



This is a repository copy of *Accuracy of the National Early Warning Score version 2 (NEWS2) in predicting need for time-critical treatment: retrospective observational cohort study*.

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/id/eprint/221984/>

Version: Published Version

---

**Article:**

Goodacre, S. [orcid.org/0000-0003-0803-8444](https://orcid.org/0000-0003-0803-8444), Sutton, L., Fuller, G. [orcid.org/0000-0001-8532-3500](https://orcid.org/0000-0001-8532-3500) et al. (2 more authors) (2025) Accuracy of the National Early Warning Score version 2 (NEWS2) in predicting need for time-critical treatment: retrospective observational cohort study. *Emergency Medicine Journal*, 42 (6). pp. 396-402. ISSN 1472-0205

<https://doi.org/10.1136/emered-2024-214562>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:  
<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.






[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>



OPEN ACCESS

# Accuracy of the National Early Warning Score version 2 (NEWS2) in predicting need for time-critical treatment: retrospective observational cohort study

Steve Goodacre <sup>1,2</sup>, Laura Sutton,<sup>1</sup> Gordon Fuller <sup>1,2</sup>, Ashleigh Trimble,<sup>2</sup> Richard Pilbery <sup>3</sup>

**Handling editor** Edward Carlton

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/emmermed-2024-214562>).

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

<sup>2</sup>Emergency Department, Northern General Hospital, Sheffield, UK

<sup>3</sup>Yorkshire Ambulance Service NHS Trust, Wakefield, UK

## Correspondence to

Professor Steve Goodacre;  
[s.goodacre@sheffield.ac.uk](mailto:s.goodacre@sheffield.ac.uk)

Received 20 September 2024  
Accepted 17 December 2024

## ABSTRACT

**Background** Initial ED assessment can use early warning scores to identify and prioritise patients who need time-critical treatment. We aimed to determine the accuracy of the National Early Warning Score version 2 (NEWS2) for predicting the need for time-critical treatment.

**Methods** We undertook a single-centre retrospective observational cohort study. We randomly selected 4000 adults who attended a tertiary hospital ED in England from 1 January 2022 to 31 December 2022 and had NEWS2 routinely recorded on electronic patient records. The first NEWS2 and vital signs were extracted from electronic records. Research nurses selected cases that received a potentially time-critical treatment. Two independent clinical experts then determined whether time-critical treatment was or should have been received using an expert consensus derived list of interventions. We used receiver operating characteristic analysis and calculated sensitivity and specificity at predefined thresholds to evaluate the accuracy of NEWS2 for predicting need for time-critical intervention and, as a secondary outcome, mortality at 7 days.

**Results** After excluding 10 patients who received their intervention before NEWS2 recording, 164/3990 (4.1%) needed time-critical treatment and 71/3990 (1.8%) died within 7 days. NEWS2 predicted need for time-critical treatment with a c-statistic of 0.807 (95% CI 0.765 to 0.849) and death within 7 days with a c-statistic of 0.865 (95% CI 0.813, 0.917). NEWS2 $\geq$ 4 predicted need for time-critical treatment with sensitivity of 51.8% (95% CI 44.2%, 59.3%) and positive predictive value of 25.8% (95% CI 21.3%, 30.7%). 37 of the 45 patients needing emergency surgery, antibiotics for open fractures, insulin infusion or manipulation of limb-threatening injuries had NEWS2 $\leq$ 4. Patients with NEWS2 $\geq$ 4 who did not need time-critical treatment frequently scored maximum points on NEWS2 for their respiratory rate, conscious level or receiving supplemental oxygen.

**Conclusion** NEWS2 has limited accuracy for predicting need for time-critical treatment. We have identified time-critical interventions that frequently have low NEWS2 scores and NEWS2 parameters that may overestimate need for time-critical intervention.

**Trial registration number** Research Registry 10450

## BACKGROUND

Patients attending the ED can face frequent and sometimes prolonged waits before they receive

## WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The National Early Warning Score version 2 (NEWS2) is used to identify patients needing time-critical treatment at initial ED assessment, but NEWS2 validation has focused on predicting the risk of death.

## WHAT THIS STUDY ADDS

⇒ This retrospective observational cohort study of 4000 patients showed that NEWS2 predicted the need for time-critical treatment with limited accuracy (c-statistic 0.807). Half of patients needing time-critical treatment and most patients needing emergency surgery, antibiotics for open fractures, insulin infusion or manipulation of limb-threatening injuries had a low NEWS2 score (0–4).

## HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ ED clinicians can use our data to recognise the limitations of NEWS2 as an initial assessment tool and identify cases where NEWS2 may underestimate or overestimate the need for time-critical treatment.

definitive assessment and treatment. This can result in avoidable harm if urgent treatment is delayed. Initial assessment of patients arriving at the ED aims to reduce this risk by ensuring that patients with time-critical conditions are identified and prioritised.<sup>1 2</sup>

Early warning scores use physiological measurements to produce a composite score reflecting illness severity that can assist initial assessment. In the UK, guidance from the Royal College of Emergency Medicine supports the use of early warning scores as part of initial assessment but advises against their use as a sole measure of acuity.<sup>1</sup> The Royal College of Physicians (RCP) developed the National Early Warning Score version 2 (NEWS2) to standardise the assessment and response to acute illness in adults and recommends that ED staff use NEWS2 to aid the initial assessment of adult patients.<sup>3</sup> NHS England has endorsed the use of NEWS2 and provided guidance to support adoption in acute and ambulance settings.<sup>4</sup>

Systematic reviews and meta-analysis have shown that early warning scores have good prediction for



© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY. Published by BMJ Group.

**To cite:** Goodacre S, Sutton L, Fuller G, *et al.* *Emerg Med J* Epub ahead of print: [please include Day Month Year]. doi:10.1136/emmermed-2024-214562



mortality but adequate to poor prediction of intensive care or hospital admission.<sup>5–8</sup> Mortality reflects illness severity, including frailty and long-term conditions, but early warning scores need to predict illness acuity—the need for time-critical treatment. Mortality identifies deaths that occurred despite treatment, and may therefore have been inevitable, but does not identify cases where treatment prevented death.<sup>9</sup> Mortality may therefore fail to identify cases most likely to benefit from time-critical treatment. A recent review of ED acuity assessment tools<sup>10</sup> identified only one small study that directly measured accuracy for time-critical treatment<sup>11</sup> and a systematic review that included studies measuring accuracy for time-critical diagnoses.<sup>12</sup>

We aimed to determine the accuracy with which NEWS2 predicts the need for time-critical treatment among adults attending the ED and characterise presentations where NEWS2 has poor accuracy.

## METHODS

### Design and setting

We undertook a single-centre retrospective observational cohort study at the Northern General Hospital ED in Sheffield, UK. This is the only adult ED serving the 530 000 population of Sheffield and the adult major trauma centre for the 1.8 million population of South Yorkshire. The ED receives all undifferentiated adult emergencies except ambulance arrivals with stroke or ST-elevation myocardial infarction requiring reperfusion, which are taken directly to specialist services. At initial assessment, nurses record vital signs on the ED information system for all patients considered to be at risk of physiological deterioration, which then generates a NEWS2 score for patients with a complete set of vital signs. Oxygen saturation is recorded using scale 1 unless the patient is known to have hypercapnic respiratory failure. Scale 1 allocates NEWS2 points only according to the severity of hypoxia, whereas scale 2 also allocates points to patients with confirmed hypercapnic respiratory failure if they have oxygen saturation exceeding 92% and are receiving supplemental oxygen. The patient tracking system and electronic record flag any patient with a total NEWS2 > 4 or a single parameter of 3 as requiring an urgent response and any patient with a total NEWS2 > 6 as requiring an emergency response, in accordance with RCP guidance.<sup>3</sup> Patients with minor injuries or primary care complaints are registered with the ED but may then be referred to services located alongside the ED without NEWS2 being recorded. The attending clinician completes a coding form when ED assessment is complete that records standardised ED diagnoses and treatments.

### Case selection

We used routine hospital data to identify all adult (aged 16 or over) ED attendances from 1 January 2022 to 31 December 2022 that had a NEWS2 score recorded and randomly sampled 4000 attendances to account for seasonality, having excluded repeat attendances and patients who had opted out of allowing their data to be used for research. Each eligible attendance was given a random number on an Excel spreadsheet, attendances were sorted in order of the random number and then consecutively selected. We extracted the following routine hospital data: age, sex, ethnicity, the first recorded NEWS2 score, heart rate, respiratory rate, temperature, peripheral oxygen saturation, blood pressure and conscious level, ED diagnoses and treatments, hospital admission, and attendances, admissions and deaths over the following week.

A research nurse reviewed the ED records, the initial inpatient records for admitted patients and hospital discharge summaries of the selected attendances to identify patients who received any treatment that could potentially be considered time critical or suffered an adverse outcome (death or safety incident) that could have been prevented by time-critical treatment. Two independent experts then reviewed hospital records of the patients selected by the nurses and used their clinical expertise to determine whether the treatment fulfilled the definition of being time critical outlined below, with a third expert resolving any disagreements. The experts (SG, AT, GF) had all completed specialist training in emergency medicine. The NEWS2 score was recorded on an observations chart that was not part of the hospital records reviewed by the experts, so outcome adjudication was undertaken by observers who were not aware of the NEWS2 score but could have estimated or calculated the score from the clinical observations.

