

Aggregate National Early Warning Score (NEWS) values are more important than high scores for a single vital signs parameter for discriminating the risk of adverse outcomes

Dr. Stuart Jarvis PhD.^{1,2}

Mrs. Caroline Kovacs, BSc¹

Dr. Jim Briggs, BA, DPhil¹

Dr. Paul Meredith, PhD³

Dr. Paul E Schmidt, MRCP, B.Med.Sc, MBA³

Dr. Peter I Featherstone, FRCP³

Professor David R Prytherch, PhD, MIPEM, CSci^{1,3}

Professor Gary B Smith, FRCA, FRCP⁴

1. Centre for Healthcare Modelling and Informatics, University of Portsmouth, Portsmouth, UK
2. Department of Health Sciences, University of York, York, UK
3. Portsmouth Hospitals NHS Trust, Portsmouth, UK
4. School of Health & Social Care, University of Bournemouth, Bournemouth, UK

Correspondence to:

Professor G B Smith, FRCA, FRCP,

Centre of Postgraduate Medical Research & Education (CoPMRE),

The School of Health & Social Care, Bournemouth University,

Royal London House, Christchurch Road, Bournemouth, Dorset BH1 3LT, United Kingdom

Tel: +44 (0) 1202 962782

Fax: +44 (0) 1202 962218

Email: gbsresearch@virginmedia.com

Word count = 3123 Number of references = 21 Figures = 3 Tables = 3

Keywords

Monitoring; Physiologic; Vital signs; Failure to rescue; Hospital Rapid Response Team

ABSTRACT

Introduction:

The Royal College of Physicians (RCPL) National Early Warning Score (NEWS) escalates care to a doctor at NEWS values of ≥ 5 and when the score for any single vital sign is 3.

Methods:

We calculated the 24-hour risk of serious clinical outcomes for vital signs observation sets with NEWS values of 3, 4 and 5, separately determining risks when the score did/did not include a single score of 3. We compared workloads generated by the RCPL's escalation protocol and for aggregate NEWS value alone.

Results:

Aggregate NEWS values of 3 or 4 ($n = 142,282$) formed 15.1% of all vital signs sets measured; those containing a single vital sign scoring 3 ($n=36,207$) constituted 3.8% of all sets. Aggregate NEWS values of either 3 or 4 with a component score of 3 have significantly lower risks (OR: 0.26 and 0.53) than an aggregate value of 5 (OR: 1.0). Escalating care to a doctor when any single component of NEWS scores 3 compared to when aggregate NEWS values ≥ 5 , would have increased doctors' workload by 40% with only a small increase in detected adverse outcomes from 2.99 to 3.08 per day (a 3% improvement in detection).

Conclusions:

The recommended NEWS escalation protocol produces additional work for the bedside nurse and responding doctor, disproportionate to a modest benefit in increased detection of adverse outcomes. It may have significant ramifications for

efficient staff resource allocation, distort patient safety focus and risk alarm fatigue.

Our findings suggest that the RCPL escalation guidance warrants review.

BACKGROUND

The focus on the prevention and recognition of patient deterioration¹⁻⁵ has led to the use of early warning scores (EWS) in many countries.⁶⁻⁹ EWS systems allocate points in a weighted manner, based on the derangement of a predetermined set of patient vital signs from an arbitrarily agreed “normal” range. The sum of the allocated points is used to indicate a patient’s severity of illness, and to inform the need to increase the patient’s physiological monitoring or deliver expert help to their bedside. The response to escalation is usually graded, based on the aggregate EWS value,² although in some EWS systems additional rules are applied when the points awarded for a single vital sign exceed a predetermined value.

In 2012, the Royal College of Physicians of London (RCPL) published a standardised National Early Warning Score¹⁰ (NEWS) (Table 1), which has been widely adopted.¹¹ The ability of NEWS and many other EWS to discriminate adverse outcomes has been evaluated,¹²⁻¹⁴ but published descriptions of EWS systems almost never describe the EWS’ escalation protocol. This makes it impossible to determine the sensitivity, specificity and the associated workload, at the EWS’ escalation point. Ideally, at its escalation point, an EWS would have high specificity and sensitivity, ensuring correct identification of ‘at risk’ patients whilst avoiding excessive workload, ‘alarm fatigue’ and staff desensitization – issues that can compromise quality of care and patient safety.^{15,16}

For NEWS, the RCPL recommends escalation of care to a doctor at NEWS values of 5 or greater, and also when 3 points are awarded for any single vital sign.¹⁰ At the same time, the RCPL recommends that the frequency of vital signs monitoring is increased from 4, 6 or 12 hourly to an hourly schedule.¹⁰ A NEWS value of 3 is

possible for severe derangement of six of the seven composite vital signs (i.e., respiration rate, oxygen saturation, temperature, systolic blood pressure, pulse rate and level of consciousness). The remaining vital sign - whether the patient is on supplemental oxygen - has a maximum score of 2.

