OPEN ACCESS

ORIGINAL ARTICLE

Comparison of early warning scores in patients with COPD exacerbation: DECAF and NEWS score

Carlos Echevarria,[1,2] John Steer,[1,3] Stephen C Bourke[1,3]

Linked

► http://dx.doi.org/10.1136/thoraxjnl-2019-213788

Check for updates

ABSTRACT

Background The National Early Warning Score 2 (NEWS2) includes two oxygen saturation scales; the second adjusts target saturations to 88%–92% for those with hypercapnic respiratory failure. Using this second scale in all patients with COPD exacerbation ('NEWS2$_{All COPD}$') would simplify practice, but the impact on alert frequency and prognostic performance is unknown. Admission NEWS2 score has not been compared with DECAF (dyspnoea, eosinopenia, consolidation, acidaemia, atrial fibrillation) for inpatient mortality prediction.

Methods NEWS, NEWS2 and NEWS2$_{All COPD}$ and DECAF were calculated at admission in 2645 patients with COPD exacerbation attending consecutively to one of six UK hospitals, all of whom met spirometry criteria for COPD. Alert frequency and appropriateness were assessed for all NEWS iterations. Prognostic performance was compared using the area under the receiver operating characteristic (AUROC) curve. Missing data were imputed using multiple imputation.

Findings Compared with NEWS, NEWS2 reclassified 3.1% patients as not requiring review by a senior clinician (score≥5). NEWS2$_{All COPD}$ reduced alerts by 12.6%, or 16.1% if scoring for injudicious use of oxygen was exempted. Mortality was low in reclassified patients, with no patients dying the same day as being identified as low risk. NEWS2$_{All COPD}$ was a better prognostic score than NEWS (AUROC 0.72 vs 0.65, p<0.001), with similar performance to NEWS2 (AUROC 0.72 vs 0.70, p=0.090). DECAF was superior to all scores (validation cohort AUROC 0.82) and offered a more clinically useful range of risk stratification (DECAF=1.2%–25.5%; NEWS2=3.5%–15.4%).

Conclusion NEWS2$_{All COPD}$ safely reduces the alert frequency compared with NEWS2. DECAF offers superior prognostic performance to guide clinical decision-making on admission, but does not replace repeated measures of NEWS2 during hospitalisation to detect the deteriorating patient.

Key messages

What is the key question?

► Does National Early Warning Score 2 (NEWS2) reduce alert frequency, and offer superior prognostic performance to other versions of NEWS and the DECAF (dyspnoea, eosinopenia, consolidation, acidaemia, atrial fibrillation) score in 2645 patients admitted with COPD exacerbation who have spirometry-proven COPD?

What is the bottom line?

► NEWS2 is superior to the original NEWS, but using the second oxygen scale within NEWS2 (target oxygen saturations 88%–92%) in all patients with COPD exacerbation substantially reduces false alerts without increased risk. DECAF offers superior risk-stratification on admission, but does not replace early warning scores to detect subsequent deterioration.

Why read on?

► Our data support monitoring all patients with COPD exacerbation using the second NEWS2 oxygen scale, and this advice should override the current Royal College of Physicians recommendation that the second scale should only be used in hypercapnic patients and at the discretion of a clinician.

BACKGROUND

The National Early Warning Score (NEWS) facilitates prompt recognition of clinical deterioration in patients hospitalised with a wide range of conditions.[1] It is composed of bedside indices (respiratory rate, oxygen saturations, systolic blood pressure, pulse, level of consciousness and temperature) with clear thresholds commanding the urgency of clinical response. Early warning scores, such as NEWS, are intended to be scored repeatedly during the inpatient stay to monitor patients for deterioration.

