BMJ Open Respiratory Research

Predictive value of the National Early Warning Score 2 for hospitalised patients with viral respiratory illness is improved by the addition of inspired oxygen fraction as a weighted variable

Jonathan Clarke,[1] Jack Gallifant ,[2] David Grant,[3] Nishita Desai,[1] Guy Glover[1]

To cite: Clarke J, Gallifant J, Grant D, *et al*. Predictive value of the National Early Warning Score 2 for hospitalised patients with viral respiratory illness is improved by the addition of inspired oxygen fraction as a weighted variable. *BMJ Open Respir Res* 2023;**10**:e001657. doi:10.1136/bmjresp-2023-001657

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjresp-2023-001657).

Received 6 February 2023
Accepted 30 November 2023

[1]Department of Critical Care, Guy's and St Thomas' NHS Foundation Trust, London, UK
[2]Department of Critical Care, Imperial College Healthcare NHS Trust, London, UK
[3]Department of Medical Informatics, Guy's and St Thomas' NHS Foundation Trust, London, UK

Correspondence to
Dr Guy Glover;
guy.glover@gstt.nhs.uk

ABSTRACT

Objectives The National Early Warning Score 2 (NEWS2) is validated for predicting acute deterioration, however, the binary grading of inspired oxygen fraction (FiO_2) may limit performance. We evaluated the incorporation of FiO_2 as a weighted categorical variable on NEWS2 prediction of patient deterioration.

Setting Two hospitals at a single medical centre, Guy's and St Thomas' NHS Foundation Trust.

Design Retrospective cohort of all ward admissions, with a viral respiratory infection (SARS-CoV-2/influenza).

Participants 3704 adult ward admissions were analysed between 01 January 2017 and 31 December 2021.

Methods The NEWS-FiO_2 score transformed FiO_2 into a weighted categorical variable, from 0 to 3 points, substituting the original 0/2 points. The primary outcome was a composite of cardiac arrest, unplanned critical care admission or death within 24 hours of the observation. Sensitivity, positive predictive value (PPV), number needed to evaluate (NNE) and area under the receiver operating characteristic curve (AUROC) were calculated. Failure analysis for the time from trigger to outcome was compared by log-rank test.

Results The mean age was 60.4±19.4 years, 52.6% were men, with a median Charlson Comorbidity of 0 (IQR 3). The primary outcome occurred in 493 (13.3%) patients, and the weighted FiO_2 score was strongly associated with the outcome (p=<0.001). In patients receiving supplemental oxygen, 78.5% of scores were reclassified correctly and the AUROC was 0.81 (95% CI 0.81 to 0.81) for NEWS-FiO_2 versus 0.77 (95% CI 0.77 to 0.77) for NEWS2. This improvement persisted in the whole cohort with a significantly higher failure rate for NEWS-FiO2 (p=<0.001). At the 5-point threshold, the PPV increased by 22.0% (NNE 6.7) for only a 3.9% decrease in sensitivity.

Conclusion Transforming FiO_2 into a weighted categorical variable improved NEWS2 prediction for patient deterioration, significantly improving the PPV. Prospective external validation is required before institutional implementation.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Previous work has demonstrated the significant limitations in the specificity of the National Early Warning Score 2 (NEWS2) score and the suitability of NEWS-inspired oxygen fraction (FiO_2) in patients receiving supplemental oxygen. However, this tool has not been generalised to all patients, which is crucial if this score is to be translated into clinical practice.

WHAT THIS STUDY ADDS

⇒ In a cohort of over 3700 patients, we demonstrate the benefit of transforming FiO_2 from a binary to a weighted categorical variable. The NEWS-FiO_2 better stratifies patient risk of deterioration and offers a clinically meaningful improvement in positive predictive value and number needed to evaluate, corresponding to roughly two fewer alarms per day, with a limited decrease in sensitivity.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This improvement provides evidence that NEWS-FiO_2 could be clinically operationalised using already collected clinical data. Our results have national and international importance in fulfilling the Royal College of Physicians' call for further research on this topic and laying the foundations for the introduction of NEWS-FiO_2 into clinical practice.

INTRODUCTION

Health agencies worldwide have advocated the benefits of early identification of the deteriorating patient for over two decades.[1 2] Similarly, calls for an in-hospital 'Chain of Prevention' model, mirroring that employed by the Resuscitation Council, have led to the adoption of both early warning scores (EWS) and rapid response systems (RRS) in hospitals.[1–4] The National Early Warning Score 2 (NEWS2) is in widespread use throughout the UK and internationally as the second iteration of a physiological aggregate weighted

track and trigger scoring system using six simple parameters: respiration rate, oxygen saturation, systolic blood pressure, pulse rate, level of consciousness (including new confusion) and temperature.[5] Each parameter is attributed a score based on the degree of variation from 'normal ranges', and the total score is augmented by two points if supplemental oxygen is required. The aggregate score is then used to trigger specific responses. A threshold of five points is defined as the point at which an urgent review by a clinician trained in managing the acutely ill patient is required. A score of 7 or more should prompt an emergency review by a team with critical care competencies.

NEWS2 has been validated for the prediction of, and clinical response to, the deteriorating patient.[6]

However, the current binary grading of inspired oxygen fraction (FiO_2) may limit the predictive power of NEWS2 by failing to provide additional weight as the patient deteriorates with rising oxygen requirements. This has been particularly pronounced during the COVID-19 pandemic, where patients frequently maintained stable parameters despite a marked increase in their FiO_2 requirements.[7 8] We have previously demonstrated that FiO_2 as a continuous variable had greater predictive validity than binary grading in patients with COVID-19.[9] This limitation has been acknowledged by the Royal College of Physicians, who subsequently issued additional guidance and called for further research into the impact of FiO_2 on NEWS2 performance.[10] In the 24 hours prior to an unplanned intensive care unit (ICU) admission, up to 75% of patients have been shown to have a new or increasing oxygen requirement.[11] Improving discrimination in oxygen administration could offer greater stratification of patient risk.