NEWS is derived from the VitalPAC Early Warning Score (ViEWS),¹⁷ which was designed so that the aggregate ViEWS value would be a good discriminator of risk of death within 24 hours of an observation set. ViEWS was not designed so that a weighting of 3 for a single vital sign should indicate a greater risk of death than, say, an aggregate score of 3 from component scores of (2 + 1) or (1 + 1 + 1). The decision to escalate when any weighting in a single component of NEWS is 3 was a pragmatic, consensus view of the NEWS Development and Implementation Group (NEWSDIG).

As NEWS is closely related to ViEWS, we hypothesised that it is the aggregate NEWS value, rather than the weighting for a single vital sign, which is most important in discriminating risk. In this paper, we studied the risks and workload associated with NEWS values in the range 3-5 to determine whether it might be appropriate to escalate care based only on the aggregate NEWS value, rather than also when a single vital sign has a score of 3. We also investigated which components most frequently contribute a NEWS value of 3 (and calculated their associated risks).

METHOD

Ethical Committee Approval

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

Vital signs test results database and its development

We searched the computerised hospital records of all discharged adult patients admitted to Portsmouth Hospitals NHS Trust (PHT) on or after 21 April 2010 and discharged on or before 23 May 2011. We excluded data from patients aged <16 years at hospital admission and patients discharged alive on the day of admission. Vital signs data were recorded in real-time at the bedside using handheld electronic equipment running the VitalPAC software (The Learning Clinic, London).¹⁸ Each set of vital signs measurements contained: pulse rate, breathing rate, systolic and diastolic blood pressure, temperature, S_pO_2 , the inspired gas (i.e., oxygen or air) at the time of S_pO_2 measurement, and the patient's conscious level. Conscious level was recorded as alert (A); responds to voice (V); responds to pain (P); or unresponsive (U). Observation sets for which one or more of the vital signs measurements were absent or physiologically impossible (i.e., recorded in error) or for which the patient was sedated – and thus their underlying consciousness level was unknown - were excluded from further analysis in this study. Practice within PHT during the period of data acquisition was such that escalation of care was mandatory only when the aggregate NEWS value was ≥ 6 . No escalation was made based on a score of 3 in a single component of NEWS.

Outcomes

We studied the following outcomes: death, cardiac arrest and unanticipated ICU

admission, each within 24 hours of an observation set. Patient outcomes were identified using the hospital's patient administration system (for the outcome of death), and its cardiac arrest and ICU admission databases. We used precedence rules so that, when multiple adverse outcomes occurred within 24 hours of an observation set, only the first outcome was counted (i.e. a cardiac arrest, followed by an ICU admission, followed by death – all within 24 hours of an observation set – was recorded as a cardiac arrest only).

Analysis of the data

We calculated the observed risk of death, cardiac arrest, unanticipated ICU admission and any of those outcomes within 24 hours of a vital signs dataset, for observations with NEWS values of 3, 4 and 5. For observations with NEWS values of 3 or 4, we separately determined these risks when the score did and did not include a single vital sign scoring 3. We did not do this for aggregate scores of ≥ 5 as, under RCPL guidance,¹⁰ these trigger anyway; neither did we consider aggregate scores < 3 as these could not contain a component scoring 3. When the aggregate NEWS value was 3 or 4, we also determined the risks associated with and prevalence of a score of 3 for each of the six vital sign components in which that score is possible.

In the above analyses, to eliminate any effects of different numbers of observations for sicker patients or those with longer stays, we calculated the risks by randomly selecting one set of vital signs from each episode of care, thus giving equal weight to each episode. This process was repeated 10,000 times (with independent random selections of 45,678 observations - one observation from each episodes – each

time). 95% confidence intervals (CI) were estimated from the 2.5-97.5 centiles in risks determined across the 10,000 sets.

We assessed the significance of differences between risks by comparing the differences in each of the 10,000 sample sets. We determined the 95% CI (from the 2.5-97.5 centiles) of the difference between risks associated with pairs of scores and considered differences significant if the CI did not include zero. All use of 'significant' hereinafter refers to this test.

