Towards NEWS 3.0: Development and validation of an oxygen therapy adjusted National Early Warning Score. A statistical analysis plan.

Date: 14th October 2024

Chief investigator:

Name	Charlotte Harrison
Position	Clinical Research Fellow
Employing organisation	University of Oxford
Address	Critical Care Research Group Kadoorie Centre for Critical Care Research and Education Level 3, John Radcliffe Hospital Headley Way Oxford OX3 9DU
Email	Charlotte.harrison@ndcn.ox.ac.uk

Chief investigator signature:

Investigators:

Jonathan Bedford, University of Oxford

Stephen Gerry, University of Oxford

Oliver Redfern, University of Oxford

Chris Subbe, Ysbyty Gwynedd, Bangor University

Peter Watkinson, University of Oxford

Funder: This research is supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Confidentiality statement:

Contents

1. Abstract

1.1 Introduction

Early warning scores (EWS) use routinely measured vital signs to identify patients at risk of deterioration. The National Early Warning Score 2 (NEWS2) is mandated for use by NHS England in all acute trusts. Within NEWS2, oxygen therapy is scored using a binary approach — patients receive 2 points if they are receiving supplemental oxygen and 0 points if they are breathing room air. Examples of binary oxygen EWS are widespread, including other nationally employed EWS. Some have expressed concerns that deterioration is missed when patients require an escalating amount of supplemental oxygen in the absence of other vital sign derangements.

This study aims to adapt and validate a version of NEWS2 that accounts for the amount of supplemental oxygen therapy the patient receives (NEWS2-O2).

1.2 Methods

We will use data from four NHS healthcare trusts across the UK to develop and validate an oxygen therapy adjusted NEWS2 score. We will also explore the time trends of oxygen requirement and other vital signs prior to deterioration. We will perform an external validation of other oxygen adjusted vital sign only EWS identified in a prior systematic review.

1.3 Generalisability and implications

This work will explore the quantification and handling of oxygen therapy within an EWS. We aim to provide an example to the NEWS Development Group on how oxygen can be incorporated into the next iteration of NEWS.

2. Introduction

2.1 Background and rationale

Early warning scores (EWS) are used in hospitals to recognise deteriorating patients(1, 2). The National Early Warning Score, now in its second iteration (NEWS2), is mandated across all acute hospitals in England(3). NEWS2 assigns points to vital signs: respiratory rate, oxygen saturations, oxygen therapy, systolic blood pressure, heart rate, consciousness level and temperature(3). The number of assigned points increases when vital signs are more abnormal. NEWS2 assigns two points if patients are receiving any supplemental oxygen(3). However, the amount of required oxygen is not considered i.e. patients requiring two litres/minute would receive the same number of points as those requiring fifteen.

The binary approach to oxygen supplementation may limit recognition of the deteriorating patient (4, 5). A deteriorating patient may have an escalating oxygen requirement with minimal changes to other vital signs, causing no change to their NEWS2 score, and hence delaying recognition of deterioration(6). This limitation became particularly apparent during the COVID-19 pandemic(6). Since, there have been continued calls to develop the oxygen therapy component in NEWS2(4-10). Previously published models that increase the complexity of oxygen scoring (beyond the binary system in NEWS2) report an improved prediction of deterioration(4, 5).

However, given the widespread use of NEWS2 across the UK(11), any updates that increase complexity require strong evidence that they will improve patient care. In practice, ward staff may already be escalating patients with an increasing oxygen requirement for clinical review, regardless of change to total NEWS2 score. Indeed, this approach is currently recommended by the RCP(3), supported by recent studies exploring factors leading to escalation(12).

We have recently undertaken a systematic review and meta-analysis of oxygen delivery in early warning scores (PROSPERO ID:CRD42024443362) to inform this statistical analysis plan. The review identified 16 published vital sign only EWS that incorporate oxygen therapy in a more detailed manner than the binary scoring seen in NEWS2. These 16 EWSs used a range of methods to define oxygen requirement (e.g. flow rate, delivery device type, estimated fraction of inspired oxygen (FiO2)), and 15/16 models were at high risk of bias due to their development methods. Meta-analysis demonstrated a significant improvement in model performance in graded oxygen EWS than binary oxygen EWS. The review demonstrates that predictive performance of an EWS increases with graded oxygen, highlights precedence for how oxygen therapy can be handled in a model and reveals the limitations of their development.

