BMJ Open

Ability of the National Early Warning Score and its respiratory and haemodynamic subcomponents to predict short-term mortality on general wards: a prospective three-centre observational study in Finland

Eetu Loisa, Antti Kallonen, Sanna Hoppu, Joonas Tirkkonen

ABSTRACT

Objectives To validate the ability of the National Early Warning Score (NEWS) to predict short-term mortality on hospital wards, with a special reference to the NEWS’s respiratory and haemodynamic subcomponents.

Design A large, 1-year, prospective, observational three-centre study. First measured vital sign datasets on general wards were prospectively collected using a mobile solution system during routine patient care. Area under receiver operator characteristic curves were constructed, and comparisons between ROC curves were conducted with Delong’s test for two correlated ROC curves.

Setting One university hospital and two regional hospitals in Finland.

Participants All 19001 adult patients admitted to 45 general wards in the three hospitals over the 1-year study period. After excluding 102/19 001 patients (0.53%) with data on some vital signs missing, the final cohort consisted of 18889 patients with full datasets.

Primary and secondary outcome measures The primary outcome measure was 1-day mortality and secondary outcomes were 2-day and 30-day mortality rates.

Results Patients’ median age was 70 years, 51% were male and 31% had a surgical reason for admission. The 1-day mortality was 0.36% and the 30-day mortality was 3.9%. The NEWS discriminated 1-day non-survivors with excellent accuracy (AUROC 0.91, 95% CI 0.87 to 0.95) and 30-day mortality with acceptable accuracy (0.75, 95% CI 0.73 to 0.77). The NEWS’s respiratory rate component discriminated 1-day non-survivors better (0.78, 95% CI 0.72 to 0.84) as compared with the oxygen saturation (0.66, 95% CI 0.59 to 0.73), systolic blood pressure (0.65, 95% CI 0.59 to 0.72) and heart rate (0.67, 95% CI 0.61 to 0.74) subcomponents (p<0.01 in all ROC comparisons). As with the total NEWS, the discriminative performance of the individual score components decreased substantially for the 30-day mortality.

Conclusions NEWS discriminated general ward patients at risk for acute death with excellent statistical accuracy. The respiratory rate component is especially strongly associated with short-term mortality.

Strengths and limitations of this study

- This study is highlighted by a large and heterogeneous sample size with near 19 000 patients and a prospective multicentre design.
- Pragmatic study protocol as all vital sign measurements were conducted by the ward nurses themselves as part of the normal clinical routines.
- Proportion of missing data is minor with exclusion percentage being only 0.5%.
- We do not have detailed patient demographics on comorbidities and specific admission diagnoses.
- The study hospitals had different response protocols for patient deterioration.

Trial registration number NCT04055350.

INTRODUCTION

Early warning scores (EWSs) are used to identify patients at risk of deterioration in hospitals; this practice is also endorsed in the recently updated European Resuscitation Council Guidelines on advanced life support. EWSs, such as the widely adopted national early warning score (NEWS), are based on vital signs that are scored from 0 to 3 depending on how gravely they deviate from the ‘normal’ range. However, a recent systematic review raised concerns on the quality of trials validating the EWSs. It found that most EWSs have only been validated in small or retrospect datasets including preselected patient cohorts and insufficient numbers of event patients. The NEWS was originally validated using the same medical assessment unit patient cohort that was used in the development phase, although the cohort was large and the outcome measure within a meaningful time horizon.

© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to Dr Eetu Loisa; eetu.loisa@tuni.fi
EWS vital sign components have been considered equal, as in using plain statistical percentiles to define score categories for each vital sign, for example. However, some studies suggest that the respiratory parameters may be more strongly associated with morbid patient outcomes. This phenomenon, however, has not been thoroughly investigated among mixed general ward patients.

The objective of this study was to validate the ability of the NEWS to discriminate between short-term survivors and non-survivors among heterogeneous general ward patient population in a large, prospective and pragmatic, three-centre trial. We further hypothesised that the respiratory components of the NEWS may be more relevant to patients’ prognosis as compared with the haemodynamic components.

