

REVISÃO DE LITERATURA IDENTIFICADO

INTELIGÊNCIA ARTIFICIAL NA PREDIÇÃO PRECOCE DE SEPSE EM AMBIENTE HOSPITALAR: UMA REVISÃO INTEGRATIVA DA LITERATURA

ARTIFICIAL INTELLIGENCE IN THE EARLY PREDICTION OF SEPSIS IN THE HOSPITAL SETTING: AN INTEGRATIVE LITERATURE REVIEW

Larissa Leite de Souza Leão^{1*}, Gabriel Henrique Carvalho Leite Romeiro¹, Vitor Comassetto Paes¹, Jussara Guimarães da Rocha Lima¹, Joana Santos Cesar de Carvalho¹, Reinaldo dos Santos Moura²

¹Acadêmicos de Medicina, Afya Centro Universitário de Maceió (UNIMA); ²Docente do Curso de Medicina, Afya Centro Universitário de Maceió (UNIMA).

*Autora principal. E-mail: larissaleao@alunos.afya.com.br

Eixo: 03 - Tecnologia e Inovação

RESUMO

A sepse é síndrome de disfunção orgânica ameaçadora à vida associada a mortalidade hospitalar elevada quando o reconhecimento e o tratamento são tardios. Modelos de inteligência artificial (IA) e aprendizado de máquina vêm sendo aplicados a dados de prontuário eletrônico e sinais vitais para prever sepse antes da manifestação clínica plena, por meio de sistemas de alerta precoce. Objetivo: sintetizar as evidências disponíveis sobre desempenho técnico, implementação em ambiente real e barreiras clínicas desses sistemas. Método: revisão integrativa da literatura, com busca em PubMed/MEDLINE, JAMIA, Nature Medicine, NEJM AI e PLOS Digital Health, priorizando publicações de 2020 a 2025. Resultados: revisões sistemáticas e metanálises mostram acurácia diagnóstica (AUROC) entre 0,68 e 0,99 conforme o cenário assistencial, com modelos de aprendizado profundo e de conjunto superando escores clínicos tradicionais como SOFA e qSOFA. Sistemas implementados em ambiente real, como o TREWS e o Sepsis Watch, associam-se a redução de mortalidade quando o alerta é confirmado rapidamente pelo clínico, mas modelos amplamente adotados, como o Epic Sepsis Model, mostraram baixa sensibilidade e alta taxa de alarme falso em validação externa, gerando fadiga de alerta. Conclusão: a IA apresenta potencial técnico real para predição precoce de sepse, mas seu benefício clínico depende de validação externa robusta, calibração cuidadosa e gestão ativa da fadiga de alerta, não apenas do desempenho estatístico do modelo isoladamente.

Palavras-chave: Inteligência Artificial; Sepse; Diagnóstico Precoce; Aprendizado de Máquina; Sistemas de Alerta.

ABSTRACT

Sepsis is a life-threatening organ dysfunction syndrome associated with high hospital mortality when recognition and treatment are delayed. Artificial intelligence (AI) and machine learning models have been applied to electronic health record data and vital signs to predict sepsis before full clinical manifestation, through early warning systems. Objective: to synthesize available evidence on the technical performance, real-world implementation, and clinical barriers of these systems. Methods: integrative literature review, with searches in PubMed/MEDLINE, JAMIA,

Nature Medicine, NEJM AI, and PLOS Digital Health, prioritizing publications from 2020 to 2025. Results: systematic reviews and meta-analyses show diagnostic accuracy (AUROC) between 0.68 and 0.99 depending on the care setting, with deep learning and ensemble models outperforming traditional clinical scores such as SOFA and qSOFA. Systems implemented in real-world settings, such as TREWS and Sepsis Watch, are associated with reduced mortality when the alert is promptly confirmed by the clinician, but widely adopted models, such as the Epic Sepsis Model, showed low sensitivity and a high false alarm rate in external validation, generating alert fatigue. Conclusion: AI shows real technical potential for early sepsis prediction, but its clinical benefit depends on robust external validation, careful calibration, and active management of alert fatigue, not solely on the model's statistical performance.

