Title page

Title:- Early identification of those acute medical admissions that will require critical care

Corresponding author:-

Dr Hugh Alisdair Carmichael

Vale of Leven Hospital, Main Street, Alexandria, Dunbartonshire, G83 0UA, Scotland, UK

Preferred email address:- hughcarmichael@hotmail.com

(Hospital email address Hugh.Carmichael@vol.scot.nhs.uk)

Home phone number:- 0141 942 0161; Hospital Secretary Tel. no.:- 01389 817239; Fax no.:- 01389 817319

Co-authors:-

Dr Edwin Robertson

Alexandria Medical Centre, 46-62 Bank Street, Alexandria, Dunbartonshire, G83 0LS, Scotland, UK.

Dr Jill Austin

Anaesthetics Department, Aberdeen Royal Infirmary, Foresterhill, Aberdeen, AB25 2ZN, Scotland, UK.

Dr Douglas McCruden

Vale of Leven Hospital, Main Street, Alexandria, Dunbartonshire, G83 0UA, Scotland, UK

Dr Claudia-Martina Messow

Robertson Centre for Biostatistics, Boyd Orr Building, University of Glasgow, G12 8QQ, Scotland, UK

Mr Philip R Belcher

Division of Cardiovascular Sciences, University of Glasgow, G12 8QQ, Scotland, UK

Keywords and keyphrases:- Early warning scores, triaging scores, acute medical admissions, intensive care, high dependency care.

Conflicts of interest We declare that there are no conflicts of interest.

Funding The former NHS Argyll and Clyde Health Board funded the audit and analysis costs.

Word count:- 2493

Early identification of those acute medical admissions that will require critical care

Hugh A Carmichael, Edwin Robertson, Jill Austin, Douglas C McCruden, Claudia- Martina Messow, Philip R Belcher

ABSTRACT

Background Removal of the Intensive Care Unit (ICU) at the Vale of Leven Hospital mandated the identification and transfer out of those acute medical admissions with a high risk of requiring ICU.

Aims To develop triaging tools that identified such patients and compare them with other scoring systems.

Methods A retrospective analysis of physiological and arterial gas measurements from 1976 acute medical admissions produced PREEMPT-1. A simpler one for ambulance use (PREAMBLE-1) was produced by the addition of peripheral oxygen saturation to a modification of MEWS (Modified Early Warning Score). Prospective application of these tools produced a larger database of 4447 acute admissions from which logistic regression models produced PREEMPT-2 and PREAMBLE-2 which were then compared to the original systems and seven other early warning scoring systems.

Results In patients with arterial gases the area under the Receiver Operator Characteristic curve (ROC AUC) was significantly higher in PREEMPT-2 (89·1%) and PREAMBLE-2 (84.4%) than all other scoring systems. Similarly, in all patients, it was higher in PREAMBLE-2 (92·4%) than PREAMBLE-1 (88·1%) and the other scoring systems.

Conclusions Risk of requiring ICU can be more accurately predicted using PREEMPT-2 and PREAMBLE-2, as described here, than by other early warning scoring systems developed over recent years.

Introduction

In October 2003, acute surgical and Emergency Department (ED) services at the Vale of Leven Hospital were transferred to the Royal Alexandra Hospital (RAH) in Paisley. “Stand-alone” acute medical services continued with on-site anaesthetic cover, and the ED was replaced by a Medical Assessment Unit (MAU). The planned removal of the remaining Intensive Care Unit (ICU) bed stimulated the desire to develop tools to help identify patients at greatest risk of requiring ICU care, to allow transfer elsewhere as early as possible. A review of the literature in late 2004 showed only two of the earliest versions of the scoring systems assessed below, and none that had risk of requiring ICU as the primary target. The retrospective audit described below was therefore undertaken to develop suitable triaging tools which are described here. The **PRE**-critical **E**mergency **M**edical **P**atient **T**riage (**PREEMPT**) tool was designed to be used in the hospital based on routine physiological and arterial blood gas measurements readily available in the emergency department. The **PRE**-**A**dmission **M**edical **B**lue-**L**ight **E**mergency (**PREAMBLE)** tool was designed for ambulance use, only using routinely collected physiological measurements, to allow a rapid risk assessment and decision regarding by-pass to another hospital. As described below, the subsequent implementation of these scoring systems led to a larger, more complete, database that allowed the further development of these scoring systems.

Leading up to and during the timeframe of this project a number of other scoring systems have been developed aimed at identifying the "at-risk" patient and including those discussed below[1-10], These arose out of the increasing concern about delays in identifying the deteriorating acute medical admission and the resultant poorer prognosis as outlined in several national reports that led to national recommendations about the use of early warning scores[11-13]. The results of our work therefore closely mirror these initiatives.

