

# National Early Warning Score 2 (NEWS2) to identify inpatient COVID-19 deterioration: a retrospective analysis

**Authors:** Kenneth F Baker,<sup>A</sup> Aidan T Hanrath,<sup>B</sup> Ina Schim van der Loeff,<sup>B</sup> Lesley J Kay,<sup>C</sup> Jonathan Back,<sup>D</sup> and Christopher JA Duncan<sup>E</sup>

## ABSTRACT

### Introduction

We sought to provide the first report of the use of NEWS2 monitoring to pre-emptively identify clinical deterioration within hospitalised COVID-19 patients.

### Methods

Consecutive adult admissions with PCR-confirmed COVID-19 were included in this single-centre retrospective UK cohort study. We analysed all electronic clinical observations recorded within 28 days of admission until discharge or occurrence of a serious event, defined as any of the following: initiation of respiratory support, admission to intensive care, initiation of end of life care, or in-hospital death.

### Results

133/296 (44.9%) patients experienced at least one serious event. NEWS2  $\geq 5$  heralded the first occurrence of a serious event with sensitivity 0.98 (95% CI 0.96–1.00), specificity 0.28 (0.21–0.35), positive predictive value (PPV) 0.53 (0.47–0.59), and negative predictive value (NPV) 0.96 (0.90–1.00). The NPV (but not PPV) of NEWS2 monitoring exceeded that of other early warning scores including the Modified Early Warning Score (MEWS) (0.59 [0.52–0.66],  $p < 0.001$ ) and quick Sepsis Related Organ Failure Assessment (qSOFA) score (0.58 [0.51–0.65],  $p < 0.001$ ).

### Conclusion

Our results support the use of NEWS2 monitoring as a sensitive method to identify deterioration of hospitalised COVID-19 patients, albeit at the expense of a relatively high false-trigger rate.

**KEYWORDS:** early warning score; COVID-19; SARS-CoV-2; deterioration; monitoring

**DOI:** 10.7861/clinmed.2020-0688

## Summary

### What is known?

The NEWS2 scoring system is widely used throughout the UK NHS to monitor physiological parameters in order to enable the early detection of clinical deterioration. However, its performance in COVID-19 has not been validated and concerns have been raised about its sensitivity.

### What is the question?

We aimed to ascertain whether longitudinal NEWS2 monitoring can pre-emptively identify clinical deterioration in patients hospitalised with COVID-19.

### What was found?

NEWS2  $\geq 5$  had an excellent sensitivity to detect deteriorating COVID-19 patients, albeit at the expense of a relatively high false-trigger rate. Longitudinal trends in NEWS2 scores increased many hours before serious clinical events, and baseline NEWS2 was also modestly predictive of future clinical deterioration.

### What is the implication for practice now?

NEWS2 monitoring is an appropriately sensitive method for identifying the potential for clinical deterioration of hospitalised COVID-19 patients and should continue to be used alongside clinical judgement.

**Authors:** <sup>A</sup>NIHR clinical lecturer, Translational and Clinical Research Institute, Newcastle University, UK and Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK; <sup>B</sup>academic clinical fellow, Translational and Clinical Research Institute, Newcastle University, UK; <sup>C</sup>honorary senior clinical lecturer, Translational and Clinical Research Institute, Newcastle University, UK, consultant rheumatologist, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK, and deputy medical director, Healthcare Safety Investigation Branch (HSIB), Farnborough, UK; <sup>D</sup>intelligence analyst, Healthcare Safety Investigation Branch (HSIB), Farnborough, UK; <sup>E</sup>Wellcome Trust clinical research career development fellow, Translational and Clinical Research Institute, Newcastle University, UK, and honorary consultant in infectious diseases, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK

## Introduction

Healthcare systems face many challenges in responding to the SARS-CoV-2 pandemic, not least the issue of how to best direct finite resources towards those patients in greatest clinical need. Many patients hospitalised with COVID-19 require non-invasive pressure support, invasive ventilation, or critical care admission, and

identifying these patients early is important.<sup>1</sup> Necessary structural reorganisation of healthcare systems has led to redeployment of medical and nursing staff, who are faced with an unfamiliar disease and in some cases are operating with limited acute medical experience.<sup>2</sup> Simple and effective tools to identify deteriorating patients are needed.

