CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Bournemouth University Research Online

Is the UK National Early Warning Score suitable for use in surgical patients?

Mrs. Caroline Kovacs, BSc, Research Associate, Centre for Healthcare Modelling and Informatics, University of Portsmouth, Portsmouth, UK

Dr. Stuart W Jarvis, PhD, Visiting Research Fellow, Centre for Healthcare Modelling and Informatics, University of Portsmouth, Portsmouth, UK, & Research Fellow, Department of Health Sciences, University of York

Professor David R Prytherch, PhD, MIPEM, CSci, Clinical Scientist, Research and Innovation Department, Portsmouth Hospitals NHS Trust & Visiting Professor, Centre for Healthcare Modelling and Informatics, University of Portsmouth, Portsmouth, UK

Dr. Paul Meredith, PhD, Data Analyst, Research and Innovation Department, Portsmouth Hospitals NHS Trust, Portsmouth, UK

Dr. Paul E Schmidt, MRCP, B.Med.Sc, MBA, Consultant In Acute Medicine, Portsmouth Hospitals NHS Trust

Dr. James S Briggs, BA, DPhil, CEng, CITP, MBCS, FHEA, Principal Lecturer & Director, Centre for Healthcare Modelling and Informatics, University of Portsmouth, Portsmouth, UK

Professor Gary B Smith, FRCA, FRCP, Visiting Professor, Faculty of Health & Social Sciences, University of Bournemouth, Bournemouth, UK

Correspondence and requests for reprints to:
Mrs Caroline Kovacs, BSc,
Centre for Healthcare Modelling and Informatics (CHMI),
School of Computing,
University of Portsmouth,
Buckingham Building, Lion Terrace
Portsmouth PO1 3HE
United Kingdom

Email: caroline.kovacs@port.ac.uk

Word count = 2 677 (main article) Number of references = 29 Figures = 2 Tables = 3

Funding: Nil
Submitted as: Original article
Previous communication to a society or meeting: None

ABSTRACT

Background:

The National Early Warning Score (NEWS) is used to identify deteriorating patients in hospital. NEWS is a better discriminator of outcomes than other early warning scores (EWS) in acute medical admissions, but it has not been evaluated in a surgical population. The study aims were to (i) evaluate the ability of NEWS to discriminate cardiac arrest, death, and unanticipated ICU admission in patients admitted to surgical specialties, and (ii) compare the performance of NEWS in admissions to medical and surgical specialties.

Method:

Hospital-wide data over 31 months, from adult in-patients, who stayed at least one night or died on the day of admission were analysed. The data were categorized as elective or non-elective surgical or medical admissions. The ability of NEWS to discriminate the outcomes above in these different groups was assessed using the area under the ROC curve.

Results:

There were too few outcomes to permit meaningful comparison of elective admissions so the analysis was constrained to comparison of non-elective admissions. NEWS performs equally well or better for surgical as for medical patients. For death within 24 hours the AUROC for surgical admissions was 0.914 (0.907 - 0.922) compared to 0.902 (0.898 - 0.905) for medical admissions. For the combined outcome of any of death, cardiac arrest or unanticipated ICU admission the AUROC for surgical admissions was 0.874 (0.868 - 0.880) and for medical was 0.874 (0.871 to 0.877).

Conclusion:

NEWS discriminates deterioration in surgical patients at least as well as in medical patients.

Key Words:

Deteriorating patient, early warning scores**,** National Early Warning Score (NEWS), surgical patient safety, medical deteriorating patient, observational study.

INTRODUCTION

Early warning scoring systems[1] are commonly used to identify sick or deteriorating patients, including those admitted to surgical areas of hospitals[2–13]. Several publications suggest that early warning scores (EWS) predict outcomes (death, cardiac arrest, unanticipated intensive care unit (ICU) admission) in surgical and trauma patients[3,7,9,12]. Some researchers claim that EWS use reduces unanticipated ICU admissions[6] and cardiac arrests[13]; others report no impact on mortality[5].

In July 2012, the Royal College of Physicians of London (RCPL) recommended the use of a standardised National EWS (NEWS) for the initial assessment of acute illness and for continuous monitoring of all adult inpatients, excluding women who are pregnant (supplementary material, Table 1)[14]. NEWS uses seven parameters: pulse rate; breathing rate; systolic blood pressure; body temperature; neurological status using the Alert-Verbal-Painful-Unresponsive (AVPU) scale; peripheral oxygen saturation (S_pO_2); and the inspired gas (i.e., air or oxygen) at the time of S_pO_2 measurement.

