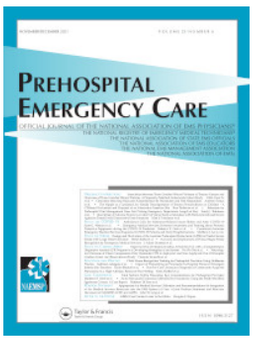

Prehospital Emergency Care

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ipec20

Tracking the National Early Warning Score 2 from prehospital care to the emergency department: A prospective, ambulance-based, observational study

Francisco Martín-Rodríguez, Ancor Sanz-García, Guillermo J. Ortega, Juan F. Delgado Benito, Silvia Aparicio Obregon, Francisco T. Martínez Fernández, Pilar González Crespo, Santiago Otero de la Torre, Miguel A. Castro Villamor & Raúl López-Izquierdo

To cite this article: Francisco Martín-Rodríguez, Ancor Sanz-García, Guillermo J. Ortega, Juan F. Delgado Benito, Silvia Aparicio Obregon, Francisco T. Martínez Fernández, Pilar González Crespo, Santiago Otero de la Torre, Miguel A. Castro Villamor & Raúl López-Izquierdo (2021): Tracking the National Early Warning Score 2 from prehospital care to the emergency department: A prospective, ambulance-based, observational study, Prehospital Emergency Care, DOI: 10.1080/10903127.2021.2011995

To link to this article: https://doi.org/10.1080/10903127.2021.2011995

Accepted author version posted online: 30 Nov 2021.

Tracking the National Early Warning Score 2 from prehospital care to the emergency department: A prospective, ambulance-based, observational study

Francisco Martín-Rodríguez, PhD[a,*], Ancor Sanz-García[b,*], PhD, Guillermo J. Ortega, PhD[b,g], Juan F. Delgado Benito, MD[c], Silvia Aparicio Obregon, PhD[d], Francisco T. Martínez Fernández, RN[c], Pilar González Crespo, RN[c], Santiago Otero de la Torre, MD[c], Miguel A. Castro Villamor, PhD[e], Raúl López-Izquierdo, PhD[f]

[a]Centro de Simulación Clínica Avanzada, Departamento de Medicina, Dermatología y Toxicología, Universidad de Valladolid. Gerencia de Emergencias Sanitarias, Gerencia Regional de Salud de Castilla y León (SACYL), Spain.

[b]Unidad de Análisis de Datos (UAD) del Instituto de Investigación Sanitaria del Hospital de la Princesa (IIS-IP), Madrid, Spain.

[c]Gerencia de Emergencias Sanitarias, Gerencia Regional de Salud de Castilla y León (SACYL), Spain.

[d]Universidad Europea del Atlántico. Santander., Spain.

[e]Centro de Simulación Clínica Avanzada, Departamento de Medicina, Dermatología y Toxicología, Universidad de Valladolid., Spain.

[f]Servicio de Urgencias, Hospital Universitario Rio Hortega de Valladolid, Gerencia Regional de Salud de Castilla y León (SACYL), Spain.

[g]Consejo Nacional de Investigaciones Científicas y Técnicas, CONICET, Argentina.

[*]Dr. Martín-Rodríguez and Sanz-García are joint first authors.

Author to whom any correspondence should be addressed: Francisco Martín-Rodríguez. Advanced Clinical Simulation Center, Faculty of Medicine, Valladolid University, Avda. Ramón y Cajal, 7, 47005 Valladolid., Spain; e-mail: fmartin@saludcastillayleon.es

Abstract

Aim of the study: To assess the prognostic ability of the National Early Warning Score 2 (NEWS2) at three time points of care -at the emergency scene (NEWS2-1), just before starting the transfer by ambulance to the hospital (NEWS2- 2), and at the hospital triage box (NEWS2-3)- to estimate in-hospital mortality after two days since the index event.

Methods: Prospective, multicenter, ambulance-based, cohort ongoing study in adults (>18 years) consecutively attended by advanced life support (ALS) and evacuated with high-priority to the emergency departments (ED) between October 2018 and May 2021. Vital sign measures were used to calculate the NEWS2 score at each time point, then this score was entered in a logistic regression model as the single predictor. Two outcomes were considered: first, all-cause mortality of the patients within 2 days of presentation to EMS, and second, unplanned ICU admission. **The calibration and scores comparison was performed by representing the predicted vs the observed risk curves according to NEWS score value.**

Results: 4943 patients were enrolled. Median age was 69 years (interquartile range 53- 81). The NEWS2-3 presented the better performance for all-cause two-day in-hospital mortality with an AUC of 0.941 (95% CI: 0.917-0.964), showing statistical differences with both the NEWS2-1 (0.872 (95% CI: 0.833-0.911); $p<0.003$) and with the NEWS2- 2 (0.895 (95% CI: 0.866-0.925; $p<0.05$). The calibration and scores comparison results showed that the NEWS2-3 was the best predictive score followed by the NEWS2-2 and the NEWS2-1, respectively.

Conclusions: The NEWS2 has an excellent predictive performance. The score showed a very consistent response over time with the difference between “at the emergency scene” and “pre-evacuation” presenting the sharpest change with decreased threshold values, thus displaying a drop in the risk of acute clinical impairment.

Keywords Clinical skills; Early warning score; In-hospital mortality; Physiological monitoring; Prehospital care; Vital sings

Introduction

Identifying the risk of clinical impairment in patients with acute diseases is not an easy task. In this sense, the role of the track-and-trigger system in recognizing and activating the appropriate response is well-known (1). Vital signs measurement at the emergency scene is a routine and mandatory practice that can be performed by untrained personnel providing considerable information about the patient's condition. The decontextualized interpretation of vital signs may result in important loss of information and failure to understand the actual patient's pathophysiological situation. The use of track-and-trigger

systems provides an aggregate interpretation of this set of vital signs that can make a critical difference by identifying sudden and subtle changes that may go unnoticed in isolation (2).

The implementation of scoring systems, in particular the National Early Warning Score (NEWS) developed in 2012 by the Royal College of Physicians of London (RCP) (3), is a clinical reality in diverse healthcare situations: emergency departments (ED), intensive care units (ICU), inpatient wards, nursing homes, emergency medical services (EMS), and tactical medicine (4) (5). The NEWS evaluates respiratory rate, pulse oxygen saturation, supplemental oxygen, systolic blood pressure, pulse, level of consciousness, and temperature, assigning different weights to individual measures and obtaining the total rating by adding them all together. In 2017, an update appeared -the National Early Warning Score 2 (NEWS2) (6)- implementing two pulse oxygen saturation scales, one for patients with type II respiratory failure and the other for the rest (7).

