параметра (`post_admission_days` и `post_admission_hours`). Дополнительно провели процедуру псевдонимизации с заменой идентифицирующих данных (ID пациента, ID госпитализации/№ истории болезни) псевдонимами (уникальными идентификаторами: `new_patient_id`, `new_hosp_id`), которые не могут быть связаны с исходными данными без дополнительной информации, хранящейся отдельно.

### Описание базы данных

**Описательные характеристики.** RICD представляет данные о 7730 госпитализациях 5115 пациентов поступивших в ФНКЦ РР в период 2017–2023 гг., включая 3291 госпитализации в ОРИТ (табл. 1). Средний возраст реанимационных пациентов: 57,8 лет (SD 17,7, диапазон от 19 до 97 лет), медиана длительности госпитализации в ОРИТ составила 32 дня (22–50), максимум — 320 суток. За весь период регистрировали 405 случаев летального исхода. Общее число записей во всех таблицах RICD составило 14 356 553.

**Формат данных.** База данных представили в нескольких форматах:

1. файлы в формате .csv, соединенные по ключевым полям (frictionless data);
2. формат .db (для SQLite);
3. формат .pbix (для системы Microsoft PowerBI).

Дополнительно представили метаданные в формате .json.

Файлы RICD могут быть загружены в любую реляционную базу данных или BI-систему.

**Структура базы данных.** Созданная база данных состоит из 16 связанных между собой таблиц (подробное описание представлено в Приложении). Объединяющие идентификаторы всех таблиц: `new_patient_id` (модифицированный ID пациента), `new_hosp_id` (модифицированный ID истории болезни).

**Таблица 1. Описательная характеристика пациентов RICD госпитального и реанимационного профилей.**

Характеристика	Стационар	Отделения реанимации
Количество госпитализаций	4339	3291
Количество пациентов*	3033	2562
Возраст, среднее (SD)	59,2 (15,2)	57,8 (17,7)
Пол, М (%)	2385 (55,0)	1864 (56,6)
Длительность госп., сут (Q1–Q3)	14,0 (14,0–18,0)	32,0 (22,0–50,0)

**Примечание.**\* — количество пациентов превышает 5115, т. к. часть пациентов госпитализировались в разные периоды времени и в отделение реанимации, и в стационар.

**Таблица 2. Краткое описание таблиц RICD.**

Таблица	Число записей	Описание
1. all_patients	7730	Медико-антропометрические характеристики пациентов, маршрут движения, исходы госпитализации.
2. ICD10_diagnoses	1058929	Диагнозы пациентов согласно классификации МКБ-10.
3. therapy_prescriptions	550880	Терапевтические назначения.
4. clinical_notes	304680	Регулярная оценка состояния пациента и диагноза МКБ-10.
5. monitoring_data	12198188	Мониторимые параметры (витальные характеристики и параметры жидкостного баланса).
6. all_scales	19297	Оценки по шкалам.
7. detailed_sofa	14859	Подробная структура оценки по шкале SOFA.
8. complete_blood_count	43584	Результаты клинического анализа крови.
9. urinalysis	37307	Результаты клинического анализа мочи.
10. blood_biochemistry	65882	Результаты биохимического анализа крови.
11. urine_biochemistry	1389	Результаты биохимического анализа мочи.
12. coagulation_profile	33727	Результаты оценки системы гемостаза.
13. acid_base_balance	6185	Результаты оценки кислотно-основного баланса и газов крови.
14. antibiotic_resistance	6794	Результаты оценки антибиотикорезистентности.
15. bacteria_culture_test	6101	Результаты культуральных исследований (посевы).
16. cerebrospinal_fluid_analysis	1021	Ликворограмма.

В таблицах представили медико-антропометрические данные пациентов, информацию об их движении внутри учреждения, диагнозы, данные о проводимой терапии, результаты лабораторных исследований, оценки по шкалам, витальные и жидкостные параметры в динамике и исходы госпитализации (табл. 2).

В RICD 85% всех записей представили оценками мониторируемых витальных показателей реанимационных пациентов (результаты пульсоксиметрии (SpO<sub>2</sub>), температура тела, частота дыхания, частота сердечных сокращений, систолическое и диастолическое АД, среднее АД, центральное венозное давление) и параметров жидкостного баланса.

#### Ключевые особенности RICD.

1. Наличие мониторируемых данных реанимационных пациентов. RICD является первой в Российской Федерации базой, в которой представлены данные, собираемые с прикроватных мониторов и иного оборудования отделений реанимации и интенсивной терапии. Частота дискретизации (оценки данных) составляет до 10 оценок в час.

