

Original Article

Logistic early warning scores to predict death, cardiac arrest or unplanned intensive care unit re-admission after cardiac surgery

Y.-D. Chiu,^{1,2} S. S. Villar,² J. W. Brand,³ M. V. Patteril,⁴ D. J. Morrice,⁵ J. Clayton⁶ and J. H. Mackay⁷

1 Statistician, 6 Senior Analyst, Clinical Governance, 7 Consultant, Department of Anaesthesia and Intensive Care, Royal Papworth Hospital, Cambridge, UK

2 Statistician, MRC Biostatistics Unit, School of Clinical Medicine, University of Cambridge, UK

3 Consultant, Department of Anaesthesia and Critical Care, James Cook University Hospital, Middlesbrough, UK

4 Consultant, Department of Anaesthesia and Critical Care, University Hospitals Coventry and Warwickshire, Coventry, UK

5 Consultant, Department of Anaesthesia and Critical Care, New Cross Hospital, Wolverhampton, UK

Summary

NHS England recently mandated that the National Early Warning Score of vital signs be used in all acute hospital trusts in the UK despite limited validation in the postoperative setting. We undertook a multicentre UK study of 13,631 patients discharged from intensive care after risk-stratified cardiac surgery in four centres, all of which used VitalPAC™ to electronically collect postoperative National Early Warning Score vital signs. We analysed 540,127 sets of vital signs to generate a logistic score, the discrimination of which we compared with the national additive score for the composite outcome of: in-hospital death; cardiac arrest; or unplanned intensive care admission. There were 578 patients (4.2%) with an outcome that followed 4300 sets of observations (0.8%) in the preceding 24 h: 499 out of 578 (86%) patients had unplanned re-admissions to intensive care. Discrimination by the logistic score was significantly better than the additive score. Respective areas (95%CI) under the receiver-operating characteristic curve with 24-h and 6-h vital signs were: 0.779 (0.771–0.786) vs. 0.754 (0.746–0.761), $p < 0.001$; and 0.841 (0.829–0.853) vs. 0.813 (0.800–0.825), $p < 0.001$, respectively. Our proposed logistic Early Warning Score was better than the current National Early Warning Score at discriminating patients who had an event after cardiac surgery from those who did not.

Correspondence to: J. H. Mackay

Email: jonmackay@doctors.org.uk

Accepted: 4 June 2019

Keywords: cardiac surgery; early warning scores; ICU re-admission; logistic regression; postoperative deterioration

Introduction

Physiological deterioration usually precedes serious patient events such as death, cardiac arrest and intensive care unit (ICU) admission. Additive early warning scores (EWS) of physiological variables are an attempt to predict and prevent these events [1, 2]. In April 2018, NHS England mandated that an updated

National Early Warning Score (NEWS) should be used by all acute hospital and ambulance trusts by March 2019; failure to comply is penalised by fines and loss of a Commissioning and Quality Innovation incentive payment [3, 4].

The NEWS has been extensively validated in the acute medical and pre-hospital settings, but the

postoperative surgical population has been subject to much less scrutiny [5–7]. Two key features of the cardiac surgical population lend themselves to address this knowledge gap. Firstly, the incidences of postoperative events are higher than other surgical specialties. Secondly, surgical outcomes are tightly scrutinised, with all UK centres mandated to return key information on all patients and their outcomes.

The simple additive NEWS was conceived in an era of 'pen and paper' observation charts and has several limitations [2]. The discrimination of NEWS is limited because: it weights five physiological variables identically; the values of which are combined in 4–6 relatively wide physiological 'dividing bins'; and the values of neurological status and oxygen therapy are dichotomised with a binary response only (Table 1). In addition, the NEWS is an isolated physiological snapshot – scores do not account for whether the patient is improving or deteriorating – or the rate of that change over time.

The dramatic recent shift towards electronic data-capture in UK hospitals makes calculation of logistic EWS at the bed-side readily achievable. In future, it will also be feasible for individual patient trajectories to be factored into the model, by giving physiological derangement additional weight for the deteriorating patient and reduced weight for the improving patient.

Our primary objective was to use simple logistic regression to model the association of the NEWS physiological variables with a serious patient event in the subsequent 24 h. Secondary objectives included comparing the discriminatory power of each model for events in the next 6 h or 12 h. Finally, we used more complex statistical techniques to explore the impact of utilising individual patient-identity information to take into account both improving or deteriorating physiology.

