**COVER LETTER**

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

**Study design:** Observational cohort external validation study

**Authors:**

Luke E Hodgson, Borislav B Dimitrov, Jo Congleton, Richard Venn, Lui G Forni, Paul J Roderick.

Names and contact details (including email addresses) of suitable peer reviewers:

1. Professor Tom Fahey tomfahey@rcsi.ie, RCSI, Dublin, Ireland.

2. Dr James Calvert james.calvert1@nhs.net, Respiratory Consultant, North Bristol NHS Trust.

3. Dr Nick Hopkinson n.hopkinson@imperial.ac.uk Clinical Senior Lecturer at Imperial College & Honorary Consultant Chest Physician at The Royal Brompton Hospital.

**ABSTRACT**

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

**Methods** An observational cohort study (2012-2014, two UK acute medical units [AMUs]), compared AECOPD (2,361 admissions, 942 individuals, ICD-10 J40-44 codes) with AMU patients (37,109 admissions, 20,415 individuals). Outcome: in-hospital mortality prediction by admission NEWS, CREWS and Salford-NEWS assessed by discrimination (area under receiver operating characteristic curves [AUCs]) and calibration (plots and Hosmer-Lemeshow [H-L] goodness-of-fit).

**Results** Median admission NEWS in AECOPD was 4 (IQR 2-6) vs 1 (0-3) for AMU (p=<0.001), despite mortality 4.5% in both. AECOPD AUCs were: NEWS 0.74 (95% confidence intervals 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 AUC 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) 98% (97 to 99) vs AMU patients PPV 17% (16 to 19) and NPV 97% (97 to 97). For NEWS in AECOPD H-L P-value =0.202.

**Conclusions** This first validation of the NEWS in AECOPD found modest discrimination to predict mortality. Lower specificity of NEWS in AECOPD patients vs 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 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 COPD exacerbation (AECOPD)?

**What is the bottom line?**

The NEWS shows acceptable discrimination in AECOPD and at a cut-off of 5 points has high NPV and low PPV 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.

**TITLE PAGE**

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

**Corresponding author:** 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.

Email: drlhodgson@gmail.com

Telephone: 07734883040

**Authors:**

Hodgson LE,1,2 *Respiratory Registrar*, Dimitrov BD,1*Associate Professor in Medical Statistics,* Congleton J,3 *Consultant Respiratory Physician,* Venn R2, *Consultant Intensivist,* Forni LG4, *Professor Intensive Care Medicine,* Roderick PJ1, *Professor Public Health.*

1. Academic Unit of Primary Care and Population Sciences, Faculty of Medicine, University of Southampton, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD.
2. Western Sussex Hospitals NHS Foundation Trust, Anaesthetics Department, Worthing Hospital, Lyndhurst Rd, Worthing, BN11 2DH.
3. Royal Sussex County Hospital, Brighton & Sussex Hospitals NHS Trust, Eastern Rd, Brighton, BN2 5BE.
4. The Royal Surrey County Hospital NHS Foundation Trust, Egerton Road, Guildford, GU2 7XX.

**Word count abstract: 250**

**Word count main text: 3596**

**Key words:** chronic obstructive pulmonary disease, validation study, prognostic prediction model, national early warning score.

“I Dr Luke Eliot Hodgson, The Corresponding Author of this article contained within the original manuscript which includes any diagrams & photographs and any related or stand alone film submitted (the Contribution”) has the right to grant on behalf of all authors and does grant on behalf of all authors, a licence to the “BMJ Publishing Group Ltd” (“BMJ”) and its licensees, to permit this Contribution (if accepted) to be published in any BMJ products and to exploit all subsidiary rights, as set out in our licence set out at:

http://group.bmj.com/products/journals/instructions-for- authors/wholly\_owned\_licence.pdf

