v7.n10.2026

REVISTA CIENTÍFICA - RECIMA21 ISSN 2675-6218

DESEMPENHO E APLICABILIDADE DE ALGORITMOS DE APRENDIZADO DE MÁQUINA NA PREDIÇÃO DE SEPSE E SUA MORTALIDADE: UMA REVISÃO SISTEMÁTICA

PERFORMANCE AND APPLICABILITY OF MACHINE LEARNING ALGORITHMS FOR SEPSIS AND SEPSIS-RELATED MORTALITY PREDICTION: A SYSTEMATIC REVIEW

RENDIMIENTO Y APLICABILIDAD DE LOS ALGORITMOS DE APRENDIZAJE AUTOMÁTICO PARA LA PREDICCIÓN DE LA SEPSIS Y LA MORTALIDAD: UNA REVISIÓN SISTEMÁTICA

Luiza Stéfani Brune Leidemer[1], Cláudio Fernando Resin Geyer[2]

e7109014
https://doi.org/10.47820/recima21.v7i10.9014

PUBLICADO: 10/2026

RESUMO

Este estudo analisou técnicas de aprendizado de máquina na predição de sepse e mortalidade. A revisão foi conduzida nas bases ACM, Springer, ScienceDirect e IEEE, seguindo as diretrizes de Kitchenham e Charters e o protocolo PRISMA 2020, resultando na inclusão de 46 artigos. A qualidade metodológica dos estudos foi avaliada com a ferramenta PROBAST+AI. As questões de pesquisa estão relacionadas à comparação entre os modelos de ML e métodos tradicionais, ao desempenho dos algoritmos e a sua aplicabilidade clínica. Os resultados mostraram que Random Forest, XGBoost e aprendizado profundo apresentam desempenho superior aos escores convencionais. Conclui-se que esses modelos podem contribuir para a detecção precoce e potencialmente apoiar estratégias voltadas à redução da mortalidade.

PALAVRAS-CHAVE: Aprendizado de Máquina. Sepse. Mortalidade. Predição.

ABSTRACT

This study analyzed machine learning techniques for sepsis and mortality prediction. The review was conducted using the ACM, Springer, ScienceDirect, and IEEE databases, following the guidelines proposed by Kitchenham and Charters and the PRISMA 2020 protocol, resulting in the inclusion of 46 articles. The methodological quality of the studies was assessed using the PROBAST+AI tool. The research questions addressed the comparison between ML models and traditional methods, algorithm performance, and their clinical applicability. The results showed that Random Forest, XGBoost, and deep learning achieved superior performance compared to conventional scoring systems. It was concluded that these models may contribute to early detection and potentially support strategies aimed at reducing mortality in clinical settings.

KEYWORDS: *Machine Learning. Sepsis. Mortality. Prediction.*

RESUMEN

Este estudio analizó técnicas de aprendizaje automático para la predicción de la sepsis y la mortalidad. La revisión se realizó utilizando las bases de datos ACM, Springer, ScienceDirect e IEEE, siguiendo las directrices propuestas por Kitchenham y Charters y el protocolo PRISMA 2020, dando como resultado la inclusión de 46 artículos.

[1] Graduada em Física Médica pela UFCSPA, Mestranda em Computação pela UFRGS.
[2] Graduado em Engenharia Mecânica pela UFRGS, mestre em Ciências da Computação pela UFRGS e doutor em Informática pela Université de Grenoble I. Professor da Universidade Federal do Rio Grande do Sul.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

La calidad metodológica de los estudios se evaluó mediante la herramienta PROBAST+AI. Las preguntas de investigación abordaron la comparación entre los modelos de aprendizaje automático y los métodos tradicionales, el rendimiento de los algoritmos y su aplicabilidad clínica. Los resultados mostraron que Random Forest, XGBoost y los modelos de aprendizaje profundo alcanzaron un rendimiento superior al de los sistemas de puntuación convencionales. Se concluyó que estos modelos pueden contribuir a la detección temprana y a la reducción de la mortalidad en entornos clínicos.

***PALABRAS CLAVE**: Aprendizaje automático. Sepsis. Mortalidad. Predicción.*

INTRODUÇÃO

A sepse, também conhecida como infecção generalizada, pode ser definida como uma resposta sistêmica a uma doença infecciosa, seja ela causada por bactérias, vírus, fungos ou protozoários. É uma resposta inadequada do organismo a uma infecção, levando à disfunção orgânica progressiva e, em casos graves, podendo ser fatal (ILAS, 2018).

No mundo, a sepse é responsável por aproximadamente 50 milhões de casos por ano e, ao menos, 11 milhões de mortes. Já no Brasil, são registrados cerca de 400 mil casos de sepse e 240 mil mortes por ano. É uma condição clínica prevalente, com altos índices de morbidade e mortalidade, além de gerar custos consideráveis para o sistema de saúde; por isso, o reconhecimento e a intervenção terapêutica precoce são determinantes para o prognóstico clínico. Atualmente, o Brasil tem uma taxa de mortalidade por sepse maior que a de países desenvolvidos, exigindo mais atenção ao problema (Brasil, 2021).

Algoritmos de aprendizado de máquina são amplamente utilizados em pesquisas para prever sepse, além de auxiliar condutas médicas e ações preventivas (Bezerra *et al.*, 2023). Ademais, modelos de aprendizado de máquina podem auxiliar a equipe médica na identificação precoce do início da sepse e de pacientes com maior risco de morte, contribuindo para a redução da mortalidade e melhor alocação de recursos hospitalares, além de serem alternativas aos sistemas tradicionais de pontuação como SOFA (*Sequential Organ Failure Assessment*), qSOFA (*quick Sequential Organ Failure Assessment*), SIRS (S*ystemic Inflammatory Response Syndrome*) e MEWS (*Modified Early Warning Score*) (Islam *et al.*, 2019).

Apesar do aumento do número de pesquisas sobre o tema, ainda existem incertezas quanto aos algoritmos que apresentam os melhores resultados, à sua capacidade de superar os sistemas tradicionais de pontuação e à forma como esses modelos estão sendo utilizados na prática clínica. Nesse contexto, o problema que orienta esta pesquisa consiste em compreender: qual é o desempenho e a aplicabilidade clínica dos algoritmos de aprendizado de máquina utilizados na predição da sepse e de sua mortalidade?