Methods

The Health Research Authority approved this study and determined ethics approval was unnecessary. We studied adults undergoing risk-stratified major cardiac surgery from 1 April 2014 to 31 March 2017 in four UK adult cardiac surgical centres: James Cook University Hospital, Middlesbrough; New Cross Hospital, Wolverhampton; Royal Papworth Hospital, Cambridge; and University Hospitals Coventry and Warwickshire, Coventry. All centres use VitalPAC™ (CareFlows Vitals, System C Healthcare, Maidstone, Kent, UK) to electronically capture patients' vital signs on the postoperative surgical wards. We recorded the date and time of observations and the patients' respiratory rate, oxygen saturations, the device and/or flow used to deliver supplemental oxygen, systolic blood pressure, heart rate, conscious level and temperature. For each patient we recorded the dates of surgery and hospital discharge and the date and time of in-hospital death, cardiac arrest and re-admission to cardiac critical care. We did not analyse patients who died in the operating theatre or in the ICU before discharge back to the general postoperative ward.

We used hospital databases to identify serious patient events: in-hospital death; cardiac arrest; and unanticipated ICU re-admission. We applied additive and logistic models to predict these outcomes. We analysed the first of multiple outcomes that happened within 24 h of an observation. We increased the number of categories for oxygen therapy from two used by NEWS to four: category 0, room air; category 1, $F_{I}O_2$ 0.25–0.34, Venturi mask or nasal cannulae with oxygen flow $< 5 \text{ l.min}^{-1}$; category 2, $F_{I}O_2$ 0.35–0.44, standard oxygen facemask or nasal cannulae with oxygen flow $\geq 5 \text{ l.min}^{-1}$; and category 3, $F_{I}O_2 \geq 0.45$ or reservoir oxygen mask. We similarly increased categories of conscious level from two to four: category 0, alert; category 1, responds to voice or confused; category 2, responds to pain; or drowsy and category 3, unresponsive.

Table 1 The original Royal College of Physicians' NEWS scoring system (2012). Adapted from [1].

Variable	Score						
	3	2	1	0	1	2	3
Respiratory rate; min^{-1}	≤ 8		9–11	12–20		21–24	≥ 25
Oxygen saturation; %	≤ 91	92–93	94–95	≥ 96			
Supplemental oxygen		Yes		No			
Systolic blood pressure; mmHg	≤ 90	91–100	101–110	111–219			≥ 220
Heart rate; min^{-1}	≤ 40		41–50	51–90	91–110	111–130	≥ 131
Alert				Yes			No
Temperature; $^{\circ}\text{C}$	≤ 35.0		35.1–36.0	36.1–38.0	38.1–39.0	≥ 39.1	

The logistic regression model fitted to the data was of the following form:

$$\text{Predicted probability of an event} = \frac{e(\beta_0 + \sum_{i=1}^p \beta_i X_i)}{1 + e(\beta_0 + \sum_{i=1}^p \beta_i X_i)},$$

where β_0 is the constant of the logistic regression, β_i is the coefficient corresponding to the X_i predictor in the logistic regression and p is the number of predictor variables included in the model.

We derived a new logistic early warning score with the seven variables used by the simple additive NEWS. We analysed the distribution of variables, categorised by whether they did or did not precede an outcome.

We used formulae to characterise non-linear associations of four variables with outcomes, with separate formulae for values more than the median and less than the median heart rate, respiratory rate, temperature and systolic blood pressure. We evaluated the risk of individual physiobiological variables based on the model estimated coefficients and the predicted probability formula. We controlled other continuous variables at their median value and categorical variables at the most frequent category. Although a physiobiological variable has a value on its median, the corresponding model estimated coefficients about increment and decrement do not contribute towards calculating predicted probability.

We used receiver-operating characteristic (ROC) curves to evaluate model discrimination, reported as the area under the curve and 95%CI. We also assessed the effect of the suggested thresholds for patient review (NEWS ≥ 5 and NEWS ≥ 7) by reporting sensitivity, specificity and predicted rate of events for each model. For the logistic model we considered two possible thresholds: an optimal one that gives equal weight to specificity and sensitivity; and a threshold that matches the specificity level of NEWS (with a threshold of 5 and 7). We derived models from two-thirds of the dataset and then validated the fitted model with the remaining third. We used four types of validation to evaluate the predictive performance of the fitted model [9–11]. We used R statistical software version 3.5.1, with the R package 'pROC' and others related to particular methods [8].