This failure to separate patients is problematic in patients with low supplemental oxygen requirements, for example, during the management of small deficits in gas exchange (eg, postoperative atelectasis) or as part of 'routine patient care', for example, during concurrent patient-controlled analgesia systems. These patients score the same two additional points as those receiving high FiO_2 via non-rebreather masks. This may increase the patient's overall NEWS2 score disproportionately compared with their true risk of an adverse event. Furthermore, this may contribute to one of the key limitations of NEWS2, namely the low positive predictive value[12] which is a function of the low prevalence of the outcome. This feature increases the rate of 'false alerts' and may contribute to alarm fatigue in clinical practice.[13]

One previous study demonstrated that a NEWS-FiO_2 model is associated with improved performance[14] but the FiO_2 thresholds developed have yet to be externally validated. This study addresses this gap by evaluating whether the incorporation of FiO_2 transformed into an additional weighed parameter in a NEWS-FiO_2 model improved the predictive validity in a cohort of patients with viral respiratory illness (COVID-19 and influenza).

METHODS

A retrospective observational cohort study was performed in two hospitals at a single medical centre (Guy's and St Thomas' NHS Foundation Trust, (GSTT)) between 01 January 2017 and 31 December 2021. The hospitals employ an electronic health record (EHR) system for recording physiological observations in general inpatient wards, with NEWS fully embedded in practice since 2012 and updated to NEWS2 in 2017. NEWS2 includes a modified peripheral oxygen saturation (SpO_2) scale for patients with chronic hypercapnic respiratory failure, formalising a modification that had been introduced with NEWS as a local exception in our system (the so called Chronic Respiratory Early Warning Score (CREWS)).[15] Observation frequency and escalation were protocolised in accordance with NEWS2 guidance. There is a dedicated 24/7 critical care outreach team (CCOT) staffed by advanced nurse practitioners with dedicated middle and senior-grade critical care doctors which is triggered by elevated NEWS2 as per national guidance. High flow nasal oxygen use is limited to specific wards with the support of the CCOT, and was not used outside of critical care in the COVID-19 period due to concerns about efficacy, oxygen resilience and infection control. Continuous positive airway pressure is not routinely used outside of critical care and therefore is a criterion for critical care admission.

The study cohort included all adult (≥17 years) patients admitted to general inpatient wards with a diagnosis of SARS-CoV-2 infection (COVID-19) or influenza, as two common causes of hypoxic respiratory failure. The COVID-19 cohort was defined by patients with a positive test by PCR from nasopharyngeal or throat swab, within 4 days of the index admission or in the 7 days prior to admission between 01 March 2020 and 30 March 2021. We used these dates to focus on patients in the first two waves of the COVID-19 pandemic only, as thereafter there was increasing heterogeneity due to patients being diagnosed with COVID-19 which was not the primary diagnosis or was not a cause of hypoxaemic respiratory failure. We excluded positive tests after 4 days to avoid confounding by patients who may have had a nosocomial acquisition and therefore for whom COVID-19 may not be the primary cause of illness. In addition to a COVID-19 cohort we included an influenza cohort, in recognition that the pathophysiology of COVID-19 may be novel and to test our hypothesis in another common cause of hypoxaemic respiratory failure. For the diagnosis of influenza, we included all adult (≥17 years) patients admitted to a general inpatient ward, with a discharge International Classification of Diseases 10th Revision (ICD-10) code (any coding position) of influenza (J09, J10, J11 or J12)—(excluding J12.8 and 12.9 if admitted after 1/1/20 as this was found to have been applied for emergency use in COVID-19 during the early phase of the pandemic) from 01 January 2017 (the onset of the study database) to 30 December 2021 inclusive. After excluding J12.8 and

Table 1 Baseline characteristics of the whole cohort, presented by the occurrence of the primary outcome

	All (3704 patients; 133 349 observations)	Primary outcome (493 patients; 17 864 observations)	No primary outcome (3211 patients; 115 485 observations)
Age (median (IQR))	61.2 (29.0)	68.8 (24.5)	60.2 (29.7)
Male gender (n (%))	1949 (52.6)	294 (59.6)	1655 (51.5)
Ethnicity (n (%))			
White	1666 (45.0)	246 (49.9)	1420 (44.2)
Black	861 (23.2)	113 (22.9)	748 (23.3)
Asian	227 (6.1)	33 (6.7)	194 (6.0)
Mixed/other	349 (9.4)	38 (7.7)	311 (9.7)
Not specified	601 (16.2)	63 (12.8)	538 (16.8)
Body mass index (median (IQR)) (missing 907, 24.5%)	26.0 (8.2)	25.5 (8.5)	26.1 (8.1)
Charlson Comorbidity (median (IQR))	0 (3)	1 (3)	0 (2)
Global Frailty Score (median (IQR))	0 (1.9)	1.6 (3.1)	0.0 (1.9)
Respiratory rate (min^{-1}) (median (IQR))	18 (17–20)	20 (18–24)	18 (17–20)
SpO_2 (%) (median (IQR))	96 (94–97)	95 (93–96)	96 (94–97)
Number of observations with supplemental oxygen (n (%))	55 751 (41.8)	12 667 (70.9)	43 084 (37.3)
FiO_2 (median (IQR))	0.21 (0.21–0.28)	0.21 (0.28–0.40)	0.21 (0.21–0.28)
FiO_2 in observations with supplemental oxygen (median (IQR))	0.28 (0.28–0.36)	0.35 (0.28–0.40)	0.28 (0.24–0.35)
Systolic blood pressure (mm Hg) (median (IQR))	125 (112–140)	126 (112–144)	124 (112–139)
Heart rate ((min^{-1}) (median (IQR))	80 (70–90)	85 (73–97)	80 (70–90)
Level of consciousness (n (%))			
Alert	132 349 (99.25)	17 323 (96.74)	115 026 (99.60)
New confusion	324 (0.24)	133 (0.74)	191 (0.17)
Verbal	528 (0.4)	329 (1.84)	199 (0.17)
Pain	122 (0.09)	66 (0.37%)	56 (0.05)
Unresponsive	26 (0.02)	13 (0.07)	13 (0.01)
Temperature °C (median (IQR))	36.5 (36.2–37.0)	36.7 (36.3–37.3)	36.5 (36.2–36.9)
NEWS2 score (median (IQR))	2 (1–4)	4 (2–6)	2 (1–3)
NEWS-FiO_2 score (median (IQR))	2 (1–3)	3 (2–6)	2 (1–3)