In the UK NHS, the recently updated National Early Warning Score (NEWS2) is widely used to identify deteriorating hospitalised patients, in particular to identify those requiring escalation to a higher level of care (ie from general ward to a critical care setting). The NEWS2 system applies an aggregative weighted ordinal stratification to routinely measured physiological parameters including heart rate, blood pressure, temperature, respiratory rate, oxygen saturations and consciousness level to generate a score from 0 to 23.<sup>3</sup> A threshold of NEWS2  $\geq 5$  is used as a trigger for immediate clinical review, and has been validated in acute medical admissions and other settings.<sup>4</sup> Ease of calculation, propensity to be incorporated into electronic monitoring systems and a desire for standardisation have led to the mandatory adoption of NEWS2 across all NHS inpatient settings, forming a fundamental component of clinical escalation and intensive care unit (ICU) outreach services.<sup>3</sup> Alternative rapid scoring systems have also been developed, notably the Modified Early Warning Score (MEWS)<sup>5</sup> and quick Sepsis Related Organ Failure Assessment (qSOFA)<sup>6</sup> scores, which have been validated in the management of acute medical admissions and sepsis respectively.

No early warning score system has been validated for use in COVID-19. Indeed, use of these scores to identify clinical deterioration in this setting has been called into question.<sup>7</sup> COVID-19 often causes hypoxaemia without substantial perturbation of other physiological parameters, known as 'silent hypoxaemia'.<sup>8</sup> This, combined with a relative underscoring of hypoxaemia by NEWS2, has led to concerns that this scoring system may be inadequately sensitive in detecting need for escalation in COVID-19.<sup>7,9</sup>

In this study, we provide the first report of the performance of longitudinal NEWS2 monitoring to identify clinical deterioration within a hospitalised COVID-19 patient cohort.

## Methods

### Setting

The study setting was the Newcastle upon Tyne Hospitals (NUTH) NHS Foundation Trust, a large tertiary academic medical centre, with a High Consequences Infectious Diseases unit, in the North East of England, and the first in the UK to manage COVID-19 patients from admission to discharge.<sup>10</sup> We have previously described the clinical characteristics and outcomes of patients admitted to our hospital Trust with COVID-19.<sup>11</sup> Briefly, consecutive patients >18 years old, admitted between 31 January and 16 April 2020 inclusive, and with a positive SARS-CoV-2 nasal and/or oropharyngeal PCR swab were included. All re-admissions within 28 days of date of initial hospitalisation were included.

### Data collection and clinical definitions

For this analysis, we collected for these patients all clinical observations (heart rate, blood pressure, oxygen saturations, respiratory rate, temperature, and consciousness level) from admission until the occurrence of a serious clinical event. We defined serious events prior to data collection as any one of the following: initiation of respiratory support (ie continuous positive

airway pressure [CPAP], bilevel positive airway pressure [BiPAP], high flow nasal cannula [HFNC], or invasive ventilation), admission to the ICU, initiation of end of care (EoLC), or in-hospital death. Where none of these events occurred, observations were recorded up to the point of discharge. Data collection was censored at 28 days for those patients who remained admitted at this time. All clinical observations, together with contemporaneous oxygen delivery device and flow rate, were recorded in real-time on our electronic care records system. Observations that were recorded within 5 minutes of a subsequent observation set were excluded in order to remove potential data transcription errors at the point of clinical care. NEWS2 scores were calculated automatically and available in real time to clinical teams. To enable comparative analysis we retrospectively calculated NEWS2, qSOFA and MEWS scores from raw observation values, with exclusion of data where missing parameters prevented score calculation. Patients who were admitted for an alternative reason prior to onset of COVID-19 symptoms were excluded from analysis, owing to potential spurious perturbations of physiological parameters attributable to non-COVID-19 pathology. Patients who had already fulfilled the definition of a serious clinical event prior to admission by virtue of domiciliary non-invasive ventilation (NIV) which was continued during admission were also excluded. Disease severity at admission was defined according to the World Health Organization definition of severe COVID-19 pneumonia<sup>12</sup> (respiratory rate >30, or oxygen saturations <90% on room air) and/or a new supplemental oxygen requirement.

### Statistical analysis

Analysis was performed in R (version 3.6.0). Tests of differences in proportions ( $\chi^2$  test) and continuous data (Wilcoxon rank sum test) were performed between contrast groups where stated. Time before event versus NEWS2/qSOFA score was plotted using loess regression with a data span of 75%, quadratic polynomial, and Gaussian kernel. Differences in trigger-to-event time (Wilcoxon signed rank test), sensitivity/specificity (exact binomial test), positive/negative predictive values (Kosinski weighted generalised score statistic), and positive/negative likelihood ratios (generalised estimating equations [GEE] logistic regression) for future clinical event between NEWS2 versus qSOFA scores were compared using the 'DTComPair' package.<sup>13</sup> Area under the receiver operating characteristic curve (ROC<sub>AUC</sub>) together with 95% confidence intervals and statistical significance of difference between AUCs (using the DeLong procedure) were calculated using the 'pROC' package.<sup>14</sup> An  $\alpha < 0.05$  was considered statistically significant.