Our aim is to develop upon the findings of the systematic review to develop an oxygen therapy adjusted NEWS2 score.

2.2 Aims and Objectives

Primary objective:

1. To use the multi-site HAVEN data and best-practice prediction model methods to develop and validate an oxygen therapy adjusted NEWS2 score (NEWS2-O2), to predict deterioration in hospitalised adults.

Secondary objectives:

1. To explore the temporal relationships between oxygen requirement, vital signs and NEWS2 score preceding deterioration in hospitalised adults.
2. To externally validate a series of early warning scores with oxygen graded beyond binary scoring and compare their ability to predict deterioration to NEWS2.

3. Methods

Study design

We will perform a retrospective cohort study using the multi-centre data collected as part of the "Hospital Alerting Via Electronic Noticeboard (HAVEN)" project.

We will report the study in line with the Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement(13).

Setting

Four hospital trusts in the UK, between 2014 and 2022.

- Oxford University Hospitals NHS Foundation Trust (OUH): January 2014 - October 2021

OUH is a hospital group serving a local population of approximately 655,000 people with around 1,465 beds. The trust is composed of the John Radcliffe Hospital (the main university teaching hospital, including the tertiary referral centre for trauma, cardiology, neurosurgery and general medical and surgical admissions), the Churchill Hospital (specialist renal transplant and cancer referral centre), the Horton General Hospital (the district general hospital) and the Nuffield Orthopaedic Centre (elective orthopaedic surgery, bone infection unit and rheumatology centre).

- Royal Berkshire NHS Foundation Trust (RBH): October 2016 – December 2022

RBH serves approximately 1,000,000 people in West Berkshire, Reading and Wokingham. The trust provides acute medical and surgical services, as well as specialist cancer, dialysis and eye surgery services(14).

- Lancashire Teaching Hospitals NHS Foundation Trust (LTH): January 2015 – March 2021

LTH provides acute and specialist care services to the population of Lancashire and South Cumbria, of approximately 1.5 million people. The trust includes the Royal Preston Hospital and Chorley & South Ribble Hospital. The trust acts as a regional specialist centre for allergy and immunology, cancer, disablement services, trauma, neurology, neurosurgery, renal and vascular surgery(15).

- South Warwickshire NHS Foundation Trust (SWFT): April 2016 - June 2021

SWFT is composed of four hospitals; Warwick Hospital, Leamington Spa Hospital, Stratford Hospital and Ellen Badger Hospital, providing approximately 500 beds to a population of 270,000 people of South Warwickshire and surrounding areas(16). The sites from SWFT contributing data include Warwick Hospital, Stratford Hospital and Ellen Badger Hospital.

Data sources

The routinely collected data were stored across different clinical information systems in all four trusts. Data extracted included; admission administrative information (time and date of admission, discharge, transfers within the hospital), diagnoses (according to the International

statistical Classification of Disease and related health problems (ICD-10) codes), vital sign observations, details of oxygen therapy and patient demographics.

Primary Objective:

To develop and validate an oxygen therapy adjusted NEWS2 score at predicting deterioration in hospitalised adults.

Study Population

We will include all adults (age ≥ 16) admitted to hospital. Admissions with no recorded vital signs will be excluded. We will exclude admissions which were < 24 hours in duration, to account for emergency department attendances, same day ambulatory unit attendances and patients admitted for palliation. We will also exclude admissions to the emergency department and admissions direct to the ICU from ED.

We will also exclude observation sets taken after admission to the ICU.

Each patient admission will be handled as an individual source of data, meaning there could be multiple admissions analysed for the same patient within the study period.

We will consider a subgroup analysis in patients at risk of type 2 respiratory failure, and only those with an oxygen requirement.

Repeated measures

We will use a random sampling method of one observation set per admission. In a sensitivity analysis, we will include all observations, essentially treating them as independent(17). Other resampling procedures will be evaluated as part of a sensitivity analysis. Multiple imputation will be performed prior to random sampling.

Outcome

Our primary outcome of deterioration will be a composite of unplanned ICU admission and death. Where a patient experienced more than one type of deterioration, the first event will be used. Secondary outcomes will be unplanned ICU admission and death treated separately. We will use a time horizon of 24 hours, consistent with other literature on NEWS/NEWS2(7, 18-20).