Keywords: Artificial Intelligence; Sepsis; Early Diagnosis; Machine Learning; Clinical Alarms.

1. INTRODUÇÃO

A sepse permanece causa importante de mortalidade hospitalar em todo o mundo, e o intervalo entre a instalação do quadro e o início da antibioticoterapia e do controle de foco é determinante direto do desfecho clínico. Modelos de aprendizado de máquina aplicados a dados de prontuário eletrônico e sinais vitais vêm sendo explorados como sistemas de alerta precoce capazes de identificar sepse antes da manifestação clínica plena, com acurácia diagnóstica relatada entre AUROC 0,68 e 0,99 conforme o cenário assistencial (UTI, enfermaria ou pronto-socorro)¹.

Revisões sistemáticas subsequentes confirmaram essa tendência: modelos de aprendizado profundo superam de forma consistente escores clínicos tradicionais na predição precoce de sepse em UTI². Metanálise em rede recente, reunindo 256 modelos derivados de 73 estudos e mais de 450 mil pacientes com sepse, encontrou área sob a curva agregada de 0,825 para modelos de aprendizado de máquina, contra 0,667 do escore SOFA e 0,612 do qSOFA³. Em linha semelhante, revisão sistemática com metanálise de 23 estudos e mais de 4,3 milhões de pacientes identificou Random Forest e XGBoost como os algoritmos de maior acurácia para predição de sepse em ambiente de UTI⁴.

Apesar desse desempenho técnico consistente em bancos de dados retrospectivos, a tradução desses modelos para a prática clínica real enfrenta desafios distintos: sistemas já implementados em larga escala mostraram tanto ganhos reais de desfecho quanto fracassos relevantes de validação externa, evidenciando lacuna entre desempenho estatístico isolado e utilidade clínica efetiva. Há, portanto, necessidade de sintetizar criticamente não apenas a acurácia relatada desses modelos, mas também sua experiência de implementação em ambiente hospitalar real.

2. OBJETIVO

Sintetizar as evidências científicas disponíveis sobre o desempenho técnico, a implementação em ambiente hospitalar real e as principais barreiras clínicas e operacionais de sistemas de inteligência artificial para predição precoce de sepse.

3. MÉTODO

Trata-se de revisão integrativa da literatura, de caráter narrativo, dada a heterogeneidade de desenhos de estudo (metanálises de acurácia diagnóstica, estudos de implementação prospectiva e estudos qualitativos) incluídos. A busca foi realizada em agosto de 2026 nas bases PubMed/MEDLINE, Journal of the American Medical Informatics Association (JAMIA), Nature Medicine, NEJM AI, PLOS Digital Health e NPJ Digital Medicine, combinando os descritores “artificial intelligence”, “machine learning”, “sepsis”, “early warning system”, “prediction model” e “alert fatigue”, priorizando publicações de 2020 a 2025, sem exclusão de estudo anterior considerado fundamental.

Foram incluídos estudos primários, revisões sistemáticas, metanálises e estudos de implementação/validação externa que avaliassem sistemas de IA para predição ou alerta precoce de sepse em ambiente hospitalar adulto. Foram excluídos estudos sem relação direta com predição de sepse por IA, editoriais sem dado original e conteúdos sem revisão por pares. Não foi aplicado protocolo PRISMA formal nem cálculo de escore de qualidade metodológica por estudo, dada a natureza panorâmica desta revisão, limitação declarada.

Como se trata de revisão de literatura, sem coleta direta de dados de pacientes, o presente trabalho está dispensado de submissão a Comitê de Ética em Pesquisa, conforme item 4.3 do edital do 6º AFYA Global Meeting, que exige parecer ético apenas para estudos com envolvimento direto ou indireto de seres humanos.