### Study approvals

The study was registered with NUTH as a clinical service evaluation, and was exempt from ethical approval and the requirement for patient consent according to UK Government guidelines.<sup>15</sup> Approval for local data collection and analysis of anonymised clinical data was approved by the NUTH Caldicott Guardian.

## Results

### Cohort characteristics

296 patients (162 [54.7%] male, median [IQR] age 75 [62–84] years) were included in the analysis. Ethnicity was recorded for 283 patients, of which 34 (12.0%) were of black, Asian or minority ethnic (BAME) background, in keeping with local population

**Table 1. Cohort clinical characteristics, stratified by occurrence of serious event**

Demographic	N	Cohort (n=296)	Any event (n=133)	No event (n=163)	p
Male sex	296	162 (54.7)	76 (57.1)	86 (52.8)	0.524
Age (years)	296	75 (62–84) [23–101]	76 (63–84) [32–97]	74 (60–83) [23–101]	0.132
Symptom duration at admission (days)	296	4 (2–9) [0–42]	5 (2–9) [0–21]	4 (2–9) [0–42]	0.386
Clinical frailty score	294	4 (2–6) [1–9]	5 (2–6) [1–9]	4 (2–6) [1–7]	0.103
Nursing / residential home	296	60 (20.3)	33 (24.8)	27 (16.6)	0.107
WHO severe COVID-19 at admission	293	131 (44.7)	86 / 131 (65.6)	45 / 162 (27.7)	<0.001
CURB-65 score at admission	287	2 (1–2) (0–4)	2 (1–3) [0–4]	1 (1–2) [0–4]	0.003
Pneumonia on admission chest radiograph	288	114 (39.6)	71 / 132 (53.8)	43 / 156 (27.6)	0.004
Cumulative admission duration (days)	296	8 (4–14) [0–28]	9 (5–18) [0–28]	8 (3–12) [0–28]	0.011
Baseline NEWS2 score	296	5 (2–7) [0–16]	6 (4–8) [0–16]	3 (2–5) [0–11]	<0.001
Baseline MEWS score	296	2 (1–3) [0–9]	3 (2–4) [0–9]	2 (1–3) [0–9]	<0.001
Baseline qSOFA score	296	1 (0–1) [0–3]	1 (0–1) [0–3]	0 (0–1) [0–2]	<0.001

Data are median (IQR) [range] for continuous variables, and n (%) or n/N (%) for categorical variables

demographics. A total of 15,565 observation time points were available, yielding 14,336 (92.1%) NEWS2 scores, 14,339 (92.1%) MEWS scores, and 14,377 qSOFA scores (92.4%, Fig 1). 133 (44.9%) patients experienced at least one serious event, including initiation of respiratory support (73 events), admission to ICU (55 events), or in-hospital death/initiation of EoLC (76 events). No significant difference in sex, age, symptom duration, clinical frailty score or care home residence was observed between those patients with and without an event (Table 1), hereafter referred to as the 'deterioration' and 'stable' groups respectively. Significantly more patients had severe COVID-19 on admission in the deterioration vs stable group (86/131 [65.6%] vs 45/162 [27.7%],  $p<0.001$ ). Similarly, pneumonia on baseline chest X-ray was more commonly seen in the deterioration group (71/132 [53.8%] vs 43/156 [27.6%],  $p=0.004$ ), reflecting the greater requirement for respiratory support in this group.

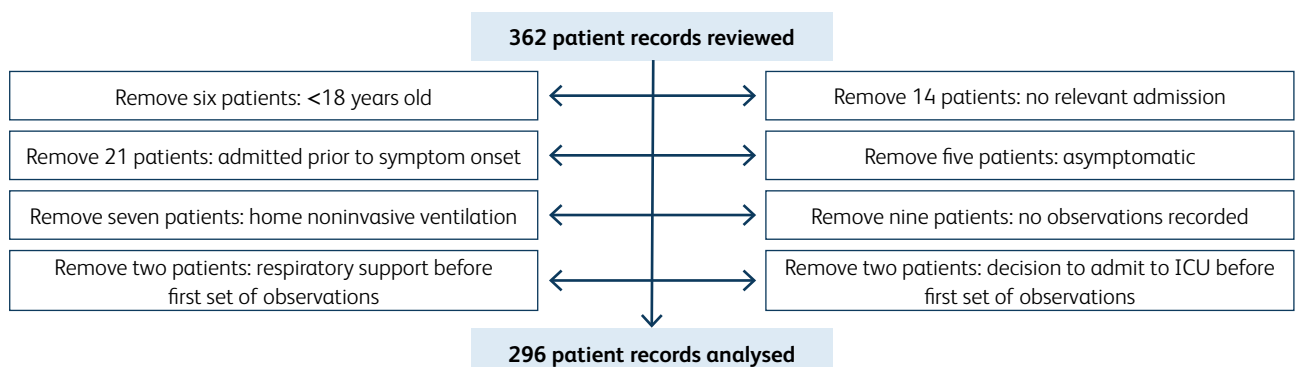
### Clinical score trigger prior to serious event