Also using one randomly selected observation per episode, we determined the workload (triggering rate) compared to sensitivity under the different escalation criteria. The resulting EWS efficiency curves¹⁷ compare the triggering rate (share of observations that result in a trigger for escalation) with the sensitivity (share of observations that are followed by an adverse outcome within 24 hours that would trigger) for each of the outcomes. We present the mean curves from the 10,000 sample sets when triggering only on aggregate NEWS ≥ 5 and when also triggering on single component scores of 3.

The efficiency curves represent an idealised situation in which all patients receive the same number of observations. In practice, the number of escalations and detected adverse outcomes that would result from the escalation criteria being applied to all collected observations is of greater interest. Therefore, we also determined the number of escalations and the number of detected adverse outcomes per day (mean values over the 488 day study period) in our study hospital. To do this, we determined the total number of escalations under each escalation strategy and also the number of episodes in which there was an adverse outcome

for which at least one escalation would have occurred within the 24 hours preceding the adverse outcome. We excluded episodes for which there was no observation set taken within the last 24 hours before an adverse outcome. In line with the hospital's observation protocol, this should include all patients except those on an end of life pathway.

All data manipulation was performed using Microsoft® Visual FoxPro 9.0. All analyses were undertaken in R version 3.02.¹⁹

RESULTS

After excluding 0.14% of observation sets for incomplete and incorrect vital signs observations, and those for which the patient was sedated, there were 942,887 complete, valid observation sets from 45,678 completed episodes of care (mean 20.6 observation sets per episode) where the patient was aged ≥ 16 and was not discharged alive on the day of admission. Of these, 142,282 observation sets (15.1%) had a NEWS value of 3 or 4, and 36,207 (3.8% of the total) included a single vital sign scoring 3. The risks of outcomes for NEWS values of 3, 4 and 5 with different component scores are summarised in Figure 1.

Risks for all outcomes studied increase with greater aggregate NEWS scores (Figure 1). There is an approximate doubling of risk for each increase of 1 in aggregate NEWS value (Table 2) for each outcome, although differences were not always significant. Scoring 3 in a single component when aggregate NEWS is 3 or 4 indicates a higher risk of an adverse outcome within 24 hours than the same NEWS value with no component scoring 3, for all outcomes (Figure 1; Table 2). However, the differences for this are not statistically significant.

An aggregate score of 5 represents a significantly higher risk than an aggregate score of 3 with 3 in a single component for all the outcomes studied and a significantly higher risk than an aggregate score of 4 with 3 in a single component for death or any adverse outcome (Figure 1, Table 2). An aggregate score of 4 with no component scoring 3 indicates higher risk than an aggregate score of 3 with a component scoring 3, for all outcomes (Table 2), but the differences in risk are not statistically significant.

Figure 2 considers the prevalence of scores of 3 in a single component of NEWS (when aggregate NEWS is 3 or 4) and the associated risks. Only a score of 3 for temperature (i.e., $\leq 35.0^{\circ}\text{C}$) represents a risk – of cardiac arrest – that is significantly higher than the risk for aggregate NEWS = 5. Low temperature is rare, so conscious level (where scores of 3 are much more prevalent) may be a more useful indicator of increased risk of adverse outcomes, although the risks are not shown to be significantly higher than the risks associated with an aggregate NEWS value of 5.

Figure 3 and table 3 illustrate that the workload generated is greater if escalation occurs when any single component of NEWS scores 3 in addition to the aggregate NEWS value of 5 than when only the aggregate NEWS value of 5 is used. In the idealised setting of equal numbers of observations for each episode (observations chosen at random, one per episode) escalating care on single component scores of 3 in addition to aggregate NEWS ≥ 5 ,¹⁰ results in escalation for 11.3% of observations compared to 7.5% when triggering only at aggregate NEWS ≥ 5 (Figure 3). The workload increases by 51%. In contrast, the additional outcome detection (for any outcome) only increases from 70.5% to 75.2% (a proportional increase in detection of 6.7%). Using the full set of observations from our hospital over a 488 day period (with differing numbers of observations per episode – a more realistic situation) workload would increase by 40% from 220 to 307 escalations per day with only a small increase in detected adverse outcomes from 2.99 (93.6% of adverse outcomes) to 3.08 (96.3%) (Table 3). This corresponds to an increase in detection of 3.0%.

