ORIGINAL ARTICLE

A validation of the National Early Warning Score to predict outcome in patients with COPD exacerbation

Luke E Hodgson,[1,2] Borislav D Dimitrov,[1] Jo Congleton,[3] Richard Venn,[2] Lui G Forni,[4] Paul J Roderick[1]

[1]Academic Unit of Primary Care and Population Sciences, Faculty of Medicine, University of Southampton, Southampton General Hospital, Southampton, UK
[2]Anaesthetics Department, Western Sussex Hospitals NHS Foundation Trust, Worthing Hospital, Worthing, UK
[3]Royal Sussex County Hospital, Brighton & Sussex Hospitals NHS Trust, Brighton, UK
[4]The Royal Surrey County Hospital NHS Foundation Trust, Guildford, UK

Correspondence to
Dr Luke E Hodgson, Academic Unit of Primary Care and Population Sciences, Faculty of Medicine, University of Southampton, Southampton General Hospital, Tremona Road, Southampton SO16 6YD, UK; drlhodgson@gmail.com

► http://dx.doi.org/10.1136/thoraxjnl-2016-208820
► http://dx.doi.org/10.1136/thoraxjnl-2016-209212

ABSTRACT

Background The National Early Warning Score (NEWS), proposed as a standardised track and trigger system, may perform less well in acute exacerbation of COPD (AECOPD). This study externally validated NEWS and modifications (Chronic Respiratory Early Warning Score (CREWS) and Salford-NEWS) in AECOPD.

Methods An observational cohort study (2012–2014, two UK acute medical units (AMUs)), compared AECOPD (2361 admissions, 942 individuals, International Statistical Classification of Diseases and Related Health Problems-10 J40–J44 codes) with AMU patients (37 109 admissions, 20 415 individuals).

Outcome In-hospital mortality prediction was done by admission NEWS, CREWS and Salford-NEWS assessed by discrimination (area under receiver operating characteristic curves (AUROCs)) and calibration (plots and Hosmer-Lemeshow (H-L) goodness-of-fit).

Results Median admission NEWS in AECOPD was 4 (IQR 2–6) versus 1 (0–3) in AMUs (p≤0.001), despite mortality of 4.5% in both. AECOPD AUROCs were NEWS 0.74 (95% CI 0.66 to 0.82), CREWS 0.72 (0.63 to 0.80) and Salford-NEWS 0.62 (0.53 to 0.70). AMU NEWS AUROC was 0.77 (0.75 to 0.78). At threshold NEWS=5 for AECOPD (44% of admissions), positive predictive value (PPV) of death was 8% (5 to 11) and negative predictive value (NPV) was 98% (97 to 99) versus AMU patients PPV of 17% (16 to 19) and NPV of 97% (97 to 97). For NEWS in AECOPD H-L p value=0.202.

Conclusion This first validation of the NEWS in AECOPD found modest discrimination to predict mortality. Lower specificity of NEWS in patients with AECOPD versus other AMU patients reflects acute and chronic respiratory physiological disturbance (including hypoxia), with resultant low PPV at NEWS=5. CREWS and Salford-NEWS, adjusting for chronic hypoxia, increased the specificity and PPV but there was no gain in discrimination.

Key messages

What is the key question?
► How does the National Early Warning Score (NEWS) perform in predicting mortality for acute exacerbation of COPD (AECOPD)?

What is the bottom line?
► The NEWS shows acceptable discrimination in AECOPD and at a cut-off of 5 points has high negative predictive value and low positive predictive value and therefore can predict survival but not mortality.

Why read on?
► This large dual-centre cohort study is the first validation of the NEWS in AECOPD.

INTRODUCTION

Hospitalised patients are frequently exposed to avoidable harm.[1] Adverse trends in clinical observations are often missed or misinterpreted, while shortcomings in management reflect poor organisation, appreciation of urgency or lack of supervision.[1][2] One response to this knowledge has been the introduction of early warning scores (EWSs). Incorporating observations, with points aggregated in a weighted manner, depending on the degree of abnormality, EWSs are examples of prognostic prediction models. Derivation studies of EWSs, usually predicting inpatient mortality, have used observations at admission[3] or at 24 hours prior to the outcome.[4] By 2008, over 30 EWSs for acute hospital admissions had been published, with variable implementation.

The Royal College of Physicians (RCP) published a National Early Warning Score (NEWS) in 2012,[5] aiming to standardise practice (see online supplementary figure A1 for further details). The NEWS is based on the validated ViEWS score derived at a single UK hospital (35 585 general acute medical unit (AMU) admissions, median age 73, mortality 5.6%),[4] reporting an area under the receiver operating characteristic curve (AUROC) of 0.89 for predicting in-hospital mortality within 24 hours of the observation set.[4] The RCP report suggested a range of thresholds, to standardise the frequency of observations, responder personnel and timing of senior review. For example, at a NEWS of 5 points, monitoring is suggested at least hourly, along with an urgent clinical review. The report noted that patients with COPD may have chronically disturbed physiology, potentially altering NEWS performance, but did not quantify this or suggest a way this should be compensated for. Two groups have proposed adjusted NEWS for patients with chronic respiratory disease, addressing this issue.[6][7] First, the Chronic Respiratory Early Warning Score (CREWS)[6] assigns risk points at a lower oxygen saturation threshold compared with patients without COPD. Second, Salford-NEWS combines this

lower threshold with risk points for the use of supplemental oxygen in the context of higher oxygen saturation levels, reflecting a concern of hyperoxia-induced hypercapnic respiratory failure[7] (see online supplementary figure A2).

