Paper

The use of combined physiological parameters in the early recognition of the deteriorating acute medical patient

[1]BH Cuthbertson, [2]M Boroujerdi, [3]G Prescott
[1]*Professor of Critical Care, Department of Critical Care Medicine, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada; [2]Lecturer in Medical Statistics; [3]Senior Lecturer in Medical Statistics, Department of Public Health, University of Aberdeen, Aberdeen, UK*

ABSTRACT

Background: Early warning scores (EWS) are widely used to allow early recognition of the deteriorating patient. We aimed to test their ability to predict major deterioration in medical patients.
Methods: Two cohorts were prospectively identified who were admitted to an acute medical admissions unit and to the respiratory unit but not admitted to the intensive care unit (ICU): medical-non ICU and respiratory-non ICU groups. Two further cohorts were retrospectively identified that required ICU admission from these units (medical-ICU and respiratory-ICU groups). Discriminant analysis and receiver operating characteristic curves were used to discriminate between groups, and time relationships were analysed.
Results: Heart rate (HR), respiratory rate (RR) and oxygen saturation (SaO$_2$) were significantly higher in the medical-ICU group than the medical-non ICU group and significantly higher in the respiratory-ICU group than in the respiratory-non ICU group. Discriminant functions incorporating HR, RR and SaO$_2$ performed at least as well as existing EWS systems in predicting ICU admission.
Conclusions: Commonly used physiological parameters and existing EWS systems are useful at identifying sick patients. The discriminant functions described here appear to have a role in this setting but require validation in future studies.

KEYWORDS Early warning scores, intensive care, medicine, risk prediction, scoring systems

DECLARATION OF INTERESTS This work was funded by a grant from the Chief Scientist's Office for Scotland.

Correspondence to **BH Cuthbertson, Department of Critical Care Medicine, Sunnybrook Health Sciences Centre, 2075 Bayview Avenue, Toronto, Ontario, Canada**

tel. **+1 416 480 6100**
e-mail
brian.cuthbertson@sunnybrook.ca

In the UK large national surgical audits conclude that current systems of post-operative care fail to detect or respond appropriately to early signs of critical illness.[1,2] Recently such national audits have focused on acute general medical care in the UK and demonstrate similar findings.[3] Failure to recognise and intervene during clinical deterioration and failure to use senior help have been repeatedly identified as problems. Such failures in care can lead to deterioration in a patient's condition and admission to the intensive care unit (ICU) or death.

It is recognised that late admission to ICU is associated with higher severity of illness and inevitably leads to a higher mortality.[1–11] It is also recognised that patients who die in ICU often have very high hospital costs even when the obvious human cost is excluded.[5,7] In-hospital cardiac arrest has also been identified as often being a late and predictable event with an appalling prognosis that is preceded by 8–12 hours of clinical deterioration.[11] The link between abnormal physiology and outcome from critical illness has been clearly demonstrated by the Acute Physiology And Chronic Health Evaluation II (APACHE II) system and in work looking at the physiological changes over time.[12–18] A confidential enquiry into the quality of care before admission to intensive care demonstrated

that the care offered to critically ill patients was sub-optimal in the majority of cases.[4,5] Thus, inability to recognise and intervene during the early signs of deterioration leads to late ICU admission, excess mortality and increased hospital costs.

Recent UK guidelines state that 'track and trigger' systems should be used in identifying the acutely deteriorating hospital patient[19–21] but also state that further work is required to develop and validate the most accurate scores for differing clinical environments. Most existing early warning scoring (EWS) systems use routine physiological criteria with cut-off points to alert medical staff of a critical deterioration in the patient's condition. Although these parameters seem clinically intuitive and rational, they comprise best-guess choices of parameters and cut-off points and lack adequate prospective clinical validation.[12–15,21,22] For this work we hypothesised that to allow effective early recognition and intervention of impending critical illness we must develop a validated physiological scoring system to alert ward staff of worsening clinical condition and thus allow early, appropriate intervention. We also hypothesised that these scores must be developed and validated for specific patient groups to allow optimal accuracy.