Sample size

Model development:

For model development, we used the sample size formula for a binary outcome as per Riley et al(21, 22), employed via the "pmsampsize" package in R statistical software. In the previously published HAVEN dataset(23), the outcome prevalence was 1.2% in the model development cohort. Given the current NEWS2 has 25 potential variable categories(3), we will consider up to 35 model parameters, and anticipate a c-statistic of 0.8. Overall, this will require a minimum of 19,390 patients with 233 events. We will not constrain the sample where more patients/admissions are available.

Model validation:

We used guidance provided by Riley et al(24) to calculate the sample size required for external validation, using the "pmvalsampsize" package in R statistical software. The calculation is based on the following assumptions, taken from an FiO2 adjusted NEWS2 score published on the HAVEN dataset(4).

- The O/E ratio. The O/E will be based on the outcome prevalence (3.5%), with a 95% confidence interval width of 0.2.
- Calibration slope. We will assume good model calibration with intercept = 0 and slope = 1.
- Linear predictor. We will assume the linear predictor is normally distributed
- C-statistic. We will assume a target c-statistic of 0.876

This estimates a final minimum sample size required for validation of 10,601, including 372 events. We will not constrain the sample where more patients/admissions are available. These assumptions will be updated to calculate an accurate sample size once model development is completed.

Statistical analysis methods

1. Missing data

Distribution of variables will be inspected, and any biologically implausible results (according to clinical expertise) will be regarded as missing. It is common for prediction model studies of deterioration to perform complete case analysis(2), despite the implications this has on the overall model(25). Instead, we will use multiple imputation, which is widely considered the best approach(2, 26, 27), and can account for data missing completely at random (MCAR) and data missing not at random (MNAR). Multiple imputation will be completed prior to random sampling.

2. Model building

The new model will keep the intrinsic NEWS2 structure, keeping the categories and points assigned to heart rate, respiratory rate, consciousness, blood pressure and temperature unchanged.

Overall, we aim to expand upon the supplemental oxygen category, by increasing the number of points available for increasing oxygen requirement.

We will aim to quantify supplemental oxygen requirement based on estimated FiO2. As part of preliminary analyses (see secondary objective 1), we will explore different formulae(28-30) for estimating FiO2 for devices where FiO2 is not established (e.g. nasal cannulae). We will choose the formula which causes the earliest rise in FiO2 in patients requiring supplemental oxygen, in order to increase speed of detection of deterioration.

Preliminary analyses to be undertaken in secondary object 2 will benchmark the performance of a range of oxygen graded EWS. We will use the top performing (according to discrimination and decision curve analysis) five models to guide the number of categories, points and potential interactions with oxygen saturations (e.g. the SpO2/FiO2 ratio) applied to supplemental oxygen therapy. If the SpO2/FiO2 ratio is selected based on the benchmarking exercise, SpO2 will only be handled within the ratio, and will not also be counted separately.

The new model will be developed using multivariable logistic regression, allowing for non-linear terms, and we may consider interactions between oxygen therapy and other variables.

After modelling, we will simplify the oxygen therapy variable into categories that reflect the NEWS2 structure (e.g. 0, 1, 2, 3). We will explore a range of approaches to categorise oxygen therapy, including pre-specification on clinical grounds, inflexion analysis of non-linear variables and decision tree analysis.

3. Internal and external validation

We will develop our model on all eligible admissions in all four sites for model development. Then, we will assess performance using 4-fold internal-external cross validation(31, 32); by rerunning the model building procedure on three sites and assessing performance on the held-out site, repeated four times. Clustering will be at the site level (e.g. OUH, RBH, SWFT, LTH). This is in line with best practice guidance for model development and evaluation(32-34), as it maximises the use of data for model development and provides an assessment of heterogeneity and generalisability in model performance across sites.

4. Performance assessment

We will report patient demographics and clinical information using descriptive statistics. We will compare the performance of the final model compared to NEWS2.

We will report an assessment of calibration (calibration plots, calibration-in-the-large), discrimination (area under receiver operating characteristic), precision-recall curves, performance analysis at key NEWS2 thresholds (score ≥5, score ≥ 7) and decision curve analysis.

On the final model, we will also plot the relationship between the score and risk of outcome. We will explore agreement between models, such as back-to-back histograms and likelihood ratio tests for the fraction of new prognostic information provided in the new model versus NEWS2, (e.g. (35, 36)).