DISCUSSION

The results show that scoring 3 in a single component when the aggregate NEWS value is 3 or 4 can indicate a higher risk of an adverse outcome within 24 hours than the same NEWS value with no component scoring 3. However, the differences are not statistically significant and are smaller than differences in risk between different aggregate scores in NEWS.

When considering care escalation using RCPL guidance,¹⁰ an aggregate NEWS value of 3 with a component score of 3 is treated the same as an aggregate NEWS value of 5. However, our data reveal that, for all outcomes, the latter indicates significantly higher risk – typically around four times the risk (Table 2). An aggregate NEWS value of 4 with no component scoring 3 (which would not trigger escalation) also indicates higher risk than an aggregate NEWS value of 3 with a component scoring 3 (which would trigger escalation) for all outcomes, although the differences are not significant. These findings imply that the RCPL guidance should be reviewed.

Considering the balance of workload against sensitivity (Figure 3 and Table 3), it is clear that triggering whenever a single component in NEWS scores 3 increases the workload required to detect the same number of adverse outcomes. Whether in the idealised situation of equal numbers of observations per patient, or based on actual observation patterns from our study hospital, workload (as measured by escalations) is greatly increased for small improvements in detection. When using all observations within the hospital, detection typically improved by fewer than 3 percentage points, although it increased by 5.1 percentage points for cardiac arrest, the least common adverse outcome. This represents evidence that triggering for NEWS values of 3 or 4 when any individual component scores 3 may have

significant quality and safety ramifications for staff, patients and the hospital in general. Increased vital sign monitoring produces additional work for the bedside nurse, and distracts attention from other important duties. Increased call outs to doctors risks alarm fatigue^{15, 16} and diverts medical care from patients at greater need. Increased monitoring and assessment disturbs patient sleep,²⁰ meals, etc.

However, NEWSDIG's recommendation that NEWS values of 3 or 4 with a component scoring 3 should escalate care is clinically, not statistically, based. A highly deranged value of a single vital sign (e.g., low blood pressure, high respiratory rate) may indicate that a patient is at imminent risk of death or other adverse outcome and it would be wrong not to escalate care in such cases. Extreme values should lead to staff concern and escalation of care, irrespective of the use of any early warning score.¹⁰ Our study suggests that the level at which a score of 3 is awarded to a single vital sign in NEWS is too low, by itself, to indicate a risk of imminent adverse outcome similar to that associated with the aggregate triggering score of 5. Scoring 3 for extreme temperature (i.e., $\leq 35.0^{\circ}\text{C}$) indicates significantly higher risk for cardiac arrest than an aggregate score of 5, but a score of 3 in other components does not indicate significantly elevated risk of any of the outcomes.

There may be a case for defining extreme values for each vital sign beyond which care should be escalated to more experienced staff, irrespective of the aggregate NEWS score. These extreme values should be more severely deranged than those that presently score a 3 in NEWS. The level of derangement would have to be separately determined (perhaps based on the levels at which risk of adverse outcome is similar to an aggregate NEWS value of 5). An alternative would be to preserve existing observation frequency protocols, i.e. increasing frequency of

observation following a score of 3 in a single component, but not to escalate solely on such a score. This would ensure that patients scoring 3 in a single component – who appear to have slightly elevated risks – would be closely monitored, without generating extra escalations that may be seen as unnecessary false alarms and potentially divert interventions from higher risk patients.

Finally, it should be appreciated that the findings of our analysis of the NEWS escalation protocol may have significance for other EWS systems that choose to escalate care based on an extreme value of a single vital sign. Whilst a different EWS using a different vital sign weighting system and different escalation point might result in a different workload and detection profile, the principles that we have investigated are pertinent.

The study has several strengths. It uses a large database from over a year of completed, hospital wide inpatient admissions from a period when escalation of care was mandatory only when the aggregate NEWS value was ≥ 6 . However, there are also weaknesses. During the study period, staff were encouraged to call for help when worried irrespective of the aggregate NEWS score. Therefore, it is possible that some observation sets of NEWS values of 3, 4 or 5 that generated 'staff concern' resulted in escalation, and that this may have influenced the risks that we have calculated. It is also possible that staff concern might have been more likely for a patient with a single highly deranged vital signs measurement than one with several mildly deranged measurements. This may have led to a greater number of interventions for those patients scoring 3 in a single component, compared to those who did not. Further, it is also easier to detect, and select the correct specific therapy for, a patient with overt single organ derangement than one with multiple subtle

physiological abnormalities. We are unable to quantify such effects in our data. Other weaknesses of the study are the lack of information on interventions that may have taken place and the fact that it is a single centre study. There is no guarantee that similar results would be obtained in other institutions. However, NEWS was derived directly from ViEWS, for which there is increasing evidence of its performance outside the UK²¹. As physiology across the world seems similar, there is no reason to suggest that our findings would not be applicable to other countries where NEWS is used. An external validation exercise is nonetheless desirable.

