

ЯДГАРОВ М.Я.,

к.м.н., ФГБНУ "Федеральный научно-клинический центр реаниматологии и реабилитологии" (ФНКЦ РР), г. Москва, Россия; e-mail: myadgarov@fnkcrr.ru. ORCID: 0000-0003-3792-1682

РЕБРОВА О.Ю.,

д.м.н., ФГАОУ ВО "Российский национальный исследовательский медицинский университет им. Н.И. Пирогова" Министерства здравоохранения Российской Федерации, г. Москва, Россия; e-mail: o.yu.rebrova@gmail.com. ORCID: 0000-0002-6733-0958

ШВЫРЕВ С.Л.,

к.м.н., ФГБУ «ЦНИИОИЗ» Министерства здравоохранения Российской Федерации, г. Москва, Россия; e-mail: shvyrev@progcom.ru. ORCID: 0009-0004-9093-6765

БЕРИКАШВИЛИ Л.Б.,

к.м.н., ФГБНУ "Федеральный научно-клинический центр реаниматологии и реабилитологии" (ФНКЦ РР), г. Москва, Россия; e-mail: lberikashvili@fnkcrr.ru. ORCID: 0000-0001-9267-3664

ПОЛЯКОВ П.А.,

ФГБНУ "Федеральный научно-клинический центр реаниматологии и реабилитологии" (ФНКЦ РР), г. Москва, Россия; e-mail: p.polyakov@fnkcrr.ru. ORCID: 0009-0009-6185-349X

ЯКОВЛЕВ А.А.,

ФГБНУ "Федеральный научно-клинический центр реаниматологии и реабилитологии" (ФНКЦ РР), г. Москва, Россия; e-mail: ayakovlev@fnkcrr.ru. ORCID: 0000-0002-8482-1249

ЛИХВАНЦЕВ В.В.,

д.м.н., профессор, ФГБНУ "Федеральный научно-клинический центр реаниматологии и реабилитологии" (ФНКЦ РР), г. Москва, Россия; e-mail: lik0704@gmail.com. ORCID: 0000-0002-5442-6950

ПРОСПЕКТИВНАЯ ВНУТРЕННЯЯ ВАЛИДАЦИЯ XGBOOST-МОДЕЛИ ПРОГНОЗИРОВАНИЯ СЕПСИСА В УСЛОВИЯХ ДЛИТЕЛЬНОГО ПРЕБЫВАНИЯ ПАЦИЕНТОВ В ОРИТ

DOI: 10.25881/18110193_2026_3_150

Аннотация. Цель исследования: провести проспективную валидацию ранее разработанной модели машинного обучения для прогнозирования развития сепсиса у данной категории пациентов.

Материалы и методы: Проспективная валидация выполнена в соответствии с рекомендациями TRIPOD+AI на данных базы "Russian Intensive Care Dataset" (RICD v3.0). Использовалась модель XGBoost, оценивающая риск развития сепсиса каждый час госпитализации с периодом прогнозирования 6 часов на основе антропометрических, клинических, лабораторных и витальных показателей, агрегированных в динамические временные окна. Эффективность модели оценивалась по площади под ROC-кривой (AUROC) с 95% доверительным интервалом (ДИ), а также по диагностическим метрикам при оптимальной отрезной точке.

Результаты: Валидация проведена на когорте пациентов, госпитализированных в ОРИТ ФНКЦ РР в период с октября 2024 по сентябрь 2025 года ($n = 103$; 53 561 пациенто-час наблюдения). Длительность пребывания в ОРИТ составила 32 (27; 38) суток. Модель XGBoost продемонстрировала AUROC 0,737, что соответствовало исходным результатам внутреннего тестирования (AUROC 0,752; 95% ДИ 0,730; 0,767). Ранее определенная оптимальная точка отсечения (0,5651) обеспечила чувствительность 60,3%, специфичность 81,5% и высокую прогностическую ценность отрицательного результата (97,4%). Ввиду высокой доли ложноположительных срабатываний модель показала наибольшую клиническую применимость для исключения сепсиса в ближайшие 6 часов. Вместе с тем, использование многоуровневого подхода к интерпретации делает модель потенциально полезной для стратификации пациентов по вероятности развития сепсиса.

Заключение: Проспективная валидация подтвердила устойчивость и воспроизводимость модели машинного обучения для прогнозирования сепсиса у пациентов, длительно находящихся в ОРИТ. Результаты свидетельствуют о возможности интеграции модели в медицинскую информационную систему как компонента системы поддержки принятия врачебных решений для последующей оценки в условиях реальной клинической практики.

Ключевые слова: сепсис; машинное обучение; хроническое критическое состояние; ОРИТ; прогнозирование; XGBoost.

Для цитирования: Ядгаров М.Я., Реброва О.Ю., Швырев С.Л., Берикашвили Л.Б., Поляков П.А., Яковлев А.А., Лихванцев В.В. Проспективная внутренняя валидация XGBoost-модели прогнозирования сепсиса в условиях длительного пребывания пациентов в ОРИТ. Врач и информационные технологии. 2026; 3: 150-163. DOI: 10.25881/18110193_2026_3_150.

YADGAROV M.YA.,

PhD, Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology, Moscow, Russia;
e-mail: myadgarov@fnkcrr.ru. ORCID: 0000-0003-3792-1682

REBROVA O.YU.,

DSc, Pirogov Russian National Research Medical University, Moscow, Russia;
e-mail: o.yu.rebrova@gmail.com. ORCID: 0000-0002-6733-0958

SHVYREV S.L.,

PhD, Russian Research Institute of Health, Moscow, Russia;
e-mail: shvyrev@progcom.ru. ORCID: 0009-0004-9093-6765

BERIKASHVILI L.B.,

PhD, Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology, Moscow, Russia;
e-mail: Iberikashvili@fnkcrr.ru. ORCID: 0000-0001-9267-3664

POLYAKOV P.A.,

Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology, Moscow, Russia;
e-mail: p.polyakov@fnkcrr.ru. ORCID: 0009-0009-6185-349X

YAKOVLEV A.A.,

PhD, Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology, Moscow, Russia;
e-mail: ayakovlev@fnkcrr.ru. ORCID: 0000-0002-8482-1249

LIKHVANTSEV V.V.,

DSc, Professor, Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology, Moscow, Russia; e-mail: lik0704@gmail.com. ORCID: 0000-0002-5442-6950

PROSPECTIVE INTERNAL VALIDATION OF AN XGBOOST-BASED MACHINE LEARNING MODEL FOR SEPSIS PREDICTION IN PATIENTS WITH PROLONGED ICU STAYS

DOI: 10.25881/18110193_2026_3_150

Abstract. *Introduction: to prospectively validate a previously developed machine learning model for predicting sepsis onset in this patient population.*

Methods: Prospective validation was conducted in accordance with TRIPOD+AI recommendations using data from the Russian Intensive Care Dataset (RICD v3.0). The XGBoost-based model estimated the risk of sepsis onset on an hourly basis with a 6-hour prediction window, using demographic, clinical, laboratory, and vital parameters aggregated into dynamic time windows. Model performance was evaluated using the area under the receiver operating characteristic curve (AUROC) with a 95% confidence interval (CI), as well as diagnostic metrics at the optimal cutoff point.