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

O objetivo desta pesquisa é realizar uma análise sobre o uso de técnicas de aprendizado de máquina na predição de sepse e de sua mortalidade, explorando tanto seu desempenho quanto sua aplicabilidade prática. Busca-se compreender quais algoritmos são mais empregados na detecção da doença e na estimativa de mortalidade, avaliando seu desempenho por meio de métricas de acurácia, sensibilidade, especificidade e outras relevantes. Além disso, pretende-se identificar os modelos que apresentam maior robustez e eficiência no contexto clínico. A pesquisa também examina de que forma essas ferramentas estão sendo aplicadas na prática hospitalar, destacando os benefícios e as limitações observadas.

1. REFERENCIAL TEÓRICO

O estudo de Nikravangolsefid *et al.* (2024) realizou uma revisão sistemática seguindo as diretrizes PRISMA, na qual foram analisados 31 estudos publicados entre 2014 e 2024 sobre o uso de algoritmos de aprendizado de máquina na predição da mortalidade hospitalar de pacientes com sepse. Os autores avaliaram aspectos metodológicos, desempenho dos modelos e risco de viés, observando resultados geralmente superiores aos obtidos pelos escores clínicos tradicionais. No entanto, a revisão restringiu-se à predição da mortalidade, enquanto a presente revisão também aborda a predição da sepse e a aplicabilidade clínica dos modelos.

A revisão sistemática de Deng *et al.* (2022) teve como objetivo propor novos critérios de avaliação e padrões de relato para modelos de predição de sepse baseados em aprendizado de máquina. Foram analisados 21 modelos selecionados de acordo com a metodologia PRISMA. O estudo investigou as definições de sepse, as fontes de dados, os métodos de treinamento e validação, os valores de métricas como a AUROC e a utilização de técnicas de *deep learning* (aprendizado profundo). Embora apresente uma avaliação técnica detalhada dos modelos, o trabalho concentra-se apenas na predição da sepse e não inclui análises relacionadas à mortalidade. A presente revisão amplia esse escopo ao incluir a predição da mortalidade, a comparação com métodos clínicos tradicionais e a análise das implicações práticas no contexto hospitalar.

Zhang *et al.* (2023) buscaram avaliar o valor preditivo dos métodos de aprendizado de máquina na predição da mortalidade relacionada à sepse em diferentes períodos. Foram incluídos 50 estudos, envolvendo 104 modelos preditivos, e o risco de viés foi avaliado por meio da ferramenta PROBAST. Os resultados indicaram desempenho superior dos modelos de aprendizado de máquina em comparação com ferramentas clínicas convencionais, com destaque para os algoritmos *Random Forest* (Floresta Aleatória) e XGBoost. A análise também

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

identificou o lactato como uma variável preditora relevante. Apesar de seu rigor e profundidade, o estudo limitou-se à predição da mortalidade. A presente revisão, além de abordar a mortalidade, considera a predição da sepse, o desempenho dos modelos em comparação com abordagens tradicionais e suas aplicações clínicas, oferecendo uma visão mais abrangente.

O estudo de Yang, Cui e Song (2023) avaliou o desempenho de modelos de aprendizado de máquina na predição do início da sepse em pacientes internados em unidades de terapia intensiva. Foram selecionados 23 estudos, que totalizaram mais de quatro milhões de pacientes e 26 modelos diferentes. Os algoritmos *Random Forest* e XGBoost destacaram-se pelo elevado desempenho, alcançando acurácia superior a 90% em muitos casos. Embora apresente contribuições relevantes sobre a eficácia dos algoritmos na detecção precoce da sepse, o estudo não aborda a predição da mortalidade nem discute a integração dos modelos ao cenário clínico. Em contraste, a presente revisão considera as duas tarefas de predição e discute de forma mais ampla as implicações clínicas dos resultados.

A Tabela 1 apresenta uma síntese comparativa entre esta revisão e estudos relevantes previamente publicados na literatura. Foram considerados aspectos como o tipo de revisão conduzida, o período de busca adotado, o foco principal de cada estudo, a abordagem clínica descrita e as principais diferenças em relação ao presente trabalho.

Tabela 1. Comparação entre esta revisão e outros estudos selecionados

Artigo	Autores (Ano)	Tipo de Revisão	Período de Busca	Foco Principal	Abordagem Clínica	Diferenças em relação a esta revisão
Machine learning for predicting mortality in adult critically ill patients with Sepsis: A systematic review	Nikravangolsefid *et al.* (2024)	Revisão Sistemática	2014–2024	Mortalidade e qualidade dos relatos	Comparação com escores clínicos	Não avalia predição da sepse; foca apenas em mortalidade; não explora aplicação clínica detalhada
Evaluating machine learning models for sepsis prediction: A systematic review of methodologies	Deng *et al.* (2022)	Revisão Sistemática	2012–2021	Avaliação metodológica de modelos	Critérios de validação e reprodutibilidade	Não foca em aplicação clínica nem em comparação com modelos tradicionais

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Machine learning for the prediction of sepsis-related death: a systematic review and meta-analysis	Zhang *et al.* (2023)	Revisão Sistemática com Meta-análise	Até ago. 2022	Mortalidade relacionada à sepse	Avaliação de subgrupos clínicos	Não contempla predição da sepse; foco apenas em morte; não analisa aplicação prática
Predicting sepsis onset in ICU using machine learning models: a systematic review and meta-analysis	Yang, Cui e Song (2023)	Revisão Sistemática com Meta-análise	Até nov. 2022	Início da sepse em UTI	Aplicação em UTI com dados em larga escala	Foco exclusivo no início da sepse; não analisa mortalidade nem comparação com métodos tradicionais
Esta revisão	Autores (2026)	Revisão Sistemática da Literatura	2019–2026	Predição de sepse e mortalidade com ML	Aplicação clínica e comparação com modelos tradicionais	Integra sepse e morte; foca em aplicação clínica e comparação com modelos tradicionais; abordagem exploratória ampla

Fonte: Elaboração própria (2026).

Em contraste com esses estudos, o presente trabalho diferencia-se ao realizar uma revisão com foco ampliado tanto na predição da sepse quanto da mortalidade, incluindo a avaliação do desempenho dos algoritmos em comparação a modelos tradicionais e investigando como essas soluções têm sido aplicadas na prática clínica.