Results

We analysed 540,127/580,961 (93%) observations on 13,631 patients (summary data Table 2 and distribution histograms on left-side of Fig. 1), 4300 (0.8%) of which preceded an outcome by less than 24 h in 568 (4.2%) patients: 87 (0.02%) observations preceded the in-hospital

deaths of 25 patients (0.2%); 288 (0.05%) observations preceded cardiac arrest in 54 (0.4%) patients; and 3925 (0.73%) observations preceded unplanned ICU re-admission in 499 (3.7%) patients. Ten patients had multiple events. We did not analyse 7% of observations due to missing values, software errors, rare outliers and unused oxygen delivery values and alert system. Detailed exclusion criterion are in the Supporting Information Appendix S1.

Figure 2 and Table 3 detail increased rates of events with preceding tachypnoea, hypoxaemia, hypotension, tachycardia and hypothermia. The logistic model indicates that scores assigned by the NEWS should be increased for tachypnoea, hypotension, tachycardia and hypothermia, and be decreased for and hypoxaemia, hypertension and hyperthermia.

The discrimination of the logistic score was better than the additive NEWS when observations were limited to 6 h or 24 h preceding an event (Tables 4 and 5 and Fig. 3). The discrimination of the logistic model exceeded that of the additive model with three disparate methods of deriving and testing the models (Table 6 and also see Supporting Information, Appendix S1). The distributions of some physiological measures differed between hospitals (see also Supporting Information, Table S1 and Appendix S1). Validated results for this last method suggest that the AUROC could be well above 0.9, and in most cases it was well above 0.8. (See Fig. 4 and additional results in appendix). The incidences of extremely high logistic scores ($> 50\%$) and NEWS scores (≥ 12) were 100 and 87,

Table 2 Summary of 540,127 observations in 13,631 patients after cardiac surgery. Values are mean (SD) or number (proportion).

Variable	
Respiratory rate; min ⁻¹	17.2 (2.4)
Oxygen saturation; %	96.2 (2.0)
Supplemental oxygen category	
Room air	388,732 (72.0%)
Low F _I O ₂ – (%)	130,793 (24.2%)
Medium F _I O ₂ – (%)	20,211 (3.7%)
High F _I O ₂ – (%)s	391 (0.1%)
Systolic blood pressure; mmHg	121.2 (18.6)
Heart rate; min ⁻¹	80.4 (16.1)
Category of consciousness	
Alert	538,716 (99.7%)
Responds to voice or confused	1016 (0.2%)
Responds to pain or drowsy	358 (0.1%)
Unresponsive	37 (0.0%)
Temperature; °C	36.6 (0.5)

