Is NEWS2 the optimal evidence-based surveillance tool for all respiratory patients or does it just represent the beginning of an iterative development process?

Dominick Shaw[1,2]

Andrew Fogarty[3]

1 Department of Respiratory Sciences, University of Leicester, Leicester, UK.

2 Institute for Lung Health, NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Leicester, UK

3 NIHR Nottingham Biomedical Research Centre (BRC), Nottingham University Hospitals NHS Trust, University of Nottingham, Nottingham, UK

Corresponding author- Dominick Shaw: des21@le.ac.uk

Summary

Medical practice is built on foundations of evidence-based medicine. Hence, the more common the clinical intervention, the more comprehensive the evidence on which that intervention should be based. Although the widespread adoption of a national early warning score in the United Kingdom has led to improvements in the delivery of care, it should be considered as providing a foundation that can be refined and developed, and there is still a need for critical reflection and evaluation of early warning scores, particularly for individuals with chronic respiratory disease, in order to optimise patient monitoring, predict deterioration and guide intervention.

Background

Early warning systems are used globally to predict clinical deterioration, with different parameters and triggers. In 2017 the UK introduced NEWS2, the second iteration of the National Early Warning Score [1]. It was designed in conjunction with the Royal College of Physicians London, and uses a point-based system in its afferent limb to allocate a score of 0-3 depending on how deranged a patient's basic physiology is from normality (see Figure 1). It was designed as a screening tool to enable early identification of patients at risk of deterioration by highlighting deviation of vital sign parameters from a predefined normal physiological range, and is applied regularly throughout a patient's hospital admission. It is used in most Hospital Trusts and all ambulance services in the UK and it is also used in parts of Europe, the USA, Canada and Asia.

Figure 1 Overview of the NEWS2 score

NEWS2 is uniformly applied to adult medical and surgical patients from initial contact throughout the course of the illness and is the most common medical assessment performed. This combined score is interpreted into an efferent limb that mandates patient review based on staff competency and urgency. Individual hospitals have their own response systems based on NEWS, but many interpret this with an increasing seniority of staff corresponding to the score e.g., 0-4 ward-based response; 5-6 medical review, ≥7, urgent response by senior doctor. In the UK NEWS is part of a wider rapid response system within hospitals.

The idea behind early warning scores and NEWS is logical. It was known that patients with potentially treatable conditions were not being escalated and were suffering harm from what was termed "failure to recognise and rescue", and that these patients could be flagged by routine observations in the period prior to deterioration [2]. For example there are 290000 in-hospital adult cardiac arrests each year in the United States.

However, the logic was derived mostly from post-operative complications in surgical patients who had deranged physiology which was not returning to baseline. The concept was disseminated into medical specialities with a much wider range of chronic conditions and not tested in any prospective or randomised controlled studies, although studies with linked interventions including a pragmatic medical response have been performed [3,4].

The initial uptake of early warning scores was pushed by medical device companies selling hardware and software solutions and catalysed by guidelines, commissioning requirements and quality standards. Together, this led to the incremental adoption of electronic vital sign measurement, which allowed simple construction of early warning scores as a composite of these measurements. NEWS became firmly embedded in the National Health Service when insurance premiums for hospitals were reduced for those employing automated electronic early warning systems, and almost overnight senior physicians found themselves puzzled when their junior colleagues started saying "this patient is scoring" whilst looking at their (hospital) mobile phone.

The introduction of a standardised approach to clinical deterioration has led to improved education and training, more robust clinical handover, improved acuity metrics and more efficient bed management, but these benefits occurred in conjunction with an increased workload [5] with a lack of qualitative analysis on the downstream effects (staff satisfaction, workforce requirements, alarm fatigue etc) of introducing an automated early warning score. A recent review highlighted that 'Early warning scores might not perform as well as expected and therefore they could have a detrimental effect on patient care'[6]. This partly reflects the difficulty in constructing well designed and controlled studies, the relative rarity of the outcome measure (ICU admission, death or cardiorespiratory arrest), the problem translating observational *post-hoc* population derived evidence into individual outcomes, the fact that scoring systems were originally only designed to show a patient was sick, not predict a change in clinical management, and the somewhat tautological practice of repeatedly using a score to both predict and prevent the same outcome.

The widespread adoption and enthusiasm for early warning scores hides several other issues; specifically there are problems with co-linearity of the component parts in early warning scores, and score performance varies markedly across different patient groups. This is a particular problem in patients with chronic respiratory disease given the interdependent relationship between blood oxygen levels, oxygen administration and respiratory rate (which are all part of the NEWS score). New data highlighting problems with pulse oximetry accuracy [7,8] and questioning the optimal target level for oxygen saturation [9] only add to the need for the evidence-base of NEWS2 to be revisited for patients with respiratory disease.

