

UNIVERSIDADE FEDERAL DO AMAZONAS
PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU EM CIÊNCIAS DA SAÚDE

KERZIA FERREIRA LIMA DINIZ

**EFETIVIDADE DO *NATIONAL EARLY WARNING SCORE* (NEWS 2) NA DETECÇÃO
PRECOCE DA DETERIORAÇÃO CLÍNICA DE PACIENTES HOSPITALIZADOS:
REVISÃO SISTEMÁTICA COM METANÁLISE**

MANAUS
2025

UNIVERSIDADE FEDERAL DO AMAZONAS

PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU EM CIÊNCIAS DA SAÚDE

KERZIA FERREIRA LIMA DINIZ

**EFETIVIDADE DO *NATIONAL EARLY WARNING SCORE* (NEWS 2) NA DETECÇÃO
PRECOCE DA DETERIORAÇÃO CLÍNICA DE PACIENTES HOSPITALIZADOS:
REVISÃO SISTEMÁTICA COM METANÁLISE**

Dissertação apresentada como requisito para a defesa de título de Mestre em Ciências da Saúde, no Programa de Pós-Graduação Stricto Sensu em Ciências da Saúde da Universidade Federal do Amazonas, nível Mestrado Acadêmico. Linha de Pesquisa: Saúde baseada em evidências.

Orientadora: Prof.^a Dr.^a Roberta Lins Gonçalves

MANAUS
2025

Ficha Catalográfica

Elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

-
- D585e Diniz, Kerzia Ferreira Lima
 Efetividade do national early warning score (NEWS 2) na detecção precoce da deterioração clínica de pacientes hospitalizados: revisão sistemática com metanálise / Kerzia Ferreira Lima Diniz. - 2025.
 80 f. : il., color. ; 31 cm.
- Orientador(a): Roberta Lins Gonçalves .
 Dissertação (mestrado) - Universidade Federal do Amazonas, Programa de Pós-Graduação em Ciências da Saúde, Manaus, 2025.
1. Pontuação de alerta precoce. 2. Deterioração clínica. 3. Intervenção médica precoce. 4. Previsão de risco. 5. Suporte à decisão clínica. I. Gonçalves, Roberta Lins.II. Universidade Federal do Amazonas. Programa de Pós-Graduação em Ciências da Saúde. III. Título
-

KERZIA FERREIRA LIMA DINIZ

**EFETIVIDADE DO *NATIONAL EARLY WARNING SCORE* (NEWS 2) NA DETECÇÃO
PRECOCE DA DETERIORAÇÃO CLÍNICA DE PACIENTES HOSPITALIZADOS:
REVISÃO SISTEMÁTICA COM METANÁLISE**

Dissertação apresentada como requisito para a defesa de título de Mestre em Ciências da Saúde, no Programa de Pós-Graduação Stricto Sensu em Ciências da Saúde da Universidade Federal do Amazonas, nível Mestrado Acadêmico. Linha de Pesquisa: Saúde baseada em evidências.

Aprovada em: 09/06/2025

BANCA EXAMINADORA

Prof.^a Dr.^a Roberta Lins Gonçalves - Orientadora; Presidente da banca – Membro Interno titular - PPGCIS/UFAM

Dra^a Narjara de Oliveira Boechat – Membro Titular Externo - Hospital Adventista

Prof^a Dr^a Lysien Ivania Zambrano – Membro Titular – Universidad Nacional de Honduras - UNAH

Suplentes

Dr. Luiz Francisco Rocha e Silva – Hospital Adventista de Manaus - HAM

Dr. Francisco Mateus João – Universidade Estadual do Amazonas - UEA

AGRADECIMENTOS

Primeiramente, agradeço a Deus, que tem sido minha força e minha sustentação. Mesmo nos momentos de incerteza e quando me sentia incapaz, Ele sempre esteve ao meu lado, me guiando e me dizendo para continuar, me dando forças para não desistir. Quando meu mundo parecia desabar, Ele me dizia: "Continue, eu estou aqui e sempre estarei". Sou imensamente grata por sua presença constante em minha vida.

Agradeço à minha colega de trabalho Kelcia, que, em uma conversa simples, me incentivou a seguir o caminho do mestrado. Foi ela quem me impulsionou a acreditar que eu seria capaz. Também agradeço ao coordenador do Centro de Ensino e Pesquisa do Hospital Adventista de Manaus, Dr. Luís Rocha, que, com suas palavras e apoio, fez com que eu tomasse a decisão de seguir em frente.

Sou grata pela oportunidade de dividir essa jornada com minha parceira de curso, Ericka Campos. Sua amizade e parceria foram fundamentais. Agradeço, também, a todas as pessoas que Deus colocou em meu caminho, pessoas que andaram comigo a segunda milha, me ajudando a superar os desafios e celebrando as conquistas ao meu lado.

Agradeço profundamente à minha família, em especial à minha filha Isadora, por todo o amor, compreensão e apoio incondicional. Mesmo nas ausências momentâneas, ela sempre esteve ao meu lado, com orgulho e carinho, me dando forças para seguir em frente.

A minha orientadora, Professora Roberta Lins, que sempre acreditou no meu potencial e me desafiou a buscar a excelência. Sua postura exigente, que muitas vezes me puxava a "orelha" e me instiga a ir além, foi fundamental para o meu crescimento acadêmico. Embora fosse rigorosa quando necessário, sua compreensão e apoio estavam sempre presentes, o que fez toda a diferença em minha jornada. Sou imensamente grata pela sua pelo seu compromisso com o meu desenvolvimento.

Agradeço aos meus colegas de trabalho, que sempre torceram por mim, me incentivando e acreditando em minha capacidade. Agradeço também à Tayla, secretária da Gerência de

Enfermagem, por sua ajuda constante sempre que precisei. Sua colaboração e dedicação foram de grande importância ao longo dessa jornada.

Agradeço aos meus colegas de trabalho, que sempre torceram por mim, me incentivando e acreditando em minha capacidade.

Agradeço profundamente ao Mateus Araújo, cuja simplicidade e inteligência tornaram tudo mais possível. Sua dedicação, comprometimento e disposição para contribuir foram fundamentais para o sucesso da nossa pesquisa. Ele sempre esteve presente, acreditando e apoiando quando solicitado, tornando cada desafio mais fácil de ser enfrentado. Sua participação foi essencial e, sem dúvida, a parceria com ele fez toda a diferença nesse processo.

A Jacimara Gomes, para mim, é a pessoa que mais acreditou em mim. Seu apoio incondicional e suas palavras de afirmação foram fontes de força e motivação nos momentos mais difíceis. A cada dia de pesquisa e estudo, ela esteve ao meu lado, me fazendo sentir forte e capaz, mesmo quando eu duvidava de mim mesma. Foram dias e noites de trabalho conjunto que se tornaram mais leves e prazerosos pela sua amizade, carinho e confiança. Obrigada, Jacimara, por ser essa pessoa tão especial na minha vida e na minha trajetória acadêmica.

Agradeço também aos professores e colegas do Programa de Pós-Graduação em Ciências da Saúde (PPGCIS) pela troca de conhecimento e pela convivência enriquecedora ao longo do curso. Um agradecimento especial a Denilsi e à Taina, cujos apoios e palavras de incentivo foram de grande importância para minha jornada na UFAM.

Agradeço ao Dr. Boechat, que nos incentivou a seguir em frente e, com sua orientação, nos apresentou à Professora Roberta Lins, que se tornou uma peça-chave nesse processo.

Agradeço também aos professores e colegas do Programa de Pós-Graduação em Ciências da Saúde (PPGCIS) pela troca de conhecimento e pela convivência enriquecedora ao longo do curso.

Agradeço ao Hospital Adventista, meu local de trabalho, por me permitir cursar o mestrado, equilibrando as demandas do trabalho e os desafios acadêmicos. A liberdade e o apoio institucional foram cruciais para o meu desenvolvimento como pesquisadora.

Por fim, agradeço à Universidade Federal do Amazonas (UFAM) e ao Programa de Pós-Graduação por me proporcionarem esta oportunidade de crescimento acadêmico e profissional. Sem o apoio de todos os envolvidos, este trabalho não teria sido possível.

A todos, o meu mais sincero e profundo agradecimento.

RESUMO

Introdução: A detecção precoce da piora clínica de pacientes hospitalizados é essencial para reduzir complicações e mortalidade. Contudo, apesar de existirem ferramentas para triagem hospitalar, como o *National Early Warning Score* (NEWS 2), não existe uma síntese abrangente que analise sua efetividade e precisão prognóstica em diferentes populações.

Objetivo: Avaliar a efetividade do NEWS 2 na detecção precoce da deterioração clínica e sua acurácia na predição de eventos adversos, comparando-o com outras escalas prognósticas.

Método: Realizamos uma revisão sistemática com metanálise seguindo as diretrizes PRISMA 2020. Conduzimos a busca nas bases PubMed, *Web of Science* (WOS), Embase, LILACS, Google Scholar e literatura cinzenta para identificar estudos publicados entre 2017 e 2024, em inglês e português. Consideramos elegíveis estudos observacionais e ensaios clínicos que avaliaram a efetividade do NEWS 2 na detecção precoce da deterioração clínica de pacientes adultos hospitalizados, utilizando os desfechos predição de eventos adversos (óbito, necessidade de UTI e ventilação mecânica). Excluímos estudos transversais, relatos de casos, revisões narrativas e estudos sem dados suficientes para a extração. Avaliamos o risco de viés com a ferramenta Joanna Briggs *Institute* e a heterogeneidade pelo I^2 de Higgins e Thompson. Utilizamos o modelo de efeitos aleatórios para a metanálise, considerando intervalos de confiança (IC 95%) para estimativas de sensibilidade e especificidade, e para comparar a eficácia relativa das diferentes escalas prognósticas utilizamos o P-score.

Resultado: Nós incluímos 10 estudos totalizando 339.216 pacientes. Observamos heterogeneidade alta para internação em UTI ($I^2 = 79,88\%$) e mortalidade por intercorrências ($I^2 = 75,2\%$), e moderada para mortalidade hospitalar ($I^2 = 45,2\%$). O NEWS 2 apresentou sensibilidade combinada: 88% (IC 95%: 85%-91%), especificidade combinada: 75% (IC 95%: 70%-80%), área sob a curva (AUC) ROC: 0,85 (IC 95%: 0,82-0,88), taxa de internação na UTI: 20% (IC 95%: 9,0%-30,7%), taxa de mortalidade hospitalar: 9,2% (IC 95%: 4,3%-14,1%). Os resultados apontaram que escalas como SMART-COP e CRB-65 apresentaram os maiores P-scores, evidenciando melhor performance em desfechos como necessidade de ventilação invasiva e mortalidade hospitalar. Em seguida, vieram as escalas NEWS 2 e PSI, que demonstraram desempenho intermediário. Escalas como MEWS e qSOFA apresentaram os menores P-scores, refletindo menor precisão preditiva nos contextos clínicos analisados.

Conclusão: Os achados confirmaram a alta sensibilidade do NEWS 2, especialmente para condições críticas como sepse e doenças respiratórias agudas, o que contribui para a redução da mortalidade hospitalar e melhora no manejo inicial. No entanto, a especificidade foi inferior a outras escalas, o que pode gerar alarmes falsos e intervenções desnecessárias. A análise da curva AUC ROC mostrou um desempenho razoável do NEWS 2 na predição de mortalidade, necessidade de ventilação e internação prolongada, mas inferior a escalas especializadas como TRIAGE e DECAF. Apesar das limitações, o NEWS 2 permanece uma ferramenta útil, especialmente quando combinado com outras escalas.

Limitações da Pesquisa: A heterogeneidade entre as populações e cenários clínicos pode ter influenciado os resultados da metanálise.

Registro e Financiamento: Esta revisão sistemática foi registrada no PROSPERO (CRD420223530). O estudo recebeu apoio da Fundação de Amparo à Pesquisa do Estado do

Amazonas (FAPEAM) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Palavras-chave: Pontuação de Alerta Precoce; Deterioração Clínica; Intervenção Médica Precoce; Previsão de Risco; Suporte à Decisão Clínica.

ABSTRACT

Introduction: Early detection of clinical deterioration in hospitalized patients is essential to reduce complications and mortality. The National Early Warning Score (NEWS 2) is widely used for hospital triage, but the literature lacks a comprehensive synthesis of its effectiveness and prognostic accuracy in different populations.

Objective: To evaluate the effectiveness of NEWS 2 in the early detection of clinical deterioration and its accuracy in predicting adverse events, comparing it with other prognostic scales.

Method: We conducted a systematic review with meta-analysis following the PRISMA 2020 guidelines. We searched PubMed, Web of Science (WOS), Embase, LILACS, Google Scholar, and gray literature to identify studies published between 2017 and 2024 in English and Portuguese. We considered eligible observational studies and clinical trials that evaluated the effectiveness of NEWS 2 in the early detection of clinical deterioration in hospitalized adult patients, using the outcomes predicting adverse events (death, need for ICU, and mechanical ventilation). We excluded case reports, narrative reviews, and studies with insufficient data for extraction. We assessed the risk of bias using the Joanna Briggs Institute tool and heterogeneity using Higgins and Thompson's I^2 . We used the random-effects model for meta-analysis, considering confidence intervals (95% CI) for sensitivity and specificity estimates, and to compare the relative efficacy of the different prognostic scales, we used the P-score.

Results: We included 10 studies totaling 339,216 patients. High heterogeneity was observed for ICU admission ($I^2 = 79.88\%$) and mortality due to adverse events ($I^2 = 75.2\%$), and moderate heterogeneity for in-hospital mortality ($I^2 = 45.2\%$). NEWS 2 presented pooled sensitivity: 88% (95% CI: 85%-91%), pooled specificity: 75% (95% CI: 70%-80%), area under the ROC curve (AUC): 0.85 (95% CI: 0.82-0.88), ICU admission rate: 20% (95% CI: 9.0%-30.7%), in-hospital mortality rate: 9.2% (95% CI: 4.3%-14.1%). The results showed that scales such as SMART-COP and CRB-65 presented the highest P-scores, demonstrating better performance in outcomes such as the need for invasive ventilation and in-hospital mortality. Next came the NEWS 2 and PSI scales, which demonstrated intermediate performance. Scales such as MEWS and qSOFA presented the lowest P-scores, reflecting lower predictive accuracy in the clinical contexts analyzed. Conclusion: The findings confirmed the high sensitivity of NEWS 2, especially for critical conditions such as sepsis and acute respiratory diseases, which reduces hospital mortality and improves initial management. However, specificity was lower than other scales, which may generate false alarms and unnecessary interventions. The AUC ROC curve analysis showed a reasonable performance of NEWS 2 in predicting mortality, need for ventilation, and prolonged hospitalization, but lower than specialized scales such as TRIAGE and DECAF. Despite the limitations, NEWS 2 remains a valuable tool, especially when combined with other scales. Research Limitations: Heterogeneity between populations and clinical settings may have influenced the meta-analysis results.

Registration and Funding: This systematic review was registered with PROSPERO (CRD420223530). The Amazonas State Research Support Foundation (FAPEAM) and the Coordination for the Improvement of Higher Education Personnel (CAPES) supported the study.

Keywords: Early Warning Score; Clinical Deterioration; Early Medical Intervention; Risk Prediction; Clinical Decision Support.

LISTA DE ILUSTRAÇÕES

QUADRO

Quadro 1. <i>National Early Warning Score</i> - Versão Brasileira -----	18
Quadro 2. Descrição da estratégia PEOTS para formulação da pergunta de pesquisa-----	22
Quadro 3. Estratégia PEOTS e termos MeSH/Emtree/DeSC -----	25

LISTA DE ABREVIATURAS E SIGLAS

ARDS	Síndrome do Desconforto Respiratório Agudo
CNPq	Conselho Nacional de desenvolvimento Científico e Tecnológico
DeCS	Descritores em Ciências da Saúde
EWS	<i>Early Warning Score</i>
FAPEAM	Fundação de Amparo à pesquisa do Estado do Amazonas
JBÍ	<i>Joanna Briggs Institute</i>
LILACS	Literatura Latino-americana e do Caribe em Ciências da Saúde
MeSH	<i>Medical Subject Heading</i>
NEWS	<i>National Early Warning Score</i>
NHS	National Health Service (Reino Unido)
PEOT/PIOT	<i>Population, Exposition, Intervention, Outcome, Time</i>
PPG	Programa de Pós-graduação
PPGCIS/UFAM	Programa de Pós-graduação em Ciências da Saúde/ Universidade Federal do Amazonas
PRISMA	<i>Preferred Reporting Items for Systematic Reviews and Meta-Analyses</i>
PROSPERO	<i>International Prospective Register of Systematic Reviews</i>
POSGRADE	Programa de apoio a pós-graduação stricto sensu
QCRI	<i>Qatar computing and Research Institute</i>
qSOFA	<i>Quick Sequential Organ Failure Assessment</i>
sPESI	<i>Simplified Pulmonary Embolism Severity Index</i>
REMS	<i>Rapid Emergency Medicine Score</i>
UTI	<i>Unidade de terapia intensiva</i>
WMD	<i>Weighted Mean Difference</i>
WOS	<i>Web of Science</i>

SUMÁRIO

1 INTRODUÇÃO	16
1.1 JUSTIFICATIVA	18
1.2 OBJETIVO	19
1.3 HIPÓTESE DA PESQUISA	20
2 ESTRATÉGIA PARA ELABORAÇÃO DA PESQUISA	20
2.1 FASES DA REVISÃO	20
2.2 CONSIDERAÇÕES ÉTICA DA PESQUISA	20
2.3 FONTES DE FINANCIAMENTO	20
2.4 PROTOCOLO E REGISTRO	21
3 MÉTODO	21
3.1 DESENHO DO ESTUDO	21
3.2 PERGUNTA DA PESQUISA	21
3.3 DESFECHOS	22
3.3.1 Desfecho primário	22
3.3.2 Desfechos secundários	22
3.4 VARIÁVEIS ANALISADAS	22
3.4.1 Variável Independente (preditora)	22
3.4.2 Variáveis dependentes (desfecho)	22
3.5 POPULAÇÃO DE ESTUDO	23
3.6 CRITÉRIOS DE ELEGIBILIDADE	23
3.6.1 Critérios de Inclusão	23
3.6.2 Critérios de Exclusão	23
3.7 ESTRATÉGIA DE PESQUISA	23
3.7.1 Qualificação dos revisores	23
3.7.2 Estratégia de elaboração a pesquisa	23
3.7.3 Estratégia de busca e fontes de informação	24
3.7.4 Avaliação da heterogeneidade entre os estudos	26
3.7.5 Análise de risco de viés e qualidade metodológica dos estudos incluídos	26

3.8 SELEÇÃO DOS ESTUDOS E EXTRAÇÃO DOS DADOS	26
3.9 GERENCIAMENTO DE DADOS	27
4 RESULTADOS	27
4.1 Artigo científico	31
5 CONCLUSÃO	67
6 REFERÊNCIAS	68

1 INTRODUÇÃO

Garantir cuidados eficazes e seguros é uma prioridade fundamental nos ambientes de saúde e, neste processo, o papel das unidades de internação é crucial (CUTHBERTSON *et al.*, 2007; LEUNG *et al.*, 2016). No entanto, mesmo em pacientes hospitalizados, a detecção precoce da deterioração clínica pode ser um desafio que pode comprometer a intervenção oportuna, e os melhores resultados clínicos (SMITH *et al.*, 2013).

