



OPEN ACCESS

Prehospital early warning scores for adults with suspected sepsis: retrospective diagnostic cohort study

Steve Goodacre ,¹ Laura Sutton,¹ Ben Thomas ,¹ Olivia Hawksworth ,¹ Khurram Iftikhar,² Susan Croft,² Gordon Fuller,¹ Simon Waterhouse,¹ Daniel Hind,¹ Mike Bradburn,¹ Michael Anthony Smyth ,³ Gavin D Perkins,³ Mark Millins,⁴ Andy Rosser,⁵ Jon M Dickson ,¹ Matthew Joseph Wilson¹

Handling editor Edward Carlton

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/emered-2023-213315>).

¹Sheffield Centre for Health and Related Research (SCHARR), The University of Sheffield, Sheffield, UK

²Emergency Department, Northern General Hospital, Sheffield, UK

³Clinical Trials Unit, University of Warwick, Coventry, UK

⁴Yorkshire Ambulance Service NHS Trust, Wakefield, UK

⁵West Midlands Ambulance Service, West Midlands, UK

Correspondence to

Professor Steve Goodacre, The University of Sheffield, Sheffield S10 2TN, UK; s.goodacre@sheffield.ac.uk

Received 28 April 2023

Accepted 18 August 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY. Published by BMJ.

To cite: Goodacre S, Sutton L, Thomas B, et al. *Emerg Med J* Epub ahead of print: [please include Day Month Year]. doi:10.1136/emered-2023-213315

ABSTRACT

Background Ambulance services need to identify and prioritise patients with sepsis for early hospital assessment. We aimed to determine the accuracy of early warning scores alongside paramedic diagnostic impression to identify sepsis that required urgent treatment.

Methods We undertook a retrospective diagnostic cohort study involving adult emergency medical cases transported to Sheffield Teaching Hospitals ED by Yorkshire Ambulance Service in 2019. We used routine ambulance service data to calculate 21 early warning scores and categorise paramedic diagnostic impressions as sepsis, infection, non-specific presentation or other presentation. We linked cases to hospital records and identified those meeting the sepsis-3 definition who received urgent hospital treatment for sepsis (reference standard). Analysis determined the accuracy of strategies that combined early warning scores at varying thresholds for positivity with paramedic diagnostic impression.

Results We linked 12 870/24 955 (51.6%) cases and identified 348/12 870 (2.7%) with a positive reference standard. None of the strategies provided sensitivity greater than 0.80 with positive predictive value greater than 0.15. The area under the receiver operating characteristic curve for the National Early Warning Score, version 2 (NEWS2) applied to patients with a diagnostic impression of sepsis or infection was 0.756 (95% CI 0.729, 0.783). No other early warning score provided clearly superior accuracy to NEWS2. Paramedic impression of sepsis or infection had sensitivity of 0.572 (0.519, 0.623) and positive predictive value of 0.156 (0.137, 0.176). NEWS2 thresholds of >4, >6 and >8 applied to patients with a diagnostic impression of sepsis or infection, respectively, provided sensitivities and positive predictive values of 0.522 (0.469, 0.574) and 0.216 (0.189, 0.245), 0.447 (0.395, 0.499) and 0.274 (0.239, 0.313), and 0.314 (0.268, 0.365) and 0.333 (0.284, 0.386).

Conclusion No strategy is ideal but using NEWS2 alongside paramedic diagnostic impression of infection or sepsis could identify one-third to half of sepsis cases without prioritising unmanageable numbers. No other score provided clearly superior accuracy to NEWS2.

Trial registration number [researchregistry5268](https://www.researchregistry.com/browse-the-registry#/home/registrationdetails/5de7bbd97ca5b50015041c33/), <https://www.researchregistry.com/browse-the-registry#/home/registrationdetails/5de7bbd97ca5b50015041c33/>

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Guidelines for sepsis recommend urgent treatment within 1 hour for people with suspected sepsis who are at highest risk. Ambulance services can use early warning scores alongside paramedic diagnostic impression to identify and prioritise people with suspected sepsis.

WHAT THIS STUDY ADDS

⇒ This retrospective diagnostic cohort study of 12 870 patients showed that no combination of early warning score alongside diagnostic impression provides sensitivity greater than 0.80 with positive predictive value greater than 0.15. Using the National Early Warning Score, version 2 (NEWS2) at thresholds of >4 to >8 in patients with a diagnostic impression of infection or sepsis could identify one-third to half of sepsis cases without prioritising unmanageable numbers. No alternative early warning score provided clearly superior accuracy to NEWS2.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Ambulance services and hospitals can use the estimates of NEWS2 sensitivity and positive predictive value to identify an appropriate NEWS2 threshold score to guide the use of prealerts for patients with suspected sepsis.