Although NEWS is widely used in the UK, two key limitations have been highlighted. First, patients with COPD and chronic hypoxaemia raise false alerts, which may lead to alert fatigue and complacency. Second, in severe COPD exacerbations (ECOPD), excess oxygen is associated with increased need for ventilation and mortality,[2–4] but this risk is not identified by NEWS,[5 6] which may encourage unsafe oxygen prescribing to achieve a lower score. The recently updated NEWS2 includes a second oxygen saturation scale which is intended for patients with hypercapnic respiratory failure (table 1). New onset of confusion has also been added to the level of consciousness index.[1] The new oxygen saturation scale encourages the prescription and delivery of oxygen to target saturations of 88%–92% in those with hypercapnia,[7] while the original oxygen saturation scale is retained for patients without hypercapnia. This incorrectly assumes that patients with ECOPD without hypercapnia are not at risk of harm from excess oxygen;

Table 1 National Early Warning Score 2 (NEWS2) with risk score for oxygen saturations alone shown for scale 1 (normocapnic patients) and for scale 2 (hypercapnic patients)

NEWS2 scale 1 and original NEWS (normocapnic patients)		NEWS2 scale 2 (hypercapnic patients)	
Oxygen saturations, %	Score	Oxygen saturations, %	Score
91 or less	3	83 or less	3
92–93	2	84–85	2
94–95	1	86–87	1
96 or more	0	88–92	0
		93 or more on air	0
		93–94 on O_2	1
		95–96 on O_2	2
		97 or more on O_2	3

Scale 1 is the same as the original NEWS oxygen saturation index and is recommended for most patients including patients with COPD without hypercapnic respiratory failure. Scale 2 is recommended for patients with hypercapnic respiratory failure, whose target saturations are set lower (ie, 88%–92%).

this can lead to CO_2 retention[5] and is associated with increased mortality risk.[8]

Despite clear guidelines by the British Thoracic Society (BTS) promoting target saturations of 88%–92% in most patients with COPD,[7] this is frequently not achieved in clinical practice.[9 10] By assigning risk points to those with excess oxygen, NEWS2 will encourage the correct titration of oxygen in patients identified as at risk. Unfortunately, this is limited to patients confirmed to have hypercapnic respiratory failure on blood gas analysis and with clinician approval. Patients with hypercapnia will be at risk until blood gas analysis and clinical review, while those with normocapnia will remain at risk throughout. The BTS guidelines limit risk by recommending initial target saturations of 88%–92%, but risk is not avoided as the target can be increased to 94%–98% following confirmation of normocapnia. In unselected patients with a clinical diagnosis of ECOPD, there is randomised controlled trial (RCT) evidence that target saturations of 88%–92% improve mortality.[6] This favours adopting the second NEWS2 oxygen scale (88%–92%) in all patients with ECOPD, which would simplify the pathway facilitating implementation, and should greatly reduce the risk of excess oxygen use.

The DECAF (dyspnoea, eosinopenia, consolidation, acidaemia, atrial fibrillation) score is a risk stratification tool intended to be scored on admission and accurately predicts risk of death.[11 12] It can be easily calculated at the bedside to guide treatment, such as hospital at home for low-risk patients.[13] DECAF and NEWS2 serve separate purposes in clinical practice: the former to risk assess patients accurately at admission and the latter to monitor patients for deterioration throughout admission. However, admission NEWS2 has not been compared with DECAF. The UK National Asthma and COPD Audit Programme currently relies on admission NEWS2 to allow case-mix adjustment for mortality risk; the superiority of DECAF would support its routine use for this purpose.[14] Furthermore, if admission NEWS2 offered similar performance to DECAF in ECOPD, then the principle of parsimony would favour reliance on NEWS2 alone and render the DECAF score redundant.

The primary aim of this study was to assess the effect of modifying NEWS2 to use the second oxygen saturation scale in all patients with ECOPD (target saturations of 88%–92%) with respect to inpatient alert frequency and appropriateness, and mortality prediction. We refer to this adapted version of NEWS2 as 'NEWS2$_{All COPD}$'. Furthermore, we compare the performances of admission NEWS2$_{All COPD}$, NEWS2, NEWS and the DECAF score.

Table 2 Indices that comprise news and DECAF scores, and rates of missing data in each cohort

	Derivation cohort n=920	Validation cohort n=1725
NEWS and NEWS2		
Respiratory rate	0.33%	1.8%
Oxygen saturation	0%	1.8%
Supplemental oxygen	1.5%	0%
Systolic blood pressure	0%	1.7%
Pulse	0%	1.5%
Level of consciousness	0%	1.8%
Confusion (NEWS2 only)	0%	2.6%
Temperature	0.54%	2.6%
DECAF score		
eMRCD score	0%	0%
Eosinopenia	0%	0%
Chest X-ray consolidation	0%	0%
Acidaemia (pH <7.3)	0%*	0%*
Atrial fibrillation	0%	0%

*In patients without an arterial blood gas, it was assumed that none had an acidaemia of <7.30 provided oxygen saturations were 93% or more while breathing room air. This was based on data from the derivation cohort: of 118 patients meeting these criteria, none had an arterial pH of <7.30.