Over 900 000 people have been diagnosed with COPD in the UK[8] and around 110 000 emergency hospital admissions in England each year are due to acute exacerbation of COPD (AECOPD), a frequency second only to pneumonia.[9] Concern has been expressed that the NEWS lacks specificity in AECOPD, over-alerting relatively stable patients, especially due to the weighting of 'chronic' hypoxia, with the potential for inappropriate diversion of resources and potentially encouraging (inappropriate) oxygen therapy.[6 7] The study, performed in two adult AMUs, has two aims:

1. To externally validated the performance of the NEWS in terms of inpatient death among:
   a. patients admitted to hospital with an AECOPD ('AECOPD cohort') for their first admission during the study period;
   b. unselected patients admitted to hospital ('AMU cohort') for their first admission during the study period;
   c. the AECOPD cohort for all inpatient episodes during the study period and
   d. the AMU cohort for all inpatient episodes during the study period.
2. To externally validate alternative EWSs (CREWS and Salford-NEWS).[6 7] Clauses 1a and 1b represent the primary analysis and clauses 1c and 1d form the sensitivity analysis. Published guidance for reporting was followed.[10]

METHODS

An observational retrospective cohort external validation study of the NEWS was performed in the adult AMUs of Worthing Hospital and St Richard's Hospital sites of Western Sussex Hospitals NHS Foundation Trust (WSHFT), for the period from March 2012 to February 2014. WSHFT is an 870-bed Trust on the South Coast of England with a combined annual emergency department attendance over 150 000 and 50–60 acute medical admissions per 24-hour period. There is a separate admissions unit for complex elderly patients on the Worthing site (not included in the analysis). Ethical approval was given by NHS Research Ethics Committee London—South East (REC reference 13/LO/0884).

At admission, all inpatients have physiological observations measured and entered via handheld systems into the clinical data software system (Patientrack Sydney, New South Wales, Australia), with the NEWS automatically calculated. Criteria for the AECOPD cohort were as follows: age over 40 years, admitted to one of the AMUs, staying for at least one night over the 24-month period (2012–2014), as identified by a primary diagnosis from the International Statistical Classification of Diseases and Related Health Problems-10 (ICD-10) classification J40–J44 (88% coded J44.0 or J44.1).[11] For comparison, over the same period, data were extracted on all other patients aged ≥18 years admitted for at least one night, through the two AMUs (AMU cohort). Exclusion criteria were as follows: patients moved directly from the accident and emergency (A&E) department to the Intensive Care Unit (ICU) (as neither area uses the Patientrack data system), aged <18 years or those discharged without spending a night in hospital.

Patients were followed up until discharge from hospital, or death, during the 24 months observation period. The primary analysis was performed for the first inpatient admission. A sensitivity analysis was performed, including all episodes during the study period, in an attempt to account for countervailing prognostic factors such as survivor bias and the effects of repeated admissions. Analysing all episodes also aids generalisability of results; as in clinical practice the NEWS is applied whatever the number of prior admissions. Furthermore, the data were analysed 'per admission' by creating a multilevel (hierarchical) multiple logistic regression model in which we adjusted the main (fixed) effect of the NEWS score for the number of admissions (level 1) per patient (level 2). 25 second-level clusters or "types of COPD patients" were formed. The main effect of NEWS score as predictor was adjusted for its random slope at level 1 (as nested within 25 clusters at level 2) and the number of admissions as a random intercept. Finally, a multiple regression model was run to test whether any difference seen in NEWS between AECOPD and AMU cohorts might have been due to the older age of the AECOPD cohort. As all patients had a NEWS calculated automatically by the Patientrack system before any outcome had occurred, there were no missing data at admission. The outcome predicted by admission NEWS was inpatient mortality. None of the researchers involved in analysis of the data were involved in the management of the patients. Data for the CREWS and Salford-NEWS prediction scores were collected and elaborated in the same way. The research team members responsible for data analysis had access only to the fully anonymised individual-level data and were blinded to any other patient data, as well as to the components of the calculated scores in the hospital information system. Since there is no consensus on how to determine what counts as an adequate sample size in such studies,[10] all available 39 470 hospital episodes for the period 2012–2014 were included in the analysis.

Performance of the score as predictor is assessed by discrimination and calibration.[10] Discrimination is demonstrated by the AUROC of the receiver operating characteristic curve, representing how well a model separates patients who experienced the outcome (in this case mortality) from those who did not. Calibration describes how well predicted results from a logistic regression model agree with the observed results. Over the entire range of prediction, this is referred to as goodness-of-fit. The Hosmer-Lemeshow (H-L) test is the most commonly used statistics in this field.[12] The H-L test-associated p value, when significant (<0.05), may indicate poor fit.[12 13] It is also recommended to graphically plot predicted against observed outcomes, for example, with a calibration slope.[10] The agreement between the predicted probabilities and the observed frequencies for calibration was evaluated graphically by plotting the predicted probabilities (x-axis) by the observed event rate (y-axis) of the outcome (at each level of the score). The association between predicted probabilities and observed event rate can be described by a line with an intercept and a slope. An intercept of zero and a slope of one indicate perfect calibration. Predictive values were also calculated at suggested NEWS call-out thresholds, to further inform on the way model performance could impact on clinical workload. Following extraction, all data were fully anonymised on Microsoft Excel and analyses were performed on SPSS (V.22) and STATA SE (V.14).

RESULTS

Over the 2-year study period, there were 2361 AECOPD inpatient episodes (123 inpatient deaths, median of 3 admissions (IQR 2–5)) and 37 109 non-COPD AMU episodes (1911 deaths). For the primary analysis (first admission), there were 942 patients in the AECOPD cohort and 20 415 patients in the AMU cohort. The AECOPD cohort had a median age of 74 (67–82) versus 71 (55–82) in the AMU cohort (p<0.001).

Median admission NEWS was significantly different—AECOPD 4 points (2–6) versus AMU 1 point (0–3) (p≤0.001). Inpatient mortality for first admission did not differ (4.5% in both cohorts). Table 1 summarises admission clinical demographic variables.