2. Основная когорта пациентов представлена пациентами, длительно пребывающими в ОРИТ. Более 60% пациентов имели сниженный уровень сознания (оценка по шкале комы Глазго < 15 баллов).

3. Наличие данных культуральных исследований (посевы) с оценкой антибиотикорезистентности.

**Техническая валидация.** Для оценки разработанной БД привлекали междисциплинарную команду сотрудников ФНКЦ РР, которая проводила валидацию данных, а именно — оценивала целостность данных путем перекрестной проверки таблиц, определяла допустимые диапазоны значений данных, а также определяла полноту анонимизации данных.

**Платформа доступа.** С целью предоставления доступа к проекту RICD разработали платформу (веб-сайт: <https://fnkcr-database.ru>). На сайте представили полную информацию о проекте RICD, описали структуру базы данных и предоставили возможность получения демо-версии базы данных, представляющей данные госпитализации 10 пациентов (более 60000 оценок мониторируемых параметров в ОРИТ). Дополнительно разместили материалы по работе с биомедицинскими данными RICD, которые могут быть использованы для изучения базы данных и ознакомления с возможными инструментами ее анализа.

При подаче заявки предоставляется возможность выбрать опцию «участие в проекте RICD», что создает предпосылки для интеграции и объединения баз данных различных медицинских центров.

**Демонстрация аналитических возможностей.** На рисунке представили параметры пациентки, мониторируемые в течение 15 сут госпитализации в ФНКЦ РР. В указанный период имеется свыше 1000 оценок по каждому из витальных

параметров (в т. ч. SpO<sub>2</sub>, частоте дыхания и ЧСС), динамика изменения лабораторных параметров и оценок по клиническим шкалам. Также RICD предоставляет возможность оценивать проведение ИВЛ и отслеживать медикаментозные назначения, в том числе применение вазопрессорных и инотропных препаратов.

### Заключение

Представленная RICD позволяет проводить глубокий анализ и исследования клинических практик в области интенсивной терапии, разрабатывать инструменты поддержки принятия клинических решений и применять методы машинного обучения для решения задач диагностики и улучшения исходов лечения пациентов. Благодаря доступности и детальной структуризации данных, база станет полезным инструментом как для научных, так и для практических задач.

Проект RICD рекомендуется к использованию научным исследователям, специалистам в области data science и машинного обучения и разработчикам передовых решений в области цифрового здравоохранения.