**INTRODUCTION**

Hospitalised patients are frequently exposed to avoidable harm.[1] Adverse trends in clinical observations are often missed or misinterpreted, whilst 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 (EWS). Incorporating observations, with points aggregated in a weighted manner, depending on the degree of abnormality, EWS are examples of prognostic prediction models. Derivation studies of EWS, usually predicting in-patient mortality, have used observations at admission[3] or at 24-hours prior to the outcome.[4] By 2008, over 30 EWS 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 (supplementary file 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 (AUC) 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, nor suggest a way this should be compensated for. Two groups have proposed adjusted NEWS scores for patients with chronic respiratory disease addressing this issue.[6 7] Firstly, the Chronic Respiratory Early Warning Score (CREWS)[6] assigns risk points at a lower oxygen saturation threshold compared to non COPD patients. Secondly, 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] (supplementary file)

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 COPD exacerbation (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 score in terms of in-patient death amongst:

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 in-patient episodes during the study period

d) The AMU cohort for all inpatient episodes during the study period.

2) To externally validate alternative Early Warning Scores (CREWS and Salford-NEWS).[6 7] 1a and 1b represent the primary analysis, and 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 on the adult AMUs of Worthing Hospital and St Richard’s Hospital sites of Western Sussex Hospitals NHS Foundation Trust (WSHFT), for the period 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 in-patients have physiological observations measured and entered via handheld systems into the clinical data software system (Patientrack**©** Sydney, NSW, Australia), with the NEWS automatically calculated. Criteria for the AECOPD cohort were: 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 ICD-10 classification J40-44 (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: patients moved directly from the A&E department to the Intensive Care Unit (ICU) (as neither area uses the Patientrack**©** data system), aged <18 years or 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 in-patient admission. A sensitivity analysis was performed, including all episodes during the study period, to attempt to account for countervailing prognostic factors such as survivor bias and the effects of repeated admissions. Analyzing all episodes also aids generalizability of results, as in clinical practice the NEWS is applied whatever the number of prior admissions. Furthermore, the data was analysed “per admission” by creating a multilevel multiple logistic regression model (MLM), in which we adjusted the main (fixed) effect of the NEWS score for the number of admissions per patient (27 second-level clusters of patients were formed). The main effect was adjusted for the number of admissions as a random intercept, as well as for the NEWS score as predictor with a random slope (at first-level level as nested within the 27 clusters as a second level). 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 in-patient 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 AUC of the ROC 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 statistic in this field.[12] The H-L test associated p-value, when significant (<0.05), may indicate poor fit.[12 13] It is recommended to also 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 performed on SPSS**®** v22 and STATA**®** SE v14.

**RESULTS**

Over the two-year study period there were 2,361 AECOPD in-patient episodes (123 in-patient deaths, median of 3 admissions [interquartile range 2-5]) and 37,109 non-COPD AMU episodes (1,911 deaths). For the primary analysis (first admission), there were 942 patients in the AECOPD cohort and 20,415 in the AMU cohort. The AECOPD cohort had a median age of 74 (67-82) vs 71 (55-82) in the AMU cohort (P <0.001). Median admission NEWS was significantly different - AECOPD 4 points (2-6) vs AMU 1 point (0-3) (P=<0.001). In-patient mortality for first admission did not differ (4.5% in both cohorts). Table 1 summarises admission clinical-demographic variables.

Table 1. Demographics and outcomes of AECOPD and AMU cohorts (first admission during the study period & all episodes).

|  |  |  |  |
| --- | --- | --- | --- |
| First Admission | COPD (n=942) | AMU(n=20,415) | P-value |
| Age | 74 (67-82) | 71 (55-82) | <0.001\* |
| In-patient 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=2,361)** | **AMU** **(n=37,109)**  |  |
| Age | 74 (67-82) | 73 (57-83) | <0.001\* |
| In-patient Mortality | 5.2% (n=123) | 5.1%(n=1,911) | 0.47\*\* |
| Admission NEWS | 4 (2-6) | 1 (0-3) | <0.001 |

Median values (interquartile range), \*Mann-Whitney U test, \*\*Chi-squared. NEWS - national early warning score, AMU - Acute Medical Unit.

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

Table 2. Prediction of in-patient mortality by admission score.

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

Sensitivity, specificity, PPV, NPV (with 95% confidence intervals) at RCP suggested NEWS thresholds of 5 points & 7 points for 1st AECOPD admission (n=942) using NEWS, CREWS, Salford-NEWS scores & the NEWS for the AMU cohort (n=20,415). AMU - acute medical unit patients, PPV - positive predictive value, NPV - negative predictive value.