The spread of scores for the AECOPD cohort can be seen to be bell-shaped, in contrast with the AMU cohort, where the data are right-skewed (figure 1). In the AECOPD cohort, 44% had a score of ≥5 points on admission, compared with only 11% in the AMU cohort. Using a NEWS threshold of 5 points, to predict inpatient mortality in the AECOPD cohort, sensitivity was 76% (95% CI 61% to 88%) specificity was 57% (54% to 61%), positive predictive value (PPV) was 8% (5% to 11%) and negative predictive value (NPV) was 98% (97% to 99%). In contrast, in the AMU cohort, sensitivity was 43% (40% to 46%), specificity was 90% (90% to 91%), PPV was 17% (16% to 19%) and NPV was 97% (97% to 97%). (See table 2, which includes threshold of NEWS of 7 points.)

In the AECOPD cohort, for their first admission, the AUROCs for predicting inpatient mortality for the three prediction scores were as follows: NEWS=0.74 (95% CI 0.66 to 0.82), CREWS 0.72 (0.63 to 0.80) and Salford-NEWS 0.62 (0.53 to 0.70). In the AMU cohort, for their first admission, the AUROC for the NEWS was 0.77 (0.75 to 0.78) (figure 2). In the AECOPD cohort, the H-L test p value was 0.202 for NEWS, 0.399 for CREWS and 0.08 for Salford-NEWS (see online supplementary table A1 for H-L observed and expected results for NEWS). Calibration plots (shown in figure 3) suggest no improvement in calibration with the alternative scores, which both underpredicted mortality, though the number of deaths was small. In lower-risk groups, NEWS in the AMU cohort also underpredicted observed mortality. By assigning less points for hypoxaemia, CREWS and Salford-NEWS increased specificity, with an accompanying decrease in sensitivity in the AECOPD cohort. For example, at a call-out threshold of 5 points, sensitivity (to predict mortality) was 76% for NEWS, 48% for CREWS and 24% for Salford-NEWS, respectively, while specificity was 57%, 88% and 91%. CREWS and Salford-NEWS in the AECOPD cohort performed similarly to the NEWS in the AMU cohort.

In the AECOPD cohort, for all inpatient episode over the 2-year study period (n=2361), AUROCs for predicting inpatient mortality for the three prediction scores were as follows:

NEWS=0.69 (0.64 to 0.75), CREWS 0.70 (0.64 to 0.75) and Salford-NEWS 0.67 (0.61 to 0.72). In the AMU cohort, for all inpatient episodes (n=37 109) using NEWS the AUROC was 0.75 (0.74 to 0.76) (see online supplementary table A2 and figures A3–4 for further details on all episodes). After adjusting for the number of admissions per patient (using 'admissions' as a random intercept), all p values for the main effect remained significant at p<0.05. After further adjusting the main effect (admission NEWS), as a random slope (at first level) as nested within the number of admissions (at second level), the main (fixed) effect of the NEWS as a predictor remained significant at p<0.05 (see online supplementary file). A further multiple regression model revealed that both AECOPD and age were independent predictors of NEWS, suggesting that the increased age in the AECOPD cohort did not account for the increase in NEWS seen in this cohort.

DISCUSSION
Statement of principal findings
This is the first validation study of the NEWS, CREWS and Salford-NEWS in AECOPD admissions. To predict inpatient mortality, admission NEWS in a AECOPD cohort demonstrated similar discrimination to an AMU cohort (AUROC 0.74 (66 to 82) vs 0.77 (75 to 78)). However, at suggested RCP cut-offs of 5 and 7 points (to predict mortality) in the AECOPD cohort, specificity and PPV values of the NEWS were lower compared with the AMU cohort, though sensitivity at the same cut-offs was higher. Modified scores have been suggested to account for chronically altered physiology in AECOPD.[6 7] However, this goes against the premise that a universal scoring system (with potential significant advantages) should be employed throughout NHS hospitals. Furthermore, patients with COPD were included in the original derivation cohort for the NEWS. Assigning lower oxygen saturation thresholds for scoring could result in patients at high risk of death being categorised into a lower-risk group, thereby missing opportunities to intervene early. As predictors of mortality on admission (assessed by respective AUROCs), CREWS (0.72) and Salford-NEWS (0.66)
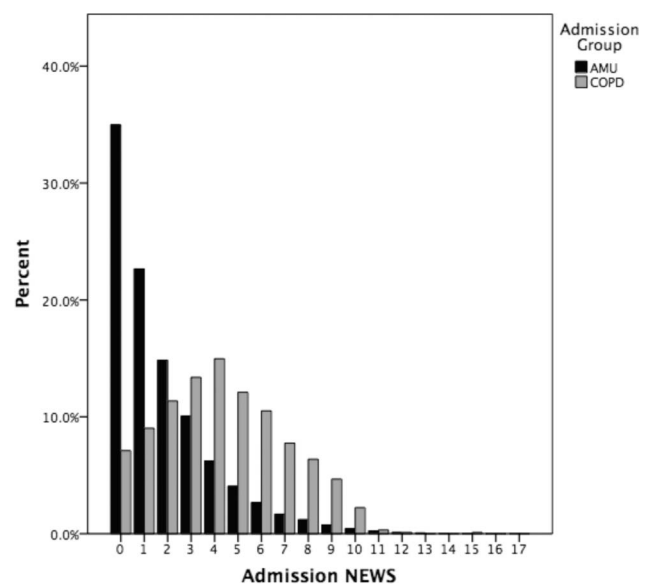
Table 1 Demographics and outcomes of AECOPD and AMU cohorts (first admission during the study period and all episodes)

First admission	COPD (n=942)	AMU (n=20 415)	p Value
Age	74 (67–82)	71 (55–82)	<0.001*
Inpatient mortality	4.5% (n=42)	4.5% (n=911)	0.967†
Admission NEWS	4 (2–6)	1 (0–3)	<0.001*
All episodes	**COPD (n=2361)**	**AMU (n=37 109)**	
Age	74 (67–82)	73 (57–83)	<0.001*
Inpatient mortality	5.2% (n=123)	5.1% (n=1911)	0.47†
Admission NEWS	4 (2–6)	1 (0–3)	<0.001

Median values (IQR).
*Mann-Whitney U test.
†χ² test.
AECOPD, acute exacerbation of COPD; AMU, acute medical unit; NEWS, National Early Warning Score.