Specific problems of early warning scores in respiratory disease

Initially, NEWS was not tailored for respiratory disease despite the obvious problems of extrapolating an early warning score designed for post-surgical patients to populations with chronic disease. Particular concerns were raised about the risks from excess oxygen administration to patients at risk of hypercapnic respiratory failure (in COPD, severe obesity, neuromuscular disease) especially given that COPD exacerbations are the second most common cause of hospital admission in the UK, with 20% of patients admitted with COPD having a respiratory acidosis at presentation and evidence showing that targeted oxygen saturations are associated with reduced mortality [10].

Consequently in 2017 following discussion with the British Thoracic Society, NEWS2 was introduced. This had two scales for oxygen saturation with the second scale reserved for patients with confirmed hypercapnia. In scale 2 the target oxygen saturations were set at a lower level (88%–92%) and the scoring adjusted accordingly, although its introduction was not based on any new data, nor formally prospectively assessed or piloted (see Figure 2) and led to debate [11–13].

A comparison of NEWS variants with the DECAF score (dyspnoea, eosinopenia, consolidation, acidaemia, atrial fibrillation) in patients with COPD found that adapting NEWS2 to apply scale 2 in all patients, including those with normocapnia, improved performance compared to NEWS2 and NEWS, whilst also safely reducing alerts. Although the admission DECAF score was the superior prognostic tool, and guided care which included hospital at home, it complemented rather than replaced NEWS2 [14].

NEWS2 requires confirmation of hypercapnia before applying scale 2 in COPD and other conditions at heightened risk, although evidence suggests that treating all patients with COPD with target saturations of 88%-92% may improve outcome [15] and that one in four cases of respiratory acidaemia requiring NIV are caused by oxygen toxicity [16].

NEWS2 is now mandated in the management of patients with respiratory disease (and all other adult medical in-patients), although its performance is difficult to delineate from external factors. For example a UK audit found that mortality in patients with respiratory disease was significantly lower in hospitals with respiratory support units (23%) than in those without (35%) [17], despite all utilising NEWS2.

Figure 2 Respiratory aspects of NEWS2

Many respiratory diseases cause deranged vital signs including chronic hypoxia and tachycardia which are measured using pulse oximetry. However not every health care worker recognises that despite its ubiquity and simplicity pulse oximetry (SpO_2) can be affected by hypoperfusion and smoking related elevated carbon monoxide levels. Moreover it has now been recognised once again that SpO_2 accuracy can also be affected by darker skin pigmentation [18] specifically giving higher readings of blood oxygen saturation than the true value. This limitation may have contributed to the later escalation and transfer of patients from non-white ethnic communities for intensive care treatment during the Covid-19 pandemic [19].

Given that the use of oxygen as a drug with both therapeutic and side effects, it is also surprising that there is no widely used standardised method for accurately recording oxygen delivery. Although NEWS2 recommends standardising and recording oxygen delivery using the British Thoracic Society oxygen delivery device codes, this is often not completed, possibly because these codes are not on the main NEWS observation chart. The issue of coding supplementary oxygen for an early warning score is challenging as there are numerous methods for oxygen delivery which include devices providing low flow (less than 15L/min of oxygen), high flow, nasal high flow, non-invasive ventilation and controlled flow devices such as Venturi masks are available with hospitals using different devices. It does require consideration as the NEWS score itself only delineates between air or oxygen administration (with two points added for oxygen) and clinical experience is required to discern between patients receiving 1L oxygen/min as opposed to 15L/min.

What are the solutions?

In an ideal world a predictive score to improve clinical outcomes would be easy to measure, derive and implement, have a high specificity and sensitivity for potentially preventable clinical deterioration, and would not consume excess resources, or lead to alert fatigue. It would target the correct oxygen saturation target for specific patient populations and be fully tested prospectively for efficacy and effectiveness. Importantly it would be easy to standardise across different organisations and institutions, a key benefit of NEWS2, and not replace clinical experience [20].

Firstly, standardising oxygen delivery (and terminology) may reduce some of the inherent score variability. Delineating between “on oxygen”, rather than “needs oxygen”, and classifying mask type, flow rate and need for non-invasive ventilation would be feasible, especially with the advent of electronic patient records. Studies have shown that a weighted fraction of inspired oxygen (FiO_2) parameter incorporated into a NEWS-FiO_2 model has better predictive ability than NEWS2 alone [13,21]. Standardising oxygen delivery also allows changes in oxygen requirements to be assessed; a study of patients on a respiratory unit found that of 765 admission episodes, 8971 out of 35991 observation sets (25%) had a NEWS2 ≥5, but only 586 (7%) of these led to an urgent medical review. In contrast 15% of interventions were associated with a NEWS2 of <5, below the normal trigger for a clinical review. A rising oxygen requirement rather than the absolute oxygen requirement appeared to influence the decision to intervene and request a medical review [22].