Methods

(i) Retrospective study:- This included 1843 non-ICU medical admissions over the six months from 1st December 2004 to 31st May 2005 and, to allow sufficient numbers of ICU patients (defined as requiring ventilation and APACHE-II monitoring) for comparison, 133 out of the 205 ICU medical admissions over a five year period from 1st January 2000; 72 of the 205 were excluded because either the case notes had been destroyed, no pre-ICU data were available, or the patient had been brought in from another ICU. The routine physiological variables recorded for each admission were age (years), sex, pulse rate (beats/minute), respiratory rate (breaths/minute), systolic and diastolic blood pressure (SBP and DBP in mmHg), temperature (°C), conscious level (AVPU:– A= alert, V= responds to verbal, P= responds to pain, U= unresponsive), peripheral oxygen saturation (SpO_2) on air (%), the composite Modified Early Warning Score (MEWS)[1], and, where measured on clinical grounds, arterial blood gases (H^+, nmol/l; $PaCO_2$, kPa). For those patients admitted to ICU more than 24 hours after MAU presentation (15 patients, 11% of ICU admissions), only data that had been available during the 24 hours before ICU admission was used in the analysis. Variables were compared between ICU and non-ICU patients using, as appropriate, t-test (T), Wilcoxon-Mann-Whitney test (W), Fisher's Exact test (F), and Chi-square test (C) to identify potential predictors for ITU admission. A subset of variables was selected for multivariate analysis.

(ii) Prospective study:- The much larger, more complete prospective database consisted of 4314 acute admissions between 30th January 2006 and 31st March 2007. The same variables were recorded as for the retrospective study. In MAU, a high MEWS ($\geq$ 4 in total, or 3 in any one category), hypoxia (SpO_2 < 95%), or clinical concern, triggered arterial gas measurement and generation of a PREEMPT-1 score which, along with clinical judgement, was used to select a small group of patients, judged to be at risk of respiratory failure, for urgent transfer to a hospital with ICU facilities. All patient journeys were tracked to ascertain who required ICU admission with invasive ventilation (37) and High Dependency Unit (HDU) admission (145). To allow a more meaningful analysis a combined dataset was created from all prospective study patients plus the 133 ICU patients, that required similar intervention, from the retrospective study, giving a total of 170 ICU patients for analysis, of whom 169 had arterial gas results available during the 24 hours prior to admission to ICU.

Using this combined dataset, logistic regression models were used to develop new versions of the triaging tools. A bootstrap variable selection method was used[14]. For continuous variables, a range of transformations was considered: x,

$1/(x^2)$, $1/x$, $1/(x^{1/2})$, $\ln(x)$, $x^{1/2}$, x^2, and x^3, to allow for possible nonlinear relationships. Including these in the full set of possible predictor variables, a forward stepwise procedure was applied to 500 bootstrap datasets. A final stepwise procedure was applied to the subset of predictor variables that appeared most often in the final models for these bootstrap datasets. The resulting logistic regression models led to PREEMPT-2 and PREAMBLE-2.

The combined dataset allowed comparison of the original and new versions of both PREEMPT and PREAMBLE. Receiver Operator Characteristic (ROC) curves were created and the area under curve (ROC AUC) estimated. Alternative cut-offs for the new scores were chosen to give sensitivities (a) of 90%, and (b) equal to those achieved by PREEMPT-1 and PREAMBLE-1. Performance statistics are presented with bootstrap 95% confidence limits based on 1000 bootstrap samples. The two PREAMBLE scores are reported within the subset of patients with arterial gases, as well as the full combined dataset. All four scoring systems were also assessed regarding predicting need for HDU care, in the complete dataset minus ICU admissions, in terms of ROC AUC performance.

The combined dataset was also used to see how effective a wide range of other early warning scoring systems were at detecting risk of requiring ICU or HDU compared to the PREEMPT and PREAMBLE systems. Most of these systems have been developed over the same timeframe, and use very similar variables (with the exception of arterial gases), as PREEMPT and PREAMBLE. They included MEWS (Modified Early Warning Score)[3], REMS (Rapid Emergency Medicine Score)[4], PARS (Patient At Risk Score)[5], MREMS (Modified Rapid Emergency Medicine Score)[7], SEWS (Standardised Early Warning Score)[8], WPSS (Worthing Physiological Scoring System)[9], and AEWS(surg) (Aberdeen Early Warning Score(surgical function))[10].

Results

(i) Retrospective study:- Table 1 shows data collected for the retrospective study.

PREEMPT-1:- Although ICU patients were younger than non-ICU patients it was felt, perhaps inappropriately, that age should not be used in the initial triaging tool. The sex distribution was similar in all groups. As well as MEWS, four of its components (pulse, respiration rate, SBP and AVPU), were associated with ICU admission. MEWS was preferred over these individual variables to allow as simple a model as possible. Both H^+ and $PaCO_2$ were higher in ICU patients ($p<0.001$) and were included in the model. SpO_2 values were lower in ICU patients than all non-ICU patients ($p<0.001$) though not for the subgroup with arterial gas measurements ($p=0.66$). They nonetheless appeared to make a contribution as, after a process of examining different coefficients, the resulting model was:- **PREEMPT-1 = 4 x MEWS+ 4 x $PaCO_2$ + H^+ – ½ x SpO_2**. A score of >25 had a sensitivity of 92·4% and a specificity of 49·8% for detecting risk of requiring ICU when applied to the 15% of patients with arterial gas results.