2. METODOLOGIA

Para a condução desta revisão, adotaram-se as diretrizes metodológicas propostas por Kitchenham e Charters (2007). Para a formulação das perguntas de pesquisa e definição dos critérios de elegibilidade dos estudos, utilizou-se uma adaptação do *framework* PICO (População, Interesse, Comparação e Desfecho) (Santos; Pimenta; Nobre, 2007) ao contexto desta revisão, e PROBAST+AI (*Prediction Model Risk of Bias Assessment Tool – Artificial Intelligence Extension*) (Moons *et al.*, 2025), respectivamente. A seguir, a Tabela 2 apresenta a

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

estrutura das perguntas de pesquisa formuladas para esta revisão. Para cada questão de pesquisa (RQ1 a RQ3), são detalhados os elementos de população, intervenção/interesse, comparador e desfecho. As seguintes questões de pesquisa nortearam a revisão:

- RQ1: Como o uso de técnicas de aprendizado de máquina pode aprimorar a predição de sepse em comparação aos modelos tradicionais?
- RQ2: Quais são os algoritmos utilizados na predição da doença ou da mortalidade e seu desempenho?
- RQ3: Como os modelos utilizados para a predição da mortalidade por sepse são aplicados na área clínica?

Tabela 2. Estruturação das perguntas de pesquisa no formato PICO

RQ	P (População)	I (Interesse)	C (Comparação)	O (*Outcome*/Desfecho)
RQ1	Pacientes com sepse	Algoritmos de ML	Modelos e escores tradicionais	Melhoria na predição da sepse
RQ2	Pacientes com sepse	Algoritmos de ML	Diferentes algoritmos	Desempenho dos algoritmos
RQ3	Profissionais clínicos	Algoritmos de ML	Diferentes algoritmos	Aplicabilidade clínica

Fonte: Elaboração própria, adaptada de PICO (2026).

A busca bibliográfica foi conduzida nas bases de dados ACM, Springer, ScienceDirect e IEEE. Para a formulação da busca, utilizou-se a sintaxe ("Sepsis" AND ("Prediction" OR "Predictive Analytics") AND ("Machine Learning")). A seleção dos estudos ocorreu em três etapas: leitura dos títulos, leitura dos resumos e leitura integral dos textos completos. A triagem foi realizada de forma sistemática e com base em critérios previamente definidos.

Os seguintes critérios de inclusão foram seguidos: os artigos deveriam estar publicados em inglês ou português; publicados entre 2019 e 2026; o conteúdo deveria estar relacionado ao tema da pesquisa, conforme indicado no título, palavras-chave ou resumo; os estudos deveriam possuir DOI (Identificador de Objeto Digital) e ser artigos completos; os estudos deveriam ser publicados em periódicos ou conferências; e os artigos deveriam estar disponíveis na íntegra.

Quanto aos critérios de exclusão, foram desconsiderados: estudos duplicados ou semelhantes, sendo considerado apenas o mais recente; artigos que não estavam relacionados ao tema e aos objetivos desta revisão; trabalhos que não utilizam aprendizado de máquina como

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

abordagem principal; e artigos de opinião, cartas ao editor, editoriais ou resumos sem o artigo completo.

A extração dos dados foi realizada manualmente com o apoio de um formulário, que contemplou as informações de cada estudo. Além disso, foram analisadas as estratégias de avaliação e validação dos modelos, a comparação do desempenho com abordagens convencionais e as principais conclusões relatadas pelos autores.

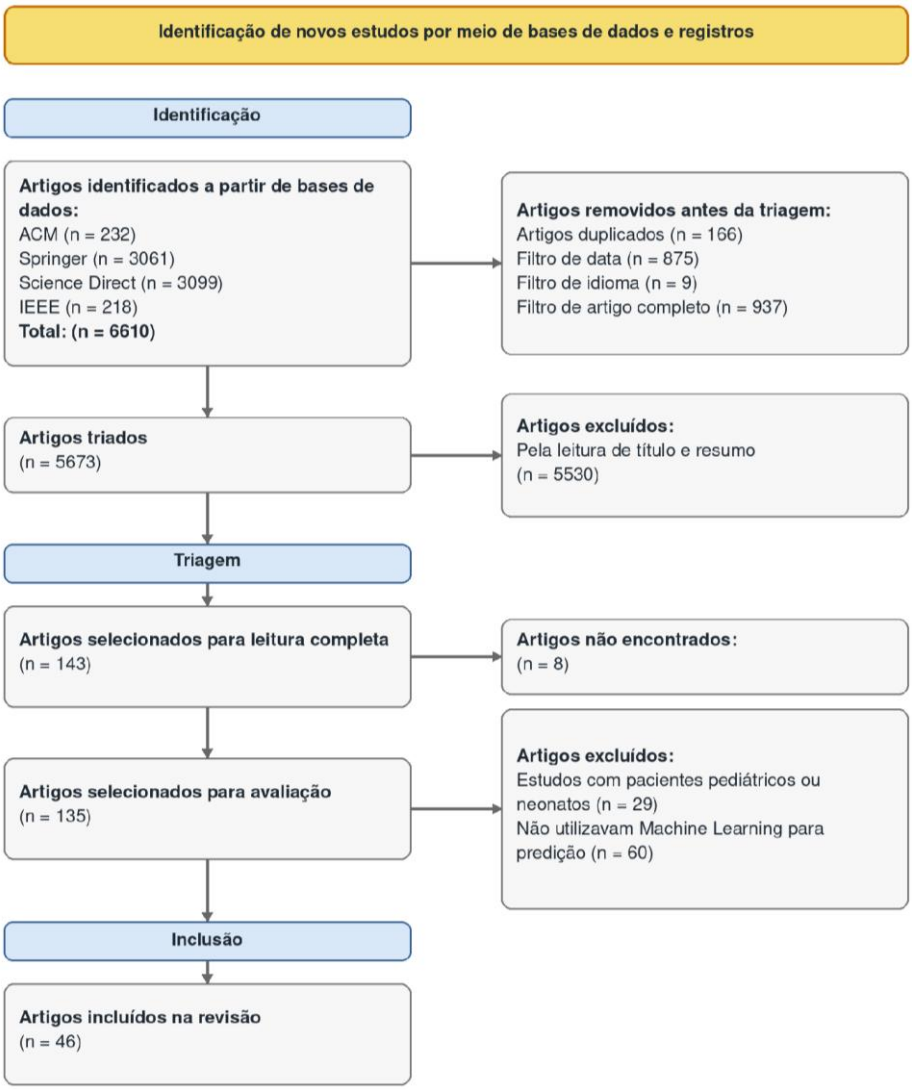
3. RESULTADOS E DISCUSSÃO

A seleção dos estudos seguiu as recomendações do modelo PRISMA 2020 *(Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) (Page *et al.*, 2021), conforme ilustrado na Figura 1. Inicialmente, foram identificados 6.610 artigos nas bases de dados. Após a remoção de duplicados e a aplicação de filtros por data, idioma e tipo de publicação, 5.673 artigos foram considerados para triagem. Nessa etapa, 5.530 estudos foram excluídos após a leitura de título e resumo, restando 143 para leitura completa. Desses, 8 não foram encontrados em sua íntegra, totalizando 135 artigos avaliados. Em seguida, 89 foram excluídos por não atenderem aos critérios de elegibilidade. Ao final, 46 artigos[1] foram incluídos na revisão.