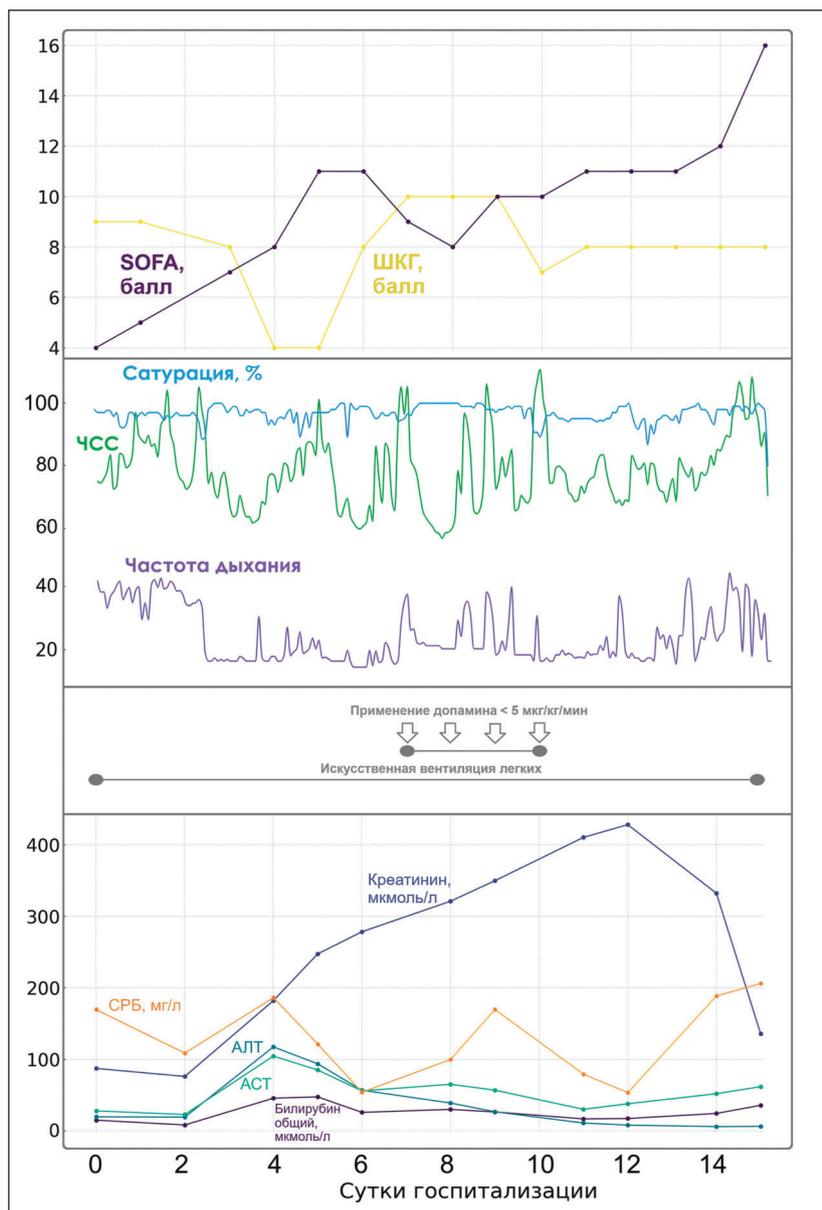


Рис. Параметры, мониторируемые в течение 15 суток госпитализации пациентки в ФНКЦ РР.

Примечание. Использовали данные таблиц all\_scales, monitoring\_data, blood\_biochemistry, и detailed\_sofa.

## Приложение. Подробное описание таблиц RICD

### 1. all\_patients (7,730 записей)

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
sex	Пол пациента	—
body_weight	Масса тела пациента на момент поступления	кг
height	Рост пациента на момент поступления	см
BMI	Индекс массы тела при поступлении	кг/м <sup>2</sup>
age	Возраст пациента на момент госпитализации	лет
patient_condition	Состояние пациента при поступлении	—
transfer	Факт перевода из другого учреждения	—
adm_year	Год поступления	—
admission_department	Отделение поступления	—
discharge_department	Отделение выбытия	—
ICU_patients	Факт прохождения через отделение реанимации	—
length_of_stay	Длительность госпитализации	дни
fatal_outcome	Летальный исход	—

**2. ICD10\_diagnoses (1,058,929 записей)**

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
document_type	Тип документа	—
diagnosis_type	Тип диагноза	—
ICD_10	Код МКБ-10	—

**3. therapy\_prescriptions (550,880 записей)**

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
prescription_rus	Назначение на русском языке	—
prescription_eng	Назначение на английском языке	—
Time (12:00_am-11:00_pm)	Время оценки статуса выполнения/назначения	V назначено + выполнено X отменено — не выполнено

**4. clinical\_notes (304,680 записей)**

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
patient_condition	Состояние пациента на момент оценки	—
ICD_10	Коды МКБ-10	—

**5. monitoring\_data (12,198,188 записей)**

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
parameter	Мониторируемый параметр*	—
unit	Единицы измерения	—
hour	Время оценки	—
value	Значение параметра	—

\***Витальные параметры:** saturation (SpO<sub>2</sub>), temperature, respiratory rate, heart rate, diastolic BP, systolic BP, mean AP, body mass, glucose, respiratory volume, central venous pressure, BIS, EtCO<sub>2</sub>. \***Жидкостные параметры:** fluid intake *per os*, diuresis, enteral feeding, infusion, liquor, stool/stoma, other intake, other output, gastrostoma, drainages volume, nasogastric tube/vomitus, cystostomy.

**6. all\_scales (19,297 записей)**

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
scales	Шкалы* и результаты оценки	—

\***Шкалы:** CRS-R, APACHE II, CHA<sub>2</sub>DS<sub>2</sub>-VASc, DRS, FIM, FOUR, GRACE, HAS-BLED, NIHSS, POSSUM, SAPS II, SOFA, Barthel, Waterlow (pressure ulcer risk), Geneva score, Glasgow coma scale, Classification of surgical and anesthesia risk (MNOAR), Ashworth scale, Caprini (DVT/PE risk), Palliative performance scale, Rehabilitation routing scale, Rivermead mobility index, Modified Rankin scale, Wells' Criteria for DVT.