The main utility of early warning scores is to identify deteriorating patients in advance of the occurrence of a serious event, in order that intervention can occur. To address the clinical utility of NEWS2 monitoring in COVID-19, we used the published thresholds for immediate clinician review based on non-COVID sepsis studies,

namely NEWS2  $\geq 5$ .<sup>3</sup> For comparison, we also studied the performance of MEWS and qSOFA using their published thresholds of MEWS  $\geq 5$ <sup>5</sup> and qSOFA  $\geq 2$ .<sup>6</sup> In order to ensure a temporal link with occurrence of serious events, we defined a 'trigger' as any score meeting the threshold within the 24-hour period immediately prior to serious event in the deterioration group. Within the stable group, we defined a trigger as any score meeting the threshold occurring at any time during admission.

A significantly greater number of patients in the deterioration group recorded at least one 'true positive' NEWS2 trigger (131/133 patients; sensitivity 0.98 [95% CI 0.96–1.00]) when compared to MEWS (52 patients; 0.39 [0.31–0.47],  $p<0.001$ ) and qSOFA (42 patients; 0.32 [0.24–0.39],  $p<0.001$ ) (Table 2). Consequently, the negative predictive value of NEWS2 (0.96 [95% CI 0.90–1.00]) was significantly greater than that of MEWS (0.59 [0.52–0.66],  $p<0.001$ ) and qSOFA (0.58 [0.51–0.65],  $p<0.001$ ). Furthermore, where triggers were recorded within 24 hours before first event, the first trigger occurred significantly earlier for NEWS2 (median [IQR] 11.4 [4.4–20.6] hours before event) versus MEWS (6.7 [2.9–14.3] hours,  $p=0.010$ ) and qSOFA (5.6 [3.2–12.4],  $p=0.003$ ) (supplementary material S1).

Only 2/133 (1.5%) patients deteriorated without prior NEWS2 trigger. One of these patients was admitted in extremis, and EoLC was initiated within 3 hours of hospitalisation and after only two



**Fig 1. Exclusion of patients prior to analysis.**

**Table 2. Diagnostic metrics of NEWS2, MEWS and qSOFA score triggers for serious event within 24 hours**

Metric	NEWS2	MEWS	qSOFA
Sensitivity	0.98 (0.96–1.00)	0.39 (0.31–0.47)*	0.32 (0.24–0.39)*
Specificity	0.28 (0.21–0.35)	0.71 (0.64–0.78)*	0.77 (0.71–0.84)*
Positive predictive value	0.53 (0.47–0.59)	0.53 (0.43–0.62) <sup>ns</sup>	0.53 (0.42–0.64) <sup>ns</sup>
Negative predictive value	0.96 (0.90–1.00)	0.59 (0.52–0.66)*	0.58 (0.51–0.65)*
Positive likelihood ratio	1.37 (1.24–1.51)	1.36 (0.98–1.87) <sup>ns</sup>	1.39 (0.95–2.03) <sup>ns</sup>
Negative likelihood ratio	0.05 (0.01–0.22)	0.86 (0.72–1.01)*	0.89 (0.77–1.02)*

\* $p < 0.001$ ; ns = not significant ( $p > 0.05$ ) versus NEWS2

sets of observations. In the other case, the patient was admitted to ICU as a precautionary measure while on 10 L/min oxygen, though subsequently survived to discharge without requiring additional respiratory support.

Conversely, a significantly greater number of patients in the stable group recorded at least one 'false positive' NEWS2 trigger (117/163 patients, specificity 0.28 [95% CI 0.21–0.35]), compared to MEWS (47 patients, 0.71 [0.64–0.78],  $p < 0.001$ ) and qSOFA (37 patients, 0.77 [0.71–0.84],  $p < 0.001$ ). Nevertheless, the positive predictive value of NEWS2 (0.53, 95% CI 0.47–0.59) was identical to both MEWS (0.53 [0.43–0.62],  $p = 0.94$ ) and qSOFA (0.53 [0.42–0.64],  $p = 0.90$ ) (Table 2).