METHODS
Study design and setting
We conducted a 1-year pragmatic, prospective observational study in the Pirkanmaa Hospital District's three hospitals in Finland: Tampere University Hospital (Tays), and two regional hospitals Valkeakoski Hospital (VALS) and Hatanpää Hospital (HASA). Together these hospitals provide ward care for patients from all medical and surgical specialties; psychiatric care is provided in a separate hospital. The Pirkanmaa Hospital District provides hospital services for 530,000 citizens and the most advanced care (tertiary level care) for a catchment population of 900,000 citizens.

All three hospitals use the NEWS for the follow-up of patients' vital signs. The NEWS tool is presented in online supplemental file A. The haemodynamic and respiratory subcomponents of the score may be scored from 0 to 3, 3 presenting the extreme deviation from the ‘normal’ range. The response for patient deterioration is tailored according to each hospitals’ resources available 24/7. Tays and HASA have medical emergency teams that attend both medical emergencies and cardiac arrests, while VALS has a dedicated response team operating from the emergency department.

Since 2016, the Pirkanmaa Hospital District’s hospitals begun to implement the Medanets mobile solution system for clinical nursing to emergency departments and general wards. The mobile solution app enables nurses to record all clinical measurements bedside; the system automatically calculates patient’s current NEWS, shows preceding NEWS and trends, and records all the measurements and NEWS to the hospital’s electronic vital signs datasheet. The mobile system requires all vital signs to be measured. The frequency of the measurements depends on the patients’ condition that are labelled from ‘green’ to ‘red’ according to the NEWS. For example, the minimum vital signs measurement frequency for the ‘green’ patients is every 12 hours. The hospitals’ general guidelines recommend that the first set of vital signs are controlled immediately on patient’s arrival to the ward (always within the first hour). The mobile system itself does not alert hospital’s emergency teams or treating physicians; the purpose of the system is to facilitate bedside nursing work and standardise patient follow-up.

The Royal College of Physicians implemented the NEWS2 in 2017. The NEWS2 has a modified blood oxygen saturation scale for those patients with confirmed hypercapnic respiratory failure. However, this updated NEWS2 has not yet shown any benefits as compared with the original NEWS and raised concerns of its feasibility altogether. Therefore, the Pirkanmaa Hospital District hospitals use the original NEWS.

Participants
All adult patients (≥18 years) admitted to the 45 surgical or medical general wards of the three hospitals that had implemented the Medanets mobile solution system at least 4 months before the study period (1 January 2019–31 December 2019) begun. These somatic wards have approximately 800 beds. Thus, emergency departments, intensive care wards, high dependency units, operation rooms and postanaesthetic units and general wards that had not implemented/were still in the implementation phase with the mobile solution system were excluded. Patients with missing data on any of the vital signs were excluded. Each patient’s first general ward vital signs—and thus NEWS—dataset was included and association with subsequent mortality was studied.

Data collection
The pragmatic design means that all vital signs measurements were conducted as part of nurses’ normal clinical routines. For the study purposes, the recordings were also prospectively collected to a separate database in the Pirkanmaa Hospital District’s internal secure data server. All the datasets were time-labelled and included patients’ social security numbers. This enabled the recording of age, gender and mortality from the patients’ medical records. As stated before, only the first recorded NEWS dataset at the beginning of the admission to general ward was included per patient in the analysis phase. Patients were followed up to 3 months after their first vital signs recordings on general wards as mortality data for those patients that die after hospital discharge may take 30–60 days to be updated to the system. The Strengthening the Reporting of Observational Studies in Epidemiology checklist for observational studies was followed.

Patient and public involvement
This study was done without patient or public involvement. Patients were not involved in the study design, data collection, interpretation of the results or writing or editing the manuscript.