No início do século XX, tornou-se evidente que os pacientes nas enfermarias dos hospitais sofriam danos evitáveis e até mesmo a morte devido à falha na identificação precoce da deterioração clínica (GERRY *et al.*, 2020; HWANG *et al.*, 2023). Para enfrentar esta questão, o primeiro sistema de pontuação de alerta precoce, conhecido como *Early Warning Score* (EWS), foi desenvolvido em 1997 como uma ferramenta à beira do leito para auxiliar os profissionais de saúde na identificação e resposta oportunas à deterioração clínica dos pacientes. A partir daí, várias ferramentas de alerta precoce foram desenvolvidas com o objetivo de aprimorar a vigilância clínica (GERRY *et al.*, 2020; HWANG *et al.*, 2023). Em 2000 o EWS foi adaptado para *Modified Early Warning Score* (MEWS), incorporando parâmetros adicionais e ajustes para diferentes grupos de pacientes. Em 2012, o *Royal College of Physicians* da Inglaterra criou o *National Early Warning Score* (NEWS) (BOULITSAKIS LOGOTHETIS *et al.*, 2023), que foi atualizado em 2017 como NEWS 2, e validado e endossado pelo maior e mais antigo sistema público de saúde do mundo *National Health Service* (NHS) para uso nas unidades de terapia intensiva (UTI) e ambulâncias (BOULITSAKIS LOGOTHETIS *et al.*, 2023; COLOMBO *et al.*, 2021; GERRY *et al.*, 2017; HWANG *et al.*, 2023; ZHANG *et al.*, 2021) 2023).

NEWS 2 é um sistema de pontuação baseado em parâmetros fisiológicos de rotina, como frequência cardíaca, frequência respiratória, saturação de oxigênio, pressão arterial sistólica, nível de consciência e temperatura, que podem facilmente e rapidamente ser obtidos à beira do leito (LALUEZA *et al.*, 2022; MARTINS; PORTO RUTZ; BRAZOROTTO SOARES, 2022; PIRNESKOSKI *et al.*, 2019). Ele oferece vantagens em relação ao seu predecessor, incluindo a capacidade aprimorada de identificar pacientes com sepse, a administração mais segura de suplementação de oxigênio em pacientes que estão enfrentando problemas respiratórios devido ao acúmulo excessivo de dióxido de

carbono no sangue, e a inclusão de avaliação de confusão e delírio como parte da análise do estado de consciência do paciente (BOULITSAKIS LOGOTHETIS *et al.*, 2023).

A efetividade do NEWS 2 em unidades de internação tem sido objeto de crescente interesse, pois pode impactar na qualidade do cuidado e assistência prestada (CANDEL *et al.*, 2023; COLOMBO *et al.*, 2021; MARTINS; PORTO RUTZ; BRAZOROTTO SOARES, 2022; SILCOCK *et al.*, 2015; SMITH *et al.*, 2013). No entanto, sua efetividade pode variar conforme fatores contextuais, como a qualificação da equipe de saúde, o ambiente clínico em que é aplicado e a disponibilidade de recursos institucionais (CUMMINGS *et al.*, 2023). Apesar de existirem estudos avaliando seu desempenho em cenários específicos, ainda não há uma síntese abrangente que consolide as evidências sobre sua aplicabilidade em diferentes contextos clínicos. Diante disso, o presente estudo teve como objetivo realizar uma revisão sistemática com metanálise para avaliar a efetividade do NEWS 2 na detecção precoce da deterioração clínica em pacientes hospitalizados. O quadro 1 ilustra os parâmetros avaliados no NEWS 2 de acordo com a adaptação para a versão brasileira.

Quadro 1. National Early Warning Score 2 (News 2) - Versão Brasileira

Parâmetros Fisiológicos	Pontuação						
	3	2	1	0	1	2	3
Frequência Respiratória (/minuto)	< 08		9 - 11	12 - 20		21 - 24	> 25
SatO ₂ % (escala 1)	< 91	92 - 93	94 - 95	>96			
SatO ₂ % (escala 2)	< 83	84 - 85	86 - 87	88 - 92 >93 ar ambiente	93 - 94 Com O ₂	95 - 96 Com O ₂	> 97 Com O ₂
Ar ambiente ou uso de O ₂		O ₂		Ar Ambiente			
Pressão Arterial Sistólica (mmHg)	< 90	91 - 100	101 - 110	111 - 219			> 220
Pulso (/minuto)	<40		41-50	51-90	91-110	111-130	>131
Nível de Consciência				Alerta			Confusão aguda Resposta a voz ou dor irresponsivo
Temperatura (°C)	< 35		35,1 - 36	36,1 - 38	38,1 - 39	> 39,1	

Fonte: Adaptado de *National Early Warning Score 2 - NEWS 2 - versão brasileira*. Porto Alegre/RS, 2018.

1.1 JUSTIFICATIVA

A proposta desta revisão sistemática sobre a efetividade do NEWS 2 na identificação precoce da deterioração clínica em pacientes hospitalizados parte da necessidade de compreender sua real aplicação em contextos clínicos diversos. Diferente da eficácia, geralmente avaliada em condições ideais de pesquisa, a efetividade refere-se ao desempenho da ferramenta na prática cotidiana, influenciada por fatores como capacitação das equipes, complexidade dos pacientes e estrutura dos serviços.

No caso de ferramentas para análise de deterioração clínica em pacientes hospitalizados, realizar uma síntese, comparando a efetividade apontada em diferentes estudos em metanálise, é altamente relevante, uma vez que a piora clínica não detectada, continua sendo uma das principais causas de agravamento e desfechos adversos evitáveis em pacientes hospitalizados, como paradas cardiorrespiratórias, admissões não

planejadas em unidades de terapia intensiva e aumento da mortalidade. Portanto, a identificação precoce desses eventos é essencial para mitigar riscos, garantir a segurança do paciente e direcionar intervenções eficazes.

Embora o NEWS 2 seja amplamente utilizado e recomendado por sistemas de saúde de referência, como o NHS do Reino Unido, ainda existem lacunas relevantes na literatura quanto à sua performance em cenários variados e sua comparação com outras escalas prognósticas. Dessa forma, esta revisão visa consolidar evidências disponíveis, avaliando o desempenho preditivo do NEWS 2 e sua aplicabilidade clínica no mundo real, fornecendo subsídios para decisões baseadas em evidência.

Além de sua relevância clínica, a implementação adequada de ferramentas pode contribuir para o uso mais racional dos recursos em saúde, evitando internações desnecessárias, reduzindo custos e melhorando a qualidade do cuidado. A análise crítica e sistematizada da literatura permitirá não apenas avaliar sua efetividade, mas também identificar limitações metodológicas dos estudos existentes e oportunidades para futuras investigações.

Apesar da existência de estudos de impacto que analisam a eficácia do NEWS 2, ainda não há uma revisão sistemática abrangente e atualizada que integre e avalie criticamente essas evidências. Desta forma, este estudo visa preencher essa lacuna, contribuindo para a base de conhecimento científico, oferecendo uma ferramenta de apoio à tomada de decisão para gestores, pesquisadores e profissionais da linha de frente no cuidado hospitalar.

1.2 OBJETIVO

Avaliar a efetividade e a precisão prognóstica do NEWS 2 na detecção precoce de deterioração clínica entre pacientes adultos hospitalizados em cenários do mundo real, comparando seu desempenho preditivo para eventos adversos com outras escalas prognósticas validadas.

1.3 HIPÓTESE DA PESQUISA

A hipótese deste estudo é que o NEWS 2 é efetivo para a detecção precoce da deterioração clínica de pacientes hospitalizados adultos.

2 ESTRATÉGIA PARA ELABORAÇÃO DA PESQUISA

2.1 FASES DA REVISÃO

- I. Identificação da necessidade de revisão;
- II. Elaboração do projeto de revisão;
- III. Desenvolvimento do protocolo de revisão;
- IV. Registro do protocolo no PROSPERO;
- V. Identificação da pesquisa – busca nas bases de dados;
- VI. Seleção de estudos;
- VII. Análise da qualidade dos estudos selecionados;
- VIII. Extração de dados;
- IX. Síntese de dados;
- X. Relatório e recomendações.

2.2 CONSIDERAÇÕES ÉTICA DA PESQUISA

Não houve conflito de interesses na condução deste estudo. As revisões sistemáticas não requerem aprovação de um comitê de ética em pesquisa, uma vez que são estudos secundários que não ferem a integridade dos participantes.

2.3 FONTES DE FINANCIAMENTO

O estudo teve apoio da Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM) por meio do Programa de Apoio à Pós-Graduação Stricto Sensu (POSGRAD), da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), e da Universidade Federal do Amazonas – UFAM.

2.4 PROTOCOLO E REGISTRO

Para o registro foram utilizadas as diretrizes PRISMA (*Preferred Reporting Items for Systematic Review and Meta-Analysis*) (ANEXO A). O protocolo do estudo detalhado foi submetido e registrado no Registro Prospectivo Internacional de Revisões

Sistemáticas PROSPERO - CRD420223530 e será submetido a publicação em revista científica.

3 MÉTODO

3.1 DESENHO DO ESTUDO

Este estudo foi uma revisão sistemática com metanálise.

3.2 PERGUNTA DA PESQUISA

A pergunta norteadora do estudo foi: O NEWS 2 é efetivo para prever a deterioração clínica precoce em adultos hospitalizados em unidade de clínica médica (quadro 2)?

Quadro 2. Descrição da Estratégia PEOTS para Formulação da Pergunta de Pesquisa

População (Population)	Pacientes adultos em unidades de internação hospitalar da clínica médica
Exposição (Exposure)	Escala NEWS 2
Desfechos (Outcome)	Primário: Capacidade de identificar precocemente a deterioração clínica do paciente. Secundário: Capacidade de prever parada cardiorrespiratória Capacidade de prever transferência para UTI Capacidade de prever óbitos
Tempo (Time)	Estudos publicados a partir de dezembro de 2017
Desenho do estudo (Study design)	Ensaio clínico, estudos de coorte, estudos caso-controle.

Fonte: Autoria dos pesquisadores (2024).

3.3 DESFECHOS

3.3.1 Desfecho primário

Capacidade de identificar precocemente a deterioração clínica do paciente.

3.3.2 Desfechos secundários

- ✓ Capacidade de prever parada cardiorrespiratória
- ✓ Capacidade de prever transferência para UTI
- ✓ Capacidade de prever óbitos

3.4 VARIÁVEIS ANALISADAS

3.4.1 Variável Independente (preditora)

Escore do NEWS 2.

3.4.2 Variáveis dependentes (desfecho)

Identificação precoce da deterioração clínica em pacientes hospitalizados;

As variáveis dependentes foram analisadas pela frequência, percentual ou como variáveis categóricas, conforme estivessem nos estudos primários;

3.5 POPULAÇÃO DE ESTUDO

Pacientes adultos em unidades de internação hospitalar da clínica médica

3.6 CRITÉRIOS DE ELEGIBILIDADE

3.6.1 Critérios de Inclusão

Foram elegíveis para a inclusão, estudos de coorte, estudos caso-controle, e estudos que utilizaram métodos qualitativos de coleta e análise de dados, que abordarem a questão PEOTS (quadro 2), publicados em inglês, espanhol ou português, de 2017 (a partir da atualização do NEWS) a 2024.

3.6.2 Critérios de Exclusão

Foram excluídos estudos que não abordaram a questão de pesquisa, artigos incompletos, resumos, artigos de revisão, opinião de especialistas, editoriais, livros, trabalhos acadêmicos, dissertações, teses, anais de eventos científicos, artigos não disponíveis online.

3.7 ESTRATÉGIA DE PESQUISA

3.7.1 Qualificação dos revisores

A equipe de pesquisa responsável pela realização deste estudo foi composta por pesquisadores da área da saúde, incluindo mestres e doutores, com experiência relevante na área de estudo.

3.7.2 Estratégia de elaboração a pesquisa

A estratégia para a elaboração da pesquisa seguiu-se de uma revisão não sistemática da literatura para aprofundamento do tema e construção da pergunta da pesquisa, com várias discussões com o grupo de pesquisa para a elaboração do protocolo e a partir daí, uma revisão sistemática de acordo com as recomendações PRISMA.

3.7.3 Estratégia de busca e fontes de informação

A seleção e análise dos estudos foi conduzida por 02 revisores independentes utilizando as bases de dados PubMed, Literatura Latino-Americana e do Caribe em Ciências da Saúde (LILACS) via BVS, Excerpta Médica dataBASE (Embase), *Web of Science* (WOS), Google scholar e a literatura cinzenta. A estratégia de busca utilizou os descritores *Medical Subject Heading* (MeSH) e os seus correspondentes dos Descritores em Ciências da Saúde (DeCS): *Patient Acuity; Inpatients; Early Warning Score; Clinical Deterioration; Disease Progression; Hospitalization; Vital Signs; Clinical Evolution; Decision Support Systems, Clinical; Clinical Decision-Making; Early Medical Intervention; Patient Safety; Effectiveness* (txt); e seus sinônimos conforme vocabulário oficial da base. A base de dados embase utiliza termos MeSH/DeCS e sugere termos específicos da base, os *Embase Thesaurus* (EMTREE): *Deterioration; Disease*

Exacerbation; Hospitalization; Vital Sign; Clinical Decision Support System; Clinical Decision Making; Medical Decision Making; Early Intervention; Patient Safety; Clinical Effectiveness. Para a criação da estratégia de busca, os descritores foram combinados entre si por meio dos operadores booleanos “AND” e “OR” para cada uma das bases (APÊNDICE A, B, C, D). Quando possível, os seguintes filtros foram empregados: ano de publicação 2017-2024; idioma de publicação: inglês, português ou espanhol; tipo de estudo: apenas em humanos; adultos (18 + anos); desenho metodológico: ensaios clínicos randomizados, ensaios clínicos e estudos observacionais. O quadro 3 apresenta a descrição dos componentes da estratégia PEOTS e os termos *MeSH/Emtree/DeSC* principais, empregados para compor a estratégia da pesquisa bibliográfica.

Quadro 3. Estratégia PEOTS e Termos Mesh/Entre/Decs

		<i>MeSH</i>	<i>Entree (específicos)</i>	<i>DeCS</i>
P	População	• <i>Patient Acuity</i> • <i>Inpatients</i>	• <i>Patient Acuity</i> • <i>Disease Severity</i> • <i>Hospital Patient</i>	• <i>Gravidade do Paciente</i> • <i>Pacientes Internados</i>
E	Exposição	• <i>Early Warning Score</i>	<i>Early Warning Score</i>	• <i>Escore de Alerta Precoce</i>
O	Desfechos	• <i>Clinical Deterioration</i> • <i>Disease Progression</i> • <i>Hospitalization</i> • <i>Vital Signs</i> • <i>Clinical Evolution</i> • <i>Decision Support Systems, Clinical</i> • <i>Clinical Decision-Making</i> • <i>Early Medical Intervention</i> • <i>Patient Safety</i> • <i>Effectiveness</i>	• <i>Deterioration</i> • <i>Disease Exacerbation</i> • <i>Hospitalization</i> • <i>Vital Sign</i> • <i>Clinical Decision Support System</i> • <i>Clinical Decision Making</i> • <i>Medical Decision Making</i> • <i>Early Intervention</i> • <i>Patient Safety</i> • <i>Clinical Effectiveness</i>	• <i>Deterioração Clínica</i> • <i>Progressão da Doença</i> • <i>Hospitalização</i> • <i>Sinais Vitais</i> • <i>Evolução Clínica</i> • <i>Suporte à decisão clínica</i> • <i>Tomada de Decisão Clínica</i> • <i>Intervenção Médica Precoce</i> • <i>Segurança do Paciente</i> • <i>Efetividade</i>
T	Tempo	-----	-----	-----
S	Desenho dos Estudos	-----	-----	-----

Fonte: autores, 2025.

Nota: Foram empregados na busca os sinônimos oficiais disponíveis para cada um dos termos, conforme metodologia própria de cada plataforma de busca, utilizando-se o operador booleano OR e AND.

Os quantitativos detalhados para os estudos identificados, bem como a demonstração das informações de cada fase, serão apresentados em um fluxograma de acordo com o método *Preferred Reporting Items for Systematic Reviews and Meta-Analysis* (PRISMA), como mostrado na **Figura 1**.

3.7.4 Avaliação da heterogeneidade entre os estudos

Para testar a heterogeneidade, foram utilizado o teste I^2 de Higgins e Thompson (HIGGINS et al., 2003). Eles sugerem uma escala em que um valor de I^2 próximo a 0% indica não heterogeneidade entre os estudos, próximo a 25% indica baixa heterogeneidade, próximo a 50% indica heterogeneidade moderada e próximo a 75% indica alta heterogeneidade entre estudos.

3.7.5 Análise de risco de viés e qualidade metodológica dos estudos incluídos

Para os estudos observacionais, foram utilizadas a ferramenta de avaliação crítica The Joanna Briggs Institute (JBI) (MUNN et al., 2020), que tem como objetivo avaliar a qualidade metodológica dos estudos, além de determinar a extensão das possibilidades de viés, condução e análise. Constitui-se numa ferramenta com 11 perguntas, que podem ser respondidas em: Sim, Não, Não está claro e Não aplicável. Ao finalizar é realizado o julgamento do estudo, que consiste em: Incluir, Excluir ou Buscar mais informações.

3.8 SELEÇÃO DOS ESTUDOS E EXTRAÇÃO DOS DADOS

O processo de seleção do estudo foi realizado por 2 revisores de saúde de forma independente e dividido em fases:

Fase 1 – Identificação:

Consistiu na busca nas bases de dados utilizando os descritores e os filtros. Nessa fase foram excluídas as duplicatas.

Fase 2 – Seleção:

Consistiu na leitura dos títulos e resumos que continham o desenho do estudo e as questões principais. O *software online Rayyan* do Instituto de Pesquisa em Computação do Catar para Análise de Dados, foi utilizado nas fases 1 e 2 para remoção de duplicatas e leitura de títulos e resumo.

Fase 3 – Elegibilidade:

Consistiu na leitura completa dos estudos para seleção baseada nos critérios de elegibilidade. As características relevantes de cada artigo extraídas foram: base de dados, autor/ano, local/país, idioma, nº de amostra, objetivo do estudo, critérios de inclusão, tipo de intervenção, tempo de intervenção, instrumento de avaliação, desfecho principal, desfecho secundário e resultado do estudo. Nessa etapa, os dados dos estudos foram extraídos conforme os desfechos e organizados separadamente para cada estudo.

Fase 4 – Inclusão:

Consistiu na identificação dos estudos, já concluídos, que foram incluídos em síntese qualitativa e em síntese quantitativa.

3.9 GERENCIAMENTO DE DADOS

O gerenciamento dos dados e a remoção de duplicatas durante a seleção cega dos estudos foram realizados utilizando o *software Rayyan*, desenvolvido pelo *Qatar Computing and Research Institute (QCRI)*. Esse processo facilitou a organização e a exclusão de estudos duplicados, além de permitir a exportação das informações para compor a tabela de estudos excluídos e suas respectivas razões de exclusão. Após a seleção, os dados foram extraídos e organizados em uma tabela contendo as características dos estudos, utilizando o *Microsoft Excel 2013* para a síntese e análise das informações. Para o gerenciamento do banco de dados dos estudos incluídos na seleção e para a organização das referências bibliográficas, foi utilizado o *software Mendeley Desktop*, versão 1.19.8.

4 RESULTADOS

O tópico resultados desta dissertação, será apresentado como um artigo pronto, em português, redigido para a revista *Critical Care Medicine*.

Critical Care Medicine Cover Letter

To the Editor-in-Chief Critical Care Medicine Dear Editor,

It is with great pleasure that we submit for consideration by Critical Care Medicine our manuscript entitled **Effectiveness of the National Early Warning Score (NEWS 2) for Early Detection of Clinical Deterioration in Hospitalized Patients: A Systematic Review and Meta-analysis**.

The study presents a rigorously conducted systematic review, including a robust meta-analysis, to synthesize evidence on the prognostic effectiveness of NEWS 2, especially in the early identification of clinical deterioration in hospitalized adult patients.

Our work covered ten observational studies, totaling 339,216 patients. It used advanced statistical methods to ensure a rigorous analysis, including the support of a specialized statistical consultant, from the design phase to the final analysis of the meta-analysis.

We want to highlight that our team of authors broadly reflects the values of diversity and inclusion strongly encouraged by Critical Care Medicine.

Our group includes female researchers, professionals from the Amazon region, and members of diverse ethnicities, bringing to the study unique perspectives related to geographic and cultural challenges typical of the Brazilian Amazon region.

We believe this diversity amplifies our analytical capacity and reinforces the global reach and relevance of the work, especially in contexts of limited resources and high social complexity. We understand that the diversity of our team adds a valuable perspective, particularly considering the unique challenges faced by professionals and patients in healthcare settings in the Amazon, one of the most diverse and challenging regions in the world, especially in critical settings. We hope that our manuscript aligns with the editorial scope of Critical Care Medicine, and we look forward to your feedback. We thank the journal for considering our work and remain at your disposal for any additional clarifications that may be necessary.

