



The ROX index has greater predictive validity than NEWS2 for deterioration in Covid-19

Emma Prower^a, David Grant^b, Alessandra Bisquera^c, Cormac P Breen^d, Luigi Camporota^e, Maja Gavrilovski^f, Megan Pontin^g, Abdel Douiri^h, Guy W Glover^{i,*}

^a Department of Critical Care, Kings College Hospital, Denmark Hill, London SE5 9RS, UK

^b Department of clinical informatics for health informatics, Guys and St Thomas NHS Foundation Trust, Westminster Bridge Road, London SE1 7EH, UK

^c Department of Primary Care and Public Health Sciences, Kings College London, Guy's Campus, Addison House, London SE1 1UL, UK

^d Department of Nephrology, Guys and St Thomas NHS Foundation Trust, Westminster Bridge Road, London SE1 7EH, UK

^e Department of Critical Care, Guys and St Thomas NHS Foundation Trust, Westminster Bridge Road, London SE1 7EH, UK

^f Department of Emergency Medicine, Guys and St Thomas NHS Foundation Trust, Westminster Bridge Road, London, SE1 7EH, UK

^g Department of Quality and Assurance, Guy's and St Thomas NHS Foundation Trust, Westminster Bridge Road, London, SE1 7EH, UK

^h Department of Primary Care and Public Health Sciences, Kings College London, Guy's Campus, Addison House, London, SE1 1UL, UK

ⁱ Department of Critical Care, Guys and St Thomas NHS Foundation Trust, Westminster Bridge Road, London, SE1 7EH, UK

ARTICLE INFO

Article History:

Received 11 December 2020

Revised 8 March 2021

Accepted 23 March 2021

Available online xxx

Keywords:

Covid-19

Coronavirus

NEWS2

ROX

Rapid response systems

Predicting deterioration

ABSTRACT

Background: Patients admitted to hospital with Covid-19 are at risk of deterioration. The National Early Warning Score (NEWS2) is widely recommended, however its validity in Covid-19 is not established and indices more specific for respiratory failure may be more appropriate. We aim to describe the physiological antecedents to deterioration, test the predictive validity of NEWS2 and compare this to the ROX index ($[\text{SpO}_2/\text{FiO}_2]/\text{respiratory rate}$).

Method: A single centre retrospective cohort study of adult patients who were admitted to a medical ward, between 1/3/20 and 30/5/20, with positive results for SARS-CoV-2 RNA. Physiological observations and the NEWS2 were extracted and analysed. The primary outcome was a composite of cardiac arrest, unplanned critical care admission or death within 24 hours. A generalized linear model was used to assess the association of physiological values, NEWS2 and ROX with the outcome.

Findings: The primary outcome occurred in 186 patients (26%). In the preceding 24 hours, deterioration was most marked in respiratory parameters, specifically in escalating oxygen requirement; tachypnoea was a late sign, whilst cardiovascular observations remained stable. The area under the receiver operating curve was 0.815 (95% CI 0.804–0.826) for NEWS2 and 0.848 (95% CI 0.837–0.858) for ROX. Applying the optimal level of ROX, the majority of patients triggered four hours earlier than with NEWS2 of 5.

Interpretation: NEWS2 may under-perform in Covid-19 due to intrinsic limitations of the design and the unique pathophysiology of the disease. A simple index utilising respiratory parameters can outperform NEWS2 in predicting the occurrence of adverse events.

© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

1. Introduction

The SARS-CoV-2 (Covid-19) pandemic has put strain on health-care systems across the globe. In the UK, between 23rd March and 2nd July 2020, there were 125,315 Covid-19 patients admitted to United Kingdom National Health Service (NHS) hospitals. At the peak there were over 3000 admissions per day with 15,700 patients

admitted in acute and general wards, and more than 3,100 ventilated Covid-19 patients [1,2].

Additionally, the severity of the disease amongst hospitalised patients presents a major challenge. In the International Severe Acute Respiratory and Emerging Infections Consortium (ISARIC) data set of more than 20000 UK cases, 55% required high flow oxygen and 17% patients were admitted to critical care (CC) [3]. In this context, there is a pressing need to evaluate current escalation pathways. The challenge is to find innovative ways to optimise identification of deterioration with Covid-19, whilst working within resource-constrained systems due to the pandemic.

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

* Corresponding author.

E-mail address: guy.glover@gstt.nhs.uk (G.W. Glover).

<https://doi.org/10.1016/j.eclinm.2021.100828>

2589-5370/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Research in context

Evidence before this study

NEWS2 is an established system for detection of the deteriorating patient, however it has not been fully validated in Covid-19. Specifically, there is a need to understand the impact of the binary grading system for oxygen requirement, which has been postulated to be a potential limitation. Current data demonstrates that hypoxaemia is a very characteristic feature of Covid-19 and therefore a system more specific for respiratory failure may be more appropriate; the ROX index ($[\text{SpO}_2/\text{FiO}_2]/\text{respiratory rate}$) has previously been demonstrated to be predictive of the need for intubation in a cohort of patients with hypoxic respiratory failure.

Added value of this study

This study demonstrates that in the 24 hours preceding cardiac arrest, critical care admission or death, deterioration is most marked in respiratory parameters, specifically in escalating oxygen requirement; tachypnoea is a late sign, whilst cardiovascular observations remain stable. The predictive validity of NEWS2 is lower than for the ROX index and applying the optimal level of ROX, the majority of patients triggered four hours earlier than with a NEWS2 of five.

Implications of all the available evidence

The available evidence suggests that NEWS2, when used as a continuous track and trigger system has good discrimination for the prediction of adverse events, however it may not be the optimal early warning score given the particular physiology in Covid-19. The ROX index may be derived from similar physiological data and holds promise as an earlier marker of deterioration. In addition, the potential benefits of adding O_2 as a continuous variable to the NEWS2 should be modelled and tested in future studies.

predictive validity of NEWS2 in predicting an adverse outcome and to compare this to the ROX index, in adult patients with Covid-19.

2. Method

A retrospective observational cohort study at an academic medical centre was conducted between 1/3/20 and 30/5/20. The hospital has a fully deployed NEWS2 protocol and critical care outreach team (CCOT). High Flow Nasal Oxygen was not utilised in line with available guidance [20,21]. Continuous Positive Airway Pressure (CPAP) / Non-Invasive Ventilation (NIV) was provided only in CC and was limited due to safety concerns. An awake proning protocol [22] and a strategy of early invasive mechanical ventilation (IMV) was employed [23]. SpO_2 targets were reduced to 92–96% in line with NHS England guidance [24].

Patients were tested for Covid-19 at presentation in the Emergency Department (ED). Patients who achieved target SpO_2 with $\text{FiO}_2 \leq 0.4$ and $\text{RR} \leq 24/\text{minute}$ were generally admitted to a medical ward. Those who had $\text{FiO}_2 \geq 60\%$ or $\text{RR} \geq 25$ were escalated to the CCOT for review unless there were pre-existing limitation of therapy orders.