Figure 1 Spread of National Early Warning Score (NEWS) on admission for acute exacerbation of COPD (grey n=942 patients) and acute medical unit (AMU) patients (black n=20 415 patients) as percentage of total.

Table 2 Prediction of inpatient mortality by admission score

Call-out	Group and score	Sensitivity	Specificity	PPV	NPV
Score ≥5	COPD				
	NEWS	76% (61 to 88)	57% (54 to 61)	8% (5 to 11)	98% (97 to 99)
	CREWS	48% (32 to 64)	88% (85 to 90)	15% (10 to 23)	97% (96 to 98)
	Salford-NEWS	24% (12 to 39)	91% (89 to 93)	11% (5 to 19)	96% (95 to 97)
	AMU NEWS	43% (40 to 46)	90% (90 to 91)	17% (16 to 19)	97% (97 to 97)
Score ≥7	COPD				
	NEWS	60% (43 to 74)	80% (77 to 83)	12% (8 to 18)	98% (96 to 99)
	CREWS	13% (6 to 23)	96% (95 to 97)	21% (10 to 37)	93% (91 to 95)
	Salford-NEWS	14% (5 to 29)	95% (94 to 97)	12% (5 to 25)	96% (95 to 97)
	AMU NEWS	25% (23 to 28)	96% (96 to 97)	25% (22 to 28)	96% (96 to 97)

Sensitivity, specificity, PPV, NPV (with 95% CIs) at RCP suggested NEWS thresholds of 5 points and 7 points for first AECOPD admission (n=942) using NEWS, CREWS, Salford-NEWS scores and the NEWS for the AMU cohort (n=20 415).
AECOPD, acute exacerbation of COPD; AMU, acute medical unit; CREWS, Chronic Respiratory Early Warning Score; NEWS, National Early Warning Score; NPV, negative predictive value; PPV, positive predictive value; RCP, Royal College of Physicians.

Figure 2 (A) Area under the receiver operating curves (AUROC) for three prediction scores (National Early Warning Score (NEWS)=blue curve, Chronic Respiratory Early Warning Score (CREWS)=green, Salford-NEWS=yellow, reference line=black) on admission to predict inpatient mortality in patients with acute COPD exacerbation (first admission during study period, n=942). NEWS 0.74 (95% CI 0.66 to 0.82), CREWS 0.72 (0.63 to 0.80) and Salford-NEWS 0.62 (0.53 to 0.70). (B) AUROC for NEWS (blue line, reference line=black) to predict inpatient mortality in the acute medical unit cohort (first admission during study period, n=20 415)—0.77 (0.75 to 0.78).
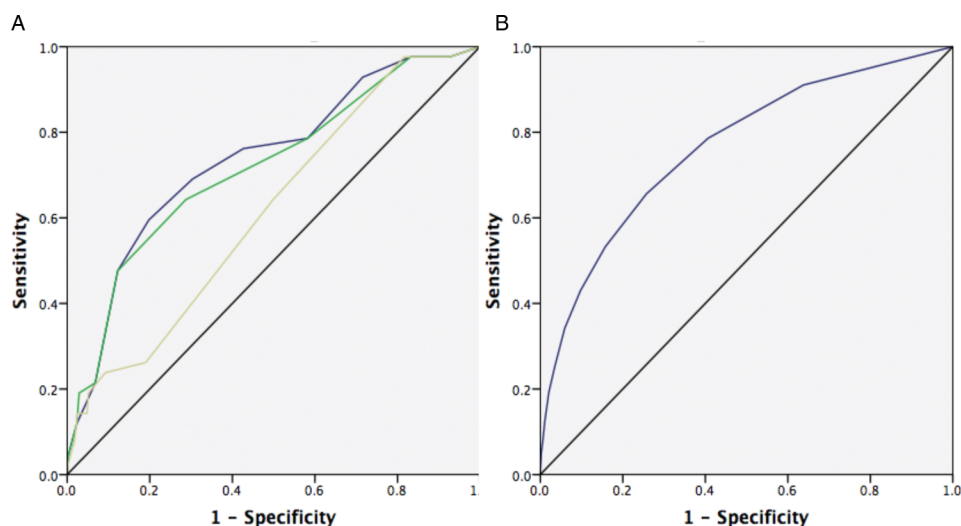
did not improve discrimination compared with NEWS (0.74). At a threshold of 5 points, both alternatives improved specificity and PPV though sensitivity was reduced.