Best regards,

Title Page

Effectiveness of the National Early Warning Score (NEWS 2) for Early Detection of Clinical Deterioration in Hospitalized Patients: A Systematic Review and Meta-analysis

Authors:

- Kezia Ferreira Lima Diniz, NP, MSc¹

First author responsible for Investigation, data curation, formal analysis, methodology, Writing – original draft.

- Éricka de Campos Correa, NP, MSc¹

The author is responsible for Investigation, data curation, formal analysis, methodology.

- Jacimara Vasques Gomes, NP, MSc¹

The author is responsible for Investigation, data curation, formal analysis, methodology.

- Antonio Mateus Araujo Teixeira, NP¹

The author is responsible for Investigation, data curation, formal analysis, methodology.

- Roberta Lins Gonçalves, PT, MSc, PhD¹

Supervisor and leading researcher responsible for conceptualization, data curation, formal analysis, funding acquisition, methodology, project administration, supervision, review, and editing.

¹ Graduate Program in Health Sciences, Federal University of Amazonas (UFAM), Manaus, Amazonas, Brazil.

Registration and Funding:

This systematic review was registered with PROSPERO (CRD420223530). The Amazonas State Research Support Foundation (FAPEAM) and the Coordination for the Improvement of Higher Education Personnel (CAPES) provided financial support for this study. No additional institutional or departmental funds were used.

Conflicts of Interest and Financial Disclosures:

The authors declare no financial disclosures or potential conflicts of interest.

Corresponding Author:

Roberta Lins Gonçalves, PT, MSc, PhD
Graduate Program in Health Sciences
Federal University of Amazonas (UFAM)

Av. General Rodrigo Octávio Jordão Ramos, 1200 – Coroado I
Manaus, Amazonas, Brazil – CEP 69067-005
Email: betalinsfisio@yahoo.com.br

ABSTRACT

Effectiveness of the National Early Warning Score (NEWS 2) for Early Detection of Clinical Deterioration in Hospitalized Patients: A Systematic Review and Meta-analysis

Kezia Ferreira Lima Diniz, NP, MSc; Éricka de Campos Correa, NP, MSc; Jacimara Vasques Gomes, NP, MSc; Antonio Mateus Araújo Teixeira, NP, MSc; Roberta Lins Gonçalves, PT, MSc, PhD.

Objective: To evaluate the effectiveness and prognostic accuracy of NEWS 2 in the early detection of clinical deterioration among hospitalized adult patients in real-world settings, comparing its predictive performance for adverse events with other validated prognostic scales.

Data Sources: Systematic searches were conducted in PubMed, Web of Science, Embase, LILACS, Google Scholar, and grey literature databases for articles published between 2017 and 2024 in English and Portuguese.

Study Selection: Observational cohort studies and clinical trials that evaluated the performance of NEWS 2 in predicting clinical deterioration and adverse outcomes (mortality, ICU admission, or mechanical ventilation) among hospitalized adult patients. Cross-sectional studies, case reports, narrative reviews, and studies lacking sufficient quantitative data for meta-analysis were excluded. Of the 94 studies selected for full-text evaluation, 10 met the inclusion criteria and were included for qualitative synthesis and meta-analysis, comprising 339,216 patients.

Data Extraction: Data were extracted on study year, design, setting, sample size, prognostic scales, outcomes, and predictive accuracy. Two reviewers independently extracted data; a third reviewer resolved discrepancies.

Data Synthesis: NEWS 2 demonstrated high pooled sensitivity (88%; 95% CI, 85–91%) and moderate specificity (75%; 95% CI, 70–80%), with an ROC-AUC of 0.85 (95% CI, 0.82–0.88; $I^2 = 45\%$), supporting clinical utility despite moderate heterogeneity. However, specialized scales (SMART-COP and CRB-65) exhibited superior predictive accuracy. Clinical applicability may vary due to population heterogeneity, suggesting that

combining NEWS 2 with complementary prognostic tools could enhance its effectiveness in clinical practice.

Conclusions: NEWS 2 has been shown to be effective in detecting early clinical deterioration with high sensitivity, particularly in critically ill patients. However, its specificity was lower than that of other scales, which may generate false alarms and unnecessary interventions. Future studies should examine the combination of NEWS 2 with additional prognostic scales or complementary assessments.

Key Words: Early Warning Score; Clinical Deterioration; Early Medical Intervention; Risk Prediction; Clinical Decision Support.

KEY POINTS

- **Question:** Is the NEWS 2 score effective in detecting early clinical deterioration in hospitalized adult patients across different settings?
- **Findings:** This systematic review and meta-analysis of 10 studies involving 339,216 patients found that NEWS 2 demonstrated high sensitivity (88%) and moderate specificity (75%) for predicting clinical deterioration. It performed comparably to other early warning scores.
- **Meaning:** NEWS 2 is an effective tool for early identification of deterioration, especially when combined with other prognostic scores or clinical assessments.

SUMMARY FOR SOCIAL MEDIA

NEWS 2 shows high sensitivity (88%) for detecting clinical deterioration in over 339,000 patients. This meta-analysis confirms its utility in early warning, though specificity is moderate. Combining it with other tools may enhance decision-making.

#CriticalCare #NEWS2

INTRODUÇÃO

A identificação oportuna da deterioração clínica de pacientes hospitalizados é essencial para melhorar os resultados e reduzir danos evitáveis, incluindo admissões não planejadas na unidade de terapia intensiva (UTI) e mortalidade hospitalar (1–16). Contudo, apesar dos crescentes investimentos em protocolos de segurança hospitalar e estratégias de monitoramento contínuo, o reconhecimento precoce desta condição continua sendo um desafio persistente em diversos ambientes de assistência à saúde (1,3,4). Clínicos experientes podem detectar anormalidades precoces em sinais vitais, mas a capacidade de interpretar e agir sobre essas descobertas consistentemente, é crítica para uma tomada de decisão oportuna e eficaz (1-3).

Em resposta, os *Early Warning Scores* (EWS) foram desenvolvidos para aprimorar a detecção de declínio fisiológico à beira do leito, integrando os principais sinais vitais em sistemas de pontuação padronizados e acionáveis (5–9). Desde sua introdução em 1997, várias atualizações foram desenvolvidas para melhorar a sua sensibilidade e precisão preditiva. Em 2000, o EWS foi adaptado para *Modified Early Warning Score* (MEWS), incorporando mais parâmetros e ajustes para diferentes grupos de pacientes. Em 2012, o *Royal College of Physicians* introduziu o *National Early Warning Score* (NEWS), que foi atualizado para NEWS 2 em 2017 (7-9). Esta versão atualizada incluiu seis parâmetros fisiológicos rotineiramente medidos: frequência respiratória, saturação de oxigênio, pressão arterial sistólica, frequência cardíaca, temperatura e nível de consciência ou nova confusão, produzindo uma pontuação total de 0 a 23 (9-13,16). O NEWS 2 também incorporou refinamentos importantes, como ajustes para insuficiência respiratória hipercápnica crônica e avaliação aprimorada do estado mental, incluindo delírio (10–13,16).

Amplamente adotado e endossado pelo *National Health Service* (NHS), o NEWS 2 é agora uma ferramenta de triagem padrão em enfermarias hospitalares e atendimento pré-hospitalar em todo o Reino Unido, com implementação internacional crescente. No entanto, sua eficácia na prática clínica do mundo real, particularmente para além dos cenários de pesquisa altamente controlados, permanece incerta (16). Além disso, seu desempenho comparativo em relação a outras ferramentas de prognóstico — como SMART-COP, CRB-65, PSI, MEWS e qSOFA — ainda não foi claramente estabelecido.

Sendo assim, o objetivo desta revisão sistemática com metanálise foi avaliar a efetividade e a precisão prognóstica do NEWS 2 na detecção precoce de deterioração clínica em pacientes adultos hospitalizados em cenários do mundo real, comparando seu desempenho preditivo para eventos adversos com outras escalas prognósticas validadas.

Nossa hipótese é que o NEWS 2 é efetivo na detecção precoce de deterioração clínica em adultos hospitalizados e oferece precisão preditiva suficiente para dar suporte à tomada de decisões clínicas em uma variedade de contextos de atendimento hospitalar.

MÉTODO

Conduzimos uma revisão sistemática com metanálise seguindo as recomendações do Manual da Colaboração Cochrane (17). Relatamos os resultados de acordo com a declaração *Preferred Reporting Items for Systematic Reviews and Meta-analysis* (PRISMA) (18) e registramos o protocolo do estudo no PROSPERO (CRD420223530).

Estratégia de busca e critérios de seleção

Três revisores independentes realizaram busca bibliográfica sistemática e abrangente da literatura científica publicada a partir de 2017, quando o NEWS 2 foi desenvolvido, a fim de encontrar ensaios clínicos, estudos de coorte e estudos caso-controle, que investigaram a efetividade da escala NEWS 2 na identificação precoce da deterioração clínica de pacientes adultos em unidades de internação hospitalar da clínica médica.

A estratégia de busca foi projetada para PubMed, Literatura Latino-Americana e do Caribe em Ciências da Saúde (LILACS) via BVS, Excerpta Médica dataBASE (Embase), *Web of Science* (WOS), Google scholar e a literatura cinzenta. A busca foi conduzida entre 10/09/2024 e 15/01/2025.

Os critérios de elegibilidade foram baseados na estrutura Paciente/População-Exposição-Resultado-Tempo-Desenho de estudo (PEOTS): **P – Population:** Pacientes adultos em unidades de internação hospitalar da clínica médica; **E – Exposure:** Escala NEWS 2; **O – Outcomes:** Desfecho primário: Capacidade de identificar precocemente a deterioração clínica do paciente. Desfechos secundários: Capacidade de prever parada cardiorrespiratória, capacidade de prever transferência para UTI, capacidade de prever

óbitos; **T – Time:** Estudos publicados a partir de 2017; **S – Study design:** Ensaios clínicos, estudos de coorte, estudos caso-controle.

Utilizamos os termos do *Medical Subject Headings* (MeSH) e palavras-chave apropriadas para cada base de dados, combinados entre si por meio dos operadores booleanos “AND” e “OR” para cada uma das bases (SUPLEMENTO A, B, C, D, E), conforme quadro 1. Quando possível, utilizamos os filtros: ano de publicação 2017-2024; idioma de publicação: inglês, português ou espanhol; tipo de estudo: apenas em humanos; adultos (18 + anos); desenho metodológico: ensaios clínicos randomizados, ensaios clínicos e estudos observacionais.

Para incluir os estudos, seguimos um processo de três fases com identificação, triagem e inclusão.

Na fase de identificação, realizamos a busca nas bases de dados por meio de descritores e filtros. Na fase de triagem, selecionamos os estudos com base em seus títulos, resumos e texto completo com base nos critérios de inclusão/exclusão. Resolvemos as discordâncias entre os avaliadores com discussão até que um consenso fosse alcançado. Na fase de inclusão, análise os estudos, e excluimos artigos incompletos, resumos, artigos de revisão, opinião de especialistas, editoriais, livros, trabalhos acadêmicos, dissertações, teses, anais de eventos científicos, artigos não disponíveis online, e aqueles que, por ventura foram incluídos, mas não responderam as questões de pesquisa. Os artigos excluídos foram relatados, juntamente com os motivos de sua exclusão.

Para o gerenciamento dos dados, incluindo a desduplicação, utilizamos o *software Rayyan do Qatar Computing the Research Institute* (QCRI) (19).

A demonstração das informações de cada fase está apresentada no fluxograma PRISMA, como mostrado na **Figura 1**.

Quadro 1. Estratégia PEOTS e Termos Mesh/Entre/DeCs

		<i>MeSH</i>	<i>Emtree (específicos)</i>	<i>DeCS</i>
P	População	• <i>Patient Acuity</i> • <i>Inpatients</i>	• <i>Patient Acuity</i> • <i>Disease Severity</i> • <i>Hospital Patient</i>	• <i>Gravidade do Paciente</i> • <i>Pacientes Internados</i>
E	Exposição	• <i>Early Warning Score</i>	<i>Early Warning Score</i>	• <i>Escore de Alerta Precoce</i>
O	Desfechos	• <i>Clinical Deterioration</i> • <i>Disease Progression</i> • <i>Hospitalization</i> • <i>Vital Signs</i> • <i>Clinical Evolution</i> • <i>Decision Support Systems, Clinical</i> • <i>Clinical Decision-Making</i> • <i>Early Medical Intervention</i> • <i>Patient Safety</i> • <i>Effectiveness</i>	• <i>Deterioration</i> • <i>Disease Exacerbation</i> • <i>Hospitalization</i> • <i>Vital Sign</i> • <i>Clinical Decision Support System</i> • <i>Clinical Decision Making</i> • <i>Medical Decision Making</i> • <i>Early Intervention</i> • <i>Patient Safety</i> • <i>Clinical Effectiveness</i>	• <i>Deterioração Clínica</i> • <i>Progressão da Doença</i> • <i>Hospitalização</i> • <i>Sinais Vitais</i> • <i>Evolução Clínica</i> • <i>Suporte à decisão clínica</i> • <i>Tomada de Decisão Clínica</i> • <i>Intervenção Médica Precoce</i> • <i>Segurança do Paciente</i> • <i>Efetividade</i>
T	Tempo	-----	-----	-----
S	Desenho dos Estudos	-----	-----	-----

Nota: Foram empregados na busca os sinônimos oficiais disponíveis para cada um dos termos, conforme metodologia própria de cada plataforma de busca, utilizando-se o operador booleano OR e AND.

Análise dos dados

Nós extraímos os dados dos artigos incluídos e os organizamos em um modelo padronizado, contendo informações como título do estudo, periódico científico, ano de publicação, delineamento metodológico, tipo de intervenção, desfechos avaliados, bem como outros achados relevantes para a presente revisão.

O risco de viés e a qualidade metodológica dos estudos incluídos foram avaliados por meio da ferramenta do *Joanna Briggs Institute* (JBI) [20]. Os critérios de avaliação (Q1 a Q11) foram aplicados para cada estudo, sendo as respostas classificadas como "Y" (Sim)

quando os critérios foram atendidos, "NA" (Não se Aplica) para itens irrelevantes ao estudo e "N" (Não) caso o critério não tenha sido cumprido.

Consideramos como indicadores de efetividade das diversas escalas utilizadas na detecção precoce da deterioração clínica em adultos hospitalizados os seguintes desfechos: a taxa de internação em unidade de terapia intensiva (UTI), a taxa de mortalidade hospitalar, a taxa de mortalidade entre pacientes que apresentaram intercorrência clínica e a área sob a curva ROC (AUROC) para a predição de mortalidade.

Estimamos estas taxas por meio de metanálises com cálculo do *efeito pooled*, considerando os estudos incluídos. As estimativas combinadas foram apresentadas com seus respectivos intervalos de confiança de 95% (IC 95%), sendo os resultados sintetizados graficamente por meio de *forest plots*, nos quais o efeito combinado foi representado por um losango.

Analizamos e comparamos a sensibilidade de diferentes escalas, incluindo NEWS 2, qSOFA, MEWS, PESI, A-DROP, CURB-65, PSI, SMART-COP, CRB-65 e REMS, por meio de metanálise, combinando as estimativas de sensibilidade extraídas dos estudos incluídos, considerando seus respectivos IC 95%.

Calculamos o valor médio ponderado das sensibilidades com base no tamanho amostral ou na precisão dos intervalos, além do erro padrão (EP) de cada estimativa de sensibilidade.

Fórmula para o EP:

$$EP = \frac{IC_{máximo} - IC_{mínimo}}{2 \times Z_{\frac{\alpha}{2}}}$$

Onde $Z_{\frac{\alpha}{2}}$ é o valor z correspondente a um intervalo de confiança de 95% (aproximadamente 1,96).

A sensibilidade combinada foi obtida por meio de modelos de efeitos aleatórios, conforme o grau de heterogeneidade entre os estudos, a qual foi testada utilizando o índice I^2 de Higgins e Thompson (21), considerando: 0–25% - Heterogeneidade baixa ou nula; 25–50% - Heterogeneidade moderada; 50–75% - Heterogeneidade substancial; e >75% - Heterogeneidade elevada.

O modelo de efeitos aleatórios nos permitiu estimar um efeito médio mais robusto mesmo diante da heterogeneidade entre os estudos. Portanto, a sensibilidade combinada levou em conta a média ponderada da sensibilidade, os IC combinados, e a heterogeneidade.

Para quantificar a heterogeneidade entre estudos, garantindo transparência na modelagem estatística e maior confiabilidade na estimativa do efeito combinado, estimamos o parâmetro Tau^2 (τ^2), que representa a variância dos efeitos verdadeiros. A estimativa foi obtida por meio do método da Máxima Verossimilhança Restrita (*Restricted Maximum Likelihood – REML*).

Nós usamos a estatística Z para determinar se uma estimativa foi significativamente diferente de um valor de referência, considerando: $|Z| > 1,96 \rightarrow$ resultado significativo ao nível de 5% ($p < 0,05$); $|Z| > 2,58 \rightarrow$ significativo ao nível de 1% ($p < 0,01$); $|Z| > 3,29 \rightarrow$ altamente significativo ao nível de 0,1% ($p < 0,001$). O valor Z indica quantos desvios padrão a estimativa está distante do valor nulo esperado sob a hipótese nula (em geral, zero).

A partir das estimativas de efeito e seus respectivos erros padrão, nós calculamos o P-score, o que nos possibilitou uma avaliação comparativa robusta entre as escalas analisadas, estimando a probabilidade de uma determinada escala ser superior às demais, assumindo a consistência dos dados entre os estudos. Essa métrica foi aplicada às medidas de sensibilidade, especificidade e AUROC, permitindo a hierarquização do desempenho preditivo de cada escala. Quanto maior o valor do P-score (variando de 0 a 1), melhor foi o desempenho da escala na predição de desfechos clínicos adversos.

Além disso, classificamos a força da recomendação como forte ou fraca, seguindo a abordagem GRADE (*Grading of Recommendations Assessment, Development and Evaluation*) (22).