[1] Artigos selecionados podem ser visualizados no link: https://bit.ly/rsl-sepse

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Figura 1. Fluxograma do processo de identificação, triagem e inclusão dos estudos



Fonte: Elaboração própria, adaptada de PRISMA (2026).

Para avaliar a qualidade, risco de viés e aplicabilidade dos estudos incluídos na revisão, foi utilizada a ferramenta PROBAST+AI. O PROBAST+AI consiste em duas partes distintas: desenvolvimento do modelo e avaliação do modelo. Para esta revisão, foi adotada apenas a parte referente à avaliação do modelo, que consiste em 18 questões divididas em quatro domínios: participantes e fontes de dados, preditores, desfecho e análise; cada questão deve ser respondida com: sim, provavelmente sim, provavelmente não, não, sem informação ou não

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

aplicável. A Tabela 3 apresenta a avaliação dos quatro domínios e a avaliação geral dos artigos selecionados.

Tabela 3. Avaliação do risco de viés segundo os domínios do PROBAST+AI

Artigo	Participantes e fontes de dados	Preditores	Desfecho	Análises	Total
1	LOW	LOW	LOW	LOW–MODERATE	LOW
2	LOW	LOW	LOW	LOW	LOW
3	LOW	LOW	LOW	LOW	LOW
4	LOW	LOW	LOW	LOW–MODERATE	LOW
5	LOW	LOW	LOW	LOW	LOW
6	LOW	LOW	LOW	LOW–MODERATE	LOW
7	LOW	LOW	LOW	LOW	LOW
8	LOW	LOW	LOW	LOW–MODERATE	LOW
9	LOW	LOW	LOW	MODERATE	LOW
10	LOW	LOW	LOW	LOW	LOW
11	LOW	LOW	LOW	LOW–MODERATE	LOW
12	LOW	LOW	LOW	MODERATE	LOW–MODERATE
13	LOW	LOW	LOW	LOW	LOW
14	LOW	LOW	LOW	LOW	LOW
15	LOW	LOW	LOW	LOW–MODERATE	LOW
16	LOW	LOW	LOW	MODERATE	LOW–MODERATE
17	LOW	LOW	LOW	LOW	LOW
18	LOW	LOW	LOW	MODERATE	LOW–MODERATE
19	LOW	LOW	LOW	LOW	LOW
20	LOW	LOW	LOW	LOW	LOW
21	LOW	LOW	LOW	LOW–MODERATE	LOW
22	LOW	LOW	LOW	LOW	LOW
23	LOW	LOW	LOW	LOW–MODERATE	LOW
24	LOW	LOW	LOW	LOW	LOW
25	LOW	LOW	LOW	LOW	LOW
26	LOW	LOW	LOW	LOW–MODERATE	LOW
27	LOW	LOW	LOW	LOW	LOW
28	LOW	LOW	LOW	LOW	LOW

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

29	LOW	LOW	LOW	LOW	LOW
30	LOW	LOW	LOW	MODERATE	LOW–MODERATE
31	LOW	LOW	LOW	LOW	LOW
32	LOW	LOW	LOW	LOW	LOW
33	LOW	LOW	LOW	LOW	LOW
34	LOW	LOW	LOW	LOW	LOW
35	LOW	LOW	LOW	LOW	LOW
36	LOW	LOW	LOW	MODERATE	LOW–MODERATE
37	LOW	LOW	LOW	MODERATE	LOW–MODERATE
38	LOW	LOW	LOW	LOW	LOW
39	LOW	LOW	LOW	LOW	LOW
40	LOW	LOW	LOW	LOW	LOW
41	LOW	LOW	LOW	LOW	LOW
42	LOW	LOW	LOW	LOW	LOW
43	LOW	LOW	LOW	LOW	LOW
44	LOW	LOW	LOW	LOW	LOW
45	LOW	LOW	LOW	LOW	LOW
46	LOW	LOW	LOW	LOW	LOW

Fonte: Elaboração própria, adaptada de PROBAST+AI (2026).

De modo geral, os artigos apresentaram risco de viés classificado como *LOW* (baixo), especialmente nos aspectos relacionados a participantes e fonte de dados, preditores e desfechos. Essa constatação indica que a maioria dos estudos utilizou bases de dados adequadas e definiu de forma consistente as variáveis preditoras e os desfechos. Em contrapartida, o domínio de análises demonstrou maior variabilidade entre os estudos, com classificações adaptadas que incluíram *MODERATE* (moderado) e *LOW–MODERATE* (baixo-moderado). Essa variação decorre, principalmente, da falta de informações que permitissem responder de forma completa às questões do *framework*, resultando frequentemente em respostas como “provavelmente sim”, “provavelmente não” ou “sem informação”.

3.1. Resultados para RQ1

Tradicionalmente, escores clínicos como SIRS, SOFA, qSOFA, NEWS e MEWS são utilizados para identificar pacientes em risco de sepse ou pior evolução clínica. Essas ferramentas utilizam sinais vitais, exames laboratoriais e indicadores de disfunção orgânica para

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

calcular uma pontuação que auxilia na avaliação do risco e na tomada de decisão clínica. No entanto, esses instrumentos apresentam limitações importantes, incluindo baixa sensibilidade em fases iniciais, dependência de critérios estáticos e menor capacidade de capturar relações não lineares entre variáveis clínicas e laboratoriais. Em contraste, técnicas de aprendizado de máquina podem processar dados mais complexos e dinâmicos, oferecendo maior precisão, melhor discriminação e potencial para detecção precoce.