We will also perform a subgroup analysis in patients with confirmed hypercapnia, as we anticipate their interaction with oxygen therapy and risk of deterioration to be different to patients without hypercapnia.

Analysis will be repeated for the secondary outcomes.

Secondary Objective 1:

To explore the temporal relationships between oxygen requirement, vital signs and NEWS2 score preceding deterioration in hospitalised adults.

Study Population

We will include all adults (age ≥ 16) admitted to hospital. Admissions with no recorded vital signs will be excluded.

We will exclude admissions which were < 24 hours in duration, admissions to the emergency department and admissions direct to the ICU from ED.

We will include all vital sign observation sets from 7 days prior to hospital discharge or deterioration (defined below). We will exclude any vital signs occurring after the first event.

Each new patient admission will be handled as an individual source of data, meaning there could be multiple admissions analysed for the same patient within the study period.

Repeated measures

We anticipate that there will be multiple observation sets per patient. We will explore different approaches to handling this, as per previous studies in this area(17). We will employ the following approaches:

1. Using all observation sets, making no adjustments, essentially treating all observations as independent.
2. As per (1) but weighting according to number of vital signs per patient within the GAMLSS package.
3. Random sampling of two sets of observations per day per admission.

Outcome

We will model time to deterioration, a composite of unplanned ICU admission or death. Where a patient experiences more than one deterioration event, we will use the first event. Our secondary outcome will be modelling time to unplanned ICU admission or death treated separately.

In those that do not experience a deterioration, we will model time to hospital discharge.

Sample Size

This aspect of the study is purely descriptive; hence the sample size will be determined by the number of deteriorations available to the research team at the time of analysis.

Statistical analysis methods

We will summarise patient demographics and clinical information using descriptive statistics.

We will model changes in individual vital signs over time . The vital signs of interest include oxygen therapy (yes/no), oxygen flow rate, estimated FiO2, heart rate, respiratory rate, systolic blood pressure, diastolic blood pressure, oxygen saturations, temperature and total NEWS2 score.

We will use the time to event as the covariate. We will present the trajectories of each vital sign in the 7 days prior to deterioration or discharge graphically. All time dependent distributions will be modelled using the GAMLSS package(37). We will repeat the analysis for the secondary outcomes.

Sensitivity analysis will explore the range of methods to estimate FiO2 from oxygen flow rate, such as the “conventional prediction rule” (1l/min = +4% FiO2)(29), the Bateman equation(30) and the Wettstein calculations(28) as identified in the preceding systematic review.

Secondary objective 2:

To externally validate a series of early warning scores with oxygen graded beyond binary scoring and compare their ability to predict deterioration to NEWS2.

Study Population

We will include all adults (age ≥ 16) admitted to hospital. Admissions with no recorded vital signs will be excluded. We will exclude admissions which were < 24 hours in duration, admissions to the emergency department and admissions direct to the ICU from ED.

We will exclude observation sets taken after admission to the ICU.

Each new patient admission will be handled as an individual source of data, meaning there could be multiple admissions analysed for the same patient within the study period.

We will consider a subgroup analysis in patients at risk of type 2 respiratory failure, and only those with an oxygen requirement.

Repeated measures

We will use a random sampling method of one observation set per admission. In a sensitivity analysis, we will include all observations, essentially treating them as independent(17). Other resampling procedures will be evaluated as part of a sensitivity analysis. Multiple imputation will be performed prior to random sampling.

Models

We will use the HAVEN dataset to externally validate models identified in our previous systematic review. We will not evaluate models which handle oxygen as a continuous variable or incorporate time-trends into weighting variables. We will also externally validate NEWS2.

Outcome

Our primary outcome of deterioration will be a composite of unplanned ICU admission and death. Where a patient experienced more than one type of deterioration, the first event will be used. Secondary outcomes will be unplanned ICU admission and death treated separately. We will use a time horizon of 24 hours, consistent with other literature on NEWS/NEWS2(7, 18-20).

Sample size

We used guidance provided by Riley et al(24) to calculate the sample size required for external validation, using the "pmvalsampsize" package in R statistical software. The calculation is based on the following assumptions, taken from an FiO2 adjusted NEWS2 score published on the HAVEN dataset(4).