BH Cuthbertson, M Boroujerdi, G Prescott

CLINICAL

METHODS

The requirement for ethical approval was waived by the local research ethics committee. Aberdeen Royal Infirmary is a 900-bed acute hospital offering all acute clinical services. The hospital does not have an outreach team. The acute medical admissions unit (AMAU) admits most general medical admissions, excluding patients who are admitted to the coronary care unit, gastrointestinal bleeding unit and the respiratory unit. The AMAU supplies level 1 care for acute medical patients. The respiratory unit admits most acute respiratory admissions and acute exacerbations of chronic chest disease. The respiratory unit supplies level 1 and 2 care for patients, including the provision of non-invasive ventilation. The intensive care unit (ICU) in Aberdeen has 14 fully funded level 3 ICU beds.

The data set thus has two broad groups of patients, those admitted to the AMAU and those admitted to the respiratory unit. This is further divided into two sub-groups determined by the group's ICU admission status. This leaves the following four groups: medical- non ICU, medical-ICU, respiratory-non ICU and respiratory-ICU.

For the medical-non ICU and respiratory-non ICU groups, patients were recruited from 21 July until 2 September 2005 (six weeks). These patients were non-consecutive, which explains why the study number is lower than the normal admission number to the unit in such a period. Data from the first 48 hours in hospital were collected prospectively and patients were only recruited when data collectors were able to identify the patients at the time of admission to allow data quality control and assurance.

The medical-ICU and respiratory-ICU groups were identified retrospectively between January 2005 and December 2005 from the ICU clinical database. There was no overlap between cohorts. Physiology data were collected for the last 24 hours before ICU admission (if patients were admitted to hospital more than 24 hours before admission to ICU) or for the full period of their pre-ICU hospital admission (if this period was less than 24 hours).

Data collected included the following physiological parameters: heart rate (HR), respiratory rate (RR), systolic blood pressure (SBP), temperature (temp), oxygen saturation (SaO_2). All data points were taken from routine clinical records and charts collected by clinical staff. An a priori decision was made to exclude individual physiological parameters from the analysis that had less than 60% complete data points according to standard protocol. Existing discrete EWS scores were also calculated from this data using standard methods.[13,14,23,24]

TABLE 1 Patient baseline demographics

Group	n (%)	Age years (SD)	Male n (%)
Medical-non ICU	230 (49)	68.2 (17.2)	105 (45.7)
Medical-ICU	61 (13)	61.6 (16.6)	34 (55.7)
Respiratory-non ICU	107 (23)	64.3 (19.1)	59 (55.1)
Respiratory-ICU	68 (15)	60.9 (14.5)	38 (55.9)
Total	466 (100)	65.4 (17.4)	236 (50.6)

Due to the 'noisy' nature of the data we summarised the data set and 'collapsed' by patient identity number to produce summary statistics for each patient. Previous work by the group revealed that maximum and minimum data had similar or lower predictive accuracy.[23] Other summary statistics representing change over time within a patient were investigated, including change between any two successive measurements. Since the four groups were predefined with obvious comparator groups, logistic regression was used to assess the differences between these groups. The logistic regression enabled us to compute sensitivity and specificity for each of our measures. Cut-points used represent the best compromise between sensitivity and specificity, as determined by receiver operator characteristic (ROC) curves.

Discriminant analysis was deemed the appropriate methodology due to the nature of the data. A linear discriminant function was used to separate the groups. We collapsed the data set to obtain discriminant functions and then applied these discriminant functions to all measurements. At first all of the five parameters that made the a priori inclusion criteria were used in the analysis. As expected, from previous studies and the results of the univariate statistical tests, HR, followed by RR and SaO_2, had the highest structure matrix values and lowest p-values, both of which indicate the statistical importance of these parameters. Data were collected on a Microsoft Access™ database and analysed on a STATA™ statistical package.