NEWS is based on the VitalPAC EWS (ViEWS)[15], and is now widely employed in clinical practice in the UK[16,17]. NEWS has been shown to be a better discriminator of outcomes than other EWS systems in acute medical admissions[18], but it has not been evaluated in a surgical population. The aims of this study were to (i) evaluate the ability of NEWS to discriminate cardiac arrest, death, and unanticipated ICU admission in patients admitted to surgical specialties, and (ii) compare the performance of NEWS in admissions to surgical specialties with those to medical specialties.

METHOD

This research is covered by ethical approval from the Isle of Wight, Portsmouth and South East Hampshire Research Ethics Committee (study ref 08/02/1394*).*

Vital signs database and its development

A database of vital sign observations was developed, which were collected in real-time from adults (≥ 16 years) admitted to Portsmouth Hospitals NHS Trust on or after 25/05/2011 and discharged on or before 31/12/2013. Data from patients who were discharged alive from hospital before midnight on the day of admission were excluded. Staff entered patients' vital signs data at the bedside into handheld devices running the VitalPAC software following the hospital protocol[19–21]. The following data were recorded: date/time of observation (automatically by VitalPAC); pulse rate (bpm); systolic and diastolic blood pressure (mmHg); breathing rate (bpm); body temperature (ºC); neurological status using the Alert-Verbal-Painful-Unresponsive (AVPU) scale; peripheral oxygen saturation (S_pO_2) (%); and the inspired gas (i.e., air or oxygen) at the time of S_pO_2 measurement. Vital signs sets for which any vital signs measurements were absent or physiologically impossible (i.e., obviously recorded in error) were excluded. The average time between observation sets was determined by the VitalPAC software, with an average interval of 6-7 hours[21].

Two data sets were extracted from this database.

Where the specialty at admission to the hospital was either General Surgery; Urology; Renal Surgery and Transplantation; Colorectal; Hepatobiliary and Pancreatic; Upper Gastrointestinal; Vascular; or Trauma and Orthopaedics the admission was categorised as surgical. Where the speciality was General Medicine; Gastroenterology; Endocrinology; Clinical Haematology; Rehabilitation; Cardiology; Respiratory (Thoracic) Medicine; Nephrology; Medical Oncology; Neurology; Rheumatology or Geriatric Medicine the admission was medical. Admissions were further categorised by mode of admission, i.e. elective or non-elective.

No theatres data were available, so it was not possible to identify which patients admitted to surgical specialties underwent surgery, or whether an observation set was pre- or post- operative.

Outcomes

For each admission (or episode of care) three outcomes (death, cardiac arrest and unanticipated ICU admission) were extracted from the appropriate hospital databases. Analysis was limited to the first of any of these three outcomes or, if none of the outcomes occurred, to the point of live discharge from hospital. Data were not analysed from patients admitted directly on to critical care areas of the hospital. In addition, episodes of care were excluded where (i) the episode had a first event before the first observation set was recorded or (ii) the episode did not have an observation set within the last 24 hours before the outcome (see supplementary material, Figure 1).

Data analysis

The mean and standard deviation (SD) were calculated for: patient age at admission; NEWS value; pulse rate; respiration rate; temperature; systolic BP; and S_pO_2 for each of four admission groups – non-elective surgical; elective surgical; non-elective medical; elective medical. Welch’s t-test s used to compare the episode mean values for the non-elective surgical, elective surgical, non-elective medical, and elective medical groups.

The outcomes used in the analysis were: death, unanticipated ICU admission and cardiac arrest, within 24 hours of the observation set, and the combined outcome of any of these three (Combined Outcome). Where there were multiple outcomes, for example cardiac arrest, followed by unanticipated ICU admission followed by death, only the first event to occur (cardiac arrest in this example) was used for analysis.