The case definition was adult patients (≥ 17 years) who were admitted to a medical ward, with positive results for SARS-CoV-2 RNA using multiplexed-tandem polymerase chain reaction, performed on nasopharyngeal and throat swab. We included cases where the Covid-19 swab was taken up to 7 days prior to the admission (due to the prevalence of community testing or testing in ED in patients who were subsequently discharged, but then readmitted, and because re-testing was not always routine upon admission when the patient was already known to be Covid-19 positive) and within the first 4 days of admission to hospital. 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. Only physiological observations that were taken on the ward were utilised. We excluded observations made before the swab was taken (in order to be certain the physiology recorded was related to the Covid-19 diagnosis), or after the primary outcome had occurred. A convenience sample was utilised between the dates of 1/3/20 and 30/5/20, reflecting the main surge of Covid-19 admissions to our hospital during the first wave of the pandemic.

The primary outcome was a composite of peri/cardiac arrest, unplanned CC admission or death, in line with previous studies of early warning scores [4]. Secondary outcomes were designed to distinguish patients who were 'for full escalation of therapy' (for which we used the composite of peri/cardiac arrest OR unplanned CC admission) as compared to patients who had ward-based limitations of therapy who would not receive either of the above escalations and for whom death on the ward was the most relevant secondary outcome. Finally we analysed the requirement for advanced respiratory support (CPAP, NIV or IMV) in recognition of the fact that admission to critical care may occur at different thresholds and can have limitations as a surrogate of severe illness. For all repeated measures analyses, a patient was determined to have an outcome if it occurred within 24 hours of the relevant observation set. An example of data used for 1 patient is shown in the supplementary appendix (Supplementary Figure 1).

Data was extracted from the electronic medical record utilising an enterprise data warehouse including demographics, comorbidities, date and time of Covid-19 diagnosis, longitudinal physiological observations (FiO_2 , SpO_2 , respiratory rate, heart rate, systolic blood pressure, level of consciousness and temperature) and clinical outcomes. FiO_2 was treated as a continuous variable but was also dichotomised into whether supplementary oxygen was being administered or not. Because it was not possible to measure the true value of FiO_2 , a reference conversion table was utilised to estimate the FiO_2 from the device and oxygen flow rate (Supplementary Figure 2). Level of

NEWS2 is an established system for detection of the deteriorating patient, in use across the NHS since 2012 and adopted worldwide [4,5,6]. It was designed to standardise grading of severity of acute illness, and has been shown to predict adverse outcomes.⁷ Initial reports of NEWS2 in Covid-19 have been limited by small samples size and / or have been restricted to the use of physiology on admission only, which does not reflect the use of a 'track and trigger' system in practice. [8–13]. Whilst the NEWS2 has recently been well validated in the setting of its original derivation [14], there is acknowledgement that it requires testing in other settings, and specifically to understand the impact of the binary grading system for oxygen requirement, which has been postulated to be a potential limitation due to the unique pathophysiology of the disease [15]. Whilst attempts to create prognostic tools for use in Covid-19 have recognised fraction of inspired concentration of oxygen (FiO_2) as a crucial component [16,17], previous attempts to add FiO_2 to the NEWS in a general population of hospitalised patients on oxygen have led to only small increases in predictive validity [18].

The ROX index (the ratio of peripheral oxygen saturation (SpO_2) to FiO_2 , divided by the respiratory rate (RR)) has previously been shown to predict the need for intubation in a cohort of patients with hypoxaemic respiratory failure [19]. Due to the predominant respiratory features of Covid-19, it is possible that the ROX index may have validity in predicting adverse outcomes. The primary aim of this study is to describe the physiological antecedents to deterioration, to test the

consciousness used a modified five-point ordinal scale: alert, new confusion, response to voice, responsive to pain, unresponsive ('ACVPU'). Co-morbidity was assessed using the Elixhauser co-morbidity index and frailty was assessed using the Dr Foster Global Frailty Score [25,26]. Time of critical care admission was defined as the time of the first recorded heart rate observation, to avoid errors associated with administrative processes.

Categorical data is presented as count (%) and continuous data is reported as median and interquartile range [IQR] or mean and standard deviation (SD) according to the distribution. The worst physiological values were compared between groups, and the physiological trend was evaluated for the 24 hours prior to the outcome occurring. Generalized linear mixed models were used to model the odds of each outcome, specifying a binomial family, with a random intercept to account for the clustering of observations within the same patient. The physiological values were normalised using Z scores to be able to make direct comparisons. A sensitivity analysis was performed adjusting for clinically relevant baseline characteristics. Optimal sensitivity and specificity was identified using Youden's Index. Observation sets with any missing parameters did not generate a NEWS2 score and therefore were excluded (4%). For FiO_2 , oxygen administered was recorded as 'Null' in 20.4% of observation sets. Commonly this occurred in a sequence of values of FiO_2 0.21, suggesting that in many instances no value was recorded as the patient was not receiving supplemental oxygen. Therefore, for FiO_2 only, data was imputed using last value carried forward (LVCF), but restricted to looking back only one observation set. If a subsequent row was also 'Null' for FiO_2 then this data was treated as missing. Sensitivity analysis was also performed without FiO_2 imputation. Statistical analysis was performed using R (version 4.0.0, <http://www.r-project.org>) using the pROC and lme4 packages. All tests were two sided and a p value of

<0.05 was considered significant. TRIPOD guidance for prediction model validation, and STROBE guidance for reporting cohort studies was followed.

This study was approved by the Ethics Committee of Guys and St Thomas' NHS Trust (REC Number: 20/HRA/1871) and the requirement for written informed consent was waived as the study used anonymised and routinely collected data. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

2.1. Role of Funders

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

3. Results

Between 1/3/20 and 30/5/20, 1122 patients were identified as being positive for Covid-19. Subsequent exclusions are shown in Figure 1, leaving 708 admissions which were included in the analysis (accounting for 693 patients with 15 re-admissions). The median time from admission to the Covid-19 swab was 0.6 [0.2–0.8] days with only a minority (12%) of cases having positive swabs identified prior to hospital admission. The total number of observation sets analysed was 29783. The number of each physiological value available for analysis, before and after imputation are included in Table 2.

Baseline characteristics are shown in Table 1.

For the 708 patients included in the study, mean (SD) age on admission was 62.2 (18.1) years and the majority were male. The population was overweight and almost two fifths had diabetes. Of the 603 patients who had comorbidity status recorded, all had at least