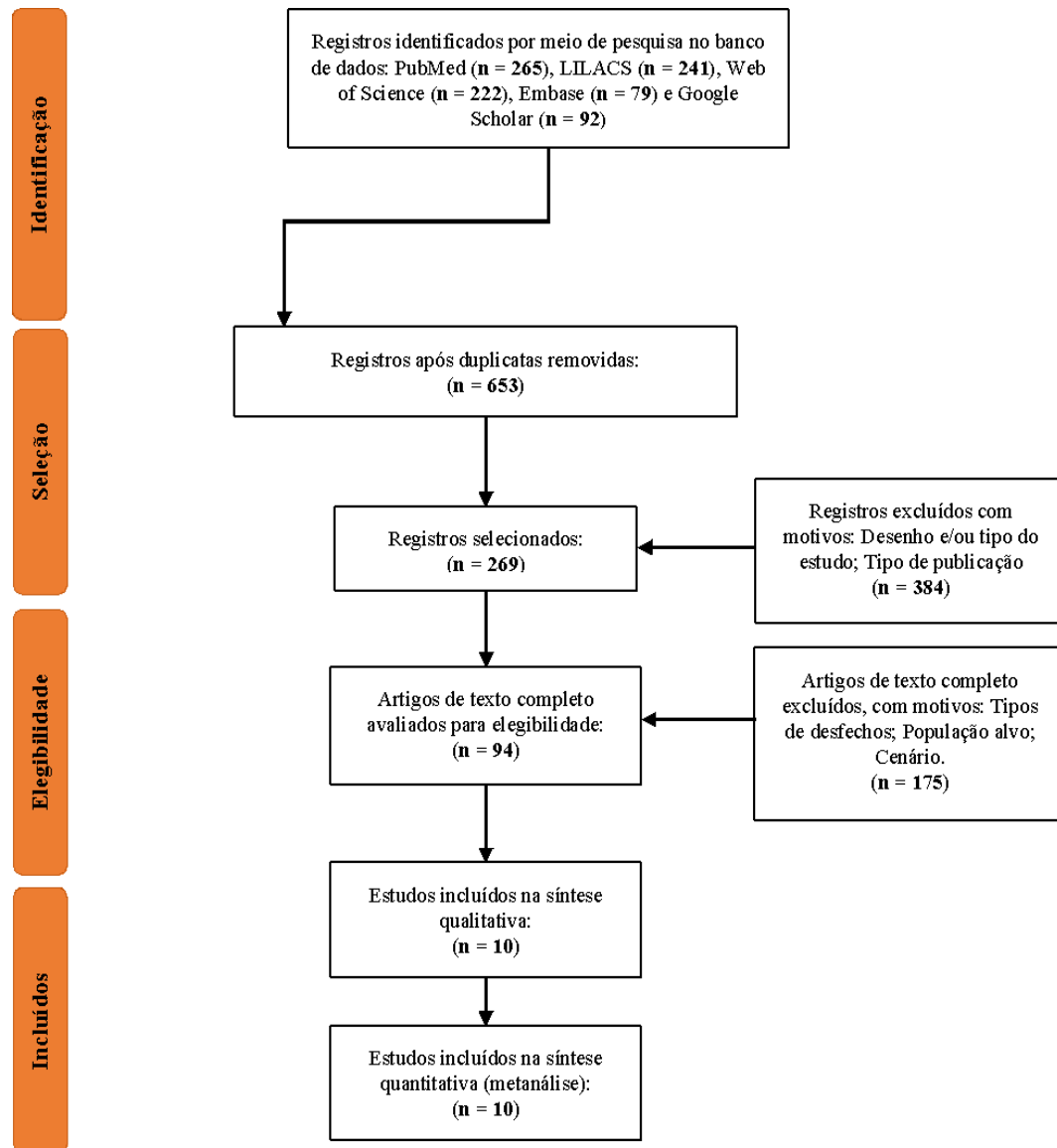
Resultados

Estudos incluídos

A busca nas bases de dados identificou 899 estudos. Removemos 246 duplicatas, restando 653 elegíveis para triagem. Excluimos 384 estudos após a leitura dos títulos, resultando em 269 estudos para análise de resumos. Após essa etapa, excluimos 175 estudos (suplemento E), restando 94 para leitura completa. Após a leitura completa, incluimos 10 estudos para a metanálise.

O fluxograma 1 resume nossa estratégia de busca (Fig. 1).

Figura 1. Fluxograma de seleção dos estudos para a revisão sistemática e metanálise



Principais características dos estudos incluídos

Os dez estudos incluídos na metanálise foram publicados entre os anos de 2020 e 2023 (23-32), e desenvolvidos em sete países diferentes, com predominância de nações de alta renda: Reino Unido (2 estudos) (23,38), Espanha (2 estudos) (25,32), Itália (2 estudos) (26,31), China (1 estudo) (24), Brasil (1 estudo) (29), Portugal (1 estudo) (30) e Índia (1 estudo) (27).

A maioria dos estudos analisados utilizou o escore NEWS2 para prever a mortalidade hospitalar em curto ou médio prazo (23, 24, 25, 26, 27, 29, 30, 31, 32). Além disso, dois estudos investigaram sua capacidade de identificar precocemente a deterioração clínica (23, 28), três investigaram a necessidade de ventilação mecânica invasiva ou não invasiva (27, 30, 31), três investigaram a admissão em unidade de terapia intensiva (UTI) (23, 26, 28, 31) e um estudo investigou o tempo prolongado de internação hospitalar (30). Apenas um estudo avaliou os impactos operacionais da implementação do NEWS2 no sistema de escalonamento clínico em uma grande coorte institucional (23).

Todos os estudos incluídos na metanálise foram do tipo observacionais, com boa aplicabilidade clínica, sendo seis de coorte retrospectiva (23,24,25,26,28) e quatro de coorte prospectiva (27,29,30,32).

A metanálise reuniu dados de uma amostra de 339.216 pacientes adultos hospitalizados em diferentes contextos clínicos. As populações variaram desde amostras reduzidas, como no estudo de Gomes et al. (2023) com 119 pacientes com DPOC agudizada, até grandes coortes institucionais, como no estudo de Forster et al. (2022), que analisou mais de 332 mil internações hospitalares. Também se destacam estudos com perfis variados, como os de Fan et al. (2020), Rodríguez et al. (2023), Covino et al. (2020), Baker et al. (2021), Aliberti et al. (2021), De Socio et al. (2021) e Martín-Rodríguez et al. (2020), abrangendo diferentes sistemas de saúde e realidades epidemiológicas.

A maioria dos pacientes foi atendida em hospitais universitários ou instituições de referência, com destaque para a inclusão de idosos e indivíduos com comorbidades, o que favoreceu a aplicabilidade dos achados em contextos clínicos diversos.

Os estudos incluídos na metanálise utilizaram o NEWS2 como principal ferramenta prognóstica para estratificação de risco clínico, sendo também objeto de comparação com outros escores amplamente utilizados.

O estudo de Fan et al. (2020) comparou sete escores — A-DROP, CURB-65, PSI, SMART-COP, NEWS2, CRB-65 e qSOFA — em pacientes com pneumonia por COVID-19, sugerindo que, embora o NEWS2 seja útil na identificação precoce de pacientes graves, sua acurácia pode variar conforme o contexto clínico e a população analisada. De forma semelhante, Covino et al. (2020) analisaram a acurácia de NEWS, NEWS2, NEWS-C, MEWS, qSOFA e REMS na predição de mortalidade precoce e admissão em UTI. Gomes et al. (2023) avaliaram o desempenho do NEWS2, NEWS2 modificado, DECAF e mDECAF em pacientes com exacerbação de DPOC, indicando que, apesar de o NEWS2 ser útil para prever necessidade de ventilação não invasiva e tempo de internação prolongado, escores como DECAF e mDECAF demonstraram maior especificidade para essa população. De Socio et al. (2021) compararam o NEWS2 ao COVID-GRAM, demonstrando seu desempenho superior na identificação de casos críticos. Aliberti et al. (2021), por sua vez, observaram que a adição de um escore de vulnerabilidade (PRO-AGE modificado) ao NEWS2 aumentou significativamente sua capacidade preditiva em idosos, reforçando a importância de estratégias combinadas. Esses achados indicam que o uso isolado do NEWS2 pode não ser ideal em todas as situações clínicas e que sua integração com outras variáveis pode ampliar sua aplicabilidade e precisão.

No que diz respeito ao desempenho prognóstico do NEWS2, os estudos analisaram desfechos clínicos relevantes, como mortalidade hospitalar, necessidade de ventilação mecânica ou não invasiva e admissão em UTI. Em pacientes com embolia pulmonar, Rodríguez et al. (2023) observaram que o NEWS2 apresentou menor valor preditivo negativo em comparação ao escore sPESI, sugerindo uma possível superestimação do risco em determinados cenários. Em contrapartida, Baker et al. (2021) relataram alta sensibilidade do NEWS2 (0,98) para detectar deterioração clínica em pacientes com COVID-19, com desempenho superior ao MEWS e qSOFA. Estudos como Martín-Rodríguez et al. (2020) e Pugazhvannan et al., (2021) reportaram elevadas áreas sob a curva ($AUC > 0,88$) na predição de mortalidade precoce e necessidade de ventilação mecânica, reforçando a robustez do escore como instrumento de triagem. Além disso, Pugazhvannan et al. (2021) validaram o uso do NEWS2 para prever desfechos adversos como ventilação mecânica e óbito em pacientes internados na Índia.

O estudo de Forster et al. (2022) foi o único, entre os incluídos na metanálise, que avaliou os impactos operacionais da implementação do protocolo de escalonamento clínico com

base no escore NEWS2 em um grande hospital do sistema de saúde do Reino Unido, em uma coorte de mais de 300 mil pacientes. A análise retrospectiva demonstrou que, embora o NEWS2 mantivesse capacidade moderada de predição de mortalidade em 24 horas, sua implementação gerou aumento significativo na carga de trabalho da equipe clínica, com elevado número de alertas para pacientes que não evoluíram para deterioração. O estudo concluiu que a aplicação indiscriminada de escalas de alerta precoce com pontos de corte rígidos pode resultar em sobrecarga operacional, sem ganhos proporcionais na acurácia preditiva. Os autores sugerem a necessidade de calibrar os protocolos de resposta clínica com base em critérios adicionais, como comorbidades e julgamento clínico, para otimizar o uso do NEWS2 na prática hospitalar.

O quadro 2 descreve as principais características dos estudos incluídos.

Quadro 2 - Distribuição dos estudos incluídos na metanálise, segundo as características dos estudos e os principais resultados obtidos

Autor/Ano	Local/País	Nº amostra	Tipo de Intervenção	Período de Intervenção	Desfecho	Principais Resultados
Rodríguez <i>et al.</i> (2023) ³²	Espanha	848	Avaliação e comparação de escores clínicos	Não especificado diretamente, mas o acompanhamento foi realizado por um mês após o diagnóstico.	Avaliar a ocorrência de eventos adversos em 30 dias utilizando os escores NEWS2 e sPESI.	Pacientes classificados como de baixo risco de acordo com o escore NEWS2 tiveram uma maior probabilidade de desfecho clínico complicado em comparação com aqueles classificados como de baixo risco de acordo com o escore sPESI (5,6% [21/377] vs 1,6% [5/313]) (P < 0,01). Quando a mortalidade por todas as causas em 30 dias foi examinada, o valor preditivo negativo para o escore sPESI foi de 99,7%, em comparação com 96,0% para a pontuação NEWS2.
Baker <i>et al.</i> (2020) ²⁸	Reino Unido	296	No estudo não houve uma intervenção direcionada como parte do estudo.	Dentro de 28 dias da admissão até a alta ou ocorrência de um evento grave	Avaliar a capacidade do NEWS 2 pode identificar previamente a deterioração clínica de paciente com COVID 19	NEWS2 ≥ 5 teve uma excelente sensibilidade para detectar deterioração pacientes com COVID-19. Um número significativamente maior de pacientes no grupo de deterioração registrou pelo menos um “gatilho verdadeiro positivo” no NEWS2 (131/133 pacientes; sensibilidade de 0,98 [IC 95% 0,96–1,00]) em comparação com o MEWS (52 pacientes; 0,39 [0,31–0,47], p<0,001) e o qSOFA (42 pacientes; 0,32 [0,24–0,39], p<0,001).
Aliberti <i>et al.</i> (2021) ²⁹	Brasil	1428	Intervenção diagnóstica ou prognóstica	Entre março e junho de 2020	Tempo até a morte dentro de 30 e 60 dias da admissão hospitalar	A incidência de mortalidade em 60 dias variou de 22% a 69% nos quartis do PRO-AGE modificado. O PRO-AGE modificado previu diferentes níveis de

Autor/Ano	Local/País	Nº amostra	Tipo de Intervenção	Período de Intervenção	Desfecho	Principais Resultados
						risco de mortalidade dentro de cada estrato do NEWS e melhorou a discriminação dos modelos de previsão de mortalidade.
Gomes <i>et al.</i> (2023) ³⁰	Portugal	119	Intervenção prognóstica ou de avaliação de risco	Janeiro de 2017 a novembro de 2018	Capacidade preditiva dos escores NEWS2, DECAF, NEWS288-92% e mDECAF para identificar eventos adversos em pacientes hospitalizados com exacerbação aguda de DPOC (EADPOC)	O escore NEWS2 foram considerados adequados para prever a necessidade de VNI aguda e internações hospitalares mais longas.
Forster <i>et al.</i> (2022) ²³	Reino Unido	332.682	Implementação de um protocolo clínico com base no NEWS2	Janeiro de 2016 a maio de 2019	Mortalidade dos pacientes internados, a carga de trabalho associada à introdução do National Early Warning Score 2 (NEWS2) e sua capacidade de prever a morte em 24 horas	Um aumento na atividade hospitalar, uma redução na mortalidade e um aumento nas escalões registradas, especialmente após a implementação do NEWS2
Fan <i>et al.</i> (2020) ²⁴	China	654	O estudo não incluiu uma intervenção direta, pois foi um estudo observacional retrospectivo. No entanto, a análise focou na avaliação de diferentes regras de pontuação de gravidade clínica (como NEWS2, qSOFA, PSI, CURB-65, entre outras) para prever o risco de morte	29 de dezembro de 2019 e 15 de fevereiro de 2020	Avaliar e comparar a precisão das regras de pontuação de gravidade (incluindo NEWS2, qSOFA, PSI, CURB-65, CRB-65, A-DROP e SMART-COP) na previsão do risco de morte em pacientes hospitalizados	Ele provou ser uma ferramenta válida para identificação precoce em pacientes gravemente doentes com infecção.
Martin Rodriguez	Espanha	2335	Intervenção prognóstica ou de avaliação de risco	1º de março de 2018 e 30 de maio de 2019	Mortalidade precoce nos primeiros 48 horas após o atendimento inicial	O NEWS2 apresentou uma área sob a curva (AUC) de 0,896 (IC 95% 0,82–

Autor/Ano	Local/País	Nº amostra	Tipo de Intervenção	Período de Intervenção	Desfecho	Principais Resultados
<i>et al.</i> (2020) ³²						0,97), indicando excelente capacidade preditiva.
Pugazhvanan et al., (2021) ²⁸	Índia	399	Intervenção prognóstica ou de avaliação de risco	Agosto de 2020 e dezembro de 2020	O desfecho primário foi a capacidade do escore NEWS2 na admissão de prever desfecho ruim em pacientes com doença COVID 19, definida como necessidade de ventilação mecânica ou morte em 28 dias	A mediana (intervalo interquartil) do escore NEWS2 na admissão foi de 2 (0–6). A sensibilidade e a especificidade do NEWS 2 de 5 ou mais na predição de desfechos ruins foram de 93,3% (IC 95%: 76,5–98,8) e 70,7% (IC 95%: 65,7–75,3), respectivamente [área sob a curva 0,88 (IC 95%: 0,847–0,927)]
Covino <i>et al.</i> (2020) ²⁶	China	334	Intervenção prognóstica ou de avaliação de risco	1º de março a 15 de abril de 2020	Os desfechos analisados neste estudo foram: Admissão na UTI dentro de 7 dias e Óbito dentro de 7 dias. Além disso, o estudo também avaliou esses desfechos em um intervalo mais curto (48 horas), mas o foco principal foi na previsão de eventos adversos dentro de uma semana.	NEWS e REMS foram os escores mais sensíveis para prever internação em UTI e óbito em 7 dias, respectivamente. Isso sugere que esses escores podem ser úteis para identificar pacientes de baixo risco no atendimento inicial do pronto-socorro.
De Socio <i>et al.</i> (2021) ³¹	Itália	121	Intervenção prognóstica ou de avaliação de risco	01 de março de 2020 a 15 de junho de 2020	O desfecho primário do estudo foi a doença crítica por COVID-19, definida como: Admissão na UTI Necessidade de ventilação invasiva Óbito Esse desfecho foi utilizado para avaliar a acurácia dos escores NEWS2 e COVID-GRAM na predição da gravidade da doença.	O NEWS2 demonstrou melhor desempenho na predição de COVID-19 grave do que o COVID-GRAM, sendo mais útil para identificar pacientes críticos.

Avaliação da qualidade metodológica

Ao avaliar o risco de viés e a qualidade metodológica, todos os estudos apresentaram 90% de respostas "Sim", indicando baixo risco de viés, conforme demonstrado na figura 2. Nenhum estudo foi classificado como de risco moderado ou alto, o que sugere uma boa qualidade metodológica entre os estudos avaliados. Além disso, todos os estudos obtiveram o mesmo percentual de respostas "Sim" (90%), o que pode sugerir uma padronização nos métodos dos estudos observacionais incluídos na revisão, aumentando a robustez dos resultados.

O critério Q5 (*Confounding factors identified and strategies to deal with them stated*) foi classificado como "NA" para todos os estudos. O critério Q10 (*Follow-up was complete and if not, reasons were explained and analysis to address follow-up loss was described*) também foi considerado "NA" em todos os estudos. A ausência de viés de detalhamento sobre fatores de confusão e acompanhamento incompleto pode limitar a interpretação de alguns resultados, especialmente em estudos de coorte.

Figura 2 - Análise De Risco De Viés - Estudos Observacionais - Ferramenta Jbi

Análise de Risco de Viés - Estudos Observacionais																
Autores	Ano	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Total "SIM" N.	%	Risco de Viés	
Rodriguez et al.	2023	Y	Y	Y	Y	NA	Y	Y	Y	Y	NA	Y	9	90%	BAIXO	
Baker et al.	2020	Y	Y	Y	Y	NA	Y	Y	Y	Y	NA	Y	9	90%	BAIXO	
Aliberti et al.	2021	Y	Y	Y	Y	NA	Y	Y	Y	Y	NA	Y	9	90%	BAIXO	
Gomes et al.	2023	Y	Y	Y	Y	NA	Y	Y	Y	Y	NA	Y	9	90%	BAIXO	
Forster et al.	2022	Y	Y	Y	Y	NA	Y	Y	Y	Y	NA	Y	9	90%	BAIXO	
Fan et al.	2020	Y	Y	Y	Y	NA	Y	Y	Y	Y	NA	Y	9	90%	BAIXO	
Martin Rodriguez et al.	2020	Y	Y	Y	Y	NA	Y	Y	Y	Y	NA	Y	9	90%	BAIXO	
Pugazhvanman et al.	2021	Y	Y	Y	Y	NA	Y	Y	Y	Y	NA	Y	9	90%	BAIXO	
Covino et al.	2020	Y	Y	Y	Y	NA	Y	Y	Y	Y	NA	Y	9	90%	BAIXO	
De Socio et al.	2021	Y	Y	Y	Y	NA	Y	Y	Y	Y	NA	Y	9	90%	BAIXO	
Y	Sim														RESPOSTAS SIM	
U	Não claro														RISCO ALTO	ATÉ 49%
N	Não														MODERADO	50 A 69%
NA	Não se aplica														BAIXO	> 70%

Nota: Respostas Sim – Risco Alto (Até 49%); Moderado (50 A 69%); Baixo (>70%). **Fonte:** Adaptado Pelos Autores Moola, S.Et Al. 2024. Jbi Critical Appraisal Checklist For Cohort Studies (2017).

Análise da heterogeneidade

Os desfechos analisados apresentaram diferentes níveis de heterogeneidade. A taxa de internação em UTI apresentou alta heterogeneidade ($I^2 = 79,88\%$), indicando grande variação entre os estudos. A mortalidade hospitalar teve heterogeneidade moderada ($I^2 = 45,2\%$), enquanto a mortalidade por intercorrências mostrou heterogeneidade elevada ($I^2 = 75,2\%$). Esses resultados sugerem variabilidade importante nos dados, devendo-se considerar possíveis fontes dessa inconsistência na interpretação dos achados.

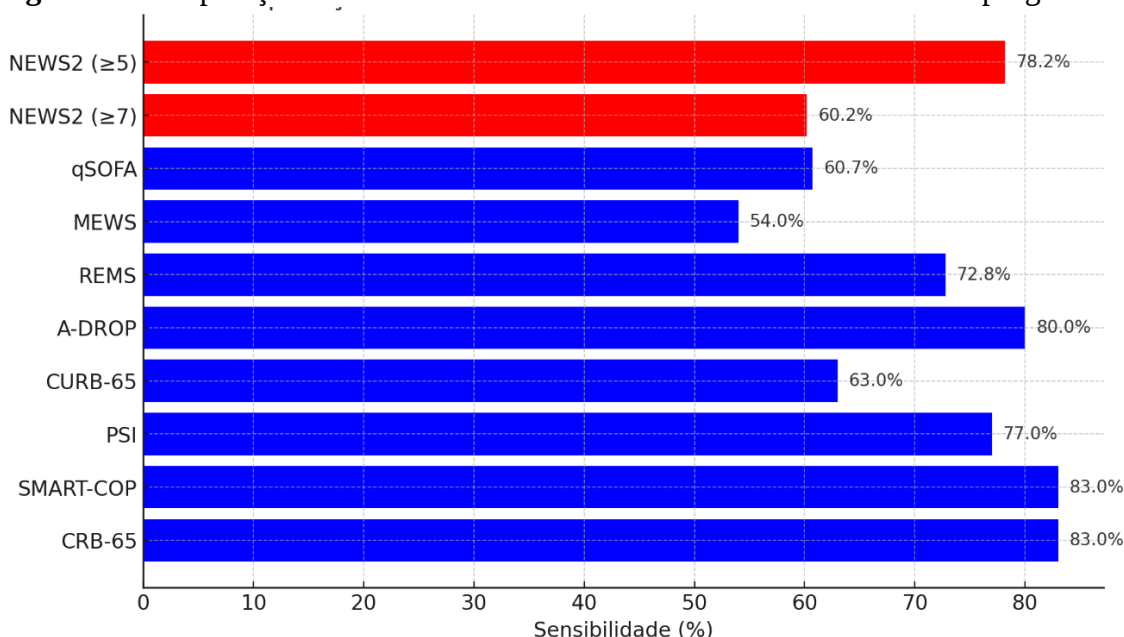
Metanálise da sensibilidade

As escalas NEWS2 (≥ 5 pontos), SMART-COP, CRB-65 e A-DROP apresentaram alta sensibilidade, indicando sua robustez na identificação de pacientes em risco. Por outro lado, escalas como qSOFA e MEWS tiveram sensibilidade mais baixa, o que pode limitar seu uso isolado na avaliação de risco (tabela 1, figura 3).