INTRODUCTION

Sepsis is a life-threatening response to a severe infection, which can lead to tissue damage, organ failure and death.¹ Guidelines for sepsis highlight the importance of early recognition and treatment, with treatment recommended within 1 hour of presentation for those at highest risk.¹⁻⁴ The emergency care system can only achieve this if sepsis is recognised and prioritised. This may involve ambulance services prealerting the ED that they are transporting a patient with suspected sepsis. However, prioritising too many patients with suspected sepsis may delay assessment of other urgent cases or may result in a lack of meaningful prioritisation.

the sepsis-3 definition³⁰) within 4 hours of initial assessment; and (2) treatment for sepsis given within 4 hours. Evidence of infection could include microbiology reports identifying organisms, radiology reports identifying infective changes or other markers strongly suggesting infection. Organ dysfunction was defined as a Sequential (sepsis-related) Organ Failure Assessment (SOFA) score of 2 or more points worse than normal. We estimated the SOFA score using the ED observations chart and first blood results after admission. In accordance with the sepsis-3 definition,³⁰ we assumed the normal SOFA score would be 0 unless there was evidence in the hospital records to suggest otherwise. Treatment for sepsis was based on relevant guidelines¹² and typically involved intravenous antibiotic therapy. One of the experts also estimated the Clinical Frailty Score using information in the hospital records.³¹

If the two reviewers disagreed on the overall sepsis-3 judgement or whether urgent treatment for sepsis was given, then a consensus decision was reached through discussion. Disagreements over an element of the sepsis-3 definition (evidence of infection or change in SOFA score) were left unresolved if they did not affect the overall judgement.

We used the patient as the unit of analysis and only included the first eligible episode per patient. Kappa scores were calculated to determine the agreement between reference standard adjudicators. We constructed receiver operating characteristic (ROC) curves to evaluate sensitivity and specificity over the range of each score. We calculated the area under the ROC curve

and sensitivities, specificities and positive and negative predictive values at key cut-points, each with a 95% CI.

We anticipated a low prevalence of reference standard positive cases, based on data from Smyth *et al*,²⁷ so we based the sample size on identifying at least 200 reference standard positive cases. Collins *et al*³² recommend basing external validation studies on a minimum of 100–200 events.³² Our sample size would allow us to estimate the sensitivity of an early warning score with an SE of 2.1% assuming sensitivity of 90%, and the area under the ROC curve with an SE of 2% assuming an area under the ROC curve of at least 0.75.³³

Clinical experts in the research team reviewed ED attendance data and determined that a positive predictive value of 0.15 or lower would result in too many positive cases for meaningful prioritisation and that sensitivity exceeding 0.8 would be considered good.

Patient and public involvement

The Sheffield Emergency Care Forum (SECF) is a public representative group interested in emergency care research.³⁴ Two members of SECF joined the project management group and helped develop and deliver the project. Public representatives supported the use of patient data without consent and reviewed the early warning scores to determine patient and public acceptability, resulting in one score being modified to remove care home residence as a variable. Patients were not involved in the recruitment to and conduct of the study. We are unable to disseminate the findings to study participants directly.

RESULTS

Figure 1 shows the flow of eligible cases. We identified 24 955 cases transported to Sheffield Teaching Hospitals ED in 2019, of whom 14 050 (56.3%) had NHS numbers and no opt-out. Table 1 shows the characteristics of the 14 050 patients and compares them to those unavailable for linkage. Included patients were markedly older (median age 71 vs 55 years) and more likely to be female (54.7% vs 53.0%) and white ethnicity (95.7% vs 91.8%). We linked 12 870/14 050 cases (91.6%) with a hospital attendance or admission, which comprised the study cohort.

There were 684/12 870 episodes with an admission or ED coding for sepsis. The research nurses referred 655/684 (95.8%) for expert review. The experts judged that 368/655 (56.2%) episodes met the sepsis-3 definition and 348/368 (94.6%) of these received urgent treatment for sepsis. Therefore, 348/12 870 (2.7%) met the reference standard definition. Online supplemental table 2 shows the agreement between the reference standard adjudicators. Agreement was moderate (kappa=0.62) for evidence of infection but disagreements tended to occur in cases that did not meet the SOFA score criterion, so overall judgement on the sepsis-3 definition was good (kappa=0.89), as was agreement for whether urgent treatment was given (kappa=0.87).

There was radiological evidence of infection in 175/348 (50.1%) cases, microbiological evidence in 171 (49.0%) and other clinical evidence in 328 (94.0%). The sites of suspected infection were chest (155, 44.4%), urine (78, 22.3%), biliary (43, 12.3%), abdominal (16, 4.6%), skin (25, 7.2%), other (6, 1.7%) and unknown (26, 7.4%). Mean Clinical Frailty Score was 5.6 (median 6.0, range 2.0–9.0) and mean SOFA score was 3.9 (median 3.0, range 2.0–14.0). Some 28 (8.0%) were admitted to