**7. detailed\_sofa (14,859 записей)**

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
report	Текстовое заключение	—
sofa_score	Оценка по шкале SOFA	—
FiO <sub>2</sub> O <sub>2</sub> %	FiO <sub>2</sub> %	%
PaO <sub>2</sub> O <sub>2</sub>	PaO <sub>2</sub>	мм рт. ст.
PaO <sub>2</sub> O <sub>2</sub> /FiO <sub>2</sub> O <sub>2</sub>	PaO <sub>2</sub> /FiO <sub>2</sub>	—
mechanical_ventilation	Искусственная вентиляция легких	—
platelets	Тромбоциты	10 <sup>9</sup> /л
GSC_score	Оценка по шкале комы Глазго	—
bilirubin	Билирубин	мкмоль/л
systolic_AP	Систолическое артериальное давление	мм рт. ст.
diastolic_AP	Диастолическое артериальное давление	мм рт. ст.
mean_AP	Среднее артериальное давление	мм рт. ст.
vasoactive_drugs	Применение вазоактивных препаратов	—
creatinine	Креатинин	мкмоль/л
daily_diuresis	Суточный диурез	мл

**8. complete\_blood\_count (43,584 записей)**

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
WBC	Лейкоциты	10 <sup>9</sup> /л
RBC	Эритроциты	10 <sup>12</sup> /л
HGB	Гемоглобин	г/л
HCT	Гематокрит	%
MCV	Средний объем эритроцита	фл
MCH	Среднее содержание гемоглобина в эритроците	пг
MCHC	Средняя концентрация гемоглобина в эритроците	г/л
RDW	Ширина распределения эритроцитов по объему	%
RDW-SD	Стандартное отклонение ширины распределения эритроцитов по объему	фл
PLT	Тромбоциты	10 <sup>9</sup> /л
PCT	Тромбокрит	%
MPV	Средний объем тромбоцита	фл
PDW	Ширина распределения тромбоцитов по объему	%
NEU%	Процент нейтрофилов	%
NEU	Абсолютное количество нейтрофилов	10 <sup>9</sup> /л
LYM%	Процент лимфоцитов	%
LYM	Абсолютное количество лимфоцитов	10 <sup>9</sup> /л
MONO%	Процент моноцитов	%
MONO	Абсолютное количество моноцитов	10 <sup>9</sup> /л
EOS%	Процент эозинофилов	%
EOS	Абсолютное количество эозинофилов	10 <sup>9</sup> /л
BASO%	Процент базофилов	%
BASO	Абсолютное количество базофилов	10 <sup>9</sup> /л
NRBC%	Процент нормобластов	%
NRBC#	Абсолютное количество нормобластов	10 <sup>9</sup> /л
%RETIC	Процент ретикулоцитов	%
RETIC	Абсолютное количество ретикулоцитов	10 <sup>9</sup> /л
IRF	Фракция незрелых ретикулоцитов	%
MRV	Средний объем ретикулоцитов	фл
band_neutrophil	Палочкоядерные нейтрофилы	%
segmented_neutrophil	Сегментоядерные нейтрофилы	%
ESR	Скорость оседания эритроцитов	mm/h

**9. urinalysis (37,307 записей)**

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
urine_color	Цвет мочи	—
clarity/turbidity	Прозрачность мочи	—
specific_gravity	Относительная плотность	отн. ед.
PH	pH мочи	—
protein	Белок	ммоль/л
glucose	Глюкоза	ммоль/л
nitrites	Нитриты	—
ketones	Кетоны	ммоль/л
bilirubin	Билирубин	мкмоль/л
ascorbic_acid	Аскорбиновая кислота	ммоль/л
urobilinogen	Уробилиноген	мкмоль/л
squamous_epithelial_cells	Плоский эпителий	—
transitional_epithelial_cells	Переходный эпителий	—
WBCs	Лейкоциты	—
RBCs	Эритроциты	—
bacteria	Бактерии	—
hyaline_casts	Гиалиновые цилиндры	—
nonclassified_casts	Неклассифицированные цилиндры	—
mucus	Слизь	—