FiO_2, inspired oxygen fraction ; NEWS2, National Early Warning Score 2 .

J12.9 we did not identify any patient who fulfilled criteria for both COVID-19 and influenza cohorts.

The following criteria were applied for the exclusion of physiological observations: preceding the confirmation of COVID-19; erroneously attributed to have occurred after critical care admission (due to mistiming of administrative processes of ward transfer); occurring after the primary outcome event in either cohort; incomplete physiological observation set which hence could not generate a NEWS2 score.

The primary outcome was a composite of peri-arrest, cardiac arrest, unplanned critical care admission or death within 24 hours of the observation, in keeping with previous studies,[16] with each EWS value treated as an individual predictor. Hence, risk was assessed continually throughout the patient's length of stay. The time of the primary outcome was determined by the occurrence of the first event for patients who experienced more than one of the three individual outcomes. Accordingly, any subsequent events including critical care readmissions were not included. Peri-arrest or cardiac arrest were defined as a CCOT review coded as '2222 emergency' (2222 being the generic telephone number for an emergency call in UK hospitals), and that were followed by a critical care admission or death within 24 hours—a definition that had been previously internally validated (data unpublished). The timing of critical care admission was defined by the time of the recording of the first heart

rate to minimise errors associated with administrative processes.

Data is extracted from the GSTT Data Warehouse, which serves as an aggregate repository of data from multiple electronic sources, developed using the Health Catalyst Data Operating System (Health Catalyst, Salt Lake City, Utah, USA) and continuous since 01 January 2017. Variables included demographics, comorbidities, date and time of COVID-19 diagnosis, longitudinal physiological observations (respiratory rate, oxygen saturation, systolic blood pressure, pulse rate, level of consciousness, temperature and oxygen delivered), Charlson Comorbidity Index[17] and the Dr Foster Global Frailty Score.[18] Data on the use of the CREWS/scale 2 modification was not available, however, the NEWS score used already reflected the use of the modification in clinical practice.

FiO_2 was transformed into a weighted categorical variable and assigned 0–3 points as previously published,[14] substituting the original 0 points for no oxygen therapy or 2 points for those requiring any supplementary oxygen. In the proposed NEWS-FiO_2 model, 0 points were attributed to those patients not requiring supplemental oxygen, 1 point for an FiO_2 between 0.22 and 0.37, 2 points for an FiO_2 between 0.371 and 0.53 and 3 points for an FiO_2 of greater than 0.53. For fixed-performance oxygen devices, the set FiO_2 value was used, while for variable-performance devices the FiO_2 was estimated from the flow rate, as previously described[19] (online supplemental table S1). The impact of an alternative method of estimating FiO_2 was also tested, which attempts to account for the impact of the patient's respiratory minute volume on ambient air entrainment and hence on the FiO_2, referred to as the 'Bateman formula'.[20] This equation states that FiO_2=(O_2 flow rate+0.21 (minute volume–O_2 flow rate))/minute volume. The minute volume was estimated using a fixed tidal volume (450 mL) multiplied by the recorded respiratory rate as previously published.[14]

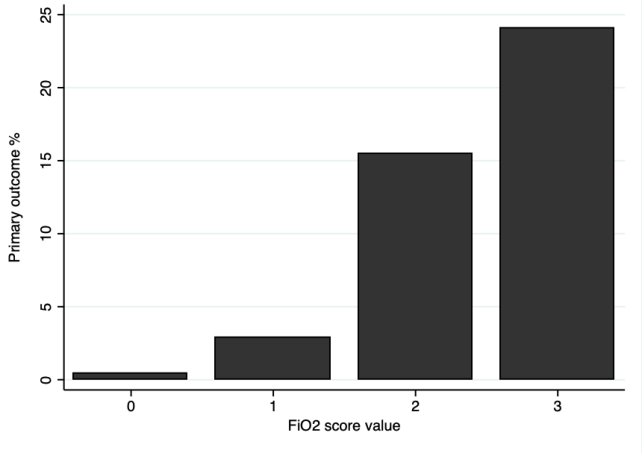

Figure 1 The rate of the primary outcome according to the value assigned to the FiO_2 in the NEWS-FiO_2 model. FiO_2, fractional inspired concentration of oxygen; NEWS2, National Early Warning Score 2.

Statistical methods

Data is presented as mean±SD, median (IQR) or number (%) and comparisons were performed using t-tests, Wilcoxon rank or χ^2 tests as appropriate. The rate of the primary outcome, stratified by the value score of the weighted FiO_2 parameter, was compared across groups with the χ^2 test. The change in score from NEWS2 to NEWS-FiO_2 is calculated, in addition to the change in classification of the primary outcome.