RESULTS

The database had 236 medical-non ICU group patients of which six were dropped due to missing data, 70 medical-ICU group patients of which nine were dropped, 120 respiratory-non ICU group patients of which 13 were dropped and 70 respiratory-ICU group patients of which two were dropped. All data relating to observations for more than 24 hours in the ICU groups and 48 hours for the non ICU groups were also dropped in line with data collection and analysis plans. Baseline characteristics of analysed patients are shown in Table 1.

Summarisation of data produced one median measure per physiological parameter per patient. The mean and standard deviation (SD) within each group of these summary median values, as well as probabilities, area

TABLE 2 The differences in the physiological variables (such as heart rate, respiratory rate, etc.) between the ICU and non-ICU admissions in both the medical and respiratory groups using the median value for each parameter for each patient

Variable	n (ICU group)	n (non ICU group)	Mean (CI) (ICU group)	SD	Mean (CI) (non ICU group)	SD	p	AUC	Cut
ACUTE MEDICAL ADMISSIONS									
HR (min^{-1})	60	230	108 (101–115)	27	80 (78–82)	16	<0.0001	0.80	93
RR (min^{-1})	52	223	25 (23–27)	7	18 (18–19)	3	<0.0001	0.84	22
sysBP (mmHg)	61	230	119 (111–127)	31	125 (122–127)	19	0.069	0.56	na
Temp (°C)	60	230	36.4 (36.1–36.8)	1.3	36.5 (36.4–36.6)	0.5	0.551	0.44	na
SaO$_2$ (%)	59	230	91 (90–93)	6	95 (95–96)	2	<0.0001	0.73	94
ACUTE RESPIRATORY ADMISSIONS									
HR (min^{-1})	68	107	112 (107–117)	21	90 (86–94)	18	<0.0001	0.80	105
RR (min^{-1})	52	106	29 (27–32)	9	22 (21–23)	5	<0.0001	0.77	24
sysBP (mmHg)	68	107	127 (120–134)	28	125 (121–8)	17	0.49	0.55	na
Temp (°C)	59	107	36.6 (36.4–36.8)	0.8	36.4 (36.3–36.6)	0.5	0.10	0.62	na
SaO$_2$ (%)	66	107	89 (87–91)	8	94 (93–95)	3	<0.0001	0.74	93

Abbreviations: AUC – area under receiver operator characteristic curve; CI – 95% confidence interval; cut – chosen cut-off point derived from receiver operator characteristic curves; HR – heart rate; ICU – intensive care unit; n – number; na – not applicable; p – probability; RR – respiratory rate; SaO$_2$ – oxygen saturation; SD – standard deviation; sysBP – systolic blood pressure; temp – temperature

under ROC curve (AUC) and cut-points for the physiological parameters, are shown in Table 2 for each group. Discriminant analysis was used to derive a function that maximised the differences between groups. The discriminant functions with five variables are shown in Table 3. The variables HR, RR and SaO$_2$ were found to have the highest discriminatory power. The discriminant analysis was repeated with these variables only.

The canonical discriminant functions are presented for the two groups (medical and respiratory) with f representing discriminant functions and the number representing the number of variables within the function (1 used all five parameters, 2 used three parameters [HR, RR and SaO$_2$] and 3 used two parameters [HR and RR]). Thus, for example, f2med represents the discriminant function derived from analysis of HR, RR and SaO$_2$ in medical patients. Data for f2med and f2resp as well as the function derived previously from a cohort of surgical patients (f2surg)[23] can be seen in Table 3. The discriminant analysis was repeated using the parameters HR and RR only, represented by f3med and f3resp in Table 3. The functions were applied to the median data set.

In addition, four other early warning systems were also applied: the EWS, Modified Early Warning System[13] (MEWS), Patient At Risk Team Score[14,15] (PART) and Scottish Early Warning System[25] (SEWS). By assigning our groups the values 0 and 1 and treating this as a dependent variable (for example, medical-non ICU and medical-ICU) and the corresponding function values or early warning scores as independent variables, we could use logistic regression to identify the differences between groups. This was in addition to using a conventional t-test. Table 4 shows the sensitivity and specificity for each of the

functions together with a 'cut-point' which corresponds to the top left-hand corner point of the ROC curve. This point gives equal weight to sensitivity and specificity when balancing these two. This corner point is one of many allowable in a ROC curve and the choice would ultimately depend on the value and purpose of a test. In a test where sensitivity or specificity presides (to allow detection of all deteriorating patients), the cut-point could be altered to reflect this precedence.