Tabela 1 - Sensibilidade agrupada das principais escalas

Escala	Sensibilidade média (%)	IC 95%
NEWS2 (≥ 5 pontos)	78,2	72,0 - 84,5
NEWS2 (≥ 7 pontos)	60,2	51,0 - 68,5
qSOFA	60,7	49,5 - 71,2
MEWS	54,0	43,5 - 64,2
REMS	72,8	62,5 - 81,5
A-DROP (>2)	80,0	73,0 - 87,0
CURB-65 (>2)	63,0	55,0 - 71,0
PSI (>3)	77,0	70,0 - 84,0
SMART-COP (>2)	83,0	77,0 - 89,0
CRB-65 (>1)	83,0	77,0 - 89,0

Figura 3 - Comparação da Sensibilidade do NEWS2 entre diferentes escalas prognósticas



Legenda: Gráfico de barras horizontais comparando os valores de sensibilidade (%) de diferentes escalas prognósticas utilizadas na detecção precoce de deterioração clínica em pacientes hospitalizados. As escalas NEWS2 com pontos de corte ≥ 5 (IC 72,0-84,5) e ≥ 7 (IC 51,0 - 68,5) estão representadas em vermelho, enquanto as demais escalas (qSOFA (IC 49,5- 71,2), MEWS (IC 43,5- 64,2), REMS (IC 62,5- 81,5), A-DROP (IC 73,0- 87,0), CURB-65 (IC 55,0 - 71,0), PSI (IC 70,0 - 84,0) , SMART-COP (77,0 - 89,0) e CRB-65 (IC 77,0- 89,0) estão representadas em azul.

Metanálise da taxa de internação na UTI

A Tabela 2 apresenta os resultados da metanálise realizada com modelo de efeitos aleatórios, incluindo sete estudos ($k = 7$): Forster et al. (2022) (23), Fan et al. (2020) (24), Rodríguez et al. (2023) (25), Covino et al. (2020) (26), Pugazhvannan et al. (2021) (27), De Socio et al. (2021) (31) e Martín-Rodríguez et al. (2020) (32).

Foi possível observar, com forte evidência estatística e 95% de confiança, que a taxa verdadeira de internação na UTI situou-se entre 9,2% e 30,7%, com uma proporção geral estimada de 20% (SE = 0,0549; Z = 3,63; $p < 0,001$; IC 95% [0,092; 0,307]).

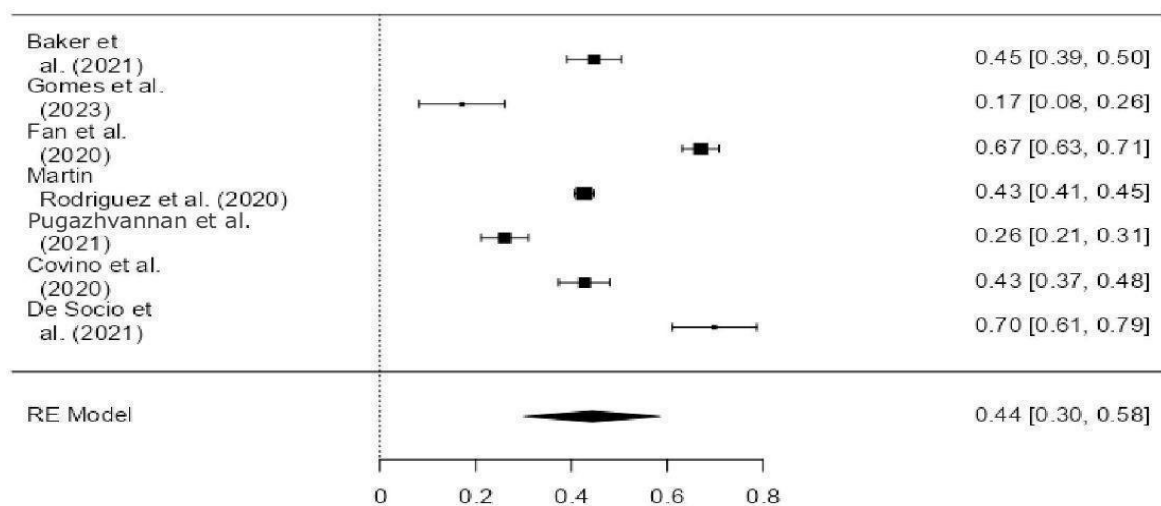
Embora a taxa média de internação na UTI seja considerável, as estimativas individuais variaram amplamente, conforme demonstrado na figura 4, sugerindo possíveis diferenças entre as populações analisadas ou nos critérios utilizados para definir a necessidade de internação em UTI (figura 5). Adicionalmente, fatores como gravidade clínica e critérios institucionais podem ter influenciado substancialmente essa taxa, recomendando-se análises adicionais para investigação dessas variáveis.

Tabela 2 - Metanálise da Taxa de Internação na Unidade de Terapia Intensiva

Modelo de Efeitos Aleatórios (k = 7)	Estimativa	Se	Z	p	Limite Inferior do Intervalo de Confiança	Limite Superior do IC
Intercepto	0.200	0.0549	3.63	< .001*	0.092	0.307

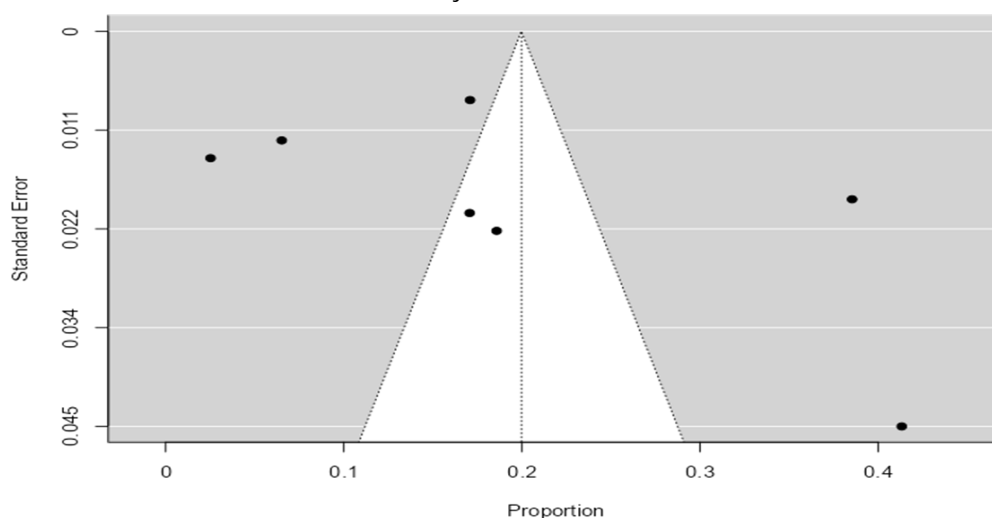
Nota: *Tau² Estimator: Restricted Maximum-Likelihood.* Metanálise da Taxa de Internação na Unidade de Terapia Intensiva: Modelo de Efeitos Aleatórios (k = 7) e Análise de Efeito Pooled entre Estudos, fornecendo a estimativa média do efeito e os principais parâmetros estatísticos associados: erro padrão (SE), estatística Z, valor de p e Intervalo de confiança (IC de 95%).

Figura 4 - Metanálise da Taxa de Internação na UTI



Legenda: Forest plot da metanálise da taxa de internação em unidade de terapia intensiva (UTI), utilizando modelo de efeitos aleatórios. Cada estudo incluído foi representado por um quadrado preto proporcional ao seu peso estatístico, com linhas horizontais indicando os respectivos intervalos de confiança de 95% (IC 95%) da estimativa pontual. O efeito combinado da metanálise está representado pelo losango na parte inferior do gráfico, refletindo a estimativa média agrupada da taxa de internação em UTI entre os estudos analisados.

Figura 5 -Metanálise da Taxa de Internação na UTI



Legenda: Funnel plot utilizado para avaliação da presença de viés de publicação na metanálise da taxa de internação em UTI ($k = 7$). Os pontos representam os estudos incluídos, posicionados com base em seus respectivos erros padrão (eixo y) e estimativas de efeito (eixo x). A linha vertical indica o valor combinado da metanálise, enquanto o triângulo invertido central delimita a região onde os estudos são esperados na ausência de viés e heterogeneidade substancial. A distribuição assimétrica dos pontos, com maior concentração à esquerda e dispersão na extremidade inferior direita, sugere possível viés de publicação ou heterogeneidade metodológica entre os estudos incluídos. Esses achados reforçam a importância de cautela na interpretação dos resultados

Metanálise da taxa de Mortalidade Hospitalar

A Tabela 3 apresenta os resultados da metanálise da taxa de mortalidade hospitalar, conduzida com modelo de efeitos aleatórios, incluindo sete estudos ($k = 7$): Forster et al. (2022) (23), Fan et al. (2020) (24), Rodríguez et al. (2023) (25), Covino et al. (2020) (26), Pugazhvanan et al. (2021) (27), Aliberti et al. (2021) (29) e De Socio et al. (2021) (31).

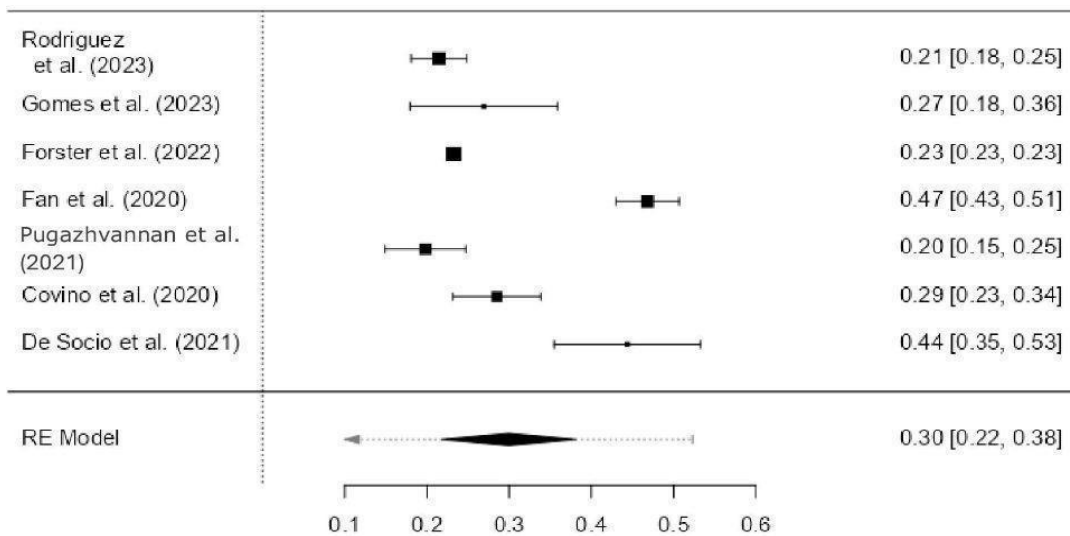
Foi possível observar, com forte evidência estatística e alto grau de confiança (IC 95% de 0,043 a 0,141 (ou seja, entre 4,3% e 14,1%)), que a taxa média de mortalidade hospitalar entre os estudos analisados foi de aproximadamente 9,2% (SE associado de 0,0252, estatística $Z = 3,66$, e $p < 0,001$) (tabela 3, figura 6, figura 7). A estimativa de heterogeneidade (Tau^2) para este desfecho indicou baixa variabilidade entre os estudos, o que fortalece a confiabilidade do efeito combinado obtido.

Tabela 3: Metanálise da Taxa de Mortalidade Hospitalar

Modelo de Efeitos Aleatórios ($k = 7$)	Estimativa	Se	Z	p	Limite Inferior do Intervalo de Confiança	Limite Superior do Intervalo de Confiança
Intercepto	0.0920	0.0252	3.66	< .001	0.043	0.141

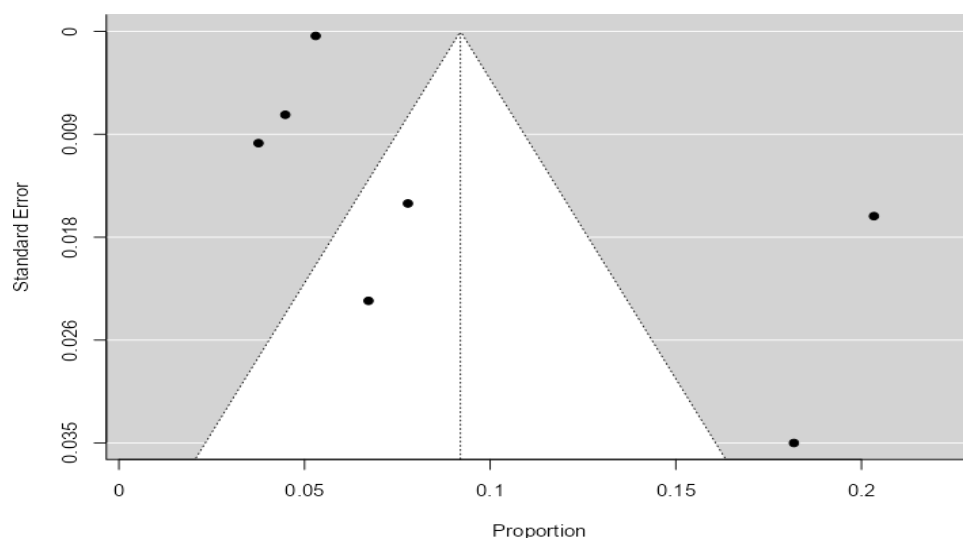
Nota: Tau^2 Estimator: Restricted Maximum-Likelihood. Essa análise forneceu a estimativa média do efeito e os principais parâmetros estatísticos associados: erro padrão (SE), estatística Z, valor de p e IC de 95%.

Figura 6 - Metanálise da Taxa de Mortalidade Hospitalar



Legenda: Forest plot da metanálise da taxa de mortalidade hospitalar, utilizando modelo de efeitos aleatórios. Cada estudo é representado por um quadrado proporcional ao seu peso na análise, com barras horizontais indicando o intervalo de confiança de 95% (IC 95%) da estimativa. O losango na parte inferior do gráfico representa a taxa média combinada de mortalidade hospitalar nos estudos incluídos.

Figura 7 - Metanálise da Taxa de Mortalidade Hospitalar



Legenda: Funnel plot representando a distribuição dos estudos incluídos ($k = 5$) na metanálise da taxa de mortalidade em pacientes que apresentaram intercorrência. Cada ponto indica um estudo, posicionado com base no erro padrão (eixo y) e na estimativa de efeito (eixo x). A linha vertical representa a estimativa combinada da metanálise, enquanto o triângulo invertido delimita a zona de dispersão esperada na ausência de viés. Observa-se assimetria evidente, com concentração dos estudos à esquerda e maior dispersão dos pontos, o que pode sugerir viés de publicação ou heterogeneidade metodológica. A interpretação desses resultados deve ser feita com cautela, especialmente considerando o número limitado de estudos incluídos.

Metanálise da taxa de mortalidade em pacientes que apresentaram intercorrência

A Tabela 5 apresenta os resultados da metanálise da taxa de mortalidade em pacientes que apresentaram intercorrência, conduzida com modelo de efeitos aleatórios, considerando cinco estudos ($k = 5$): Covino et al. (2020) (26), Pugazhvannan et al. (2021) (27), Aliberti et al. (2021) (29), Gomes et al. (2023) (30) e De Socio et al. (2021) (31).

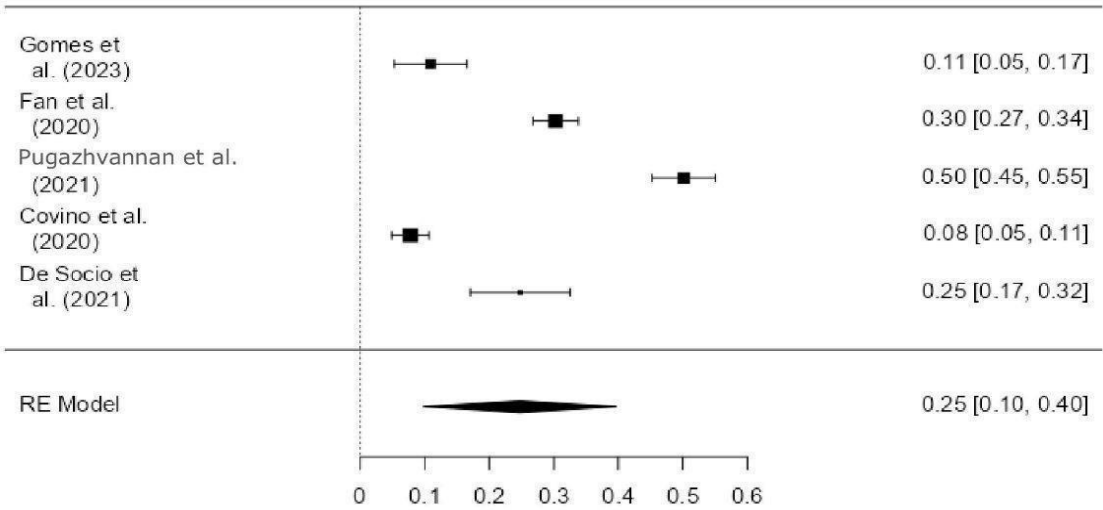
Foi possível observar, com forte evidência estatística e alto grau de confiança que a taxa de mortalidade ficou entre 9,8% e 39,7%, indicando que, em média, 24,8% dos pacientes com intercorrência evoluíram para óbito (SE associado = 0,0763, $Z = 3,24$, $p = 0,001$, IC 95% = [0,098; 0,397]) (tabela 5, figura 8, figura 9). Apesar da amplitude do IC sugerir heterogeneidade entre os estudos incluídos, corroborado pela estimativa de Tau^2 , ainda assim, o efeito combinado permanece estatisticamente confiável e relevante, representando uma estimativa robusta da mortalidade em pacientes com intercorrência clínica.