The frequency of adverse outcomes: death, cardiac arrest or unanticipated ICU admission among elective admissions to medical and surgical specialties was too low to perform a valid analysis. Therefore, the ability of NEWS to discriminate outcomes within 24 hours post vital signs observation was only evaluated for the non-elective admissions to surgical and medical specialties, using the area under the receiver-operating characteristic (ROC) curve[22]. The area under the ROC curve (AUROC) was calculated using (a) all observation sets in the data set, and (b) 10 000 sample sets of one observation per episode by first picking at random a time during each episode and selecting the observation set closest to that time[23]. When using all observations, a 95% confidence interval was calculated for the AUROC and assessed the significance of the differences using the methods set out by DeLong et al[24], as implemented in the pROC[25] package in R[26]. When using one observation per episode, the AUROCs were calculated for each of the 10 000 sample sets and the 2.5 and 97.5 centiles were taken as the 95% confidence interval. The significance of differences was tested by computing the difference between AUROCs reported for non-elective surgical and medical admissions in each of the 10 000 sample sets. The 95% confidence interval of the difference was estimated from the 2.5 and 97.5 centiles – if the 95% confidence interval of the difference did not include zero then the AUROCs were significantly different at the $\alpha = 0.05$ level. Finally, p-values for the differences were estimated by determining the largest symmetric confidence interval at which the range of the lower and upper confidence interval did not include zero (i.e., beginning from the 0.01 and 100.00 centiles, moving in to the 0.02 and 99.99 centiles and so on until the interval did not include zero). The p-value was then calculated as 1 minus the confidence interval divided by 100. ROC curves were plotted for death, cardiac arrest, unanticipated ICU admission and the Combined Outcome (i.e. any of these) within 24 hours of a vital signs observation set.

For the non-elective surgical and medical admissions, the performance of NEWS was also analysed using the EWS efficiency curve[15], which plots the percentage of the total number of observations at or above a threshold EWS value for escalation (a measure of workload assigned to responders) against sensitivity for the Combined Outcome.

RESULTS

Comparison of admissions to surgical specialties with those to medical specialties

The relative composition of admissions across all medical and surgical specialties is described fully in Table 1. In summary, there were 87 399 admissions, comprising 2 017 455 observation sets (see supplementary material, Figure 1). Of these, 35 174 were admissions to surgical specialities (792 889 observation sets) and 52 225 admissions to medical specialties (1 174 574 observation sets).

A substantially higher proportion of medical admissions were non-elective than for surgical admissions (93.3% vs 58.6%; $p<0.001$). On average, these non-elective patients were younger for surgical specialties (mean age = 56.6 vs. 67.5) and were more likely to be female (52.6% vs 49.2%; $p < 0.001$). Overall, the rate of combined outcomes for admissions to medical specialties was higher than for admissions to surgical specialties Table 2). However, the number of such outcomes was very low for both medical and surgical elective admissions (1.5%, 0.7%) and consequently did not permit a full analysis (see Discussion).

Comparison of vital signs in surgical specialties with those in medical specialties

The distribution of outcomes between the four patient groups, by number of vital signs observation sets and patient care episodes is shown in Table 2. Non-elective surgical admissions and non-elective medical admissions had similar mean numbers of observation sets per admission (25 vs. 24). The mean (SD) NEWS value for the whole hospital stay was lower for non-elective admissions to surgical specialties than for non-elective admissions to medical specialties (1.5 (1.7) vs. 2.0 (2.1), $p < 0.001$. The percentage of observation sets where patients were receiving supplementary oxygen was 14.8% for the non-elective surgical group and 19.1% for the non-elective medical group.

Figure 1 shows the distribution of NEWS values for all observations for the non-elective medical and non-elective surgical groups, against the Combined Outcome. The proportion of observations of NEWS values ≥ 3 is higher for admissions to medical specialties, however observed risk is higher in patients admitted to surgical specialties. The increased observed risk results from an increased risk of unanticipated ICU admission rather than because of an increased risk of death or cardiac arrest (see supplementary material, 7).

Table 3 shows the AUROCs for NEWS for death, cardiac arrest, unanticipated ICU admission and the Combined Outcome (of any of these) for the non-elective surgical and non-elective medical groups. No matter whether all observations or one observation per episode was used, NEWS always produced AUROC values exceeding 0.80 for each category of patient for all of the outcomes except cardiac arrest. Discrimination of the risk of death within 24 hours was particularly high, with AUROC ≥ 0.9 for both groups.