PREAMBLE-1:- As stated above, subsequent admission to ICU was associated with abnormal values of the 4 MEWS components of pulse, respiration rate, SBP and AVPU as well as high MEWS and low SpO_2 values. Using the four MEWS components, a modified version of MEWS was agreed following a multidisciplinary round-table discussion of their relative role in identifying risk of requiring ICU, as reflected in their performance in the retrospective dataset. Aimed at being used in the ambulance setting, this was simpler and easier to apply than MEWS and, in the retrospective audit, appeared at least equivalent to MEWS, with a score of 1 having a sensitivity of 83.9% and specificity of 59.0% compared to a score of 2 for MEWS giving 81.5% and 59.8% respectively. This was further improved subsequently by adding SpO_2 and became known as PREAMBLE-1. Table 2 shows how this score is calculated. A score of >1 had a sensitivity of 92·0% and a specificity of 58·8% for detecting risk of requiring ICU in the whole dataset. Results were 92.0% and 29.1% respectively when applied to the 15% of patients with arterial gas results, significantly poorer than PREEMPT-1.

(ii) Prospective study:-Table 3 summarises the combined dataset. Patterns of association were the same as for the retrospective dataset except, in patients with arterial gases, pulse rates were similar and SpO_2 levels higher in the non-ICU compared to ICU patients. Table 4 shows the formulae of the logistic regression models derived for calculating PREEMPT-2 and PREAMBLE-2.

Table 5 shows that, in those with arterial gases, PREEMPT-2 is significantly better than PREEMPT-1 regarding ROC AUC (89.1% versus 79.8%) and, when comparing a similar sensitivity of 92.9%, specificity (61.1% versus 46.9%) results for detecting ICU need. Also this subgroup shows PREEMPT-2 has significantly better results compared to PREAMBLE-2, as does PREEMPT-1 compared to PREAMBLE-1. PREAMBLE-2 appears to be as powerful as PREEMPT-1. In all patients, and at a sensitivity of 91·8%, PREAMBLE-2 has a mildly, but still significantly, improved specificity compared to PREAMBLE-1 (69·6% versus 66.7%), although the 95% CIs for ROC AUC just overlap. Figure I illustrates the associated ROC plots.

Figure 1 Receiver Operator Characteristic (ROC) curves for the PREEMPT and PREAMBLE scoring tools examined using the combined dataset and looking at risk of requiring ICU. The left-hand set looks at the 868 patients with arterial gases comparing PREEMPT-1 and PREEMPT-2. The right-hand set looks at all 4447 patients comparing PREAMBLE-1 and PREAMBLE-2.

When compared to the 7 other previously reported scoring systems, PREEMPT and PREAMBLE appear better at identifying risk of requiring ICU care, as Table 6 demonstrates. This is particularly so with PREEMPT-2 and PREAMBLE-2 in the most at-risk group, i.e. those with arterial gases, with no overlap between the 95% CIs for ROC AUC between these and all other scoring systems. When looking at all patients, although PREAMBLE-2 has a significantly greater ROC AUC than most other scoring systems, there is a slight overlap in the 95% CI between it and PARS and SEWS.

Table 6 also shows that the ability of all scoring systems to identify HDU patients is much poorer than their ability to identify ICU patients, with up to 20% comparative reduction in ROC AUC levels. Some scoring systems, particularly when applied to the subgroup with arterial gases, are almost no better than guessing. PREEMPT-2 and PREAMBLE-2 have highest ROC AUC values but 95% CIs often overlap with others.

DISCUSSION

Over recent years, scoring systems have been developed in a number of specific clinical situations to identify poor prognosis[15-19]. Delays in recognising deteriorating acute medical patients led to attempts to introduce Early Warning Scores (EWS and MEWS) and the development of outreach critical care teams to detect and manage these patients[1-3, 20-22]. The reports from the National Confidential Enquiry into Patient Outcome and Death (NCEPOD) and National Institute for Health and Clinical Excellence (NICE) describe these scoring systems in detail[11-12]. More recently, and during the period of development of our own triaging tools, other such scoring systems have been reported including several of those analysed in this paper using our combined dataset[3-5, 7-10]. The systems so far described have concentrated mainly, but not always exclusively, on risk of dying as a surrogate for requiring critical care. The exception to this is the Aberdeen Early Warning Score (AEWS)[10] which, like our own, was designed to identify risk of requiring ICU care. In an attempt to improve specificity some systems have become complex and could be unwieldy in the emergency situation[6].

In this paper, we describe a process through which the PREEMPT and PREAMBLE triaging tools have been developed aimed at detecting those emergency admissions at risk of requiring ICU care. They have good diagnostic performance despite coming from routine clinical databases, and reflecting “the real world” of data collection with the potential for more errors than in formal studies. As such they could be used in a wide range of acute clinical settings. PREAMBLE-2 could be generated by the referring GP or en route by ambulance staff to assist decisions regarding urgency and the need to alert ED. It may also prove to have a role in selecting those requiring arterial blood gas measurements on arrival at ED which would thus generate a PREEMPT-2 score. By requiring arterial gas results availability, PREEMPT-2 is aimed at a small, more at-risk, group of admissions. In our experience about 15 – 17% of acute general medical admissions are felt to be unwell enough to require arterial gas analysis. The score generated allows the selection of a subgroup (in our experience about 1/3rd of the 15%–17% i.e. 5%–6% of all admissions) that contains about 90% of those that will require ICU after admission. Its easy application on arrival, if necessary by non-medical staff, would allow rapid early triaging of this small group that could be more intensively monitored. We have found, but not reported here, that sequential application of these tools may allow nearly every patient to be identified before ICU. In some, their deterioration is clearly obvious but in others this is much less so.