**10. blood\_biochemistry (65,882 записей)**

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
HbA1C_NGSP	Гликированный гемоглобин по методу NGSP	%
troponin	Тропонин (I, T)	пг/мл
total_bilirubin	Общий билирубин	мкмоль/л
direct_bilirubin	Прямой билирубин	мкмоль/л
total_protein	Общий белок	г/л
albumin	Альбумин	г/л
prealbumin	Преальбумин	г/л
urea	Мочевина	ммоль/л
creatinine	Креатинин	мкмоль/л
glucose	Глюкоза	ммоль/л
triglycerides	Триглицериды	ммоль/л
cholesterol	Холестерин	ммоль/л
HDL	Липопротеины высокой плотности	ммоль/л
LDL	Липопротеины низкой плотности	ммоль/л
atherogenic_coefficient	Коэффициент атерогенности	—
VLDL	Липопротеины очень низкой плотности	ммоль/л
magnesium	Магний	ммоль/л
calcium	Кальций	ммоль/л
phosphorus	Фосфор	ммоль/л
iron	Железо	мкмоль/л
latent_iron_binding_capacity	Латентная железосвязывающая способность сыворотки	мкмоль/л
transferrin	Трансферрин	мг/дл
potassium	Калий	ммоль/л
sodium	Натрий	ммоль/л
procalcitonin	Прокальцитонин	нг/мл
chlorides	Хлориды	ммоль/л
LDH	Лактатдегидрогеназа	Ед/л
ALT	Аланинаминотрансфераза	Ед/л
AST	Аспаратаминотрансфераза	Ед/л
GGT	Гамма-глутамилтрансфераза	Ед/л
alkaline_phosphatase	Щелочная фосфатаза	Ед/л
amylase	Амилаза	Ед/л
CRP	С-реактивный белок	мг/л
uric_acid	Мочевая кислота	мкмоль/л
rheumatoid_factor	Ревматоидный фактор	Ед/мл
cholinesterase	Холинэстераза	Ед/л
creatine_kinase	Креатинкиназа	Ед/л
anti_streptolysin_O	Антистрептолизин О	Ед/мл
HBDH	Гидроксibuтиратдегидрогеназа	Ед/л
CK_MB	Креатинкиназа MB	Ед/л
APO_A1	Аполипопротеин А1	г/л
APO_B	Аполипопротеин В	г/л
APO_B_APO_A1_ratio	Соотношение АпоВ/АпоА1	—
ferritin	Ферритин	мкг/л

**11. urine\_biochemistry (1,389 записей)**

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
Ca+24	Кальций в суточной моче	ммоль/сутки
Cl+24	Хлориды в суточной моче	ммоль/сутки
CREAT24	Креатинин в суточной моче	ммоль/сутки
GLU24	Глюкоза в суточной моче	г/сутки
K+24	Калий в суточной моче	ммоль/сутки
Mg24	Магний в суточной моче	ммоль/сутки
Na+24	Натрий в суточной моче	ммоль/сутки
urine_albumin_24h	Альбумин в суточной моче	мгг/сутки
24h_urine_urea	Мочевина в суточной моче	ммоль/сутки
urea_in_urine	Мочевина в моче	ммоль/сутки
nitrogen_loss_per_day	Потеря азота в день	г/сутки
phosphorus24	Фосфор в суточной моче	ммоль/сутки
urine_volume	Объем мочи за 24 часа	л/сутки
urine_glucose	Глюкоза в моче	ммоль/л
microalbumin	Микроальбумин в моче	мг/л
microalbumin24	Микроальбумин в суточной моче	ммоль/сутки
total_protein_urine	Общий белок в моче	г/сутки
PROT24	Белок в суточной моче	мг/сутки
urine_calcium	Кальций в моче	ммоль/л
urine_magnesium	Магний в моче	ммоль/л
uric_acid_urine	Мочевая кислота в моче	мг/л
urine_sodium	Натрий в моче	ммоль/л
urine_phosphorus	Фосфор в моче	ммоль/л
urine_chlorides	Хлориды в моче	ммоль/л
urine_potassium	Калий в моче	ммоль/л
urine_amylase	Амилаза в моче	Ед/л
urine_creatinine	Креатинин в моче	мг/л

**12. coagulation\_profile (33,727 записей)**

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
anti-Xa	Анти-Xa	Ед/мл
PT_seconds	Протромбиновое время	сек
quick_test_prothrombin_time	Протромбиновое время по Квику	%
prothrombin_ratio	Протромбиновое отношение	—
international_normalized_ratio	Международное нормализованное отношение	—
fibrinogen_calculated	Рассчитанный фибриноген	г/л
fibrinogen_Clauss	Фибриноген по Клауссу	г/л
activated_partial_thromboplastin_time	Активированное частичное тромбопластиновое время	сек
thrombin_time	Тромбиновое время	сек
D-dimer	Д-димер	мг/л