Outcomes
The main outcome measure was 1-day mortality. The secondary outcome was 2-day and 30-day mortality rates.
Statistical analysis
Statistical programming was done with Python V.3.8.6 and R V.3.6.3. Data are represented as counts and percentages and continuous data as medians and percentiles. Area under receiver operating characteristic curves (AUROCs) were determined for the first measured NEWS and the haemodynamic and respiratory subcomponents that can be scored between 0 and 3 (respiratory rate, peripheral blood oxygen saturation, heart rate and systolic blood pressure). ROC curves were compared using Delong’s test for two correlated ROC curves. All tests were two-sided, 95% CIs were used and a p value of <0.05 was considered statistically significant.

RESULTS
Study cohort
During the study period there were 182 223 datasets. After excluding measurements outside the 45 included general wards and repeat measurements, there were 19 001 patients fulfilling the inclusion criteria. However, in 102 cases data on either level of consciousness or oxygen supplementation were missing (0.53%). Thus, the final cohort included 18 899 patients (figure 1).

Patient demographics
The median age of the patients was 70 (57–80) years (table 1). Half (51%) of the patients were male, four-fifths (81%) were admitted to the university hospital and two-thirds (69%) had a medical reason for admission.

Patient outcomes
Table 1 presents patient outcomes for the study cohort. The 1-day mortality rate was 0.36%, 2-day mortality was 0.61% and 30-day mortality was 3.9%.

Discriminative value of the NEWS
AUROC values for the NEWS to discriminate mortality at three different end points are presented in table 2 and figure 2. The NEWS had excellent discriminative performance for 1-day and 2-day mortality. The discriminative performance for long-term mortality decreased substantially. The discriminative performance of the NEWS did not differ between the university-level hospital and the regional hospitals.

Discriminative value of the NEWS’s haemodynamic and respiratory components
AUROC values for the NEWS’s haemodynamic and respiratory subcomponents are presented in table 2 (and as a figure for 1-day mortality in online supplemental file B). The respiratory rate component discriminated 1-day survivors better as compared with the blood oxygen saturation component (p=0.005), heart rate component (p=0.007) and systolic blood pressure component (p=0.002). There were no differences in the discriminative performance between the blood oxygen saturation component, heart rate component and systolic blood pressure component (comprehensive data on the Delong’s test results are presented in online supplemental file C). Figures 3 and 4 present the NEWS’s subcomponents score distributions for patients who died at 1 day and 30 days after the measurements. The figures read from left to right, presenting the percentage (and thus distribution) of different scores (0, 1, 2 or 3) for those patients that died at 1 day (figure 3) and within 30 days (figure 4). The respiratory components were more frequently scored as

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Patient characteristics, admission vital signs and outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>70 (57, 80)</td>
</tr>
<tr>
<td>Sex (male)</td>
<td>9623 (51)</td>
</tr>
<tr>
<td>Hospital level (tertiary centre)</td>
<td>15323 (81)</td>
</tr>
<tr>
<td>Admission ward (medical)</td>
<td>13064 (69)</td>
</tr>
<tr>
<td><strong>First measured general ward vital signs and NEWS</strong></td>
<td></td>
</tr>
<tr>
<td>Respiratory rate (breaths/min)</td>
<td>16 (15, 19)</td>
</tr>
<tr>
<td>Oxygen saturation (%)</td>
<td>96 (95, 98)</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>36.8 (36.5, 37.2)</td>
</tr>
<tr>
<td>Systolic blood pressure (mm Hg)</td>
<td>134 (119, 151)</td>
</tr>
<tr>
<td>Heart rate (beats/min)</td>
<td>74 (64, 85)</td>
</tr>
<tr>
<td>Level of consciousness (normal)</td>
<td>18524 (98)</td>
</tr>
<tr>
<td>Need for supplementary oxygen</td>
<td>3719 (20)</td>
</tr>
<tr>
<td>NEWS</td>
<td>1.0 (0.0, 3.0)</td>
</tr>
<tr>
<td><strong>Outcome</strong></td>
<td></td>
</tr>
<tr>
<td>1-day mortality</td>
<td>68 (0.36)</td>
</tr>
<tr>
<td>2-day mortality</td>
<td>115 (0.61)</td>
</tr>
<tr>
<td>30-day mortality</td>
<td>746 (3.9)</td>
</tr>
</tbody>
</table>

Data are represented as counts with percentages and continuous data as medians with percentiles (Q1, Q3).
extremely deviating (score 3) among the patients who died as compared with the haemodynamic components.