Results: Validation was performed on a cohort of patients admitted to the ICU of the Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology between October 2024 and September 2025 (n = 103; 53,561 patient-hours). The median ICU length of stay was 32 (27; 38) days. The XGBoost model achieved an AUROC of 0.737, consistent with the internal testing results (AUROC 0.752; 95% CI 0.730; 0.767). The optimal cutoff (≥ 0.5651) yielded a sensitivity of 60.3%, specificity of 81.5%, and a high negative predictive value (97.4%). Although the model showed a high rate of false positives, it demonstrated strong clinical utility for ruling out sepsis within the next 6 hours. Moreover, a multilevel interpretation approach allows risk stratification across low, intermediate, and high probability zones of sepsis development.

Conclusion: Prospective validation confirmed the stability and reproducibility of the machine learning model for sepsis prediction in patients with prolonged ICU stays. The results support the feasibility of integrating the model into an electronic health system as a component of a decision support framework for subsequent real-world evaluation.

Keywords: *sepsis; machine learning; chronic critical illness; intensive care unit; prediction; XGBoost.*

For citation: *Yadgarov M.Ya., Rebrova O.Yu., Shvyrev S.L., Berikashvili L.B., Polyakov P.A., Yakovlev A.A., Likhvantsev V.V. Prospective internal validation of an XGBoost-based machine learning model for sepsis prediction in patients with prolonged ICU stays. Medical doctor and information technology. 2026; 3: 150-163. DOI: 10.25881/18110193_2026_3_150.*

ВВЕДЕНИЕ

Сепсис остаётся одной из ведущих причин летальности и инвалидизации среди пациентов отделений реанимации и интенсивной терапии (ОРИТ), несмотря на значительный прогресс в области антибактериальной терапии, контроля гемодинамики и использования методов экстракорпоральной гемокоррекции [1–3]. Ежегодная глобальная заболеваемость сепсисом превышает 30 миллионов случаев, при этом летальный исход отмечается более чем у 5 миллионов пациентов [4]. По другим данным, на долю септических пациентов приходится от 13 до 40% всех госпитализаций в ОРИТ, и летальность остаётся стабильно высокой – от 25% до 50%, в зависимости от исходной тяжести состояния, возраста, наличия сопутствующих заболеваний и иных факторов [5–10].

Своевременность начала терапии при сепсисе имеет решающее значение для прогноза. В соответствии с концепцией «золотого часа», задержка начала антибактериальной и инфузионной терапии при сепсисе прямо связана с увеличением летальности [11–13]. Однако раннее выявление сепсиса остаётся сложной задачей. Клинические признаки зачастую неспецифичны, а лабораторные биомаркеры характеризуются ограниченной эффективностью [14,15].

Для оценки тяжести состояния и риска летального исхода традиционно применяются клинические шкалы – SOFA, APACHE II, SAPS II, qSOFA и др. Тем не менее, эти инструменты не разрабатывались для динамического мониторинга, и их эффективность для прогнозирования развития сепсиса при использовании в реальных клинических условиях ОРИТ является неудовлетворительной [15–19].

Актуальные рекомендации Surviving Sepsis Campaign (SSC, 2021) подчёркивают необходимость внедрения автоматизированных систем для раннего выявления сепсиса и указывают на возможность использования методов машинного обучения (МО) для решения этой задачи [13]. За последние годы разработан целый ряд правоориентированных моделей (моделей прогнозирования в реальном времени), основанных на непрерывном сборе витальных параметров и лабораторных данных, позволяющих прогнозировать развитие сепсиса за несколько часов до его клинической манифестации, тем самым обеспечивая возможность превентивных вмешательств [19–21].

Несмотря на подтверждённую эффективность моделей МО для прогнозирования развития сепсиса, все они были разработаны и валидированы на выборках пациентов в остром критическом состоянии – в первые несколько суток пребывания в ОРИТ. Отдельную клиническую категорию представляют пациенты, находящиеся в продлённом или хроническом критическом состоянии (ПКС/ХКС), для которых характерны длительное пребывание в ОРИТ, персистирующая органная дисфункция и высокая частота нозокомиальных инфекций [22,23]. В условиях длительной госпитализации острый компонент критического состояния, как правило, утрачивает определяющее значение после девятих суток пребывания в ОРИТ, и дальнейшая клиническая траектория определяется совокупным влиянием коморбидного фона, медико-демографических характеристик, развивающихся осложнений и особенностей проводимой терапии [24]. Несмотря на растущее внимание к данной популяции, единое определение ПКС/ХКС до сих пор не выработано, а специализированные прогностические модели для этой группы пациентов до недавнего времени отсутствовали.

Ранее нами впервые была разработана модель для прогнозирования развития сепсиса у пациентов, находящихся в ПКС/ХКС [25]. Модель, основанная на алгоритме XGBoost, продемонстрировала среднюю дискриминационную способность (AUROC = 0,752 на внутреннем тесте) и показала клиническую применимость для динамической оценки риска сепсиса в условиях длительного пребывания пациента в ОРИТ. Поскольку характеристики пациентов, методы мониторинга и набор регистрируемых параметров со временем изменяются, проспективная валидация необходима для подтверждения устойчивости и актуальности модели и оценки необходимости её актуализации для последующей интеграции в медицинскую информационную систему медицинской организации (МИС) как компонента системы поддержки принятия врачебных решений (СППВР).

Цель настоящего исследования – провести проспективную валидацию ранее разработанной модели прогнозирования развития сепсиса у пациентов, длительно находящихся в ОРИТ, и оценить её прогностическую эффективность, клиническую применимость и необходимость адаптации.