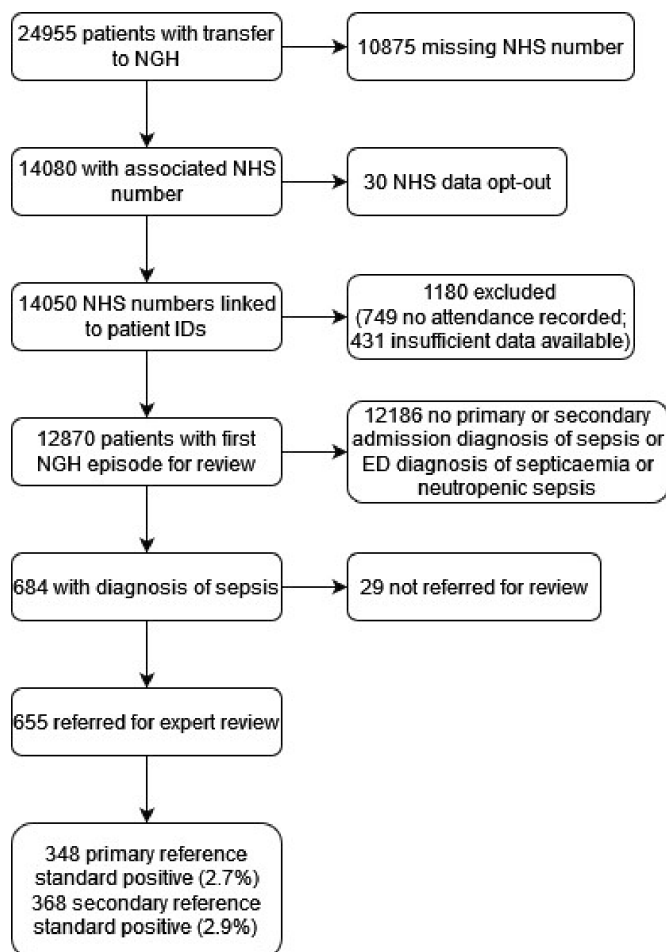


Figure 1 Participant flow through the study. NGH, Northern General Hospital.

Table 1 Characteristics of patients available for linkage with hospital data

	Not linked (n=10 905)	Linked (n=14 050)	Total (n=24 955)
Age (years)			
Mean (SD)	55.2 (23.3)	65.3 (21.2)	60.9 (22.7)
Median (IQR)	55.0 (34.0, 76.0)	71.0 (51.0, 82.0)	65.0 (42.0, 80.0)
Range	16.0–102.0	16.0–105.0	16.0–105.0
Sex			
Missing	0	22	22
Female	5484 (50.3%)	7672 (54.7%)	13 156 (52.8%)
Male	5421 (49.7%)	6356 (45.3%)	11 777 (47.2%)
Ethnicity			
Missing	5290	6880	12 170
White	5153 (91.8%)	6860 (95.7%)	12 013 (94.0%)
Asian	136 (2.4%)	122 (1.7%)	258 (2.0%)
Black	73 (1.3%)	55 (0.8%)	128 (1.0%)
Mixed	49 (0.9%)	32 (0.4%)	81 (0.6%)
Other	204 (3.6%)	101 (1.4%)	305 (2.4%)
ACVPU			
Missing	0	0	0
Alert	9754 (89.4%)	13 232 (94.2%)	22 986 (92.1%)
Confusion	341 (3.1%)	387 (2.8%)	728 (2.9%)
Voice	386 (3.5%)	257 (1.8%)	643 (2.6%)
Pain	192 (1.8%)	107 (0.8%)	299 (1.2%)
Unresponsive	232 (2.1%)	67 (0.5%)	299 (1.2%)
GCS			
Mean (SD)	14.4 (2.0)	14.7 (1.2)	14.5 (1.6)
Median (IQR)	15.0 (15.0, 15.0)	15.0 (15.0, 15.0)	15.0 (15.0, 15.0)
Range	3.0–15.0	3.0–15.0	3.0–15.0
Diastolic BP (mm Hg)			
Mean (SD)	83.1 (17.5)	82.1 (17.2)	82.6 (17.4)
Median (IQR)	83.0 (72.0, 94.0)	82.0 (71.0, 93.0)	82.0 (71.0, 93.0)
Range	0.0–190.0	5.0–195.0	0.0–195.0
Systolic BP (mm Hg)			
Mean (SD)	139.0 (26.5)	142.1 (27.4)	140.8 (27.1)
Median (IQR)	138.0 (122.0, 153.0)	140.0 (124.0, 158.0)	139.0 (123.0, 156.0)
Range	53.0–257.0	43.0–285.0	43.0–285.0
HR (beats/min)			
Mean (SD)	89.5 (22.8)	88.7 (21.9)	89.1 (22.3)
Median (IQR)	87.0 (74.0, 103.0)	86.0 (73.0, 102.0)	86.0 (74.0, 102.0)
Range	0.0–218.0	0.0–216.0	0.0–218.0
Oxygen saturation (%)			
Mean (SD)	96.0 (4.9)	95.6 (4.9)	95.8 (4.9)
Median (IQR)	97.0 (95.0, 98.0)	97.0 (95.0, 98.0)	97.0 (95.0, 98.0)
Range	18.0–100.0	10.0–100.0	10.0–100.0
Supplemental oxygen			
Missing	18	27	45
No	10 345 (95.0%)	13 252 (94.5%)	23 597 (94.7%)
Yes	542 (5.0%)	771 (5.5%)	1313 (5.3%)
Respiration (breath/min)			
Mean (SD)	19.7 (6.0)	20.5 (6.1)	20.1 (6.0)
Median (IQR)	18.0 (16.0, 20.0)	18.0 (16.0, 22.0)	18.0 (16.0, 22.0)
Range	0.0–93.0	0.0–91.0	0.0–93.0
Temperature (°C)			
Mean (SD)	36.8 (1.0)	37.0 (1.0)	36.9 (1.0)
Median (IQR)	36.8 (36.2, 37.3)	36.9 (36.4, 37.4)	36.8 (36.4, 37.4)
Range	26.0–41.3	27.1–41.8	26.0–41.8
Glucose (mmol/L)			
Mean (SD)	7.1 (3.2)	7.4 (3.4)	7.2 (3.3)