In the AECOPD cohort for their first admission the AUCs for predicting in-patient mortality for the three prediction scores were: NEWS = 0.74 (95% confidence intervals 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 AUC 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. Calibration plots (shown in Figure 3**)**,suggest no improvement in calibration with the alternative scores, whichboth under-predicted mortality, though the number of deaths was small. In lower risk groups, NEWS in the AMU cohort also under predicted 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, whilst 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 in-patient episode over the 2-year study period (n=2,361), AUCs for predicting in-patient mortality for the three prediction scores were: 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 in-patient episodes (n=37,109) using NEWS the AUC was 0.75 (0.74 to 0.76). (supplementary file has 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 having adjusted further 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 score as a predictor remained significant at p<0.05 (see 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 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 in-patient mortality, admission NEWS in a AECOPD cohort demonstrated similar discrimination to an AMU cohort (AUC 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 to 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, missing opportunities to intervene early. As predictors of mortality on admission (assessed by respective AUCs) CREWS (0.72) and Salford-NEWS (0.66), did not improve discrimination vs 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 in-patient mortality, admission NEWS discriminated satisfactorily (AUC 0.77). A trade-off between sensitivity and specificity must be noted, for example, at a cut-off of NEWS 7 points, sensitivity was only 25% for in-patient mortality. The AUC for the NEWS is similar to a previously described admission prediction model by Duckitt and colleagues (AUC 0.74).[3] The higher AUC in the original derivation study for NEWS (0.89), is explained by prediction timeframe (mortality 24 hours from observation set vs 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, upon which the NEWS is based. One Canadian study found an abbreviated ViEWS gave an AUCof 0.81 (0.80–0.82) to predict 30-day mortality;[14] a US study reported an AUC of 0.86 (timing of outcome not reported);[15] and a Ugandan study reported an AUC 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, in-patient mortality, duration of stay and number of readmissions closely resemble the most recent UK 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 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 heterogenous 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] Secondly, 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. Thirdly, 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 though 100 events have been suggested, this is based on limited simulation studies[10] and the presented overall sample size is large. Fourthly, 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 been recognized 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**

The AUC is used widely in predictive models, yet 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 statistic (H-L test) has shortcomings depending on sample size.[12 13] Groups should also be visualized to interpret calibration (plotting predicted vs observed outcomes).[10] For clinical use a strong model should lead to a wide spread of predicted values and accurately stratify individuals into higher or lower risk categories. The calibration plots in the presented study for the AECOPD cohort suggests 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 (i.e. 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 ≥5 points at admission, has high specificity and PPV (90% and 17% respectively) for mortality, making the requirement for senior input entirely appropriate. The study, however, highlights how concentrating just on such patients would miss the majority (57%) of in-patient deaths.