МАТЕРИАЛЫ И МЕТОДЫ

Дизайн исследования

Проведена проспективная валидация ранее разработанной модели прогнозирования развития сепсиса [25] у пациентов, длительно находящихся в ОРИТ Федерального научно-клинического центра реаниматологии и реабилитологии (ФНКЦ РР). Для проверки модели использована база данных "Russian Intensive Care Dataset" (RICD, <https://fnkcrr-database.ru/>) версии 3.0 [26,27], формируемая на основе данных МИС ФНКЦ РР (функционирует на базе госпитальной автоматизированной информационной системы Асклепиус) и включающая демографические и антропометрические сведения, клинические и лабораторные показатели, витальные параметры, сведения о проводимой терапии и исходы госпитализации. В валидационную когорту включены все пациенты, поступившие в ОРИТ ФНКЦ РР с 1 октября 2024 г. по 30 сентября 2025 г.; критерии отбора, состав предикторов, правила обработки данных и определения исходов соответствовали исходной публикации [25]. Модель применялась без дообучения и без изменения набора признаков; пороговые значения (отрезные точки) для клинической интерпретации результатов соответствовали исходной работе. Исследование подготовлено в соответствии с рекомендациями TRIPOD + AI [28,29]. Протокол одобрен локальным этическим комитетом ФНКЦ РР (№ 1/24/1 от 24 апреля 2024 г.).

Критерии соответствия

Критерии отбора полностью соответствовали исходной публикации [25]. В анализ включались пациенты, находившиеся в ОРИТ не менее 24 часов; обязательным условием было наличие хотя бы одной оценки сепсиса по критериям Sepsis-3 [13] в течение пребывания в ОРИТ. Из рассмотрения исключались госпитализации, для которых отсутствовали записи непрерывно мониторируемых витальных показателей или сведения о проводимой терапии. При наличии повторных госпитализаций учитывалось только первое поступление в анализируемый период.

Описание модели

Использовалась модель XGBoost (eXtreme Gradient Boosting), включающая 26 предикторов, для оценки вероятности развития сепсиса в

последующие 6 часов пребывания в ОРИТ. Данные организованы с выравниванием по правому краю, в формате со смещением окна прогнозирования в 1 час. Витальные показатели агрегировались за 3 часа, предшествующие моменту прогнозирования (среднее, минимум, максимум, стандартное отклонение и разница за 3 часа), лабораторные – за 12 часов; также учитывались пол, возраст и сопутствующие патологии. Отрицательные временные окна формировались как у пациентов, у которых сепсис развился позднее, так и у пациентов без сепсиса (подход case-crossover-control или full-window) [30,31].

Модель XGBoost была обучена на наборе данных 345 пациентов (238 396 пациенто-часов), с внутренней валидацией на 115 пациентах (81 621 пациенто-час) и независимым внутренним тестированием на 115 пациентах (68 897 пациенто-часов). Высокая доля пациентов с сепсисом (более 50%) в обучающей и тестовых выборках обусловлена критериями отбора; при обучении модели применялась коррекция дисбаланса классов, что нивелировало влияние этого фактора при анализе данных реальной клинической практики.

По результатам анализа значимости признаков (SHAP-анализ) наибольший вклад в формирование прогноза вносили такие параметры как частота дыхания, возраст, частота сердечных сокращений и температура тела. Подробное описание процедуры отбора признаков, предобработки данных, методов импутации и калибровки приведено в исходной публикации [25]; в настоящей работе эти параметры полностью воспроизведены без изменений.

Статистический анализ

Статистическая обработка данных выполнялась с использованием программ SPSS Statistics 29.0 (IBM, США) и Python 3.11 с библиотеками pandas, numpy, scikit-learn, xgboost и statsmodels. Количественные показатели представлены в виде медианы (Me) и межквартильного размаха (Q1; Q3), категориальные – в виде абсолютных значений и долей (%). Сравнение наборов данных проводилось с использованием критерия Манна-Уитни для количественных и критерия χ^2 Пирсона для категориальных переменных.

Прогностическая эффективность модели оценивалась по площади под ROC-кривой (AUROC) с 95% доверительным интервалом (ДИ).

Критерием успешной проспективной валидации считалось сохранение значения AUROC, полученного на новой когорте, в пределах 95% ДИ, установленного при внутреннем тестировании модели в исходной работе [25]. Для уточнения оптимальной отрезной точки использовался критерий Юдена. Для отрезных точек рассчитывались чувствительность, специфичность, прогностические ценности положительного (ПЦПР) и отрицательного результатов (ПЦОР), а также точность с 95% ДИ (ПЦПР, ПЦОР и точность с поправкой на частоту (преваленс) сепсиса в полном наборе данных). Статистически значимыми считались различия при $p < 0,05$ (двусторонний критерий).

РЕЗУЛЬТАТЫ

Характеристика пациентов

В когорту проспективной валидации были включены 103 пациента (рис. 1). Пациенты находились преимущественно в ПКС/ХКС: длительность пребывания в ОРИТ составила 32 (27; 38) суток [22], госпитальная летальность – 10,7 %.

Общая длительность наблюдения составила 53 561 пациенто-час. Мужчины составляли 52,0 % ($n = 54$), медианный возраст – 63 (49; 75) года. Эпизод сепсиса в ходе госпитализации зарегистрирован у 54 пациентов (52,4 %), медианное время от поступления в ОРИТ до развития септического эпизода – 165 (92; 335) часов.

Основные характеристики проспективной когорты представлены в таблице 1. При сравнении с выборкой, использованной на этапе разработки модели (ретроспективная когорта) [25], статистически значимых различий по возрасту, полу, частоте сепсиса и госпитальной летальности не выявлено ($p > 0,05$). Различие отмечено лишь по длительности пребывания в ОРИТ – медиана 32 сут против 42 сут в исходной когорте ($p < 0,001$).