Continued

Table 1 Continued

	Not linked (n=10905)	Linked (n=14050)	Total (n=24955)
Median (IQR)	6.2 (5.4, 7.6)	6.4 (5.5, 8.0)	6.3 (5.5, 7.8)
Range	0.5–36.6	0.9–49.0	0.5–49.0
Prealerted			
No	10307 (94.5%)	13419 (95.5%)	23726 (95.1%)
Yes	598 (5.5%)	631 (4.5%)	1229 (4.9%)
Impression			
1—Sepsis	222 (2.0%)	407 (2.9%)	629 (2.5%)
2—Infection	471 (4.3%)	912 (6.5%)	1383 (5.5%)
3—Non-specific	3494 (32.0%)	4962 (35.3%)	8456 (33.9%)
4—Other	6718 (61.6%)	7769 (55.3%)	14487 (58.1%)

critical care and 261 (74.8%) survived to hospital discharge or 30 days after attendance, whichever was sooner.

Paramedic diagnostic impression of sepsis had sensitivity (95% CI) of 0.328 (0.28, 0.379) and positive predictive value of 0.285 (0.243, 0.331); infection or sepsis had sensitivity of 0.572 (0.519, 0.623) and positive predictive value of 0.156 (0.137, 0.176); and non-specific, infection or sepsis had sensitivity of 0.897 (0.86, 0.924) and positive predictive value of 0.053 (0.048, 0.059). Online supplemental table 3 shows the full details.

Figures 2–5 show the ROC curves for each score alongside diagnostic impression of sepsis, infection, non-specific and all cases. Online supplemental table 4 reports the areas under each ROC curve and online supplemental tables 5–12 show the accuracy parameters behind the ROC curves. The area under the ROC curve is greater when the scores are used less selectively with paramedic diagnostic impression. However, the accuracy parameters in the online supplemental tables show that the positive predictive value is low (<0.15) if specificity is below 0.9. The area under the ROC curve is therefore a poor reflection of accuracy at the thresholds that yield acceptable positive

predictive value (ie, specificity >0.9). Figures 2–5 show that none of the alternative scores had superior accuracy to NEWS2. The possible exception is the Screening to Enhance Prehospital Identification of Sepsis (SEPSIS) score that has a higher area under the ROC curve when applied to non-specific or all cases, but has similar accuracy to NEWS2 at thresholds that provide specificity greater than 0.9.

Table 2 shows the accuracy parameters (reproduced from online supplemental tables 5–12) for early warning scores at specified thresholds, selected on the basis of their use in sepsis guidelines^{3 4 17 30} in patients with an impression of infection or sepsis. NEWS2>4, NEWS2>6 and quick SOFA (qSOFA)>1 provide a range of options with varying sensitivity and positive predictive value. qSOFA>1 provides similar accuracy to NEWS2>8 (also included in the table). The modified NHS prealert criteria¹⁷ provide slightly inferior accuracy to NEWS2>6. The modified UK Sepsis Trust criteria³ provide similar accuracy to NEWS2>4.