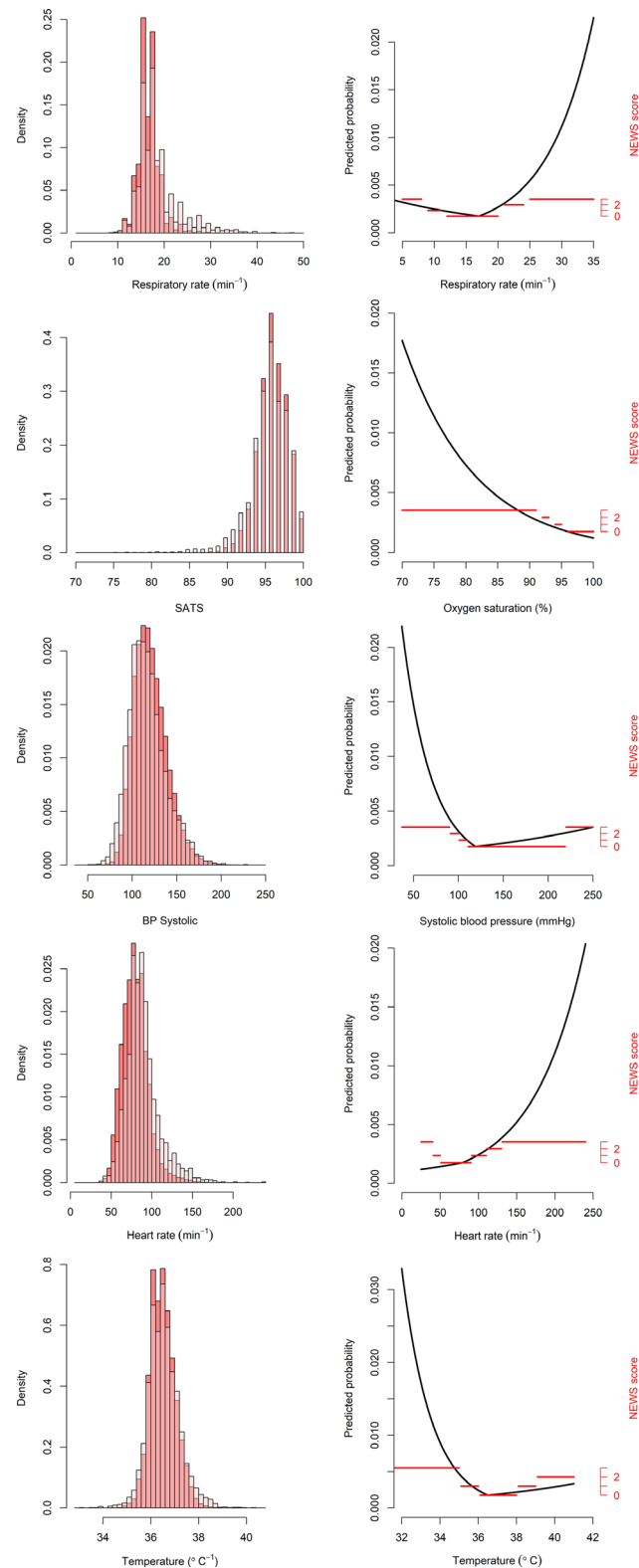


Figure 1 Left–distribution of the five physiological variables in all measurements (dark) and in the measurements with serious adverse events (light)/Right: black curves represent predicted probability of the physiological variable given the other predictors being controlled for logistic EWS. Horizontal red lines represent individual parameter dividing bins used by NEWS (right axis). Note that the scales of figures on the both sides are different for show patterns of interest.