Проспективная валидация

На этапе проспективной валидации модель XGBoost продемонстрировала AUROC 0,737. Данное значение находилось в пределах доверительного интервала, определённого на этапе внутреннего тестирования (0,730; 0,767), что

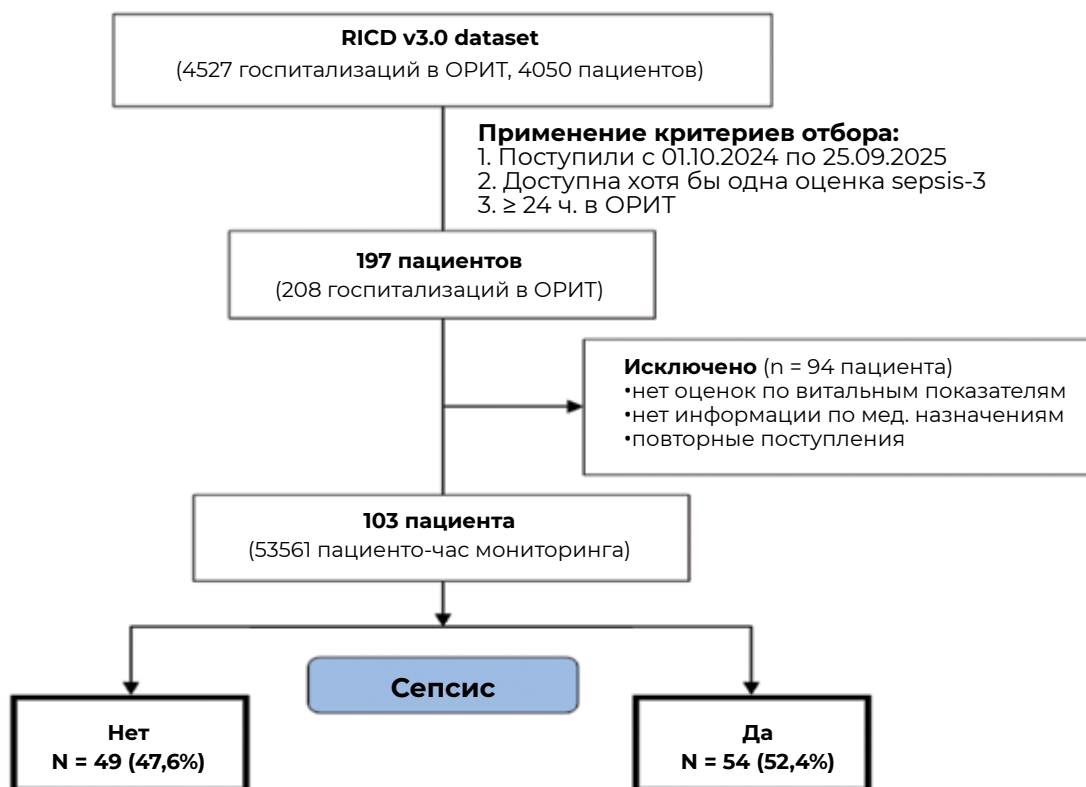


Рисунок 1 – Диаграмма отбора пациентов для проспективной валидации модели.

соответствовало установленному критерию успешности валидации. Результаты ROC-анализа на этапах разработки и проверки модели представлены в таблице 2.

Графическое сравнение ROC-кривых, полученных на различных этапах оценки модели, представлено на рисунке 2.

Отрезные точки и их характеристики

В исходной работе [25] были предложены три отрезные точки: оптимальная – по критерию Юдена (0,5651), а также две дополнительные, определённые по максимальному значению LR+ при сохранении чувствительности выше 60 % и 70 % соответственно. Для оптимальной точки 0,5651 на проспективной когорте

чувствительность составила 60,3 %, специфичность – 81,5 %, ПЦПР – 15,4 %, ПЦОР – 97,4 %, доля верных классификаций – 80,4 % (таблица 3).

При повторении ROC-анализа на проспективной когорте оптимальная точка отсечения оставалась близкой – 0,5747, обеспечивая сопоставимые характеристики: чувствительность 59,5 %, специфичность 83,0 %, ПЦПР – 16,4 %, ПЦОР – 97,3 % (таблица 3). Для обеих оптимальных отрезных точек отмечалась высокая ПЦОР (свыше 97 %) и низкая ПЦПР (15-16 %), что указывает на применимость модели прежде всего для исключения сепсиса в ближайшие 6 часов. Показатель NNE (number needed to evaluate) составил 48,1.

Возможна многоуровневая интерпретация результатов с использованием нескольких

Таблица 1 – Сравнительная характеристика пациентов

Характеристики	Ретроспективная когорта (обучение, валидация, внутр. тестирование)	Проспективная когорта	p-value
Набор данных	RICD v.2.0 (12.2017-09.2024)	RICD v.3.0 (10.2024-09.2025)	-
Общее число пациентов	575	103	
Пол, М	N = 320, 55,7%	N = 52, 50,5%	0,3 ¹
Возраст, лет	61 (48; 73)	63 (49; 75)	0,5 ²
SOFA при поступлении	N = 561, 4 (3; 5)	N = 101, 3 (2; 4)	
Число септических пациентов	336	54	-
Пациенто-часов	388914	53561	
Доля септических пациентов	58,4%	52,4%	0,3 ¹
Время развития сепсиса, часы (от поступления в ОРИТ)	N = 336, 233 (103; 402), от 27 до 4044	N = 54, 165 (92; 335), от 25 до 1007	0,11 ²
Длительность госпитализации в ОРИТ, сут	N = 575, 42 (30; 59)	N = 103, 32 (27; 38)	<0,001 ²
Летальность от всех причин	65, 11,3%	11, 10,7%	0,9 ¹

Сокращения: RICD, Российская база данных реанимационных пациентов; SOFA, Sequential Organ Failure Assessment; ОРИТ, отделение реанимации и интенсивной терапии. ¹ – хи-квадрат критерий; ² – U-критерий Манна-Уитни.

Таблица 2 – ROC-анализ модели XGBoost для прогнозирования развития сепсиса в период 6 часов

Этапы	N	Количество предикторов	AUROC	95% ДИ
Ретроспективная когорта	Обучение	26	0,879	0,873; 0,886
	Валидация		0,711	0,684; 0,729
	Внутренний тест		0,752	0,730; 0,767
	Все наблюдения суммарно		0,819	0,812; 0,826
Внешняя валидация (набор данных Challenge-2)	20000		0,474	0,467; 0,472
Проспективная когорта	103		0,737	0,711; 0,763

Сокращения: AUROC, площадь под ROC кривой; ДИ, доверительный интервал.