4. RESULTADOS

Desempenho técnico dos modelos preditivos

Os achados de acurácia diagnóstica descritos nas metanálises acima¹⁻⁴ situam consistentemente os modelos de aprendizado de máquina, sobretudo redes neurais e métodos de conjunto (ensemble), acima dos escores clínicos tradicionais. Esse desempenho, contudo, é predominantemente medido em bancos de dados retrospectivos, com validação externa ainda menos frequente do que a validação interna, padrão que se repete nos demais campos de IA em saúde.

Sistemas implementados em ambiente hospitalar real

O TREWS (Targeted Real-time Early Warning System) foi avaliado prospectivamente em cinco hospitais, monitorando mais de 590 mil pacientes; entre os identificados com sepse pelo alerta antes do início da antibioticoterapia, a confirmação do alerta pelo clínico em até três horas associou-se a redução absoluta de 3,3 pontos percentuais na mortalidade hospitalar⁵. O Sepsis Watch, primeiro modelo de aprendizado profundo integrado à rotina clínica nos Estados Unidos, opera de forma contínua no sistema Duke Health desde 2018, com previsões horárias e desempenho reportado de até 88% de sensibilidade em baixas taxas de falso positivo⁶; validação externa multicêntrica recente confirmou desempenho robusto (AUROC entre 0,906 e 0,960) em novos contextos geográficos e demográficos⁷.

Em contraste, o Epic Sepsis Model, um dos sistemas proprietários mais amplamente implementados nos Estados Unidos, mostrou em validação externa independente sensibilidade baixa: identificou apenas 7% dos casos de sepse não detectados clinicamente e deixou de sinalizar 67% dos pacientes com sepse, apesar de gerar alertas em 18% de todas as internações⁸. Mais recentemente, o Sepsis ImmunoScore tornou-se o primeiro dispositivo de IA/ML para previsão de sepse autorizado pela FDA norte-americana, com AUC de 0,81 em coorte de validação externa multicêntrica⁹, marco regulatório relevante para o campo, mas que não elimina a necessidade de validação local antes da adoção institucional.

Alarme falso e fadiga de alerta

A discrepância entre desempenho relatado e desempenho real observada no Epic Sepsis Model⁸ ilustra um problema central e bem documentado neste campo: a alta taxa de alarme falso gera fadiga de alerta, reduzindo a responsividade clínica e, paradoxalmente, podendo atrasar a intervenção que o sistema deveria antecipar. Modelos mais recentes, como o SepsisAI, foram desenhados especificamente com foco em minimizar alarmes falsos, priorizando alta especificidade em detrimento de sensibilidade máxima, como estratégia direta de mitigação da fadiga de alerta¹⁰. Revisão sistemática que aplicou o framework SALIENT de implementação de IA clínica identificou a fadiga de alerta e a desconfiança do clínico como duas das principais barreiras à adoção efetiva desses sistemas, mesmo entre aqueles com desempenho estatístico favorável¹¹.

Confiança do clínico e barreiras de implementação

Estudo qualitativo conduzido em hospitais da Inglaterra identificou que a confiança do clínico nos alertas digitais de sepse depende fortemente de fatores organizacionais, como quadro de

pessoal, treinamento da equipe e envolvimento dos profissionais na otimização do sistema, mais do que da acurácia estatística isolada do modelo¹². Esse achado reforça que a barreira à adoção de IA não é apenas técnica, é também de governança clínica e integração ao fluxo de trabalho.

No contexto brasileiro, estudo em unidade de terapia intensiva de hospital público de referência, com dados de 659 pacientes, testou cinco algoritmos de aprendizado de máquina para predição de sepse, com Random Forest e Gradient Boosting apresentando os melhores desempenhos (F1-score 0,98; acurácia 98,48%)¹³. Trata-se, contudo, de estudo unicêntrico, sem validação externa multicêntrica nem avaliação de implementação em fluxo clínico real, o que limita a extrapolação para outros serviços do Sistema Único de Saúde.