The earlier versions have been, and continue to be, used in the triaging of acute medical emergencies arising within the catchment of Vale of Leven Hospital, both in the ambulances (PREAMBLE-1) and in the MAU (PREEMPT-1). The apparent superiority of PREEMPT-2 and PREAMBLE-2 to the earlier versions and to the other scoring tools mentioned here would favour their use, but their complexity would require pre-programmed desktop or hand-held computers/calculators. This would however improve accuracy and increase speed of calculation as well as allowing collection of data to download on to a dataset so allowing audit and further development of the tools as suggested below. Electronic recording of data, and automatic generation of scores, is in the process of happening in the Vale of Leven Hospital and, given the increasing trend to electronic rather than paper data collection in the NHS as a whole, this is seen as an appropriate way for such tools to develop in future[23].

From our experience of their use, the PREAMBLE and PREEMPT tools for identifying risk of requiring ICU, as described in this paper, could significantly reduce the morbidity and mortality of deteriorating acute medical patients by their earlier identification and thus more intensive monitoring and management, as has been advocated in the NCEPOD, NICE and Royal College of Physicians of London reports[11-13]. The other early warning scoring tools that have been developed in recent years have gone some way to achieving this goal but, as seen here, additional prognostic markers, as in PREEMPT, may be required to improve their accuracy . Avenues for further development include assessment of additions or alternatives to these prognostic markers (e.g. venous rather than arterial gas measurements), prospective modelling studies of the implications of using them in isolation, in combination, or sequentially, and cluster randomised studies to evaluate their impact on patient outcomes. This may help achieve a consensus on the best way forward for early identification of the at-risk acute hospital admission.

Acknowledgements

This work could not have been carried out without the enthusiastic commitment of a wide range of healthcare professionals in Clyde Division of NHS Greater Glasgow and Clyde, especially the Vale of Leven Hospital and Royal Alexandra Hospital, and is presented on behalf of the Lomond Integrated Care Implementation Group made up of physicians, general practitioners, anaesthetists, A&E consultants, nurse managers and managers. Funding of the retrospective and prospective data gathering and analysis was provided by the former NHS Argyll and Clyde Health Board.

References

1 Morgan RJM, Williams F, Wright MM. An early warning score for the early detection of patients with impending illness. *Clin Intens Care* 1997; 8: 100.

1 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 Anaesthesia* 2000;84: 663P.

3 Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning Score in medical admissions. *QJM* 2001; 94: 521–6.

4 Olsson T, Terent A, Lind L. Rapid Emergency Medicine score: a new prognostic tool for in-hospital mortality in nonsurgical emergency department patients. *J Intern Med* 2004; 255: 579–87.

5 Goldhill DR, McNarry AF, Mandersloot G, McGinley A. A physiologically-based early warning score for ward patients: the association between score and outcome. *Anaesthesia* 2005; 60(6): 547–53.

6 Kellett J, Deane B. The Simple Clinical Score predicts mortality for 30 days after admission to an acute medical unit. *Q J Med* 2006: 99: 771–81.

7 Goodacre S, Turner J, Nicholl J. Prediction of mortality among emergency medical admissions. *Emerg Med J* 2006; 23: 372–5.

8 Paterson R, MacLeod DC, Thetford D et al. Prediction of in-hospital mortality and length of stay using an early warning scoring system; clinical audit. *Clin Med* 2006; 6(3): 281-4.

9 Duckitt RW, Buxton-Thomas R, Walker J et al. Worthing physiological scoring system: derivation and validation of a physiological early-warning system for medical admissions. An observational, population-based single-centre study. *Br J Anaesthesia* 2007; 98(6): 769–94.

10 Cuthbertson BH, Boroujerdi M, Prescott G. The use of combined physiological parameters in the early recognition of the deteriorating acute medical patient. *J R Coll Physicians Edinb* 2010; 40: 19 25.

11 National Confidential Enquiry into Patient Outcome and Death (NCEPOD). An Acute Problem? www.ncepod.org.uk/2005report May 2005. Accessed July 2005.

12 National Institute for Health and Clinical Excellence. NICE clinical guideline 50 - Acutely ill patients in hospital: recognition of and response to acute illness in adults in hospital. www.nice.org.uk July 2007. Accessed October 2007.

13 Royal College of Physicians. Acute medical care. The right person, in the right setting – first time. Report of the Acute Medical Task Force. London: RCP, 2007. ISBN 978–1–86016–321–0.

14 Sauerbrei W, Schumacher M. A bootstrap resampling procedure for model building: Application to the Cox regression model. *Statistics in Medicine* 1992; 11: 2093–109.

15 Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet* 1974; 2(7872): 81–4.

16 Knaus WA, Draper EA, Wagner DP, Zimmerman JE. Apache II: a severity of disease classification system. *Crit Care Med* 1985; 12: 818–29.