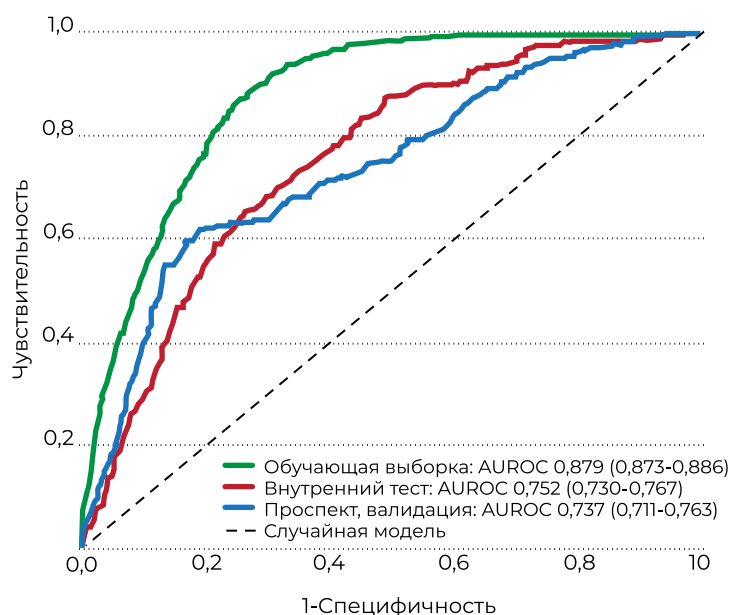


Рисунок 2 – ROC-кривые для модели XGBoost на обучающей выборке, внутреннем тестировании и проспективной валидации.

пороговых зон (рис. 3). При значениях sepsis score менее 0,5 ПЦОР превышала 98 %, при этом доля ложных положительных тревог составляла более 90 % – в данной зоне можно с высокой вероятностью прогнозировать отсутствие сепсиса. Напротив, среди пациентов со значениями sepsis score 0,7 и выше доля тех, у кого действительно развился сепсис в течение 6 часов, составляет 18 % и более, что может служить ориентиром для усиления клинического наблюдения и переоценки критериев Sepsis-3 в течение последующих часов госпитализации (рис. 3).

ОБСУЖДЕНИЕ

Проспективная валидация модели XGBoost проведена на когорте из 103 пациентов (53 561 пациенто-час наблюдения), поступивших в ОПИТ ФНКЦ РР за период 12 месяцев. Модель разрабатывалась для оценки риска развития сепсиса в течение 6 часов от момента прогнозирования. Для формирования признаков использовались временные окна сбора данных длительностью 12 часов для лабораторных показателей и 3 часа для витальных параметров, со смещением окна прогнозирования на 1 час. Модель продемонстрировала сохранение прогностической эффективности, и значение AUROC находилось в пределах доверительного интервала,

определённого при внутреннем тестировании. Оптимальная отрезная точка (0,5747) была близка к выявленной ранее при разработке модели (0,5651), и обеспечила чувствительность 59,5 %, специфичность 83,0 %, высокую ПЦОР (97,3 %) при относительно низкой ПЦПР (16,4 %).

Низкая ПЦПР является распространенным ограничением прогностических моделей, особенно при применении в клинических условиях с невысокой распространённостью прогнозируемого исхода (до 50 %), когда доля ложноположительных срабатываний является значительной, что может снижать воспринимаемую клиническую полезность алгоритма [32]. Аналогичные результаты продемонстрированы и в других работах, включая исследование Yu и соавт. (2022), где модель МО, предназначенная для прогнозирования сепсиса за 6 часов до клинической манифестации, показала сопоставимо низкую ПЦПР 29,1 % при проведении псевдопроспективной оценки в условиях стационара [33]. Высокая ПЦОР при этом позволяет использовать модель преимущественно для исключения сепсиса в ближайшие часы. Возможна и организация многоуровневого подхода к оценке риска, при котором низкие значения sepsis score служат основанием для снижения интенсивности наблюдения, а высокие – для его усиления.

Таблица 3 – Характеристики точек отсечения модели XGBoost для прогнозирования развития сепсиса в период 6 часов

Модель	N	AUROC	p-value	Cutoff для sepsis score	Чувст.	Спец.	ПЦПР#	ПЦОР#	Асс.#
XGBoost (ретроспективная когорта)	575	0,819 (0,812; 0,826)	<0,001	0,4725 (индекс Юдена)†	86,5 (84,5; 88,2)	75,0 (74,9; 75,2)	16,2 (15,9; 16,5)	99,0 (98,9; 99,1)	75,6 (75,5; 75,8)
				0,5651 (макс LR+ и чувст. ≥70%)†	70,9 (68,4; 73,3)	83,1 (82,9; 83,2)	19,0 (18,5; 19,5)	98,1 (97,9; 98,2)	82,4 (82,3; 82,6)
				0,6331 (макс LR+ и чувст. ≥60%)†	60,0 (57,4; 62,6)	87,4 (87,2; 87,5)	21,0 (20,3; 21,7)	97,5 (97,3; 97,7)	85,9 (85,8; 86,0)
XGBoost (проспективная когорта)	103	0,737 (0,711; 0,763)	<0,001	0,4725†	71,4 (66,6; 75,9)	59,05 (49,1; 59,9)	9,0 (8,5; 9,5)	97,4 (96,9; 97,8)	60,1 (59,7; 60,5)
				0,5651†	60,3 (55,2; 65,3)	81,5 (81,2; 81,8)	15,4 (14,4; 16,6)	97,4 (97,0; 97,7)	80,4 (80,1; 80,7)
				0,6331†	48,7 (43,5; 53,8)	87,3 (87,1; 87,6)	17,7 (16,2; 19,3)	96,8 (96,5; 97,1)	85,3 (85,0; 85,6)
				0,5747 (индекс Юдена)	59,5 (54,4; 64,5)	83,0 (82,6; 83,3)	16,4 (15,2; 17,6)	97,3 (97,0; 97,6)	81,7 (81,4; 82,0)

Сокращения: Асс, прогностическая точность; AUROC, площадь под ROC кривой; LR+, отношение правдоподобия положительного результата теста; ДИ, доверительный интервал; ПЦПР, прогностическая ценность положительного результата; ПЦОР, прогностическая ценность отрицательного результата. #Расчет с поправкой на распространённость/преваленс (5,3% – 13316/252836 пациенто-часов). †Отрезная точка из статьи с описанием разработки модели [25].