5. CONCLUSÃO

A literatura disponível indica que modelos de inteligência artificial apresentam desempenho técnico consistentemente superior a escores clínicos tradicionais na predição precoce de sepse, e que sistemas implementados prospectivamente, como o TREWS e o Sepsis Watch, associam-se a desfechos clínicos favoráveis quando o alerta é confirmado rapidamente pelo profissional de saúde.

Contudo, esta revisão também evidencia que desempenho estatístico elevado não garante, isoladamente, utilidade clínica: o caso do Epic Sepsis Model demonstra que um sistema amplamente adotado pode apresentar baixa sensibilidade e gerar fadiga de alerta em validação externa. A implementação responsável de IA para predição de sepse depende de validação externa robusta, calibração cuidadosa do limiar de alerta e governança clínica que envolva ativamente a equipe assistencial, sendo a validação em contexto brasileiro, incluindo no Sistema Único de Saúde, ainda uma lacuna a ser preenchida por estudos multicêntricos futuros.

REFERÊNCIAS

1. Fleuren LM, Klausch TLT, Zwager CL, Schoonmade LJ, Guo T, Roggeveen LF, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive Care Med*. 2020;46(3):383-400.
2. Moor M, Rieck B, Horn M, Jutzeler CR, Borgwardt K. Early prediction of sepsis in the ICU using machine learning: a systematic review. *Front Med (Lausanne)*. 2021;8:607952.
3. Yadgarov MY, Landoni G, Berikashvili LB, Polyakov PA, Kadantseva KK, Smirnova AV, et al. Early detection of sepsis using machine learning algorithms: a systematic review and network meta-analysis. *Front Med (Lausanne)*. 2024;11:1491358.

4. Yang Z, Cui X, Song Z. Predicting sepsis onset in ICU using machine learning models: a systematic review and meta-analysis. *BMC Infect Dis.* 2023;23:635.
5. Adams R, Henry KE, Sridharan A, Soleimani H, Zhan A, Rawat N, 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(7):1455-1460.
6. Sendak MP, Ratliff W, Sarro D, Alderton E, Futoma J, Gao M, et al. Real-World Integration of a Sepsis Deep Learning Technology Into Routine Clinical Care: Implementation Study. *JMIR Med Inform.* 2020;8(7):e15182.
7. Valan B, Prakash A, Ratliff W, Gao M, Muthya S, Thomas A, et al. Evaluating sepsis watch generalizability through multisite external validation of a sepsis machine learning model. *NPJ Digit Med.* 2025;8:350.
8. Wong A, Otles E, Donnelly JP, Krumm A, McCullough J, DeTroyer-Cooley O, et al. External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients. *JAMA Intern Med.* 2021;181(8):1065-1070.
9. Bhargava A, Lopez-Espina C, Schmalz L, Khan S, et al. FDA-Authorized AI/ML Tool for Sepsis Prediction: Development and Validation. *NEJM AI.* 2024. doi:10.1056/AIoa2400867.
10. Gupta A, Chauhan R, G S, Shreekumar A. Improving sepsis prediction in intensive care with SepsisAI: A clinical decision support system with a focus on minimizing false alarms. *PLOS Digit Health.* 2024;3(8):e0000569.
11. van der Vegt AH, Scott IA, Dermawan K, Schnetler RJ, Kalke VR, Lane PJ. Deployment of machine learning algorithms to predict sepsis: systematic review and application of the SALIENT clinical AI implementation framework. *J Am Med Inform Assoc.* 2023;30(7):1349-1361.
12. Lazzarino R, Borek AJ, Honeyford K, Welch J, Brent AJ, Kinderlerer A, et al. Views and Uses of Sepsis Digital Alerts in National Health Service Trusts in England: Qualitative Study With Health Care Professionals. *JMIR Hum Factors.* 2024;11(1):e56949.
13. Bezerra ADC, Sales da Silva LM, Maia Pinto FJ, Oliveira Gonçalves RV. Predição de sepse em pacientes de unidade de terapia intensiva: desenvolvimento e avaliação de algoritmos de inteligência artificial. *Rev Ciênc Méd Biol.* 2025;24(1):111-118.