The most common gateway for medical emergencies is through the EMS, performing the appropriate basic and/or advanced life support techniques at the scene or *en route* to the ED. In this sense, clinical conditions such as cardiorespiratory arrest, acute myocardial infarction, stroke, major trauma, and interventions are automatically conducted in accordance with international guidelines (8) (9). However, most interventions are carried out under considerable levels of uncertainty.

Quick recognition and handling of critical medical emergencies is a major challenge for EMS (10), although low-data decision support tools like NEWS and NEWS2, which are validated for use in prehospital care (11) (12), can make a critical difference. Several studies have analyzed the performance of prehospital NEWS2 as a predictor of unplanned ICU-admission, or as a trigger of short- and long-term mortalities (13) (14). However, all of them present a static picture of the patient state rather than consider the temporal dynamic associated to the patients changes and interventions undertaken during the prehospital care.

The primary target of this research was to determine the evolving prognostic ability of NEWS2 by monitoring it at three time points: at the scene during the first contact with the patient (NEWS2-1), just before starting the transfer by ambulance to the hospital (NEWS2-2), and at the hospital triage box (NEWS2-3), with the objective of estimating in-hospital mortality two days since the index event. A secondary target of his work was to assess the ability of the NEWS2 tracking to identify the unplanned ICU-admission.

Methods

Study design and ethical issues

This is a prospective, multicenter, ambulance-based, cohort ongoing study in adults (>18 years) evacuated with high-priority to the emergency departments (ED) either in basic life support (BLS) or advance life support (ALS) teams. Data came from three consecutive studies conducted under the same criteria, starting in October 2018 and finishing in May 2021 (ISRCTN17676798, ISRCTN48326533, and ISRCTN49321933).

The study was validated by the institutional review committees of all involved institutions; it was conducted in compliance with the Declaration of Helsinki and following the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline (15).

Study setting

The study was conducted in four provinces (Burgos, Salamanca, Segovia, and Valladolid) of the Community of Castilla y León (Spain) with a reference target of 1,364,952 inhabitants. The project was conducted by the EMS with the collaboration of eight ALS and fifty-one BLS teams, one minor general district hospital, and four university tertiary care centers; these facilities came from the Castilla y León Public Health System.

The ALS staff is composed of a physician, an emergency registered nurse (ERN), and two emergency medical technician (EMT); the BLS is composed of two EMT. All resources operate in 24/7 non-stop mode, adhering to current clinical guidelines, and applying state-of-the-art basic and advanced life support protocols.

Inclusion and exclusion criteria

Inclusion criteria involved adult patients (> 18 years), firstly evaluated by ALS and later transferred to the ED by ALS or BLS, following the physician decision. Cases of cardiorespiratory arrest not recovered at the scene, end-of-life care situations, pregnant women, discharged in situ (after evaluation by the ALS physician), incident with danger to healthcare personnel at the scene (e.g., gunshots, bladed weapons, assaults in progress), or inability to collect informed consent at the site, *en route* or in the ED, were excluded.

To obtain informed consent, all patients, or family member or legal guardian in the patient absence, signed the informed consent forms which were collected by the ERN

and considered valid for the entire study and its later follow-up. When it was impossible obtaining consents during the prehospital care, a physician in each ED was assigned with the responsibility of procuring permissions. If consent was not received in any of the previous ways, the subject was removed from the study.

Outcomes

The primary outcome was all-cause 2-day in-hospital mortality, in line with similar studies (16) (17) (18). The secondary outcome was unplanned ICU-admission. Finally, a composite outcome was considered as the addition of 2-day mortality and ICU-admission.

Data collection and predictors

Before starting data collection, several meetings were held with the team of investigators to standardize the monitoring process.

Demographic and epidemiologic variables (sex, age, urban or rural area, nursing home derivation, vector of transfer, and intervention timing) and clinical input variables needed to perform baseline NEWS2 during first contact with the patient at the emergency scene -NEWS2-1 from now on- were recorded by the ERN. The respiratory rate was monitored by observation of the respiratory cycles for 30 seconds; in the case of very fast or slow breathing or irregular or shallow breathing, the respiratory cycles were monitored by direct auscultation with a stethoscope for one minute. Pulse oxygen saturation, systolic blood pressure, and pulse were determined with the LifePAK® 15 monitor-defibrillator (Physio-Control, Inc., Redmond, USA), and temperature with the ThermoScan® PRO 6000 thermometer (Welch Allyn, Inc, Skaneateles Falls, USA).

The ambulance physician registered the advanced life support procedures: intravenous medication, advanced airway management (non-invasive or invasive mechanical ventilation), and electrical treatments (defibrillation, cardioversion, or transcutaneous external pacemaker) as well as the prehospital diagnosis based on the International Classification of Diseases 11th Revision. Lastly, the ERN re-takes the variables to calculate another NEWS2 during the evacuation just before starting the transfer to the hospital in the ambulance -NEWS2-2 from now on-, irrespective of the way the patient was transferred, either by BLS or ALS.

For precise data linkage between EMS records and hospital's electronic medical records, a co-researcher from each ED matched at least five of the following identifiers: age, sex, ambulance code, date, time of admission, and health care card number (unique) from both records.

The follow-up period included 90-day since the ambulance pick-up. After the transfer of the inpatient to the ED, an ERN obtained again the variables required to calculate the NEWS2 at the hospital triage box -NEWS2-3 from now on-. Pulse oximetry saturation, systolic blood pressure, pulse, and temperature were measured using the Connex® Vital Signs Monitor (Welch Allyn, Inc, Skaneateles Falls, USA).

After completion the follow-up period, an associate investigator from each center checked the electronic medical records and collected hospital variables including hospital-inpatient, ICU-admissions, and all-cause 2-day in-hospital mortality.

Scoring calculation

The NEWS2 was estimated at the three time points based on the RCP guidelines (6) -baseline, evacuation, and hospital-. The researchers made an additional effort to locate patients with type II respiratory failure (95 patients described in the clinical history with a specialist's report), and the pulse oximetry saturation scale 2 was applied in these cases (19) (20). The use of supplemental oxygen was indirectly estimated by the fraction of inspired oxygen (FiO2) value; a FiO2 superior to 0.21 was recorded as supplemental oxygen use. Finally, the level of consciousness was derived from the Glasgow Coma Scale (GCS). A GCS less than 15 points was declared abnormal and categorized for the NEWS2 calculation as an altered level of consciousness. GCS levels have been described as equivalent to the alert/verbal/painful/unresponsive scale ones (21)

The categorical clinical risk derived from the NEWS2, as described in (6), was obtained in the following way: Low-risk (aggregate score 0-4 points), low-medium-risk (score of 3 points in any individual parameter), medium-risk (aggregate score of 5-6 points), and high-risk (aggregate score of 7 or more points). For the predictive validity and calibration calculations, NEWS2 points were used instead of the categorical levels.