METHODS
The DECAF derivation and validation cohorts are composed of 2645 consecutive admissions of unique patients with ECOPD to six UK hospitals with preadmission obstructive spirometry, age of 35 or older, and smoking history of 10 or more cigarette pack-years (ISRCTN13946813 and ISRCTN29082260).[11 12]

Collection of the DECAF and NEWS score indices, including confusion, was pre-specified in the original study protocols. The first set of NEWS indices was used, either from the accident and emergency (A+E) department, or the medical admissions unit if the patient bypassed A+E. Missing data were handled using multiple imputation as described previously.[15 16] NEWS2 was initially calculated using scale 1 and scale 2 for normocapnic and hypercapnic patients, respectively ('NEWS2'). NEWS2 was then re-calculated with the assumption that all patients with COPD should have oxygen saturations of 88%–92% ('NEWS2$_{All COPD}$').

Patients were assigned to standard risk groups for each early warning score. For NEWS, NEWS2 and NEWS2$_{All COPD}$, a score of 0 to 4 is low risk, 5 or 6 is moderate risk, and 7 or more is high risk. A patient with a NEWS of 5 or more should have at least hourly observations and be reviewed by a clinician with competencies in dealing with acute illness, such as a ward doctor. A score of 7 or more should prompt an emergency review, with continuous monitoring of observations.[1] For the DECAF score, 0 to 1 equates to a low in-hospital mortality risk, 2 is moderate risk, and 3 or more is high risk. Alert frequencies (NEWS, NEWS2, NEWS2$_{All COPD}$ score ≥5) and other proportions were compared with Fisher's exact test. The performance of NEWS, NEWS2, NEWS2$_{All COPD}$ and DECAF for the prediction of

Table 3 Numbers of deaths by risk group for NEWS2, NEWS2 88%–92% and NEWS

All cohorts	Low-risk groups (0–4)			Moderate-risk groups (5–6)			High-risk groups (7+)		
	Died	All	%	Died	All	%	Died	All	%
NEWS2$_{All COPD}$	35	1003	3.5	39	628	6.2	154	1014	15.2
NEWS2	23	752	3.1	45	671	6.7	160	1222	13.1
NEWS	25	670	3.7	57	730	7.8	146	1245	11.7

NEWS, NEWS2 and NEWS2 88%–92%: a score of 0 to 4 is low risk, 5 or 6 is moderate risk, and 7 or more is high risk.

inpatient mortality was assessed and compared using the area under the receiver operating characteristic (AUROC) curve.[17] Calibration was assessed using the Hosmer-Lemeshow (HL) test and by visual assessment in calibration plots.[18 19] Analyses were performed on IBM SPSS Statistics V.22 and SigmaPlot V.12.3.

RESULTS

Population and missing data

In 2645 patients, the mean (SD) age was 73.1 (10.2), 54% were women and the mean (SD) FEV$_1$ was 45.1% (SD 18.2) predicted. Almost half were unable to leave the house unassisted (Extended Medical Research Council Dyspnoea score 5a or 5b=47%), and 29.8% had consolidation on the admission chest X-ray. Furthermore, 17.9% of patients were acidaemic on the admission arterial blood gas (ABG) analysis, and 10.1% had a pH of less than 7.3. Comorbid conditions were common; 55.4% had a Charlson index of 2 or more, and 28.9% scored 3 or more. Overall inpatient mortality was 8.6%; 96 in 920 (10.4%) in the derivation cohort and 132 in 1725 (7.7%) in the validation cohort.

A more detailed breakdown of patient characteristics by site and cohort is published elsewhere.[11 12] There were no missing data for the primary outcome, inpatient mortality. Rates of missing data for indices were low and are shown in table 2.