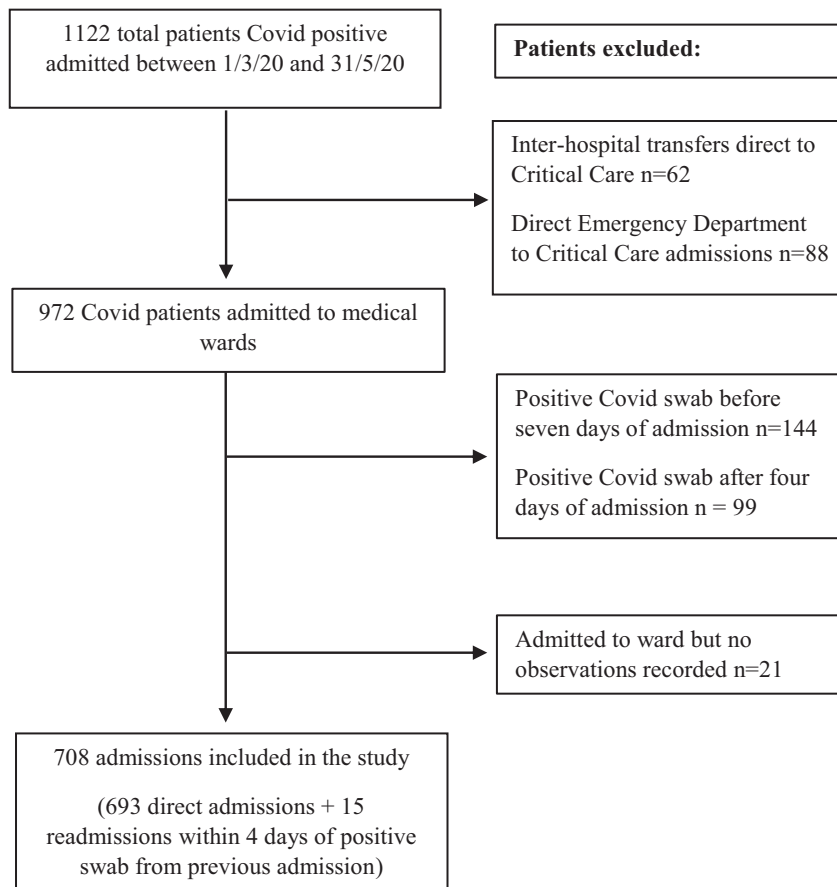


Figure 1. Flow diagram for the study cohort.

Table 1
Baseline characteristics of cohort presented as overall and by the occurrence of the primary outcome.

	Overall n=708	Any adverse event n=186	No adverse event n=522	Difference (95% CI)
Age (Mean (SD))	62.2 (18.1)	67.1 (16.8)	60.4 (18.3)	6.7 (3.8,9.6)
Male gender (n [%])	408 [57.6]	116 [62.4]	292 [55.9]	6.4 (-2.1, 14.9)
Ethnicity (n [%])				
White	288[40.7]	86 [46.2]	202 [38.7]	7.5 (-1.1, 16.2)
Black	196[27.7]	46 [24.7]	150 [28.7]	-4.0 (-11.7, 3.7)
Asian	55 [7.8]	19 [10.2]	36 [6.9]	3.3 (-1.9, 8.5)
Mixed/Other	51[7.2]	16 [8.6]	35 [6.7]	1.9 (-3.0, 6.8)
Not specified	118 [16.7]	19 [10.2]	99 [19.0]	-8.8 (-14.6, -2.9)
Body Mass Index (Mean (SD))	27.4 (6.8)	27.5 (6.7)	27.3 (6.9)	0.2 (-0.9, 1.3)
Elixhauser Comorbidity (n [%])				
Total condition count of 10+	53 [7.5]	20 [10.8]	33 [6.3]	4.4 (-0.8, 9.7)
Total condition count of 7-9	82 [11.6]	20 [10.8]	62 [11.9]	-1.1 (-6.7, 4.5)
Total condition count of 4-6	158 [22.3]	46 [24.7]	112 [21.5]	3.3 (-4.2, 10.8)
Total condition count of 1-3	310 [43.8]	64 [34.4]	246 [47.1]	-12.7 (-21.1, -4.3)
Missing	105 [14.8]	36 [19.4]	69 [13.2]	6.1 (-0.6, 12.9)
Diabetes (n [%])	272 [38.4]	72 [38.7]	200 [38.3]	0.4 (-8.1, 8.9)
Global Frailty Score (Mean (SD))	3.5 (2.2)	3.7 (2.1)	3.5 (2.3)	0.2 (-0.2,0.6)

one comorbidity. Half the patients were from black, asian or mixed racial backgrounds. Hospital length of stay was 6.8 [3.6 - 13.6] days and overall hospital mortality for this cohort was 15%. Patients who suffered the primary outcome were older (67.1 (16.8) years versus 60.4 (18.3) years, mean difference 6.7 years (95% CI 3.8 - 9.6)), however other baseline characteristics were similar. The group whose first adverse event was peri/cardiac arrest or CC admission were younger and with lower co-morbidity and frailty as compared to those who died on the ward. (Table S1)

Overall, the composite primary outcome occurred in 186/708 (26%) admissions, which were peri/cardiac arrest (5, 2.7%), admission to CC (98, 52.7%) or death on a ward (83, 44.6%). Of those admitted to CC, 77 (78.6%) received IMV and 4 (0.6%) received NIV. Time from hospital admission to critical care admission was 2.2 [1.2 - 3.2] days and time to death as the first adverse event was 8.0 [4.0 - 10.0] days.

3.1. Physiology and trends prior to adverse events

The most deranged physiological values for the group that had the primary outcome and for that where the patient's stay was uncomplicated are shown in Table 2

Trends in physiological deterioration in the final 24 hours prior to the adverse event occurring are shown in Fig 2. There is a notable inflection in the FiO₂ up to 12 hours before the event, whilst the RR did not increase until approximately 5 hours before. SpO₂ was generally well maintained and the cardiovascular observations remain relatively constant with no clear indication of the impending adverse event.

3.2. Associations with primary and secondary outcomes

The odds ratio for the occurrence of the primary outcome for all standardised physiological variables is shown in Table 3, both unadjusted, and adjusted for baseline variables, age, gender, ethnicity and body mass index (BMI).

A standardised unit increase in NEWS2 is associated with a significant increase in odds of having an adverse event within 24 hours, however the odds ratio for the ROX index is more than three times greater (2.43 (95% CI 2.15-2.74) versus 7.87 (95% CI 6.31-9.82)). The predominant physiological variables exerting this influence are the FiO₂, with a significant but lesser influence from SpO₂, RR and HR. An increase in temperature was associated with a reduced odds of an adverse event. Additional data for associations with baseline characteristics, absolute unit changes in physiological variables, including

analysis for the secondary outcomes are shown in the Supplementary Tables S2, S3, S4, and S5, where similar associations are demonstrated. None of the baseline characteristics that we tested were statistically associated with either the primary or secondary outcomes, and sensitivity analysis generated similar co-efficients.

3.3. Sensitivity and specificity for predicting first adverse events

The cut-point with the optimal sensitivity and specificity for NEWS2 is 4.5 with sensitivity of 0.72 and specificity of 0.79 (Table 4).

For the ROX index, optimal sensitivity (0.85) and specificity (0.75) for the primary outcome occurs at a value of 14.91, however only integer values are shown.

The area under the receiving operating characteristics curve (AUROC) for the occurrence of the primary outcome within 24 hours is 0.815 (95% CI 0.804-0.826) for NEWS2 and 0.848 (95% CI 0.837-0.858) for the ROX index (Figure 3). The AUROC for all physiological parameters and for secondary outcomes are shown in Figures S3-6. For the primary outcome, analysis of individual physiological parameters (Figure S3) shows that the AUROC for FiO₂, when treated as a continuous variable alone is 0.827 (95% CI 0.815-0.837), thus outperforming NEWS2. This compares to an AUROC of 0.701 (0.95% CI 0.693-0.708) when FiO₂ is analysed as a binary variable. The AUROC for SpO₂ is low at 0.568 ((95% CI 0.551-0.585) and thus the addition of SpO₂ to FiO₂ does not increase predictive validity. RR performs less well than other respiratory parameters (AUROC 0.784, (95% CI 0.770-0.798)). Other parameters including HR, SBP, temperature and conscious level all perform poorly. Sensitivity analysis without imputed FiO₂ data showed the AUROC for ROX was less still greater than NEWS2 (0.832, 95% CI 0.822-0.842).