For the "all observations" analysis, NEWS was a better discriminator of death for non-elective admissions to surgical specialties than for non-elective admissions to medical specialties (AUROC 0.914 (0.907–0.922) vs. 0.902 (0.898–0.905), p=0.003). There was no difference in the ability of NEWS to discriminate cardiac arrest (p= 0.345), unanticipated ICU admission (p=0.555), or the Combined Outcome (p= 0.874) in either of the two groups (Table 3). When one observation set per episode of patient care was selected at random, NEWS performed slightly worse for unanticipated ICU admission and the Combined Outcome for non-elective admissions to surgical specialties than for non-elective admissions, but performed the same in the two groups for death and cardiac arrest (Table 3).

Figure 2 shows the "EWS efficiency curve" for NEWS using the Combined Outcome for non-elective admissions to surgical specialties and non-elective admissions to medical specialties. This provides a measure of the proportion of vital signs observation sets that would trigger at a range of NEWS values and provides a measure of the sensitivity at that point. The RCPL recommends that a NEWS value of ≥5 should trigger an urgent assessment of the patient by a clinician with core competencies to assess acutely ill patients, and a NEWS value of ≥7 should prompt emergency assessment by a clinical team with critical-care competencies. Figure 2 shows that a NEWS value of 5 would trigger such an assessment for 12.3% (NEWS of ≥7: 4.4%) of observations performed on non-elective medical admissions, and this would result in the detection of 70.2% (NEWS of ≥7: 48.7%) of Combined Outcomes. In these data, a NEWS value of 4 would have a similar efficiency (11.0% of observations detecting 70.9% of Combined Outcomes) for non-elective admissions to surgical specialties.

DISCUSSION

This study demonstrates that NEWS performed well in admissions to a group of common surgical specialties, despite having been developed in a population of general medical hospital admissions[18]. Irrespective of the sampling methods used to build a database for analysis, NEWS consistently produced AUROC values in non-elective admissions to surgical specialties that exceeded 0.800 for all outcomes except cardiac arrest. The

AUROCs for non-elective admissions to surgical specialties are similar to those for medical specialties, indicating almost identical performance. The finding that NEWS is a poorer discriminator of cardiac arrest, compared with the other outcomes studied, also replicates the findings of the original NEWS validation study in medicine[18]. This could be because cardiac arrest is less 'predictable' than unanticipated ICU admission and death, sometimes occurring in the absence of antecedent physiological disturbance as an unanticipated, sudden event. Moreover, cardiac arrest is a relatively 'artificial' outcome, being indistinguishable from death for most patients, except for the fact that a cardiac arrest team is called to the former but not the latter, and is also affected by the variable use of Do Not Attempt CardioPulmonary Resuscitation (DNACPR) orders. These inevitably introduce an element of randomness, which can impact upon the discriminative power of any EWS system.

Another important consideration in evaluating the performance of any EWS is the clinical workload that it creates. The efficiency curves for NEWS for non-elective surgical admissions are virtually identical to non-elective medical admissions. Analysis of these curves shows that using a NEWS threshold of 5 generates different workload and detection rates for the two patient groups (medical, workload 12.3%, detection 70.2%; surgical, workload 6.1%, detection 60.6%) and similarly for a NEWS value of 7 (medical, workload 4.4%, detection 48.7%; surgical, workload 1.8%, detection 36.9%).

Some authors have argued that disease-specific or condition-specific EWS scores might be required when EWS systems are used outside the patient population in which they were developed[27–29]. The findings of this study suggest that, for admissions to surgical specialties, a change in trigger level could be all that is required as a NEWS value of 4 for non-elective admissions to surgical specialties has equivalent efficiency to NEWS of 5 for medicine. However, it would be premature to recommend such a change, as further work is required to remove the potential confounders of age, gender and whether they underwent an operation or not.

This study has several strengths. It considers completed non-elective admissions to common surgical and medical specialties at a UK district general hospital over 31 months. For all observation sets, all necessary vital signs variables were collected simultaneously in a standardised manner as part of the clinical process using an electronic device [19–21]. In addition, each vital signs observation set contained all of the necessary factors to calculate a NEWS value. The study also has several weaknesses. For reasons already outlined in a previous publication[15], the recorded date/time of death used in this research is likely to be systematically late and, therefore, the number of observations followed by death within 24 hours is liable to be an underestimate. Such errors are less likely in the database used to identify ICU admission. Other weaknesses are: this is a single centre evaluation, which requires external validation; it was assumed that the hospital treatment of all patients studied was optimal and equitable; data from patients who were admitted directly to critical care areas of the hospital have not been evaluated; and patients on an end-of-life pathway could not be explicitly excluded from analysis. It is standard practice within the study hospital to cease vital signs collection once a patient enters the end-of-life pathway. Consequently, the study attempted to minimise the impact of retaining these patients by the selection technique, through excluding admissions where there was no observation set within the last 24 hours before the outcome, which would exclude a large proportion of such patients. Finally, no account was taken of whether vital signs for surgical patients were measured pre- or post-operatively, and age and gender were not analysed. These analyses will be undertaken in subsequent research.