**Models to predict mortality or adverse events in AECOPD**

As the NEWS has limitations in adequately risk stratifying AECOPD patients, investigating more specific prediction models using electronic hospital records is of interest. A systematic review in 2013[28] noted ten prediction model studies in AECOPD, though only four were new models to predict in-patient mortality in general cohorts.[29-32] Table 3 summarizes these studies and a further four studies recently published.[33-36] Three have external validation evidence[29 31 32] with discrimination for the outcome studied, as assessed by the AUC, 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 DECAF score has been recommended by the most recent BTS COPD Audit, which also recommended that the dyspnea 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 randomized controlled trial (ISRCTN29082260) comparing home to standard in-patient management of patients with AECOPD.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Study | Predictors | AUC | Calibration | External validation |
| [29]Roche, 2008,103 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 | - | [38]Roche 2014 0.78 |
| [30]Ruiz, 2008, Spanishsingle 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 ≥50mg/L, current smoker, ≥2 comorbidities & confusion | 0.80, (0.72 to 0.88). | - | - |
| [31]Tabak (BAP-65), 2009US multi-centre, 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 | [39]Shorr 2011, 0.79 (0.78 to 0.80);[40]Hu 2015 0.67 (0.59 to 0.74);[32]Steer 0.68 |
| [33]Asiimwe, 2011, UK single-centre, derivation n=4,986.Outcome: in-hospital mortality, mortality (15.5%). | albumin, urea, & arterial pCO2 | 0.73;IV 0.70 | - | - |
| [32]Steer (DECAF), 2012, 2 UK centres; n=920. Outcome: in hospital mortality (10.4%). | Extended MRC Dyspnoea Score, eosinopenia, consolidation, acidaemia & atrial fibrillation | 0.86 (0.82 to 0.89)IV bootstrap 0.86 (0.82 to 0.89) | H-L test P=0.39 | [41]Echevarria 2016, 0.82 (0.77 to 0.87) |
| [34]Tabak, 2013, derivation 172 US hospitals n=69,299 (2005-6), validation (2007) n=33,327.Outcome: mortality (3.2% derivation vs 2.9% validation). | Age, Urea, pH, Albumin, WBC, pCO2, Troponin I or CK-MB, CPK, PO2, 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 | - |
| [35]Stiell, 2014, 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 <100g/L, Urea ≥12 mmol/L, serum CO2 ≥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 | - |
| [36]Quintana 2014,Spain, n=2487. Outcome: mortality in-hospital or <7 days of discharge (2.4%). | Age, baseline dyspnea, 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 | - |

Table 3. Mortality prediction models in AECOPD.

\*Death <30 days or <14 days & admission to monitored unit, MV or NIV, MI, CABG, PCI, new hemodialysis or, if discharged re-admission <14 days. AUC – area under the receiver operating characteristic curve (95% confidence intervals, where reported), Bands – immature bands on full blood count, BP – blood pressure, CABG – coronary artery by-pass grafting, CK-MB – creatine kinase MB isoenzyme, CPK – creatine phosphokinase, ED – emergency department, Hb – haemoglobin, H-L test - Hosmer-Lemeshow test (of calibration), IV – internal validation, LTOT – long-term oxygen therapy, MI – myocardial infarction, MV – mechanical ventilation, Na – Sodium, NIV – non-invasive ventilation, PT time – prothrombin time or INR – international normalized ratio, RR – respiratory rate, PVD – peripheral vascular disease, WBC – white blood cells.

**NEWS and AECOPD**

We propose at present, 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 (e.g., the British Thoracic Society [BTS] recommend 88-92%)[42] for supplemental oxygen delivery, that 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 patients’ baseline function, to individually titrate treatment and observation frequency, during subsequent admissions. This study demonstrates 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 patients’ admission, observations 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 in-patient mortality in patients with AECOPD similarly to general medical patients. The AECOPD group have a significantly higher NEWS, despite the same in-patient 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 COPD patients 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]

**COMPETING INTERESTS STATEMENT**

*“All authors have completed the Unified Competing Interest form at* [*www.icmje.org/coi\_disclosure.pdf*](http://www.icmje.org/coi_disclosure.pdf) *and declare that all have no relationships with companies that might have an interest in the submitted work in the previous 3 years; their spouses, partners, or children have no financial relationships that may be relevant to the submitted work; and have no non-financial interests that may be relevant to the submitted work.”*

All authors had full access to all of the data (including statistical reports and tables) in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported and no important aspects of the study have been omitted.

Data sharing: individual consent was not obtained but the presented data are anonymised and risk of identification is low.

**Details of contributors, and the name of the guarantor**

LH - Study design, data collection, data analysis and writing up of the paper.

BDD Study design, data analysis and writing up of the paper.

JC - Study design, data analysis and writing up of the paper.

RV - Study design, data analysis and writing up of the paper.

LF Study design, data analysis and writing up of the paper.

PR Study design, and writing up of the paper.

**Ethics**

A favourable ethical opinion for the study was given by NHS Research Ethics Committee London - South East (REC reference 13/LO/0884).

**No funding applicable**

**Headings to figures**

**Figure 1**. Spread of NEWS score on admission for AECOPD (green n=942 patients) and AMU patients (blue n=20,415 patients) as percentage of total.