Essa vantagem é observada principalmente nos estudos que realizaram comparações diretas com escores clínicos tradicionais. Delahanty *et al.* (2019) desenvolveram o Risk of Sepsis (RoS), baseado em *gradient boosting*, e o compararam com SIRS, SOFA, qSOFA, NEWS e MEWS. O RoS apresentou AUROC entre 0,93 e 0,97, superando todos os escores avaliados. Considerando as informações disponíveis nas primeiras 24 horas após a chegada do paciente ao pronto-socorro, o RoS obteve AUROC de 0,97, enquanto SOFA, NEWS, qSOFA, MEWS e SIRS apresentaram valores de 0,90, 0,84, 0,80, 0,79 e 0,77, respectivamente. Em comparação ao SOFA, que apresentou o melhor desempenho entre os escores tradicionais, o RoS alcançou maior sensibilidade na identificação dos pacientes com sepse durante a primeira hora após a chegada ao pronto-socorro (67,7% contra 49,2%) e nas primeiras 24 horas de atendimento (84,6% contra 80,4%). O qSOFA apresentou sensibilidade de apenas 3,7% na primeira hora e de 23,5% nas primeiras 24 horas, indicando baixa capacidade de identificar precocemente os pacientes com sepse.

Resultados semelhantes foram observados por Barton *et al.* (2019), que avaliaram um modelo baseado em *gradient boosted ensemble*. O algoritmo apresentou AUROC de 0,88, mantendo desempenho de 0,84 em 24 horas antes da sepse e 0,83 em 48 horas antes do evento, além de AUROC de 0,89 na validação externa. Esses valores foram superiores aos escores tradicionais avaliados, incluindo SOFA (0,72), SIRS (0,66), MEWS (0,61) e qSOFA (0,60). O estudo também evidenciou redução de alarmes falsos, uma vez que SIRS e SOFA produziram, respectivamente, 122% e 86% mais falsos alertas do que o algoritmo proposto.

Burdick *et al.* (2020) também validaram retrospectivamente um algoritmo baseado em *gradient boosted trees* para predição de sepse grave com até 48 horas de antecedência, utilizando dados de 461 hospitais dos Estados Unidos. O modelo apresentou AUROC de 0,930, sensibilidade de 0,800, especificidade de 0,933 e acurácia de 0,929, superando MEWS, SOFA e SIRS, que obtiveram AUROC de 0,725, 0,716 e 0,655, respectivamente.

No contexto da predição de choque séptico, Yun *et al.* (2021) compararam o NEWS isolado com modelos de aprendizado de máquina baseados em informações de triagem inicial. O NEWS apresentou AUROC de 0,804, enquanto a regressão logística e o XGBoost alcançaram

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

AUROC de 0,844 e 0,845, respectivamente. De forma semelhante, Ullah *et al.* (2021) observaram desempenho superior do XGBoost em comparação a outros algoritmos e escores clínicos, com acurácia de 92,31%, enquanto MEWS apresentou sensibilidade de 73% e especificidade de 52%, e NEWS apresentou sensibilidade de 30% e especificidade de 89%.

Essa diferença de desempenho pode estar associada ao fato de que escores como SOFA, APACHE-II e PIRO foram desenvolvidos para avaliação geral de pacientes críticos, e não especificamente para a predição precoce da sepse. Tang, Zhang e Li (2024) destacam essa limitação e relatam evidências de melhor desempenho dos modelos de aprendizado de máquina, como o XGBoost, que apresentou AUC de 0,835, superando a regressão logística (0,737) e o SOFA (0,621) na identificação de pacientes sépticos de alto risco.

Na predição de mortalidade por sepse, também foram observados ganhos em relação aos escores tradicionais de gravidade. Wang *et al.* (2024) desenvolveram um modelo interpretável baseado em XGBoost para predição precoce de mortalidade em sepse, com AUROC de 0,873 na validação interna e 0,844 na validação externa. Esse desempenho foi superior aos escores SAPS II (0,728) e SOFA (0,728).

Outros estudos também compararam modelos de aprendizado de máquina com o SOFA na predição de mortalidade. Pan *et al.* (2023) observaram que regressão logística e Gaussian Naive Bayes apresentaram AUROC de 0,76, ligeiramente superior ao SOFA tradicional, que obteve AUROC de 0,742. Zhang *et al.* (2024) relataram desempenho mais expressivo do XGBoost, com AUC de 0,940 e F1-score de 0,937, em comparação ao SOFA, que apresentou AUC de 0,687 e F1-score de 0,914. Diwan *et al.* (2025) também observaram desempenho superior de modelos explicáveis, como Extra Trees (AUROC 0,87), *Random Forest* (0,86), XGBoost (0,84) e Gradient Boosting (0,81), em relação ao SOFA, que apresentou AUROC de 0,77.

De forma geral, os estudos indicam que modelos de aprendizado de máquina superam escores tradicionais na predição precoce de sepse e de mortalidade, com melhores valores de AUROC, sensibilidade, especificidade ou calibração. Esses resultados sugerem que o aprendizado de máquina pode aprimorar a predição ao integrar mais variáveis, capturar relações não lineares e gerar estimativas de risco mais individualizadas, configurando uma alternativa promissora aos escores clínicos baseados em regras fixas.

3.2. Resultados para RQ2

Os estudos incluídos nesta revisão empregaram uma ampla variedade de algoritmos de aprendizado de máquina para a predição da sepse e de sua mortalidade. Ao todo, foram

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

identificados 30 algoritmos diferentes nos 46 estudos analisados, evidenciando a diversidade de abordagens utilizadas na literatura. Entre os métodos mais recorrentes, destacaram-se o *Random Forest*, com 28 menções, o XGBoost, com 25 menções, e a Regressão Logística, com 24 menções. Também foram frequentemente utilizados algoritmos como Support Vector Machine (SVM), Árvores de Decisão, Gradient Boosting, LightGBM, KNN, Naive Bayes e MLP. A predominância desses modelos indica uma preferência por técnicas supervisionadas, especialmente aquelas voltadas à classificação de pacientes em grupos de risco. A Tabela 4 apresenta os algoritmos identificados nos estudos analisados.

Tabela 4. Frequência dos algoritmos utilizados nos estudos analisados

Algoritmo	Frequência
Random Forest	28
XGBoost	25
Regressão Logística	24
SVM/SVR	16
Decision Tree	14
Gradient Boosting	13
LightGBM	10
KNN	10
MLP	8
Naive Bayes	8
AdaBoost	6
LASSO	4
LDA	3
Extra Trees	3
Regressão Linear	2
QDA	2
LSTM	2
CatBoost	2
ANN	2
HMM	1
SGD	1
Easy Ensemble	1
Bagging	1
CNN	1
Boosted Tree	1
CART	1

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Bootstrap Forest	1
RoS	1
NAVOY® Sepsis	1
RSBoost	1

Fonte: Elaboração própria (2026).