In a large dual-centre adult AMU cohort, to predict inpatient mortality, admission NEWS discriminated satisfactorily (AUROC 0.77). A trade-off between sensitivity and specificity must be noted, for example, at a cut-off NEWS of 7 points; sensitivity was only 25% for inpatient mortality. The AUROC for the NEWS is similar to a previously described admission prediction model by Duckitt *et al* (AUROC 0.74).[3] The higher AUROC in the original derivation study for NEWS (0.89) is explained by prediction time frame (mortality 24 hours from observation set versus admission score),[4] and derivation studies usually perform better than validation studies, making the later crucial to perform.[10] Admission NEWS was analysed here, as it can be used to triage the patient and facilitate early physician review in higher-risk patients. This study complements three external validations of the ViEWS, on which the NEWS is based. One Canadian study found that an abbreviated ViEWS gave an AUROC of 0.81 (0.80–0.82) to predict 30-day mortality;[14] a US study reported an AUROC of 0.86 (timing of outcome not reported)[15] and a Ugandan study reported an AUROC of 0.89 (0.83 to 0.95) to predict mortality within

24 hours of admission.[16] Overall mortality was lower than in the original ViEWS derivation study, reflecting a decrease across the NHS for emergency admissions, previously reported.[17 18]

Strengths of the study

This large observational cohort validation study using automatic (electronic) collection of observation information provides novel insights into the performance of the NEWS, CREWS and Salford-NEWS in a specialist group of patients, who nevertheless account for a large number of acute hospital admissions. Population characteristics, inpatient mortality, duration of stay and number of readmissions closely resemble the most recent UK British Thoracic Society (BTS) Audit.[19] The study also provides the first UK external validation of the NEWS on admission. The additional analyses, adjusting for number of admissions, using multilevel modelling strengthen the findings of the primary analysis that NEWS is a strong predictor of mortality in this group.

Limitations of the study

The study relied on the ICD-10 coding of diagnosis, which has potential shortcomings,[20] and using a primary diagnosis could have missed or misdiagnosed episodes. As there is no single
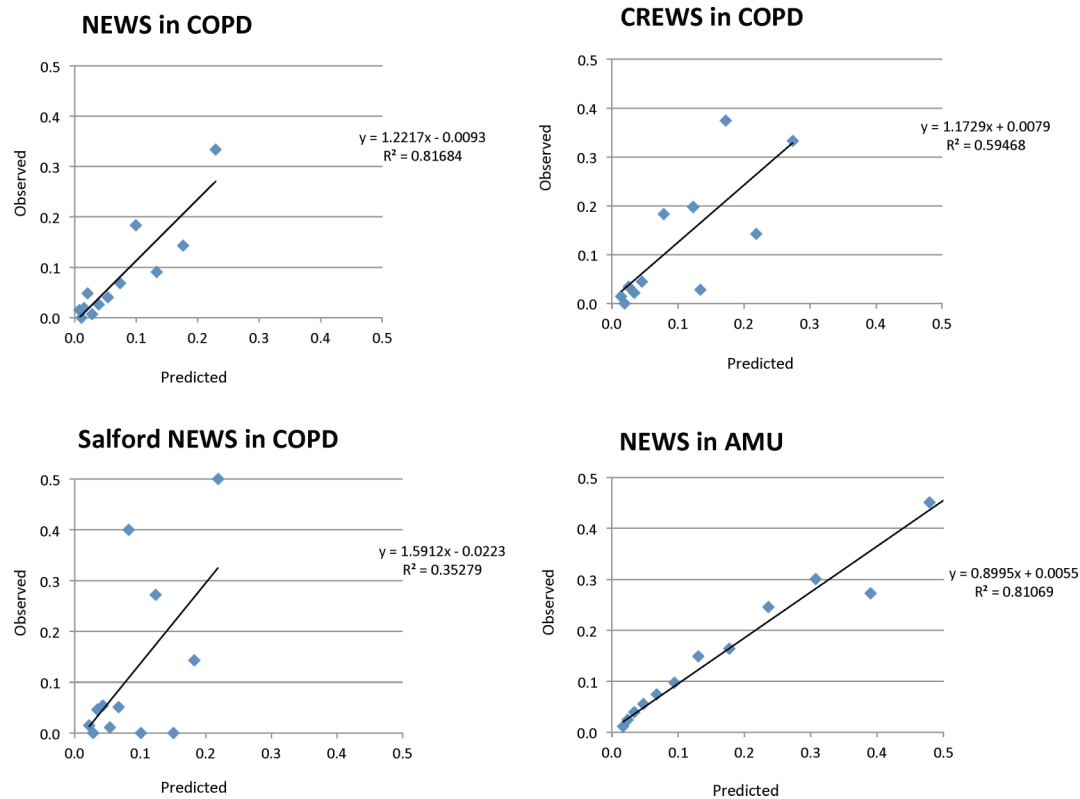
Figure 3 Calibration plots for National Early Warning Score (NEWS), Chronic Respiratory Early Warning Score (CREWS) and Salford-NEWS in acute COPD exacerbation cohort and NEWS in acute medical unit (AMU) cohort for predicted probability versus observed event rate (inpatient mortality) at each level of the score.

diagnostic test for COPD, clinical judgement based on history, physical examination and confirmation of airflow obstruction on spirometry is used. It follows that any study on an often heterogeneous group of patients will be open to criticism of inclusion criteria. Indeed, current COPD guidelines may overdiagnose older men and underdiagnose young women.[21] In one recent study, using a diagnostic code alone, PPV for COPD was 87% (78% to 92%), while adding spirometry plus specific medication only marginally increased the PPV to 89% (81% to 95%).[22] Second, as only two hospital sites (including one with a separate admissions unit for complex elderly patients) were used, with a largely white, elderly demographic, in one South-East England county, cautions over generalisability must be noted. Third, the number of outcome events in the AECOPD cohort for the primary (first admission) analysis was limited (n=42) due to the relatively short observation period, though a sensitivity analysis on all admissions over the study period (n=123 deaths) produced similar results. There is little empirical evidence to guide sample size in validation studies and though 100 events have been suggested, this is based on limited simulation studies[10] and the presented overall sample size is large. Fourth, no A&E observations were available, as the electronic system is only for patients admitted to a hospital ward (including AMU), so for the proportion of patients admitted via A&E, initial observations (up to four hours) would not be reflected in the data. Patients directly admitted to ICU were not included, either, though in both hospitals these account for only <1% of all admissions; such patients, by definition, have already been promptly recognised as critically unwell. Finally, using in-hospital as the outcome measure could potentially miss a patient who subsequently died in the

community, though a minority of patients with respiratory disease die at home.[23]