Summary:

The recommended NEWS escalation protocol produces additional work for the bedside nurse and responding doctor, disproportionate to a modest benefit in increased detection of adverse outcomes. It may have significant ramifications for efficient staff resource allocation, distort patient safety focus and risk alarm fatigue. There may be a case for defining extreme values for each vital sign beyond which care should be escalated to more experienced staff, irrespective of the aggregate NEWS score. These extreme values should be more severely deranged than those that presently score 3 in NEWS. Our findings suggest that the RCPL escalation guidance warrants review.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the efforts of the medical, nursing and administrative staff at Portsmouth Hospitals NHS Trust who collected the data used in this study. Dr. Stuart Jarvis takes responsibility for the integrity and the accuracy of the data analysis.

COMPETING INTERESTS

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. PM, DP, PF and PS are employed by PHT. GS was an employee of PHT until 31/03/2011. PS, PF, and the wives of GS and DP, are minority shareholders in TLC. GS, DP, and PS are unpaid research advisors to TLC, and have received reimbursement of travel expenses from TLC for attending symposia in the UK. JB's research has previously received funding from TLC through a Knowledge Transfer Partnership.

GS, DP, PS and PF were members of the clinical team that developed the VitalPAC Early Warning Score (ViEWS). GS was also a member of the Royal College of Physicians of London's National Early Warning Score (NEWS) Development and Implementation Group. DP assisted the Royal College of Physicians of London in the analysis of data validating NEWS. SJ and CK have no conflicts of interest.

FUNDING

None.

REFERENCES

1. National Patient Safety Agency. Recognising and responding appropriately to early signs of deterioration in hospitalised patients. NPSA, London, 2007.
2. National Institute for Health and Clinical Excellence: Acutely ill patients in hospital: recognition of and response to acute illness in adults in hospital. NICE clinical guideline No.50. London; 2007.
3. DeVita MA, Bellomo R, Hillman K, Kellum J, Rotondi A, Teres D, et al. Findings of the first consensus conference on medical emergency teams. *Crit Care Med.* 2006;34:2463-78.
4. DeVita MA, Smith GB, Adam S et al. Consensus Conference on the Afferent Limb: Identifying Hospitalised Patients in Crisis. *Resuscitation.* 2010; 81:375-382.
5. Australian Commission on Safety and Quality in Health Care (2010). National Consensus Statement: Essential Elements for Recognising and Responding to Clinical Deterioration. Sydney, ACSQHC.
6. Kyriacos U, Jelsma J, James M, Jordan S. Monitoring Vital Signs: Development of a Modified Early Warning Scoring (MEWS) System for General Wards in a Developing Country. *PLoS One.* 2014;24;9:e87073.
7. De Meester K, Das T, Hellemans K, Verbrugghe W, Jorens PG, Verpooten GA et al. Impact of a standardized nurse observation protocol including MEWS after Intensive Care Unit discharge. *Resuscitation.* 2013;84:184-8.
8. Smith T, Den Hartog D, Moerman T, Patka P, Van Lieshout EM, Schep NW. Accuracy of an expanded early warning score for patients in general and trauma surgery wards. *Br J Surg.* 2012;99:192-7.

9. Cei M, Bartolomei C, Mumoli N. In-hospital mortality and morbidity of elderly medical patients can be predicted at admission by the Modified Early Warning Score: a prospective study. *Int J ClinPract* 2009; 63:591-595.
10. National Early Warning Score (NEWS): Standardising the assessment of acute-illness severity in the NHS. Report of a working party. Royal College of Physicians, London, 2012.
11. Sprinks J. Swift take-up of standardised early warning system across NHS trusts. *Nursing Standard* 2013; 27:7.
12. 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;84:465-470.
13. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI. Review and performance evaluation of aggregate weighted 'track and trigger' systems. *Resuscitation*. 2008;77:170-179.
14. Tirkkonen J, Olkkola KT, Huhtala H, Tenhunen J, Hoppu S. Medical emergency team activation: performance of conventional dichotomised criteria versus national early warning score. *ActaAnaesthesiol Scand*. 2014;58:411–419.
15. Sendelbach S, Funk M. Alarm Fatigue: A Patient Safety Concern. *AACN AdvCrit Care*. 2013;24:378-386.
16. Gazarian PK. Nurses' response to frequency and types of electrocardiography alarms in a non-critical care setting: A descriptive study. *Int J Nurs Stud*. 2014;51:190-197.
17. Prytherch DR, Smith GB, Schmidt PE, Featherstone PI. ViEWS - towards a national Early Warning Score for detecting adult inpatient deterioration. *Resuscitation*. 2010;81:932–937.