### Longitudinal trends in clinical scores

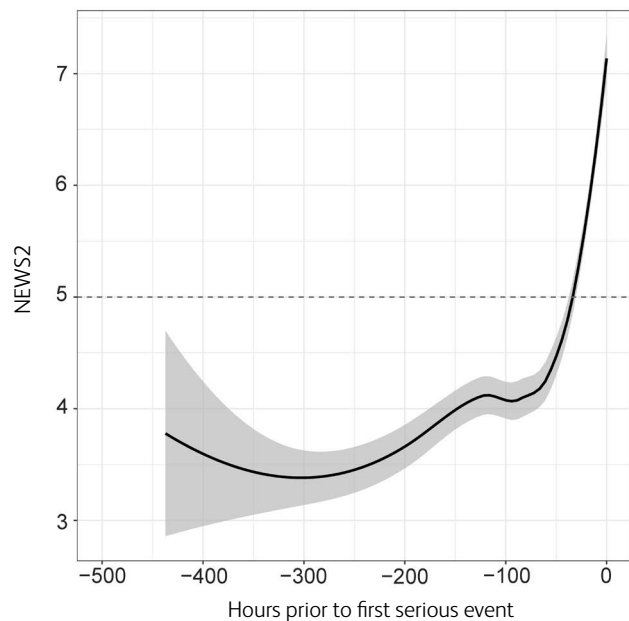
We next asked whether there were any discernible trends in early warning scores in the approach towards a serious event in patients within the deterioration group. We observed that NEWS2 scores increased in the lead up to all events, with a fitted average trend exceeding the NEWS2  $\geq 5$  trigger threshold at 33.6 hours before occurrence of first event (Fig 2). In comparison, MEWS and qSOFA scores showed only modest upward trends prior to first event, with average fitted trends failing to exceed the trigger thresholds (supplementary material S2).

### Baseline clinical scores modestly predicted subsequent deterioration

In an exploratory analysis, we investigated the prognostic utility of baseline early warning scores to predict future deterioration, together with the CURB-65 score – a widely used prognostic score in the management of community-acquired bacterial pneumonia.<sup>16</sup> NEWS2, MEWS, qSOFA and CURB-65 scores were significantly elevated at baseline in the deterioration versus stable groups (Table 1). Nevertheless, baseline NEWS2 showed only modest prognostic utility for any future serious event, with a  $ROC_{AUC}$  for NEWS2 of 0.71 (95% CI 0.65–0.77). This was significantly greater than that of MEWS (0.63 [0.57–0.69],  $p < 0.001$ ), qSOFA (0.62 [0.56–0.68],  $p < 0.001$ ), and CURB-65 (0.60 [0.53–0.66],  $p = 0.003$ ) (supplementary material S3). The prognostic metrics for future deterioration event based on previously validated pre-COVID-19 thresholds for each clinical score at baseline are shown in (supplementary material S4).

### Discussion

There is an urgent need for robust mechanisms to quickly and reliably identify those at risk of imminent clinical deterioration in COVID-19. In this study, we examined the utility of the widely



**Fig 2. Longitudinal trend in NEWS2 score in the deterioration group prior to occurrence of first serious event.** Solid line shows fitted trend (loess regression), grey shading depicts 95% confidence intervals, dashed line shows score trigger threshold.

deployed NEWS2 system to identify deteriorating inpatients admitted with COVID-19. We showed that although baseline NEWS2 is only modestly predictive, longitudinal monitoring of NEWS2 is a highly sensitive tool to identify those at risk of clinical deterioration and outperforms the alternative MEWS and qSOFA scores – albeit at the expense of relatively high numbers of 'false-positive' triggers.

The original iteration of NEWS was released in 2012, and aimed to create a standardised early warning score to replace the then fragmented and duplicatory scoring systems used across various NHS organisations.<sup>17</sup> A later update to the score – NEWS2 – was released in 2017, notably incorporating new onset confusion and highlighting NEWS2  $\geq 5$  as the threshold for urgent clinical response.<sup>3</sup> Studies of acute medical hospitalisations have validated the utility of NEWS and NEWS2 monitoring in inpatient settings to predict adverse clinical outcomes including cardiac arrest, ICU admission and death.<sup>4,18</sup> Further studies have demonstrated that a single measurement of NEWS2 at presentation to hospital emergency departments can predict important clinical outcomes

including severe sepsis, ICU admission, length of hospital stay, and mortality.<sup>19,20</sup> Furthermore, pre-hospital NEWS2 measurement by ambulance crews has also been demonstrated to predict ICU admission and mortality.<sup>21</sup> Given these data, NEWS2 is now the recommended inpatient NHS early warning score system used across all four devolved UK nations.<sup>3</sup>