Risk groups and alerts

Inpatient mortality by risk group is shown in table 3. The percentage of patients classified as requiring an urgent review (score of 5 or more) for NEWS2$_{All COPD}$, NEWS2 and NEWS was 62.1%, 71.6% and 74.7% (middle column figure 1).
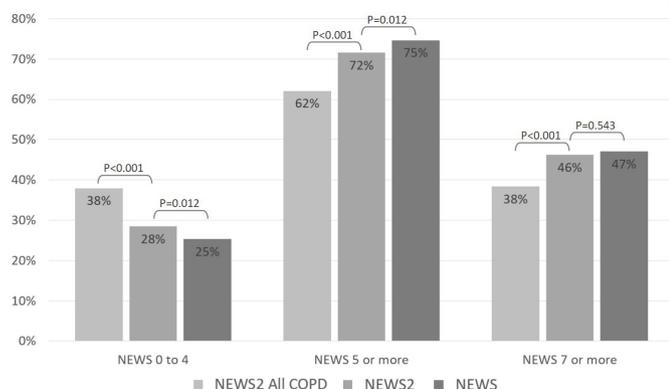

Figure 1 Frequency of alerts for NEWS2$_{All COPD}$, NEWS2 and NEWS. Figure shows the percentage of patients in each risk category, grouped together by early warning score. The first column in each group is NEWS2$_{All COPD}$, the second is NEWS2 and the third is NEWS. P values were calculated using Fisher's exact test. NEWS 5–6=medium risk, prompting urgent response by clinician/clinical team and minimum hourly observations. NEWS 7=high risk. Urgent response by clinician/clinical team, which may include critical care, and continuous monitoring or vital signs.

NEWS2$_{All COPD}$ resulted in a shift of total scores towards the lower risk range compared with NEWS2 (figure 2). NEWS2$_{All COPD}$ identified fewer patients requiring medical review relative to NEWS2 (9.5% absolute reduction), but the risk of death in the low-risk group was similar (3.5% and 3.1%, p=0.686).

A patient receiving excess oxygen may not necessarily require an urgent clinical review and should first have their oxygen titrated down to target range. For NEWS2, if excess oxygen scores were discounted, there was a significant reduction in the frequency of alerts for patients with a score of 5 or more (71.6% to 68.0%, p=0.005). For NEWS2$_{All COPD}$, discounting excess oxygen reduced reviews in this group from 62.1% to 58.6% (p=0.010), with little difference in the percentage of deaths in the low-risk patient group (counting excess oxygen=3.5% vs discounting excess oxygen=3.6%, p=0.905). Of those patients who were assigned a low-risk NEWS2 or NEWS2$_{All COPD}$ score, none died that same day. This suggests that all of these patients were correctly categorised as low risk at the time; early warning scores are repeated with each set of observations and are likely to have risen in those who subsequently died.

Performance of admission DECAF and early warning scores for inpatient mortality

For all patients (n=2645), NEWS2$_{All COPD}$ was a stronger mortality predictor than NEWS (AUROC NEWS2$_{All COPD}$=0.72, 95% CI 0.68 to 0.76, vs NEWS=0.65, 95% CI 0.61 to 0.68; p<0.001). NEWS2$_{All COPD}$ showed a trend towards superiority compared with NEWS2 (AUROC NEWS2$_{All COPD}$=0.72, 95% CI 0.68 to 0.76, vs NEWS2=0.70, 95% CI 0.67 to 0.74; p=0.090).

Of interest, the addition of confusion to the 'level of consciousness' component of NEWS2 accounted for part of the improved performance in NEWS2 compared with NEWS. For example,
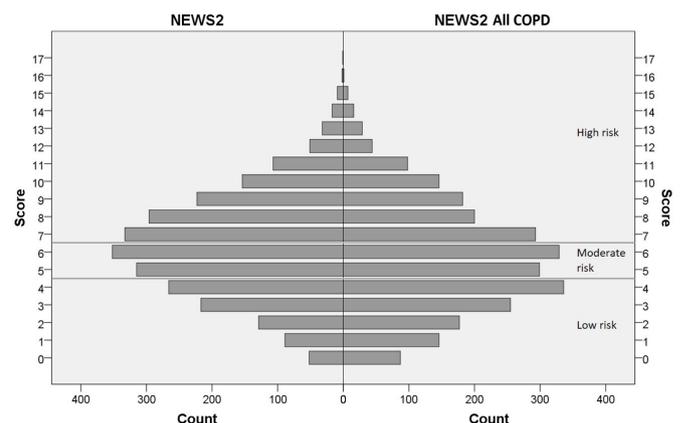

Figure 2 Histogram of NEWS scores. Figure shows all patients from the derivation and validation cohorts. The number of individuals is shown in the x-axis ('count') and the total NEWS2 score is shown on the y-axis. The grey lines separate low-risk (0–4), moderate-risk (5–6) and high-risk (7 or more) groups.