Tabela 5: Metanálise da taxa de mortalidade em pacientes que apresentaram intercorrência

Modelo de Efeitos Aleatórios ($k = 5$)	Estimativa	Se	Z	p	Limite Inferior do IC	Limite Superior do IC
Intercepto	0.248	0.0763	3.24	0.001	0.098	0.397

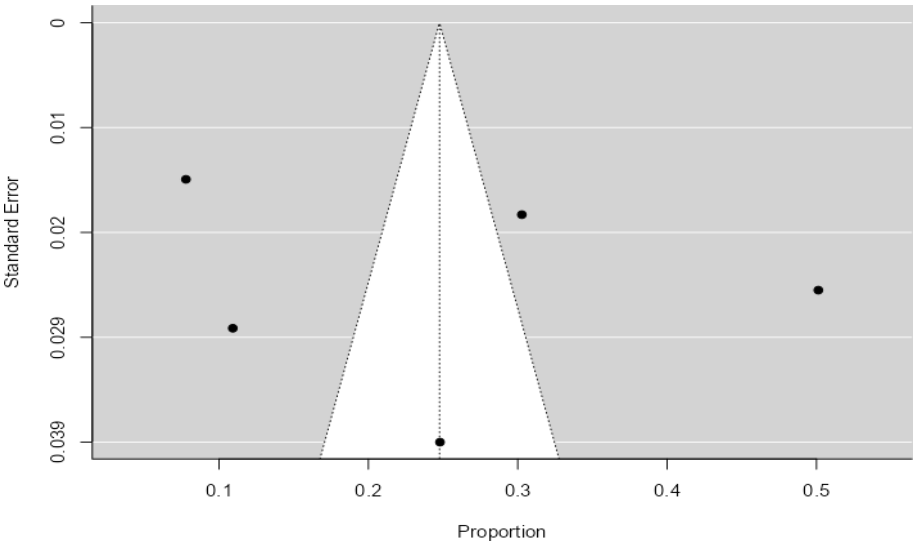
Legenda: Tau^2 Estimator: Restricted Maximum-Likelihood. Essa análise forneceu a estimativa média do efeito e os principais parâmetros estatísticos associados: erro padrão (SE), estatística Z, valor de p e Intervalo de confiança (IC de 95%). Modelo de Efeitos Aleatórios ($k = 5$) e Análise de Efeito Pooled entre Estudos.

Figura 8 - Metanálise da Taxa de Mortalidade Hospitalar em pacientes que apresentaram intercorrência



Legenda: Forest plot da metanálise da Taxa de Mortalidade Hospitalar, utilizando modelo de efeitos aleatórios. Cada estudo incluído foi representado por um quadrado preto proporcional ao seu peso estatístico, com linhas horizontais indicando os respectivos intervalos de confiança de 95% (IC 95%) da estimativa pontual. O efeito combinado da metanálise está representado pelo losango na parte inferior do gráfico, refletindo a estimativa média agrupada da taxa de Mortalidade Hospitalar entre os estudos analisados.

Figura 9 - Metanálise da Taxa de Mortalidade Hospitalar



Legenda: Funnel plot utilizado para avaliação da presença de viés de publicação na metanálise da taxa de mortalidade hospitalar: Análise de Efeito Pooled entre Estudos. Os pontos representam os estudos incluídos ($k = 7$), distribuídos em torno da estimativa combinada de efeito (linha vertical central). O triângulo invertido delimita a região esperada de dispersão dos estudos sob ausência de viés e heterogeneidade, com base no erro padrão. A distribuição visual dos pontos sugere leve assimetria, com maior concentração de estudos à esquerda do gráfico, o que pode indicar viés de publicação ou variabilidade metodológica. Recomenda-se cautela na interpretação, sobretudo devido ao número reduzido de estudos.

Metanálise da Curva AUC ROC

Analizamos o desempenho discriminativo da escala NEWS2 na predição de desfechos clínicos adversos, como mortalidade hospitalar, necessidade de ventilação invasiva, internação prolongada e admissão em UTI, com base na área sob a curva ROC (AUROC).

Para esta análise incluímos cinco estudos que compararam a acurácia do NEWS2 com outras escalas prognósticas validadas, com cálculo das médias ponderadas das AUROCs quando aplicável: Fan et al. (2020), que comparou NEWS2 com A-DROP, CURB-65, PSI, SMART-COP, CRB-65 e qSOFA em pacientes com pneumonia por COVID-19; Aliberti et al. (2021), que avaliou NEWS2 e o escore PRO-AGE para prever mortalidade em idosos hospitalizados; De Socio et al. (2021), que comparou a capacidade preditiva do NEWS2 com o escore COVID-GRAM para casos críticos de COVID-19; Rodríguez et al. (2023), que comparou NEWS2 e sPESI na estratificação de risco em pacientes com embolia pulmonar; e Covino et al. (2020), que comparou NEWS2 com TRIAGE, MEWS, NEWS, NEWS-C, qSOFA e REMS em pacientes com COVID-19 no pronto-socorro.

A análise revelou que o desempenho do NEWS2 na predição de mortalidade foi moderado (AUROC média = 0,76), sendo inferior a escores como A-DROP (0,87), CURB-65 (0,85), PSI (0,85) e REMS (0,882).

Em relação à admissão em UTI, NEWS2 foi superado por escalas como TRIAGE e NEWS original.

Para a ventilação mecânica, o NEWS2 teve desempenho inferior aos escores específicos para DPOC, como DECAF e mDECAF.

Portanto, foi possível observar que, embora o NEWS2 apresente boa performance geral e seja útil como ferramenta de triagem hospitalar, sua acurácia tende a ser inferior à de escalas especializadas para desfechos clínicos específicos. Esses resultados sugerem que o uso isolado do NEWS2 pode não ser ideal para decisões clínicas críticas, e que estratégias combinadas podem ampliar sua eficácia. Estudos futuros devem explorar a integração do NEWS2 com outras escalas para melhorar a capacidade preditiva em diferentes contextos assistenciais.

Tabela 6 - Metanálise da curva AUROC para Previsão de Mortalidade

Estudo	NEWS2	Outras Escalas Comparadas
Fan <i>et al.</i> (2020)	0.81	A-DROP (0.87), CURB-65 (0.85), PSI (0.85), SMART-COP (0.84), CRB-65 (0.80), qSOFA (0.73)
Aliberti <i>et al.</i> (2021)	0.71	PRO-AGE (0.70)
De Socio <i>et al.</i> (2021)	0.87	COVID-GRAM (0.77)
Rodriguez <i>et al.</i> (2023)	0.59	sPESI (0.74)
Covino <i>et al.</i> (2020)	0.753	TRIAGE (0.849), MEWS (0.630), NEWS (0.829), NEWS-C (0.773), qSOFA (0.808), REMS (0.882)

Legenda: Metanálise da curva AUROC para Previsão de Mortalidade: Modelo de Efeitos Aleatórios ($k = 5$) e Análise de Efeito *Pooled* entre Estudos Média Ponderada NEWS2: 0.76

A Tabela 7 apresenta a comparação da AUROC para previsão de admissão em UTI em dois intervalos de tempo (48 horas e 7 dias). A média ponderada da AUROC para a escala NEWS2 foi de 0,77, enquanto a escala TRIAGE apresentou desempenho ligeiramente superior, com AUROC de 0,839 para 48h e 0,818 para 7 dias. Apesar dessa diferença, o NEWS2 manteve um bom desempenho discriminativo na predição de admissões em UTI.

Embora a análise incluía apenas dois estudos, o agrupamento dos dados permitiu uma avaliação exploratória da AUROC do NEWS2 na previsão de necessidade de ventilação, uma vez que os dois estudos são de alta qualidade metodológica, e avaliaram desfechos idênticos, com medidas estatísticas comparáveis (ex.: AUROC). Contudo, não se pode calcular a heterogeneidade com precisão, sendo que a interpretação destes resultados deve levar em conta o viés de publicação e a possibilidade da estimativa combinada ser instável e não representativa.

Sendo assim, os resultados da metanálise com dois estudos devem ser interpretados com cautela, dado o número limitado de estudos disponíveis e a consequente limitação do poder estatístico da síntese.

Tabela 7 - Metanálise da curva AUROC para Previsão de Admissão na UTI

Estudo	NEWS2	Outras Escalas Comparadas
Covino <i>et al.</i> (2020) (48h)	0.780	TRIAGE (0.839), MEWS (0.652), NEWS (0.802), NEWS-C (0.771), qSOFA (0.566), REMS (0.742)
Covino <i>et al.</i> (2020) (7 dias)	0.762	TRIAGE (0.818), MEWS (0.618), NEWS (0.783), NEWS-C (0.755), qSOFA (0.557), REMS (0.735)

Legenda: Metanálise da curva AUROC para Previsão de Admissão na UTI: Modelo de Efeitos Aleatórios ($k = 2$) e Análise de Efeito Pooled entre Estudos. Média Ponderada NEWS2: 0.77.

A Tabela 8 apresenta a comparação da AUROC da escala NEWS2 com outras escalas prognósticas aplicadas a diferentes medidas de desempenho clínico. A média ponderada da AUROC não foi calculada nesta análise devido à heterogeneidade dos desfechos analisados, à variação dos contextos clínicos entre os estudos e à inconsistência nos dados reportados (como erro padrão, IC e tamanhos amostrais). Por isso, optou-se por apresentar os resultados de forma descritiva e exploratória, sem agrupamento estatístico formal.

Os dados sugerem que o NEWS2 apresentou AUROCs competitivas, especialmente em comparação com escalas amplamente utilizadas, como MEWS (0,63) e qSOFA (0,62). Um exemplo notável é o estudo de Forster *et al.* (2022), no qual o NEWS2 apresentou AUROCs variando de 0,909 a 0,921, desempenho comparável ao de outras escalas utilizadas em contextos cirúrgicos e clínicos. Já o estudo de Martín-Rodríguez *et al.* (2020) relatou AUROCs associadas a diferentes janelas temporais (24h, 48h, 7 dias e 30 dias), com valores geralmente acima de 0,80, o que reforça o bom desempenho geral do NEWS2 na predição de desfechos clínicos adversos.

Contudo, esta análise apresenta limitações metodológicas relevantes: além da ausência de cálculo de média ponderada, o número reduzido de estudos ($k = 3$) restringe o poder de generalização dos achados e limita a interpretação estatística formal, como a avaliação da heterogeneidade entre os estudos. Dessa forma, os resultados devem ser interpretados de forma exploratória e descritiva, sendo recomendada a condução de futuras investigações para validação desses achados em diferentes contextos clínicos e populacionais.

Tabela 8: Metanálise da curva AUROC para Outras Medidas de Desempenho

Estudo	NEWS2	Outras Escalas Comparadas
Baker et al. (2021)	0.71	MEWS (0.63), qSOFA (0.62), AUC geral (0.88)
Forster <i>et al.</i> (2022)	0.909-0.921	Cirurgia (0.891-0.928), Medicina (0.904-0.915)
Martin Rodriguez <i>et al.</i> (2020)	-	24h (0.862), 48h (0.885), 7 dias (0.835), 30 dias (0.807)
Pugazhvannan et al., (2021)	-	AUC geral (0.88)

Legenda: Metanálise da curva AUROC para Outras Medidas de Desempenho: Modelo de Efeitos Aleatórios (k = 3) e Análise de Efeito Pooled entre Estudos

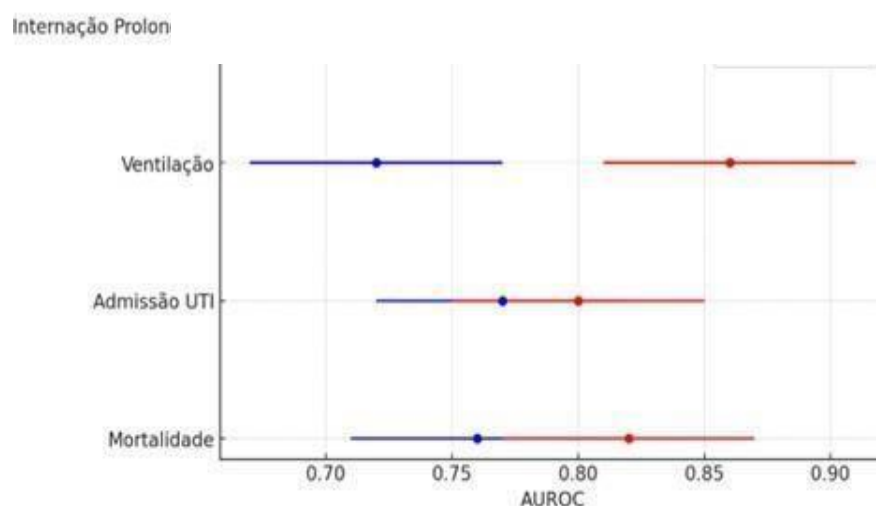
A Figura 10 apresenta a comparação da acurácia preditiva da escala NEWS2 com outras escalas prognósticas, utilizando os valores da AUROC para quatro desfechos clínicos relevantes: internação prolongada, ventilação, admissão em UTI e mortalidade. O gráfico ilustra os valores médios de AUROC, acompanhados de seus IC 95%, permitindo uma avaliação direta do desempenho discriminativo de cada escala em relação aos desfechos analisados.

De modo geral, observa-se que as outras escalas prognósticas (em vermelho) apresentaram AUROCs superiores em todos os desfechos, especialmente na previsão de ventilação e mortalidade, onde a diferença em relação ao NEWS2 (em azul) foi mais pronunciada.

Para admissão na UTI e internação prolongada, a diferença entre as escalas foi mais discreta, com desempenho moderadamente semelhante.

Portanto, foi possível observar que apesar de a escala NEWS2 demonstrar boa acurácia preditiva geral, os resultados reforçam que escalas específicas para determinados contextos clínicos podem oferecer melhor performance na predição de desfechos críticos, sugerindo que o uso isolado do NEWS2 deve ser avaliado com cautela, principalmente quando decisões clínicas complexas estão envolvidas.

Figura 10. Comparação da acurácia preditiva (AUROC) entre a escala NEWS2 e outras escalas prognósticas para diferentes desfechos clínicos



Legenda: Gráfico de dispersão com barras de erro horizontais comparando os valores médios da área sob a curva ROC (AUROC) da escala NEWS2 (em azul) com outras escalas prognósticas validadas (em vermelho) para quatro desfechos clínicos: internação prolongada, necessidade de ventilação, admissão em UTI e mortalidade hospitalar. Cada ponto representa a AUROC média estimada para o respectivo desfecho, e as barras horizontais indicam os intervalos de confiança (IC 95%) das estimativas. O gráfico evidencia que, embora o NEWS2 apresente desempenho razoável, algumas escalas especializadas demonstraram maior acurácia preditiva, sobretudo nos desfechos de ventilação e mortalidade.

Avaliação da qualidade da evidência (GRADE)

Nós avaliamos a qualidade da evidência dos principais desfechos avaliados nesta metanálise com base na abordagem GRADE, considerando a certeza global dos achados, a magnitude do efeito, a consistência entre os estudos e a aplicabilidade clínica, conforme demonstrado no Quadro 2.

Para o desfecho de mortalidade hospitalar, a proporção combinada encontrada foi de 9,2% (IC 95%: 4,3%–14,1%). Embora o escore NEWS2 tenha demonstrado utilidade clínica, não apresentou desempenho consistentemente superior a outras escalas comparadoras, como A-DROP, sPESI ou PSI. Em diversos estudos, o NEWS2 exibiu AUROC inferior para predição de mortalidade. Assim, a certeza da evidência foi classificada como moderada, em razão da ausência de superioridade sustentada e da heterogeneidade metodológica entre os estudos incluídos.

No desfecho de sensibilidade para detecção de deterioração clínica, a média agrupada para o NEWS2 foi de 78,2%, enquanto escalas como SMART-COP e A-DROP apresentaram

sensibilidade de até 83%. Apesar de o NEWS2 ser adequado para triagem, não apresentou desempenho significativamente melhor, e em alguns contextos foi inferior a escalas específicas. A certeza da evidência foi considerada baixa a moderada, principalmente devido à variabilidade entre os estudos e à ausência de um benefício clínico consistente frente às escalas alternativas.

Em relação ao desfecho de admissão em UTI, a proporção combinada observada foi de 20% (IC 95%: 9,2%–30,7%). Embora o NEWS2 tenha se mostrado clinicamente útil, estudos comparativos demonstraram desempenho inferior ao da escala TRIAGE e à versão original do próprio NEWS. Portanto, a certeza da evidência foi classificada como moderada, uma vez que, apesar da utilidade prática, não foi observada superioridade estatisticamente robusta em relação às escalas comparadoras.

Quadro 2. Avaliação da qualidade da evidência (GRADE)

Desfecho	Estimativa do efeito (RR ou Proporção)	Comparação com outras escalas	Certeza da evidência (GRADE)	Importância para decisão clínica
Mortalidade hospitalar	Proporção combinada: 9,2% (IC 95%: 4,3%–14,1%)	NEWS2 teve desempenho inferior em AUROC em vários estudos; não há RR agrupado entre escalas	Moderada	Crítico
Sensibilidade do NEWS2	Sensibilidade média: 78,2%	NEWS2 similar ou inferior a SMART-COP e A-DROP (ambos com 83%)	Baixa a moderada	Crítico
Admissão em UTI	Proporção combinada: 20% (IC 95%: 9,2%–30,7%)	Desempenho inferior a TRIAGE e NEWS original em alguns estudos	Moderada	Crítico

DISCUSSÃO

Esta metanálise avaliou a efetividade do NEWS2 na detecção precoce de deterioração clínica em adultos hospitalizados, considerando desfechos clínicos críticos como mortalidade hospitalar, necessidade de ventilação, admissão em UTI e sensibilidade diagnóstica em comparação com outras escalas prognósticas.

A análise conjunta dos dados de dez estudos observacionais revelou que, embora o NEWS2 seja uma ferramenta útil e amplamente utilizada em serviços de urgência e internação clínica, seus resultados não demonstraram superioridade consistente em relação a outras escalas clínicas validadas.

O NEWS2 apresentou sensibilidade média de 78,2%, um valor satisfatório, mas inferior ao de escalas como SMART-COP e A-DROP, que atingiram até 83% em determinadas populações. Além disso, a área sob a curva ROC (AUROC) média do NEWS2 para predição de mortalidade de 0,76, indicou desempenho moderado. Escalas como A-DROP (0,87), REMS (0,882) e PSI mostraram desempenho superior em diferentes contextos clínicos, especialmente em pacientes com condições respiratórias ou cardiovasculares específicas.

Portanto, embora o NEWS2 seja eficiente para triagem inicial, os dados não sustentam sua superioridade estatística como ferramenta preditiva universal.

Em relação à admissão em UTI, a taxa combinada foi de 20% (IC 95%: 9,2%–30,7%), com variação significativa entre os estudos, refletindo diferenças em protocolos assistenciais e perfis populacionais.

Na comparação da acurácia preditiva do NEWS2 com outras escalas prognósticas para a necessidade de ventilação assistida, a AUROC média ponderada de 0,77 apresentou bom desempenho. Contudo, escalas mais específicas, como DECAF e mDECAF, demonstraram superioridade, com AUROC média ponderada de 0,89 e 0,88, respectivamente, especialmente na previsão da necessidade de ventilação invasiva. Contudo, essa análise é exploratória, uma vez que se baseia em apenas um estudo (Gomes et al., 2023), o que limita a robustez estatística e a generalização dos achados.

Assim, embora a NEWS2 se mostre promissora, sua aplicação isolada ainda requer validação adicional em populações mais heterogêneas.

Relativo à mortalidade hospitalar, com proporção combinada de 9,2% (IC 95%: 4,3%–14,1%), não há evidências de que o NEWS2 tenha melhor capacidade discriminatória do que escalas específicas como o sPESI, amplamente validado para pacientes com embolia pulmonar, ou o DECAF, aplicado em DPOC.

A avaliação da qualidade da evidência segundo o GRADE apontou certeza moderada para os desfechos de mortalidade e admissão em UTI, e baixa a moderada para sensibilidade diagnóstica, uma vez que os estudos apresentaram elevada heterogeneidade e os dados comparativos não confirmaram superioridade robusta do NEWS2.

LIMITAÇÕES

Apesar dos importantes resultados obtidos neste estudo, a metanálise apresenta limitações que devem ser consideradas na interpretação crítica dos resultados.

Se por um lado todos os estudos incluídos possuíram delineamento observacional, que pode apresentar viés e reduzir a relação de causa e efeito, todos eram estudos de coorte, o que aumenta a robustez entre os estudos observacionais, e estudos de vida real, ponto que pode valorizar o resultado de efetividade.

Outro ponto foi a heterogeneidade metodológica relevante entre os estudos, com variações nos pontos de corte do score NEWS 2 (como ≥ 5 ou ≥ 7), nas populações estudadas e nos contextos clínicos em que o score foi aplicado. Essa variabilidade pode comprometer a comparabilidade entre os estudos e a generalização dos resultados, refletindo-se em índices elevados de heterogeneidade (I^2), especialmente nas análises de AUROC e sensibilidade.