Recent systematic reviews have concluded that there is no current reliable prognostic score that can predict clinical outcome in patients with COVID-19 in either inpatient<sup>22</sup> or pre-hospitalisation<sup>23</sup> settings. Although NEWS2 remains to be validated in the management of COVID-19, its potential in this setting has been highlighted by several organisations including NICE<sup>24</sup> and the Royal College of General Practitioners.<sup>25</sup> However, important questions remain as to the performance of NEWS2 in COVID-19. Profound hypoxemia, a hallmark of COVID-19, is disproportionate to other physiological perturbations that are commonly observed in bacterial sepsis, such as hypotension and altered consciousness.<sup>7,8</sup> Furthermore, the NEWS2 system scores supplemental oxygen as a binary variable, and thus does not differentiate between different rates of oxygen delivery. Accordingly, the Royal College of Physicians has issued additional precautionary guidance to recommend that any increase in supplemental oxygen in a patient treated for COVID-19 should trigger a medical review and enhanced monitoring.<sup>26</sup>

A few small cohort studies have explored the prognostic utility of baseline NEWS2 and other clinical scoring systems to predict clinical outcome in COVID-19 patients based on a single measurement at the point of hospital admission. In a Chinese study of 654 COVID-19 admissions,<sup>27</sup> baseline NEWS2 predicted mortality with a ROC<sub>AUC</sub> of 0.81 (95% CI 0.77–0.85), comparable to CURB-65 (0.85 [0.81–0.89]) and greater than qSOFA (0.73 [0.69–0.78]). In a Korean study of 110 COVID-19 inpatients,<sup>28</sup> baseline (original) NEWS predicted an event (defined as ICU admission and/or death) with a ROC<sub>AUC</sub> 0.92 (95% CI 0.84–1.00) versus 0.76 (0.62–0.90) for qSOFA; using a baseline threshold NEWS  $\geq 5$  yielded a negative predictive value of 0.98 and a positive predictive value of 0.59 for future event. In an Italian study of 68 inpatients,<sup>29</sup> NEWS2 at hospitalisation predicted ICU admission with a ROC<sub>AUC</sub> of 0.90 (95% CI 0.82–0.97). In a Norwegian study of 66 inpatients,<sup>30</sup> baseline NEWS2 predicted a composite adverse outcome of inpatient mortality and/or ICU admission with a ROC<sub>AUC</sub> of 0.79 (95% CI 0.66–0.91), versus 0.62 (0.45–0.81) and 0.58 (0.41–0.76) for qSOFA and CURB-65 respectively. These and our findings therefore suggest that while modestly predictive, pre-hospital and emergency department triage based solely on NEWS2, MEWS, qSOFA or CURB-65 scoring systems would fail to prioritise a substantial proportion of patients with COVID-19 who subsequently deteriorate.

In contrast, virtually no published data exists regarding the utility of NEWS2 in the management of COVID-19 for the purpose for which it was originally designed – namely longitudinal monitoring to identify clinical deterioration. In one report of 17 COVID-19 admissions to a UK hospital,<sup>31</sup> a high variability in NEWS2 (defined as a daily change in NEWS2  $\geq 5$ ) was observed in 7/10 patients who died, versus 0/7 of those who survived. However, no analysis of NEWS2 threshold triggers was included in this preliminary report. In a small French study of 27 COVID-19 admissions,<sup>32</sup> a modified version of the ViEWS score (an early warning score closely related to NEWS2) was shown to predict deterioration 12 hours before ICU admission with a sensitivity of 94% and specificity of 78%. In our study, we demonstrate increasing trends of NEWS2 beginning many hours prior to occurrence of a serious clinical deterioration event. We show that longitudinal NEWS2 monitoring (using the pre-existing  $\geq 5$  threshold) has a good sensitivity

for detection of clinical deterioration, with only 2/133 (1.5%) patients not meeting the NEWS2 threshold prior to deterioration. We did however observe a substantial false-positive trigger rate (117/248 [47.2%] patients who triggered did not develop a serious event), raising potential resource allocation issues for medical and ICU outreach review systems.