17 Champion HR, Socco WJ, Copes WS, Gann DS, Gennarelli TA, Flanagan ME. A revision of the trauma score. *J Trauma* 1989; 29: 623–9.

18 Antman EM, Cohen M, Beninck PJ, McCabe CH, Horacek T, Papuchis G, et al. The TIMI risk score for unstable angina/ non ST elevation myocardial infarction: a method for prognostication and for therapeutic decision-making. *JAMA* 2000; 284: 835–42.

19 Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 2002; 58: 377–82

20 Lee A, Bishop G, Hillman KM, Daffurn K. The medical emergency team. *Anaesthesia Intens Care* 1995; 45: 183–6.

21 Goldhiil DR, Worthington L, Mulcahy A, Tarling M, Sumner A. The patient at risk team: identifying and managing seriously ill ward patients. *Anaesthesia* 1999: 54: 853–60.

22 Comprehensive Critical Care: A Review of Adult Critical Care Services. www.doh.gov.uk May 2000. Accessed January 2008.

23 Smith GB, Prytherch D, Peet H et al. Automated calculation of early warning scores. Anaesthesia 2006; 61: 1009-10.

Variable		ICU Patients	Non-ICU Patients			
		(N_{max}=133)	All (N_{max}=1843)		With Arterial gases (N_{max}=275)	
Age (years)	Median [Range]	57.0 [17.0, 88.0]	66.0 [11.0, 99.0]	p<0.001 T	71.0 [15.0, 96.0]	p<0.001 W
Sex	N (%) Male N (%) Female	52 (40.6%) 76 (59.4%)	826 (45.8%) 977 (54.2%)	p=0.271 F	110 (41.5%) 155 (58.5%)	p=1.000 F
Temperature (°C)	Median [Range]	36.2 [25.0, 41.3]	36.4 [26.7, 39.9]	p=0.399 W	36.5 [26.7, 39.8]	p=0.196 W
Respiration Rate (/min)	Median [Range]	22.0 [0.0, 60.0]	19.0 [5.0, 69.0]	p=0.010 W	24.0 [9.0, 51.0]	p=0.048 T
Pulse (bpm)	Median [Range]	110.0 [0.0, 193.0]	88.0 [14.0, 256.0]	p<0.001 T	101.0 [39.0, 198.0]	p=0.002 T
SBP (mmHg)	Median [Range]	117.0 [0.0, 262.0]	140.0 [43.0, 279.0]	p<0.001 T	141.0 [43.0, 253.0]	p<0.001 T
DBP (mmHg)	Median [Range]	68.0 [0.0, 161.0]	75.0 [14.0, 163.0]	p<0.001 T	73.0 [29.0, 152.0]	p=0.003 T
AVPU	N (%) Alert N (%) to Voice N (%) to Pain N (%) Unresponsive	56 (42.4%) 26 (19.7%) 11 (8.3%) 39 (29.5%)	1443 (87.5%) 162 (9.8%) 17 (1.0%) 28 (1.7%)	p<0.001 C	213 (84.2%) 28 (11.1%) 4 (1.6%) 8 (3.2%)	p<0.001 C
MEWS	Median [Range]	5.0 [0.0, 11.0]	2.0 [0.0, 11.0]	p<0.001 W	3.0 [0.0, 10.0]	p<0.001 T
SpO_2 (%)	Median [Range]	93.0 [39.0, 100.0]	97.0 [53.0, 100.0]	p<0.001 W	94.0 [67.0, 100.0]	p=0.664 W
$PaCO_2$ (kPa)	Median [Range]	6.1 [1.4, 21.7]	-	-	4.9 [0.8, 15.0]	p<0.001 W
H^+ (nmol/l)	Median [Range]	50.1 [21.3, 215.0]	-	-	37.2 [18.2, 158.5]	p<0.001 W

Table I: Summary statistics from retrospective study. All 133 ICU patients had full data; the majority of the 1843 non-ICU patients had physiological and SpO_2 data available; almost all of the 275 non-ICU patients, for whom arterial gases were collected (15% of total), had physiological and SpO_2 data available. Arterial gases are summarised only for the subset of non-ICU patients in whom arterial gases were measured, not for all non-ICU patients. p-values are reported for comparisons between ICU patients and each group of non-ICU patients: T, groups compared with t-test; W, groups compared with Wilcoxon-Mann-Whitney test; F, groups compared with Fisher's Exact test; C, groups compared with Chi-square test.

Variable	Score						
	3	2	1	0	1	2	3
Heart rate	≤ 30	31 - 40	41 - 50	51 - 110	111 - 130	> 130	
Respiratory rate	≤ 8		9 - 11	12 - 20	21 - 30	31 - 40	> 40
AVPU				Alert	to Voice	to Pain	Unresponsive
Systolic BP	≤ 70	71 - 80	81 - 95	96 - 199	>199		
SpO2 on air	≤ 84	85 - 89	90 - 94	95 - 100			
Age*	> 40 - 60	> 60 - 75	0 - 40	> 75			

Table II: Scoring algorithm for PREAMBLE-1.
Each variable is assigned a score according to the above scoring table, and the item scores summed to give the total PREAMBLE-1 score.
*PREAMBLE-1 did not include age during pilot period. It has been added for subsequent analysis as shown in Table VI