Differences in parameters between time points

The differences between the temporally consecutive measurements were computed. The measurements used were HR, RR and SaO$_2$ as these were the dominant

TABLE 3 Discriminant functions for all groups. In this table f represents discriminant functions and the number represents the number of variables within the function as demonstrated in the abbreviations section below

f1med = 0.028HR + 0.131RR – 0.009sysBP – 0.237temp – 0.141SaO$_2$ + 18.128
f1resp = 0.032HR + 0.069RR – 0.008sysBP – 0.056temp – 0.093SaO$_2$ + 6.918
f2med = 0.026HR + 0.137RR – 0.137SaO$_2$ + 8.045
f2resp = 0.031HR + 0.068RR – 0.088SaO$_2$ + 3.559
f2surg = 0.027HR + 0.091RR – 0.155SaO$_2$ + 10.613
f3med = 0.029HR + 0.169RR – 5.930
f3resp = 0.034HR + 0.084RR – 5.373

Abbreviations: f – function (numbers 1 – function used all five parameters, 2 – function used three parameters and 3 – function used two parameters); HR – heart rate; med – medical; resp – respiratory; RR – respiratory rate; SaO$_2$ – oxygen saturation; surg – surgical; sysBP – systolic blood pressure; temp – temperature

TABLE 4 Sensitivity and specificity for various functions and discrete early warning scores for predicting ICU admission

Warning system	Group	Sens	Spec	Cut	AUC	PPV %	NPV %
ACUTE MEDICAL ADMISSIONS							
f2surg	Medical	0.93	0.82	0.91	0.92	95.8	74.1
f1med	Medical	0.92	0.82	0.86	0.93	95.7	70.2
f2med	Medical	0.89	0.86	0.65	0.91	96.6	63.6
f3med	Medical	0.85	0.78	0.26	0.86	94.5	53.4
EWS	Medical	0.83	0.70	2.0	0.81	91.4	53.1
PART	Medical	0.79	0.77	2.0	0.84	92.8	49.0
MEWS	Medical	0.83	0.79	2.0	0.87	93.6	55.8
SEWS	Medical	0.95	0.77	3.0	0.95	94.0	79.7
ACUTE RESPIRATORY ADMISSIONS							
f2surg	Respiratory	0.83	0.77	1.32	0.86	89.0	69.1
f1resp	Respiratory	0.90	0.63	0.78	0.86	85.7	72.9
f2resp	Respiratory	0.87	0.73	0.58	0.85	87.8	75.0
f3resp	Respiratory	0.90	0.61	0.69	0.83	82.8	75.6
EWS	Respiratory	0.66	0.59	3.0	0.69	71.7	52.6
PART	Respiratory	0.83	0.53	3.0	0.73	73.5	66.7
MEWS	Respiratory	0.77	0.68	3.0	0.77	79.0	65.7
SEWS	Respiratory	0.84	0.70	3.0	0.87	81.8	73.8

Abbreviations: AUC – area under the receiver operator characteristic curve; cut – cut-point; EWS – Early Warning Score;[13] f – function; med – medical; MEWS – Modified Early Warning Score;[13] NPV – negative predictive value; PART – Patient At Risk Team Score;[14,15] PPV – positive predictive value; resp – respiratory; SEWS – Scottish Early Warning Score;[25] sens – sensitivity; spec – specificity; surg – surgical

variables in differentiating the groups. The differences were then summarised for each individual using the median values. Discriminant functions were derived using the differences for each individual, where delta represents the change in the parameters with time. Using function f2 (delta), and the cut-point corresponding to the top-left corner from the relevant ROC curve, an accurate prediction of whether patients ended up in the ICU or not was possible for only 52.9% of original grouped cases. The discriminant analysis was repeated for the medical and respiratory groups separately. For the resulting functions (f2med [delta] and f2resp [delta]), ROC curve analysis demonstrated AUC values of 0.56 and 0.53 respectively. Therefore this method was not pursued due to the poor discriminatory power. All functions are displayed in Table 3. The sensitivity and specificity for various functions and discrete early warning scores for predicting ICU admission are shown in Table 4.