Data analysis

Absolute values and percentages were used for categorical variables. For the case of continuous variables, median and interquartile ranges (IQR) were used because they did not follow a normal distribution. The characterization of the total sample and the association between each independent variable and the outcome was assessed by the Mann-Whitney U test or chi-squared test, when necessary. The standardized difference (SDF) was used to compare groups.

The discrimination assessment was evaluated for the three NEWS2 scores (NEWS2-1, NEWS2-2, and NEWS2-3) as well as for the differences between these scores

(NEWS2-1 – NEWS2-2, NEWS2-2 – NEWS2-3, and NEWS2-1 – NEWS2-3). In every case the discrimination was assessed by the area under the receiver operating characteristic (ROC) curve (AUC). Specificity, sensitivity, positive predictive value, negative predictive value, and positive and negative likelihood ratio were also calculated by computing the mean and confidence interval of each metric from all the points of the ROC curve. Moreover, the comparison between AUCs was performed using a Delong's test.

All AUCs described in the work were determined on a validation cohort, i.e., two thirds of the sample were used to fit the model and the other third to determine the validation capacity. This derivation/validation approach was used to avoid any potential interference when using GCS instead of level of consciousness.

A calibration of the score was also performed by calculating the calibration curve, that is, plotting predicted vs observed probability of the outcome according to the score value. The comparison of scores was assessed by using the Akaike Information Criterium (AIC); given a set of models, the lowest AIC values represents a better fitted model. For the calibration curve and the scores comparison, the whole cohort was used.

Results

Subject characteristics

Based on an initial cohort of 5138 patients evacuated with high priority by EMS and later discharged to ED, 195 cases were excluded due to missing data (3.79%), resulting in a final cohort for analysis of 4943 patients (see flowchart in Figure 1).

Table 1 shows the demographic characteristics and clinical data of patients according to NEWS2-1. The median age was 69 years (25th-75th percentile: 53-81; range 18-102). The proportion of males was significantly larger than of females (58.4% vs. 41.6%). The commonest cause of emergency calls was related with cardiovascular processes (39.6%, 1957 cases) followed by neurological diseases (18.2%, 902 cases) and trauma and injuries by external agents (13.1%, 648 cases). The all-cause 2-day in-hospital related mortality rate was 4.8% (239 cases) mainly due to the following causes: cardiovascular (34.7%, 83 cases), infections (15.5%, 37 cases), and neurological pathologies (13.8%, 33 cases). Non-survivors exhibited significantly superior median age, were predominantly males, with increased need for advanced life support interventions by EMS, with more unplanned ICU-admissions, and overall elevated

scores in all NEWS2 (baseline, evacuation, and hospital). Demographic characteristics and clinical data are described in supplementary Table S1.

NEWS2 calibration and scores comparison

The mortality distribution according to NEWS2 values and the predicted probability of mortality for each time point: a) NEWS2-1, b) NEWS2-2, c) NEWS2-3 are shown in Figure 2. As can be observed in panels a), b), and c) of Figure 2, the predicted curves presented a typical sigmoid shape, reaching higher probabilities of death (80%) for higher values of NEWS2. For the NEWS2-1 we observed that there is an underestimated mortality in patients with 20% to 60% of real probability of death, for NEWS2-2 there is an overestimation in patients with 40% to 80% of real probability of death, and finally, NEWS2-3, presents the better fit between real and predicted curves. Calibration curves for the other two outcomes can be found in supplementary Figure S1 and S2. Supplementary Table S2 shows the AIC results; the models with better fitting in order of increased AIC are (starting with the lowest AIC, which corresponds to the best fit): NEWS2-3, NEWS2-2, NEWS2-1, NEWS2-2-3, NEWS2-1-3, NEWS2-1-2; this order is similar for the other outcomes.

NEWS2 discrimination

NEWS2-3 presented the better performance for mortality with an AUC of 0.941 (95% CI: 0.917-0.964), statistically different from the NEWS2-1 (0.872 (95% CI: 0.833-0.911); $p<0.003$) and from the NEWS2-2 (0.895 (95% CI: 0.866-0.925; $p<0.05$). **The AUCs of the NEWS2 time points outperformed the NEWS2 time points differences, with all differences being significant ($p<0.001$).** The calibration results reached an accuracy of 95.3%, 94.5%, and 95.5% for NEWS2-1, NEWS2-2, and NEWS2-3, respectively. No differences were found between NEWS2 time points for the other two outcomes, presenting similar AUCs (all AUCs >0.75) and accuracies (all accuracies >87%). Further details, including the likelihood ratios of both the predictive validity for mortality and the previous outcome, can be found in Table 2.

NEWS2 predictive ability

Differences between NEWS2 scores were interpreted as the effects of the received treatments and patients' evolution. Interestingly, none of the NEWS2 time points differences presented relevant AUC or accuracies for any outcome; further details on the AUC results and metrics can be found in Supplementary Table S3. This null prognostic capacity can be explained by the lack of variability between time points; the median [first quartile, third quartile] for NEWS2-1-2, NEWS2-2-3, and NEWS2-1-3

were 1 [-1, 2], 0 [-1, 1], and 0 [-1, 2], respectively. These results showed that NEWS2 is rather stable along time as can be observed in Figure 3, where the black lines are practically flat, being the difference between the first and the second NEWS2 the one presenting the highest change. These results were similar for the other two outcomes, as can be seen in Supplementary Figures S3 and S4.

Discussion

Main findings

To the best of our knowledge, this is the first prospective, multicenter, EMS-delivery study examining the NEWS2 performance across three stages, from the prehospital care to the hospital check-in, tracking in this way the patient's condition on a real-time. We noted that NEWS2-3 (hospital triage box) presents the best AUC, followed by NEWS2-2 (before evacuation) and finally by the NEWS2-1 (basal), this was also observed in the calibration and in the scores comparison by the AIC. The temporal dynamics of NEWS2 was very stable and consistent, with score ratings generally decreasing along time, suggesting that NEWS2 can be employed right from the prehospital care stage with confidence in its predictive power and its capability to detect short-term clinical impairment.