Secondly a respiratory specific score could be added to NEWS2, and the SpO_2 to FiO_2 ratio (S/F ratio) has been suggested as an alternative. This is a simple summary score with just two components that has similar prognostic properties to NEWS2 in patients with Covid-19 [23] despite having a ceiling effect, because at high saturations SpO_2 is not dependent on pulmonary oxygenation as blood is close to maximally oxygenated, and as estimating FiO_2 from nasal cannulae (the commonest method of delivery in hospital) can be difficult as the FiO_2 can vary. The ROX index ([SpO_2/FiO_2]/respiratory rate) is a variant of the S/F ratio with three components, also performed better than NEWS2 in patients with Covid-19 [24] despite the measurement error that accompanies respiratory rate measurements [25]. Alternatively, the current NEWS2 system could be modified to improve sensitivity and specificity. For example, adding the maximum score in the preceding 24 hours to the most recently recorded NEWS2 improved score performance at predicting death in 24 hours in a large respiratory population [26].

Thirdly more complicated models using observational statistical analysis of large data sets have demonstrated improved predictive performance. In one respiratory disease model, time series features were entered into logistic regression analyses to derive a dynamic early warning score which had improved predictive ability compared to NEWS2, based clinically significant deteriorations [27]. Novel machine learning techniques are also inevitably being studied to develop early warning scores, but the more granular the data for score parameters, the harder it is to apply the score across a population rather than at an individual, or specialised unit, level.

Fourthly, given that using NEWS2 Scale 2 in all presentations of COPD exacerbations, regardless of $PaCO_2$, has been shown to improve performance whilst also safely reducing false alerts, applying this protocol could remove the need for an arterial blood gas and senior review, simplifying the assessment of COPD exacerbations, and could be considered for other conditions at risk of hypercapnia.

Finally, newer oxygen delivery technologies and digital monitoring may help improve care. Constant measurement of SpO_2 could help establish individualized parameters for vital sign monitoring, and enable aggregated continuous data collection to identify subtle deviations in physiology. This could be allied to automated oxygen titration which includes a feedback loop targeting oxygen delivery to oxygen saturations. Data has shown that in hypoxaemic participants with stable cardiorespiratory disease applying automated oxygen control with bilevel or CPAP therapies took less time to reach the target SpO_2 range compared to nasal high flow therapy, while the time spent within the target range was similar [28].

Conclusion

Early warning scores and NEWS2 have led to improvements in the identification of patients at risk of deterioration in hospital settings. However whether NEWS2 is really the best approach for patients with chronic respiratory disease (or acute presentations of chronic disease) has not yet been properly questioned or critically evaluated.

The respiratory community should consider developing a more accurate tool for the identification of both types of respiratory failure, taking into account the relationship between SaO_2, FiO_2 and respiratory rate, rather than continue to passively adopt what may be a sub-optimal surveillance system and alternatives including the S/F ratio should be formally evaluated. Any future intervention needs to be standardised, appropriate for respiratory disease, scalable, tested in suitably designed studies and be interoperable with current electronic patient systems. Ideally, it must not lead to alert fatigue, nor supplant clinical acumen. Over time scoring systems, or oxygen delivery may become individualised by utilising wearable technology.

It is now time to reflect on how can we develop the research evidence base to determine the best early warning score for our patients with lung disease. Understanding the historical context of how we came to use NEWS2, and appreciating that this may be the start of an iterative process, rather than the end, is an important step in the evolution of the optimal surveillance system. The widespread adoption of NEWS2 has led to improvements in the delivery of care but as a respiratory community we should be striving to develop the best possible surveillance systems for our patients. Rigorous evaluation of new approaches is needed at the level of institutions, not individuals, to allow the design and implementation of prospective clinical trials. This will require collaboration and buy-in at a national level to move forward.

Word count 2444

a. Contributorship

DS is the guarantor and conceived the idea for the article, both authors contributed equally to the writing

b. Funding Statement

no funding declared

c. Competing Interests

none

d. Acknowledgement

The authors would like to thank Professor Tim Coates for his input and www.russanimates.co.uk for the image

e. AI use statement

AI was not used in the preparation of any part of this manuscript

Figure 1 NEWS2 Score

Royal College of Physicians, London. news2-executive-summary_0. London, England: Royal College of Physicians of London

Figure 2 Respiratory aspects of NEWS2

Royal College of Physicians, London. news2-executive-summary_0. London, England: Royal College of Physicians of London

Figure 3 How do we make sense of NEWS scores in patients with respiratory disease?

References

1 Royal College of Physicians, London. *news2-executive-summary_0*. London, England: Royal College of Physicians of London .

2 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.