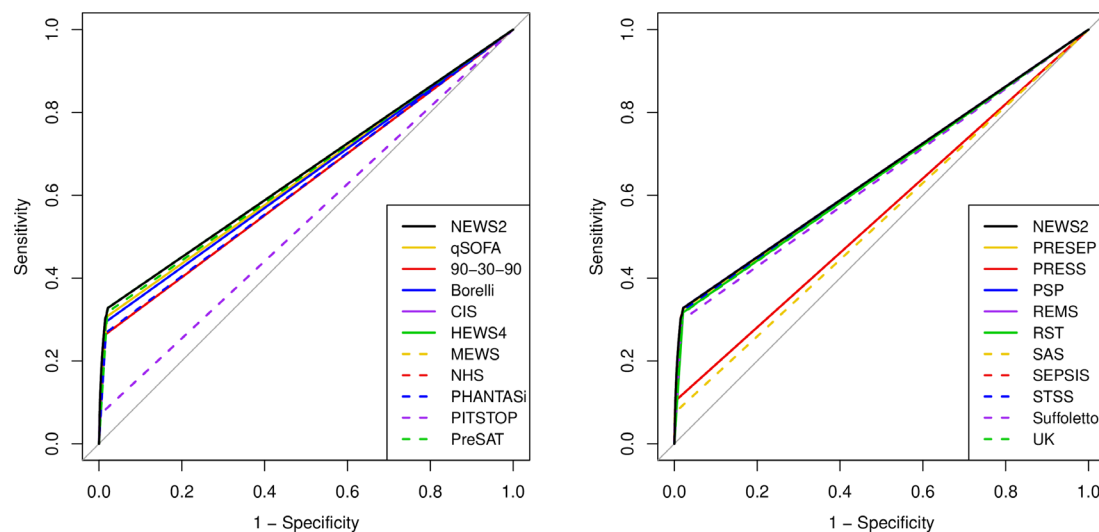


Figure 2 Receiver operating characteristic (ROC) curves for early warning scores applied to diagnostic impression of sepsis. CIS, Critical Illness Score; HEWS, Hamilton Early Warning Score; MEWS, Modified Early Warning Score; NEWS2, National Early Warning Score, version 2; PHANTASI, Prehospital Antibiotics Against Sepsis; PITSTOP, Paramedic Initiated Treatment of Sepsis Targeting Out-of-hospital Patients; PreSAT, Prehospital Sepsis Assessment Tool; PRESEP, Prehospital Early Sepsis Detection; PRESS, Prehospital Severe Sepsis; PSP, Prehospital Sepsis Project; qSOFA, quick Sequential Organ Failure Assessment; REMS, Rapid Emergency Medicine Score; RST, Robson Screening Tool; SEPSIS, Screening to Enhance Prehospital Identification of Sepsis; STSS, Simple Triage Scoring System.

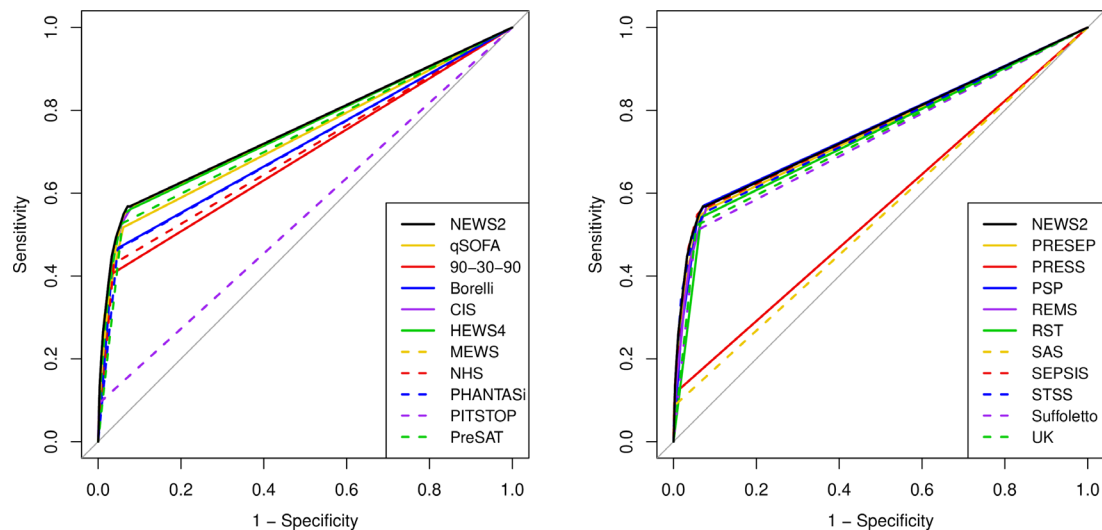


Figure 3 Receiver operating characteristic (ROC) curves for early warning scores applied to diagnostic impression of sepsis or infection. CIS, Critical Illness Score; HEWS, Hamilton Early Warning Score; MEWS, Modified Early Warning Score; NEWS2, National Early Warning Score, version 2; PHANTASi, Prehospital Antibiotics Against Sepsis; PITSTOP, Paramedic Initiated Treatment of Sepsis Targeting Out-of-hospital Patients; PreSAT, Prehospital Sepsis Assessment Tool; PRESEP, Prehospital Early Sepsis Detection; PRESS, Prehospital Severe Sepsis; PSP, Prehospital Sepsis Project; qSOFA, quick Sequential Organ Failure Assessment; REMS, Rapid Emergency Medicine Score; RST, Robson Screening Tool; SEPSIS, Screening to Enhance Prehospital Identification of Sepsis; STSS, Simple Triage Scoring System.

DISCUSSION

We found that no combination of early warning score alongside paramedic diagnostic impression provided sensitivity greater than 0.8 and positive predictive value greater than 0.15 for sepsis. The appropriate trade-off between sensitivity and positive predictive value will depend on the consequences of prioritisation. However, prioritising more than five people for each case of sepsis (which would be the consequence of using a strategy with

positive predictive value of 0.15 or lower) risks overstressing ED capacity and a loss of meaningful prioritisation.

No score had superior accuracy to NEWS2. The only possible exception was the SEPSIS score²⁷ when thresholds were used that optimised sensitivity at the expense of low positive predictive value. NEWS2 is widely used in the UK NHS, so any alternative score would need to demonstrate clear superiority to justify the additional training and documentation required in this setting.