- The O/E ratio. The O/E will be based on the outcome prevalence (3.5%), with a 95% confidence interval width of 0.2.
- Calibration slope. We will assume good model calibration with intercept = 0 and slope = 1.
- Linear predictor. We will assume the linear predictor is normally distributed
- C-statistic. We will assume a target c-statistic of 0.876

This estimates a final minimum sample size required for validation of 10,601 admissions, including 372 events.

Statistical analysis methods

We will report patient demographics and clinical information using descriptive statistics.

For each model, we will report an assessment of calibration (calibration plots, calibration-in-the-large), discrimination (c-statistic), performance analysis and key model/NEWS2 thresholds (score ≥5, score ≥ 7) and decision curve analysis.

Results will be presented for performance overall, and performance at each individual site.

Analysis will be repeated for the secondary outcomes.

4. Discussion

4.1 Strengths

The protocol for the development of NEWS2-O_2 is based on best current practice. We will use multi-site data across the UK to enhance generalisability. Our model development will be informed by first, evaluating the trajectory of rising FiO2 prior to deterioration compared to other vital signs, and secondly, to externally validate (and thus "benchmarking") relevant examples of oxygen graded models identified in a prior systematic review. This analysis will confirm if there is a need to add FiO2 as a predictor and provide insight into the best way to handle FiO2. We will then use a combination of clinically derived and data driven approaches to handle FiO2.

4.2 Limitations

This study will have limitations. We will only be able to adapt NEWS2 to be consistent with its additive categorical structure. There may be "better" ways to handle the predictors, but the end format will need to mimic NEWS2. This is in order to strike the balance between prediction of deterioration and end-user functionality that allows NEWS2 to act as a universal language across healthcare settings(7).

4.3 Generalisability and Implications

We will use best practice approaches, examples from existing literature, and data driven methodologies to develop NEWS2-O_2. We hope NEWS2-O_2will provide an example to the NEWS Development Group to incorporate oxygen therapy into NEWS3. By externally validating NEWS2-O_2 across several NHS trusts, we hope this will give an assessment of anticipated future performance, compared to current NEWS2.