For NEWS2/NEWS-FiO_2, the area under the receiving operator characteristic curves (AUROC) were calculated with 95% CIs. However, due to the limitations of using the AUROC to assess the efficacy of an EWS,[21] we also report the positive predictive value (PPV) and numbers needed to evaluate (NNE) at each integer threshold of the EWS. The NNE is the number of patients who meet the threshold and who would need to be escalated to detect one true positive case, being the reciprocal of the PPV.[21]

Failure analysis was performed to compare time from meeting an EWS threshold to the occurrence of the primary outcome for the first 24 hours, making the following comparisons: (1) NEWS2 ≥5 versus NEWS-FiO_2 ≥5 and (2) NEWS-FiO_2<5 versus ≥5; time to event was compared using the log-rank test.

For all analyses, a threshold of $p<0.05$ was considered statistically significant. All analysis was performed in Stata (StataCorp 2013 Stata Statistical Software: Release 13). Strengthening the Reporting of Observational Studies in Epidemiology guidance for reporting cohort studies was followed (online supplemental file S2). Ethics approval was granted by the Research and Ethics Committee of GSTT (REC Number: 20/HRA/1871).

Patient and public involvement

Patients and/or the public did not take part in the development, conduct, reporting or dissemination of this study.

RESULTS

Study population

After exclusions, 3704 patients were included in this analysis, of which 2468 (66.6%) were in the COVID-19 cohort and 1236 (33.3%) were in the influenza cohort (online supplemental figure 1). In the influenza cohort, 70% had the code in the first coding position and 91% in the first to third coding position, indicating that viral illness was the primary cause, or at least a significant contributor to the admission. 6020 (4.3%) of potential observations sets were excluded as they were incomplete and hence there was no associated NEWS2 score (4876 (4.1%) for patients without the primary outcome versus 1144 (6.0%) for patients with the primary outcome), hence the total number of observation sets analysed was 133 349. Baseline characteristics are shown in table 1, presented by the occurrence of the primary outcome.

Table 2 Area under receiver operating characteristic curve with 95% CIs for the predictive validity for the primary outcome; all included observations (n=133 349)

	All	COVID-19	Influenza
NEWS2	0.87 (0.87 to 0.87)	0.86 (0.86 to 0.86)	0.88 (0.88 to 0.88)
NEWS-FiO_2 (predicted)	0.88 (0.88 to 0.88)	0.88 (0.87 to 0.88)	0.87 (0.87 to 0.87)
NEWS-FiO_2 (Bateman)	0.88 (0.88 to 0.88)	0.87 (0.87 to 0.88)	0.88 (0.87 to 0.88)

FiO_2, fractional inspired concentration of oxygen; NEWS2, National Early Warning Score 2 .

Of the 3704 patients included in this study, the median age on admission was 61.2 years with just over half being men. Over one-third of the patients were of black, Asian or mixed ethnicity and there was generally a low burden of comorbidity and frailty. Those who sustained the primary outcome event tended to be older, were more likely to be male and had higher Charlson Comorbidity and Global Frailty Scores (table 1). In addition, the observations in patients with the primary outcome were more abnormal, there was a greater proportion with supplemental oxygen and the EWS were numerically higher as compared with patients without the primary outcome.

Primary outcome

Overall, an adverse event occurred in 493 (13.3%) patients, of which 11 (2.2%) suffered a peri-arrest or cardiac arrest, 298 (60.4%) had an unplanned critical care admission and 184 (37.3%) died on the ward as their first event. The primary outcome occurred following 3536 (2.65%) of the observation sets.

NEWS2 versus NEWS-FiO_2

55 751 (41.8%) of the included data sets included supplemental oxygen and therefore scored two additional points on the NEWS2. Applying the NEWS-FiO_2 model to those patients with supplementary oxygen, led to a change in score follows: in 12.3% there was no change in the total score, in 78.5% the score decreased by one point and in 9.2% the score increased by one point. 43 743 (78.5%) EWS were reclassified correctly (42 514 (76.3%) decreased when the primary outcome did not occur; 1229 (2.2%) increased when the primary outcome did occur), as compared with 5133 (9.2%) EWS which were reclassified incorrectly (3890 (7.0%) increased when the primary outcome did not occur and 1243 (2.2%) decreased when the primary outcome did occur).

The weighted FiO_2 score was strongly associated with the primary outcome (figure 1); namely, increased oxygen requirements were correlated with an increased crude risk of a primary outcome event (0, 0.4%, 1, 3.1%, 2, 15.9%, 3, 25.2%; (p=<0.001).

Compared with NEWS2 the AUROC was greater for NEWS-FiO_2 (table 2). This improvement was particularly evident when the analysis was restricted to those patients receiving supplemental oxygen (table 3).

At the clinically and operationally relevant 5-point threshold, the specificity of NEWS-FiO_2 was 89.4% versus 85.8% for NEWS2. The PPV increased by 22% from 12.2 for NEWS2 to 14.9 for NEWS-FiO_2, and the NNE decreased from 8.2 to 6.7. This is balanced with a decrease of 3.9% in sensitivity, comparing NEWS2 to NEWS-FiO_2 (online supplemental tables S2 & S3).

Estimating the FiO_2 using the Bateman formula resulted in a decrease in the AUROC when compared with the predicted model (0.80 (95% CI 0.79 to 0.80) vs 0.81 (0.81 to 0.81)). At the 5-point threshold, the Bateman method resulted in a reduced PPV (12.8 vs 14.9), an increased NNE (7.8 vs 6.7) and a reduced specificity (86.6 vs 89.4) as compared with the predicted method.