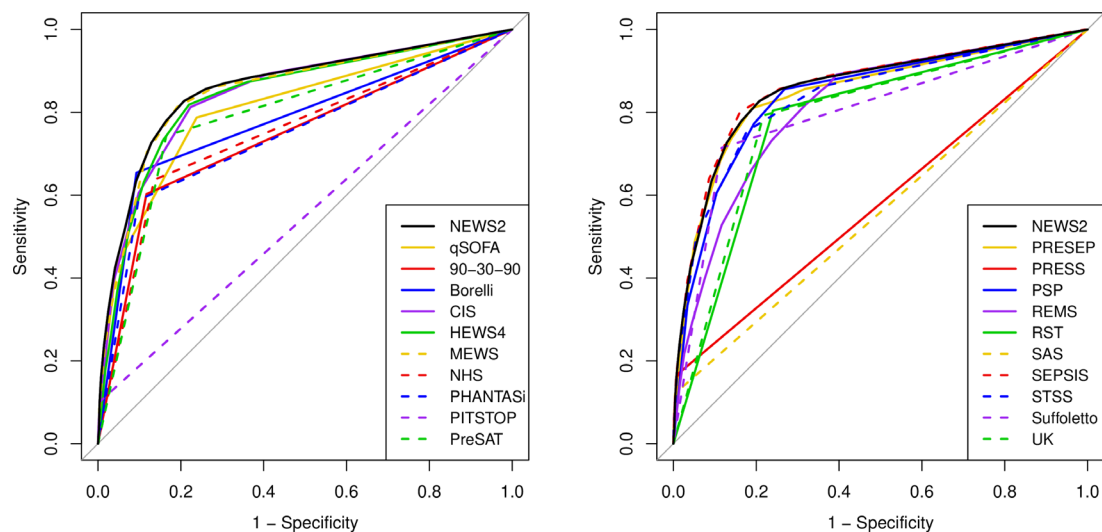


Figure 4 Receiver operating characteristic (ROC) curves for early warning scores applied to diagnostic impression of sepsis, infection or non-specific presentation. CIS, Critical Illness Score; HEWS, Hamilton Early Warning Score; MEWS, Modified Early Warning Score; NEWS2, National Early Warning Score, version 2; PHANTASi, Prehospital Antibiotics Against Sepsis; PITSTOP, Paramedic Initiated Treatment of Sepsis Targeting Out-of-hospital Patients; PreSAT, Prehospital Sepsis Assessment Tool; PRESEP, Prehospital Early Sepsis Detection; PRESS, Prehospital Severe Sepsis; PSP, Prehospital Sepsis Project; qSOFA, quick Sequential Organ Failure Assessment; REMS, Rapid Emergency Medicine Score; RST, Robson Screening Tool; SEPSIS, Screening to Enhance Prehospital Identification of Sepsis; STSS, Simple Triage Scoring System.

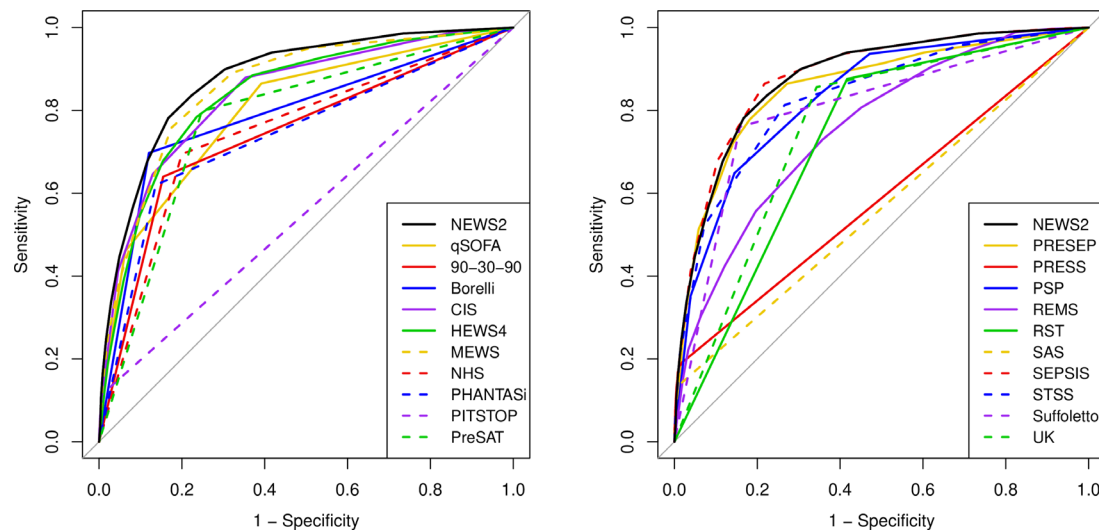


Figure 5 Receiver operating characteristic (ROC) curves for early warning score applied to all diagnostic impressions. CIS, Critical Illness Score; HEWS, Hamilton Early Warning Score; MEWS, Modified Early Warning Score; NEWS2, National Early Warning Score, version 2; PHANTASi, Prehospital Antibiotics Against Sepsis; PITSTOP, Paramedic Initiated Treatment of Sepsis Targeting Out-of-hospital Patients; PreSAT, Prehospital Sepsis Assessment Tool; PRESEP, Prehospital Early Sepsis Detection; PRESS, Prehospital Severe Sepsis; PSP, Prehospital Sepsis Project; qSOFA, quick Sequential Organ Failure Assessment; REMS, Rapid Emergency Medicine Score; RST, Robson Screening Tool; SEPSIS, Screening to Enhance Prehospital Identification of Sepsis; STSS, Simple Triage Scoring System.