Outro ponto importante foi que apenas um dos estudos incluídos avaliou diretamente os impactos organizacionais da implementação do NEWS2, como a sobrecarga de alertas, o aumento da carga de trabalho da equipe clínica e os efeitos sobre o tempo de resposta. Esse

aspecto é particularmente relevante em sistemas de saúde com recursos limitados, nos quais a adoção de sistemas de alerta deve considerar não apenas a acurácia, mas também a viabilidade assistencial.

Além disso, em algumas análises específicas, como ventilação mecânica, internação prolongada ou predição de eventos em subgrupos clínicos, a metanálise contou com número reduzido de estudos ($k = 2$ ou 3), o que limita a robustez estatística e reduz a confiabilidade das estimativas agregadas.

O potencial viés de publicação também deve ser considerado, uma vez que os gráficos de funil evidenciaram assimetria em diversos desfechos, sugerindo sub-representação de estudos com resultados neutros ou desfavoráveis.

Por fim, destaca-se que muitos estudos incluídos não relataram de forma clara estratégias para controle de fatores de confusão, nem abordaram adequadamente perdas de seguimento. Essas lacunas metodológicas podem comprometer a validade interna e a precisão das estimativas dos efeitos observados.

CONCLUSÃO

Os resultados desta revisão sistemática indicam que o escore NEWS 2 é uma ferramenta clinicamente válida para a triagem de pacientes hospitalizados, demonstrando alta sensibilidade na identificação precoce de deterioração clínica, especialmente em ambientes clínicos gerais. Contudo, com moderado nível de confiança, não evidenciamos superioridade estatisticamente consistente do NEWS 2 em comparação com escalas prognósticas específicas, como A-DROP, SMART-COP, DECAF, sPESI e TRIAGE, que apresentaram desempenho superior em cenários clínicos particulares, como DPOC, embolia pulmonar e pneumonia grave.

Diante desses achados, recomenda-se que o NEWS 2 seja utilizado como instrumento complementar, integrado aos protocolos institucionais e, sempre que possível, associado a outras escalas prognósticas adequadas ao perfil da população e ao contexto clínico específico. O uso isolado do NEWS 2, sobretudo com pontos de corte padronizados, pode aumentar a incidência de falsos positivos, gerar sobrecarga nas equipes assistenciais e comprometer a eficiência das decisões clínicas.

Futuras pesquisas devem priorizar a avaliação da performance operacional do NEWS 2 em cenários reais, incluindo sua integração a sistemas de suporte à decisão baseados em inteligência artificial, bem como sua validação em populações diversas e em contextos de recursos limitados. A personalização e o refinamento do escore, aliados a modelos preditivos

mais sofisticados, têm potencial para ampliar sua efetividade e fortalecer a segurança e a qualidade da gestão clínica em hospitais de diferentes níveis de complexidade.

Referências

1. Honarmand K, Wax RS, Penoyer D, Lighthall G, Danesh V, Rochweg B, Cheatham ML, Davis DP, DeVita M, Downar J, Edelson D, Fox-Robichaud A, Fujitani S, Fuller RM, Haskell H, Inada-Kim M, Jones D, Kumar A, Olsen KM, Rowley DD, Welch J, Baldisseri MR, Kellett J, Knowles H, Shipley JK, Kolb P, Wax SP, Hecht JD, Sebat F. Society of Critical Care Medicine Guidelines on Recognizing and Responding to Clinical Deterioration Outside the ICU: 2023. *Crit Care Med.* 2024 Feb 1;52(2):314-330. doi: 10.1097/CCM.0000000000006072. Epub 2024 Jan 19. PMID: 38240510.
2. Gerry, S., Bonnici, T., Birks, J., Kirtley, S., Virdee, P. S., Watkinson, P. J., & Collins, G. S. (2020). Early warning scores for detecting deterioration in adult hospital patients: a systematic review protocol. *BMJ Open*, 10(6), e033038. <https://doi.org/10.1136/bmjopen-2019-033038>
3. Boultsakis Logothetis, S., Li, X., Ashfaq, A., Guo, Y., & Zhang, P. (2023). Predicting acute clinical deterioration with interpretable machine learning to support emergency care decision making. *Scientific Reports*, 13(1), Article 20110. <https://doi.org/10.1038/s41598-023-47238-y>
4. Colombo, C. J., Kochav, S., Altman, M., Sulaiman, A., Lane, K., Palacios, C., ... & McDermott, M. (2021). Performance analysis of the National Early Warning Score and Modified Early Warning Score in the Adaptive COVID-19 Treatment Trial cohort. *Critical Care Explorations*, 3(7), e0474. <https://doi.org/10.1097/CCE.0000000000000474>
5. Covino, M., De Matteis, G., Santoro, M., Sabia, L., Simeoni, B., Candelli, M., ... & Franceschi, F. (2020). Predicting intensive care unit admission and death for COVID-19 patients in the emergency department using early warning scores. *Resuscitation*, 156, 84–91. <https://doi.org/10.1016/j.resuscitation.2020.09.136>
6. Cuthbertson, B. H., Boroujerdi, M., McKie, L., Aucott, L., & Prescott, G. (2007). Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? *Critical Care Medicine*, 35(2), 402–409. <https://doi.org/10.1097/01.CCM.0000254826.10520.87>
7. Cuthbertson, B. H., Boroujerdi, M., McKie, L., Aucott, L., & Prescott, G. (2007). Can physiological variables and early warning scoring systems allow early recognition of the

- deteriorating surgical patient? *Critical Care Medicine*, 35(2), 402–409. <https://doi.org/10.1097/01.CCM.0000254826.10520.87>
8. Pirneskoski, J., Kuisma, M., Nurmi, J., & Olkkola, K. T. (2019). Prehospital National Early Warning Score predicts early mortality. *Acta Anaesthesiologica Scandinavica*, 63(5), 676–683. <https://doi.org/10.1111/aas.13335>
 9. Forster, S., McKeever, T. M., & Shaw, D. (2022). Effect of implementing the NEWS2 escalation protocol in a large acute NHS trust: A retrospective cohort analysis of mortality, workload and ability of early warning score to predict death within 24 hours. *BMJ Open*, 12(11), e064579. <https://doi.org/10.1136/bmjopen-2022-064579>
 10. Martín-Rodríguez, F., López-Izquierdo, R., Delgado-Benito, J. F., Martín-Bayón, J. A., Bellido-Esteban, A., & Mohedano-Moriano, A. (2020). Can the prehospital National Early Warning Score 2 identify patients at risk of in-hospital early mortality? A prospective, multicenter cohort study. *Heart & Lung*, 49(5), 585–591. <https://doi.org/10.1016/j.hrtlng.2020.05.001>
 11. Gerry, S., Bonnici, T., Birks, J., Kirtley, S., Virdee, P. S., Watkinson, P. J., & Collins, G. S. (2017). Early warning scores for detecting deterioration in adult hospital patients: A systematic review protocol. *BMJ Open*, 7(12), e019268. <https://doi.org/10.1136/bmjopen-2017-019268>
 12. Gerry, S., Bonnici, T., Birks, J., Kirtley, S., Virdee, P. S., Watkinson, P. J., & Collins, G. S. (2020). Early warning scores for detecting deterioration in adult hospital patients: Systematic review and critical appraisal of methodology. *BMJ*, 369, m1501. <https://doi.org/10.1136/bmj.m1501>
 13. Gonem, S., Hainsworth, R., McCarthy, A., & Gautam, V. (2024). Real-world implementation of the National Early Warning Score-2 in an acute respiratory unit. *BMJ Open Respiratory Research*, 11(1), e002279. <https://doi.org/10.1136/bmjresp-2023-002279>
 14. Hwang, J. I., Kim, K. S., & Kim, J. (2023). Performance of Early Warning Scoring Systems Regarding Adverse Events of Unanticipated Clinical Deterioration in Complementary and Alternative Medicine Hospitals. *Asian Nursing Research*, 17(2), 110–117. <https://doi.org/10.1016/j.anr.2023.03.001>
 15. Cummings, B. C., Blackmer, J. M., Motyka, J. R., Farzaneh, N., Cao, L., Bisco, E. L., Glassbrook, J. D., Roebuck, M. D., Gillies, C. E., Admon, A. J., Medlin, R. P. Jr., Singh, K., Sjoding, M. W., Ward, K. R., & Ansari, S. (2023). External validation and comparison of a general ward deterioration index between diversely different health systems. *Critical Care Medicine*, 51(6), 775–786. <https://doi.org/10.1097/CCM.0000000000005837>

16. Candel, B. G. J., Nissen, S. K., Nickel, C. H., Raven, W., Thijssen, W., Gaakeer, M. I., Lassen, A. T., Brabrand, M., Steyerberg, E. W., de Jonge, E., & de Groot, B. (2023). Development and external validation of the International Early Warning Score for improved age- and sex-adjusted in-hospital mortality prediction in the emergency department. *Critical Care Medicine*, 51(7), 881–891. <https://doi.org/10.1097/CCM.0000000000005842>
17. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.3 (updated February 2022). Cochrane, 2022. Available from www.training.cochrane.org/handbook
18. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021 Mar 29;372:n71. doi: 10.1136/bmj.n71. PMID: 33782057; PMCID: PMC8005924
19. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev* Dec 5;5(1):210, 2016. doi: 10.1186/s13643-016-0384-4. PMID: 27919275; PMCID: PMC5139140
20. JOANNA BRIGGS INSTITUTE. Joanna Briggs Institute Reviewers' Manual: 2014 edition. The Joanna Briggs Institute. JBI, 2014
21. Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327(7414), 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
22. Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, 336(7650), 924–926. <https://doi.org/10.1136/bmj.39489.470347.AD>
23. Forster, S., McKeever, T. M., & Shaw, D. (2022). Effect of implementing the NEWS2 escalation protocol in a large acute NHS trust: A retrospective cohort analysis of mortality, workload and ability of early warning score to predict death within 24 hours. *BMJ Open*, 12(11), e064579. <https://doi.org/10.1136/bmjopen-2022-064579>
24. Fan, G., Tu, C., Zhou, F., Liu, Z., Wang, Y., Song, B., Gu, X., Wang, Y., Wei, Y., Li, H., Wu, X., Xu, J., Tu, S., Zhang, Y., Wu, W., & Cao, B. (2020). Comparison of severity scores for COVID-19 patients with pneumonia: a retrospective study. *European Respiratory Journal*, 56(3), 2002113. <https://doi.org/10.1183/13993003.02113-2020>
25. Rodríguez, C., Callejas, A., Jiménez, D., Tovt, I., García, L., Yusen, R. D., & Lobo, J. L. (2023). Usefulness of the National Early Warning Score for risk stratification of stable

- patients with acute symptomatic pulmonary embolism. *Archivos de Bronconeumología*, 59(3), 152–156. <https://doi.org/10.1016/j.arbres.2022.04.019>
26. Covino, M., Sandroni, C., Santoro, M., Sabia, L., Simeoni, B., Pucciarelli, A., ... & Franceschi, F. (2020). Predicting intensive care unit admission and death for COVID-19 patients in the emergency department using early warning scores. *Resuscitation*, 156, 84–91. <https://doi.org/10.1016/j.resuscitation.2020.09.126>
 27. Pugazhvannan, C. R., Vanidassane, I., Pownraj, D., Kandasamy, R., & Basheer, A. (2021). National Early Warning Score 2 (NEWS2) to predict poor outcome in hospitalised COVID-19 patients in India. *PLOS ONE*, 16(12), e0261376. <https://doi.org/10.1371/journal.pone.0261376>
 28. Baker, K. F., Hanrath, A. T., van der Loeff, I. S., Kay, L. J., Back, J., & Duncan, C. J. (2021). National Early Warning Score 2 (NEWS2) to identify inpatient COVID-19 deterioration: A retrospective analysis. *Clinical Medicine*, 21(2), 84–89. <https://doi.org/10.7861/clinmed.2020-0688>
 29. Aliberti, M. J. R., Covinsky, K. E., Garcez, F. B., Smith, A. K., Curiati, P. K., Lee, S. J., Dias, M. B., Melo, V. J. D., Rego-Júnior, O. F., Richinho, V. P., Jacob-Filho, W., & Avelino-Silva, T. J. (2021). A fuller picture of COVID-19 prognosis: The added value of vulnerability measures to predict mortality in hospitalised older adults. *Age and Ageing*, 50(1), 32–39. <https://doi.org/10.1093/ageing/afaa240>
 30. Gomes, L., Pereira, S., Sousa-Pinto, B., & Rodrigues, C. (2023). Performance of risk scores in patients with acute exacerbations of COPD. *Jornal Brasileiro de Pneumologia*, 49(5), e20230032. <https://doi.org/10.36416/1806-3756/e20230032>
 31. De Socio, G. V., Gidari, A., Sicari, F., Palumbo, M., & Francisci, D. (2021). National Early Warning Score 2 (NEWS2) better predicts critical Coronavirus Disease 2019 (COVID-19) illness than COVID-GRAM: A multi-centre study. *Infection*, 49(5), 1033–1038. <https://doi.org/10.1007/s15010-021-01620-x>
 32. Martín-Rodríguez, F., Sanz-García, A., Medina-Lozano, E., Castro Villamor, M. Á., Carbajosa Rodríguez, V., Del Pozo Vegas, C., Fadrique Millán, L. N., Rabbione, G. O., Martín-Conty, J. L., & López-Izquierdo, R. (2020). Can the prehospital National Early Warning Score 2 identify patients at risk of in-hospital early mortality? A prospective, multicenter cohort study. *Heart & Lung*, 49(5), 585–591. <https://doi.org/10.1016/j.hrtlng.2020.02.047>

33. Pugazhvannan, C. R., Vanidassane, I., Pownraj, D., Kandasamy, R., & Basheer, A. (2021). National Early Warning Score 2 (NEWS2) to predict poor outcome in hospitalised COVID-19 patients in India. PLOS ONE, 16(12), e0261376. <https://doi.org/10.1371/journal.pone.0261376>

5 CONCLUSÃO

Os resultados desta revisão sistemática com metanálise demonstraram que o NEWS2 possui aplicabilidade clínica relevante como ferramenta de triagem e detecção precoce da deterioração clínica em adultos hospitalizados. A análise integrada dos estudos indicou que o NEWS2 apresenta sensibilidade moderada e capacidade discriminativa satisfatória para predição de desfechos graves, como mortalidade hospitalar e necessidade de admissão em UTI.

Entretanto, os achados também evidenciaram que o NEWS2 não é consistentemente superior a outras escalas clínicas amplamente validadas, como A-DROP, SMART-COP, DECAF, PSI e TRIAGE. Essas escalas apresentaram melhor desempenho em desfechos específicos, especialmente em subgrupos clínicos como pacientes com doenças respiratórias ou cardiovasculares agudas.

A avaliação da qualidade da evidência pela metodologia GRADE reforçou a robustez dos achados para os desfechos de mortalidade hospitalar e admissão em UTI (certeza moderada), mas indicou certeza baixa a moderada para sensibilidade diagnóstica, principalmente devido à heterogeneidade metodológica entre os estudos e à ausência de superioridade estatística sustentada.

Portanto, conclui-se que o NEWS2 deve ser utilizado como ferramenta complementar, e não exclusiva, na estratificação de risco e tomada de decisão clínica. Seu uso deve ser integrado a outros escores prognósticos e a sistemas de alerta precoce institucionalizados, especialmente em ambientes de alta complexidade. Adicionalmente, futuros estudos devem investigar a efetividade operacional do NEWS2, incluindo seu impacto nos fluxos assistenciais, no tempo de resposta clínica e na segurança do paciente em diferentes contextos hospitalares.

REFERÊNCIAS

- ALIBERTI, Márton Juliano Romero et al. A fuller picture of COVID-19 prognosis: the added value of vulnerability measures to predict mortality in hospitalised older adults. **Age and Ageing**, v. 50, n. 1, p. 32-39, 2021.
- BAKER, Kenneth F. et al. National Early Warning Score 2 (NEWS2) to identify inpatient COVID-19 deterioration: a retrospective analysis. **Clinical Medicine**, v. 21, n. 2, p. 84-89, 2021.
- BOULITSAKIS LOGOTHETIS, S. *et al.* Predicting acute clinical deterioration with interpretable machine learning to support emergency care decision making. **Scientific Reports**, v. 13, n. 1, 1 dez. 2023.
- CANDEL, B. G. J. et al. Development and external validation of the International Early Warning Score for improved age- and sex-adjusted in-hospital mortality prediction in the emergency department. **Critical Care Medicine**, Baltimore, v. 51, n. 7, p. 881–891, 1 jul. 2023. DOI: <https://doi.org/10.1097/CCM.0000000000005842>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/36951452/>. Acesso em: 31 mar. 2025.
- COLOMBO, C. J. *et al.* Performance Analysis of the National Early Warning Score and Modified Early Warning Score in the Adaptive COVID-19 Treatment Trial Cohort. **Critical Care Explorations**, v. 3, n. 7, p. E0474, 13 jul. 2021.
- COVINO, M. *et al.* Predicting intensive care unit admission and death for COVID-19 patients in the emergency department using early warning scores. **Resuscitation**, v. 156, p. 84- 91, 2020.
- CUMMINGS, B. C. *et al.* External validation and comparison of a general ward deterioration index between diversely different health systems. **Critical Care Medicine**, Baltimore, v. 51, n. 6, p. 775–786, 1 jun. 2023. DOI: <https://doi.org/10.1097/CCM.0000000000005837>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/36927631/>. Acesso em: 31 mar. 2025.
- CUTHBERTSON, B. H. *et al.* Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? **Critical Care Medicine**, v. 35, n. 2, p. 402–409, fev. 2007.
- DE OLIVEIRA, A. P. *et al.* National Early Warning Score 2–versão brasileira: validade preditiva para adultos com COVID-19. **Revista de Enfermagem da UFSM**, v. 13, p. e14-e14, 2023.

DE SOCIO, G. V. *et al.* National Early Warning Score 2 (NEWS2) better predicts critical Coronavirus Disease 2019 (COVID-19) illness than COVID-GRAM, a multi-centre study. **Infection**, v. 49, n. 5, p. 1033-1038, 2021.

FAN, G. *et al.* Comparison of severity scores for COVID-19 patients with pneumonia: a retrospective study. **European Respiratory Journal**, v. 56, n. 3, 2020.

FORSTER, S; MCKEEVER, M.; SHAW, D. Effect of implementing the NEWS2 escalation protocol in a large acute NHS trust: a retrospective cohort analysis of mortality, workload and ability of early warning score to predict death within 24 hours. **BMJ open**, v. 12, n. 11, p. e064579, 2022.

GERRY, S. *et al.* Early warning scores for detecting deterioration in adult hospital patients: A systematic review protocol. **BMJ Open BMJ Publishing Group**, v.7, n.12, 1 dez. 2017.

GERRY, S. *et al.* Early warning scores for detecting deterioration in adult hospital patients: Systematic review and critical appraisal of methodology. **The BMJ Publishing Group**, v. 369, n. 3, p. 1-16, 20 maio 2020.

GOMES, Lúcia et al. Performance of risk scores in patients with acute exacerbations of COPD. **Jornal Brasileiro de Pneumologia**, v. 49, n. 5, p. e20230032, 2023.

GONEM, S. *et al.* **Real-world implementation of the National Early Warning Score-2 in an acute respiratory unit.** **BMJ Open Respiratory Research**, v. 11, n. 1, 31 jan. 2024.

HIGGINS, J. P. T. *et al.* Measuring inconsistency in meta-analyses. **British Medical Journal BMJ Publishing Group**, v.327, n. 7414, p.557 560, 6 set. 2003.

HWANG, J. I. *et al.* Performance of Early Warning Scoring Systems Regarding Adverse Events of Unanticipated Clinical Deterioration in Complementary and Alternative Medicine Hospitals. **Asian Nursing Research**, v. 17, n. 2, p. 110–117, 1 maio 2023.

JULIAN HIGGINS; JONATHAN DEEKS; DOUGLAS G ALTMAN. **Statistical Methods Group of The Cochrane Collaboration.** 2020. Disponível em:

<<https://training.cochrane.org/handbook/current/chapter-10>>. Acesso em: 29 dez. 2023.