Variable		ICU Patients (N$_{max}$=170)	Non-ICU Patients: All (N$_{max}$=4277)		Non-ICU Patients: With Arterial gases (N$_{max}$=699)	
Age (years)	Median [Range]	59·0 [17·0, 88·0]	67·0 [14·1, 101·2]	p<0·001 T	66·4 [15·2, 97·6]	p<0·001 T
Sex	N (%) Male N (%) Female	69 (40·6%) 101 (59·4%)	2013 (47·1%) 2264 (52·9%)	p=0·100 F	321 (45·9%) 378 (54·1%)	p=0·263 F
Temperature (°C)	Median [Range]	36·7 [25·0, 41·3]	36·5 [25·3, 43·6]	p=0·628 W	36·6 [25·5, 43·6]	p=0·074 W
Respiration Rate (/min)	Median [Range]	20·0 [3·0, 60·0]	18·0 [7·0, 58·0]	p<0·001 T	23·0 [8·0, 58·0]	p=0·074 T
Pulse (bpm)	Median [Range]	110·0 [35·0, 193·0]	88·0 [26·0, 222·0]	p<0·001 T	108·0 [33·0, 186·0]	p=0·238 T
SBP **(**mmHg)	Median [Range]	119·5 [49·0, 262·0]	136·0 [51·0, 269·0]	p<0·001 W	137·0 [51·0, 269·0]	p<0·001 W
DBP **(**mmHg)	Median [Range]	68·0 [28·0, 161·0]	73·0 [10·0, 229·0]	p=0·003 T	74·0 [19·0, 189·0]	p=0·004 T
AVPU	N (%) Alert N (%) to Voice N (%) to Pain N (%) Unresponsive	83 (48·8%) 34 (20·0%) 13 (7·6%) 40 (23·5%)	3960 (92·6%) 262 (6·1%) 34 (0·8%) 21 (0·5%)	p<0·001 C	632 (90·4%) 44 (6·3%) 13 (1·9%) 10 (1·4%)	p<0·001 C
MEWS	Median [Range]	5·0 [0·0, 11·0]	2·0 [0·0, 13·0]	p<0·001 T	4·0 [0·0, 13·0]	p<0·001 T
SpO_2 (%)	Median [Range]	93·0 [39·0, 100·0]	97·0 [60·0, 100·0]	p<0·001 W	94·0 [65·0, 100·0]	p=0·013 W
$PaCO_2$ (kPa)	Median [Range]	5·9 [1·4, 21·7]	-	-	4·8 [0·9, 13·2]	p<0·001 W
H^+ (nmol/l)	Median [Range]	49·0 [21·3, 215·0]	-	-	36·3 [17·0, 144·5]	p<0·001 W

Table III: Summary statistics from combined dataset of prospective study plus the 133 retrospective ICU patients. Summaries shown for all ICU patients, all non-ICU patients and for non-ICU patients for whom arterial gases were collected. Arterial gases are not summarised separately for all non-ICU patients. p-values are reported for comparisons between ICU patients and each group of non-ICU patients: T, groups compared with t-test; W, groups compared with Wilcoxon-Mann-Whitney test; F, groups compared with Fisher's Exact test; C, groups compared with Chi-square test.

PREEMPT-2 = $0{\cdot}39713 + f_1(x_1) + f_2(x_2) + f_3(x_3) + f_4(x_4) + f_5(x_5) + f_6(x_6) + f_7(x_7)$, where:		
x_1 = Age	$f_1(x_1) =$	$15{\cdot}03253 \times (x_1/100)^2 - 19{\cdot}37260 \times (x_1/100)^3$
x_2 = Resp rate	$f_2(x_2) =$	$123{\cdot}31922 \times (1/x_2)^2 + 19{\cdot}29794 \times (x_2/100)^3$
x_3 = SBP	$f_3(x_3) =$	$-0{\cdot}35839 \times x_3^{1/2}$
x_4 = AVPU	$f_4(x_4) =$	$1{\cdot}60142 \times (x_4 = \text{“V”}) + 1{\cdot}90037 \times (x_4 = \text{“P”}) + 3{\cdot}15080 \times (x_4 = \text{“U”})$
$x_5 = SpO_2$	$f_5(x_5) =$	$1{\cdot}90957 \times (100/x_5)^2$
$x_6 = PaCO_2$	$f_6(x_6) =$	$0{\cdot}76629 \times (x_6/10)^3$
$x_7 = H^+$	$f_7(x_7) =$	$-95{\cdot}22595 \times (1/x_7)$
PREAMBLE-2 = $2{\cdot}19253 + g_1(y_1) + g_2(y_2) + g_3(y_3) + g_4(y_4) + g_5(y_5) + g_6(y_6) + g_7(y_7)$, where:		
y_1 = Age	$g_1(y_1) =$	$0{\cdot}18183 \times (y_1/10)^2 - 0{\cdot}02295 \times (y_1/10)^3$
y_2 = Gender	$g_2(y_2) =$	$-0{\cdot}42532 \times (y_2 = \text{“Male”})$
y_3 = Resp Rate	$g_3(y_3) =$	$1{\cdot}20966 \times (10/y_3)^2 + 0{\cdot}02621 \times (y_3/10)^3$
y_4 = Pulse	$g_4(y_4) =$	$0{\cdot}36556 \times y_4^{1/2}$
y_5 = SBP	$g_5(y_5) =$	$1{\cdot}16637 \times (100/y_5)^2$
y_6 = AVPU	$g_6(y_6) =$	$1{\cdot}76085 \times (y_6 = \text{“V”}) + 3{\cdot}12845 \times (y_6 = \text{“P”}) + 4{\cdot}67077 \times (y_6 = \text{“U”})$
$y_7 = SpO_2$	$g_7(y_7) =$	$-0{\cdot}12880 \times (y_7)$