With respect to secondary outcomes, the ROX index also has the highest AUROC for predicting advanced respiratory support (0.872, 95% CI 0.864-0.881) compared to NEWS2 or any of its individual components (Figure S4). The SpO₂/FiO₂ ratio has the highest AUROC with regards to predicting death on a medical ward (AUROC 0.880, 95% CI 0.845-0.915); Figure S6), likely related to the fact that SpO₂ begins to decrease late in the physiological deterioration, or ceases to be a physiological target at the end of life.

The cumulative proportion of patients who met the threshold of five points on NEWS2, as compared to those who met the optimal threshold of sensitivity and specificity for ROX (14.55), prior to peri/cardiac arrest or escalation to CC are shown in Figure S4. Of patients who had this outcome, the majority met the ROX threshold approximately 4hrs before the widely recommended NEWS2 threshold,

Table 2

Differences in most deranged physiological values observed presented by the occurrence of the primary outcome NEWS2, National Early Warning Score 2; ROX, the ratio of peripheral oxygen saturation to fractional inspired concentration of oxygen, divided by the respiratory rate; FiO₂, fractional inspired concentration of oxygen; SpO₂, peripheral oxygen saturation; BP, blood pressure; SD, standard deviation.

	Overall n=708	Any adverse event n=186	Total available data (out of 1173) (n [%])	No adverse event n=522	Total available data (out of 28610) (n [%])	Difference in means (95% CI)
Maximum NEWS2 score (Mean (SD))	6.0 (3.2)	9.0 (2.4)	1095 [93.4]	4.9 (2.8)	27335 [95.5]	4.1 (3.7,4.5)
Minimum ROX index with imputation ^a (Mean (SD))	13.9 (7.9)	6.4 (4.3)	1131 [96.4]	16.6 (7.1)	27554 [96.3]	-10.2 (-11.1,-9.3)
Minimum ROX index without imputation ^a (Mean (SD))	14.1 (8.0)	6.5 (4.4)	1086 [92.6]	16.8 (7.2)	22291 [77.9]	-10.3 (-11.2,-9.4)
Maximum FiO ₂ with imputation ^a (Mean (SD))	0.4 (0.2)	0.6 (0.2)	1170 [99.7]	0.3 (0.2)	28385 [99.2]	0.3 (0.3,0.3)
Maximum FiO ₂ without imputation ^a (Mean (SD))	0.4 (0.2)	0.6 (0.2)	1117 [95.2]	0.3 (0.2)	22592 [78.9]	0.3 (0.3,0.3)
Minimum SpO ₂ (Mean (SD))	90.3 (7.5)	86.2 (10.3)	1160 [98.9]	91.8 (5.5)	28064 [98.1]	-5.6 (-7.2,-4)
Minimum SpO ₂ / FiO ₂ with imputation ^a (Mean (SD))	295.1 (136.7)	164.8 (81.4)	1158 [98.7]	342.1 (121.2)	27847 [97.3]	-177.3 (-193,-161.6)
Minimum SpO ₂ / FiO ₂ without imputation ^a (Mean (SD))	296.2 (136.8)	165.1 (81.7)	1113 [94.9]	343.4 (121.1)	22530 [78.7]	-178.3 (-194,-162.6)
Maximum respiratory rate (Mean (SD))	25.9 (7.5)	31.7 (8.7)	1135 [96.8]	23.8 (5.8)	27879 [97.4]	7.9 (6.6,9.2)
Maximum heart rate (Mean (SD))	101.5 (17.8)	107.6 (20.5)	1125 [95.9]	99.3 (16.2)	27807 [97.2]	8.3 (5,11.6)
Minimum systolic BP (Mean (SD))	105.0 (15.3)	106.0 (16.5)	1107 [94.4]	104.6 (14.8)	27708 [96.8]	1.4 (-1.3,4.1)
Maximum temperature (Mean (SD))	38.1 (0.9)	38.5 (1.0)	1107 [94.4]	38.0 (0.9)	27771 [97.1]	0.5 (0.3,0.7)
Minimum temperature (Mean (SD))	36.0 (0.8)	36.1 (0.8)	1107 [94.4]	35.9 (0.7)	27771 [97.1]	0.2 (0.1,0.3)
Lowest recorded level of consciousness (n [%])			1127 [96.1]		27784 [97.1]	
Alert	623[88.0]	145 [78.0]		478 [91.6]		
New confusion	21 [3.0]	8 [4.3]		13 [2.5]		
Voice	38 [5.4]	21 [11.3]		17 [3.3]		
Pain	19 [2.7]	9 [4.8]		10 [1.9]		
Unresponsive	6 [0.8]	2 [1.1]		4 [0.8]		
Missing	1 [0.1]	1 [0.5]		0 [0]		

^a Imputation of FiO₂ values using last value carried forward.

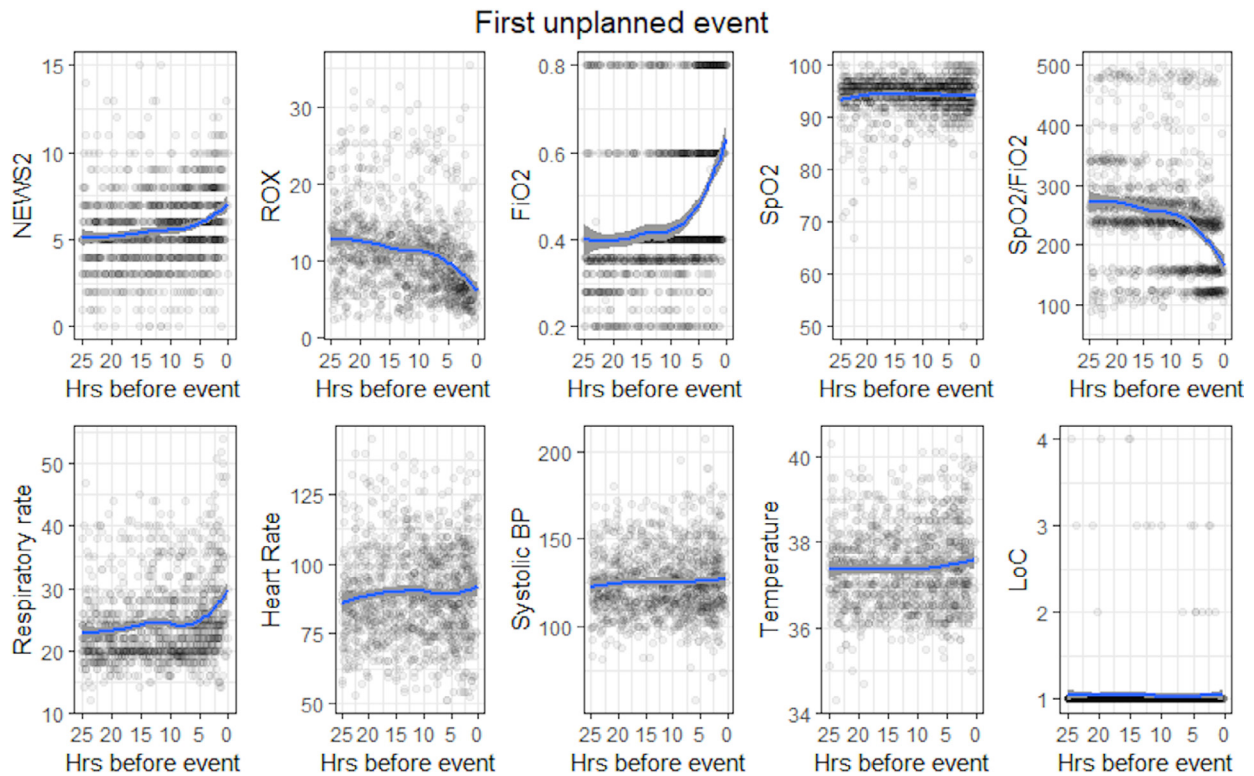


Figure 2. Physiological values in the 24 hours preceding an adverse event with LOESS smoothing curves
NEWS2, National Early Warning Score 2; ROX, ROX index - the ratio of peripheral oxygen saturation to fractional inspired concentration of oxygen, divided by the respiratory rate; FiO₂, fractional inspired concentration of oxygen; SpO₂, peripheral oxygen saturation; BP, blood pressure; LOC, level of consciousness.