Prehospital NEWS2 is an evidence-based system which effectively identify high-risk patients allowing a rapid response (14) (22). Comparable trials have shown how the use of prehospital NEWS and NEWS2 can help in anticipating complications by an early recognition of serious adverse events (23). Pirneskoski *et al*. (13) reported an AUC for one-day mortality of 0.840; in the same way, Hoikka et al. (14) presented an AUC of 0.801. Silckock *et al*. (24) showed how the NEWS has superior performance than the quick Sepsis-related Organ Failure Assessment Score for ICU-admissions in 30-day mortality (AUC of 0.740 vs. 0.679). The prehospital NEWS2 has also proved its benefits in selected cohorts, as showed by Mitsunaga et al. (25) obtaining an AUC of 0.789 for in-hospital mortality of older adults. Our findings in the prehospital NEWS2s (basal and prior evacuation) are significantly superior to those previously described (AUC=0.866 and 0.895).

The question arises, however, if a unique calculation of the score is enough to yield an adequate and reliable risk of deterioration or, conversely, subsequent calculations would improve the sensitivity to predict the short-term disease evolution. Different studies have compared the predictive capacity of continuous vs. discontinuous monitoring for a

timely recognition of impairment (26) (27). Petersen *et al*. (28) showed no significant discrepancies between alternative measurements of early warning scores, although the study encompassed only patients with initial NEWS of 0 or 1. Along the same lines, Van Velthoven *et al*. (29) and Weenk *et al*. (30) have examined the feasibility of monitoring vital signs in real time to estimate early warning scores continuously.

Different scores and diagnostic aids based on vital signs have been developed for the primary screening of patients with acute diseases (31). All of them have in common the inclusion of easily obtainable variables which, together with a quick learning curve, allows them to be used by prehospital personnel (32) (33).

Continuous monitoring may seem reasonable (2) (34) although in the prehospital care, this choice may be unavailable in all situations (traffic accidents, transfer from scene to the ambulance, movements, and vibrations during the transfer, etc.). Our findings demonstrate that a sole NEWS2 measurement has a consistently predictive capacity, even during the first contact with the patient; additionally, differences between time points do not present good predictive capacity. These facts support the validity of unique NEWS measurements.

Limitations

This study has certain limitations. First, it is a convenience cohort, **patients should provide informed consent to participate and patients with missing data were excluded, which may introduce a selection bias**. To avoid bias, subjects were recruited 24/7 during all months in rural and urban locations, with any condition, and for the complete period of the survey. Second, the study began in 2018 and is still in progress. The current coronavirus 2019 pandemic may have impacted the EMS. Consolidated data are required for comparison with similar time points to assess the actual incidence of coronavirus disease 2019. Third, the data extractors were not blind; thus, to diminish inter-observer confounding, a pre-training protocol on data collection was carried out with all the associated investigators. Also, a non-controversial primary outcome -all-cause 2-day in-hospital mortality- was adopted and, as a double-check system, the PI checked all the cases with the positive primary outcome. **Fourth, the scores are not independent from the effect of treatments, therefore, the NEWS2-3, the last one, could be best because it detects patients that are refractory to treatments. Fifth and final**, although the application of early warning scores is a straightforward procedure, it requires basic training and implementation in the

prehospital care (35) (36). The adoption of this standardized scoring system is not yet widespread in all regular EMS-systems (22).

Conclusion

The NEWS2 has an excellent predictive performance. The score showed a very consistent response over time, with the difference between emergency scene NEWS2 and pre-evacuation NEWS2 presenting the sharpest change with decreased threshold values and therefore, a drop in the risk of acute clinical impairment, partially explained by the pre-hospital care.

EMS facilities the patients gateway to health care system; initial examination on-scene and rapid identification of clinical deterioration is a key challenge. NEWS2 can help in anticipating complications by assisting in the clinical decision-making, so the track-and-trigger system should be a standardized EMS practice.

Declarations

Funding

This work was supported by the Gerencia Regional de Salud, Public Health System of Castilla y León (Spain) [grant number GRS 1678/A/18, GRS 1903/A/19 and GRS 2131/A/20].

Conflicts of interest/Competing interests

None.

Ethics approval

This study was approved by the Health Research Ethics Board from all participating centers. The study is registered in the WHO International Clinical Trials Registry Platform (ICTRP) with number [ISRCTN17676798, ISRCTN48326533 and ISRCTN49321933]. Details of the study design, statistical analysis plan and raw data are available online.

Transparency declaration

The corresponding author on behalf of the other authors guarantee the accuracy, transparency and honesty of the data and information contained in the study, that no relevant information has been omitted and that all discrepancies between authors have been adequately resolved and described.

This article is an original work, has not been published before, and is not being considered for publication elsewhere in its final form, in either printed or electronic media. It is not based on any previous communication to a society or meeting.

References

1. Campbell V, Conway R, Carey K, Tran K, Visser A, Gifford S, et al. Predicting clinical deterioration with Q-ADDS compared to NEWS, Between the Flags, and eCART track and trigger tools. Resuscitation. 2020; 153: p. 28-34.

2. Credland N, Dyson J, Johnson MJ. Do early warning track and trigger tools improve patient outcomes? A systematic synthesis without meta-analysis. J Adv Nurs. 2021; 77(2): p. 622-634.

3. Royal College of Physicians. National Early Warning Score (NEWS): Standardising the assessment of acuteillness severity in the NHS. Report of a working party. London: RCP 2. 2012.

4. Scott LJ, Redmond NM, Garrett J, Whiting P, Northstone K, Pullyblank A. Distributions of the National Early Warning Score (NEWS) across a healthcare system following a large-scale roll-out. Emerg Med J. 2019; 36(5): p. 287-292.

5. Brangan E, Banks J, Brant H, Pullyblank A, Le Roux H, Redwood S. Using the National Early Warning Score (NEWS) outside acute hospital settings: a qualitative study of staff experiences in the West of England. BMJ Open. 2018; 8(10): p. e022528.

6. 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 2. 2017.

7. Pimentel MAF, Redfern OC, Gerry S, Collins GS, Malycha J, Prytherch D, et al. A comparison of the ability of the National Early Warning Score and the National Early Warning Score 2 to identify patients at risk of in-hospital mortality: A multi-centre database study. Resuscitation. 2019; 134: p. 147-156.