We also checked all the selected attendances against the ED database of patient safety incidents recorded using the DATIX system to identify any adverse events that could have been prevented by time-critical treatment. The incident reports were used to select any potentially relevant incidents, which were then independently adjudicated by two of the clinical experts.

### Outcomes

The primary outcome, need for time-critical treatment, was defined through the expert consensus process described in the online supplemental appendix. Nine experts in emergency medicine used existing literature and their clinical experience to define 34 interventions as likely to be time critical and provided advice on how the list of interventions should be used in outcome adjudication. The clinical experts then used this list of interventions to determine whether each patient had appropriately received or should have received a time-critical treatment. We included time-critical interventions provided after leaving the ED if the intervention required the patient to be urgently moved to a specialist facility (eg, operating theatre or intensive care) or if the intervention should have been urgently provided in the ED but was delayed or missed.

We excluded cases from the analysis if they received a time-critical treatment before NEWS2 was recorded or received continuing time-critical prehospital treatment in the ED (eg, airway or breathing support that was initiated prehospital and continued in the ED). We included cases that had received time-critical prehospital intervention if the intervention had been completed and NEWS2 was recorded after completion of the prehospital intervention (eg, treated hypoglycaemia).

### Analysis

A medical statistician from the University of Sheffield (LS) undertook all analysis. NEWS2 scores generated by the hospital information system were verified against their constituent elements. The information system only generates a NEWS2 score if all constituent elements are entered and sets limits for the values entered, so there were no missing data in the study population. However, the system allocates a NEWS2 score of 3 for any variable with a zero value. We therefore checked any zero values against the hospital records to ensure that the score of 3 was appropriately allocated (eg, unrecordable temperature due to hypothermia or unrecordable blood pressure due to shock) and imputed the next available measurement if the zero value appeared to be due to equipment failure.

We calculated a kappa statistic with a 95% CI to determine the agreement between the two experts adjudicating the primary outcome. We undertook receiver operating characteristic (ROC) analysis to determine the discriminant value of NEWS2 for predicting the need for time-critical treatment across varying thresholds.<sup>13</sup> We repeated this analysis using a secondary outcome of death within 7 days and a composite secondary outcome of death within 7 days or need for time-critical treatment, and tested the hypothesis that NEWS2 prediction differs between time-critical treatment and death. We calculated the sensitivity, specificity, positive predictive and negative predictive values (with 95% CIs calculated using the Wilson score method<sup>14</sup>) for thresholds of NEWS2>4 and NEWS2>6. We repeated this analysis with the score classified as being above the threshold if any NEWS2 parameter equals 3, in accordance with RCP guidance.<sup>3</sup> Finally, we described the characteristics of the 'false negative' cases that had NEWS2≤4 and needed time-critical treatment, and the 'false positive' cases that had NEWS2>4 but did not need time-critical treatment.

We estimated that a sample size of 4000 would include 200 cases with the primary outcome, based on the subjective judgement of experienced clinicians working in the ED (SG and GF). This sample would give an estimated 95% CI of 0.73 to 0.81 for an assumed c-statistic of 0.77,<sup>13</sup> based on a previous similar study.<sup>15</sup>

### Patient and public involvement

Sheffield Emergency Care Forum is a patient and public representative group with an interest in emergency care that has extensive experience of involvement in emergency care research.<sup>16</sup> Two members of the forum joined the research team and were specifically responsible for reviewing the list of time-critical interventions to ensure that it reflected public values and would not discriminate against any patient group.

### RESULTS

Online supplemental figure 1 shows the flow of cases through the study. There were 85 499 first and 41 220 repeat ED attendances in 2022, with 56 145/85 499 (65.7%) first attendances having NEWS2 recorded. The 27 905 patients without NEWS2 at first attendance were relatively young (mean age 44.5 years), with a low admission rate (1911/27 905, 6.8%) and relatively high proportions of minor injuries (14 346/27 905, 46%) and referral to primary care (5423/27 905, 17.4%). We excluded 2689 eligible cases that had opted out of allowing data use and 46 aged <16 years, and then randomly selected 4000 from 53 410 eligible attendances for inclusion.

The research nurses selected 704/4000 cases with possible time-critical interventions for expert review, with 173/704 adjudicated as needing time-critical treatment ( $\kappa=0.89$ , 95% CI 0.86, 0.93). One additional case was identified through safety incident review, and 10 cases were excluded because the time-critical intervention was received before NEWS2 was recorded. Therefore, 164/3990 (4.1%) cases were positive for the primary outcome. There were 71 participants (1.8%) experiencing the secondary outcome of death within 7 days and 195 (4.9%) with the composite secondary outcome. Table 1 shows the characteristics of the included patients and table 2 shows the time-critical interventions received. The most frequent intervention was intravenous antibiotics for infection causing new organ dysfunction or shock (66/164).