Table 4 Sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV) at NEWS thresholds of 5 and 7

	Low-risk group (cut-off less than 5 for NEWS)			Low and moderate risk group combined (cut-off less than 7 for NEWS)		
	Sensitivity	Specificity	NPV	Sensitivity	Specificity	PPV
NEWS2$_{All COPD}$	0.85	0.40	0.97	0.68	0.64	0.15
NEWS2	0.90	0.30	0.97	0.71	0.55	0.13
NEWS	0.89	0.27	0.96	0.64	0.55	0.12
DECAF	0.93	0.52	0.99	0.73	0.80	0.25

when scored without confusion, the AUROC for NEWS2$_{All COPD}$ was only 0.68 (95% CI 0.64 to 0.71) which is significantly worse than NEWS2$_{All COPD}$ scored with confusion (p<0.001).

DECAF provided optimal prediction of inpatient mortality, and was superior to NEWS2 and NEWS2$_{All COPD}$ in terms of performance as shown in figure 3 (further data in online supplementary table 1).

Table 4 shows the performance of early warning scores at low-risk and high-risk NEWS thresholds (5 and 7), with the DECAF score the gold standard. The positive predictive value (PPV) and negative predictive value (NPV) most inform the clinician: the PPV is the chance of a patient dying if they score above the threshold, while the NPV is the chance of a patient surviving if they score below the threshold. Sensitivity is the proportion of those who died that were correctly identified as at risk (above the threshold), while the specificity looks at the proportion of those who survived, identified as not at risk (below the threshold).

In the low-risk group, a higher sensitivity and NPV is desirable. With regards to sensitivity, NEWS2 performed optimally, while both NEWS2 scores performed equally well for NPV. In the high-risk group, a high specificity and PPV is desirable, for which NEWS2$_{All COPD}$ performed optimally.

Calibration describes how well individuals' predicted risks from a logistic regression model agree with the observed risk in terms of the proportion of individuals affected. The goodness of fit can be measured with the HL test, with poor fit indicated by a small calculated p value. DECAF, NEWS2$_{All COPD}$ and NEWS were well calibrated with HL statistics of p=0.209, p=0.237 and p=0.736. NEWS2 was not well calibrated, with an HL statistic of p=0.048.
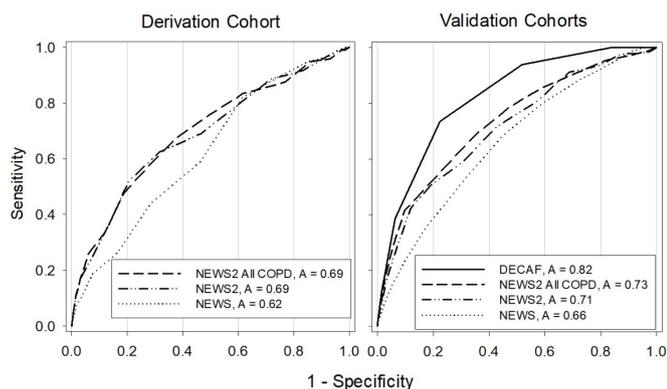

Figure 3 Receiver operator curves for the DECAF score, NEWS2$_{All COPD}$ and NEWS all cohorts. Figure shows the performance of the DECAF, NEWS2$_{All COPD}$, NEWS2 and NEWS scores. A higher area under the receiver operating characteristic curve ('A') shows better prediction. DECAF is not included in the derivation cohort, as this would unfairly favour its performance.

Calibration curves are shown in figure 4. A good prognostic tool should cover a range of risks, have an intercept close to zero and a slope of one, and have the intersection of predicted and observed risk close to the line (shown by the dots). NEWS2$_{All COPD}$ was well calibrated and covered a larger range of risks than other early warning scores.