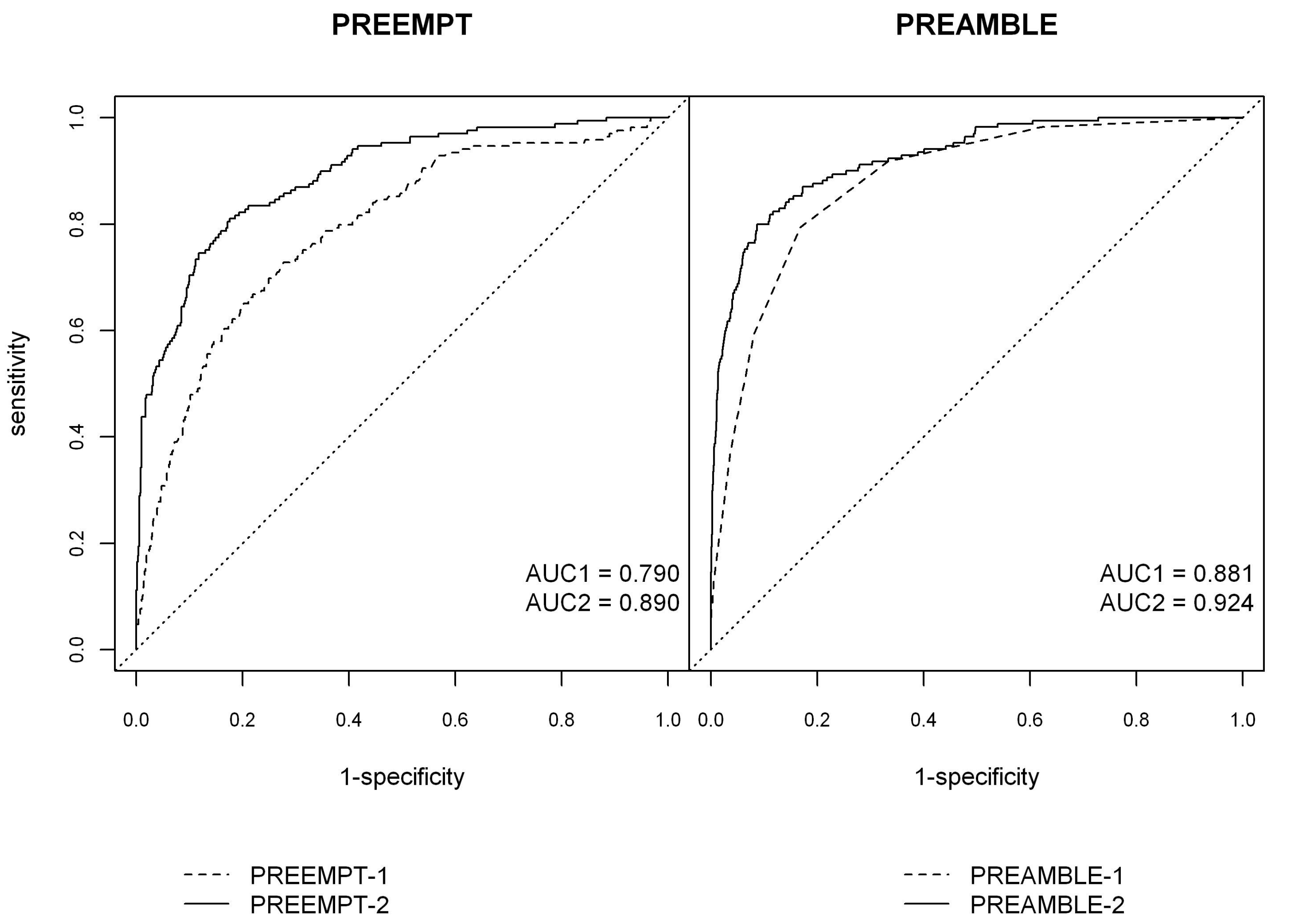
Table IV: Formulae for PREEMPT-2 and PREAMBLE-2