Failure analysis

At the 5-point threshold, the failure rate was higher for patients with NEWS-FiO_2 than with NEWS2, indicating that the novel score portends higher risk (log-rank test p=<0.001) (figure 2). Additionally, when analysing solely the NEWS-FiO_2 scores, the ≥5 threshold shows discrimination, with the higher scores being associated with a markedly higher risk of the primary outcome in the following 24 hours. (Log-rank test p=<0.001) (figure 3).

Table 3 Area under receiver operating characteristic curve with 95% CIs for the predictive validity for the primary outcome; observations with supplemental oxygen only (n=55 751)

	All	COVID-19	Influenza
NEWS2	0.77 (0.77 to 0.77)	0.77 (0.76 to 0.77)	0.78 (0.77 to 0.79)
NEWS-FiO_2 (predicted)	0.81 (0.81 to 0.81)	0.81 (0.81 to 0.82)	0.80 (0.79 to 0.80)
NEWS-FiO_2 (Bateman)	0.80 (0.79 to 0.80)	0.80 (0.79 to 0.80)	0.79 (0.78 to 0.80)

FiO_2, fractional inspired concentration of oxygen; NEWS2, National Early Warning Score 2 .

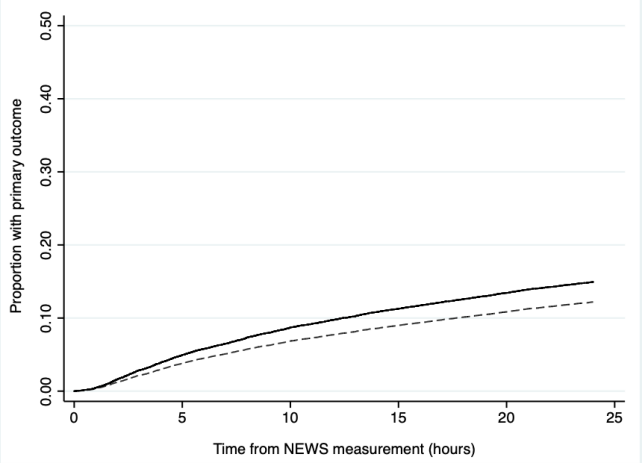

Figure 2 Failure analysis—the proportion of observations followed by the primary outcome within 24 hours; all early warning score values ≥5, comparing NEWS2 (dashed line) and NEW-FiO_2.(solid line) FiO_2, fractional inspired oxygen concentration; NEWS, National Early Warning Score.

DISCUSSION

In this cohort of patients admitted to a general inpatient ward with a diagnosis of either COVID-19 or influenza, 13.3% went on to experience a peri-arrest or cardiac arrest, unplanned critical care admission or death on the ward. This study demonstrates that the weighted FiO_2 parameter score is strongly associated with outcome and that incorporating this into a NEWS-FiO_2 model outperforms the existing NEWS2, with an increased PPV at limited cost to the overall sensitivity of the scoring system. This benefit holds true in two common viral respiratory illnesses. As expected, the effect is mediated through benefit in patients receiving supplemental oxygen, but is sufficiently robust to remain evident across the whole cohort, regardless of supplemental oxygen, in whom practically any novel EWS would be implemented. By changing from a pooled risk for all patients on oxygen regardless of FiO_2 to a stratified risk based on the FiO_2 requirement, the NEWS-FiO_2 score addresses two key limitations of NEWS2—namely, lack of PPV and the fact that it is not sensitive to deterioration in the SpO2-FiO_2 ratio. The fact that NEWS-FiO_2 confers a reduction in score in over three quarters of patients on supplemental oxygen suggests an improved predictive value of the NEWS-FiO_2 compared with NEWS2, primarily by minimising the score of those on low FiO_2 supplemental oxygen who are at lower risk of deterioration to the primary outcome.

The findings of our study have the potential to bring clinically relevant benefits. Given that the majority of patients in hospitals on supplemental oxygen actually have low FiO_2 requirements[22] and the majority of these are at low risk of deterioration, NEWS-FiO_2 appropriately decreased the EWS and, therefore the NNE. In our system, this translates to approximately one fewer referral of a patient meeting the 5-point threshold per 12-hour shift, thus increasing the efficiency of the RRS and reducing the likelihood of 'alarm fatigue'.[13] Equally, the increase in PPV means that the patient meeting the 5-point threshold for an urgent review, is more likely to genuinely warrant CCOT involvement as they are at a higher risk of deterioration, a fact supported by the higher failure rate after NEWS-FiO_2 ≥5, as compared with NEWS2. As previously discussed, patients generally have an increase in oxygen requirement as an antecedent to deterioration, and it is this population who have the greatest potential to benefit from employing NEWS-FiO_2.

Despite our findings, it is noted that the performance of NEWS2 in this study was similar to previous validation studies,[23] justifying the literature that has validated and recommended the use of NEWS2 in COVID-19.[24]

We have also demonstrated that the method of estimating FiO_2 makes a small but material difference to the performance of the scores. While the 'Bateman formula' attempts to account for the impact of the minute ventilation on the FiO_2 and may give a more accurate representation of the true FiO_2 for variable performance devices, it still relies on assumptions regarding the minute ventilation of the patient and would be significantly more challenging to operationalise without an EHR system to calculate this automatically. We opted to replicate a method previously reported, however acknowledge that the use of a fixed tidal volume is simplistic. While changing the tidal volume does inversely affect the estimated FiO_2, data of Malycha *et al*[14] demonstrates that the impact is likely to be limited. This component of our study demonstrates that the method of estimating FiO_2 is important in any future work on NEWS-FiO_2 and should be standardised.