DISCUSSION

Patients admitted with ECOPD have a modest in-hospital mortality rate (3.9%–4.3%),[10] [20] but frequently trigger alerts on the National Early Warning Scale (NEWS). This can place a substantial burden on healthcare professionals, create complacency and promote excess oxygen use in those who are at particular risk of its harmful effects.[6] In this study, NEWS2 showed superior discrimination for mortality to NEWS, but only
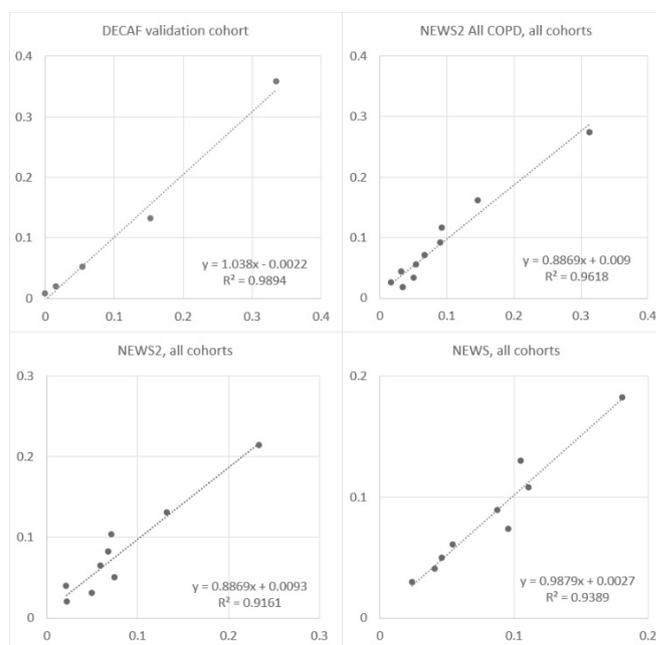

Figure 4 Calibration curves for DECAF, NEWS2$_{All COPD}$, NEWS2 and NEWS in all cohorts. Figure shows calibration curves for NEWS2$_{All COPD}$, NEWS2 and NEWS in all patients. DECAF is shown in the validation cohort, as inclusion of the derivation cohort would unfairly favour its performance. Predicted risk is shown on the y-axis which was calculated using logistic regression analysis to give individual's mortality risk as a percentage. The x-axis shows the observed risk, which is a proportion of the number of patients who died within an assigned risk range. The slope of the equation should be near 1 (eg, for DECAF, the slope is 1.038), and the intercept should be near 0 (for DECAF, this is 0.0022). The R^2 represents the correlation between the predicted risk and observed risk using Pearson's correlation, with scores closest to 1 showing maximum correlation.

reclassified 3.1% as not requiring a review, based on an alert trigger of 5 or more. NEWS2$_{All\ COPD}$, adopting target saturations of 88%–92% and scale 2 of the NEWS2 tool for all patients, led to an absolute reduction in alert frequency of 12.6%. In clinical practice, titrating down oxygen prior to alerts may be appropriate; this approach decreased alerts by 16.1%. The performance of NEWS2$_{All\ COPD}$ was similar to NEWS2, with no increased risk of death in the low-risk group. Of key importance, there were no deaths on the same day that patients were classified as not requiring review.

DECAF was prognostically superior to all scores, showed the most consistent performance and was well calibrated. Accurate mortality prediction informs clinical decision-making; the low-risk DECAF group was larger with a lower risk of death than seen with other scores, which makes DECAF suitable for the identification of patients for early discharge or hospital at home services. Additionally, in the National Asthma and COPD Audit, baseline mortality risk is being quantified by the NEWS2 score, but our results show NEWS2 should not be used in place of the DECAF score for this purpose.

Eligibility criteria were broad and rates of missing data were low, with robust data analyses, including methods to deal with missing data. The assessment of the NEWS score was pre-specified, as was the assessment of confusion; therefore, there is no bias in the subsequent re-coding of NEWS2. Furthermore, in both cohorts, the inclusion of NEWS indices was pre-specified, and data were collected prior to and/or without knowledge of the (objective) outcome. For the AUROC curve comparison, DECAF was compared with NEWS iterations only in the validation cohort; comparing performance in the derivation cohort would have introduced bias in favour of the DECAF score. A key strength is that our population was well described, with respiratory specialist confirmation of exacerbation and spirometric confirmation of COPD. These results may not apply to those with presumed ECOPD without obstructive spirometry (patients misdiagnosed with COPD). In our study, considerable efforts were made to obtain previous lung function results: we checked primary and secondary paper and electronic records before concluding spirometry was unavailable.