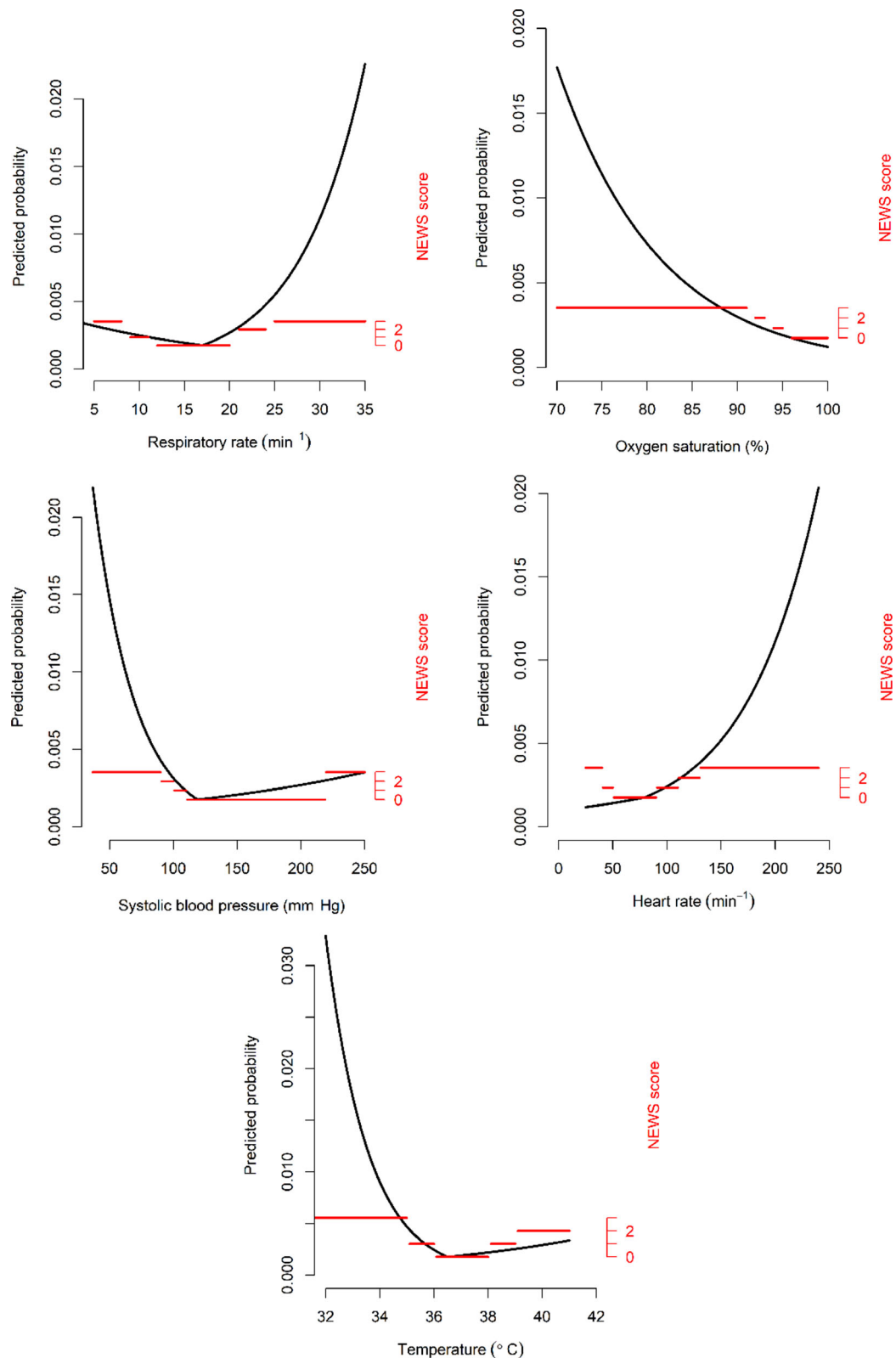


Figure 2 Black curves represent predicted probability of the physiological variable given the other predictors being controlled for logistic Early Warning Score (EWS). Horizontal red lines represent individual parameter dividing bins used by National EWS (NEWS, right axis).

Table 3 The association of logistic Early Warning Score variables with the composite outcome of in-hospital death, cardiac arrest or unplanned intensive care unit re-admission within 24 h of observation.

Variable	β	OR (95%CI)	p value
Intercept	2.259		
Respiration rate: median 17 min ⁻¹			
Increment(min ⁻¹) > 17	0.143	1.15 (1.14–1.16)	< 0.001
Decrement(min ⁻¹) < 17	0.050	1.05 (1.04–1.07)	< 0.001
Oxygen saturation (%)	−0.090	0.91 (0.90–0.93)	< 0.001
Supplemental oxygen category			
0 air		Referent	
1 low	1.30	3.68 (3.43–3.96)	< 0.001
2 medium	2.13	8.39 (7.65–9.20)	< 0.001
3 high	2.92	18.51 (13.46–25.44)	< 0.001
Systolic blood pressure: median 119 mmHg			
Increment(mmHg) > 119	0.005	1.01 (1.00–1.01)	< 0.001
Decrement(mmHg) < 119	0.031	1.03 (1.03–1.04)	< 0.001
Heart rate: median 79 min ⁻¹			
Increment(min ⁻¹) > 79	0.015	1.02 (1.01–1.02)	< 0.001
Decrement(min ⁻¹) < 79	−0.007	0.99 (0.99–1.00)	0.010
Level of consciousness			
0 Alert		Referent	
1 Responds to voice or confused	1.84	6.28 (5.03–7.85)	< 0.001
2 Responds to pain or drowsy	1.90	6.65 (4.64–9.53)	< 0.001
3 Unresponsive	3.27	26.29 (12.08–57.21)	< 0.001
Temperature: median 36.5 °C			
Increment(°C) > 36.5	0.145	1.16 (1.06–1.25)	< 0.001
Decrement(°C) < 36.5	0.659	1.93 (1.73–2.16)	< 0.001

respectively, out of 540,127 sets of observations. Calibration was excellent for logistic EWS scores of up to 50%, but less impressive in the extremely rare event (~1 in 5000 incidence) of scores > 50%. (See also Supporting Information, Appendix S5).

Discussion

This is first study to test the National Early Warning Score after cardiac surgery. We found that the logistic score was

Table 4 The discrimination of NEWS vs log EWS for a subsequent event when observations are limited to the preceding 6 h, 12 h or 24 h.

Scoring system			
	NEWS	log EWS	p value
Observation period; h			
6	0.813 (0.800–0.825)	0.841 (0.829–0.853)	<0.001
12	0.789 (0.779–0.799)	0.815 (0.806–0.824)	<0.001
24	0.754 (0.746–0.761)	0.779 (0.771–0.786)	<0.001

NEWS, National Early Warning Score; EWS, early warning scores.

significantly better at predicting deterioration than the current additive score. The logistic score performed even better if only the last 6 h of observations are used, rather than the preceding 24 h. For a given level of specificity, the logistic model offers increments in sensitivity at threshold values: the 3.7% increment at NEWS 7 represents a relevant increase in true positive cases from 17.5% to 21.2% [12]. Similarly, at a threshold of NEWS 3, sensitivity is increased to 70%; however, this would quadruple the number of clinical reviews required.