Prediction models: statistical and clinical relevance and limitations

Although the AUROC is used widely in predictive models, it has well-documented shortcomings,[24–27] including reflecting only a model's ability to rank order cases rather than be a function of actual predicted probabilities and does not inform on the consequences of using a model. Calibration is often overlooked[26] and its most commonly used statistics (H-L test) has shortcomings depending on sample size.[12 13] Groups should also be visualised to interpret calibration (plotting predicted versus observed outcomes).[10] For clinical use, a strong model should lead to a wide range of predicted values and accurately stratify individuals into higher or lower risk categories. The calibration plots in the presented study for the AECOPD cohort suggest an overestimation of risk with NEWS but a potential underestimation with suggested alternatives (and in the AMU cohort using NEWS).

The NEWS is proposed to highlight those at *highest* risk to enable early, appropriate stratification of resources (ie, frequency of performed observations and senior input). Almost half of the AECOPD cohort had observations that would prompt hourly monitoring and urgent senior input (NEWS ≥5 points). As AECOPDs account for a relatively large number of admissions (6% of the 39 470 admissions in the study period), the low PPV has the potential for alert fatigue that could in turn lead to a failure to act when urgent attention is required for patients with and without AECOPD. In the AMU cohort, NEWS of ≥5 points at admission has high specificity and PPV (90% and 17%, respectively) for mortality, making the requirement for

Table 3 Mortality prediction models in AECOPD

Study	Predictors	AUROC	Calibration	EV
Roche et al,[29] 103 French EDs, n=794 (14.5% pneumonia). Outcome: in-hospital mortality (7.4%).	Age >70, MRC score at baseline, cyanosis, use of accessory muscles, paradoxical abdominal movement, asterixis, neurological impairment, oedema.	0.79; IV 0.83	–	Roche et al[38] 0.78
Ruiz-González et al,[30] Spanish single centre, n=160. Excluded if infiltrates on CXR. Outcome: composite death in hospital or <15 days of discharge; ICU transfer or developed acute heart failure (24%).	CRP ≥50 mg/L, current smoker, ≥2 comorbidities and confusion.	0.80, (0.72 to 0.88)	–	–
Tabak et al (BAP-65),[31] US multicentre study, n=88 074. Outcome: in-hospital mortality (2%).	Urea >8 mmol/L, acute mental status change, pulse ≥110/min, age >65.	0.72 (0.70 to 0.74) IV bootstrap; EV 0.71 (0.70 to 0.73)	Plotted derivation vs validation risk groups	Shorr et al,[39] 0.79 (0.78 to 0.80); Hu et al,[40] 0.67 (0.59 to 0.74); Steer et al,[32] 0.68
Asiimwe et al,[33] UK single-centre study, derivation n=4986. Outcome: in-hospital mortality (15.5%).	Albumin, urea and arterial pCO_2.	0.73; IV 0.70	–	–
Steer et al,[32] (DECAF) 2 UK centres; n=920. Outcome: in hospital mortality (10.4%).	Extended MRC DECAF.	0.86 (0.82 to 0.89) IV bootstrap 0.86 (0.82 to 0.89)	H-L test p=0.39	Echevarria et al,[41] 0.82 (0.77 to 0.87)
Tabak et al,[34] derivation 172 US hospitals n=69 299 (2005–2006), validation (2007) n=33 327. Outcome: mortality (3.2% derivation vs 2.9% validation).	Age, urea, pH, albumin, WBC, pCO_2, Troponin I or CK-MB, CPK, pO_2, Na, Hb, PT time or INR, bands, platelets, pulse, temperature, RR, BP, altered mental state, CCF, pulmonary circulation disease, weight loss, metastatic cancer, malignancy.	0.83 (0.82 to 0.84) IV 0.84 (0.83 to 0.85)	Calibration plotted; H-L test derivation p<0.0001, validation p<0.01	–
Stiell et al,[35] 6 Canadian EDs, n=844. Outcome: serious adverse event* (7.8%).	PVD, CABG, prior intubation, ischaemic ECG, pulmonary congestion on CXR, unable to walk in ED, HR ≥110/min, Hb <100 g/L, urea ≥12 mmol/L, serum CO_2 ≥35 mmol/L.	0.80 (0.74 to 0.85); IV bootstrap 0.79 (0.79 to 0.80)	H-L test p=0.7	–
Quintana et al,[36] Spain, n=2487. Outcome: mortality in-hospital or <7 days of discharge (2.4%).	Age, baseline dyspnoea, previous LTOT or NIV, altered mental status, use of accessory muscles or paradoxical breathing on arrival.	0.85 (0.77 to 0.93)	H-L test p>0.62	–

*Death <30 or <14 days and admission to monitored unit, MV or NIV, MI, CABG, PCI, new haemodialysis or, if discharged, readmission <14 days.
AECOPD, acute exacerbation of COPD; AUROC, area under the receiver operating characteristic curve (95% CI, where reported); Bands, immature bands on full blood count; BP, blood pressure; CABG, coronary artery bypass grafting; CCF, congestive cardiac failure; CK-MB, creatine kinase MB isoenzyme; CPK, creatine kinase; CRP, C reactive protein; CXR, chest X-ray; DECAF, Dyspnoea, Eosinopenia, Consolidation, Acidaemia and Atrial Fibrillation; ED, emergency department; EV, external validation; Hb, haemoglobin; H-L test, Hosmer-Lemeshow test (of calibration); ICU, intensive care unit; INR, international normalised ratio; IV, internal validation; LTOT, long-term oxygen therapy; MI, myocardial infarction; MRC, Medical Research Council; MV, mechanical ventilation; Na, Sodium; NIV, non-invasive ventilation; PCI, percutaneous coronary intervention; PT time, prothrombin time; RR, respiratory rate; PVD, peripheral vascular disease; WBC, white blood cells.

senior input entirely appropriate. The study, however, highlights how concentrating just on such patients would miss the majority (57%) of inpatient deaths.