**Figure 2 – A (Left) -** Area under the receiver operating curves (AUC) for three prediction scores (NEWS= Green curve, CREWS= Blue, Salford-NEWS=Yellow, Reference line= Black) on admission to predict in-patient mortality in patients with AECOPD (first admission during study period, n=942). NEWS 0.74 (95% CIs 0.66 to 0.82), CREWS 0.72 (0.63 to 0.80) and Salford-NEWS 0.62 (0.53 to 0.70). **B** (**Right)** - AUC for NEWS (Blue line, Reference line= Black) to predict in-patient mortality in the AMU cohort (first admission during study period, n=20,415) - 0.77 (0.75 to 0.78).

**Figure 3**. Calibration plots for NEWS, CREWS & Salford-NEWS in AECOPD cohort and NEWS in AMU cohort for predicted probability vs observed event rate (in-patient mortality) at each level of the score.

**REFERENCES**

1. NCEPOD. National Confidential Enquiry in to Patient Outcome and Death. *Time to intervene?* London, 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**(11):1629-34 doi: 10.1007/s00134-002-1496-ypublished Online First: Epub Date]|.

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**(6):769-74 doi: 10.1093/bja/aem097published Online First: Epub Date]|.

4. Prytherch DR, Smith GB, Schmidt PE, Featherstone PI. ViEWS--Towards a national early warning score for detecting adult inpatient deterioration. Resuscitation 2010;**81**(8):932-7 doi: 10.1016/j.resuscitation.2010.04.014published Online First: Epub Date]|.

5. Physicians RCo. National Early Warning Score (NEWS) Standardising the assessment of acute-illness severity in the NHS - Report of a working party July London, 2012.

6. Eccles SR, Subbe C, Hancock D, Thomson N. CREWS: improving specificity whilst maintaining sensitivity of the National Early Warning Score in patients with chronic hypoxaemia. Resuscitation 2014;**85**(1):109-11 doi: 10.1016/j.resuscitation.2013.08.277published Online First: Epub Date]|.

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**(3):319-20 doi: 10.7861/clinmedicine.13-3-319published Online First: Epub Date]|.

8. NICE. CG101 Chronic obstructive pulmonary disease (update): full guideline. London: NICE, 2010.

9. Hospital Episode Statistics, Admitted Patient Care, England - 2012-13, 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**(1):W1-73 doi: 10.7326/M14-0698published Online First: Epub Date]|.

11. WHO. WHO International Statistical Classification of Diseases and Related Health Problems 10th Revision.

: WHO, 2010.

12. Hosmer DW LS. 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**(9):2052-6 doi: 10.1097/01.ccm.0000275267.64078.b0published Online First: Epub Date]|.

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**(3):297-302 doi: 10.1016/j.resuscitation.2011.08.022published Online First: Epub Date]|.

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**(11):1387-92 doi: 10.1016/j.resuscitation.2011.06.033published Online First: Epub Date]|.

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**(6):743-6 doi: 10.1016/j.resuscitation.2013.02.007published Online First: Epub Date]|.

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**(4):603-9 doi: 10.1093/bja/aeu107published Online First: Epub Date]|.

18. Blunt I, Bardsley M, Dixon J. Trends in emergency admissions in England 2004–2009: is greater efficiency breeding inefficiency?

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**(9):989-95 doi: 10.1164/rccm.201406-1166OCpublished Online First: Epub Date]|.

21. Miller MR, Levy ML. Chronic obstructive pulmonary disease: missed diagnosis versus misdiagnosis. BMJ 2015;**351**:h3021 doi: 10.1136/bmj.h3021published Online First: Epub Date]|.

22. Quint JK, Mullerova H, DiSantostefano RL, et al. Validation of chronic obstructive pulmonary disease recording in the Clinical Practice Research Datalink (CPRD-GOLD). BMJ Open 2014;**4**(7):e005540 doi: 10.1136/bmjopen-2014-005540published Online First: Epub Date]|.