Discrimination by NEWS, as measured by the area under the ROC curve (0.75), was less in our postoperative population than typically reported in acute medical populations (> 0.85). A recent large, single-centre North American study reported a similar area (0.76) for a general postoperative population [7]. A continuous logistic risk score has previously been demonstrated to offer better discriminatory performance than an additive score in general ward admissions [13].

Unanticipated re-admission to intensive care constituted most outcomes (86%), whilst death and cardiac

Table 5 The rate of events predicted by different score thresholds for NEWS (e.g. 4) and log EWS (e.g. 0.003), with accompanying sensitivity and specificity, when observations are limited to the preceding 6 h, 12 h or 24 h. See online Supporting Information Appendix for additional information.

Observation period	Event rate		Sensitivity		Specificity	
	NEWS	logEWS	NEWS	logEWS	NEWS	logEWS
6 h: score threshold						
4 (0.003) ^a	18%	20%	67%	74%	83%	80%
5 (0.010)	9%	9%	48%	52%	92%	92%
7 (0.017)	2%	2%	26%	34%	98%	98%
12 h: score threshold						
4 (0.005) ^a	18%	21%	61%	69%	83%	80%
5 (0.010)	9%	9%	48%	52%	92%	92%
7 (0.029)	2%	2%	24%	28%	98%	98%
24 h: score threshold						
3 (0.007) ^a	33%	29%	71%	71%	67%	72%
5 (0.018)	9%	9%	40%	43%	92%	92%
7 (0.043)	2%	2%	18%	21%	98%	98%

NEWS, National Early Warning Score; EWS, early warning scores.

^aOptimal Youden index.

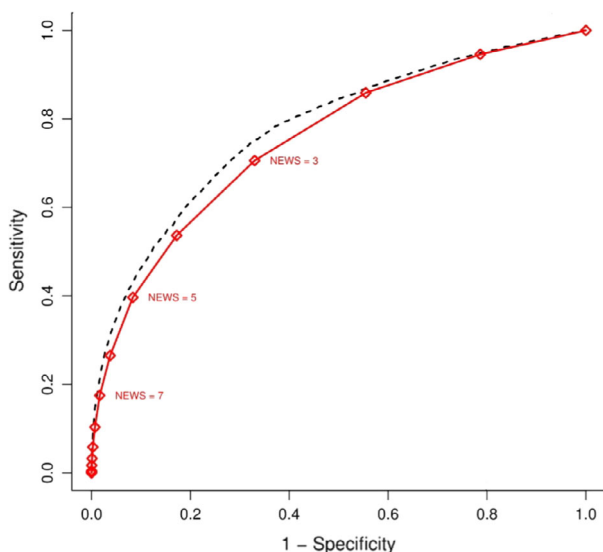


Figure 3 ROC curves for logistic early warning score (EWS) and national early warning score (NEWS) for comparison when observations are limited to the preceding 24 h. The red solid line represents NEWS; the black dashed line represents logistic EWS.

arrest accounted for 4% and 9% of outcomes, respectively. Death has been the commonest outcome in most previous studies of NEWS [3, 14, 15]. The National Early Warning Score has consistently discriminated patients who die from those who survive better than discriminating patients who are admitted to intensive care [5, 6]. The incidences of cardiac arrest and death were low in all cardiac surgical

centres. We share Schmidt's belief that hospital-wide physiological surveillance may have reduced these outcomes [16]. The majority of cardiac arrests and deaths after cardiac surgery occur in ICU, before discharge to the postoperative wards [17].

The results were extensively validated using both internal and external validation procedures. All validated results indicated the same hierarchy of discriminating performances, where NEWS was ranked last and the logistic EWS was ranked highest. We would, therefore, recommend logistic EWS for predicting serious adverse events in hospitals with similar populations to this paper.

A simple additive model like NEWS – with low discriminatory power – is unlikely to achieve a good predictive performance in postoperative surgical populations with very low incidences of adverse events. We have preliminarily tested more complex methods, including naïve Bayes classifier, classification trees, random forest, gradient boosting and neural network with a single hidden layer (results not presented in the main paper) [18]. These models did not offer significant advantages over the logistic model. The only method that offered significant and impressive predictive gains was a multilevel logistic regression model in which the patient-identity information and temporal evolution are taken into account to make predictions.