Several limitations to this retrospective cohort study should be acknowledged. It is probable that some patients with a NEWS2  $\geq 5$  did not go on to experience a serious event due to prompt medical review and appropriate clinical intervention. This confounding effect may have inflated the false-trigger rate, biasing the estimate of NEWS2 specificity and negative predictive value. Further research is needed, beyond the scope of this report, to determine whether these were true false positives, or were patients who responded to appropriate intervention. Similarly, we do not have the necessary data to estimate the health economic implications of such false triggers when implementing NEWS2 monitoring in practice. Data were not collected to allow us to include the ISARIC 4C mortality score,<sup>33</sup> published while this manuscript was in review, in the exploratory analysis at admission. Our results may also not be applicable to patient populations under-represented within this cohort, such as younger patients or those from BAME backgrounds. It is possible that prognostic performance could be improved through adjustment of score parameter thresholds and/or supplementation with additional metrics (such as laboratory blood tests). However, our aim was to assess the utility of the existing NEWS2 scoring system in the setting of COVID-19 as is currently implemented in clinical practice, rather than develop a new early warning score – a process that would require a larger sample size and an external validation cohort.

## Conclusion

We provide the first report to examine the utility of longitudinal NEWS2 monitoring to identify deteriorating patients hospitalised with COVID-19. Our results show that NEWS2 has adequate sensitivity to detect deteriorating patients, outperforming both MEWS and qSOFA scores in this setting. Furthermore, we show a modest prognostic value of NEWS2 at admission in predicting subsequent inpatient clinical deterioration, in keeping with preliminary results from other smaller studies. However, the reduced specificity as a result of a high proportion of seemingly ‘false-positive’ triggers raises potential resource issues in the routine implementation of NEWS2 scoring systems in COVID-19 management. Our results support the use of NEWS2 monitoring of hospitalised COVID-19 patients, as a sensitive method for identifying clinical deterioration. ■

## Supplementary material

Additional supplementary material may be found in the online version of this article at [www.rcpjournals.org/clinmedicine](http://www.rcpjournals.org/clinmedicine):

S1 – Time from first score trigger to occurrence of first serious event, within 24 hour period before first event.

S2 – Longitudinal trend in MEWS and qSOFA scores in the deterioration group prior to occurrence of first serious event.

S3 – Receiver operating characteristic (ROC) curves showing prognostic utility of a single baseline early warning score measurement on admission to predict future serious clinical event.

S4 – Comparison of diagnostic metrics of baseline clinical scores for prediction of any serious event.



## Funding

CJAD is funded by the Wellcome Trust (211153/Z/18/Z). KFB is funded by a National Institute for Health Research (NIHR) Clinical Lectureship (CL-2017-01-004), and is supported by the NIHR Newcastle Biomedical Research Centre (BRC). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. The funders had no role in study design, data collection, or decision to publish.