Figure 1 shows the proportion of cases needing time-critical treatment at each NEWS2 score. The proportion increases with

**Table 1** Participant characteristics (whole cohort)

Variable	All (n=3990)
Age (years)	
Mean (SD)	52.0 (22.8)
Median (min, max)	52.0 (16.0, 107)
Sex	
Female	2173 (54.5%)
Male	1817 (45.5%)
Ethnicity	
Asian	285 (7.6%)
Black	164 (4.4%)
Mixed	54 (1.4%)
Other	236 (6.3%)
White	3007 (80.3%)
Not stated	244
Heart rate (beats per minute)	
Mean (SD)	86.6 (19.0)
Median (min, max)	85.0 (18.0, 199)
Respiration rate (breaths per minute)	
Mean (SD)	18.2 (3.25)
Median (min, max)	18.0 (6.00, 66.0)
Oxygen saturation (%)	
Mean (SD)	97.7 (2.42)
Median (min, max)	98.0 (54.0, 100)
Level of consciousness	
Alert	3857 (96.7%)
Confused	86 (2.2%)
Voice	20 (0.5%)
Pain	17 (0.4%)
Unresponsive	10 (0.3%)
Systolic BP (mm Hg)	
Mean (SD)	139 (25.0)
Median (min, max)	137 (52.0, 271)
Diastolic BP (mm Hg)	
Mean (SD)	83.8 (16.7)
Median (min, max)	83.0 (0, 198)
Temperature (°C)	
Mean (SD)	36.6 (0.762)
Median (min, max)	36.6 (29.1, 40.6)
Supplemental oxygen	
Air	3764 (94.3%)
On O <sub>2</sub>	226 (5.7%)

NEWS2 score and exceeds a quarter of cases at NEWS2=7 and half of cases at NEWS2=11, although based on small numbers at higher NEWS2 score. Figure 2 shows the ROC curves for the primary and secondary outcomes. The c-statistic for NEWS2 prediction, the need for time-critical intervention, was 0.807 (95% CI 0.765 to 0.849), which was lower than the c-statistic for death within 7 days (0.865, 95% CI 0.813, 0.917), although the difference was not statistically significant ( $p=0.09$ ).

Table 3 shows the accuracy of NEWS2 for the primary outcome using recommended thresholds for patient prioritisation. Around half of patients needing time-critical intervention have NEWS2>4 and around a quarter with NEWS2>4 need time-critical intervention. Sensitivity can be improved to around 60%, at the expense of specificity, by including those with any NEWS2 parameter equalling 3 as equivalent to NEWS2>4 (index test positive).

**Table 2** Treatment received or should have been received in reference standard positive cases

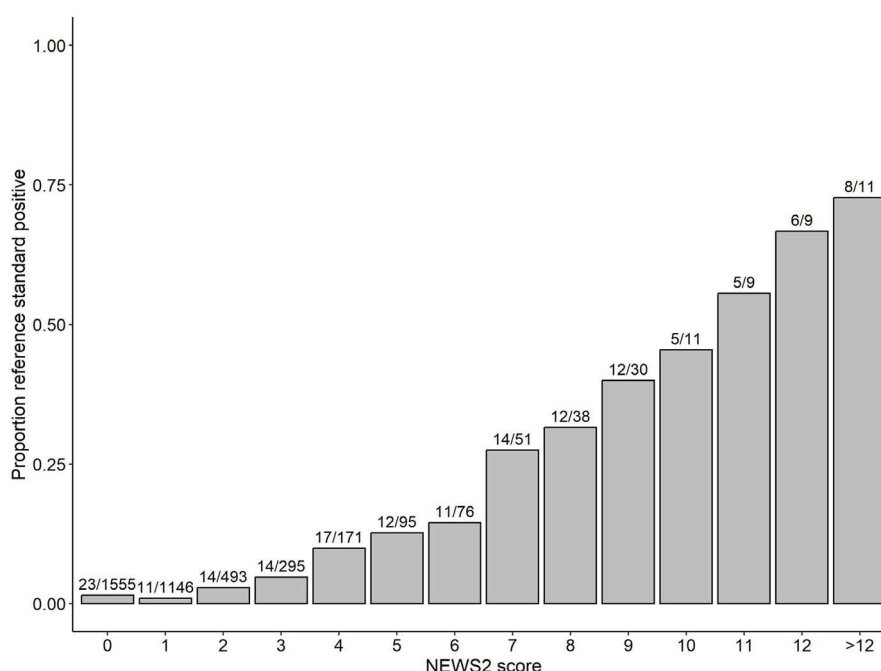
Intervention	Frequency
Intravenous antibiotics for infection causing new organ dysfunction or shock	66
Emergency surgery to avoid death or significant disability, including surgical source control for infection	21
Large-volume intravenous fluid replacement (>2 L within 2 hours or >1 L with central venous monitoring)	19
Any intervention to support ventilation (except supplemental oxygen), other than during sedation for a procedure	18
Administration of blood products for acute blood loss, to allow emergency procedures, in haematological emergencies, or severe anaemia in context of proven acute coronary syndrome	15
Unplanned airway intervention to provide a patent airway, other than during sedation for a procedure	11
Intravenous antibiotics for genuine open fractures	11
Hyperkalaemia or hypokalaemia involving intravenous treatment and cardiac monitoring	10
Insulin infusion as part of treatment protocol for diabetic ketoacidosis or hyperosmolar hyperglycaemic state	8
Antidote for life or disability threatening poisoning	7
Any intervention to support circulation (including CPR), other than intravenous fluids	7
Intravenous fluids and steroids for Addisonian crisis	6
Reduction of limb-threatening fracture or dislocation (including threat to skin, nerve or perfusion)	5
Cardioversion or rate control for life-threatening arrhythmia	3
Primary percutaneous coronary intervention or thrombolysis for myocardial infarction	3
Parenteral treatment for hypoglycaemia	3
Hypertonic saline for hyponatraemia causing a specific neurological disturbance, such as seizures or reduced conscious level	2
Any intervention to achieve haemorrhage control, other than manual pressure or a dressing	2
Intravenous nitrates for acute heart failure	2
Intravenous antibiotics for meningitis or necrotising fasciitis	2
Emergency reperfusion of an ischaemic limb or organ (eg, testicular torsion)	2
Active rewarming for hypothermia	1
Intravenous treatment to lower life or disability threatening hypertension	1
Thrombolysis for massive pulmonary embolism	1
Any other, not previously specified—acyclovir for herpes simplex meningitis	1