Models to predict mortality or adverse events in AECOPD

As the NEWS has limitations in adequately risk-stratifying patients with AECOPD, investigating more specific prediction models using electronic hospital records is of interest. A systematic review in 2013[28] noted 10 prediction model studies in AECOPD, though only 4 were new models to predict inpatient mortality in general cohorts.[29–32] table 3 summarises these studies and a further four studies were recently published.[33–36] Three have external validation evidence[29 31 32] with discrimination for the outcome studied, as assessed by the AUROC, ranging from 0.72 to 0.86. Automatic computer-based provision of recommendations as part of clinician workflow, at the time and location of decision-making, predicts successful implementation of prediction models.[37] A number of the models include subjective variables (cyanosis, use of accessory muscles)[29] or variables that would require manual input,[29 32 36] making immediate electronic automation a challenge. The externally validated Dyspnoea, Eosinopenia, Consolidation, Acidaemia and Atrial Fibrillation (DECAF) score has been recommended by the most recent BTS COPD Audit, which also recommended that the dyspnoea scale (included in the DECAF) should routinely be filled out at admission in AECOPD.[19] Unfortunately, as with other areas of prediction model studies, there are as yet no published impact analysis studies,[10] though of interest the DECAF score is undergoing analysis as a triage tool for hospital admission in A&E as part of a randomised controlled trial (ISRCTN29082260) comparing home to standard inpatient management of patients with AECOPD.

NEWS and AECOPD

We propose, at present, that the NEWS should remain unadjusted as it provides standardisation as an *aid* to clinical staff. However, as with any observation(s), its utility is limited by the individual interpreting it. Any EWS must be supported by ongoing education and training, with the understanding that certain patient groups are best managed on specialist wards (or with specialist input elsewhere). In a patient with chronically disturbed respiratory physiology, when a specialist has indicated a lower target oxygen saturation (eg, the BTS recommend 88%–92%)[42] for supplemental oxygen delivery, it should be taken into account when interpreting the NEWS. One strategy for the increasing number of hospitals with electronic observations would be to use previous

discharge oxygen saturations as a guide to a patient's baseline function, to individually titrate treatment and observation frequency, during subsequent admissions. This study demonstrates that recommended RCP call-out thresholds in AECOPD lead to hourly or continuous monitoring of a large number of relatively stable patients, diverting resources and producing alarm fatigue. Thus, local practices could be adjusted. For example, 48 hours after a patient's admission, observation frequencies could appropriately be reduced, on a respiratory ward, if observations are close to their baseline state. Alternatively, an adjustment could be made to the score to account for chronic hypoxia in the presence of a normal respiratory rate. Ongoing education in oxygen prescribing and delivery is essential; for example, a patient achieving their target range for oxygen saturation may trigger a higher NEWS, but in this case escalation of oxygen supplementation would be inappropriate.

Meaning of the study

At admission to hospital, the NEWS discriminates for risk of inpatient mortality in patients with AECOPD similarly to general medical patients. The AECOPD group have a significantly higher NEWS, despite the same inpatient mortality. Call-out thresholds for patients with chronic respiratory disease may need reviewing to avoid alert fatigue, though applying clinical acumen, training and guidelines should aid recognition of the high risk or deteriorating patient. It follows that patients with COPD are likely to be best served in a location where input from specialist staff is available. An avenue of future investigation to improve performance of the NEWS in this cohort could be to incorporate blood test parameters and previous (coded) history into an electronic prediction model. The results of the DECAF impact analysis study will also be awaited with interest.

CONCLUSION

For patients with AECOPD, the NEWS as a mortality prediction score on admission to hospital discriminates similarly to general medical patients. Adjustment of the NEWS is not without risks —for example, inappropriately assigning lower risk to hypoxia could potentially attenuate the benefits that NEWS standardisation can bring. Improving education, including utility of the NEWS, recognition of the deteriorating patient, oxygen prescribing and ensuring that such patients are managed with input by respiratory clinicians is recommended.[19]