Table 3

Results of univariable generalized linear mixed models for the primary outcome occurring within 24 hours of the observation set NEWS2. National Early Warning Score 2; ROX, the ratio of peripheral oxygen saturation to fractional inspired concentration of oxygen, divided by the respiratory rate; FiO₂, fractional inspired concentration of oxygen; SpO₂, peripheral oxygen saturation; BP, blood pressure; BMI, body mass index.

Variable	Unadjusted Odds Ratio (95% CI)	p-value	Unadjusted relative risks (95% CI)	Adjusted ^a Odds Ratio (95% CI)	p-value	Adjusted ^a relative risks (95% CI)
NEWS2 z-score	2.43 (2.15,2.74)	<0.001	2.30 (2.06,2.57)	2.36 (2.06,2.7)	<0.001	2.24 (1.98,2.53)
ROX index z-score	7.87 (6.31,9.82)	<0.001	6.19 (5.22,7.28)	8.33 (6.52,10.63)	<0.001	6.45 (5.35,7.69)
FiO ₂ z-score	3.63 (3.13,4.21)	<0.001	3.29 (2.89,3.73)	4.36 (3.63,5.23)	<0.001	3.85 (3.29,4.48)
SpO ₂ z-score	1.5 (1.37,1.65)	<0.001	1.47 (1.35,1.61)	1.5 (1.35,1.68)	<0.001	1.47 (1.33,1.63)
SpO ₂ /FiO ₂ z-score	7.19 (5.83,8.86)	<0.001	5.76 (4.89,6.75)	8.0 (6.32,10.12)	<0.001	6.25 (5.21,7.41)
Respiratory rate z-score	1.77 (1.61,1.94)	<0.001	1.71 (1.57,1.87)	1.78 (1.59,1.98)	<0.001	1.73 (1.56,1.91)
Heart rate z-score	1.47 (1.29,1.68)	<0.001	1.45 (1.28,1.63)	1.48 (1.28,1.7)	<0.001	1.45 (1.26,1.66)
Systolic BP z-score	0.93 (0.8,1.08)	0.344	1.07 (0.93,1.24)	1.11 (0.94,1.32)	0.213	1.11 (0.94,1.3)
Temperature z-score	0.87 (0.79,0.96)	0.006	0.88 (0.8,0.96)	0.89 (0.8,0.99)	0.029	0.89 (0.81,0.99)

^a Adjusted for the baseline variables age, gender, ethnicity, and BMI

Table 4

Sensitivity, specificity, negative and positive predictive values for NEWS2 and ROX index at different cut points for the occurrence of the primary outcome PPV, positive predictive value; NPV, negative predictive value.

NEWS2					ROX index				
Cut point	Sensitivity	Specificity	PPV	NPV	Cut point	Sensitivity	Specificity	PPV	NPV
0.5	0.99	0.15	0.04	1.00	3	0.02	1	0.20	0.96
1.5	0.97	0.31	0.05	1.00	4	0.05	0.99	0.20	0.96
2.5	0.93	0.49	0.07	0.99	5	0.11	0.98	0.18	0.96
3.5	0.84	0.66	0.09	0.99	6	0.18	0.96	0.17	0.97
4.5	0.72	0.79	0.12	0.99	7	0.28	0.95	0.19	0.97
5.5	0.51	0.87	0.14	0.98	8	0.35	0.93	0.18	0.97
6.5	0.31	0.93	0.15	0.97	9	0.42	0.92	0.17	0.97
7.5	0.18	0.96	0.15	0.97	10	0.49	0.9	0.17	0.98
8.5	0.09	0.98	0.15	0.96	11	0.58	0.88	0.17	0.98
9.5	0.04	0.99	0.16	0.96	12	0.68	0.85	0.16	0.98
					13	0.74	0.83	0.15	0.99
					14	0.8	0.78	0.13	0.99
					15	0.85	0.74	0.12	0.99
					16	0.87	0.71	0.11	0.99
					17	0.89	0.67	0.10	0.99
					18	0.91	0.63	0.09	0.99
					19	0.92	0.59	0.08	0.99
					20	0.93	0.54	0.08	0.99

although the curves converge again closer to the event and the overall positive predictive value of ROX is similar to NEWS2

Fig 4

4. Discussion

In this cohort of Covid-19 positive patients who were admitted onto a medical ward, 26% went on to have a peri/cardiac arrest, CC admission or died on the ward. Our data shows that there is marked derangement of normal physiology and a clear trajectory of deterioration that precedes the event. We tested the validity of NEWS2, a “track and trigger” system mandated across the NHS and found that a simple three parameter model (the ROX index), based solely on respiratory parameters and which treats FiO₂ as a continuous variable, outperformed NEWS2 for predicting adverse events, particularly the need for advanced respiratory support. These findings have the potential to be clinically important if they translate into earlier identification of deteriorating patients.

The key strength of our study is that it uses accessible, longitudinal bedside physiological data to aid risk stratification and escalation of Covid-19 patients and a composite of clinically important and comparable outcomes. Because ROX utilises the observations already being recorded, it can be easily calculated and integrated with current escalation protocols. On the other hand, this a single centre, retrospective study and the generalisability to other healthcare systems is not known. In the absence of a none-Covid-19 control group we

cannot know if the findings are unique to Covid-19 as a cause of respiratory failure. Data from the ED was not used and therefore our findings cannot necessarily be extrapolated to patients in that setting; neither are the findings applicable to patients receiving high flow oxygen therapy or CPAP / NIV. There was an appreciable amount of ‘Null’ data in the FiO₂ field, raising the potential to bias the results. However, this occurred predominantly in the group that did not suffer an adverse event (a group who would be less sick), which we believe supports our assumption that ‘Null’ may have commonly represented the absence of supplemental oxygen. For this the use of LVCF was appropriate, although sensitivity analysis without FiO₂ imputation also gave similar results. The ROX index is sensitive to the FiO₂ and this value cannot usually be measured accurately when using simple oxygen delivery devices; utilising estimated values in this study is a recognised limitation and is a reason why applying absolute values of ROX, derived from this study is not recommended.