CPR, cardiopulmonary resuscitation.

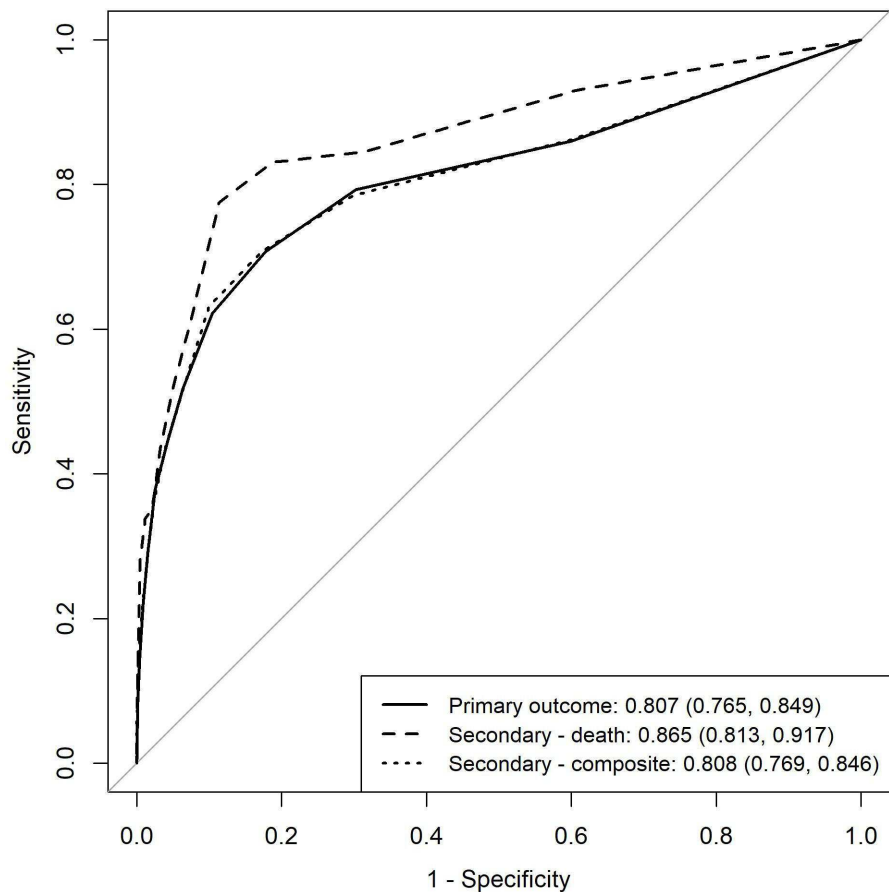
Table 4 shows the time-critical interventions that the false negative cases with NEWS2 $\leq$ 4 needed. Patients needing emergency surgery (18/21), intravenous antibiotics for open fractures (9/11), insulin infusion (5/8) and reduction of limb-threatening injury (5/5) often had NEWS2 $\leq$ 4. The frequency of intravenous

antibiotics for infection causing new organ dysfunction or shock among false negative cases reflects the high frequency of this time-critical intervention (14/66).