5. References

1. Morgan RJ, Lloyd-Williams F, Wright MM, Morgan-Warren R, editors. An early warning scoring system for detecting developing critical illness1997.
2. Gerry S, Bonnici T, Birks J, Kirtley S, Virdee PS, Watkinson PJ, et al. Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. BMJ. 2020;369:m1501.
3. Royal College of Physicians London. National Early Warning Score (NEWS): standardising the assessment of acute-illness severity in the NHS. Updated report of a working party London: Royal College of Physicians. 2017.
4. Malycha J, Farajidavar N, Pimentel MAF, Redfern O, Clifton DA, Tarassenko L, et al. The effect of fractional inspired oxygen concentration on early warning score performance: A database analysis. Resuscitation. 2019;139:192-9.
5. Clarke J, Gallifant J, Grant D, Desai N, Glover G. 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(1).
6. Pimentel MAF, Redfern OC, Hatch R, Young JD, Tarassenko L, Watkinson PJ. Trajectories of vital signs in patients with COVID-19. Resuscitation. 2020;156:99-106.
7. Holland M, Kellett J. The United Kingdom's National Early Warning Score: should everyone use it? A narrative review. Intern Emerg Med. 2023:1-11.
8. Brand JW, Morrice DJ, Patteril MV, Mackay JH. National Early Warning Score 2 (NEWS2) to identify inpatient COVID-19 deterioration: The importance of pO(2):FiO(2) ratio. Clin Med (Lond). 2021;21(3):e315-e6.
9. Zhu Y, Chiu YD, Villar SS, Brand JW, Patteril MV, Morrice DJ, et al. Dynamic individual vital sign trajectory early warning score (DyniEWS) versus snapshot national early warning score (NEWS) for predicting postoperative deterioration. Resuscitation. 2020;157:176-84.
10. Machiwenyika JMT, Zhu Y, Villar SS, Mackay JH. Trajectories of vital signs in patients with Covid-19. Resuscitation. 2021;162:449-50.
11. National Institute of Health and Care Excellence. NG50. Acutely ill adults in hospital: recognising and responding to deterioration [Available from: https://www.nice.org.uk/guidance/cg50.
12. Gonem S, Lemberger J, Baguneid A, Briggs S, McKeever TM, Shaw D. Real-world implementation of the National Early Warning Score-2 in an acute respiratory unit. BMJ Open Respiratory Research. 2024;11(1):e002095.
13. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD statement. Journal of Clinical Epidemiology. 2015;68(2):112-21.
14. Royal Berkshire NHS Foundation Trust. About Us 2024 [Available from: https://www.royalberkshire.nhs.uk/about-us.
15. Lanchashire Teaching Hospitals NHS Foundation Trust. About Us [Available from: https://www.lancsteachinghospitals.nhs.uk/about.
16. South Warwickshire University NHS Foundation Trust. About us 2024 [Available from: https://www.swft.nhs.uk/about-us.
17. Jarvis SW, Kovacs C, Briggs J, Meredith P, Schmidt PE, Featherstone PI, et al. Are observation selection methods important when comparing early warning score performance? Resuscitation. 2015;90:1-6.
18. Holland M, Kellett J. A systematic review of the discrimination and absolute mortality predicted by the National Early Warning Scores according to different cut-off values and prediction windows. European Journal of Internal Medicine. 2022;98:15-26.
19. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. 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(4):465-70.
20. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI. Review and performance evaluation of aggregate weighted ‘track and trigger’ systems. Resuscitation. 2008;77(2):170-9.
21. Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. BMJ. 2020;368:m441.
22. Riley RD, Snell KI, Ensor J, Burke DL, Harrell FE, Jr., Moons KG, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. Stat Med. 2019;38(7):1276-96.
23. Pimentel MAF, Redfern OC, Malycha J, Meredith P, Prytherch D, Briggs J, et al. Detecting Deteriorating Patients in the Hospital: Development and Validation of a Novel Scoring System. Am J Respir Crit Care Med. 2021;204(1):44-52.
24. Riley RD, Debray TPA, Collins GS, Archer L, Ensor J, van Smeden M, et al. Minimum sample size for external validation of a clinical prediction model with a binary outcome. Statistics in Medicine. 2021;40(19):4230-51.
25. Janssen KJ, Donders AR, Harrell FE, Jr., Vergouwe Y, Chen Q, Grobbee DE, et al. Missing covariate data in medical research: to impute is better than to ignore. J Clin Epidemiol. 2010;63(7):721-7.
26. Ambler G, Omar RZ, Royston P. A comparison of imputation techniques for handling missing predictor values in a risk model with a binary outcome. Statistical Methods in Medical Research. 2007;16(3):277-98.
27. Vergouwe Y, Royston P, Moons KG, Altman DG. Development and validation of a prediction model with missing predictor data: a practical approach. J Clin Epidemiol. 2010;63(2):205-14.
28. Wettstein RB, Shelledy DC, Peters JI. Delivered oxygen concentrations using low-flow and high-flow nasal cannulas. Respir Care. 2005;50(5):604-9.

29. Fuentes S, Chowdhury YS. Fraction of Inspired Oxygen. StatPearls. Treasure Island (FL): StatPearls Publishing Copyright © 2024, StatPearls Publishing LLC.; 2024.

30. Bateman NT, Leach RM. ABC of oxygen. Acute oxygen therapy. Bmj. 1998;317(7161):798-801.
31. Takada T, Nijman S, Denaxas S, Snell KIE, Uijl A, Nguyen T-L, et al. Internal-external cross-validation helped to evaluate the generalizability of prediction models in large clustered datasets. Journal of Clinical Epidemiology. 2021;137:83-91.
32. Steyerberg EW, Harrell FE. Prediction models need appropriate internal, internal–external, and external validation. Journal of Clinical Epidemiology. 2016;69:245-7.
33. Collins GS, Dhiman P, Ma J, Schlussel MM, Archer L, Van Calster B, et al. Evaluation of clinical prediction models (part 1): from development to external validation. Bmj. 2024;384:e074819.
34. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162(1):W1-73.
35. Fronczek J, Polok K, de Lange DW, Jung C, Beil M, Rhodes A, et al. Relationship between the Clinical Frailty Scale and short-term mortality in patients ≥ 80 years old acutely admitted to the ICU: a prospective cohort study. Critical Care. 2021;25(1):231.
36. Harrell Jr FE. Statistical Thinking [Internet]2020. [2nd August 2024]. Available from: https://www.fharrell.com/post/addvalue/.
37. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. Journal of the Royal Statistical Society: Series C (Applied Statistics). 2005;54(3):507-54.