Using NEWS2 at thresholds of >4 and >6 to prioritise patients with suspected infection, as recommended by the Academy of Medical Royal Colleges clinical decision support framework,⁴ would provide sensitivities of 0.522 and 0.447, respectively, and positive predictive values of 0.216 and 0.274. To prioritise fewer patients, NEWS2 could be used with a threshold of >8 , which would provide similar sensitivity and positive predictive value (0.314 and 0.333) to using qSOFA with a threshold of >1 (0.305 and 0.356). Using NEWS2 alongside paramedic diagnostic impression improves positive predictive value at the expense of sensitivity, compared with paramedic diagnostic impression alone.

We recently searched for studies validating the accuracy of early warning scores for suspected sepsis in a prehospital population and identified 13 studies evaluating the scores included in this study.⁵ There was substantial variation in the reported results, with no consistent evidence that any score was superior to the others. Variations in study populations, outcomes and the thresholds used make comparisons difficult. A systematic review of hospital studies found that at established thresholds NEWS tended to have higher sensitivity while qSOFA tended to have higher specificity.³⁵ Our study suggests that this difference

reflects the chosen threshold. The sensitivity and specificity of NEWS2 at a higher threshold than usually recommended (>8) are similar to the sensitivity and specificity of qSOFA >1 .

Our findings are similar to other studies evaluating multiple scores in a large cohort. Lane *et al* found that no single strategy had high sensitivity and specificity for classifying sepsis, but the Critical Illness Prediction score, NEWS and qSOFA showed good prediction for sepsis.³⁶ Smyth *et al* identified three strategies offering an acceptable balance between sensitivity and positive predictive value: SEPSIS >2 , Systemic Inflammatory Response Syndrome (SIRS) criteria >1 and NEWS >4 .²⁷ These studies did not identify any early warning score with clearly superior accuracy to NEWS2.

Key strengths of our study include the large sample size including sufficient cases with sepsis to estimate sensitivity with reasonable precision. The reference standard was based on an internationally recognised definition of sepsis that was adjudicated by two independent clinicians with acceptable interobserver agreement. The main limitation is that we were only able to link around half the eligible cases with hospital records. Those linked tended to be much older, possibly reflecting more frequent contact with health services. Sepsis is associated with

Table 2 Accuracy of selected early warning scores alongside paramedic impression of sepsis or infection for identifying sepsis receiving urgent treatment

Early warning score	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Paramedic impression alone	0.572 (0.519, 0.623)	0.914 (0.909, 0.919)	0.156 (0.137, 0.176)	0.987 (0.985, 0.989)
NEWS2 >4	0.522 (0.469, 0.574)	0.947 (0.943, 0.951)	0.216 (0.189, 0.245)	0.986 (0.984, 0.988)
NEWS2 >6	0.447 (0.395, 0.499)	0.967 (0.964, 0.97)	0.274 (0.239, 0.313)	0.984 (0.982, 0.986)
NEWS2 >8	0.314 (0.268, 0.365)	0.983 (0.98, 0.985)	0.333 (0.284, 0.386)	0.981 (0.978, 0.983)
qSOFA >1	0.305 (0.259, 0.355)	0.985 (0.982, 0.987)	0.356 (0.304, 0.412)	0.981 (0.978, 0.983)
NHS prealert	0.429 (0.378, 0.482)	0.962 (0.959, 0.966)	0.24 (0.208, 0.275)	0.984 (0.981, 0.986)
UK Sepsis Trust	0.522 (0.469, 0.574)	0.945 (0.941, 0.949)	0.209 (0.183, 0.237)	0.986 (0.984, 0.988)

NEWS2, National Early Warning Score, version 2; NPV, negative predictive value; PPV, positive predictive value; qSOFA, quick Sequential Organ Failure Assessment.