Online supplemental table 1 shows the characteristics of the false negative cases and online supplemental table 2 shows the

**Figure 1** Proportion of cases needing time-critical treatment at each National Early Warning Score version 2 (NEWS2) score.





**Figure 2** Area under the receiver operating characteristic (AUROC) (95% CI) for National Early Warning Score version 2 (NEWS2) and primary and secondary outcomes.

characteristics of the false positive cases with NEWS2>4 who did not need time-critical treatment. Online supplemental table 3 shows the frequency of NEWS2 parameters scoring maximum points in the false positive cases. Respiratory rate, supplemental oxygen and altered consciousness frequently contributed maximum points in false positive cases.

**DISCUSSION**

We found that NEWS2 had limited accuracy in predicting need for time-critical treatment among adult ED attenders. The c-statistic for predicting the need for time-critical treatment was 0.807 (95% CI 0.765 to 0.849). Using a threshold of NEWS2>4 would fail to predict around half of cases needing time-critical intervention and three-quarters of patients with NEWS2>4 would not need time-critical intervention. Most patients needing

emergency surgery, antibiotics for open fractures, insulin infusion or manipulation of limb-threatening injuries had NEWS2≤4. Patients with NEWS2>4 who did not need time-critical treatment frequently scored maximum points for respiratory rate, supplemental oxygen or conscious level on NEWS2.

Previous studies have shown that early warning scores have better accuracy for predicting death than for predicting hospital or intensive care admission.<sup>3–8</sup> Meta-analysis of nine studies of NEWS in ED patients reported c-statistics of 0.88, 0.86 and 0.77 for 24-hour, 48-hour and in-hospital mortality, 0.68 for hospital admission and 0.69 for intensive care admission.<sup>5</sup> A systematic review of 22 studies in acute medical units reported c-statistics of 0.7–0.9 for mortality and <0.6 for intensive care admission.<sup>6</sup> Our study showed similar prediction of mortality to previous studies.

Table 3 Diagnostic accuracy statistics (95% CI) for the primary outcome										
EWS threshold	TP	FP	FN	TN	Sensitivity	Specificity	PPV	NPV	LR+	LR–
NEWS2>4	85	245	79	3581	51.8% (44.2%, 59.3%)	93.6% (92.8%, 94.3%)	25.8% (21.3%, 30.7%)	97.8% (97.3%, 98.3%)	8.09 (6.69, 9.80)	0.51 (0.44, 0.60)
NEWS2>4 or parameter 3	98	428	66	3398	59.8% (52.1%, 67.0%)	88.8% (87.8%, 89.8%)	18.6% (15.5%, 22.2%)	98.1% (97.6%, 98.5%)	5.34 (4.58, 6.23)	0.45 (0.38, 0.55)
NEWS2>6	62	97	102	3729	37.8% (30.7%, 45.4%)	97.5% (96.9%, 97.9%)	39.0% (31.8%, 46.7%)	97.3% (96.8%, 97.8%)	14.91 (11.3, 19.69)	0.64 (0.57, 0.72)
NEWS2>6 or parameter 3	88	380	76	3446	53.7% (46.0%, 61.1%)	90.1% (89.1%, 91.0%)	18.8% (15.5%, 22.6%)	97.8% (97.3%, 98.3%)	5.40 (4.55, 6.41)	0.51 (0.44, 0.61)
EWS, early warning system; FN, false negative; FP, false positive; LR, likelihood ratio; NEWS2, National Early Warning Score version 2; NPV, negative predictive value; PPP, positive predictive value; TN, true negative; TP, true positive.										

**Table 4** Time-critical treatments that patients with NEWS2≤4 needed (n=79)

Intervention	Frequency
Emergency surgery to avoid death or significant disability, including surgical source control for infection	18
Intravenous antibiotics for infection causing new organ dysfunction or shock	14
Intravenous antibiotics for genuine open fractures	9
Administration of blood products for acute blood loss, to allow emergency procedures, in haematological emergencies, or severe anaemia in context of proven acute coronary syndrome	7
Insulin infusion as part of treatment protocol for diabetic ketoacidosis or hyperosmolar hyperglycaemic state	5
Large-volume intravenous fluid replacement (>2 L within 2 hours or >1 L with central venous monitoring)	5
Reduction of limb-threatening fracture or dislocation (including threat to skin, nerve or perfusion)	5
Antidote for life or disability threatening poisoning	5
Hyperkalaemia or hypokalaemia involving intravenous treatment and cardiac monitoring	4
Intravenous fluids and steroids for Addisonian crisis	4
Any intervention to support ventilation (except supplemental oxygen), other than during sedation for a procedure	4
Cardioversion or rate control for life-threatening arrhythmia	3
Primary percutaneous coronary intervention or thrombolysis for myocardial infarction	3
Unplanned airway intervention to provide a patent airway, other than during sedation for a procedure	2
Any intervention to achieve haemorrhage control, other than manual pressure or a dressing	2
Any intervention to support circulation (including CPR), other than intravenous fluids	2
Intravenous nitrates for acute heart failure	2
Emergency reperfusion of an ischaemic limb or organ (eg, testicular torsion)	2
Parenteral treatment for hypoglycaemia	2
Hypertonic saline for hyponatraemia causing a specific neurological disturbance, such as seizures or reduced conscious level	1
Intravenous treatment to lower life or disability threatening BP	1
Intravenous antibiotics for meningitis or necrotising fasciitis	1
Thrombolysis for massive pulmonary embolism	1
Any other, not previously specified	1