18. Smith GB, Prytherch DR, Schmidt P, et al. Hospital-wide physiological surveillance - a new approach to the early identification and management of the sick patient. *Resuscitation*. 2006;71:19-29.
19. R Core Team 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
20. Sharda S, Carter J, Wingard JR, Mehta P. Monitoring vital signs in a bone marrow transplant unit: are they needed in the middle of the night? *Bone Marrow Transplant*. 2001;27:1197-1200.
21. Smith G. Have we found the perfect early warning score? A view of ViEWS. *Resuscitation* 2013;84:707-8.

LEGENDS FOR FIGURES:

Figure 1: Risk of an adverse outcome within 24 hours of an observation set scoring 3, 4 or 5 on the National Early Warning Score. For those scoring 3 or 4, risks are split by whether the score included a single component score of 3.

Figure 2: Prevalence of and risks associated with vital sign measurements responsible for the single component score of 3 in observation sets with NEWS values of 3 or 4, for a range of outcomes. Dashed horizontal lines indicate the risk associated with an aggregate NEWS score of 5 (shaded area is the 95% CI).

Figure 3: Comparison of EWS efficiency curves for NEWS when care is escalated (a) at NEWS values ≥ 5 and (b) when NEWS values ≥ 5 or 3 points are awarded for any single vital sign. The efficiency curve plots the percentage of the observations that would trigger at, or above, a given NEWS value against the sensitivity at that NEWS value (the percentage of the observations for which the outcome within 24 hours was true). From the point at 100,100 the NEWS values are 0, 1, 2, 3 Points associated with triggering at aggregate NEWS ≥ 5 are labelled. Shaded regions indicate 95% CI.

Table 1: The National Early Warning Score (NEWS)¹⁰

Physiological parameters	3	2	1	0	1	2	3
Respiration Rate (breaths per minute)	≤8		9 - 11	12 - 20		21 - 24	≥25
SpO ₂ %	≤91	92 - 93	94 - 95	≥96			
Any supplemental oxygen?		Yes		No			
Temperature (°C)	≤35.0		35.1 - 36.0	36.1 - 38.0	38.1 - 39.0	≥39.1	
Systolic BP (mmHg)	≤90	91 - 100	101 - 110	111 - 219			≥220
Heart/pulse rate (beats per minute)	≤40		41 - 50	51 - 90	91 - 110	111 - 130	≥131
Level of consciousness using the AVPU system				A			V, P or U

Table 2: Risk (expressed as the odds ratio, compared to an aggregate NEWS value of 5) for scores that would and would not trigger under RCP guidance.¹⁰ 95% confidence intervals are in parentheses. * indicates a significantly ($p < 0.05$) higher risk than the value immediately below; ^ indicates a significantly higher risk than the value two places below.

	Odds ratio, compared to NEWS = 5 (95% CI)			
	Death	Cardiac arrest	Unanticipated ICU admission	Any of these outcomes
Triggering combinations of NEWS				
5	1.00 (0.72 - 1.29)*^	1.00 (0.59 - 1.44)^	1.00 (0.55 - 1.49)^	1.00 (0.79 - 1.22)*^
4 (includes a component = 3)	0.53 (0.25 - 0.85)	0.66 (0.17 - 1.26)	0.46 (0.00 - 0.99)	0.54 (0.32 - 0.79)*
3 (includes a component = 3)	0.26 (0.12 - 0.42)	0.24 (0.00 - 0.55)	0.23 (0.00 - 0.52)	0.25 (0.14 - 0.37)
Non-triggering combinations of NEWS				
4 (no component = 3)	0.38 (0.22 - 0.56)*	0.43 (0.14 - 0.74)	0.45 (0.13 - 0.80)	0.41 (0.27 - 0.55)*
3 (no component = 3)	0.20 (0.12 - 0.28)	0.21 (0.07 - 0.36)	0.22 (0.09 - 0.38)	0.20 (0.14 - 0.27)

Table 3: Typical numbers of escalations and detected outcomes per day (based on the study hospital results averaged over a 488 day period). Detected outcomes refer to the number of patients experiencing an adverse event for whom there would have been at least one escalation of care under each of the triggering criteria. The figures for deaths include all deaths, whether preceded or not by another of the adverse outcomes. Cardiac arrest and unanticipated ICU admission are only counted if they were the first adverse outcome within 24 hours of an observation set.