SUMMARY

This paper describes the application of NEWS to a large vital signs database from surgical admissions to a district general hospital in the UK and compares its performance with that in unselected admissions to general medicine. The results of this research suggest that NEWS has similar performance in patients admitted to surgical and medical wards, but that a change in trigger level may be required for surgical patients. These results further strengthen the argument that NEWS should be adopted as a universal early warning system for detecting deterioration in hospital.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the co-operation of the nursing and medical staff in the study hospital.

CONFLICT OF INTERESTS STATEMENT

VitalPAC is a collaborative development of The Learning Clinic Ltd (TLC) and Portsmouth Hospitals NHS Trust (PHT). At the time of the study, PHT had a royalty agreement with TLC to pay for the use of PHT intellectual property within the VitalPAC product. Professor Prytherch, Dr Schmidt and Dr Meredith are employed by PHT. Professor Smith was an employee of PHT until 31/03/2011. Professors Smith and Prytherch and Dr Schmidt are unpaid research advisors to TLC. Professors Smith and Prytherch have received reimbursement of travel expenses from TLC for attending symposia in the UK. Dr Briggs's research has previously received funding from TLC through a Knowledge Transfer Partnership.

Professor Smith acted as expert advisor to the National Institute for Health and Clinical Excellence during the development of the NICE clinical guideline 50: 'Acutely ill patients in hospital: recognition of and response to acute illness in adults in hospital'. He was also a member of the National Patient Safety Agency committee that wrote the two reports: 'Recognising and responding appropriately to early signs of deterioration in hospitalised patients' and ' Safer care for the acutely ill patient: learning from serious incidents'. He was a member of the Royal College of Physicians of London's National Early Warning Score Development and Implementation Group (NEWSDIG). Professor Prytherch assisted the Royal College of Physicians of London in the analysis of data validating NEWS.

REFERENCES

1 Smith GB, Prytherch DR, Schmidt PE, Featherstone PI. Review and performance evaluation of aggregate weighted 'track and trigger' systems. *Resuscitation*. 2008; **77**: 170–179.

2 Stenhouse C, Coates S, Tivey M, Allsop P, Parker T. Prospective evaluation of a modified Early Warning Score to aid earlier detection of patients developing critical illness on a general surgical ward. *Br J Anaesth* [Internet]. 2000; **84**: 663. Available from: http://bja.oxfordjournals.org/content/84/5/663.2.abstract

3 Garcea G, Ganga R, Neal CP, Ong SL, Dennison AR, Berry DP. Preoperative early warning scores can predict in-hospital mortality and critical care admission following emergency surgery. *J Surg Res* [Internet]. Elsevier Ltd; 2010; **159**: 729–734. Available from: http://dx.doi.org/10.1016/j.jss.2008.08.013

4 Gardner-Thorpe J, Love N, Wrightson J, Walsh S, Keeling N. The value of Modified Early Warning Score (MEWS) in surgical in-patients: A prospective observational study. *Ann R Coll Surg Engl* [Internet]. 2006 Oct [cited 2014 Nov 21]; **88**: 571–575. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1963767&tool=pmcentrez&rendertype=abstract

5 Patel MS, Jones M a, Jiggins M, Williams SC. Does the use of a 'track and trigger' warning system reduce mortality in trauma patients? *Injury* [Internet]. 2011; **42**: 1455–1459. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21696724

6 Peris a, Zagli G, Maccarrone N, Batacchi S, Cammelli R, Cecchi a, *et al.* The use of Modified Early Warning Score may help anesthesists in postoperative level of care selection in emergency abdominal surgery. *Minerva Anestesiol* [Internet]. 2012; **78**: 1034–1038. Available from: http://europepmc.org/abstract/med/22580592