REFERENCES

- National Institute for Health and Care Excellence. NICE Guideline 51 [NG51]. Sepsis: recognition, diagnosis and early management. Available: <https://www.nice.org.uk/guidance/ng51> [Accessed 14 Apr 2023].
- Evans L, Rhodes A, Alhazzani W, et al. Surviving sepsis campaign: International guidelines for management of sepsis and septic shock 2021. *Intensive Care Med* 2021;47:1181–247.
- UK Sepsis Trust. Sepsis screening tool prehospital. Available: <https://sepsistrust.org/wp-content/uploads/2020/08/Sepsis-Prehospital-12-231219.pdf> [Accessed 14 Apr 2023].
- Bion JBG, Boyle A, Carrol E, et al. *Academy of Medical Royal Colleges Statement on the Initial Antimicrobial Treatment of Sepsis*. Academy of Medical Royal Colleges, 2022.
- Goodacre S, Thomas B, Smyth M, et al. Should prehospital early warning scores be used to identify which patients need urgent treatment for sepsis? *BMJ* 2021;375:2432.
- Hamilton F, Arnold D, Baird A, et al. Early warning scores do not accurately predict mortality in sepsis: a meta-analysis and systematic review of the literature. *J Infect* 2018;76:241–8.
- Lane D, Ichelson RI, Drennan IR, et al. Prehospital management and identification of sepsis by emergency medical services: a systematic review. *Emerg Med J* 2016;33:408–13.
- Smyth MA, Brace-McDonnell SJ, Perkins GD. Identification of adults with sepsis in the prehospital environment: a systematic review. *BMJ Open* 2016;6:e011218.
- Accuracy, impact and cost-effectiveness of prehospital clinical early warning scores for adults with suspected sepsis. Available: <https://www.fundingawards.nihr.ac.uk/award/17/136/10> [Accessed 14 Apr 2023].
- Royal College of Physicians, London. National early warning score (NEWS) 2. 2017. Available: <http://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news-2> [Accessed 14 Apr 2023].
- Seymour CW, Liu VX, Iwashyna TJ, et al. Assessment of clinical criteria for sepsis: for the third International consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 2016;315:762–74.
- Wallgren UM, Castrén M, Svensson AEV, et al. Identification of adult septic patients in the prehospital setting: a comparison of two screening tools and clinical judgment. *Eur J Emerg Med* 2014;21:260–5.
- Borrelli G, Koch E, Sterk E, et al. Early recognition of sepsis through emergency medical services pre-hospital screening. *Am J Emerg Med* 2019;37:1428–32.
- Seymour CW. Prediction of critical illness during out-of-hospital emergency care. *JAMA* 2010;304:747.
- Tam BHSc MD B, Xu BHSc M, Kwong BHSc Cand. M, et al. The admission Hamilton early warning score (HEWS) predicts the risk of critical event during hospitalization. *Can Journ Gen Int Med* 2017;11.
- Jouffroy R, Saade A, Ellouze S, et al. Prehospital triage of septic patients at the SAMU regulation: comparison of qSOFA, MRST, MEWS and PRESEP scores. *Am J Emerg Med* 2018;36:820–4.
- Association of Ambulance Chief Executives and the Royal College of Emergency Medicine. UK NHS ambulance services pre-alert guideline for the deteriorating adult patient. Available: https://rcem.ac.uk/wp-content/uploads/2021/10/NHS_Ambulance_Services_pre_alert_criteria_guideline.pdf [Accessed 14 Apr 2023].
- Alam N, Oskam E, Stassen PM, et al. Prehospital antibiotics in the ambulance for sepsis: a multicentre, open label, randomised trial. *Lancet Respir Med* 2018;6:40–50.
- Scales DD. Paramedic initiated treatment of sepsis targeting out-of-hospital patients (PITSTOP). 2023. Available: clinicaltrials.gov
- Walchok JG, Pirrallo RG, Furmanek D, et al. Paramedic-initiated CMS sepsis core measure bundle prior to hospital arrival: a stepwise approach. *Prehosp Emerg Care* 2017;21:291–300.
- Bayer O, Schwarzkopf D, Stumme C, et al. An early warning scoring system to identify septic patients in the prehospital setting: the PRESEP score. *Acad Emerg Med* 2015;22:868–71.
- Polito CC, Isakov A, Yancey AH, et al. Prehospital recognition of severe sepsis: development and validation of a novel EMS screening tool. *Am J Emerg Med* 2015;33:1119–25.
- Baez AA, Cochon L. Acute care diagnostics collaboration: assessment of a Bayesian clinical decision model integrating the prehospital sepsis score and point-of-care lactate. *Am J Emerg Med* 2016;34:193–6.
- Olsson T, Terent A, Lind L. Rapid emergency medicine score: a new prognostic tool for in-hospital mortality in nonsurgical emergency department patients. *J Intern Med* 2004;255:579–87.
- Robson W, Nutbeam T, Daniels R. Sepsis: a need for prehospital intervention *Emerg Med J* 2009;26:535–8.
- Guerra WF, Mayfield TR, Meyers MS, et al. Early detection and treatment of patients with severe sepsis by prehospital personnel. *J Emerg Med* 2013;44:1116–25.
- Smyth MA, Gallacher D, Kimani PK, et al. Derivation and internal validation of the screening to enhance prehospital identification of sepsis (SEPSIS) score in adults on arrival at the emergency department. *Scand J Trauma Resusc Emerg Med* 2019;27:67.
- Talmor D, Jones AE, Rubinson L, et al. Simple triage scoring system predicting death and the need for critical care resources for use during epidemics. *Crit Care Med* 2007;35:1251–6.
- Suffoletto B, Frisch A, Prabhu A, et al. Prediction of serious infection during prehospital emergency care. *Prehosp Emerg Care* 2011;15:325–30.
- Singer M, Deutschman CS, Seymour CW, et al. The third International consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 2016;315:801.
- Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. *CMAJ* 2005;173:489–95.
- Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med* 2016;35:214–26.
- Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36.
- Hirst E, Irving A, Goodacre S. Patient and public involvement in emergency care research. *Emerg Med J* 2016;33:665–70.
- Sabir L, Ramlakhan S, Goodacre S. Comparison of qSOFA and hospital early warning scores for prognosis in suspected sepsis in emergency department patients: a systematic review. *Emerg Med J* 2022;39:284–94.
- Lane DJ, Wunsch H, Saskin R, et al. Screening strategies to identify sepsis in the Prehospital setting: a validation study. *CMAJ* 2020:E230–9.
- Goodacre S. Using clinical risk models to predict outcomes: what are we predicting and why? *Emerg Med J* 2023;emermed-2022-213057.
- Goodacre S, Fuller G, Conroy S, et al. Diagnosis and management of sepsis in the older adult. *BMJ* 2023;382:e075585.