3 Haegdorens F, Monsieurs KG, De Meester K, *et al.* An intervention including the national early warning score improves patient monitoring practice and reduces mortality: A cluster randomized controlled trial. *J Adv Nurs*. Published Online First: 22 April 2019. doi: 10.1111/jan.14034

4 Mcgaughey J, Alderdice F, Fowler R. Outreach and Early Warning Systems (EWS) for the prevention of Intensive Care admission and death of critically ill adult patients on general hospital wards. *Cochrane Database Syst Rev*. 2007.

5 Forster S, McKeever TM, Shaw D. Effect of implementing the NEWS2 escalation protocol in a large acute NHS trust: a retrospective cohort analysis of mortality, workload and ability of early warning score to predict death within 24 hours. *BMJ Open*. 2022;12:e064579.

6 Gerry S, Bonnici T, Birks J, *et al.* Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. *BMJ*. 2020;369:m1501.

7 Valbuena VSM, Seelye S, Sjoding MW, *et al.* Racial bias and reproducibility in pulse oximetry among medical and surgical inpatients in general care in the Veterans Health Administration 2013-19: multicenter, retrospective cohort study. *BMJ*. 2022;378:e069775.

8 Crooks CJ, West J, Morling JR, *et al.* A requirement for the use of supplementary oxygen to guide medical treatment decisions may introduce bias against non-white individuals. *Eur Respir J*. 2024;63:2302320.

9 Buell KG, Spicer AB, Casey JD, *et al.* Individualized treatment effects of oxygen targets in mechanically ventilated critically ill adults. *JAMA*. 2024;331:1195.

10 Turner AM, Lim WS, Rodrigo C, *et al.* A care-bundles approach to improving standard of care in AECOPD admissions: results of a national project. *Thorax*. 2015;70:992–4.

11 Pimentel MAF, Smith GB, Redfern OC, *et al.* Reply to: NEWS2 needs to be tested in prospective trials involving patients with confirmed hypercapnia. Resuscitation. 2019;139:371–2.

12 O'Driscoll R, Bakerly N, Murphy P, *et al.* NEWS2 needs to be tested in prospective trials involving patients with confirmed hypercapnia. *Resuscitation*. 2019;139:369–70.

13 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.

14 Echevarria C, Steer J, Bourke SC. Comparison of early warning scores in patients with COPD exacerbation: DECAF and NEWS score. *Thorax*. 2019;74:941–6.

15 Echevarria C, Steer J, Wason J, *et al.* Oxygen therapy and inpatient mortality in COPD exacerbation. *Emerg Med J*. 2021;38:170–7.

16 NCEPOD. Inspiring Change.

17 National Respiratory Support Audit 2023. https://www.brit-thoracic.org.uk/quality-improvement/clinical-audit/national-respiratory-support-audit-2023 (accessed 12 July 2024)

18 Crooks CJ, West J, Morling JR, *et al.* Pulse oximeter measurements vary across ethnic groups: an observational study in patients with COVID-19. *Eur Respir J*. 2022;59. doi: 10.1183/13993003.03246-2021

19 Crooks CJ, West J, Morling JR, *et al.* Differential pulse oximetry readings between ethnic groups and delayed transfer to intensive care units. *QJM*. 2023;116:63–7.

20 Knack SKS, Scott N, Driver BE, *et al.* Early physician gestalt versus usual screening tools for the prediction of sepsis in critically ill emergency patients. *Ann Emerg Med*. 2024;84:246–58.

21 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.

22 Gonem S, Lemberger J, Baguneid A, *et al.* Real-world implementation of the National Early Warning Score-2 in an acute respiratory unit. *BMJ Open Respir Res*. 2024;11:e002095.

23 Crooks CJ, West J, Morling JR, *et al.* Supplementary inspired oxygen fraction is a simple clinical tool that predicts clinical deterioration. *IJTLD OPEN*.

24 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.

25 Fogarty AW, Card T, Shaw D, *et al.* Error in respiratory rate measurement by direct observation impacts on clinical early warning score algorithms. *Emerg Med J*. Published Online First: 1 December 2022. doi: 10.1136/emermed-2022-212689

26 Forster S, McKeever TM, Churpek M, *et al.* Predicting outcome in acute respiratory admissions using patterns of National Early Warning Scores. *Clin Med*. 2022;22:409–15.

27 Gonem S, Taylor A, Figueredo G, *et al.* Dynamic early warning scores for predicting clinical deterioration in patients with respiratory disease. *Respir Res*. 2022;23:203.

28 Kirton L, Kung S, Bird G, *et al.* Automated oxygen titration with non-invasive ventilation in hypoxaemic adults with cardiorespiratory disease: a randomised cross-over trial. *BMJ Open Respir Res*. 2024;11:e002196.