7 Smith T, Den Hartog D, Moerman T, Patka P, Van Lieshout EMM, Schep NWL. Accuracy of an expanded early warning score for patients in general and trauma surgery wards. *Br J Surg* [Internet]. 2012 Feb [cited 2014 Nov 21]; **99**: 192–197. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22183685

8 Bulut M, Cebicci H, Sigirli D, Sak A, Durmus O, Top AA, *et al.* The comparison of modified early warning score with rapid emergency medicine score: a prospective multicentre observational cohort study on medical and surgical patients presenting to emergency department. *Emerg Med J* [Internet]. 2013; **202444**: 1–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23562988

9 Kellett J, Wang F, Woodworth S, Huang W. Changes and their prognostic implications in the

abbreviated VitalPAC™ Early Warning Score (ViEWS) after admission to hospital of 18,827 surgical patients. *Resuscitation* [Internet]. European Resuscitation Council, American Heart Association, Inc., and International Liaison Committee on Resuscitation.~Published by Elsevier Ireland Ltd; 2013 Apr [cited 2015 Mar 15]; **84**: 471–476. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23228559

10 Neary PM, Regan M, Joyce MJ, McAnena OJ, Callanan I. National early warning score (NEWS) - evaluation in surgery. *Int J Health Care Qual Assur* [Internet]. 2015 Jan [cited 2016 Jan 20]; **28**: 245–252. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25860921

11 Cuthbertson BH, Boroujerdi M, McKie L, Aucott L, Prescott G. Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? *Crit Care Med*. 2007; **35**: 402–409.

12 Stark AP, Maciel RC, Sheppard W, Sacks G, Hines OJ. An Early Warning Score Predicts Risk of Death after In-hospital Cardiopulmonary Arrest in Surgical Patients. *Am Surg* [Internet]. 2015 Oct [cited 2015 Dec 7]; **81**: 916–921. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26463280

13 Simmes FM, Schoonhoven L, Mintjes J, Fikkers BG, van der Hoeven JG. Incidence of cardiac arrests and unexpected deaths in surgical patients before and after implementation of a rapid response system. *Ann Intensive Care* [Internet]. 2012; **2**: 20. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3425134&tool=pmcentrez&rendertype=abstract

14 Report of a working party. National Early Warning Score (NEWS) National Early Warning Score (NEWS): Standardising the assessment of acute-illness severity in the NHS. London; 2012.

15 Prytherch DR, Smith GB, Schmidt PE, Featherstone PI. ViEWS-Towards a national early warning score for detecting adult inpatient deterioration. *Resuscitation* [Internet]. Elsevier Ireland Ltd; 2010 Aug [cited 2014 Nov 21]; **81**: 932–937. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20637974

16 Hancock C. The Good NEWS for Wales. Cardiff; 2013.

17 Sprinks J. Swift take-up of standardised early warning system across NHS trusts. *Nurs Stand*. 2013; **27**: 7.

18 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 Apr; **84**: 465–470.

19 Smith GB, Prytherch DR, Schmidt P, Featherstone PI, Knight D, Clements G, *et al.* Hospital-wide physiological surveillance-A new approach to the early identification and management of the sick

patient. *Resuscitation*. 2006; **71**: 19–28.

20 Schmidt PE, Meredith P, Prytherch DR, Watson D, Watson V, Killen RM, *et al.* Impact of introducing an electronic physiological surveillance system on hospital mortality. *BMJ Qual Saf* [Internet]. 2015 Sep 23 [cited 2014 Nov 21]; **24**: 10–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25249636

21 Hands C, Reid E, Meredith P, Smith GB, Prytherch DR, Schmidt PE, *et al.* Patterns in the recording of vital signs and early warning scores: compliance with a clinical escalation protocol. *BMJ Qual Saf* [Internet]. 2013 Sep [cited 2014 Nov 21]; **22**: 719–726. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23603474

22 Hanley J, BJ M. The meaning and use of the area under the receiver operating characteristic (ROC) curve. *Radiology*. 1982; **143**: 29–36.

23 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.

24 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametetric approach. *Biometrics*. 1988; **44**: 837–845.

25 Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011; **12**: 77.