CPR, cardiopulmonary resuscitation; NEWS2, National Early Warning Score version 2.

Few studies have examined accuracy for time-critical treatment. Hong *et al*<sup>11</sup> compared the Emergency Severity Index and Simple Triage and Rapid Treatment triage tool in predicting 21/233 cases requiring emergent intervention. Hinson *et al*<sup>12</sup> systematically reviewed triage systems and identified a few studies that evaluated prediction for specific time-critical diagnoses, such as sepsis and myocardial infarction, but none that evaluated accuracy for time-critical interventions.

Our study has shown that it is possible to measure the accuracy of early warning scores for predicting need for time-critical treatment. We developed a reproducible method for adjudicating the primary outcome based on expert consensus and previous literature that we implemented with excellent interobserver agreement. Our study also had low rates of missing data due to the ability of the research team to access hospital records and a sample size that allowed accuracy to be estimated with reasonable precision.

Our study had limitations that need to be appreciated. The outcome adjudicators were not aware of the NEWS2 score but knowledge of the observations that comprise the NEWS2 could have influenced their judgements. We only used the first recorded NEWS2 score, whereas repeated scores may provide more information. The analyses of false negative and false positive cases were based on small numbers of cases. Our definition of time-critical intervention may be contested, with a large proportion of cases involving intravenous antibiotics for infection causing new organ dysfunction or shock, which is based on limited evidence.<sup>17</sup> Greater use of scale 2 to record oxygen saturation for patients with potential rather than known hypercapnic respiratory failure could reduce the number of false positives arising from low oxygen saturation. Finally, the findings

may not be generalisable to EDs with different case mix. Stroke and ST-elevation myocardial infarction requiring reperfusion may be identified as false negative cases in EDs that do not divert these cases to specialist services. NEWS2 may have lower sensitivity for identifying need for time-critical treatment in EDs that receive all such cases. Further research is therefore required to reproduce our findings in other settings.

The implications of our findings are that NEWS2 provides some useful information in identifying need for time-critical treatment, but ED staff should avoid over-reliance on NEWS2 in initial assessment. A substantial proportion of patients needing time-critical treatment will have a low NEWS2 score and most patients with NEWS2>4 will not require time-critical treatment. Concerns about over-reliance on NEWS2 have been raised elsewhere, most notably in a recent coroner's report to prevent future deaths, which concluded that over-reliance on NEWS2 during ED assessment contributed to the outcome.<sup>18</sup> We have identified time-critical interventions that NEWS2 predicts poorly and NEWS2 parameters that may overpredict need for time-critical intervention. Further research is required to confirm these findings in other settings and then explore whether NEWS2 can be modified or augmented to improve prediction.

X Richard Pilbery @999CPD

**Acknowledgements** We thank Enid Hirst and Linda Abouzeid (public representatives from Sheffield Emergency Care Forum) for providing the patient and public involvement outlined in the paper, Martin Bayley (Healthcare Computer Scientist, Sheffield Teaching Hospitals NHS Foundation Trust) for providing the hospital data, Erica Wallis (Research Coordinator, Sheffield Teaching Hospitals NHS Foundation Trust) for assistance with research governance and regulatory approvals, Anna Wilson and Sarah Bird (Research Nurses, Sheffield Teaching Hospitals NHS Foundation Trust) for undertaking the case record screening and the expert

consensus group that developed the outcome measure (Ed Carlton, Kirsty Challen, Tim Coats, Andreas Crede, Alasdair Gray, Graham Johnson, Fiona Lecky, Matthew Reed, Andrew Tabner).

**Contributors** SG conceived the study. SG, LS, GF and RP designed the study. SG acquired the data. SG, GF and AT undertook outcome adjudication. LS analysed the data. All authors interpreted the data. All authors contributed to drafting the manuscript. SG is the guarantor of the paper.

**Funding** The study was funded by the UK National Institute for Health Research (NIHR) Research for Patient Benefit (RfPB) programme (project reference NIHR204935).

**Disclaimer** The funder played no role in the study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

**Competing interests** All authors declare funding to their employing institutions from the NIHR RfPB programme (project reference NIHR204935).