8. Panchal AR, Bartos JA, Cabañas JG, Donnino MW, Drennan IR, Hirsch KG, et al. Part 3: Adult Basic and Advanced Life Support: 2020 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation. 2020; 142(16_suppl_2): p. S366-S468.

9. Teuben M, Löhr N, Jensen KO, Brüesch M, Müller S, Pfeifer R, et al. Improved pre-hospital care efficiency due to the implementation of pre-hospital trauma life support (PHTLS) algorithms. Eur J Trauma Emerg Surg. 2020; 46(6): p. 1321-1325.

10. Magnusson C, Herlitz J, Axelsson C. Pre-hospital triage performance and emergency medical services nurse\'s field assessment in an unselected patient population attended to by the emergency medical services: a prospective observational study. Scand J Trauma Resusc Emerg Med. 2020; 28(1): p. 81.

11. Silcock DJ, Corfield AR, Gowens PA, Rooney KD. Validation of the National Early Warning Score in the prehospital setting. Resuscitation. 2015; 89: p. 31-35.

12. Martín-Rodríguez F, Sanz-García A, Medina-Lozano E, Castro Villamor MÁ, Carbajosa Rodríguez V, Del Pozo Vegas C, et al. Prehosp Emerg Care. The Value of Prehospital Early Warning Scores to Predict in - Hospital Clinical Deterioration: A Multicenter, Observational Base-Ambulance Study. 2020; Online ahead of print: p. DOI: 10.1080/10903127.2020.1813224.

13. Pirneskoski J, Kuisma M, Olkkola KT, Nurmi J. Prehospital National Early Warning Score predicts early mortality. Acta Anaesthesiol Scand. 2019; 63(5): p. 676-683.

14. Hoikka M, Silfvast T, Ala-Kokko TI. Does the prehospital National Early Warning Score predict the short-term mortality of unselected emergency patients? Scand J Trauma Resusc Emerg Med. 2018; 26(1): p. 48.

15. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007; 370(9596): p. 1453-1457.

16. Lane DJ, Wunsch H, Saskin R, Cheskes S, Lin S, Morrison LJ, et al. Assessing Severity of Illness in Patients Transported to Hospital by Paramedics: External Validation of 3 Prognostic Scores. Prehosp Emerg Care. 2020; 24(2): p. 273-281.

17. Williams TA, Tohira H, Finn J, Perkins GD, Ho KM. The ability of early warning scores (EWS) to detect critical illness in the prehospital setting: A systematic review. Resuscitation. 2016; 102: p. 35-43.

18. Nannan Panday RS, Minderhoud TC, Alam N, Nanayakkara PWB. 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: p. 20-31.

19. Williams B. The National Early Warning Score 2 (NEWS2) in patients with hypercapnic respiratory failure. Clin Med (Lond). 2019; 19(1): p. 94-95.

20. Tirkkonen J, Karlsson S, Skrifvars MB. National early warning score (NEWS) and the new alternative SpO scale during rapid response team reviews: a prospective observational study. Scand J Trauma Resusc Emerg Med. 2019; 27(1): p. 111.

21. Kelly CA, Upex A, Bateman DN. Comparison of consciousness level assessment in the poisoned patient using the alert/verbal/painful/unresponsive scale and the Glasgow Coma Scale. Ann Emerg Med. 2004; 44(2):108-13

22. Shaw J, Fothergill RT, Clark S, Moore F. Can the prehospital National Early Warning Score identify patients most at risk from subsequent deterioration? Emerg Med J. 2017; 34(8): p. 533-537.

23. Endo T, Yoshida T, Shinozaki T, Motohashi T, Hsu HC, Fukuda S, et al. Efficacy of prehospital National Early Warning Score to predict outpatient disposition at an emergency department of a Japanese tertiary hospital: a retrospective study. BMJ Open. 2020; 10(6): p. e34602.

24. Silcock DJ, Corfield AR, Staines H, Rooney KD. Superior performance of National Early Warning Score compared with quick Sepsis-related Organ Failure Assessment Score in predicting adverse outcomes: a retrospective observational study of patients in the prehospital setting. Eur J Emerg Med. 2019; 26(6): p. 433-439.

25. Mitsunaga T, Hasegawa I, Uzura M, Okuno K, Otani K, Ohtaki Y, et al. Comparison of the National Early Warning Score (NEWS) and the Modified Early Warning Score (MEWS) for predicting admission and in-hospital mortality in elderly patients in the pre-hospital setting and in the emergency department. PeerJ. 2019; 7: p. e6947.

26. Downey CL, Croft J, Ainsworth G, Buckley H, Shinkins B, Randell R, et al. Trial of remote continuous versus intermittent NEWS monitoring after major surgery (TRaCINg): a feasibility randomised controlled trial. Pilot Feasibility Stud. 2020; 6(1): p. 183.

27. Watkinson PJ, Pimentel MAF, Clifton DA, Tarassenko L. Manual centile-based early warning scores derived from statistical distributions of observational vital-sign data. Resuscitation. 2018; 129: p. 55-60.

28. Petersen JA, Antonsen K, Rasmussen LS. Frequency of early warning score assessment and clinical deterioration in hospitalized patients: A randomized trial. Resuscitation. 2016; 101: p. 91-96.

29. Van Velthoven MH, Adjei F, Vavoulis D, Wells G, Brindley D, Kardos A. ChroniSense National Early Warning Score Study (CHESS): a wearable wrist device to measure vital signs in hospitalised patients-protocol and study design. BMJ Open. 2019; 9(9): p. e028219.

30. Weenk M, Koeneman M, van de Belt TH, Engelen LJLPG, van Goor H, Bredie SJH. Wireless and continuous monitoring of vital signs in patients at the general ward. Resuscitation. 2019; 136: p. 47-53.

31. Clemency BM, Murk W, Moore A, Brown LH. The EMS Modified Early Warning Score (EMEWS): A Simple Count of Vital Signs as a Predictor of Out-of-Hospital Cardiac Arrests. Prehosp Emerg Care. 2021; Online ahead of print(DOI: 10.1080/10903127.2021.1908464): p. 1-22.

32. Patel R, Nugawela MD, Edwards HB, Richards A, Le Roux H, Pullyblank A, et al. Can early warning scores identify deteriorating patients in pre-hospital settings? A systematic review. Resuscitation. 2018; 132: p. 101-111.

33. Martín-Rodríguez F, Castro-Villamor MÁ, Del Pozo Vegas C, Martín-Conty JL, Mayo-Iscar A, Delgado Benito JF, et al. Analysis of the early warning score to detect critical or high-risk patients in the prehospital setting. Intern Emerg Med. 2019; 14(4): p. 581-589.