The strengths of this study include the relatively large sample size as well as the large volume of observations included. The inclusion of a viral pneumonia other than COVID-19 is important as an indicator of validity

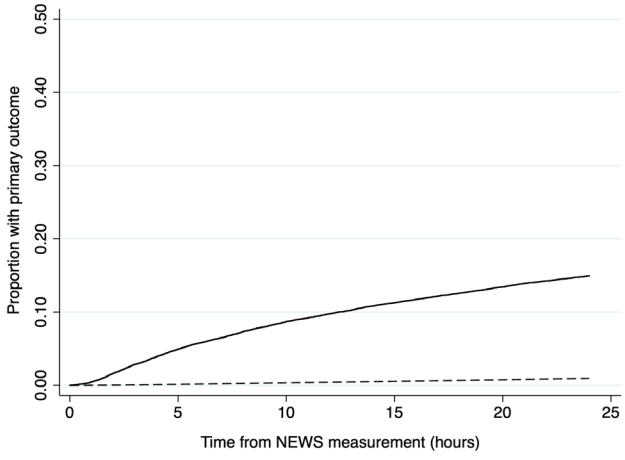

Figure 3 Failure analysis—the proportion of observations followed by the primary outcome within 24 hours; NEW-FiO_2 <5 (dashed line) versus ≥5 (solid line). FiO_2, fractional inspired oxygen concentration; NEWS, National Early Warning Score.

in a wider population of common respiratory illnesses. This study builds on an already validated EWS which is well-adopted in clinical practice, using data routinely recorded at the bedside. The addition of FiO_2 to the observations collected could be easily operationalised by minor modifications to any EHR system or a simple 'look-up' reference table for those recorded manually.

The limitations of our study are principally that it is a single-centre study, in an academic medical setting, performed in a retrospective fashion and patient inclusion was limited to those with a COVID-19 or influenza diagnosis. While this study adds external validity to the initial description of NEWS-FiO_2, it is not known if the findings are further generalisable across different healthcare settings, patient demographics and primary diagnosis.[25] This population had an uncommonly high rate of supplemental oxygen administration and thus, in a more general population, a statistically significant benefit may not be demonstrated across the whole cohort. Provided the burden of widespread implementation is limited, this ought not to discourage adoption for the benefits in the patients who do require supplemental oxygen. It is not known if the FiO_2 thresholds that have been previously published and which we used are optimal, however, a simple sensitivity analysis at different thresholds did not improve the discrimination between the FiO_2 parameter scores and the primary outcome. We have been unable to report the proportion of patients on a modified SpO_2 scale, however because we used the NEWS2 score which already reflected the modified scale and because any bias related to a reduced FiO_2 administration will have equally affected both the NEWS2 and NEWS-FiO_2, this ought not to have affected the results. A complete case analysis approach was used which is justifiable as the rate of missing data was low and similarly distributed with the occurrence of the primary outcome. The fact that the findings are consistent across both the COVID-19 and influenza cohorts suggests that there is a limited impact of pandemic specific factors such as organisational strain, but this cannot be excluded.

Further work is needed to prospectively validate this model in an unselected hospital population, across multiple sites and in different demographic groups. In particular, the impact of bias in pulse oximetry and hence oxygen administration for patients with darker skin tones should be evaluated.[26]

CONCLUSION

Incorporation of FiO_2 into NEWS2 improves predictive validity for adverse events for hospitalised patients with viral respiratory illness, particularly by improving the PPV. This benefit was mediated within those patients receiving oxygen supplementation and held true in both the COVID-19 and influenza cohorts. These findings require prospective validation across a more generalised population before consideration of inclusion in future iterations of NEWS2.

Acknowledgements We wish to acknowledge the support of Leah Evans BS, Kathleen Merkley DNP APRN ANP and Samantha Johnson MBA, of Health Catalyst, Salt Lake City, Utah, USA, for developing the HealthCatalyst Data Operating System (DOS) used for this study. We would also like to acknowledge all of the clinicians and support staff who cared for patients during the COVID-19 pandemic.

Contributors JC: conceptualisation, methodology, writing—original draft. JG: conceptualisation, methodology, writing—review and editing. DG: conceptualisation, methodology, validation, data curation. ND: conceptualisation, methodology, writing—review and editing. GG: conceptualisation, methodology, supervision, formal analysis, data curation writing, data visualisation—review and editing, guarantor. DG and GG have verified the underlying data. All authors have read and finally approved the version being submitted. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/disclosure-of-interest/ and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iD

Jack Gallifant http://orcid.org/0000-0003-1306-2334

REFERENCES

1 National Patient Safety Agency. Recognising and responding appropriately to early signs of deterioration in hospitalised patients. London NPSA; 2007.
2 National confidential enquiry into patient outcome and death. emergency admissions: a journey in the right direction. London NCEPOD; 2007.
3 Smith GB. In-hospital cardiac arrest: is it time for an in-hospital 'chain of prevention *Resuscitation* 2010;81:1209–11.
4 Hughes C, Pain C, Braithwaite J, *et al*. Between the flags': implementing a rapid response system at scale. *BMJ Qual Saf* 2014;23:714–7.
5 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.
6 Smith GB, Prytherch DR, Meredith P, *et al*. The ability of the national early warning score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. *Resuscitation* 2013;84:465–70.
7 Pimentel MAF, Redfern OC, Hatch R, *et al*. Trajectories of vital signs in patients with COVID-19. *Resuscitation* 2020;156:99–106.
8 Baker K, Hanrath A, van der Loeff I, *et al*. National early warning score 2 (News2) to identify inpatient COVID-19 deterioration: a retrospective analysis. *Clin Med* 2021;2:84–9.
9 Prower E, Grant D, Bisquera A, *et al*. The ROX index has greater predictive validity than News2 for deterioration in COVID-19. *EClinicalMedicine* 2021;35:100828.