**13. acid\_base\_balance (6,185 записей)**

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
A-aDO <sub>2</sub>	Альвеоларно-артериальный градиент по кислороду	мм рт. ст.
temperature_corrected_pH	pH венозный, скорректированный по температуре	—
temperature_corrected_pO <sub>2</sub>	pO <sub>2</sub> венозный, скорректированный по температуре	мм рт. ст.
respiratory_index	Респираторный индекс	—
pH	pH венозный	—
pH_arterial	Артериальный pH	—
pCO <sub>2</sub>	Венозный pCO <sub>2</sub>	мм рт. ст.
pCO <sub>2</sub> _arterial	Артериальный pCO <sub>2</sub>	мм рт. ст.
pCO <sub>2</sub> _capillary	Капиллярный pCO <sub>2</sub>	мм рт. ст.
temperature_corrected_pCO <sub>2</sub>	pCO <sub>2</sub> венозный, скорректированный по температуре	мм рт. ст.
pO <sub>2</sub>	pO <sub>2</sub> венозный	мм рт. ст.
pO <sub>2</sub> _arterial	Артериальный pO <sub>2</sub>	мм рт. ст.
Na <sup>+</sup>	Натрий	ммоль/л
K <sup>+</sup>	Калий	ммоль/л
Ca <sup>++</sup>	Кальций	ммоль/л
Glu	Глюкоза	ммоль/л
Hct	Гематокрит	%
Lac	Лактат	ммоль/л
total_hemoglobin	Общий гемоглобин	г/л
sO <sub>2</sub> _arterial	Артериальное насыщение кислородом	%
BE(B)	Базовый избыток (кровь венозная)	мм рт. ст.
Beecf	Базовый избыток (экстрацеллюлярная жидкость)	ммоль/л
paO <sub>2</sub> /pAO <sub>2</sub>	Соотношение артериального парциального давления кислорода к альвеоларному	—
%FiO <sub>2</sub>	%FiO <sub>2</sub>	%
HCO <sub>3</sub> <sup>-</sup> _std	Стандартное содержание бикарбоната	—
HCO <sub>3</sub> <sup>-</sup>	Бикарбонат	ммоль/л
P/F Ratio	P/F Ratio	—
temp	Температура	С

**14. antibiotic\_resistance (6,794 записей)**

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
microorganism	Микроорганизм (возбудитель)	—
biological_material	Биологический материал	—
antibiotics (MIC)*	Оценка минимальной ингибирующей концентрации для каждого антибиотика	—
antibiotics (RSI)*	Оценка антибиотикорезистентности в категориях R, S и I	R (Resistant): Устойчивый. S (Sensitive): Чувствительный. I (Intermediate): Сомнительно.

\***АНТИБИОТИКИ:** Aztreonam, Amikacin, Amoxicillin/Clavulanate (f), Ampicillin, Gentamicin, Imipenem, Colistin, Meropenem, Nitrofurantoin, Norfloxacin, Piperacillin/Tazobactam, Trimethoprim/Sulfamethoxazole, Fosfomycin with Glucose-6-Phosphate, Cefepime, Cefixime, Ceftazidime, Ceftriaxone, Ciprofloxacin, Ertapenem, Amoxicillin/Clavulanate, Vancomycin, Gentamicin-syn, Daptomycin, Clindamycin, Levofloxacin, Linezolid, Oxacillin, Penicillin G, Rifampin, Streptomycin-synergism, Teicoplanin, Tetracycline, Tigecycline, Tobramycin, Fusidic Acid, Quinupristin/Dalfopristin, Cefoxitin, Erythromycin, Netilmicin, Piperacillin, Cefuroxime, Cefoperazone, Cefotaxime, Amoxicillin, Moxifloxacin, Pristinamycin, Chloramphenicol, BMS-284756, Cefazolin, High-activity Mupirocin, Doxycycline, Mupirocin, Trimethoprim, Clarithromycin, Cefoperazone/Sulbactam, Ofloxacin, Ampicillin/Sulbactam (f), Ceftolozane-Tazobactam, Polymyxin B, Ceftazidime-Avibactam, Kanamycin, Kanamycin-syn, Moxalactam, Ceftaroline, Cefepime/Sulbactam.