There are clear parallels with current risk-stratification modelling used to predict death after cardiac surgery. Initially, simplicity and the ability to calculate bed-side

Table 6 The discrimination of NEWS vs. log EWS for a subsequent event when observations are limited to the preceding 24 h, with the models derived from: a random sample of two-thirds of the dataset (an average of 1000 resamples); 478,867 observations reported before 2017; the first 90% of observations recorded for each patient.

	Scoring system		p value
	NEWS	log EWS	
Derivation dataset			
Random two-thirds resampled	0.754 (0.745–0.763)	0.778 (0.769–0.787)	< 0.001
2014–2016 inclusive	0.717 (0.694–0.740)	0.737 (0.714–0.760)	< 0.001
First 90% each patient's data	0.833 (0.808–0.858)	0.861 (0.837–0.885)	< 0.001

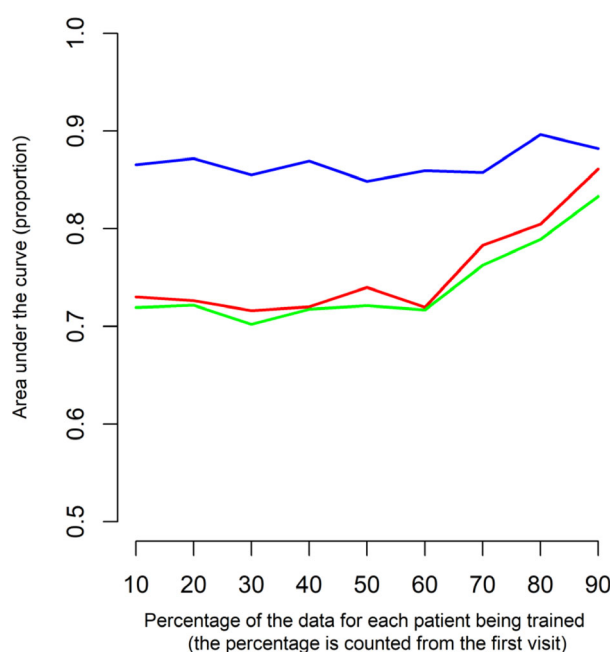


Figure 4 Comparison of area under the curve among methods. Using the first 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90% of each patient's data on x-axis for model fitting and validation every next 10% of data. The colours of the lines represent different methods: blue (for multilevel logistic regression (MLR)), red (for logistic early warning score), green (for national early warning score); the MLR model utilises patient-identity information and temporal evolution of scoring to make predictions.

scores were desirable when the additive EuroSCORE was originally conceived, however, electronic data-capture and computerised scoring led to this being superseded by the more powerful logistic EuroSCORE which better predicted risk in the high-risk groups of patients [19]. Complex sophisticated logistic EWS models will similarly only replace the current additive NEWS after demonstration of clinically meaningful performance improvement.

By 2022 (when NEWS2 is projected for its next review), it is likely that most NHS hospitals will have

electronic observation charts in place. This provides the opportunity to replace additive scores with more powerful scoring systems that would support tailoring interventions to improve patient outcomes. The clinical significance of any absolute additive NEWS score is currently very dependent on the patient population and consequently difficult to predict at the bed-side. There is recognition that NEWS is too sensitive in patients with chronic chest medical disease and a not sensitive enough in surgical patients [3, 6].

Logistic scores could be recalibrated to reduce sensitivity in the former group and increase sensitivity in postoperative patients. Substituting Glasgow Coma Scale for the less discriminatory 'AVPU' (see Table 2) in neurosurgical patients; and adding urine output as an eighth parameter in cardiothoracic surgical patients would further increase sensitivity. Logistic scores, which predict the probability of an adverse event, should therefore facilitate earlier recognition and escalation of the deteriorating patient. Logistic EWS would also enable a future paediatric EWS to be calibrated for patient age and/or weight.

Using our logistic EWS data we have also produced an App to use at the bed-side https://yidachiu.shinyapps.io/vitalpac_log_ews_app/. Seven parameters (conscious level, F_iO_2 , temperature, systolic blood pressure, heart rate, respiratory rate and oxygen saturations) are entered in turn to generate both the log EWS and NEWS scores [20]. Logistic EWS forecasts the 'positive predictive value' of a subsequent adverse event in cardiac surgical patients – with any given score representing the percentage chance of such an event. We believe this scoring system could be recalibrated for use in other surgical and medical populations.