10 NEWS2 and deterioration in Covid-19, Available: https://www.rcplondon.ac.uk/news/news2-and-deterioration-covid-19 [Accessed 1 Mar 2023].
11 Goldhill DR, White SA, Sumner A. Physiological values and procedures in the 24 H before ICU admission from the ward. *Anaesthesia* 1999;54:529–34.
12 Covino M, Sandroni C, Santoro M, *et al*. Predicting intensive care unit admission and death for COVID-19 patients in the emergency department using early warning scores. *Resuscitation* 2020;156:84–91.
13 Hannibal GB. Monitor alarms and alarm fatigue. *AACN Adv Crit Care* 2011;22:418–20.
14 Malycha J, Farajidavar N, Pimentel MAF, *et al*. The effect of fractional inspired oxygen concentration on early warning score performance: a database analysis. *Resuscitation* 2019;139:192–9.
15 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.
16 Alam N, Hobbelink EL, van Tienhoven AJ, *et al*. The impact of the use of the early warning score (EWS) on patient outcomes: a systematic review. *Resuscitation* 2014;85:587–94.
17 Charlson ME, Pompei P, Ales KL, *et al*. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83.
18 Soong JTY, Kaubryte J, Liew D, *et al*. Dr foster global frailty score: an international retrospective observational study developing and validating a risk prediction model for hospitalised older persons from administrative data SETS. *BMJ Open* 2019;9:e026759.
19 Fuentes S, Chowdhury YS. Fraction of inspired oxygen. Updated 2022 Feb 17. StatPearls Publishing; 2022. Available: https://www.ncbi.nlm.nih.gov/books/NBK560867
20 Bateman N, Leach R. Acute oxygen therapy. *BMJ* 1998;317:798.
21 Romero-Brufau S, Huddleston JM, Escobar GJ, *et al*. Why the C-Statistic is not informative to evaluate early warning scores and what metrics to use. *Crit Care* 2015;19:285.
22 O'Driscoll BR. *British Thoracic Society Emergency Oxygen Audit Report 2015*. London: British Thoracic Society, 2015: 8.
23 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.
24 Myrstad M, Ihle-Hansen H, Tveita AA, *et al*. National early warning score 2 (News2) on admission predicts severe disease and in-hospital mortality from COVID-19 - a prospective cohort study. *Scand J Trauma Resusc Emerg Med* 2020;28:66.
25 Shamout F, Zhu T, Clifton L, *et al*. Early warning score adjusted for age to predict the composite outcome of mortality, cardiac arrest or unplanned intensive care unit admission using observational vital-sign data: a multicentre development and validation. *BMJ Open* 2019;9:e033301.
26 Sjoding MW, Dickson RP, Iwashyna TJ, *et al*. Racial bias in pulse oximetry measurement. *N Engl J Med* 2020;383:2477–8.

SUPPLEMENTAL DATA

Device	Fixed performance FiO_2	Flow l/min	Predicted FiO_2
Nasal cannula		1	24
Nasal cannula		2	28
Nasal cannula		3	32
Nasal cannula		4	36
Simple mask		1	24
Simple mask		2	28
Simple mask		3	32
Simple mask		4	36
Simple mask		5	40
Simple mask		6	44
Simple mask		7	48
Simple mask		8	52
Simple mask		9	56
Simple mask		>=10	60
Venturi	24	2	
Venturi	28	4	
Venturi	35	8	
Venturi	40	10	
Venturi	60	12	
NIV OR CPAP		1	24
NIV OR CPAP		2	26
NIV OR CPAP		3	28
NIV OR CPAP		4	30
NIV OR CPAP		5	32
NIV OR CPAP		6	34
NIV OR CPAP		7	36
NIV OR CPAP		8	38
NIV OR CPAP		9	40
NIV OR CPAP		>=10	44
Humidified	28		
Humidified	35		
Humidified	40		
Humidified	60		
Tracheostomy mask	Treat as per simple face mask		
HFNC	Continuous % O_2 values from 21 - 100	Continuous flow rate values from 0 - 60	
100% Reservoir Mask 15l/min		15	80

Table S1. Assigned FiO_2 value for fixed and variable performance oxygen delivery devices

Cut point (GTE)	Sensitivity	Specificity	PPV	NPV	NNE	Likelihood ratio +'ve	Likelihood ratio –'ve
0	100.00	0.00	2.65		37.74	1.00	
1	98.98	20.32	3.27	99.86	30.57	1.24	0.05
2	97.40	40.25	4.25	99.82	23.54	1.63	0.06
3	92.99	59.20	5.84	99.68	17.12	2.28	0.12
4	83.96	74.53	8.24	99.42	12.14	3.30	0.22
5	**72.31**	**85.83**	**12.20**	**99.13**	**8.20**	**5.10**	**0.32**
6	54.52	91.88	15.46	98.67	6.47	6.72	0.49
7	**37.53**	**95.75**	**19.39**	**98.25**	**5.16**	**8.84**	**0.65**
8	24.41	97.81	23.29	97.94	4.29	11.16	0.77
9	12.22	99.03	25.59	97.64	3.91	12.64	0.89
10	6.11	99.58	28.15	97.50	3.55	14.39	0.94
11	3.22	99.81	32.10	97.43	3.12	17.37	0.97
12	1.58	99.92	34.98	97.39	2.86	19.77	0.99
13	0.74	99.97	38.22	97.37	2.62	22.73	0.99
14	0.34	99.98	35.28	97.36	2.83	20.03	1.0
15	0.20	100	69.99	97.36	1.43	85.64	1.0
18	0.03	100	100.0	97.35	1.00		1.0

Table S2. Sensitivity, specificity, PPV, NPV and NNE for NEWS2 for whole cohort for the primary outcome NB. Prevalence = 2.65%. Youden calculation: maximum value 58.49 at NEWS >=4. ***PPV, positive predictive value; NPV, negative predictive value; NNE, number needed to evaluate; LR, likelihood ratio; GTE, greater than or equal to (*****n=133349)**