Differences in parameters and functions with time

In previous sections we have been using the median value of each patient's measures to derive a discriminant function. The rules governing the scoring for the discrete EWS have been reported previously in the literature.[13,14,23,24] We compared the function values (f2) and the existing EWS measures for measurements within each hour up to the time of ICU admission in the ICU groups against all the values for the corresponding non ICU group. Figures 1 and 2 show the results of this analysis.

DISCUSSION

We believe that our cohorts are comparable with the population of patients treated in similar wards in Aberdeen and are likely to be similar to populations in other UK hospitals and thus have reasonable generalisability. We had adequate power to determine differences between groups based on the a priori power calculation. The large areas under the ROC curves suggested that high sensitivities and specificities could still be achieved in practice if the appropriate cut-points are used, despite the noisy nature of the data.

Blood pressure and temperature were excluded due to their poor discrimination power. For the remaining physiological parameters we identified cut-points based on balancing maximum sensitivity and specificity. In acute medical patients we identified cut-points of heart rate of 93 min⁻¹, which has an ROC with an AUC of 0.8 in the medical group and 105 min⁻¹ and 0.8 in the respiratory group; of 22 min⁻¹ for respiratory rate with an AUC of 0.84 in the medical group and 24 min⁻¹ and 0.77 in the respiratory group; and of 94% for oxygen saturation with an AUC of 0.73 in the medical group and 93% and 0.74 in the respiratory group. Interestingly, the cut-points for HR and RR with maximum sensitivity versus specificity play off are lower than are commonly used in clinical practice or in many existing scoring systems. A cut-point for SaO_2 of 94% is in line with existing
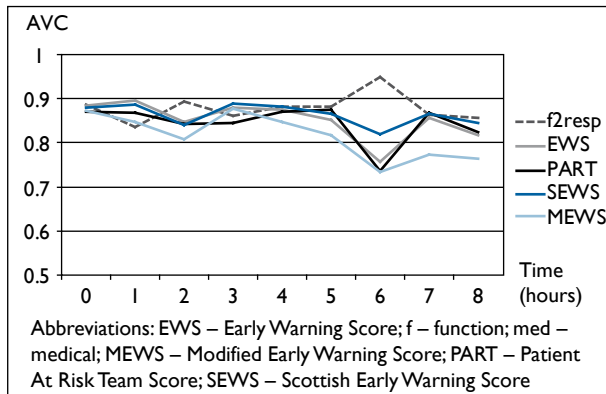
FIGURE 1 Trends in area under the curves for function 2 medical (f2med) and existing early warning scores at individual time points up to eight hours before ICU admission (time 0). Data from the medical-non ICU group at all time points were used as a comparator.

FIGURE 2 Trends in area under the curves for function 2 respiratory (f2resp) and existing early warning scores at individual time points up to eight hours before ICU admission (time 0). Data from the respiratory-non ICU group at all time points were used as a comparator.

scores for medical patients. This cohort included all patients admitted to the respiratory unit during the study period and demonstrates that a cut-point of 93% does have reasonable sensitivity and specificity even in this group who often suffer from chronic pulmonary disease. We believe that lower cut-points are commonly used in clinical practice in this specific group. This work suggests that existing scoring systems may require alterations in their thresholds for scoring.

We also analysed trends in parameters with time. It is commonly quoted that trends are more important than actual values, although little literature exists to support this statement. We were interested to see that these delta values did not have good discriminant power, with AUC values of only 0.53 to 0.56. We concluded that developing functions in this way was not useful. Although this does not prove that trends with time are not important clinically, it does further support the concept that absolute values for physiological values may be more important and more useful than trends.