Triaging tool			ICU vs. All non-ICU	ICU vs. non-ICU with Arterial Gases
PREEMPT-1	AUC (%)		-	79·8% (76·0, 83·4)
	Score > 25	Sensitivity (%)	-	92·9% (89·1, 96·3)
		Specificity (%)	-	46·9% (43·3, 50·8)
PREEMPT-2	AUC (%)		-	89·1% (86·1, 91·6)
	Score > -2·30496	Sensitivity (%)	-	90·5% (86·0, 94·6)
		Specificity (%)	-	64·5% (61·0, 67·9)
	Score > -2·39709	Sensitivity (%)	-	92·9 %(89·4, 96·8)
		Specificity (%)	-	61·1% (56·8, 63·7)
PREAMBLE-1	AUC (%)		88·1% (85·5, 90·2)	71.0% (66.7, 75.3)
	Score >1	Sensitivity (%)	91·8% (87·3, 95·7)	91·7% (87·2, 95·4)
		Specificity (%)	66·7% (65·3, 68·0)	25·6% (23·3, 29·9)
PREAMBLE-2	AUC (%)		92·4% (90·1, 94·4)	84·4% (81·0, 87·6)
	Score > -3·93852	Sensitivity (%)	90·0% (85·3, 94·2)	89·9% (85·4, 94·2)
		Specificity (%)	74·6% (73·3, 75·7)	47·6% (44·1, 51·0)
	Score > -4·10600	Sensitivity (%)	91·8% (87·4, 95·6)	91·7% (87·2, 95·5)
		Specificity (%)	69·6% (68·2, 71·0)	41·2% (37·9, 44·9)

Table V: Performance of PREEMPT and PREAMBLE scores.
For each score, the Area Under Curve (AUC) for the ROC curve is reported with a 95% bootstrap CI. For PREEMPT-1 and PREAMBLE-1, the sensitivity and specificity (with 95% bootstrap CIs) are reported for the cut-offs originally chosen. For PREEMPT-2 and PREAMBLE-2, the sensitivity and specificity (with 95% bootstrap CIs) are reported for cut-offs chosen to give 90% sensitivity in the combined dataset, and for cut-offs chosen to equal the sensitivity of the corresponding original scores. For both PREAMBLE scores, performance statistics are also reported when applied using the subset of non-ITU patients with arterial gases.

Scoring system	(i) Prediction of requiring ICU in all patients (% ROC AUC)	(ii) Prediction of requiring HDU in all patients - minus ICU (% ROC AUC)	(iii) Prediction of requiring ICU in patients with gases (% ROC AUC)	(iv) Prediction of requiring HDU in patients with gases - minus ICU (% ROC AUC)
PREEMPT-1	-	-	79•8% (76.0, 83.6)	60.8% (53.9, 67.6)
PREEMPT-2*	-	-	89•1% (86•3, 91•9)	67.5% (61.6, 73.5)*
PREAMBLE-1 (with age)	89.4% (87.2, 91.7)	69.0% (64.3, 73.6)	75.8% (71.9, 79.8)	61.6% (54.8, 68.5)
PREAMBLE-1 (without age)	88.3% (86.0, 90.7)	70.3% (65.6, 75.0)	71.0% (66.7, 75.3)	60.4% (53.7, 67.0)
PREAMBLE-2**	92•4% (90•1, 94•3)	72.2% (67.8, 76.6)**	84•8% (81.3, 88.3)	67.8% (61.9, 73.6)
MEWS	82•3% (78•8, 85•6)	69.1% (64.6, 73.7)	64•6% (59•9, 69•3)	55.1% (48.0, 62.3)
REMS	70•4% (66•2, 74•5)	57.8% (53.1, 62.6)	60•5% (55•4, 64•7)	50.4% (43.3, 57.4)
PARS	89.2% (87.1, 91.2)	71.0% (66.5, 75.4)	71.9% (67.7, 76.1)	58.5% (51.6, 65.4)
MREMS	63.2% (58.6, 67.8)	54.7% (49.8, 59.6)	58.6% (53.8, 63.4)	51.8% (44.5, 59.0)
SEWS	88.8% (86.5, 91.1)	71.2% (66.8, 75.7)	71.1% (66.7, 75.4)	57.2% (50.5, 63.8)
WPSS	87.7% (85.5, 89.9)	69.6% (65.3, 73.9)	72.1% (67.9, 76.3)	56.9% (50.5, 63.3)
AEWS (surgical)	77.9% (74.0, 81.8)	69.1% (64.6, 73.7)	55.5% (50.4, 61.0)	54.9% (48.1, 61.8)

Table VI: Performance of all scoring systems in detecting ICU or HDU patients. Results are reported as % area under the Receiver Operator Curve (ROC AUC) with bootstrap 95% confidence interval (CI) for all scores comparing (i) the prediction of ICU need in whole combined dataset (includes 170 ICU patients) with (ii) the prediction of HDU need in the whole combined dataset minus the 170 ICU admissions (includes 145 HDU patients); and comparing (iii) the prediction of ICU need in those of combined dataset with arterial gases (includes 169 ICU patients) with (iv) the prediction of HDU need in those of combined dataset with arterial gases minus the 169 ICU patients (includes 74 HDU patients).
*PREEMPT-2 designed specifically to predict HDU need in those with arterial gases gives ROC AUC of 68.4%.
**PREAMBLE-2 designed specifically to predict HDU need gives a ROC AUC of 73.7%.


PREEMPT
PREAMBLE
sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
AUC1 = 0.790
AUC2 = 0.890
AUC1 = 0.881
AUC2 = 0.924
0.0
0.2
0.4
0.6
0.8
1.0
1-specificity
1-specificity
PREEMPT-1
PREEMPT-2
PREAMBLE-1
PREAMBLE-2