**Patient and public involvement** Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants and was approved by the Health Research Authority in England (reference 23/HRA/4572) and Health and Care Research Wales without ethics committee review on the basis that it is a low-risk study involving retrospective case note review. The Health Research Authority approved the study without requiring patient consent and it was undertaken in accordance with the General Data Protection Regulations.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. Anonymised data are available from the corresponding author upon reasonable request (contact details on first page).

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

#### ORCID iDs

Steve Goodacre <http://orcid.org/0000-0003-0803-8444>  
Gordon Fuller <http://orcid.org/0000-0001-8532-3500>

Richard Pilbery <http://orcid.org/0000-0002-5797-9788>

#### REFERENCES

- 1 Smith E, Higginson I, Cleaver B, *et al*. Initial assessment of emergency department patients. Royal College of Emergency Medicine; 2017. Available: [https://www.rcem.ac.uk/docs/SDDC%20Initial%20Assessment%20\(Feb%202017\).pdf](https://www.rcem.ac.uk/docs/SDDC%20Initial%20Assessment%20(Feb%202017).pdf)
- 2 NHS England. Guidance for emergency departments: initial assessment. 2022. Available: <https://www.england.nhs.uk/guidance-for-emergency-departments-initial-assessment/> [Accessed 02 Aug 2024].
- 3 Royal College of Physicians. National early warning score (NEWS) 2: standardising the assessment of acute-illness severity in the NHS. updated report of a working party. London RCP; 2017.
- 4 NHS England Clinical Policy Unit. Resources to support the adoption of the national early warning score. 2019. Available: [https://www.england.nhs.uk/wp-content/uploads/2019/12/Resources\\_to\\_support\\_the\\_adoption\\_of\\_NEWSFINAL.pdf](https://www.england.nhs.uk/wp-content/uploads/2019/12/Resources_to_support_the_adoption_of_NEWSFINAL.pdf)
- 5 Arévalo-Buitrago P, Morales-Cané I, Olivares Luque E, *et al*. Predictive power of early-warning scores used in hospital emergency departments: a systematic review and meta-analysis. *Emerg* 2021;33:374–81.
- 6 Nannan Panday RS, Minderhoud TC, Alam N, *et al*. Prognostic value of early warning scores in the emergency department (ED) and acute medical unit (AMU): A narrative review. *Eur J Intern Med* 2017;45:20–31.
- 7 Patel R, Nugawela MD, Edwards HB, *et al*. Can early warning scores identify deteriorating patients in pre-hospital settings? A systematic review. *Resuscitation* 2018;132:101–11.
- 8 Guan G, Lee CMY, Begg S, *et al*. The use of early warning system scores in prehospital and emergency department settings to predict clinical deterioration: A systematic review and meta-analysis. *PLoS One* 2022;17:e0265559.
- 9 Goodacre S. Using clinical risk models to predict outcomes: what are we predicting and why? *Emerg Med J* 2023;40:728–30.
- 10 NHS Midlands and Lancashire Strategy Unit 2023. Emergency department acuity measurement and process: quick scoping review. 2023. Available: <https://www.strategyunitwm.nhs.uk/publications/emergency-department-acuity-measurement-and-process-quick-scoping-review> [Accessed 02 Aug 2024].
- 11 Hong R, Sexton R, Sweet B, *et al*. Comparison of START triage categories to emergency department triage levels to determine need for urgent care and to predict hospitalization. *Am J Disaster Med* 2015;10:13–21.
- 12 Hinson JS, Martinez DA, Cabral S, *et al*. Triage Performance in Emergency Medicine: A Systematic Review. *Ann Emerg Med* 2019;74:140–52.
- 13 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36.
- 14 Brown LD, Cai TT, DasGupta A. Interval Estimation for a Binomial Proportion. *Statist Sci* 2001;16:101–17.
- 15 Thomas B, Goodacre S, Lee E, *et al*. Prognostic accuracy of emergency department triage tools for adults with suspected COVID-19: the PRIEST observational cohort study. *Emerg Med J* 2021;38:587–93.
- 16 Hirst E, Irving A, Goodacre S. Patient and public involvement in emergency care research. *Emerg Med J* 2016;33:665–70.
- 17 Bion J, Barton G, Boyle A, *et al*. Academy of medical royal colleges statement on the initial antimicrobial treatment of sepsis. Academy of Medical Royal Colleges; 2022.
- 18 Irvine G. Dave Onawelo: prevention of future deaths report (ref 2024-0470). 2024. Available: [https://www.judiciary.uk/prevention-of-future-death-reports/dave-onawelo-prevention-of-future-deaths-report/#related\\_content](https://www.judiciary.uk/prevention-of-future-death-reports/dave-onawelo-prevention-of-future-deaths-report/#related_content) [accessed 03 Dec 2024]