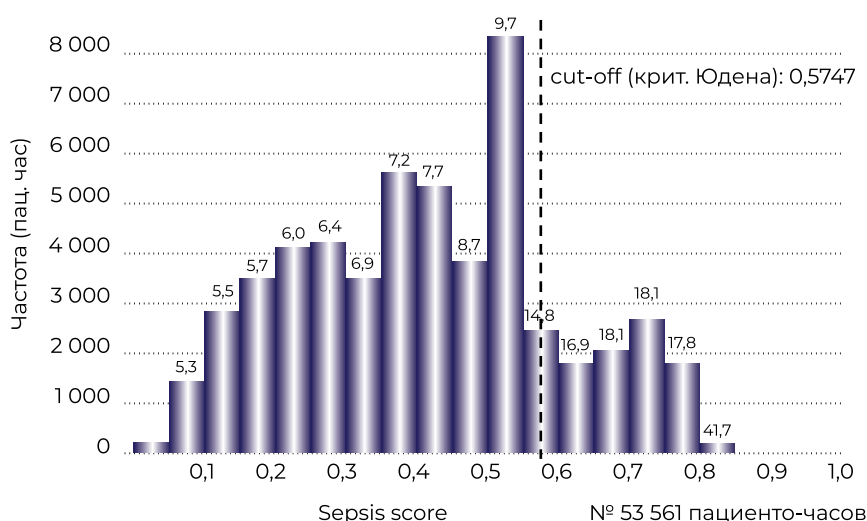


Рисунок 3 – Гистограмма распределения вероятности развития сепсиса (sepsis score) для модели XGBoost на проспективной валидации. Над столбцами указаны значения прогностической ценности положительного результата, пунктирной линией – оптимальная точка отсечения.

Следует отметить, что низкая ПЦПР может быть связана и с особенностями методологии организации временных окон, применённой в нашем исследовании. Используемый подход полного временного окна (full-window, case-crossover-control) обеспечивает более точное воспроизведение реальных условий мониторинга, однако закономерно приводит к снижению метрик, чувствительных к частоте ложноположительных срабатываний, включая ПЦПР [30,31]. Согласно систематическому обзору Wang и соавт. (2025), модели прогнозирования сепсиса в реальном времени демонстрировали медианные значения AUROC 0,886 и 0,861 при прогнозировании за 6 и 12 часов до манифестации сепсиса, которые снижались до 0,783 при внешней валидации в рамках full-window подхода, что сопоставимо с результатами нашего исследования [30].

Ещё одним важным аспектом разработки прогностических моделей является оценка устойчивости/применимости и своевременная переоценка эффективности, связанная с «устареванием» моделей. Ключевое значение, в связи с этим, имеют проведение внешней и проспективной валидации, обеспечивающих объективную оценку воспроизводимости модели в новых популяциях пациентов и реальных клинических условиях. Согласно данным 2025 г., внешняя валидация проводится менее чем в 15 % исследований, связанных с разработкой МО-моделей в интенсивной терапии [34]. Показательно, что при внешней валидации широко применяемой модели прогнозирования сепсиса в клиниках США (Epic Sepsis Model) значение AUROC составило лишь 0,63 по сравнению с исходно заявленными 0,76-0,83, что вызвало серьёзные сомнения в её применимости [35]. При внешней валидации нашей модели на открытых наборах данных острых критически больных пациентов дискриминационная способность модели являлась неудовлетворительной, что обусловлено принципиальными различиями между пациентами в остром критическом состоянии и пациентами, находящимися преимущественно в ПКС/ХКС, что подробно описано в исходной работе [25].

Применимость право-ориентированных моделей прогнозирования подтверждена рядом исследований, показавших, что их клиническая валидация и последующая интеграция в МИС

в качестве компонентов СППВР способствует более раннему началу антибактериальной терапии и улучшению исходов лечения, включая сокращение длительности госпитализации и снижение летальности у пациентов с сепсисом [21,36–39]. В то же время интеграция подобных инструментов в реальную клиническую практику крайне ограничена, в том числе по причине ограниченного доверия со стороны врачей. Отсутствие понимания механизмов принятия решения (феномен «чёрного ящика») снижает готовность клиницистов доверять результатам модели, даже при ее высокой формальной эффективности. При разработке нашей модели интерпретируемость алгоритма XGBoost обеспечивалась с помощью анализа SHAP-диаграмм (SHapley Additive exPlanations), позволяющих оценить как общий вклад каждого предиктора в модель в целом, так и интерпретировать влияние признаков на прогноз для конкретного пациента [25].

Из сотен разработанных моделей для прогнозирования развития сепсиса лишь немногие дошли до реального внедрения в клиническую практику. Это связано не только с проблемой доверия, но и с отсутствием валидации, сложностями интеграции в рабочие процессы и трудностями адаптации моделей к меняющимся условиям клинической практики. Согласно данным Agnello и соавт. (2024), лишь 3 исследования из 39 рассмотренных в обзоре сообщали о проведении внешней валидации [40]. На 2025 г. несколько моделей прогнозирования сепсиса достигли стадии практического внедрения, включая Epic Sepsis Model, которая показала неудовлетворительные результаты при внешней валидации [35], COMPOSER [41], Sepsis Watch [42] и TREWS [43]. Adams и соавт. (2022) провели проспективное нерандомизированное исследование модели TREWS в пяти клиниках, продемонстрировав статистически значимое снижение летальности на 3,3 %, частоты органной дисфункции и длительности госпитализации [43]. В рамках нерандомизированного проспективного исследования модель Sepsis Watch, протестированная в четырех ОПИТ, показала AUROC выше 0,9 [44]. COMPOSER, в свою очередь, была протестирована в двух ОПИТ, показав снижение летальности на 1,9 % [41].

Все эти модели были разработаны и валидированы на когортах острых критически больных

пациентов, чья длительность пребывания в ОРИТ не превышала в среднем 7 суток. В то же время, на сегодняшний день отсутствуют иные модели, помимо нашей, разработанные для пациентов, находящихся в ПКС/ХКС. Полноценная внешняя валидация нашей модели также не представляется возможной из-за отсутствия открытых наборов данных пациентов, длительно находящихся в ОРИТ, за исключением RICD.

Безусловно, реальный эффект от внедрения моделей прогнозирования в клиническую практику может быть оценен только в рамках крупных мультицентровых исследований. Однако перед проведением таких исследований требуется интеграция модели в МИС. В рамках проспективной валидации наша модель продемонстрировала сохранение прогностической эффективности и может быть рекомендована к внедрению в МИС для последующей оценки, однако в перспективе ее использование возможно исключительно для прогнозирования сепсиса у пациентов в ПКС/ХКС.

Ограничения исследования

Исследование характеризуется рядом ограничений, которые следует учитывать при интерпретации его результатов. Во-первых, полноценная проспективная валидация модели возможна только после её интеграции в МИС, что позволит оценить её клиническую применимость в реальных условиях. Во-вторых, размер выборки для проспективной когорты (103 пациента) может ограничивать статистическую мощность исследования, хотя объем данных для валидации (53 561 пациенто-часов) является относительно большим. Кроме того, данные для исследования были получены из одного центра (RICD v3.0), что может ограничить обобщаемость полученных результатов. Также стоит отметить, что модель обучена для

прогнозирования развития первого эпизода сепсиса у каждого пациента. Повторные эпизоды сепсиса и их временные характеристики не были учтены, что ограничивает понимание того, как модель работает в условиях многократно рецидивирующих эпизодов сепсиса в условиях длительного пребывания в ОРИТ.