Recently published studies have suggested that NEWS2 is useful and predictive of deterioration in Covid-19 patients, [9,10,11,13] whilst others have criticised its efficacy [8,12]. The majority of these studies use NEWS2 score at a single time point at hospital admission, rather than as a continuous assessment tool. The PACIFIC-19 group have recently published a comprehensive validation, reporting an AUROC of 0.882 (95% CI 0.868–0.895), and similar to a none-Covid-19 population, but acknowledge that this may be inflated as it is in the setting of the original NEWS validation, and that the impact of the binary oxygen grading system is not fully known [14].

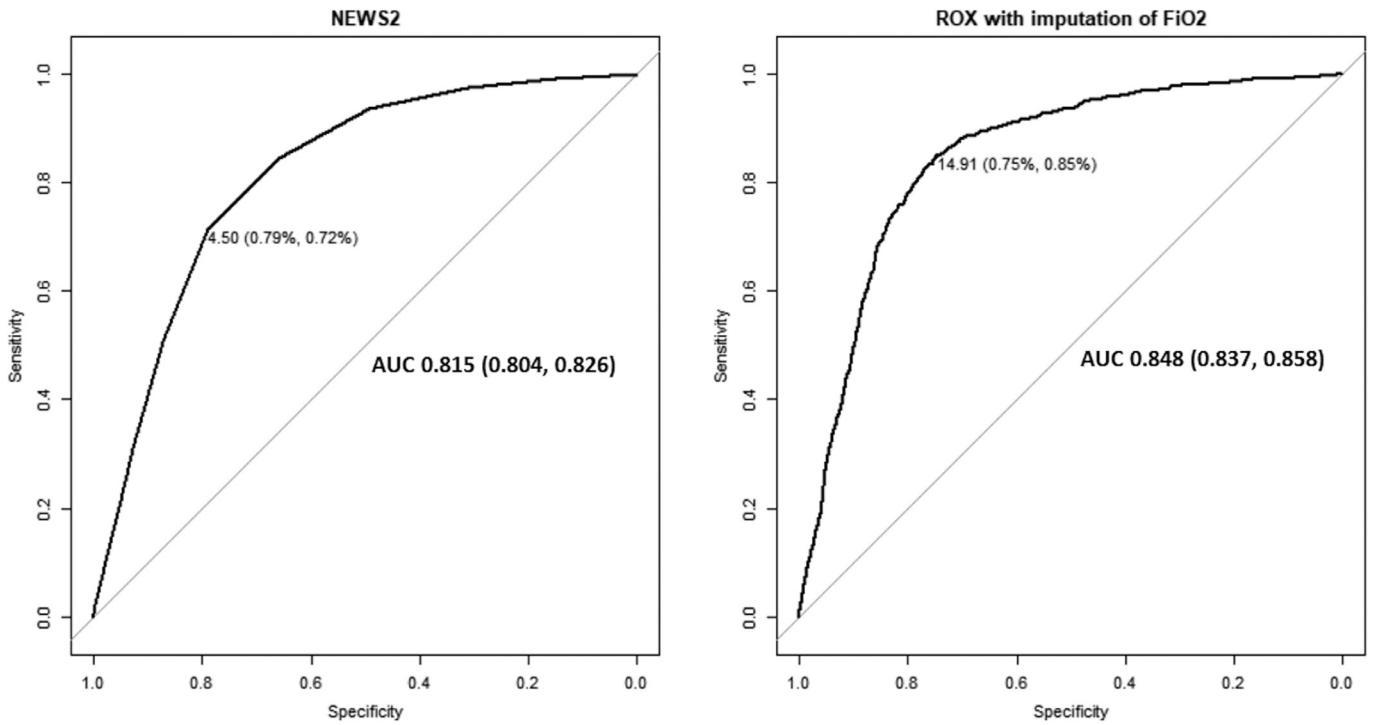


Figure 3. Receiver operating characteristic curves for NEWS2 and ROX index to predict the primary outcome within 24 hours
 NEWS2, National Early Warning Score 2; ROX, ROX index - the ratio of peripheral oxygen saturation to fractional inspired concentration of oxygen, divided by the respiratory rate; AUC, area under the receiver operating characteristic curve

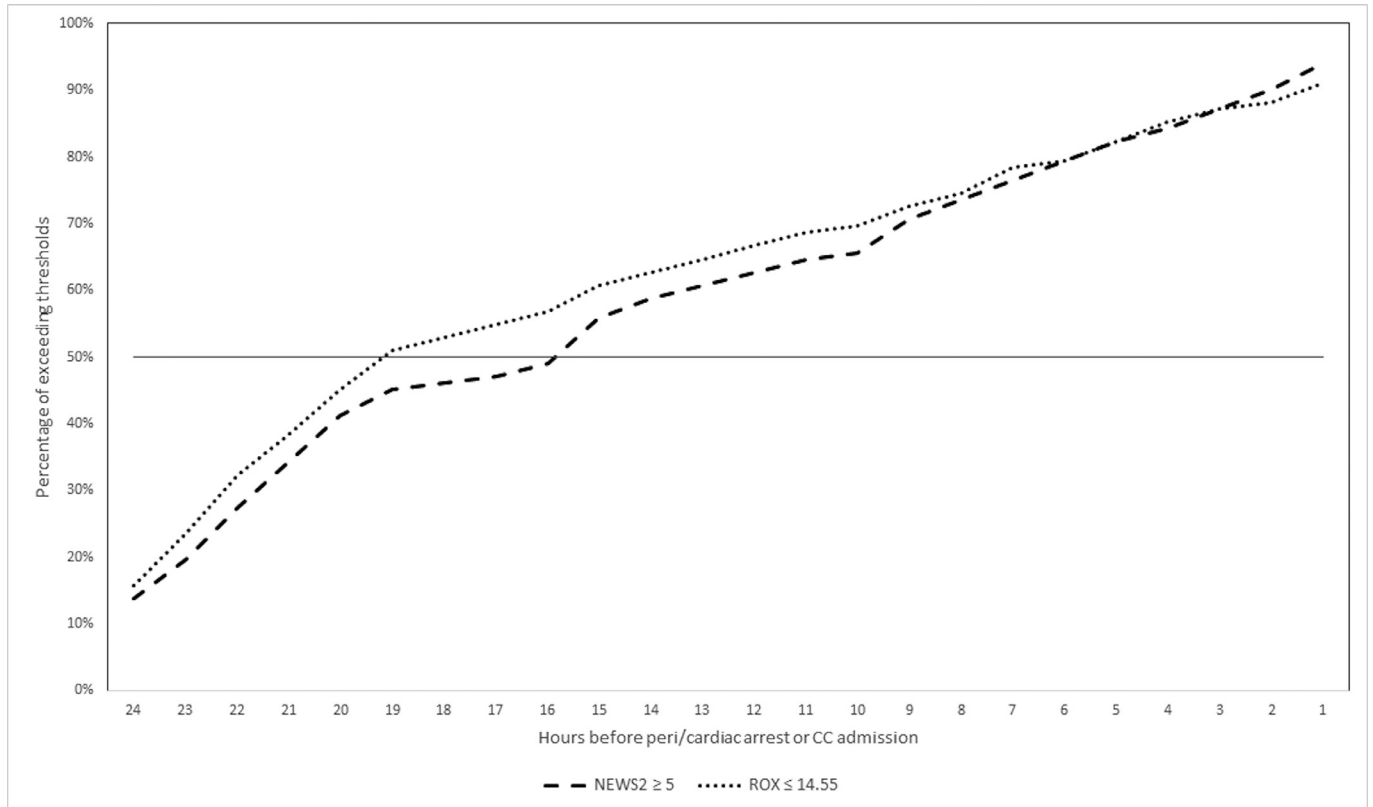


Figure 4. Cumulative occurrence of NEWS ≥ 5 (established threshold for escalation) vs ROX ≤ 14.55 (optimal sensitivity and specificity for predicting peri/cardiac arrest call or Critical Care admission 24 hrs prior to the event occurring.
 NEWS2, National Early Warning Score 2; ROX, ROX index - the ratio of peripheral oxygen saturation to fractional inspired concentration of oxygen, divided by the respiratory rate

Other published prediction models have used a more complicated composite of patient related risk factors, laboratory data and clinical parameters which are more challenging to operationalise [18,27-29]. To our knowledge, no other study has evaluated the use of the ROX index in Covid-19 although Hamovich et al. proposed the “Covid severity index” (qCSI), which also focuses on respiratory physiology [17].