De modo geral, os modelos supervisionados foram os mais empregados nos estudos analisados, tanto para a predição da doença quanto para a estimativa de mortalidade. Dentro dessa categoria, os modelos de *ensemble* baseados em árvores apresentaram papel de destaque, sobretudo por sua capacidade de lidar com dados clínicos tabulares, variáveis heterogêneas e relações não lineares. Algoritmos como *Random Forest*, XGBoost, *Gradient Boosting* e LightGBM foram amplamente explorados devido ao bom desempenho em tarefas de classificação e à robustez diante de bases de dados com diferentes níveis de complexidade.

Os modelos lineares e estatísticos, como Regressão Logística, Regressão Linear, LASSO, LDA e Naive Bayes, também apareceram nos estudos, embora, em muitos casos, tenham sido utilizados como modelos de comparação. A Regressão Logística, em particular, permanece relevante por sua simplicidade, interpretabilidade e ampla aceitação na área clínica. Já o LASSO contribui para a seleção automática de variáveis, reduzindo a complexidade do modelo e auxiliando na identificação dos preditores mais relevantes. O Naive Bayes, por sua vez, apresenta baixo custo computacional, enquanto o LDA pode ser útil quando os dados seguem distribuições aproximadamente lineares ou gaussianas. Apesar dessas vantagens, esses métodos tendem a apresentar limitações quando aplicados a relações clínicas mais complexas e não lineares.

Outros algoritmos clássicos, como SVM e KNN, também foram identificados na revisão. As SVMs são adequadas para problemas com alta dimensionalidade e podem apresentar bom desempenho em conjuntos de dados com separação complexa entre classes. O KNN, por outro lado, é um método simples e intuitivo, podendo ser útil em bases menores ou mais balanceadas. Entretanto, sua aplicação em grandes volumes de dados clínicos pode ser limitada pelo custo computacional e pela sensibilidade à escala das variáveis.

No contexto dos modelos de aprendizado profundo, foram observadas abordagens baseadas em Redes Neurais Artificiais (ANN), MLP, CNN e LSTM. Esses modelos são especialmente relevantes quando há necessidade de capturar relações complexas entre variáveis ou padrões temporais associados à evolução clínica dos pacientes. As arquiteturas LSTM, por exemplo, são frequentemente utilizadas em estudos que consideram a sequência temporal dos dados, permitindo analisar mudanças progressivas no estado do paciente ao longo

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

do tempo. Já as CNNs, embora mais comuns em aplicações envolvendo imagens ou sinais, também foram exploradas em alguns estudos com dados clínicos estruturados ou biomarcadores (Tang; Zhang; Li, 2024).

Além dos modelos individuais, os estudos analisados também apresentaram abordagens híbridas, que combinam diferentes técnicas de aprendizado de máquina com o objetivo de melhorar o desempenho preditivo. Essas estratégias buscam aproveitar as vantagens específicas de cada algoritmo, integrando, por exemplo, métodos de seleção de variáveis, modelos de *ensemble* e redes neurais. Um exemplo identificado foi o RSBoost, que combina uma camada CNN-SVM com o método LSBoost, permitindo a realização de tarefas de classificação e regressão. Essa abordagem apresentou precisão diagnóstica de 96,2% na identificação de biomarcadores em amostras de sangue em até 24 horas (Shin *et al.*, 2025).

Também foram identificadas combinações envolvendo *ensemble learning*, Regressão Logística e LASSO. Nesses casos, o *ensemble* contribui para aumentar a robustez e a precisão das previsões, enquanto o LASSO auxilia na seleção das variáveis mais relevantes, favorecendo a interpretação dos resultados (Zhang *et al.*, 2024). Outros estudos exploraram combinações entre MLP e Gradient Boosting Trees (GBT), unindo a capacidade dos modelos neurais de representar relações não lineares com a robustez dos algoritmos baseados em árvores para dados tabulares (Ying; Abu-Samah, 2022).

Em relação ao desempenho, os modelos de *ensemble* baseados em árvores apresentaram os resultados mais consistentes entre os estudos avaliados. De modo geral, esses métodos alcançaram médias de AUC entre 0,85 e 0,91, sensibilidade entre 75% e 82% e especificidade entre 87% e 93%. Além disso, algoritmos de boosting demonstraram potencial para antecipar a predição da sepse, com bons valores de acurácia e AUROC até 48 horas antes de sua manifestação clínica (Barton *et al.*, 2019; Burdick *et al.*, 2020). Nas abordagens de *deep learning*, modelos LSTM voltados à análise temporal alcançaram AUC entre 0,82 e 0,89 (Tang; Zhang; Li, 2024), enquanto arquiteturas CNN apresentaram resultados mais modestos, com AUC entre 0,75 e 0,82 em cenários com dados estáticos.

Apesar dos bons resultados, a análise das métricas também evidencia desafios comuns entre os algoritmos. A sensibilidade média tende a ser inferior à especificidade, indicando maior dificuldade na identificação de todos os casos positivos de sepse ou mortalidade sem aumentar o número de falsos positivos. Da mesma forma, os valores médios de F1-score, geralmente entre 0,83 e 0,87, sugerem que mesmo os modelos com melhor desempenho ainda enfrentam limitações no equilíbrio entre sensibilidade e precisão. Esses achados reforçam que, embora os algoritmos de aprendizado de máquina apresentem desempenho promissor, sua aplicação

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

clínica exige validação robusta, calibração adequada e avaliação cuidadosa dos riscos associados a falsos negativos e falsos positivos.

3.3. Resultados para RQ3

A aplicação clínica dos modelos preditivos ainda é limitada, mas alguns estudos descreveram implementações experimentais ou simulações em ambientes hospitalares. Os modelos foram utilizados principalmente para apoiar a tomada de decisão médica. No entanto, a integração com os sistemas de prontuário eletrônico e a validação em ambientes reais de cuidado continuam sendo desafios. Os estudos apontam que, para uma adoção mais ampla, é necessário garantir a interpretabilidade dos modelos, bem como avaliar seu impacto clínico de forma mais sistemática.