Перспективы

Поскольку модель показала сохранение прогностической эффективности при валидации на проспективной когорте, планируется её реализация в действующей МИС ФНКЦ РР для оценки применимости в условиях реальной клинической практики.

Вместе с тем для обеспечения внешней валидации, воспроизводимости и возможности сравнительного анализа прогностических моделей, разработанных для пациентов, находящихся в ПКС/ХКС, необходимо развитие и открытая доступность специализированных наборов данных для этой популяции.

ЗАКЛЮЧЕНИЕ

В настоящем исследовании была проведена проспективная валидация модели МО для прогнозирования развития сепсиса у пациентов, длительно находящихся в условиях ОРИТ. Модель продемонстрировала дискриминационную способность, сопоставимую с результатами, полученными на этапе внутреннего тестирования при разработке модели, что подтверждает её клиническую применимость в этой популяции пациентов. На основе полученных результатов можно рекомендовать внедрение модели в МИС для дальнейшей оценки в условиях реальной клинической практики.

Конфликт интересов: отсутствует

Источники финансирования: отсутствуют.

ЛИТЕРАТУРА/REFERENCES

1. Singer M, Deutschman CS, Seymour CW, et al. The third international consensus definitions for sepsis and septic shock (sepsis-3). JAMA - J Am Med Assoc. 2016; 315: 801-10. doi: 10.1001/jama.2016.0287.
2. La Via L, Sangiorgio G, Stefani S, et al. The Global Burden of Sepsis and Septic Shock. Epidemiologia. 2024; 5: 456-78. doi: 10.3390/epidemiologia5030032.
3. Arbous SM, Termorshuizen F, Brinkman S, et al. Three-year mortality of ICU survivors with sepsis, an infection or an inflammatory illness: an individually matched cohort study of ICU patients in the Netherlands from 2007 to 2019. Crit Care. 2024; 28: 374. doi: 10.1186/s13054-024-05165-x.

4. Fleischmann C, Scherag A, Adhikari NKJ, et al. Assessment of global incidence and mortality of hospital-treated sepsis current estimates and limitations. *Am J Respir Crit Care Med*. 2016; 193: 25972. doi: 10.1164/rccm.201504-0781OC.
5. Sakr Y, Jaschinski U, Wittebole X, et al. Sepsis in intensive care unit patients: Worldwide data from the intensive care over nations audit. *Open Forum Infect Dis*. 2018; 5: ofy313. doi: 10.1093/ofid/ofy313.
6. Лянгузов А.В., Лучинин А.С., Игнатьев С.В. и др. Предикторы летальности в отделении реанимации и интенсивной терапии у пациентов с гемобластозами и инфекционными осложнениями // *Анестезиология и реаниматология*. – 2023. – №1. – С.33-38. [Lyanguzov AV, Luchinin AS, Ignat'ev SV, et al. Predictors of ICU mortality in patients with hemoblastosis and infectious complications. *Anesteziologiya i Reanimatologiya*. 2023; 1: 33-38. (In Russ.)] doi: 10.17116/anaesthesiology202301133.
7. Кочкин А.А., Ядгаров М.Я., Берикашвили Л.Б. и др. Предикторы летальных исходов у пациентов с рефрактерным септическим шоком // *Вестник анестезиологии и реаниматологии*. – 2021. – Т.18. – №3. – С.30-35. [Kochkin AA, Yadgarov MY, Berikashvili LB, et al. Predictors of lethal outcomes in patients with refractory septic shock. *Vestnik anesteziologii i reanimatologii*. 2021; 18(3): 30-35. (In Russ.)] doi: 10.21292/2078-5658-2021-18-3-30-35.
8. Abe T, Ogura H, Shiraishi A, et al. Characteristics, management, and in-hospital mortality among patients with severe sepsis in intensive care units in Japan: The FORECAST study. *Crit Care*. 2018; 22: 322. doi: 10.1186/s13054-018-2186-7.
9. Mohamed AKS, Mehta AA, James P. Predictors of mortality of severe sepsis among adult patients in the medical Intensive Care Unit. *Lung India*. 2017; 34: 330-5. doi: 10.4103/lungindia.lungindia_54_16.
10. Vincent JL, Marshall JC, Namendys-Silva SA, et al. Assessment of the worldwide burden of critical illness: The Intensive Care Over Nations (ICON) audit. *Lancet Respir Med*. 2014; 2: 380-6. doi: 10.1016/S2213-2600(14)70061-X.
11. Raghavan M, Marik PE. Management of sepsis during the early “golden hours”. *J Emerg Med*. 2006; 31: 185-99. doi: 10.1016/j.jemermed.2006.05.008.
12. Im Y, Kang D, Ko RE, et al. Time-to-antibiotics and clinical outcomes in patients with sepsis and septic shock: a prospective nationwide multicenter cohort study. *Crit Care*. 2022; 26: 19. doi: 10.1186/s13054-021-03883-0.
13. Evans L, Rhodes A, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Intensive Care Med*. 2021; 47: 1181-247. doi: 10.1007/s00134-021-06506-y.
14. Vincent J-L, Beumier M. Diagnostic and prognostic markers in sepsis. *Expert Rev Anti Infect Ther*. 2013; 11: 265-75. doi: 10.1586/eri.13.9.
15. He RR, Yue GL, Dong ML, et al. Sepsis Biomarkers: Advancements and Clinical Applications—A Narrative Review. *Int J Mol Sci*. 2024; 25. doi: 10.3390/ijms25169010.
16. Adegbite BR, Edoa JR, Ndzebe Ndoumba WF, et al. A comparison of different scores for diagnosis and mortality prediction of adults with sepsis in Low-and-Middle-Income Countries: a systematic review and meta-analysis. *EClinicalMedicine*. 2021; 42: 101184. doi: 10.1016/j.eclinm.2021.101184.
17. Konjety P, Chakole VG. Beyond the Horizon: A Comprehensive Review of Contemporary Strategies in Sepsis Management Encompassing Predictors, Diagnostic Tools, and Therapeutic Advances. *Cureus*. 2024; 16: e64249. doi: 10.7759/cureus.64249.
18. Sakib N, Ahamed SI, Khan RA, et al. Unpacking prevalence and dichotomy in quick sequential organ failure assessment and systemic inflammatory response syndrome parameters: Observational data-driven approach backed by sepsis pathophysiology. *JMIR Med Informatics*. 2020; 8: e18352. doi: 10.2196/18352.
19. Islam MM, Nasrin T, Walther BA, et al. Prediction of sepsis patients using machine learning approach: A meta-analysis. *Comput Methods Programs Biomed*. 2019; 170: 1-9. doi: 10.1016/j.cmpb.2018.12.027.
20. Yadgarov MY, Landoni G, Berikashvili LB, et al. Early detection of sepsis using machine learning algorithms: a systematic review and network meta-analysis. *Front Med*. 2024; 11: 1491358. doi: 10.3389/fmed.2024.1491358.
21. Fleuren LM, Klausch TLT, Zwager CL, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive Care Med*. 2020; 46: 383-400. doi: 10.1007/s00134-019-05872-y.
22. Likhvantsev VV., Berikashvili LB, Yadgarov MY, et al. The Tri-Steps Model of Critical Conditions in Intensive Care: Introducing a New Paradigm for Chronic Critical Illness. *J Clin Med*. 2024; 13. doi: 10.3390/jcm13133683.