23. Statistics OfN. Mortality Statistics Series DH1 no 38. London 2007.

24. Moons KG, van Es GA, Deckers JW, Habbema JD, Grobbee DE. Limitations of sensitivity, specificity, likelihood ratio, and bayes' theorem in assessing diagnostic probabilities: a clinical example. Epidemiology 1997;**8**(1):12-7

25. Mallett S, Halligan S, Collins GS, Altman DG. Exploration of analysis methods for diagnostic imaging tests: problems with ROC AUC and confidence scores in CT colonography. PLoS One 2014;**9**(10):e107633 doi: 10.1371/journal.pone.0107633published Online First: Epub Date]|.

26. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;**115**(7):928-35 doi: 10.1161/circulationaha.106.672402published Online First: Epub Date]|.

27. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;**26**(6):565-74 doi: 10.1177/0272989x06295361published Online First: Epub Date]|.

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**(2):81-9 doi: 10.1513/AnnalsATS.201208-043OCpublished Online First: Epub Date]|.

29. Roche N, Zureik M, Soussan D, Neukirch F, Perrotin D. Predictors of outcomes in COPD exacerbation cases presenting to the emergency department. Eur Respir J 2008;**32**(4):953-61 doi: 10.1183/09031936.00129507published Online First: Epub Date]|.

30. Ruiz-Gonzalez A, Lacasta D, Ibarz M, Martinez-Alonso M, Falguera M, Porcel JM. C-reactive protein and other predictors of poor outcome in patients hospitalized with exacerbations of chronic obstructive pulmonary disease. Respirology 2008;**13**(7):1028-33 doi: 10.1111/j.1440-1843.2008.01403.xpublished Online First: Epub Date]|.

31. Tabak YP, Sun X, Johannes RS, Gupta V, Shorr AF. 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**(17):1595-602 doi: 10.1001/archinternmed.2009.270published Online First: Epub Date]|.

32. Steer J, Gibson J, Bourke SC. The DECAF Score: predicting hospital mortality in exacerbations of chronic obstructive pulmonary disease. Thorax 2012;**67**(11):970-6 doi: 10.1136/thoraxjnl-2012-202103published Online First: Epub Date]|.

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**(3):225-32 doi: 10.1007/s00408-011-9298-zpublished Online First: Epub Date]|.

34. Tabak YP, Sun X, Johannes RS, Hyde L, Shorr AF, Lindenauer PK. Development and validation of a mortality risk-adjustment model for patients hospitalized for exacerbations of chronic obstructive pulmonary disease. Med Care 2013;**51**(7):597-605 doi: 10.1097/MLR.0b013e3182901982published Online First: Epub Date]|.

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**(6):E193-204 doi: 10.1503/cmaj.130968published Online First: Epub Date]|.

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 doi: 10.1186/1741-7015-12-66published Online First: Epub Date]|.

37. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005;**330**(7494):765 doi: 10.1136/bmj.38398.500764.8Fpublished Online First: Epub Date]|.

38. Roche N, Chavaillon JM, Maurer C, Zureik M, Piquet J. A clinical in-hospital prognostic score for acute exacerbations of COPD. Respir Res 2014;**15**:99 doi: 10.1186/s12931-014-0099-9published Online First: Epub Date]|.

39. Shorr AF, Sun X, Johannes RS, Yaitanes A, Tabak YP. Validation of a novel risk score for severity of illness in acute exacerbations of COPD. Chest 2011;**140**(5):1177-83 doi: 10.1378/chest.10-3035published Online First: Epub Date]|.

40. Hu G, Zhou Y, Wu Y, Yu Y, Liang W, Ran P. The Pneumonia Severity Index as a Predictor of In-Hospital Mortality in Acute Exacerbation of Chronic Obstructive Pulmonary Disease. PLoS One 2015;**10**(7):e0133160 doi: 10.1371/journal.pone.0133160published Online First: Epub Date]|.

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**(2):133-40 doi: 10.1136/thoraxjnl-2015-207775published Online First: Epub Date]|.

42. O'Driscoll BR, Howard LS, Davison AG. BTS guideline for emergency oxygen use in adult patients. Thorax 2008;**63 Suppl 6**:vi1-68 doi: 10.1136/thx.2008.102947published Online First: Epub Date]|.