Cut point (GTE)	Sensitivity	Specificity	PPV	NPV	NNE	Likelihood ratio +'ve	Likelihood ratio –'ve
0	100	0.30	2.65	100	37.74	1.00	
1	98.98	20.71	3.28	99.87	30.52	1.24	0.05
2	95.33	46.87	4.64	99.73	21.53	1.79	0.10
3	88.80	67.67	6.94	99.55	14.41	2.74	0.17
4	80.49	81.91	10.77	99.36	9.28	4.45	0.24
5	**68.38**	**89.39**	**14.89**	**99.05**	**6.72**	**6.45**	**0.35**
6	55.60	93.88	19.78	98.73	5.05	9.09	0.47

7	**40.70**	**96.55**	**24.25**	**98.36**	**4.12**	**11.79**	**0.61**
8	26.64	98.13	27.87	98.01	3.59	14.23	0.75
9	15.50	99.05	30.73	97.73	3.25	16.34	0.85
10	7.86	99.54	31.84	97.54	3.14	17.21	0.93
11	3.99	99.80	35.08	97.45	2.85	19.90	0.96
12	2.01	99.91	37.88	97.40	2.64	22.47	0.98
13	1.05	99.95	35.49	97.38	2.82	20.27	0.99
14	0.59	99.98	45.56	97.36	2.19	30.84	0.99
15	0.17	99.99	33.25	97.35	3.01	18.35	1.00
16	0.08	100	74.93	97.35	1.33	119.10	1.00
17	0.03	100	100.00	97.35	1.00		1.00

Table S3. Sensitivity, specificity, PPV, NPV and NNE for NEWS-FiO_2 for whole cohort for the primary outcome. NB. Prevalence = 2.65%. Youden calculation: maximum value 62.34 at NEWS >=4. ***PPV, positive predictive value; NPV, negative predictive value; NNE, number needed to evaluate; LR, likelihood ratio; GTE, greater than or equal to (*****n=133349)**

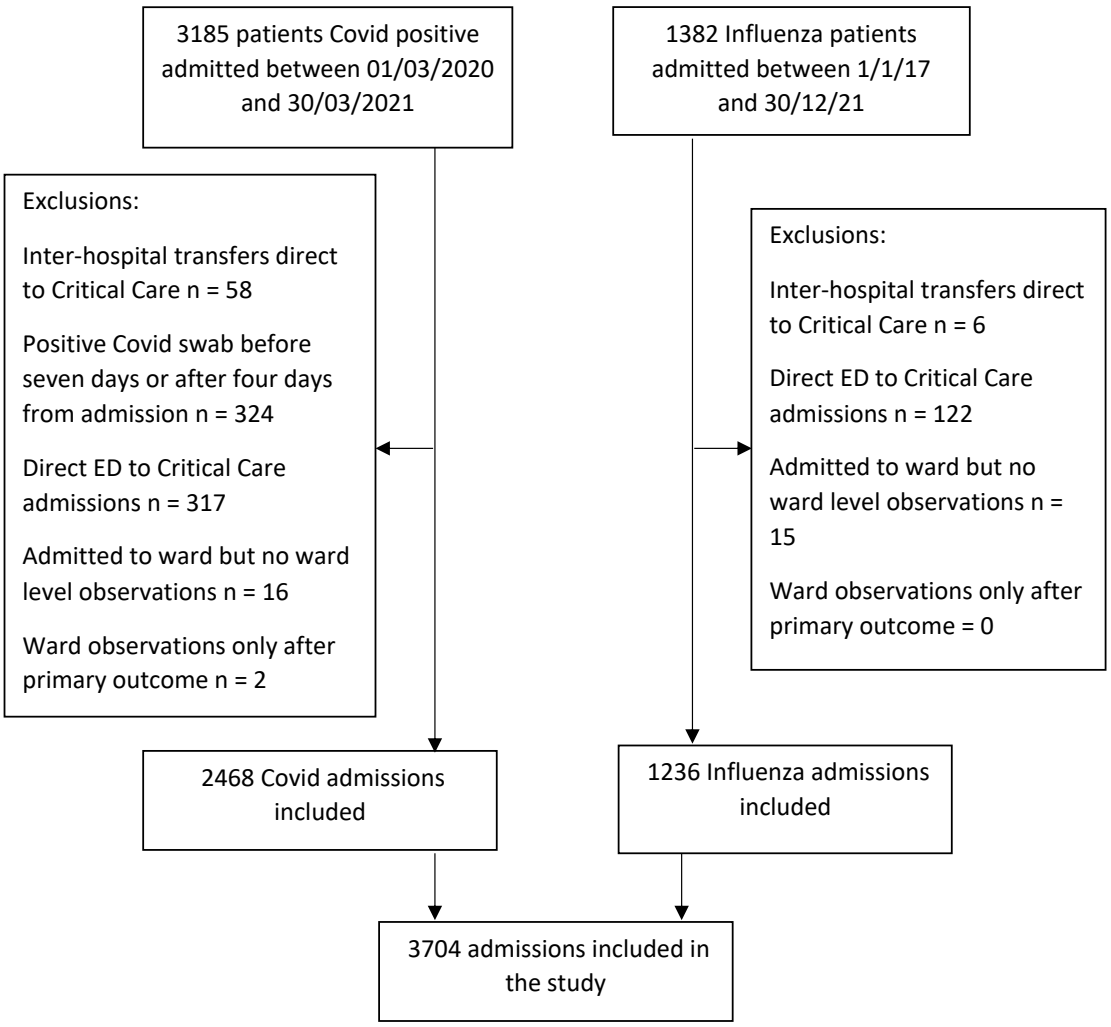

Figure S1. Flow diagram of patient study population