Our study has the following limitations. Scores were compared using admission data only. For the comparison with the DECAF score, using admission data is appropriate as we are assessing the tools' performance as a guide to inform clinical decision-making at the point of admission. However, for NEWS the inclusion of multiple measures would have added further useful data.

'NEWS2-without confusion' has been compared with NEWS by Hodgson and others, in a large cohort of patients.[21] They reported concerns that high-risk patients were inappropriately downgraded by NEWS2. However, NEWS2 was calculated from NEWS, which does not include confusion. This is an important omission as we have shown that confusion offers additional prognostic benefit, and inclusion of confusion may have resulted in fewer patients being downgraded. Furthermore, time to death was not reported: in our study, no low-risk patients died the same day, which supports correct risk categorisation. In common with other studies assessing similar early warning scores (such as Salford NEWS),[22] Hodgson et al[23] used ICD-10 codes to identify patients, which are known to be inaccurate.

A recent retrospective study by Pimentel et al[24] offers the advantage of capture of sequential early warning scores in 251 266 acute adult admissions. Again, reliance on coding limits data quality in terms of case ascertainment. Again, confusion was not included, and therefore NEWS2 was not correctly scored. In 48 898 patients with conditions that were associated with risk of hypercapnia based on ICD-10 codes, 'NEWS2-without confusion' was slightly inferior to NEWS for identifying death within 24 hours. In a subgroup of 1394 patients with confirmed hypercapnia, 'NEWS2-without confusion' offered similar discrimination and a higher PPV than NEWS. In common with our study, 'NEWS2-without confusion' generated fewer alerts, and mortality rates were low on the same day as patients having low-risk assessments. Our study differs in that we only included patients with confirmed COPD, NEWS2 was correctly scored and case ascertainment was robust.

While no patient should receive a diagnosis of COPD without obstructive spirometry, we acknowledge that this does occur and that our results may not be generalisable to such patients. However, we highlight that the Austin RCT showed that target SpO$_2$ 88%–92% in all patients with a presumed diagnosis of ECOPD at ambulance pick-up was associated with lower mortality, and this included patients without a spirometry-confirmed diagnosis of COPD.[6]

An important limitation of this and other studies of NEWS2 is that we did not observe NEWS2 in practice. Real-time recording of NEWS2 may result in changes in clinical behaviour. For example, while the performance of NEWS2$_{All\ COPD}$ and NEWS2 was similar in our study, the reduction in false alerts with NEWS2$_{All\ COPD}$ could result in the more timely clinical assessment of other patients and hence a positive clinical outcome. Prospective randomised studies are required to address these points,[25] and to see if the use of NEWS2$_{All\ COPD}$ leads to improvements in mortality by preventing excess oxygen prescribing, but would be challenging to perform. Mortality was high in this study, though appropriate given the higher levels of comorbidity, pneumonia and frailty in our cohort compared with UK national audit data collected at a similar time.

The current BTS oxygen guidelines recommend initial target saturations of 88%–92% in patients with COPD. NEWS2 requires confirmation of hypercapnia before the alternative oxygen saturation scale is activated and should only be instigated by a clinician. Of concern, the NEWS2 implementation guidance and NEWS2 chart may result in many patients receiving excess oxygen prior to ABG analysis or clinician review. RCT data show a mortality benefit in those treated with oxygen target saturations 88%–92% at the point of ambulance pick-up. These patients did not have baseline ABG analysis prior to oxygen treatment, and this approach was beneficial even though a substantial proportion of patients were subsequently described as not having COPD.[6] RCTs have identified increases in mortality with excess oxygen in patients with a range of acute medical presentations even where there is no risk of hypercapnia.[26] Using NEWS2$_{All\ COPD}$ in all patients with ECOPD could offer improved prognostic performance and calibration, with significantly fewer alerts, and lead to a reduction in harm from excess oxygen delivery in this at-risk population.