## 15. bacteria\_culture\_test (6,101 записей)

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
biomaterial_rus	Биологический материал на русском языке	—
biomaterial_eng	Биологический материал на английском языке	—
microorganism*	Микроорганизм (возбудитель)	КОЕ/мл

\***Микроорганизмы:** *A. baum/haem*, *Chry. indolog*, *Chrys. meningosept*, *Pseud. oryzihabit*, *Ser. liquefaciens*, *Cory. striatum*, *Ped. pentosaceus*, *Staph. schleiferi*, *Strep. dysgal./ca*, *Strep. gallolytic*, *Strep. mitis gr.*, *Strep. parasanguis*, *Strep. vestibular*, *G. haemolysan*, *M. lacunata*, *Moraxella. sp.*, *A. baum/calc. comp*, *A. baumannii*, *A. lwoffii/haemolyt*, *Acinetobac. sp.*, *Coryn. amycolatum*, *Coryn. urealytic*, *Corynebac. sp.*, *Corynebacterium*, *Bacil. cereus*, *P. putida*, *R. radiobacte*, *Alcalig. faecalis*, *S. aureus*, *S. epidermidi*, *S. haemolytic*, *S. capitis*, *S. coh-ss-coh*, *S. xylosus*, *Staphyl. carnosus*, *S. agalactiae*, *S. constellatus*, *Str. anginosus*, *Str. gordonii*, *Str. intermedius*, *Str. mitis*, *Str. oralis*, *Str. pneumoniae*, *Str. sanguinis*, *Leuconost. pseudom*, *Achromobacter ssp.*, *B. cepacia CF*, *Bur. cepacia*, *Bur. gladioli*, *S. maltophili*, *Sph. paucimob*, *K. oxytoca*, *K. pne-ss-oza*, *K. pne-ss-pne*, *K. pneumoxy*, *Kleb. pneumoniae*, *E. coli*, *E. coli urea+*, *E. cloacae*, *E. gergoviae*, *P. mirabilis*, *P. penn/vulg*, *P. penneri*, *P. vulgaris*, *S. marcescens*, *S. odorifera I*, *S. plymuthica*, *A. caviae*, *Lact. catenafor*, *H. alvei*, *P. multocida*, *C. violaceum*, *C. davisae*, *C. lapagei*, *C. neteri*, *C. freundii*, *C. koseri*, *K. ascorbata*, *Morganella morgan*, *P. agglomeran*, *P. alcalafaci*, *P. rettgeri*, *P. rustigian*, *P. stuartii*, *E. avium*, *E. casselifl/gall*, *E. faecalis*, *E. faecium*, *E. raffinosus*, *M. wisconsens*, *Klebsiella pneumo*, *Streptococcus con*, *Candida albicans*, *E. aerogenes*, *Pseudomonas aeruginosa*.

## 16. cerebrospinal\_fluid\_analysis (1,021 записей)

Поля	Описание	Дополнительно
new_patient_id	Уникальный идентификатор пациента	—
new_hosp_id	Уникальный идентификатор госпитализации	—
post_admission_days	Сутки с момента поступления	—
post_admission_hours	Часы с момента поступления	—
lymphocytes_csف	Лимфоциты в цереброспинальной жидкости	N/ п. зр.
eosinophils_csف	Эозинофилы в цереброспинальной жидкости	N/ п. зр.
neutrophils_csف	Нейтрофилы в цереброспинальной жидкости	N/ п. зр.
macrophages_csف	Макрофаги в цереброспинальной жидкости	N/ п. зр.
monocytes_csف	Моноциты в цереброспинальной жидкости	N/ п. зр.
arachnoid cells_csف	Арахноидальные клетки в цереброспинальной жидкости	N/ п. зр.
granular spheres_csف	Гранулярные сферы в цереброспинальной жидкости	N/ п. зр.
erythrocytes_csف	Эритроциты в цереброспинальной жидкости	N/ п. зр.