23. Hawkins RB, Raymond SL, Stortz JA, et al. Chronic Critical Illness and the Persistent Inflammation, Immunosuppression, and Catabolism Syndrome. *Front Immunol.* 2018; 9: 1511. doi: 10.3389/fimmu.2018.01511.
24. Bagshaw SM, Stelfox HT, Iwashyna TJ, et al. Timing of onset of persistent critical illness: a multi-centre retrospective cohort study. *Intensive Care Med.* 2018; 44: 2134-44. doi: 10.1007/s00134-018-5440-1.
25. Yadgarov MY, Rebrova OY, Berikashvili LB, et al. Machine Learning Model for Sepsis Prediction in Prolonged and Chronic Critical Illness: Development and Validation Using Retrospective Real-World ICU Data. *J Clin Med.* 2026; 15: 777. doi: 10.3390/jcm15020777.
26. Гречко А.В., Ядгаров М.Я., Яковлев А.А. и др. Российская база данных пациентов отделений реанимации и интенсивной терапии RICD // *Общая реаниматология.* – 2024. – Т.20. – №3. – С.22-31. [Grechko AV, Yadgarov MY, Yakovlev AA, et al. Russian Intensive Care Dataset - RICD. *Obshchaya Reanimatologiya.* 2024; 20(3): 22-31. (In Russ.)] doi: 10.15360/1813-9779-2024-3-22-31.
27. RICD - Open dataset [updated 2025 Oct 16; cited 2025 Dec 9]. Available from: <https://fnkcrr-database.ru>.
28. Collins GSS, Reitsma JBB, Altman DGG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. *BMC Med.* 2015; 13: 1. doi: 10.1186/s12916-014-0241-z.
29. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *Bmj.* 2024; 385. doi: 10.1136/bmj-2023-078378.
30. Wang Z, Wang W, Sun C, et al. A methodological systematic review of validation and performance of sepsis real-time prediction models. *Npj Digit Med.* 2025; 8: 190. doi: 10.1038/s41746-025-01587-1.
31. Schvetz M, Fuchs L, Novack V, et al. Outcomes prediction in longitudinal data: Study designs evaluation, use case in ICU acquired sepsis. *J Biomed Inform.* 2021; 117: 103734. doi: 10.1016/j.jbi.2021.103734.
32. Liu Z, Khojandi A, Mohammed A, et al. HeMA: A hierarchically enriched machine learning approach for managing false alarms in real time: A sepsis prediction case study. *Comput Biol Med.* 2021; 131: 104255. doi: 10.1016/j.compbio.2021.104255.
33. Yu SC, Gupta A, Betthausen KD, et al. Sepsis Prediction for the General Ward Setting. *Front Digit Heal.* 2022; 4: 848599. doi: 10.3389/fdgth.2022.848599.
34. Rockenschaub P, Akay EM, Carlisle BG, et al. External validation of AI-based scoring systems in the ICU: a systematic review and meta-analysis. *BMC Med Inform Decis Mak.* 2025; 25: 5. doi: 10.1186/s12911-024-02830-7.
35. Wong A, Otles E, Donnelly JP, et al. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. *JAMA Intern Med.* 2021; 181: 1065-70. doi: 10.1001/jamainternmed.2021.2626.
36. Amland RC, Haley JM, Lyons JJ. A Multidisciplinary Sepsis Program Enabled by a Two-Stage Clinical Decision Support System: Factors That Influence Patient Outcomes. *Am J Med Qual.* 2016; 31: 501-8. doi: 10.1177/1062860615606801.
37. Narayanan N, Gross AK, Pintens M, et al. Effect of an electronic medical record alert for severe sepsis among ED patients. *Am J Emerg Med.* 2016; 34: 185-8. doi: 10.1016/j.ajem.2015.10.005.
38. Tafelski S, Nachtigall I, Deja M, et al. Computer-assisted Decision Support for Changing Practice in Severe Sepsis and Septic Shock. *J Int Med Res.* 2010; 38: 1605-16. doi: 10.1177/147323001003800505.
39. Liu VX, Fielding-Singh V, Greene JD, et al. The timing of early antibiotics and hospital mortality in sepsis. *Am J Respir Crit Care Med.* 2017; 196: 856-63. doi: 10.1164/rccm.201609-1848OC.
40. Agnello L, Vidali M, Padoan A, et al. Machine learning algorithms in sepsis. *Clin Chim Acta.* 2024; 553: 117738. doi: 10.1016/j.cca.2023.117738.
41. Boussina A, Shashikumar SP, Malhotra A, et al. Impact of a deep learning sepsis prediction model on quality of care and survival. *Npj Digit Med.* 2024; 7: 14. doi: 10.1038/s41746-023-00986-6.
42. Sendak MP, Ratliff W, Sarro D, et al. Real-world integration of a sepsis deep learning technology into routine clinical care: Implementation study. *JMIR Med Informatics.* 2020; 8: e15182. doi: 10.2196/15182.
43. Adams R, Henry KE, Sridharan A, et al. Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis. *Nat Med.* 2022; 28: 1455-60. doi: 10.1038/s41591-022-01894-0.
44. Valan B, Prakash A, Ratliff W, et al. Evaluating sepsis watch generalizability through multisite external validation of a sepsis machine learning model. *Npj Digit Med.* 2025; 8: 350. doi: 10.1038/s41746-025-01664-5.