REFERENCES

1. Royal College of Physicians. *National Early Warning Score (NEWS) 2: standardising the assessment of acute-illness severity in the NHS. Updated report of a working party*. London: RCP, 2017.
2. Wedzicha JA, Seemungal TAR. COPD exacerbations: defining their cause and prevention. *Lancet* 2007;370:786–96.
3. Inspiring change. *The national confidential enquiry into patient outcome and death*. Inspiring change, 2017.
4. Alrawi YA, Potter JF, Myint PK. UK National COPD Resources and Outcomes Project (NCROP): 2008 national audit data presents an opportunity to highlight the areas for improvement in COPD care in the ageing population. *COPD* 2010;7:360–5.
5. O'Driscoll BR, Grant K, Green D, *et al*. The national early warning score gives misleading scores for oxygen saturation in patients at risk of hypercapnia. *Clin Med* 2014;14:695–6.
6. Austin MA, Wills KE, Blizzard L, *et al*. Effect of high flow oxygen on mortality in chronic obstructive pulmonary disease patients in prehospital setting: randomised controlled trial. *BMJ* 2010;341:c5462.
7. O'Driscoll BR, Howard LS, Earis J, *et al*. BTS guideline for oxygen use in adults in healthcare and emergency settings. *Thorax* 2017;72(Suppl 1):ii1–90.
8. Robinson TD, Freiberg DB, Regnis JA, *et al*. The role of hypoventilation and ventilation-perfusion redistribution in oxygen-induced hypercapnia during acute exacerbations of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2000;161:1524–9.
9. O'Driscoll R.BTS emergency oxygen audit report; 2015.
10. Stone RA, Holzhauer-Barrie J, Lowe D, *et al*. *OPD: 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*, 2015.
11. Steer J, Gibson J, Bourke SC. The DECAF score: predicting hospital mortality in exacerbations of chronic obstructive pulmonary disease. *Thorax* 2012;67:970–6.
12. 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.
13. Echevarria C, Gray J, Hartley T, *et al*. Home treatment of COPD exacerbation selected by DECAF score: a non-inferiority, randomised controlled trial and economic evaluation. *Thorax* 2018;73:713–22.
14. Royal College of Physicians. National Asthma and COPD Audit Programme (NACAP): secondary care workstream—COPD resources, 2018. Available: https://www.rcplondon.ac.uk/projects/outputs/national-asthma-and-copd-audit-programme-nacap-secondary-care-workstream-copd [Accessed 8 Nov 2018].
15. Echevarria C, Steer J, Heslop-Marshall K, *et al*. The PEARL score predicts 90-day readmission or death after hospitalisation for acute exacerbation of COPD. *Thorax* 2017;72:686–93.
16. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc* 1996;91:473–89.
17. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–45.
18. Hosmer DW, Lemesbow S. Goodness of fit tests for the multiple logistic regression model. *Commun Stat Theory Methods* 1980;9:1043–69.
19. Moons KGM, 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.
20. Stone RA MV, Mortier K, Holzhauer-Barrie J, *et al*. *COPD: working together. National chronic obstructive pulmonary disease (COPD) audit programme: clinical audit of COPD exacerbations admitted to acute hospitals in England and Wales 2017. National clinical audit report*. London: RCP, 2018.
21. Hodgson LE, Congleton J, Venn R, *et al*. NEWS 2—too little evidence to implement? *Clin Med* 2018;18:371–3.
22. Pedersen NE, Rasmussen LS, Petersen JA, *et al*. Modifications of the national early warning score for patients with chronic respiratory disease. *Acta Anaesthesiol Scand* 2018;62:242–52.
23. Hodgson LE, Dimitrov BD, Congleton J, *et al*. A validation of the national early warning score to predict outcome in patients with COPD exacerbation. *Thorax* 2017;72:23–30.
24. Pimentel MAF, Redfern OC, Gerry S, *et al*. A comparison of the ability of the National Early Warning Score and the National Early Warning Score 2 to identify patients at risk of in-hospital mortality: a multi-centre database study. *Resuscitation* 2019;134:147–56.
25. O'Driscoll R, Bakerly N, Murphy P, *et al*. NEWS2 needs to be tested in prospective trials involving patients with confirmed hypercapnia. *Resuscitation* 2019;139:369–70.
26. Chu DK, Kim LH-Y, Young PJ, *et al*. Mortality and morbidity in acutely ill adults treated with liberal versus conservative oxygen therapy (IOTA): a systematic review and meta-analysis. *Lancet* 2018;391:1693–705.