Existing early warning scoring systems were analysed in this cohort. The optimum cut-point for the MEWS score is 2 for acute medical environments and 3 in acute respiratory medicine environments. For a given scoring system, if a higher cut-point were used then the sensitivity would decrease and the specificity would increase. For instance, in medical patients the MEWS score has an optimal cut-point of 2 and a sensitivity of 79% and specificity of 83%, whereas if we choose a cut-point of 4 then the sensitivity falls to 36% and specificity rises to 98%. This will have the effect of reducing the number of calls generated but at the expense of failing to detect patients at the early stage of critical illness. Since early detection is the primary aim of these scores, increasing cut-points to manage workload should be avoided. With the optimal cut-point, all of these existing scores are able to identify deteriorating patients many hours before major deterioration.
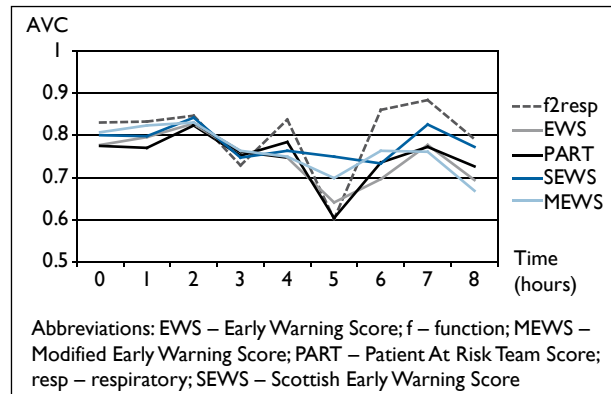
We have been able to demonstrate that a combination of three and five physiological parameter functions are able to differentiate accurately between groups. However, the contribution of temperature and systolic blood pressure were insignificant in the five-parameter function (f1) and thus the two three-parameter functions f2med and f2resp are to be preferred due to their comparative simplicity. We have also demonstrated that although f2med and f2resp perform well in differentiating between groups, f2surg (developed in a previous surgical group[23]) consistently performed at least as well in these medical cohorts but uses a different cut-point. All five of these physiological parameters are used in common early warning scoring systems in clinical practice.

This result may suggest that commonly used scoring systems that utilise these five parameters, including temperature and blood pressure, or scoring systems that give similar weighting to these parameters as to parameters such as HR, may have lower than optimal sensitivities and specificities due to the inappropriate weighting given to these parameters. Certainly, few have undergone prospective validation in patient cohorts. However, the sensitivities and specificities of these existing early warning scores seem to be good in this study. Our functions, and existing scores, are able to identify differences between groups for approximately eight hours before ICU admission with high accuracy. These data may suggest that the application of scoring systems allows earlier recognition of these patients before they deteriorate to the stage where they require ICU admission with the associated high morbidity and mortality.

In the MERIT study[26] it was found that only half of patients had the medical emergency teams (MET) criteria documented before an adverse event, suggesting that staff were not effectively using the scoring system. Poor use of scoring systems and significant inter- and intra-rater variability have been demonstrated in other studies.[27] They also found that 70% of all calls were not

CLINICAL

associated with an adverse event at all (poor specificity) despite increasing the number of calls generated. Further, the MET scoring system was only activated more than 15 minutes before the event in half of unplanned ICU admissions. Finally, the team was only called to 30% of these cases. This suggests low sensitivity of the MET scoring system as well as an inability to allow early recognition of acute deterioration.

Other studies also suggest low sensitivities of existing scores.[17,22] The PART score was 'not able to reliably predict which patients would be admitted to the intensive care unit'[14,15] and the MEWS score found an AUC of 0.67 for detecting deterioration in medical admissions and failed to demonstrate an improved outcome in this group.[28] Finally, the Rapid Emergency Medicine Score (REMS) has been shown to be able to predict hospital and longer-term mortality in non-surgical emergency admissions with an AUC of 0.85–0.92, but it does not predict in-hospital deterioration.[29–31] In a recent study it was found that 'sensitivities and positive predictive values were low', concluding that 'a wide variety of track and trigger systems were in use, with little evidence of reliability, validity and utility'.[22]

Limitations of our study include the fact that it studied small cohorts of patients in one teaching hospital in the UK and these data require testing in a prospective validation cohort. In some of the time-based analysis the number of data points was small and analysis was not performed. Our statistical methods required us to summarise data to produce one median measure per physiological parameter per patient. Although methodo-logically sound and appropriate, this could lead to an oversimplification of the richness of the data and therefore could allow inaccurate assumptions and conclusions from the data.