REFERENCES

1. NCEPOD. National Confidential Enquiry in to Patient Outcome and Death. Time to intervene? London: NCEPOD, 2012.
2. Hillman KM, Bristow PJ, Chey T, et al. Duration of life-threatening antecedents prior to intensive care admission. *Intensive Care Med* 2002;28:1629–34.
3. Duckitt RW, Buxton-Thomas R, Walker J, et al. Worthing physiological scoring system: derivation and validation of a physiological early-warning system for medical admissions. An observational, population-based single-centre study. *Br J Anaesth* 2007;98:769–74.
4. Prytherch DR, Smith GB, Schmidt PE, et al. ViEWS—Towards a national early warning score for detecting adult inpatient deterioration. *Resuscitation* 2010;81:932–7.
5. Physicians RCo. National Early Warning Score (NEWS) Standardising the assessment of acute-illness severity in the NHS—Report of a working party. London, 2012.
6. Eccles SR, Subbe C, Hancock D, et al. CREWS: improving specificity whilst maintaining sensitivity of the National Early Warning Score in patients with chronic hypoxaemia. *Resuscitation* 2014;85:109–11.
7. O'Driscoll BR, Bakerly ND, Murphy P, et al. Concerns regarding the design of the bedside monitoring chart for use with the NEWS (National Early Warning System). *Clin Med* 2013;13:319–20.
8. NICE. *CG101 Chronic obstructive pulmonary disease (update): full guideline*. London: NICE, 2010.
9. Hospital Episode Statistics, Admitted Patient Care, England—2012–2013, 2013.
10. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1–73.
11. WHO. *WHO International Statistical Classification of Diseases and Related Health Problems*. 10th Revision. WHO, 2010.
12. Hosmer DW, Lemeshow S. *Applied logistic regression*. New York: Wiley, 1989:187–215.
13. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. *Crit Care Med* 2007;35: 2052–6.
14. Kellett J, Kim A. Validation of an abbreviated Vitalpac™ Early Warning Score (ViEWS) in 75,419 consecutive admissions to a Canadian regional hospital. *Resuscitation* 2012;83:297–302.
15. Bleyer AJ, Vidya S, Russell GB, et al. Longitudinal analysis of one million vital signs in patients in an academic medical center. *Resuscitation* 2011;82:1387–92.
16. Opio MO, Nansubuga G, Kellett J. Validation of the VitalPAC Early Warning Score (ViEWS) in acutely ill medical patients attending a resource-poor hospital in sub-Saharan Africa. *Resuscitation* 2013;84:743–6.
17. Dawes TR, Cheek E, Bewick V, et al. Introduction of an electronic physiological early warning system: effects on mortality and length of stay. *Br J Anaesth* 2014;113:603–9.
18. Blunt I, Bardsley M, Dixon J. *Trends in emergency admissions in England 2004–2009: is greater efficiency breeding inefficiency?* London: The Nuffield Trust, 5 Jul 2010.
19. BTS. *COPD: Who cares matters. National Chronic Obstructive Pulmonary Disease (COPD) Audit Programme: Clinical audit of COPD exacerbations admitted to acute units in England and Wales 2014*. London: BTS, 2015.
20. Prieto-Centurion V, Rolle AJ, Au DH, et al. Multicenter study comparing case definitions used to identify patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2014;190:989–95.
21. Miller MR, Levy ML. Chronic obstructive pulmonary disease: missed diagnosis versus misdiagnosis. *BMJ* 2015;351:h3021.
22. Quint JK, Müllerova H, DiSantostefano RL, et al. Validation of chronic obstructive pulmonary disease recording in the Clinical Practice Research Datalink (CPRD-GOLD). *BMJ Open* 2014;4:e005540.
23. Statistics OfN. *Mortality Statistics Series DH1 no 38*. London: Office for National Statistics, 2007.
24. Moons KG, van Es GA, Deckers JW, et al. Limitations of sensitivity, specificity, likelihood ratio, and Bayes' theorem in assessing diagnostic probabilities: a clinical example. *Epidemiology* 1997;8:12–17.
25. Mallett S, Halligan S, Collins GS, et al. Exploration of analysis methods for diagnostic imaging tests: problems with ROC AUROC and confidence scores in CT colonography. *PLoS ONE* 2014;9:e107633.
26. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 2007;115:928–35.
27. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006;26:565–74.
28. Singanayagam A, Schembri S, Chalmers JD. Predictors of mortality in hospitalized adults with acute exacerbation of chronic obstructive pulmonary disease. *Ann Am Thorac Soc* 2013;10:81–9.
29. Roche N, Zureik M, Soussan D, et al. Predictors of outcomes in COPD exacerbation cases presenting to the emergency department. *Eur Respir J* 2008;32:953–61.
30. Ruiz-González A, Lacasta D, Ibarz M, et al. C-reactive protein and other predictors of poor outcome in patients hospitalized with exacerbations of chronic obstructive pulmonary disease. *Respirology* 2008;13:1028–33.
31. Tabak YP, Sun X, Johannes RS, et al. Mortality and need for mechanical ventilation in acute exacerbations of chronic obstructive pulmonary disease: development and validation of a simple risk score. *Arch Intern Med* 2009;169:1595–602.
32. Steer J, Gibson J, Bourke SC. The DECAF Score: predicting hospital mortality in exacerbations of chronic obstructive pulmonary disease. *Thorax* 2012;67:970–6.
33. Asiimwe AC, Brims FJ, Andrews NP, et al. Routine laboratory tests can predict in-hospital mortality in acute exacerbations of COPD. *Lung* 2011;189:225–32.
34. Tabak YP, Sun X, Johannes RS, et al. Development and validation of a mortality risk-adjustment model for patients hospitalized for exacerbations of chronic obstructive pulmonary disease. *Med Care* 2013;51:597–605.

35 Stiell IG, Clement CM, Aaron SD, et al. Clinical characteristics associated with adverse events in patients with exacerbation of chronic obstructive pulmonary disease: a prospective cohort study. CMAJ 2014;186:E193–204.

36 Quintana JM, Esteban C, Unzurrunzaga A, et al. Predictive score for mortality in patients with COPD exacerbations attending hospital emergency departments. BMC Med 2014;12:66.

37 Kawamoto K, Houlihan CA, Balas EA, et al. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005;330:765.

38 Roche N, Chavaillon JM, Maurer C, et al. A clinical in-hospital prognostic score for acute exacerbations of COPD. Respir Res 2014;15:99.

39 Shorr AF, Sun X, Johannes RS, et al. Validation of a novel risk score for severity of illness in acute exacerbations of COPD. Chest 2011;140: 1177–83.

40 Hu G, Zhou Y, Wu Y, et al. The Pneumonia Severity Index as a Predictor of In-Hospital Mortality in Acute Exacerbation of Chronic Obstructive Pulmonary Disease. PLoS ONE 2015;10:e0133160.

41 Echevarria C, Steer J, Heslop-Marshall K, et al. Validation of the DECAF score to predict hospital mortality in acute exacerbations of COPD. Thorax 2016; 71:133–40.

42 O'Driscoll BR, Howard LS, Davison AG. BTS guideline for emergency oxygen use in adult patients. Thorax 2008;63(Suppl 6):vi1–68.