34. Lyckhage LF, Hansen ML, Procida K, Wienecke T. Prehospital continuous ECG is valuable for very early detection of atrial fibrillation in patients with acute stroke. J Stroke Cerebrovasc Dis. 2020; 29(9): p. 105014.

35. Downey CL, Tahir W, Randell R, Brown JM, Jayne DG. Strengths and limitations of early warning scores: A systematic review and narrative synthesis. Int J Nurs Stud. 2017; 76: p. 106-119.

36. Wuytack F, Meskell P, Conway A, McDaid F, Santesso N, Hickey FG, et al. The effectiveness of physiologically based early warning or track and trigger systems after triage in adult patients presenting to emergency departments: a systematic review. BMC Emerg Med. 2017; 17(1): p. 38.

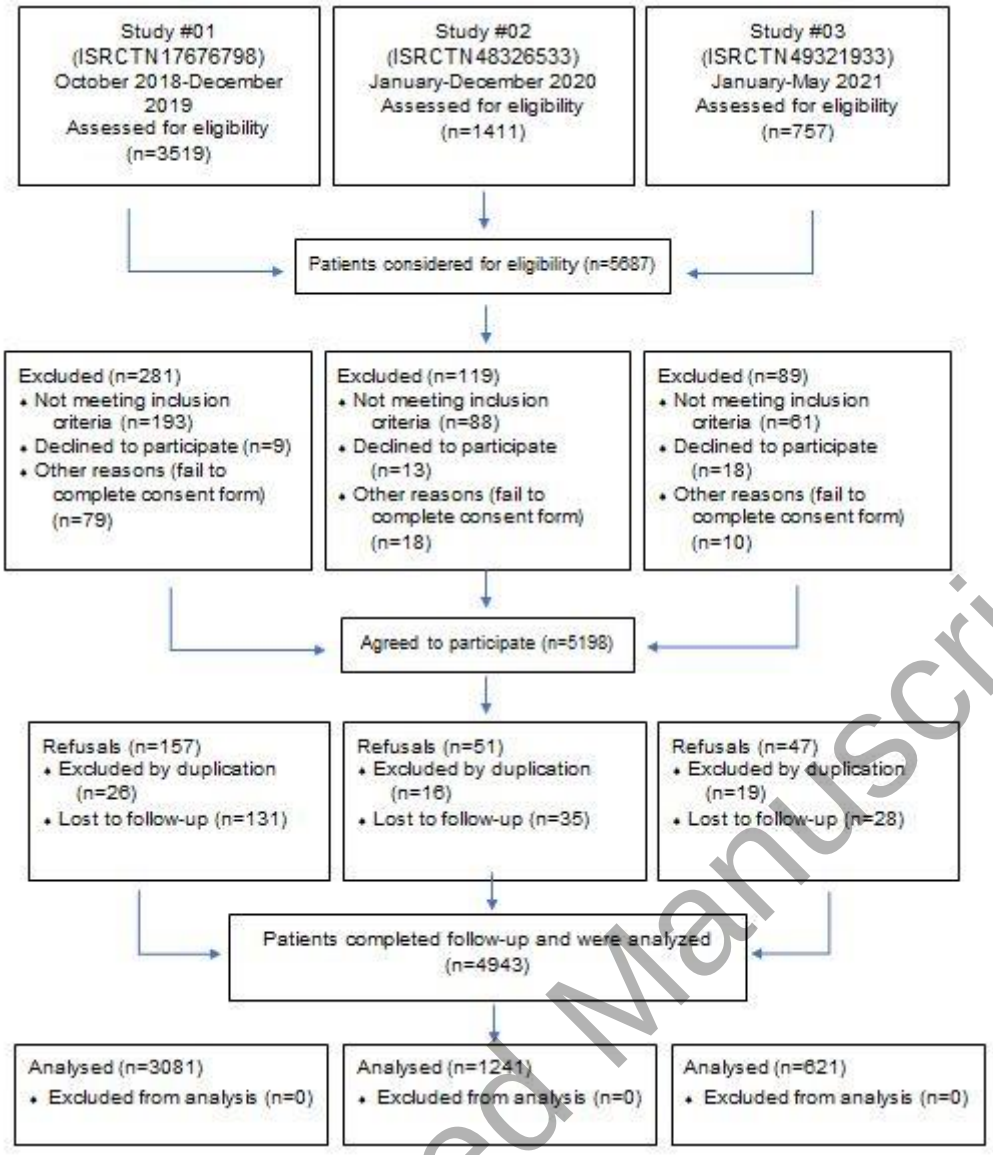

Figure 1. Flow chart of study population.

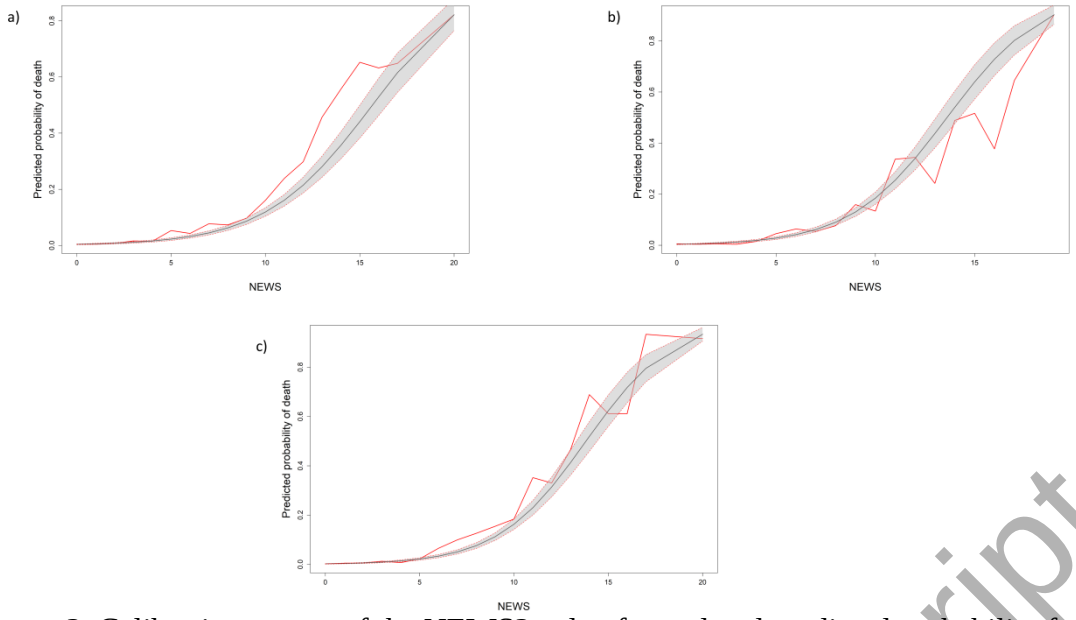

Figure 2. Calibration curves of the NEWS2 value for real and predicted probability for two day-mortality for each time point a) NEWS2-1, b) NEWS2-2, c) NEWS2-3. The grey area of the trend line corresponds to the 95% confidence interval of the predicted probability of the outcome (trend line). The red line corresponds to the real probability of the outcome.