LALUEZA, A. *et al.* The early use of sepsis scores to predict respiratory failure and mortality in non-ICU patients with COVID-19. **Revista Clínica Espanola**, v. 222, n. 5, p. 293–298, 1 maio. 2022.

LEUNG, S. C. *et al.* Can prehospital Modified Early Warning Score identify non-trauma patients requiring life-saving intervention in the emergency department? **EMA - Emergency Medicine Australasia**, v. 28, n. 1, p. 84–89, 1 fev. 2016.

MARTÍN-RODRÍGUEZ, Francisco *et al.* Can the prehospital National Early Warning Score 2 identify patients at risk of in-hospital early mortality? A prospective, multicenter cohort study. **Heart & Lung**, v. 49, n. 5, p. 585-591, 2020.

MARTINS, R. C.; PORTO RUTZ, A.; BRAZOROTTO SOARES, J. Formação Continuada Remota para os Cuidados Agudos do National Early Warning Score 2: Adaptação Transcultural. **Revista Nursing**, v. 289, n. 25, p. 7952–7959, 2022.

MCMASTER UNIVERSITY. **Guideline Development Tool**. Disponível em: <<https://cebgrade.mcmaster.ca/gradepro.html>>. Acesso em: 28 dez. 2023.

MENDELEY. **Mendeley - Reference Management Software**. <https://www.mendeley.com>, 2023. Disponível em: <Mendeley - Reference Management Software>. Acesso em: 28 dez. 2023.

MUNN, Z. *et al.* **JBIM Manual for Evidence Synthesis**. JBI, 2020. Disponível em: <https://jbi.global/critical-appraisal-tools>. Acesso em: 29 dez.2023.

OUZZANI, M. *et al.* **Rayyan-a web and mobile app for systematic reviews**. *Systematic Reviews*, v. 5, n. 1, 5 dez. 2016.

PAGE, M. J. *et al.* The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. **The BMJBMJ Publishing Group**, 29 mar. 2021.

PIRNESKOSKI, J. *et al.* Prehospital National Early Warning Score predicts early mortality. **Acta Anaesthesiologica Scandinavica**, v. 63, n. 5, p. 676–683, 1 maio. 2019.

PUGAZHVANNAN, C.R *et al.* National Early Warning Score 2 (NEWS2) to predict poor outcome in hospitalised COVID-19 patients in India. **PloS one**, v. 16, n. 12, p. e0261376, 2021.

R CORE DEVELOPMENT TEAM. R Core Team (2020) R: A Language and Environment for Statistical Computing. **R Foundation for Statistical Computing**. 2020. Disponível em: <<https://www.r-project.org/>>. Acesso em 29 dez.2023.

RAO, S. V *et al.* Prognostic Accuracy of News 2 in Geriatric Patients with COVID-19 in a Tertiary Care Hospital. **International Journal of Pharmaceutical and Clinical Research**, v. 15, n. 4, p. 427–434, 2023.

RODRÍGUEZ, Carmen et al. Usefulness of the National Early Warning Score for risk stratification of stable patients with acute symptomatic pulmonary embolism. **Archivos de Bronconeumología**, v. 59, n. 3, p. 152-156, 2023.

SILCOCK, D. J. *et al.* Validation of the National Early Warning Score in the prehospital setting. **Resuscitation**, v. 89, n. C, p. 31–35, 2015.

SMITH, G. B. *et al.* The ability of the National Early Warning Score (NEWS) to discriminate against patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. **Resuscitation**, v. 84, n. 4, p. 465–470, abr. 2013.

STROUP, D. F. *et al.* Meta-analysis of Observational Studies in Epidemiology. A Proposal for Reporting. **JAMA**, v.283, n.15, p. 2008-2012, 19 abril.2008. Disponível em:<<http://www.wesleyan.edu>>.

ZHANG, K. *et al.* The Prognostic Accuracy of National Early Warning Score 2 on Predicting Clinical Deterioration for Patients With COVID-19: A Systematic Review and Meta-Analysis. **Frontiers in Medicine**, v.8, 9 jul. 2021.

APÊNDICE A

Base de Dados: PubMed – 265 resultados

((("Patient Acuity" OR "Severity of Illness" OR "Illness Severities" OR "Illness Severity" OR "Disease Severity" OR "Disease Severities" OR "Severity, Disease" OR " Inpatients" OR " Inpatient" OR "Hospital Patient")) AND ("Early Warning Score" OR "Early Warning Scores" OR "Score, Early Warning" OR "Scores, Early Warning" OR "Warning Scores, Early")) AND ("Clinical Deterioration" OR "Clinical Deteriorations" OR "Deterioration, Clinical" OR "Disease Progression" OR "Progression, Disease" OR "Clinical Course" OR "Clinical Progression" OR "Progression, Clinical" OR "Disease Exacerbation" OR "Exacerbation, Disease" OR "Hospitalization" OR "Hospitalizations" OR "Vital Signs" OR "Sign, Vital" OR "Signs, Vital" OR "Vital Sign" OR "Clinical Evolution" OR "Decision Support Systems, Clinical" OR "Clinical Decision Support Systems" OR "Clinical Decision Support System" OR "Clinical Decision Support" OR "Clinical Decision Supports" OR "Decision Supports, Clinical" OR "Support, Clinical Decision" OR "Supports, Clinical Decision" OR "Decision Support, Clinical" OR "Clinical Decision-Making" OR "Clinical Decision Making" OR "Decision-Making, Clinical" OR "Medical Decision-Making" OR "Decision-Making, Medical" OR "Medical Decision Making" OR "Early Medical Intervention" OR "Early Medical Interventions" OR "Intervention, Early Medical" OR "Interventions, Early Medical" OR "Medical Intervention, Early" OR "Medical Interventions, Earl" OR "Patient Safety" OR "Patient Safeties" OR "Safeties, Patient" OR "Safety, Patient" OR "Effectiveness"))

APÊNDICE B

Base de Dados: Lilacs via Portal BVS – 241 resultados

("Gravidade do Paciente" OR "Pacientes Internados") AND (("Escore de Alerta Precoce" OR "Early Warning Score")) AND ("Deterioração Clínica" OR "Progressão da Doença" OR "Hospitalização" OR "Sinais Vitais" OR "Evolução Clínica" OR "Suporte à decisão clínica" OR "Tomada de Decisão Clínica" OR "Intervenção Médica Precoce" OR "Segurança do Paciente" OR "Efetividade") AND ("en" OR "es" OR "pt")

APÊNDICE C

Base de Dados: EMBASE – 79 resultados

('patient acuity'/exp OR 'patient acuity' OR 'severity of illness' OR 'illness severities' OR 'illness severity' OR 'disease severity'/exp OR 'disease severity' OR 'disease severities' OR 'severity, disease' OR 'inpatients' OR 'inpatient' OR 'hospital patient'/exp) AND ('early warning score'/exp OR 'early warning score' OR 'early warning scores' OR 'score, early warning' OR 'scores, early warning' OR 'warning scores, early') AND (('clinical deterioration' OR 'clinical deteriorations' OR 'deterioration, clinical' OR 'deterioration'/exp OR 'disease progression' OR 'progression, disease' OR 'clinical course' OR 'clinical progression' OR 'progression, clinical' OR 'disease exacerbation'/exp OR 'disease exacerbation' OR 'exacerbation, disease' OR 'hospitalization'/exp OR 'hospitalization' OR 'hospitalizations' OR 'vital signs' OR 'sign, vital' OR 'signs, vital' OR 'vital sign'/exp OR 'vital sign' OR 'clinical evolution' OR decision) AND 'support systems, clinical' OR 'clinical decision support systems' OR 'clinical decision support system'/exp OR 'clinical decision support system' OR 'clinical decision support' OR 'clinical decision supports' OR 'decision supports, clinical' OR 'support, clinical decision' OR 'supports, clinical decision' OR 'decision support, clinical' OR 'clinical decision-making' OR 'clinical decision making'/exp OR 'clinical decision making' OR 'decision-making, clinical' OR 'medical decision-making' OR 'decision-making, medical' OR 'medical decision making'/exp OR 'medical decision making' OR 'early intervention'/exp OR 'early medical intervention' OR 'early medical interventions' OR 'intervention, early medical' OR 'interventions, early medical' OR 'medical intervention, early' OR 'medical interventions, earl' OR 'patient safety'/exp OR 'patient safety' OR 'patient safeties' OR 'safeties, patient' OR 'safety, patient' OR 'clinical effectiveness'/exp OR 'effectiveness') AND [embase]/lim

APÊNDICE D

Base de Dados: Web of Science – 222 resultados

((ALL=(“Patient Acuity” OR “Severity of Illness” OR “Illness Severities” OR “Illness Severity” OR “Disease Severity” OR “Disease Severities” OR “Severity, Disease” OR “Inpatients” OR “ Inpatient” OR “Hospital Patient”)) AND ALL=(“Early Warning Score” OR “Early Warning Scores” OR “Score, Early Warning” OR “Scores, Early Warning” OR “Warning Scores, Early”)) AND ALL=(“Clinical Deterioration” OR “Clinical Deteriorations” OR “Deterioration, Clinical” OR “Disease Progression” OR “Progression, Disease” OR “Clinical Course” OR “Clinical Progression” OR “Progression, Clinical” OR “Disease Exacerbation” OR “Exacerbation, Disease” OR “Hospitalization” OR “Hospitalizations” OR “Vital Signs” OR “Sign, Vital” OR “Signs, Vital” OR “Vital Sign” OR “Clinical Evolution” OR “Decision Support Systems, Clinical” OR “Clinical Decision Support Systems” OR “Clinical Decision Support System” OR “Clinical Decision Support” OR “Clinical Decision Supports” OR “Decision Supports, Clinical” OR “Support, Clinical Decision” OR “Supports, Clinical Decision” OR “Decision Support, Clinical” OR “Clinical Decision-Making” OR “Clinical Decision Making” OR “Decision-Making, Clinical” OR “Medical Decision-Making” OR “Decision-Making, Medical” OR “Medical Decision Making” OR “Early Medical Intervention” OR “Early Medical Interventions” OR “Intervention, Early Medical” OR “Interventions, Early Medical” OR “Medical Intervention, Early” OR “Medical Interventions, Early” OR “Patient Safety” OR “Patient Safeties” OR “Safeties, Patient” OR “Safety, Patient” OR “Effectiveness”))

APÊNDICE E

Base de Dados: Google Scholar – 92 resultados

("Patient Acuity" OR "Inpatients" OR "Early Warning Score" OR "Clinical Deterioration" OR "Disease Progression" OR "Hospitalization" OR "Vital Signs" OR "Clinical Evolution" OR "Decision Support Systems, Clinical" OR "Clinical Decision-Making" OR "Early Medical Intervention" OR "Patient Safety" OR "Effectiveness" OR "Deterioration" OR "Disease Exacerbation" OR "Hospitalization" OR "Vital Sign" OR "Clinical Decision Support System" OR "Clinical Decision Making" OR "Medical Decision Making" OR "Early Intervention" OR "Patient Safety" OR "Clinical Effectiveness")
--

ANEXO A - CHECK LIST PRISMA FONTE: PRINCIPAIS ITENS PARA RELATAR

Seção/tópico	N.	Item do checklist
TÍTULO		
Título	1	Identifique o artigo como uma revisão sistemática, meta-análise, ou ambos.
RESUMO		
Resumo estruturado	2	Apresente um resumo estruturado incluindo, se aplicável: referencial teórico; objetivos; fonte de dados; critérios de elegibilidade; participantes e intervenções; avaliação do estudo e síntese dos métodos; resultados; limitações; conclusões e implicações dos achados principais; número de registro da revisão sistemática.
INTRODUÇÃO		
Racional	3	Descreva a justificativa da revisão no contexto do que já é conhecido.
Objetivos	4	Apresente uma afirmação explícita sobre as questões abordadas com referência a participantes, intervenções, comparações, resultados e delineamento dos estudos (PICOS).
MÉTODOS		
Protocolo e registro	5	Indique se existe um protocolo de revisão, se e onde pode ser acessado (ex. endereço eletrônico), e, se disponível, forneça informações sobre o registro da revisão, incluindo o número de registro.
Critérios de elegibilidade	6	Especifique características do estudo (ex.: PICOS, extensão do seguimento) e características dos relatos (ex. anos considerados, idioma, a situação da publicação) usadas como critérios de elegibilidade, apresentando justificativa.
Fontes de informação	7	Descreva todas as fontes de informação na busca (ex.: base de dados com datas de cobertura, contato com autores para identificação de estudos adicionais) e data da última busca.
Busca	8	Apresente a estratégia completa de busca eletrônica para pelo menos uma base de dados, incluindo os limites utilizados, de forma que possa ser repetida.
Seleção dos estudos	9	Apresente o processo de seleção dos estudos (isto é, rastreados, elegíveis, incluídos na revisão sistemática, e, se aplicável, incluídos na meta-análise).
Processo de coleta de dados	10	Descreva o método de extração de dados dos artigos (ex.: formulários piloto, de forma independente, em duplicata) e todos os processos para obtenção e confirmação de dados dos pesquisadores.
Lista dos dados	11	Liste e defina todas as variáveis obtidas dos dados (ex.: PICOS, fontes de financiamento) e quaisquer suposições ou simplificações realizadas.
Risco de viés em cada estudo	12	Descreva os métodos usados para avaliar o risco de viés em cada estudo (incluindo a especificação se foi feito no nível dos estudos ou dos resultados), e como esta informação foi usada na análise de dados.
Medidas de sumarização	13	Defina as principais medidas de sumarização dos resultados (ex.: risco relativo, diferença média).
Síntese dos resultados	14	Descreva os métodos de análise dos dados e combinação de resultados dos estudos, se realizados, incluindo medidas de consistência (por exemplo, I ²) para cada meta-análise.
Risco de viés entre estudos	15	Especifique qualquer avaliação do risco de viés que possa influenciar a evidência cumulativa (ex.: viés de publicação, relato seletivo nos estudos).
Análises adicionais	16	Descreva métodos de análise adicional (ex.: análise de sensibilidade ou análise de subgrupos, metarregressão), se realizados, indicando quais foram pré-especificados.
RESULTADOS		
Seleção de estudos	17	Apresente números dos estudos rastreados, avaliados para elegibilidade e incluídos na revisão, razões para exclusão em cada estágio, preferencialmente por meio de gráfico de fluxo.
Características dos estudos	18	Para cada estudo, apresente características para extração dos dados (ex.: tamanho do estudo, PICOS, período de acompanhamento) e apresente as citações.
Risco de viés em cada estudo	19	Apresente dados sobre o risco de viés em cada estudo e, se disponível, alguma avaliação em resultados (ver item 12).
Resultados de estudos individuais	20	Para todos os desfechos considerados (benefícios ou riscos), apresente para cada estudo: (a) sumário simples de dados para cada grupo de intervenção e (b) efeitos estimados e intervalos de confiança, preferencialmente por meio de gráficos de floresta.
Síntese dos resultados	21	Apresente resultados para cada meta-análise feita, incluindo intervalos de confiança e medidas de consistência.
Risco de viés entre estudos	22	Apresente resultados da avaliação de risco de viés entre os estudos (ver item 15).
Análises adicionais	23	Apresente resultados de análises adicionais, se realizadas (ex.: análise de sensibilidade ou subgrupos, metarregressão [ver item 16]).
DISCUSSÃO		
Sumário da evidência	24	Sumarize os resultados principais, incluindo a força de evidência para cada resultado; considere sua relevância para grupos-chave (ex.: profissionais da saúde, usuários e formuladores de políticas).
Limitações	25	Discuta limitações no nível dos estudos e dos desfechos (ex.: risco de viés) e no nível da revisão (ex.: obtenção incompleta de pesquisas identificadas, viés de relato).
Conclusões	26	Apresente a interpretação geral dos resultados no contexto de outras evidências e implicações para futuras pesquisas.
FINANCIAMENTO		
Financiamento	27	Descreva fontes de financiamento para a revisão sistemática e outros suportes (ex.: suprimento de dados); papel dos financiadores na revisão sistemática.

REVISÕES sistemáticas e Meta-análises: A recomendação PRISMA, 2015.

ANEXO B – JBI CRITICAL APPRAISAL CHECKLIST FOR COHORT STUDIES

JBI CRITICAL APPRAISAL CHECKLIST FOR COHORT STUDIES

Reviewer _____ Date _____

Author _____ Year _____ Record Number _____

	Yes	No	Unclear	Not applicable
1. Were the two groups similar and recruited from the same population?	☐	☐	☐	☐
2. Were the exposures measured similarly to assign people to both exposed and unexposed groups?	☐	☐	☐	☐
3. Was the exposure measured in a valid and reliable way?	☐	☐	☐	☐
4. Were confounding factors identified?	☐	☐	☐	☐
5. Were strategies to deal with confounding factors stated?	☐	☐	☐	☐
6. Were the groups/participants free of the outcome at the start of the study (or at the moment of exposure)?	☐	☐	☐	☐
7. Were the outcomes measured in a valid and reliable way?	☐	☐	☐	☐
8. Was the follow up time reported and sufficient to be long enough for outcomes to occur?	☐	☐	☐	☐
9. Was follow up complete, and if not, were the reasons to loss to follow up described and explored?	☐	☐	☐	☐
10. Were strategies to address incomplete follow up utilized?	☐	☐	☐	☐
11. Was appropriate statistical analysis used?	☐	☐	☐	☐

Overall appraisal: Include ☐ Exclude ☐ Seek further info ☐

Comments (Including reason for exclusion)

Fonte: MOOLA, S. et al. *JBI Critical Appraisal Checklist for Cohort Studies*, 2017.


NIHR | National Institute for Health and Care Research

PROSPERO
 International prospective register of systematic reviews

[Home](#) | [About PROSPERO](#) | [How to register](#) | [Service information](#)

[Search](#) | [My PROSPERO](#) | Logout: **Kerzia Diniz**

Register your review now

Edit your details

You have 1 records

My other records

These are records that have either been published or rejected and are not currently being worked on.

ID	Title	Status	Last edited
CRD42024496336	EFFECTIVENESS OF THE NATIONAL EARLY WARNING SCORE (NEWS 2) IN EARLY DETECTION OF CLINICAL DETERIORATION IN HOSPITALIZED PATIENTS: A SYSTEMATIC REVIEW	Registered	15/01/2024 

78

ANEXO D – LISTAGEM DE ESTUDOS EXCLUÍDOS

Devido a grande quantidade de estudos que compõem a lista, a mesma pode ser acessada na nuvem através do link disponibilizado abaixo:

<https://docs.google.com/spreadsheets/d/1dTIpVx2rHqQ8tLn4LNLD76zs0YSr0lbN/edit?gid=958452046#gid=958452046>

ANEXO E – RAZÕES DE EXCLUSÃO DE ARTIGOS POR ETAPA DE LEITURA

RAZÕES DE EXCLUSÃO	ETAPA DE LEITURA			TOTAL EXCLUÍDOS
	<i>Títulos</i>	<i>Resumo</i>	<i>Completa</i>	
Não se refere à pergunta de pesquisa	142	89	4	235
Não se refere à população de estudo	-	127	3	130
Não contempla a intervenção de estudo	-	66	2	68
Não é o tipo de estudo da pesquisa	15	64	2	81
Não apresenta as variáveis de desfecho	-	42	3	45
TOTAL	157	388	14	559

Fonte: A autoria dos pesquisadores (2025).

ANEXO F – ESTATÍSTICA DE HETEROGENEIDADE

Fonte: Aatoria dos pesquisadores (2025).

Estatísticas de heterogeneidade dos desfechos analisados na meta-análise

Fonte: Aatoria dos pesquisadores (2025).

Desfecho	Tau	Tau ² (SE)	I ² (%)	H ²	Q	df	p-valor	Classificação I ²
Taxa de Internação em UTI	0.130	0.0169	79,88	4.975	29.851	6	< 0.001	Alta
Taxa de Mortalidade Hospitalar	0.023	0.00053 (SE ≈ 0.0013)	45,2	1.824	10.946	6	0.090	Moderada
Mortalidade por Intercorrências	0.108	0.0117 (SE ≈ 0.0047)	75,2	4.032	20.130	4	0.0005	Alta