NEWS2 has been shown to be an indicator of increased risk of serious clinical deterioration and mortality in adult patients [4]. The original validation reported an AUROC for prediction of the same primary outcome of 0.873 (95% CI 0.866-0.879) [7]. We found the AUROC in this Covid-19 population is 0.815 (95% CI 0.804-0.826), thus underperforming in comparison. The difference in these findings is likely to be accounted for by the particular pathophysiology in Covid-19 which manifests with predominantly, though not exclusively respiratory features. Our data in Figure 2 shows that the increase in FiO_2 is the key harbinger of deterioration with a major inflection from around 10 hours before the event, with tachypnoea a later but still notable feature. There is a relative absence of cardiovascular disturbance. Heart rate is only weakly statistically associated with the outcome and in a way that may have limited clinical significance and systolic blood pressure is not associated at all. Therefore, an early warning system that applies weight to these factors is less likely to be predictive in this condition. This data supports anecdotal reports of Covid-19 and are consistent with the findings of Pimentel et al.[15]. The other key issue is that ‘oxygen requirement’ is treated as a binary variable in NEWS2 with two points allocated for oxygen supplementation, regardless of whether this is FiO_2 0.24 or 0.8. Whilst this has advantages in terms of implementation, it fails to identify the key feature of deterioration in Covid-19. By directly comparing the predictive validity of FiO_2 as a binary and as a continuous variable we have demonstrated mechanistically an important reason why NEWS2 may under-perform.

We note that the ROX index threshold we identified is markedly higher than has been described previously [19]. Our threshold is for predicting the outcome within 24 hours, thus providing an ‘early warning’, rather than dictating that the patient requires CC admission or intubation at that time. In addition, the observation that Covid-19 patients are unusually tolerant of hypoxaemia and thus tachypnoea is less of a feature may raise the ROX threshold [30,31]. Finally, because we had a low threshold for escalation and did not use CPAP on the ward, this may have produced more conservative thresholds for NEWS2 and ROX than is practical elsewhere. It is acknowledged that the incremental benefit of ROX is only small. Whilst the positive predictive value for the occurrence of the outcome within 24 hours remains relatively low and is similar to NEWS2, within that time period there is the potential for earlier detection by up to four hours. The negative predictive value, an important safety consideration for early warning scores, is high in both systems. One additional strength of the ROX index as compared to the NEWS2 is that it may allow a greater number of patient to be monitored more easily and with reduced contact for health care workers; thus it could have advantages for infection prevention and be more easily delivered in lower resources settings or are at times of overwhelming demand.

An optimal detection and escalation pathway needs to offer a high sensitivity trigger to identify a deteriorating patient in a timely manner. This will enable enhanced monitoring and consideration of treatment and escalation planning, with admission to CC if appropriate, before injurious deterioration occurs. We have demonstrated that a ROX threshold of 14.55 can identify patients up to 4 hours earlier than a NEWS2 of 5, which is likely to be clinically important. However, high sensitivity must be balanced with sufficient specificity for an adverse outcome so as to avoid alert fatigue, over-treatment or preservation of scarce resources. The thresholds identified as statistically optimum do not imply clinical safety, and need to be tested prospectively. Whilst we do not advocate abandoning NEWS2, with all

of the associated benefits of standardisation, there could be advantage from additionally incorporating the ROX index as a parallel system. Alternatively, the benefits of grading oxygen dependency as a weighed variable could be incorporated into a unified Covid-19 early warning score which would require prospective testing in parallel with established systems.

In this study NEWS2 under-performed in Covid-19, as compared the population in which it was first validated, due to intrinsic limitations of the design and the unique physiology of the disease. A simple index utilising respiratory parameters can outperform NEWS2 in predicting the occurrence of adverse events. These findings require validation in future studies.

Declaration of competing Interest

All authors declare that they have nothing to disclose

Acknowledgements

We wish to acknowledge the support of Leah Evans, BS, and Kathleen Merkley, DNP, APRN, ANP, of Health Catalyst, Inc. Salt Lake City, Utah for developing the Health Catalyst® Data Operating System (DOS™) utilised 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.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data sharing statement

Individual participant data that underlie the results reported in this article, after de-identification (text, tables, figures, and appendices), an associated data dictionary and the study protocol is available from the corresponding author, immediately following publication and with no end date. This will be made available to researchers who provide a methodologically sound proposal to achieve the aims of the proposal. Please contact guy.glover@gstt.nhs.uk

Contributors

Emma Prower: conceptualisation, methodology, writing – original draft; David Grant: conceptualisation, methodology, validation, data curation, writing – review and editing; Alessandra Bisquera: formal analysis, data curation, data visualisation, writing – review and editing; Cormac Breen: supervision, project administration; writing – review and editing; Luigi Camporota: conceptualisation, methodology; writing – review and editing; Maja Gavrilovski: conceptualisation, writing – review and editing; Meghan Pontin: conceptualisation, project administration, writing – review and editing; Abdel Douri: formal analysis, data curation, data visualisation, writing – review and editing; Guy Glover: conceptualisation, methodology, writing – review and editing.

David Grant, Alessandra Bisquera and Guy Glover have verified the underlying data. All authors have read and finally approved the version being submitted. All authors confirm that they had full access to all the data in the study and accept responsibility to submit for publication

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.eclinm.2021.100828](https://doi.org/10.1016/j.eclinm.2021.100828).