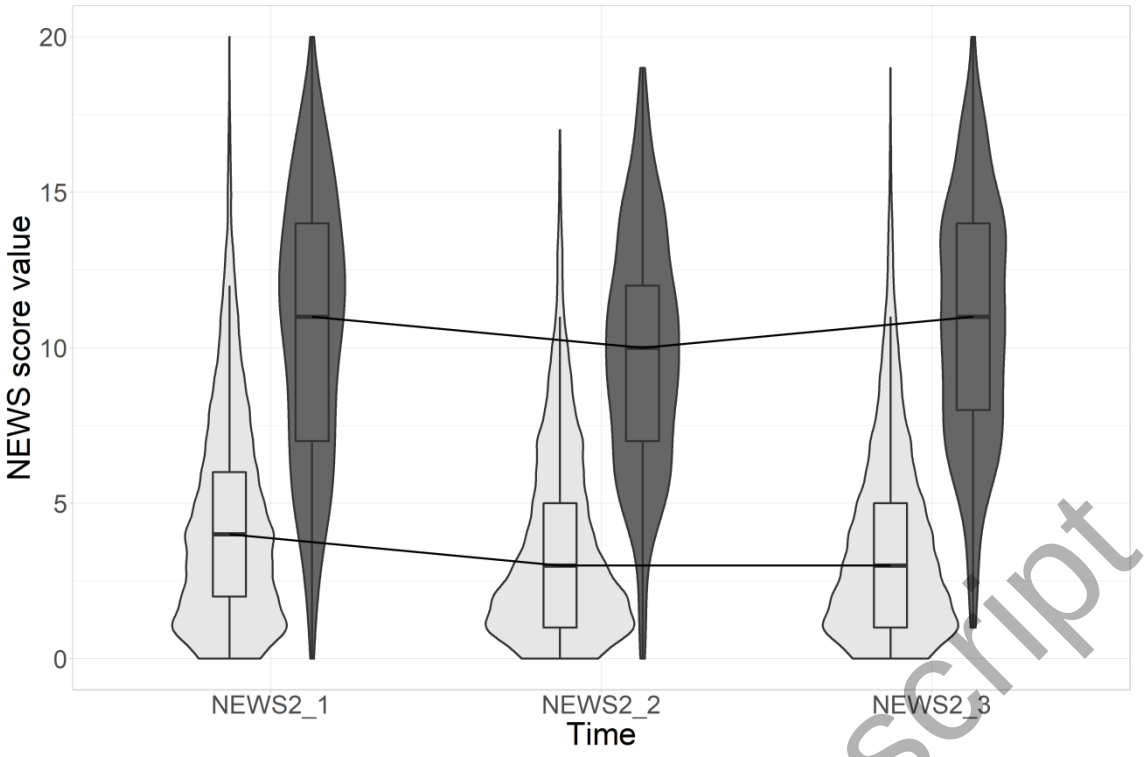

Figure 3. Boxplot and violin plot (density) of the NEWS2 value for each time point (NEWS2-1, NEWS2-2, and NEWS2-3). Grey color represents non-survivors and black color represent survivors, black line connects times from the same group.

Table 1. Balance summary across all treatment pairs according to basal NEWS2 stratification.

Variable	N	Low-risk, N = 2,154[1]	low-medium-risk, N = 650[1]	medium-risk, N = 797[1]	high-risk, N = 1,342[1]	Standardized difference
Age, years	4,943	65 (51, 78)	66 (51, 80)	71 (54, 82)	76 (60, 84)	0.4327
Sex	4,943					0.0420
Males		1,290 (60%)	362 (56%)	445 (56%)	792 (59%)	
Females		864 (40%)	288 (44%)	352 (44%)	550 (41%)	
Advanced life support	4,943					0.2370
No		1,085 (50%)	403 (62%)	546 (69%)	994 (74%)	
Yes		1,069 (50%)	247 (38%)	251 (31%)	348 (26%)	
Zone	4,943					0.0117
Urban		1,825 (85%)	552 (85%)	668 (84%)	1,124 (84%)	
Rural		329 (15%)	98 (15%)	129 (16%)	218 (16%)	
Nursing home derivation	4,943					0.0775
No		2,024 (94%)	601 (92%)	734 (92%)	1,157 (86%)	
Yes		130 (6.0%)	49 (7.5%)	63 (7.9%)	185 (14%)	
Time of Arrival, minutes	4,943	10 (8, 14)	10 (8, 14)	11 (8, 14)	10 (8, 14)	0.0558
Time of Assistance, minutes	4,942	28 (22, 35)	29 (23, 36)	29 (24, 37)	32 (25, 39)	0.3663
Time of Transfer, minutes	4,943	10 (7, 15)	10 (7, 15)	10 (7, 15)	10 (7, 15)	0.0750
Type II respiratory failure	4,943					0.0446
No		2,145 (100%)	648 (100%)	777 (97%)	1,278 (95%)	
Yes		9 (0.4%)	2 (0.3%)	20 (2.5%)	64 (4.8%)	
Intravenous medication, number	4,943	1.00 (1.00, 2.00)	1.00 (1.00, 2.00)	2.00 (1.00, 3.00)	3.00 (1.00, 4.00)	0.8418
NIMV	4,943					0.0917
No		2,154 (100%)	647 (100%)	789 (99%)	1,219 (91%)	
Yes		0 (0%)	3 (0.5%)	8 (1.0%)	123 (9.2%)	
IMV	4,943					0.1481
No		2,152 (100%)	635 (98%)	768 (96%)	1,142 (85%)	
Yes		2 (<0.1%)	15 (2.3%)	29 (3.6%)	200 (15%)	
Defibrillation	4,943					0.0097
No		2,153 (100%)	650 (100%)	797 (100%)	1,329 (99%)	
Yes		1 (<0.1%)	0 (0%)	0 (0%)	13 (1.0%)	
Cardioversion	4,943					0.0343
No		2,146 (100%)	636 (98%)	783 (98%)	1,291 (96%)	
Yes		8 (0.4%)	14 (2.2%)	14 (1.8%)	51 (3.8%)	
Pacemaker	4,943					0.0238
No		2,149 (100%)	644 (99%)	787 (99%)	1,307 (97%)	
Yes		5 (0.2%)	6 (0.9%)	10 (1.3%)	35 (2.6%)	
Hospital-inpatient	4,943					0.3563
No		1,196 (56%)	313 (48%)	300 (38%)	267 (20%)	
Yes		958 (44%)	337 (52%)	497 (62%)	1,075 (80%)	
Polyvalent-ICU admission	4,943					0.1674
No		2,094 (97%)	608 (94%)	718 (90%)	1,080 (80%)	
Yes		60 (2.8%)	42 (6.5%)	79 (9.9%)	262 (20%)	
Two days mortality	4,943					0.1392
No		2,144 (100%)	642 (99%)	769 (96%)	1,149 (86%)	
Yes		10 (0.5%)	8 (1.2%)	28 (3.5%)	193 (14%)	
Circulatory	4,943					0.1151
No		2,074 (96%)	626 (96%)	760 (95%)	1,138 (85%)	
Yes		80 (3.7%)	24 (3.7%)	37 (4.6%)	204 (15%)	
Neurol	4,943					0.0149
No		2,125 (99%)	638 (98%)	781 (98%)	1,304 (97%)	
Yes		29 (1.3%)	12 (1.8%)	16 (2.0%)	38 (2.8%)	
Trauma	4,943					0.1057
No		1,819 (84%)	482 (74%)	603 (76%)	1,137 (85%)	
Yes		335 (16%)	168 (26%)	194 (24%)	205 (15%)	