It is important to realise that exclusion of sysBP and temperature from this analysis may mean that they are not good early warning signs. This does not mean they are not important clinical signs in assessing a sick patient, but simply that they are not likely to be early signs.

Unlike in prospective studies or in clinical practice, we were able to identify a clear clinical outcome of ICU admission to use as our index event for analysis. It could be argued that this outcome is at the severe end of the spectrum and that in clinical practice we would want to identify deterioration before this event. We also demonstrated that these scores can identify these patients many hours before ICU admission.

Finally, this study looks at acute medical and respiratory admissions to hospital and therefore it could be argued that it cannot be generalised to other clinical areas. This is in part true, but our previous work has performed similar analysis in surgical patients and been able to demonstrate broadly similar results.[23] Indeed, it demonstrates that the same functions and early warning scoring systems can be used in different patient groups with only modification of the cut-points used.

Commonly measured physiological parameters such as HR, RR and SaO_2 have reasonable diagnostic accuracy in detecting acute deterioration in these medical patient groups. These parameters can be used by clinicians as early signs of clinical deterioration in the medical admissions context, using the cut-point of 93/min for HR, 22/min for RR and 94% for SaO_2 in medical admissions. However, blood pressure and temperature do not have good accuracy and should not be used by clinicians as early signs of clinical deterioration in the medical admissions unit context, although this does not mean they do not have clinical utility as late signs of deterioration.

Existing physiological scoring systems and our developed functions also have high diagnostic accuracy in these cohorts. These functions identify the deteriorating patient who requires intensive care management in the hours before routine referral through normal clinical pathways. These scores may be useful in the identification of the deteriorating patient in medical admissions units, but require further validation.

REFERENCES

1   The National Confidential Enquiry into Patient Outcome and Death (NCEPOD). *The report of the National Confidential Enquiry into Peri-operative Deaths 1996–97.* London: NCEPOD; 1997. Available from: http://www.ncepod.org.uk/1996_7.htm

2   Scottish Audit of Surgical Mortality (SASM). *Annual report 2004 data.* Paisley: SASM; Scotland 2004. Available from: http://www.sasm.org.uk/Reports/2004Report/SASM_Report_2004_data.pdf

3   The National Confidential Enquiry into Patient Outcome and Death (NCEPOD). *An acute problem?* London: NCEPOD; 2005. Available from: http://www.ncepod.org.uk/2005aap.htm

4   McQuillan P, Pilkington S, Allan A et al. Confidential enquiry into quality of care before admission to intensive care. *BMJ* 1998; 316:1853–8. Erratum in: *BMJ* 1998; 317:631.

5   Garrard C, Young D. Sub-optimal care of patients before admission to intensive care. *BMJ* 1998; 316:1841–2.

6   Subbe CP, Kruger M, Rutherford P et al. Validation of a modified early warning score in medical admissions. *QJM* 2001; 94:521–6. doi:10.1093/qjmed/94.10.521

7   Rowan KM, Kerr JH, Major E et al. The Intensive Care Society APACHE II study in Britain and Ireland – I: variations in case mix

in adult admissions to general intensive care units and impact on outcome. *BMJ* 1993; 307:972–6. doi:10.1136/bmj.307.6910.972

8 Rowan KM, Kerr JH, Major E, et al. The Intensive Care Society APACHE II study in Britain and Ireland – II: outcome comparisons of intensive care units after adjustment for case mix by the American APACHE II method. *BMJ* 1993; 307:977–81. doi:10.1136/bmj.307.6910.977

9 Bion J. Rationing intensive care. *BMJ* 1995; 310:682–3.

10 Bion J. Outcomes in intensive care. *BMJ* 1993; 307:953–4. doi:10.1136/bmj.307.6910.953

11 Schein RM, Hazday N, Pena M et al. Clinical antecedents to in-hospital cardiac arrest. *Chest* 1990; 98:1388–92. doi:10.1378/chest.98.6.1388

12 Lee A, Bishop G, Hillman KM et al. The medical emergency team. *Anaesth Intens Care* 1995; 23:183–6.

13 Morgan RJ, Williams F, Wright MM. An early warning scoring system for detecting developing critical illness. *Clin Intens Care* 1997; 8:100.