References

- [1] McCabe R, Schmit N, Christen P, et al. Adapting hospital capacity to meet changing demands during the COVID19 pandemic. *Imperial College Lond* 2020 (15-06-2020). doi: [10.25561/79837](https://doi.org/10.25561/79837).
- [2] WHO Coronavirus disease dashboard. Coronavirus-staging.data.gov.uk: Accessed online: <https://coronavirus-staging.data.gov.uk/healthcare>
- [3] Docherty A, Harrison E, Green C, et al. on behalf of ISARIC4C investigators. Features of 20 133 UK patients in hospital with Covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ* 2020;369:m1985. doi: [10.1136/bmj.m1985](https://doi.org/10.1136/bmj.m1985).
- [4] Smith G B, Prytherch D R, Meredith P, Schmidt P E, Featherstone P I. 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. doi: [10.1016/j.resuscitation.2012.12.016](https://doi.org/10.1016/j.resuscitation.2012.12.016).
- [5] 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. doi: [10.1016/j.resuscitation.2018.09.026](https://doi.org/10.1016/j.resuscitation.2018.09.026).
- [6] Mellhammar L, Linder A, Tverring J, et al. NEWS2 is Superior to qSOFA in detecting sepsis with organ dysfunction in the emergency department. *J Clin Med* 2019;8:1128. doi: [10.3390/jcm8081128](https://doi.org/10.3390/jcm8081128).
- [7] 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. Accessed online: <https://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news-2>
- [8] Carr E, Bendayan R, Bean D, et al. Evaluation and improvement of the National Early Warning Score (NEWS2) for COVID-19: a multi-hospital study. *BMC Med* 2021;19:23. doi: [10.1186/s12916-020-01893-3](https://doi.org/10.1186/s12916-020-01893-3).
- [9] Gidari A, Vittorio De Socio G, Sabbatini S, Francisci D. Predictive value of National Early Warning Score 2 (NEWS2) for intensive care unit admission in patients with SARS-CoV-2 infection. *Infect Dis* 2020;52:698–704. doi: [10.1080/23744235.2020.1784457](https://doi.org/10.1080/23744235.2020.1784457).
- [10] 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. doi: [10.1186/s13049-020-00764-3](https://doi.org/10.1186/s13049-020-00764-3).
- [11] Peng X, Subbe C, Zang L, et al. NEWS can predict deterioration of patients with COVID-19. Letter to the editor. *Resuscitation* 2020. doi: [10.1016/j.resuscitation.2020.05.011](https://doi.org/10.1016/j.resuscitation.2020.05.011).
- [12] Sze S, Pan D, Williams C, et al. Variability but not admission or trends in NEWS2 score predicts outcome in elderly hospitalised patients with COVID. *J Infect* 2020 (20):30341–8 S0163–4453. doi: [10.1016/j.jinf.2020.05.063](https://doi.org/10.1016/j.jinf.2020.05.063).
- [13] 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. doi: [10.1016/j.resuscitation.2020.08.124](https://doi.org/10.1016/j.resuscitation.2020.08.124).
- [14] Kostakis I, Smith GB, Prytherch D, Meredith P, Price C, Chauhan A. The performance of the National Early Warning Score and National Early Warning Score 2 in hospitalised patients infected by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). *Resuscitation* 2020. doi: [10.1016/j.resuscitation.2020.10.039](https://doi.org/10.1016/j.resuscitation.2020.10.039).
- [15] Pimentel M, Redfern O, Hatch R, Young D, Tarresenko L, Watkinson PJ. Trajectories of vital signs in patients with Covid-19. *Resuscitation* 2020;156:99–106. doi: [10.1016/j.resuscitation.2020.09.002](https://doi.org/10.1016/j.resuscitation.2020.09.002).
- [16] Haimovich A, Ravindra N, Stoychev S, et al. Development and validation of the COVID-19 severity index (CSI): a prognostic tool for early respiratory decompensation. *Ann Emerg Med* 2020;76:442–53. doi: [10.1016/j.annemergmed.2020.07.022](https://doi.org/10.1016/j.annemergmed.2020.07.022).
- [17] Cheng F-Y, Joshi H, Tandon P, et al. Using machine learning to predict ICU transfer in hospitalized COVID-19 patients. *J Clin Med* 2020;9:668. doi: [10.3390/jcm9061668](https://doi.org/10.3390/jcm9061668).
- [18] Malycha J, Farajidavar N, Pimentel M, et al. The effect of fractional inspired oxygen concentration on early warning score performance: A database analysis. *Resuscitation* 2019;139:192–9 Tdoi. doi: [10.1016/j.resuscitation.2019.04.002](https://doi.org/10.1016/j.resuscitation.2019.04.002).
- [19] Roca O, Caralt B, Messika J, et al. An index combining respiratory rate and oxygenation to predict outcome of nasal high-flow therapy. *Am J Resp Crit Care Med* 2019;199:1368–76.
- [20] NHS England: Specialty guides for patient management during the coronavirus pandemic. The role and use of non-invasive respiratory support in adult patients with COVID- 19 (confirmed or suspected). 6 April 2020, Version 3. Accessed online: <https://www.england.nhs.uk/coronavirus/wpcontent/uploads/sites/52/2020/03/specialty-guide-NIV-respiratory-support-and-coronavirus-v3.pdf>
- [21] NHS England. Specialty guides for patient management during the coronavirus pandemic. Clinical guide for the management of critical care patients during the coronavirus pandemic 16 March 2020 Version 1. Accessed online.
- [22] Kelly N, Curtis A, Douthwaite S, et al. Effect of awake prone positioning in hypoxaemic adult patients with COVID-19. *J Intens Care Soc* 2020. doi: [10.1177/1751143720961244](https://doi.org/10.1177/1751143720961244).
- [23] World Health Organisation. Clinical management of severe acute respiratory infection (SARI) when COVID-19 disease is suspected. Interim Guid: March 2020. Accessed online: <https://www.who.int/docs/defaultsource/coronaviruse/clinical-management-of-novel-cov.pdf>
- [24] NHS England: Specialty guides for patient management during the coronavirus pandemic. Clinical guide for the optimal use of oxygen therapy during the coronavirus pandemic. 9 April 2020 Version 1. Accessed online: <https://www.england.nhs.uk/coronavirus/wpcontent/uploads/sites/52/2020/04/C0256-specialty-guide-oxygen-therapy-and-coronavirus-9-april-2020.pdf>
- [25] Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care* 1998;36:8–27. doi: [10.1097/00005650-199801000-00004](https://doi.org/10.1097/00005650-199801000-00004).
- [26] 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. doi: [10.1136/bmjopen-2018-026759](https://doi.org/10.1136/bmjopen-2018-026759).
- [27] Zhou Y, He Y, Yang H, Yu H, Wang T, Chen Z, et al. Development and validation a nomogram for predicting the risk of severe COVID- 19: A multi-center study in Sichuan, China. *PLoS ONE* 2020;15:e0233328. doi: [10.1371/journal.pone.0233328](https://doi.org/10.1371/journal.pone.0233328).
- [28] Ji Dong, Zhang Dawei, Xu Jing, et al. Prediction for Progression Risk in Patients With COVID-19 Pneumonia: The CALL Score (2020). *Clin Infect Dis* 2020;71:1393–9. doi: [10.1093/cid/ciaa414](https://doi.org/10.1093/cid/ciaa414).
- [29] Knight SR, Ho A, Pius R, et al. Risk stratification of patients admitted to hospital with Covid-19 using the ISARIC WHO clinical characterisation protocol: development and validation of the 4C Mortality Score. *BMJ* 2020;370:m3339. doi: [10.1136/bmj.m3339](https://doi.org/10.1136/bmj.m3339).
- [30] Tobin M, Laghi F, Jubran A. Why COVID-19 silent hypoxemia is baffling to physicians. *Am J Resp Crit Care Med* 2020;202. doi: [10.1164/rccm.202006-2157CP](https://doi.org/10.1164/rccm.202006-2157CP).
- [31] Dhont S, Derom E, Van Braeckel E, et al. The pathophysiology of 'happy' hypoxemia in COVID-19. *Respir Res* 2020;21:198. doi: [10.1186/s12931-020-01462-5](https://doi.org/10.1186/s12931-020-01462-5).