Resp	4,943					0.2100
No		1,103 (51%)	371 (57%)	543 (68%)	969 (72%)	
Yes		1,051 (49%)	279 (43%)	254 (32%)	373 (28%)	
Infection	4,943					0.1884
No		2,112 (98%)	633 (97%)	710 (89%)	1,063 (79%)	
Yes		42 (1.9%)	17 (2.6%)	87 (11%)	279 (21%)	
poisoning	4,943					0.0332
No		2,023 (94%)	618 (95%)	749 (94%)	1,305 (97%)	
Yes		131 (6.1%)	32 (4.9%)	48 (6.0%)	37 (2.8%)	
Digestive	4,943					0.0755
No		1,794 (83%)	577 (89%)	705 (88%)	1,219 (91%)	
Yes		360 (17%)	73 (11%)	92 (12%)	123 (9.2%)	
Others[2]	4,943					0.0281
No		2,028 (94%)	605 (93%)	728 (91%)	1,259 (94%)	
Yes		126 (5.8%)	45 (6.9%)	69 (8.7%)	83 (6.2%)	

Abbreviations: NEWS2: National Early Warning Score 2; NIMV: non-invasive mechanical ventilation; IMV: invasive mechanical ventilation; ICU: intensive care unit.

1 Median (IQR) or n (%)

2 Other pathology: endocrine, genitourinary, diseases of the blood and the immune system.

Table 2. Statistical details of the risk-models for NEWS2 for two-day mortality, ICU-admission, and composite outcome (two-day mortality and ICU).

			2-day in-hospital mortality				
	Specificity	Sensitivity	PPV	NPV	PLR	NLR	AUC
NEWS2-1	74.3 (59.5-89.1)	55.3 (38.3-72.3)	27.3 (18.4-35.8)	97.5 (96.7-98.3)	9.46 (5.54-13.3)	0.50 (0.33-0.66)	0.872 (0.833-0.911)
NEWS2-2	77.1(62.2-91.7)	53.9 (35.8-72.1)	37.8 (23.6-52.1)	97.5 (96.6-98.3)	23.1 (7.69-38.4)	0.50 (0.33-0.68)	0.895 (0.866-0.924)
NEWS2-3	77.4 (63.1-91.8)	55.9 (38.4-73.5)	38.7 (25.8-51.5)	97.6 (96.8-98.5)	21.5 (9.70-33.3)	0.47 (0.30-0.65)	0.940 (0.916-0.964)
			ICU-admissions				
NEWS2-1	74.4 (59.7-89.1)	43.1 (26.5-59.7)	25.3 (18.9-31.7)	93.8 (92.6-94.9)	3.96 (2.65-5.26)	0.67 (0.54-0.80)	0.758 (0.718-0.797)
NEWS2-2	77.1 (62.5-91.7)	40.7 (23.1-58.4)	21.6 (16.2-27.1)	93.7 (92.4-95.1)	3.07 (2.18-3.96)	0.68 (0.54-0.83)	0.786 (0.751-0.821)
NEWS2-3	77.5 (63.3-91.7)	40.8 (23.8-57.9)	23.9 (17.3-30)	93.7 (92.5-95.1)	3.58 (2.43-4.74)	0.68 (0.53-0.82)	0.776 (0.738-0.813)
			Combined outcomes (ICU-Mortality)				
NEWS2-1	75.4 (60.7-90.2)	45.8 (29.1-62.5)	41.6 (32.1-51.1)	92.1 (90.5-93.8)	7.11 (4.4-9.7)	0.62 (0.48-0.76)	0.795 (0.762-0.829)
NEWS2-2	78.2 (63.5-92.8)	44.1 (26.3-61.7)	49.5 (36.7-62.3)	92.1 (90.3-93.9)	17.7 (4.05-31.4)	0.63 (0.48-0.78)	0.834 (0.805-0.863)
NEWS2-3	78.7 (64.4-92.9)	44.7 (27.6-61.7)	51.8 (39.3-64.4)	92.2 (90.5-94.1)	17.7 (6.74-28.8)	0.62 (0.47-0.77)	0.833 (0.802-0.864)

Abbreviations: NEWS2: National Early Warning Score 2; ICU: intensive care unit; PPV: positive predictive value; NPV: negative predictive value; PLR: positive likelihood ratio; NLR: negative likelihood ratio; AUC: area under the curve.

[a] Bracketed number indicate 95% confidence interval

[b] NEWS2-1 (bedside evaluation); NEWS2-2 (before ambulance evacuation); NEWS2-3 (emergency department triage box)