Os sistemas de alerta precoce em tempo real constituem uma das aplicações mais difundidas. Baseados em algoritmos como XGBoost e *Random Forest*, esses sistemas são integrados aos prontuários eletrônicos hospitalares para processar continuamente os dados clínicos. O estudo do sistema NAVOY® Sepsis, desenvolvido por Persson *et al.* (2024), ilustra essa abordagem, na qual o monitoramento automático de sinais vitais, parâmetros laboratoriais e características demográficas permite gerar alertas quando o risco de mortalidade atinge valores críticos. Essa capacidade de análise contínua e em tempo real supera a avaliação proporcionada pelos escores tradicionais.

Na gestão de recursos hospitalares, os modelos preditivos têm ajudado a organizar os pacientes de acordo com o risco de cada um. O estudo de Gao *et al.* (2024) mostra que essas ferramentas auxiliam na alocação de leitos de UTI, na definição de prioridades para monitoramento intensivo e na seleção de pacientes para protocolos especiais, corroborando a tomada de decisão mais informada.

Por meio da técnica SHAP, os sistemas desenvolvidos por Zhang *et al.* (2024) e Wang *et al.* (2024) incorporam abordagens para categorizar a importância de cada variável preditora na série histórica de dados dos pacientes. O método SHAP pode melhorar a interpretabilidade dos modelos de aprendizado de máquina e auxiliar a equipe clínica a compreender melhor o raciocínio por trás das previsões. Ademais, o estudo de Shakeri *et al.* (2021) descobriu que, ao longo da internação, há uma variação significativa na importância relativa das variáveis preditoras utilizadas pelo modelo, dependendo do estágio temporal considerado (por exemplo, fases iniciais versus tardias da admissão). Técnicas de interpretabilidade como o SHAP permitem quantificar essas mudanças, possibilitando a análise da contribuição individual de cada variável em diferentes janelas temporais por meio da comparação dos valores de SHAP.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

4. CONSIDERAÇÕES FINAIS

O estudo evidencia o papel dos modelos de aprendizado de máquina na predição e no manejo da sepse, condição que ainda apresenta altas taxas de mortalidade, especialmente no Brasil. Os resultados demonstram que algoritmos como *Random Forest*, XGBoost e redes neurais profundas superam os sistemas tradicionais, como SOFA, SIRS e MEWS, em métricas como AUC, sensibilidade e tempo de detecção. Essa vantagem decorre da capacidade desses modelos de analisar padrões complexos em variáveis dinâmicas, superando as limitações dos escores clínicos convencionais. Além dos algoritmos mais frequentemente utilizados, observou-se o emprego de modelos híbridos em alguns estudos, evidenciando uma tendência de combinar diferentes abordagens para aprimorar a capacidade preditiva dos sistemas de apoio à decisão clínica. Os resultados reportados para esses modelos foram, em geral, bastante satisfatórios, alcançando desempenho comparável ou superior ao de algoritmos individuais em diferentes cenários de predição da sepse e de sua mortalidade.

Casos bem-sucedidos documentados na literatura mostram que, quando adequadamente implementados, esses sistemas preditivos de ML podem identificar precocemente a sepse, contribuindo para a redução da mortalidade, desde a alocação de recursos até a priorização de intervenções. No entanto, apesar dos resultados promissores, ainda existem limitações para a adoção desses modelos na prática clínica, especialmente em relação à interpretabilidade e à validação em diferentes contextos hospitalares.

Também foi realizado um levantamento da frequência de uso dos diferentes algoritmos, com o objetivo de identificar padrões e possíveis tendências entre os estudos. Observou-se que, na predição do surgimento de sepse, o XGBoost foi o algoritmo mais utilizado, enquanto, na predição de mortalidade por sepse, houve predominância do *Random Forest*, o que pode refletir diferenças no desempenho, na interpretabilidade ou na adequação desses modelos a distintos desfechos clínicos. Assim, trabalhos futuros podem explorar de forma mais aprofundada os fatores que influenciam a escolha dos algoritmos em diferentes contextos de predição, bem como avaliar o impacto dessas escolhas nos resultados obtidos.

REFERÊNCIAS

BARTON, Christopher *et al.* Evaluation of a machine learning algorithm for up to 48-hour advance prediction of sepsis using six vital signs. **Computers in Biology and Medicine**, v. 109, p. 79–84, jun. 2019. DOI: https://doi.org/10.1016/j.compbiomed.2019.04.027.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

BEZERRA, Antônio; MACIEL, Nathanael; FILHO, Luciano; MENDES, Alice; GOIS, Francisco; SILVA, Lucilane. Efetividade de algoritmos de inteligência artificial para predição de sepse em adultos de unidades de terapia intensiva: revisão de escopo. **Revista Interfaces**, 2023.

BRASIL. Ministério da Saúde. **Dia Mundial da Sepse: Ministério da Saúde alerta para a importância do diagnóstico precoce.** 2021. Disponível em: https://bvsms.saude.gov.br/?p=5399. Acesso em: 20 mar. 2026.

BURDICK, Hoyt *et al.* Validation of a machine learning algorithm for early severe sepsis prediction: a retrospective study predicting severe sepsis up to 48 h in advance using a diverse dataset from 461 US hospitals. **BMC Medical Informatics and Decision Making**, v. 20, art. 276, 2020. DOI: https://doi.org/10.1186/s12911-020-01347-z.

DELAHANTY, Ryan J.; ALVAREZ, JoAnn; FLYNN, Lisa M.; SHERWIN, Robert L.; JONES, Spencer S. Development and evaluation of a machine learning model for the early identification of patients at risk for sepsis. **Annals of Emergency Medicine**, v. 73, n. 4, p. 334–344, 2019. DOI: https://doi.org/10.1016/j.annemergmed.2018.11.036.

DENG, Hong-Fei *et al.* Evaluating machine learning models for sepsis prediction: a systematic review of methodologies. **iScience**, v. 25, n. 1, art. 103651, jan. 2022. DOI: https://doi.org/10.1016/j.isci.2021.103651.

DIWAN, Saumya *et al.* Explainable machine learning models for mortality prediction in patients with sepsis in tertiary care hospital ICU in low- to middle-income countries. **Intensive Care Medicine Experimental**, v. 13, art. 56, 2025. DOI: https://doi.org/10.1186/s40635-025-00765-5.