## References

- Docherty AB, Harrison EM, Green CA *et al*. 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.
- Cram P, Anderson ML, Shaughnessy EE. All hands on deck: Learning to 'un-specialize' in the COVID-19 pandemic. *J Hosp Med* 2020;15:314–5.
- Royal College of Physicians. *National Early Warning Score (NEWS) 2*. RCP, 2017. Available online at [www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news-2](http://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news-2) [Accessed 14 July 2020].
- 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.
- Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning Score in medical admissions. *QJM* 2001;94:521–6.
- Singer M, Deutschman CS, Seymour CW *et al*. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016;315:801–10.
- Xie J, Tong Z, Guan X *et al*. Critical care crisis and some recommendations during the COVID-19 epidemic in China. *Intensive Care Med* 2020;46:837–40.
- Tobin MJ, Laghi F, Jubran A. Why COVID-19 silent hypoxemia is baffling to physicians. *Am J Respir Crit Care Med* 2020;202:356–60.
- Lim NTY, Pan D, Barker J. NEWS2 system requires modification to identify deteriorating patients with COVID-19. *Clin Med* 2020;20:e133.
- Lillie PJ, Samson A, Li A *et al*. Novel coronavirus disease (Covid-19): The first two patients in the UK with person to person transmission. *J Infect* 2020;80:578–606.
- Baker KF, Hanrath AT, Schim van der Loeff I *et al*. COVID-19 management in a UK NHS foundation trust with a high consequence infectious diseases centre: a detailed descriptive analysis. *medRxiv* 2020:2020.05.14.20100834.
- Goyal P, Choi JJ, Pinheiro LC *et al*. Clinical characteristics of Covid-19 in New York City. *N Engl J Med* 2020;382:2372–74.
- Stock C, Hielscher T. *DTComPair: comparison of binary diagnostic tests in a paired study design*. <http://CRAN.R-project.org/package=DTComPair> [Accessed 14 July 2020].
- Robin X, Turck N, Hainard A *et al*. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011;12:77.
- Department of Health and Social Care. *Coronavirus (COVID-19): notice under regulation 3(4) of the Health Service (Control of Patient Information) Regulations 2002 – general*. DHSC, 2020. [www.gov.uk/government/publications/coronavirus-covid-19-notification-of-data-controllers-to-share-information/coronavirus-covid-19-notice-under-regulation-34-of-the-health-service-control-of-patient-information-regulations-2002-general](http://www.gov.uk/government/publications/coronavirus-covid-19-notification-of-data-controllers-to-share-information/coronavirus-covid-19-notice-under-regulation-34-of-the-health-service-control-of-patient-information-regulations-2002-general).
- Lim WS, van der Eerden MM, Laing R *et al*. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 2003;58:377–82.
- Royal College of Physicians. *National Early Warning Score (NEWS): standardising the assessment of acute-illness severity in the NHS*. RCP, 2012. [www.rcplondon.ac.uk/file/9559/download](http://www.rcplondon.ac.uk/file/9559/download) [Accessed 14 July 2020].
- Spångfors M, Arvidsson L, Karlsson V, Samuelson K. The National Early Warning Score: Translation, testing and prediction in a Swedish setting. *Intensive Critical Care Nurs* 2016;37:62–7.
- Alam N, Vegting IL, Houben E *et al*. Exploring the performance of the National Early Warning Score (NEWS) in a European emergency department. *Resuscitation* 2015;90:111–5.
- Keep JW, Messmer AS, Sladden R *et al*. National early warning score at Emergency Department triage may allow earlier identification of patients with severe sepsis and septic shock: a retrospective observational study. *Emerg Med J* 2016;33:37.
- Silcock DJ, Corfield AR, Gowens PA, Rooney KD. Validation of the National Early Warning Score in the prehospital setting. *Resuscitation* 2015;89:31–5.
- Wynants L, Van Calster B, Collins GS *et al*. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* 2020;369:m1328.
- Greenhalgh T, Treadwell J, Burrow R, Roberts N. *NEWS (or NEWS2) score when assessing possible COVID-19 patients in primary care? Centre for Evidence Based Medicine, 2020*. [www.cebm.net/covid-19/should-we-use-the-news-or-news2-score-when-assessing-patients-with-possible-covid-19-in-primary-care/](http://www.cebm.net/covid-19/should-we-use-the-news-or-news2-score-when-assessing-patients-with-possible-covid-19-in-primary-care/) [Accessed 22 July 2020].
- National Institute for Health and Care Excellence. *COVID-19 rapid guideline: managing suspected or confirmed pneumonia in adults in the community*. NICE guideline NG165. NICE, 2020. [www.nice.org.uk/guidance/ng165](http://www.nice.org.uk/guidance/ng165) [Accessed 22 July 2020].
- Royal College of General Practitioners. *COVID-19 evidence summary: clinical management*. RCGP, 2020. Available at <https://elearning.rcgp.org.uk/mod/page/view.php?id=10570> [Accessed 22 July 2020].
- Royal College of Physicians. *NEWS2 and deterioration in COVID-19*. RCP, 2020. Available at [www.rcplondon.ac.uk/news/news2-and-deterioration-covid-19](http://www.rcplondon.ac.uk/news/news2-and-deterioration-covid-19) [Accessed 22 July 2020].
- Fan G, Tu C, Zhou F *et al*. Comparison of severity scores for COVID-19 patients with pneumonia: a retrospective study. *Eur Respir J* 2020:2002113.
- Jang JG, Hur J, Hong KS *et al*. Prognostic accuracy of the SIRS, qSOFA, and NEWS for early detection of clinical deterioration in SARS-CoV-2 infected patients. *J Korean Med Sci* 2020;35:e234.
- Gidari A, De Socio GV, 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 (Lond)* 2020;52:698–704.
- 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.
- Sze S, Pan D, Williams CML *et al*. Letter to the editor: Variability but not admission or trends in NEWS2 score predicts clinical outcome in elderly hospitalised patients with COVID-19. *J Infect* 82:159–98.
- Meylan S, Akrou R, Regina J *et al*. An Early Warning Score to predict ICU admission in COVID-19 positive patients. *J Infect* 2020;81:816–46.
- 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.

**Address for correspondence:** Christopher JA Duncan, Translational and Clinical Research Institute, Newcastle University, UK, NE2 4HH.  
**Email:** [christopher.duncan@newcastle.ac.uk](mailto:christopher.duncan@newcastle.ac.uk)  
**Twitter:** @chrisjaduncan