**DISCUSSION**

**Key findings**

This pragmatic three-centre trial with nearly 19000 general ward patients validates the NEWS’s ability to discriminate ward patients at a risk of acute death with excellent accuracy. Out of NEWS’s individual subcomponents, the respiratory rate component predicted short-term mortality better as compared with the blood oxygen saturation and the haemodynamic subcomponents.

**External validation of the NEWS**

Gerry et al found in their recent systematic review that the external validation of the NEWS has been mostly conducted in small and/or preselected patient populations, while the primary validation studies were conducted with retrospective, existing data. A further existing limitation is that many NEWS studies have been conducted among emergency department patients, although the NEWS is widely endorsed for routine patient follow-up among hospitalised general ward patients.

Previous studies have demonstrated the NEWS’s association with short-term morbidity or mortality and reported slightly lower, similar discriminative performance as compared with our results (AUROCs ranging from 0.83 to 0.91). Our results externally validate the NEWS as a feasible patient safety strategy on general wards. An AUROC of >0.9 for a patient-centred outcome that is within a meaningful timeframe (1-day mortality) with a feasible, completely non-invasive and practically cost-free test is unquestionably worth implementing to everyday practice in hospitals.

Some previous studies have investigated the NEWS’s association with mid-term to long-term prognosis.

**Table 2** AUC values for NEWS and NEWS’s subcomponents for 1-day, 2-day and 30-day mortality

<table>
<thead>
<tr>
<th></th>
<th>1-day mortality</th>
<th>2-day mortality</th>
<th>30-day mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC</td>
<td>95% CI</td>
<td>AUC</td>
</tr>
<tr>
<td>NEWS</td>
<td>0.91 (0.87 to 0.95)</td>
<td>0.89 (0.85 to 0.92)</td>
<td>0.75 (0.73 to 0.77)</td>
</tr>
<tr>
<td>NEWS SBP component</td>
<td>0.65 (0.59 to 0.72)</td>
<td>0.62 (0.57 to 0.66)</td>
<td>0.55 (0.53 to 0.56)</td>
</tr>
<tr>
<td>NEWS HR component</td>
<td>0.67 (0.61 to 0.74)</td>
<td>0.65 (0.60 to 0.70)</td>
<td>0.58 (0.56 to 0.59)</td>
</tr>
<tr>
<td>NEWS RR component</td>
<td>0.78 (0.72 to 0.84)</td>
<td>0.75 (0.70 to 0.79)</td>
<td>0.62 (0.61 to 0.64)</td>
</tr>
<tr>
<td>NEWS SpO2 component</td>
<td>0.66 (0.59 to 0.73)</td>
<td>0.68 (0.63 to 0.73)</td>
<td>0.61 (0.59 to 0.63)</td>
</tr>
</tbody>
</table>

AUC, area under the curve; HR, heart rate; NEWS, National Early Warning Score; RR, respiratory rate; SBP, systolic blood pressure; SpO2, oxygen saturation.

**Figure 2** ROC curves for NEWS and 1-day, 2-day and 30-day mortality. AUC, area under the curve; NEWS, National Early Warning Score; ROC, receiver operator characteristic.