## Литература

- Johnson A.E.W., Pollard T.J., Shen L., Lehman L.W.H., Feng M., Ghassemi M., Moody B., et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016; 3: 160035. DOI: 10.1038/sdata.2016.35. PMID: 27219127.
- Johnson A.E.W., Bulgarelli L., Shen L., Gayles A., Shammout A., Horng S., Pollard T.J., et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data*. 2023; 10 (1): 1. DOI: 10.1038/s41597-022-01899-x. PMID: 36596836.
- Pollard T.J., Johnson A.E.W., Raffa J.D., Celi L.A., Mark R.G., Badawi O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci Data*. 2018; 5: 180178. DOI: 10.1038/sdata.2018.178. PMID: 30204154.
- Chen H., Zhu Z., Zhao C., Guo Y., Chen D., Wei Y., Jin J. Central venous pressure measurement is associated with improved outcomes in septic patients: an analysis of the MIMIC-III database. *Crit Care*. 2020; 24 (1): 433. DOI: 10.1186/S13054-020-03109-9. PMID: 32665010.
- Arévalo A.R., Maley J.H., Baker L., da Silva Vieira S.M., da Costa Sousa J.M., Finkelstein S., Mateo-Collado R., et al. Data-driven curation process for describing the blood glucose management in the intensive care unit. *Sci Data*. 2021; 8 (1): 80. DOI: 10.1038/s41597-021-00864-4. PMID: 33692359.
- Liu P, Li S., Zheng T., Wu J., Fan Y., Liu X., Gong W., et al. Sub-phenotyping heterogeneous patients with chronic critical illness to guide individualised fluid balance treatment using machine learning: a retrospective cohort study. *Eclinical-Medicine*. 2023; 59: 101970. DOI: 10.1016/j.eclinm.2023.101970. PMID: 37131542.
- Tyler P.D., Du H., Feng M., Bai R., Xu Z., Horowitz G.L., Stone D.J., et al. Assessment of intensive care unit laboratory values that differ from reference ranges and association with patient mortality and length of stay. *JAMA Netw Open*. 2018; 1 (7): e184521. DOI: 10.1001/JAMANETWORKOPEN.2018.4521. PMID: 30646358.
- Sun Y., He Z., Ren J., Wu Y. 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 (1): 178. DOI: 10.1186/S12871-023-02138-5. PMID: 37231340.
- Hyland S.L., Faltys M., Hüser M., Lyu X., Gumbsch T., Esteban C., Bock C., et al. Early prediction of circulatory failure in the intensive care unit using machine learning. *Nat Med*. 2020; 26 (3): 364–373. DOI: 10.1038/s41591-020-0789-4. PMID: 32152583.
- Thorat P.J., Peppink J.M., Driessen R.H., Sijbrands E.J.G., Kompanje E.J.O., Kaplan L., Bailey H., et al. Sharing ICU patient data responsibly under the Society of Critical Care Medicine/European Society of Intensive Care Medicine Joint Data Science Collaboration: The Amsterdam University Medical Centers Database (AmsterdamUMCdb) example. *Crit Care Med*. 2021; 49 (6): e563–e577. DOI: 10.1097/CCM.0000000000004916. PMID: 33625129.
- Zeng X., Yu G., Lu Y., Tan L., Wu X., Shi S., Duan H., et al. PIC, a paediatric-specific intensive care database. *Sci Data*. 2020; 7 (1): 14. DOI: 10.1038/s41597-020-0355-4. PMID: 31932583.
- Куликов Е.С., Федорова О.С., Толмачев И.В., Рязанцева У.В., Вразнов Д.А., Губанов А.В., Нестерович С.В., с соавт. Русскоязычный репозиторий открытых клинических данных SibMED Data Clinical Repository. *Бюллетень Сибирской Медицины*. 2023; 22 (2): 182–184. Kulikov E.S., Fedorova O.S., Tolmachev I.V., Ryazantseva U.V., Vrazhnov D.A., Gubanov A.V., Nesterovich S.V., et al. Russian-language repository of the open clinical data «SibMed Data Clinical Repository». *Bulletin of Siberian Medicine=Bulleten Sibirskoy Meditsiny*. 2023; 22 (2): 182–184. (in Russ.). DOI: 10.20538/1682-0363-2023-2-182-184.
- Davenport T., Kalakota R. The potential for artificial intelligence in healthcare. *Futur Healthc J*. 2019; 6 (2): 94–98. DOI: 10.7861/futurehosp.6-2-94. PMID: 31363513.
- Реброва О.Ю. Жизненный цикл систем поддержки принятия врачебных решений как медицинских технологий. *Врач и Информационные Технологии*. 2020; (1): 27–37. Rebrova O.Yu. Life cycle of decision support systems as medical technologies. *Doctor and Information Technology=Vrach i Informatsionnye Tekhnologii*. 2020; (1): 27–37. (in Russ.). DOI: 10.37690/1811-0193-2020-1-27-37.
- Гусев А.В., Зарубина Т.В. Поддержка принятия врачебных решений в медицинских информационных системах медицинской организации. *Врач и Информационные Технологии*. 2017; (2): УДК 614.2. Gusev A.V., Zarubina T.V. Clinical decision support in medical information systems of a medical organization. *Doctor and Information Technology=Vrach i Informatsionnye Tekhnologii*. 2017; (2). (in Russ.). UDC 614.2
- Shepherd A. Frictionless Data 2016. <https://www.frictionlessdata.io/> (accessed 3 March 2024).

Поступила 21.03.2024  
Принята 29.03.2024