14 Goldhill DR, McNarry AF, Mandersloot G et al. A physiologically-based early warning score for ward patients: the association between score and outcome. *Anaesthesia* 2005; 60:547–53. doi:10.1111/j.1365-2044.2005.04186.x

15 Goldhill DR, McNarry AF. Physiological abnormalities in early warning scores are related to mortality in adult inpatients. *Br J Anaesth* 2004; 92:882–4. doi:10.1093/bja/aeh113

16 Cuthbertson BH. Outreach critical care – cash for no questions? *Br J Anaesth* 2003; 90:4–6. doi:10.1093/bja/90.1.5

17 Rogers J, Fuller HD. Use of the APACHE II score to predict individual patient survival rate. *Crit Care Med* 1994; 22:1402–5. doi:10.1097/00003246-199409000-00008

18 Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. *Br J Surg* 1991; 78:356–60. doi:10.1002/bjs.1800780327

19 Scottish Executive. *Better critical care – report of a short-life working group on ICU and HDU issues.* Edinburgh: Scottish Executive; 2000.

20 Department of Health. *Comprehensive Critical Care. A review of adult critical care services.* London: Department of Health; 2000.

21 National Institute for Health and Clinical Excellence (NICE). *Management of the acutely ill patients in hospital.* London: NICE; 2007. Available from: http://guidance.nice.org.uk/CG50

22 Gao H, McDonnell A, Harrison DA et al. Systematic review and evaluation of physiological track and trigger warning systems for identifying at-risk patients on the ward. *Intens Care Med* 2007; 33:667–79. doi:10.1007/s00134-007-0532-3

23 Cuthbertson BH, Boroujerdi M, McKie L et al. Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? *Crit Care Med* 2007; 35:402–9. doi:10.1097/01.CCM.0000254826.10520.87

24 Stenhouse C, Coates S, Tivey M et al. 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* 2000; 84:663.

25 Patterson R, MacLeod DC, Thetford D et al. Prediction of in-hospital mortality and length of stay using an early warning scoring system. *Clin Med* 2006; 6:281–4.

26 Hillman K, Chen J, Cretikos M et al. Introduction of the medical emergency team (MET) system: a cluster-randomised controlled trial. *Lancet* 2005; 365:2091–7. doi:10.1016/S0140-6736(05)66733-5

27 Subbe CP, Gao H, Harrison DA. Reproducibility of physiological track-and-trigger warning systems for identifying at-risk patients on the ward. *Intens Care Med* 2007; 33:619–24. doi:10.1007/s00134-006-0516-8

28 Subbe CP, Davies RG, Williams E et al. Effect of introducing the Modified Early Warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions. *Anaesthesia* 2003; 58:775–803. doi:10.1046/j.1365-2044.2003.03258.x

29 Olsson T, Terent A, Lind L. Rapid emergency medicine score can predict long-term mortality in non-surgical emergency department patients. *Acad Emerg Med* 2004; 11:1008–13. doi:10.1111/j.1553-2712.2004.tb00667.x

30 Olsson T, Terent A, Lind L. Rapid emergency medicine score: a new prognostic tool for in-hospital mortality in non-surgical emergency department patients. *J Intern Med* 2004; 255:579–87. doi:10.1111/j.1365-2796.2004.01321.x

31 Olsson T, Lind L. Comparison of the rapid emergency medicine score and APACHE II in non-surgical emergency department patients. *Acad Emerg Med* 2003; 10:1040–8. doi:10.1111/j.1553-2712.2003.tb00572.x