In summary, a logistic version of the National Early Warning Score, rather than the current additive model, better discriminates patients after cardiac surgery who die, have a cardiac arrest or unplanned readmission to intensive care. Logistic scores also provide a useful quantified tool of predicted risks for clinicians, which NEWS cannot.

Acknowledgements

We are grateful to many people in each study centre for their significant contributions – including those responsible for maintaining accurate surgical, ICU re-admission and cardiac arrest databases. We thank the analysts in each centre for successfully extracting their VitalPAC data: Coventry, G. Georgiades and S. Kumar; Middlesbrough, M. Ahmed, A. Goodwin, I. Pattinson, T. Smailes and C. Williams; Papworth, J. Bracken, E. Gorman, V. Hughes, J. Machiwenyika, J. Quigley and M. Sale; Wolverhampton, R. Giri, Y. Li, S. Murphy, S. Rowles and N. Wise. No external funding or competing interests declared.

References

1. Royal College of Physicians. National Early Warning Score (NEWS): Standardising the assessment of acute-illness severity in the NHS. Report of a working party. London: RCP, 2012. <https://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news-2> Archived pdf of original 2012 NEWS report available in Downloads section (accessed 01/06/2019).
2. Prytherch D, Smith GB, Schmidt PE, Featherstone PI. ViEWS—towards a national Early Warning Score for detecting adult in-patient deterioration. *Resuscitation* 2010; **81**: 932–7.
3. 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. <https://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news-2> (accessed 01/06/2019).
4. NHS Improvement. Resources to support the safe adoption of the revised National Early Warning Score (NEWS2). 2018. https://improvement.nhs.uk/documents/2508/Patient_Safety_Alert_-_adoption_of_NEWS2.pdf (accessed 07/05/2019).
5. Smith GB, Prytherch D, 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 (ICU) admission, and death. *Resuscitation* 2013; **84**: 465–70.
6. Kovacs C, Jarvis SW, Prytherch DR, et al. Comparison of the National Early Warning Score in non-elective medical and surgical patients. *British Journal of Surgery* 2016; **103**: 1385–93.
7. Bartkowiak B, Snyder AM, Benjamin A, et al. Validating the electronic cardiac arrest risk triage (eCART) score for risk stratification of surgical patients in the postoperative setting: retrospective cohort study. *Annals of Surgery* 2019; **269**: 1059–63.
8. 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.
9. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology* 2001; **54**: 774–81.
10. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine* 1996; **15**: 361–87.
11. 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**: 112–21.
12. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950; **3**: 32–5.
13. Ghosh E, Eshelman L, Yang L, Eric Carlson E, Lord B. Early Deterioration Indicator: data-driven approach to detecting deterioration in general ward. *Resuscitation* 2018; **122**: 99–105.
14. Watkinson PJ, Pimentel MAF, Clifton DA, Tarassenko L. Manual centile-based early warning scores derived from statistical distributions of observational vital-sign data. *Resuscitation* 2018; **129**: 55–60.
15. 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 multicentre database study. *Resuscitation* 2019; **134**: 147–56.
16. Schmidt PE, Meredith P, Prytherch DR, et al. Impact of introducing an electronic physiological surveillance system on hospital mortality. *BMJ Quality and Safety* 2014; **24**: 003073.
17. Mackay JH, Powell SJ, Osgathorp J, Rozario CJ. Six-year prospective audit of chest reopening after cardiac arrest. *European Journal of Cardiothoracic Surgery* 2002; **22**: 421–5.
18. James G, Witten D, Hastie T, Tibshirani R. *An introduction to statistical learning: with applications in R*. New York, NY: Springer, 2013.
19. Roques F, Michel P, Goldstone AR, Nashef SA. The logistic EuroSCORE. *European Heart Journal*. 2003; **24**: 882–3.
20. Chiu YD, Villar SS, Mackay JH. Logistic early warning score app for cardiac surgical patients. 2019. https://yidachiu.shinyapps.io/vitalpac_log_ews_app/ (accessed 01/06/2019).

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Model coefficients.

Appendix S2. Data summary.

Appendix S3. Evaluation of predictive performance and validation.

Appendix S4. Multilevel logistic regression.

Appendix S5. Logistic EWS model calibration.