**Figure 3** The NEWS’s subscore distributions for patients who died within 1 day. Read from left (percentage of patients that scored 0) to right (percentage that scored 3). NEWS, National Early Warning Score.
with discriminative values of the NEWS in identifying 30-day mortality comparable to that of our study (AUROCs ranging from 0.71 to 0.86\(^{16, 22, 23}\)). However, as we demonstrate with both the NEWS and its haemodynamic and respiratory subcomponents, the discriminative value of a physiology-based assessment decreases substantially in predicting long-term mortality. Indeed, these outcomes are more dependent on other factors such as patient’s age, gender and comorbidity.\(^{24}\) Gerry et al found that 52% of the EWS external validation studies included long-term time horizons, such as in-hospital and/or 30-day mortality, as outcome measures.\(^4\) In the context of using the NEWS in hospital floors to identify patient deterioration in time, these end points seem irrelevant.

**Respiratory and haemodynamic components of the NEWS**

To no surprise, we found that the predictive value the NEWS’s respiratory and haemodynamic subcomponents were significantly lower as compared with the total NEWS. Previous studies comparing single parameter systems have found that any individual with altered vital sign alone owns poor sensitivity to identify patients at a risk of death as compared with the comprehensive clinical assessment of all vital signs.\(^{25, 26}\)

When comparing the respiratory and haemodynamic subcomponents of the NEWS between each other, however, we found that the respiratory rate component discriminated patients at a risk of acute death better as compared with the blood oxygen saturation component and the haemodynamic components. It was also more frequently scored as extremely deviating as compared with the other subcomponents among those patients that died within 1 day and 30 days after the measurements. The respiratory rate has previously been found as one of the key parameters in predicting patient at a risk of acute death.\(^8, 9, 27–29\)

In some subcohorts, such as patients with COVID-19 and patients reviewed by hospitals’ rapid response teams, abnormal haemodynamic parameters have not been associated with morbidity at all or they have represented in late phase of deterioration, whereas abnormal respiratory parameters have reliably discriminated patients at risk.\(^7, 10\) In general, studies across a variety of different settings have somewhat consistently found that of individual vital signs, the respiratory rate, the blood pressure and the level of consciousness are more strongly associated with mortality as compared with the heart rate and the blood oxygen saturation.\(^26, 30–32\) However, in most studies comparing individual vital signs, arbitrary dichotomised thresholds have defined the individual vital signs as normal/abnormal before statistical comparisons.\(^26, 30–32\) Differences and discrepancies between the significance of individual vital signs may further indicate that the prognostic value of the individual vital signs are time dependent, as in some vital signs may be early indicators of deterioration while other vital signs may be late indicators of deterioration. Here, we investigated the respiratory and haemodynamic parameters scored from 0 to 3 according to the NEWS and reported the results for both 1-day and 2-day mortalities. Thus, our results provide a different point of view for the individual vital signs and further underline the importance of the repeated respiratory rate measurements on general wards.

**Strengths and limitations**

The internal validity of this trial is strengthened by the following factors: a prospective design, a study population of nearly 19,000 patients with no vital signs data missing and the exclusion of patients with any missing data (exclusion percentage being only 0.5%). The internal validity is limited by the facts that we were unable to identify patients with do-not-resuscitate orders and we were not able to capture specific patient demographic data, such as comorbidities. While our guidelines expect that the first set of vital signs are measured within the first hour (preferably immediately) when a patient arrives to the ward, we do not know whether delays existed in some cases. We are further unable to comment whether high NEWS values were acted on and some patients salvaged by the hospitals’ emergency response teams. However, after a decade of endorsement on appropriate vital signs monitoring and rapid response teams, it is practically impossible to find hospitals that have ignored these patient safety aspects.

The strengths related to external validity include the three-centre platform including both university-level and local hospitals, a highly heterogeneous
CONCLUSION
The NEWS discriminates ward patients at a risk of acute death with excellent accuracy. Out of NEWS’s haemodynamic and respiratory subcomponents, the respiratory rate predicts short-term mortality better as compared with the blood oxygen saturation and the haemodynamic components.

REFERENCES
18. Mitsunaga T, Hasegawa I, Uzura M, 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:e6947.