GAO, Jie *et al.* Prediction of sepsis mortality in ICU patients using machine learning methods. **BMC Medical Informatics and Decision Making**, v. 24, art. 228, 2024. DOI: https://doi.org/10.1186/s12911-024-02630-z.

INSTITUTO LATINO AMERICANO DE SEPSE (ILAS). **Implementação de protocolo gerenciado de sepse: protocolo clínico: atendimento ao paciente adulto com sepse/choque séptico.** São Paulo: ILAS, 2018. Revisado em agosto de 2018. Disponível em: https://ilas.org.br/wp-content/uploads/2022/02/protocolo-de-tratamento.pdf. Acesso em: 26 mar. 2026.

ISLAM, Md. Mohaimenul; NASRIN, Tahmina; WALTHER, Bruno Andreas; WU, Chieh-Chen; YANG, Hsuan-Chia; LI, Yu-Chuan. Prediction of sepsis patients using machine learning approach: a meta-analysis. **Computer Methods and Programs in Biomedicine**, v. 170, p. 1–9, 2019. DOI: https://doi.org/10.1016/j.cmpb.2018.12.027.

KITCHENHAM, Barbara; CHARTERS, Stuart. **Guidelines for performing systematic literature reviews in software engineering.** Keele: Keele University; Durham: University of Durham, 2007. (EBSE Technical Report, EBSE-2007-01).

MOONS, Karel G. M. *et al.* PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. **BMJ**, v. 388, e082505, 2025. DOI: https://doi.org/10.1136/bmj-2024-082505.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

NIKRAVANGOLSEFID, Niloofar *et al.* Machine learning for predicting mortality in adult critically ill patients with sepsis: a systematic review. **Journal of Critical Care**, v. 84, art. 154889, dez. 2024. DOI: https://doi.org/10.1016/j.jcrc.2024.154889.

PAGE, Matthew J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. **BMJ**, v. 372, art. n71, 2021. DOI: https://doi.org/10.1136/bmj.n71.

PAN, Xiaobin *et al.* Evaluate prognostic accuracy of SOFA component score for mortality among adults with sepsis by machine learning method. **BMC Infectious Diseases**, v. 23, art. 76, 2023. DOI: https://doi.org/10.1186/s12879-023-08045-x.

PERSSON, Inger; MACURA, Andreas; BECEDAS, David; SJÖVALL, Fredrik. Early prediction of sepsis in intensive care patients using the machine learning algorithm NAVOY® Sepsis: a prospective randomized clinical validation study. **Journal of Critical Care**, v. 83, art. 154400, 2024. DOI: https://doi.org/10.1016/j.jcrc.2023.154400.

SANTOS, Cristina Mamédio da Costa; PIMENTA, Cibele Andrucioli de Mattos; NOBRE, Moacyr Roberto Cuce. The PICO strategy for the research question construction and evidence search. **Revista Latino-Americana de Enfermagem**, Ribeirão Preto, v. 15, n. 3, p. 508–511, maio/jun. 2007. DOI: https://doi.org/10.1590/S0104-11692007000300023.

SHAKERI, Esmaeil; MOHAMMED, Emad A.; SHAKERI H. A., Zahra; FAR, Behrouz. Exploring features contributing to the early prediction of sepsis using machine learning. In: ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC), 43., 2021. **Proceedings [...]**. Piscataway: IEEE, 2021. p. 2472–2475. DOI: https://doi.org/10.1109/EMBC46164.2021.9630317.

SHIN, Joonchul *et al.* Artificially intelligent nasal perception for rapid sepsis diagnostics. **npj Digital Medicine**, v. 8, n. 1, art. 476, 2025. DOI: https://doi.org/10.1038/s41746-025-01851-4.

TANG, Yan; ZHANG, Yu; LI, Jiaxi. A time series driven model for early sepsis prediction based on transformer module. **BMC Medical Research Methodology**, v. 24, art. 23, 2024. DOI: https://doi.org/10.1186/s12874-024-02075-3.

ULLAH, Asad; QAYYUM, Huma; KHAN, Muhammad Khateeb; AHMAD, Fawad. Sepsis detection using Extreme Gradient Boost (XGB): a supervised learning approach. In: MOHAMMAD ALI JINNAH UNIVERSITY INTERNATIONAL CONFERENCE ON COMPUTING (MAJICC), 2021. **Proceedings [...]**. Piscataway: IEEE, 2021. p. 1–6. DOI: https://doi.org/10.1109/MAJICC53071.2021.9526260.

WANG, Yiping; GAO, Zhihong; ZHANG, Yang; LU, Zhongqiu; SUN, Fangyuan. Early sepsis mortality prediction model based on interpretable machine learning approach: development and validation study. **Internal and Emergency Medicine**, v. 20, n. 3, p. 909–918, 2024. DOI: https://doi.org/10.1007/s11739-024-03732-2.

YANG, Zhen; CUI, Xiaojun; SONG, Zhen. Predicting sepsis onset in ICU using machine learning models: a systematic review and meta-analysis. **BMC Infectious Diseases**, v. 23, art. 635, 2023. DOI: https://doi.org/10.1186/s12879-023-08614-0.

YING, Teh Xuan; ABU-SAMAH, Asma. Early prediction of sepsis for ICU patients using Gradient Boosted Tree. In: IEEE INTERNATIONAL CONFERENCE ON AUTOMATIC CONTROL AND

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

INTELLIGENT SYSTEMS (I2CACIS), 2022. **Proceedings [...]**. Piscataway: IEEE, 2022. p. 78–83. DOI: https://doi.org/10.1109/I2CACIS54679.2022.9815467.

YUN, Hyoungju *et al.* Enhancement in performance of septic shock prediction using National Early Warning Score, initial triage information, and machine learning analysis. **The Journal of Emergency Medicine**, v. 61, n. 1, p. 1–11, jul. 2021. DOI: https://doi.org/10.1016/j.jemermed.2021.01.038.

ZHANG, Guyu *et al.* Predicting sepsis in-hospital mortality with machine learning: a multi-center study using clinical and inflammatory biomarkers. **European Journal of Medical Research**, v. 29, n. 1, art. 156, 2024. DOI: https://doi.org/10.1186/s40001-024-01447-x.

ZHANG, Yaru *et al.* Machine learning for the prediction of sepsis-related death: a systematic review and meta-analysis. **BMC Medical Informatics and Decision Making**, v. 23, art. 283, 2023. DOI: https://doi.org/10.1186/s12911-023-02383-1.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.