26 R Development Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2014. Available from: http://www.r-project.org./

27 Eccles SR, Subbe C, Hancock D, Thomson N. CREWS: Improving specificity whilst maintaining sensitivity of the National Early Warning Score in patients with chronic hypoxaemia. *Resuscitation* [Internet]. European Resuscitation Council, American Heart Association, Inc., and International Liaison Committee on Resuscitation.~Published by Elsevier Ireland Ltd; 2014 Jan [cited 2014 Nov 21]; **85**: 109–111. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24056390

28 Corfield a. R, Lees F, Zealley I, Houston G, Dickie S, Ward K, *et al.* Utility of a single early warning score in patients with sepsis in the emergency department. *Emerg Med J* [Internet]. 2014; **31**: 482–487. Available from: http://emj.bmj.com/cgi/doi/10.1136/emermed-2012-202186

29 Cattermole GN, Liow ECH, Graham C a., Rainer TH. THERM: the Resuscitation Management score. A prognostic tool to identify critically ill patients in the emergency department. *Emerg Med J* [Internet]. 2014; **31**: 803–807. Available from: http://emj.bmj.com/cgi/doi/10.1136/emermed-2013-202772

LEGENDS FOR FIGURES:

Figure 1

The distribution of NEWS values for non-elective admissions to medical specialties, and non-elective admissions to surgical specialties using the combined outcome of any of death, cardiac arrest or unanticipated ICU admission within 24 hours of a vital signs dataset. Observed risk shown with 95% confidence intervals (CI).

Figure 2

"EWS efficiency curve" for NEWS using the combined outcome of any of death, cardiac arrest or unanticipated ICU admission within 24 hours of a vital signs dataset for non-elective admissions to medical specialties, and non-elective admissions to surgical specialties.

The percentage of observations at, or above a given NEWS value, or Activity (%) = $\frac{(TP+FP)}{(TP+FP+TN+FN)}$ plotted against the percentage of observations which were followed by the Combined Outcome within 24 hours, at or above a given NEWS value, or Sensitivity (%) = $\frac{(TP)}{(TP+FN)}$

Legends for Tables

Table 1: Patient Demographic and observation data regarding admission to one of four patient groups: non-elective or elective surgical specialties, non-elective or elective medical specialties.

Table 2: Distribution of outcomes between non-elective and elective admissions to surgical and medical specialties, by numbers of admissions and observation sets.

Table 3. Area under the Receiver Operator Characteristic (AUROC) curves for outcomes studied (death, cardiac arrest, unanticipated ICU admission and any of these, each within 24 hours of a vital signs dataset) in the patient groups studied: elective or non-elective admissions to surgical or medical specialties. Figures in brackets indicate 95% confidence intervals (CIs)

Supplementary Material:

1. Figure 1: Flow diagram for the study
2. Table 1: The National Early Warning Score (NEWS)
3. Table 2: Table showing sensitivity, specificity, positive predictive value and negative predictive value of NEWS for non-elective admissions to surgical specialties for the Combined Outcome as first event.
4. Table 3: Table showing sensitivity, specificity, positive predictive value and negative predictive value of NEWS for non-elective admissions to medical specialties for the Combined Outcome as first event.
5. The relationship between NEWS values and the incidence of death within 24 hours of a vital signs dataset for non-elective admissions to surgical and medical specialties.
6. The relationship between NEWS values and the incidence of cardiac arrest within 24 hours of a vital signs dataset for non-elective admissions to surgical and medical specialties.
7. The relationship between NEWS values and the incidence of unanticipated ICU admission within 24 hours of a vital signs dataset for non-elective admissions to surgical and medical specialties.
8. "EWS efficiency curve" for NEWS for death within 24 hours of a vital signs dataset for non-elective admissions to surgical and medical specialties.
9. "EWS efficiency curve" for NEWS for cardiac arrest within 24 hours of a vital signs dataset for non-elective admissions to surgical and medical specialties.
10. "EWS efficiency curve" for NEWS for unanticipated ICU admission within 24 hours of a vital signs dataset for non-elective admissions to surgical and medical specialties.
11. ROC curves for non-elective admissions to surgical and medical specialties using the Combined Outcome of any of death, cardiac arrest or unanticipated ICU admission within 24 hours of a vital signs dataset
12. ROC curves for non-elective admissions to surgical and medical specialties for death within 24 hours of a vital signs dataset.
13. ROC curves for non-elective admissions to surgical and medical specialties for cardiac arrest within 24 hours of a vital signs dataset.
14. ROC curves for non-elective admissions to surgical and medical specialties unanticipated ICU admission within 24 hours of a vital signs dataset.