Triggering criteria	Number of escalations/day	Number of detected deaths/day	Number of detected cardiac arrests/day	Number of detected unanticipated ICU admissions/day	Number of detected adverse outcomes/day
Trigger only on aggregate NEWS ≥ 5	220	2.24 (94.6%)	0.35 (87.3%)	0.78 (92.2%)	2.99 (93.6%)
Trigger on aggregate NEWS ≥ 5 and any single component score of 3	307	2.30 (96.9%)	0.37 (92.4%)	0.80 (95.1%)	3.08 (96.3%)

Figure 1

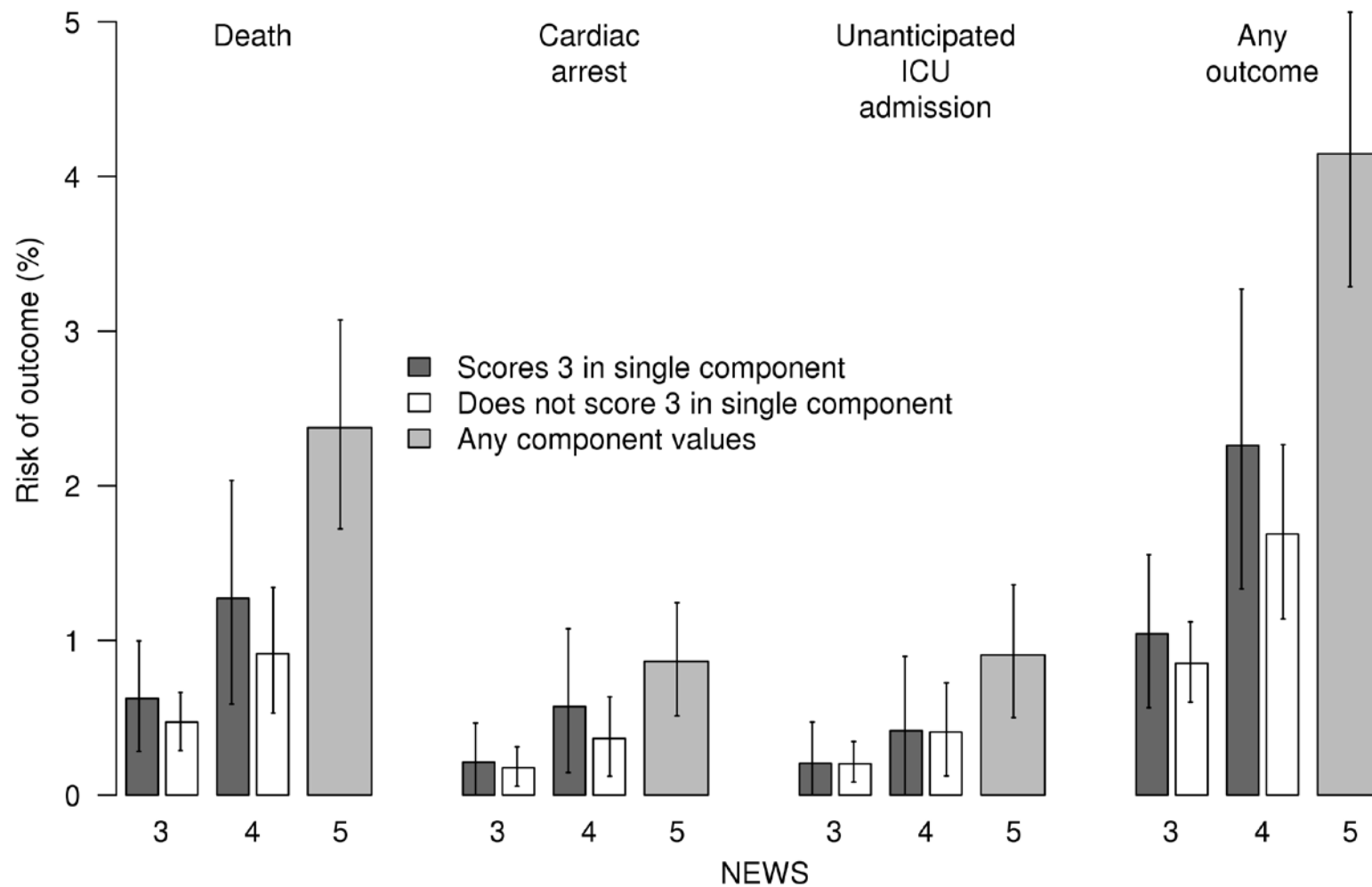
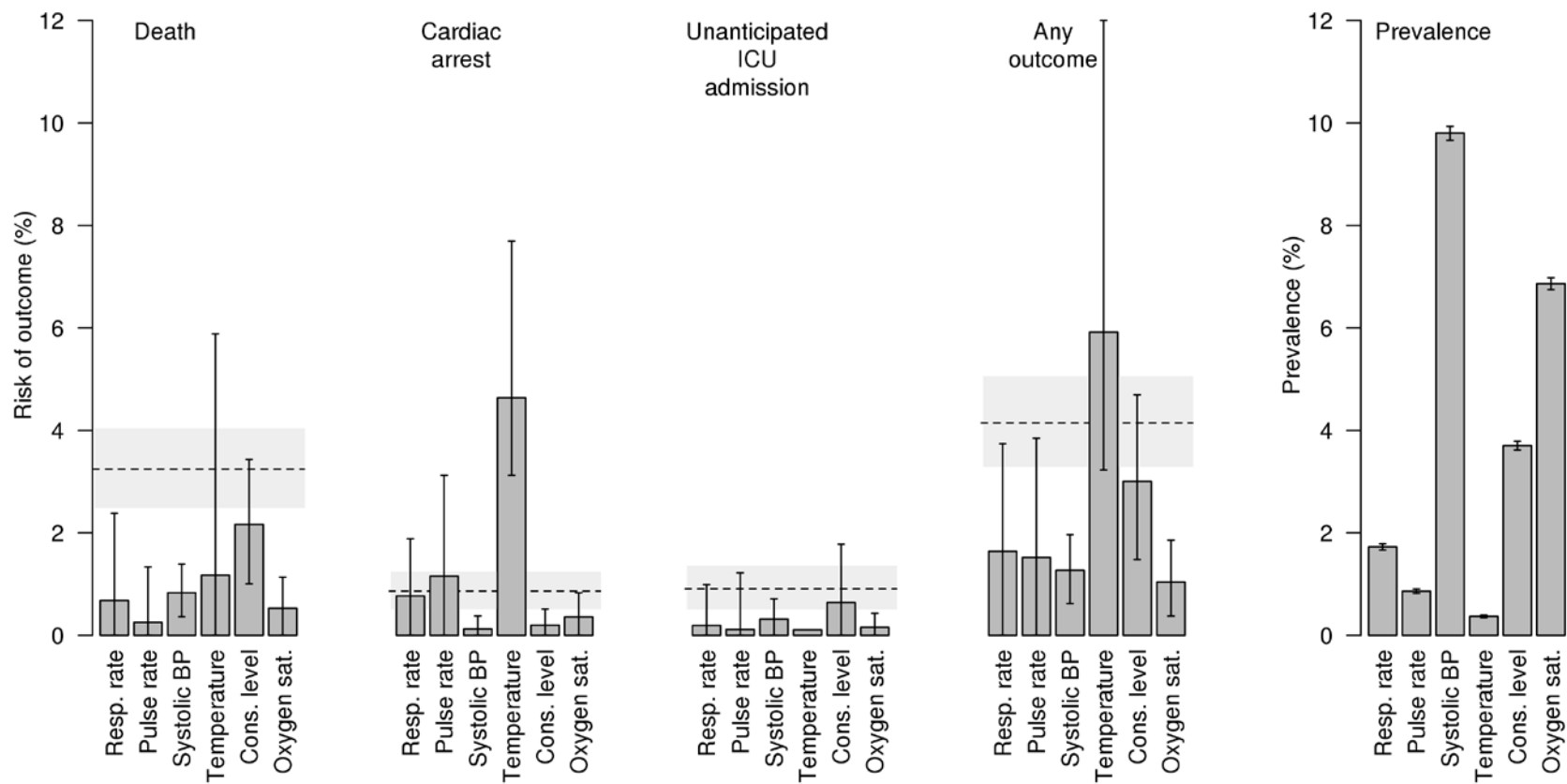


Figure 2



Component scoring 3

Figure 3

