Editorial

An early warning on early warning systems!

Brian H Cuthbertson

Health Services Research Unit

Health Sciences Building

University of Aberdeen

Aberdeen

Scotland, UK

(44) (0)1224 552730

b.h.cuthbertson@abdn.ac.uk

It is clinically intuitive that physiological deterioration proceeds critical illness. There is also a growing evidence base to support this statement [1-5]. The use of early warning systems to facilitate the early recognition of the acutely deteriorating hospital patient is based on this premise. Indeed it is often presumed that these systems have a higher diagnostic accuracy to detect the deteriorating patient than is offered by clinical judgement alone. However, this may not always be the case.

The first such scoring system was the Medical Emergency Team (MET) Criteria developed in Liverpool, Australia [6]. This was followed by the development of a large numbers of scoring systems which have become widely used in clinical practice in the UK and around the world [7]. They utilise a variety of physiological parameters with a range of cut points and weightings to develop a score designed to trigger activation of a response system. Most parameters and cut points are based on clinical intuition rather than derived using rigorous derivation and validation methodologies. The appropriateness of the choice of parameters, the cut points as well as the accuracy of these scores was rarely questioned nor prospectively tested.

Recently work has appeared that casts significant doubt on the clinical utility of these scoring systems. The MERIT study failed to demonstrate benefit of the MET system partly due to the poor sensitivity and specificity of the MET calling criteria [MERIT]. The MET score was shown to have sensitivities and specificities

of less than 50% and did not allow early detection of the deteriorating patient. The majority of patients were not detected until less than 15 minutes before they suffered the primary outcome event. This lack of efficacy for the MET calling criteria led the Outreach movement in the UK rapidly distancing itself from it's founding father. Sadly, recent publications have also failed to demonstrate that the early warning systems used in the UK allow early detection of the deteriorating patient. When applied in an ideal mathematical way these scores have reasonable sensitivity and specificity [Cuthbertson]. However, when applied in clinical practice there are major concerns over their accuracy and a recent study concluded that there was "little evidence of reliability, validity and utility" [8]. They also concluded that the "sensitivity was poor, which might be due in part to the nature of the physiology monitored or to the choice of trigger threshold" [8]. Further work from the same group concluded "there was significant variation in the reproducibility of different track-and-trigger warning systems" with marked problems with both intra and inter-rater reliability [8]. Unless scoring systems have appropriate sensitivity and specificity and minimise errors associated with documentation and scoring they will fail to identify patients who needs assessment and intervention and will incorrectly trigger for patient in whom no intervention is required. There is a clear need to reassess the use of early warning scoring systems in the role in which they are currently utilised and there are a variety of possible ways to progress.

Firstly, we need to accurately identify which physiological parameters and cut points which actually act as early identifiers of deterioration. The ACADEMIA study demonstrated a range of clinical antecedents (similar to the MET criteria) that were very commonly present between 15 minutes and 24 hours before cardiac arrest, ICU admission or death [4]. Of course the presence of these antecedents for more than 15 minutes does not necessarily mean they have accuracy as early warning signs. The SOCCER study identified early and late predictors of critical illness [5]. It was interesting to note that the parameter cut points of the MET score were nearly all late signs of critical illness although their presence was strongly associated with a very poor outcome. The authors suggested that this"provides a possible explanation for the failure to demonstrate efficacy of a MET in some trials because current call criteria maybe too late in the progress of the patient's critical condition". This work suggests that a heart rate of 120/min is an early sign where as a heart rate of 140/min is late, that a respiratory rate of 30/min is an early sign where as 40/min is late and that an oxygen saturation of less than 96% is early and less than 90% is late. A recent case controlled study used different level of physiological abnormality for commonly used parameters and tested their diagnostic accuracy against the MET criteria. They found that the sensitivity and specificity of the MET calling criteria was low and that lower levels of physiological abnormality had the greatest sensitivity and specificities in an early warning scoring system [9]. Other

work in surgical patients suggests that only heart rate (with a cut point of 90/min, Area under the receiver operator curve (AUC) 0.74), respiratory rate (cut point 20/min, AUC 0.82) and oxygen saturation (cut point 96%, AUC 0.79) are early predictors of ICU admission and that blood pressure (AUC 0.51) and temperature (AUC 0.51) are not [10]. Similar findings have been made in other studies [11, 12, 12]. Presented in this edition of the journal are results of a study from Duckit and colleagues who also studied the accuracy of individual physiological parameters in a medical admissions setting [14]. Importantly, they utilized a derivation and validation cohort methodology. They found that heart rate greater than 101/min, respiratory rate greater than 19/min, oxygen saturation below 96% and blood pressure below 100mmHg were early predictors. They go onto develop and validate the Worthing Physiological Scoring System (see below). Together, this body of work suggests that many existing scoring systems use the incorrect parameters (temperature)[10,11] or fail to use all appropriate ones (oxygen saturation)[10-14], or indeed, use parameters with the wrong cut points (oxygen saturation)[5,10-13].

Next is the requirement to derive and validate a scoring system do novo from the physiological parameters that are known markers of early deterioration. The above study in surgical patients also attempted to derive a score de novo in a cohort of surgical patients [10]. A discriminate function utilising only heart rate, respiratory rate and oxygen saturation had a very high predictive accuracy for ICU admission (AUC 0.88) which was higher than existing early warning scores.

These functions still require prospective validation in a variety of clinical situations. The Worthing group, using logistic regression methods, derived and weighted parameters into a scoring system which they tested in a validation cohort [14]. This scoring system has reasonable diagnostic accuracy and is more accurate than another commonly used scoring system. Higher scores are also associated with higher mortalities and longer hospital length of stay. The authors are to be congratulated for taking such a robust approach and their study adds to the understanding in this field. It could be concluded that scoring systems of high accuracy can be developed de novo and may, after appropriate development and validation in varied patient groups, be able to replace existing scores that have low or poor diagnostic accuracy.

Thirdly, to overcome inaccuracies and miscalculations related to manual data collection we need to consider adoption of electronic data management and scoring systems [15]. Such systems should minimize intra and inter rater reliability issues and may also allow the use of potentially more accurate discriminate functions that are less amenable to manual calculation at the bedside [10]. Although not yet widely used in clinical practice, commercial systems are now available and their cost is not preclusive. Such systems have not been demonstrated to improve outcome and further work is needed in this area [15,16].

In conclusion, to date existing scoring systems and Outreach / MET interventions have failed to demonstrate improved outcomes in the studies which attempt to eliminate bias. Recent data has suggested that one of the reasons this may be the case is that currently used scoring systems lack sensitivity and specificity to allow accurate detection of early critical illness. If existing scores use parameters and cut points that are inappropriate then perhaps their use should be discontinued and scores replaced by systems that are derived using reliable methodologies with good quality validation data and have high diagnostic accuracy. Further work is required to derive and validate such scores in the full breadth of clinical practice and to explore the role of electronic data capture and scoring systems. Until this work is completed we should not allow existing early warning scoring system to replace high quality clinical assessment and judgment by appropriately skilled and experienced personnel.

Reference-

1. Schein RMH, Hazday N, Pena M, Ruben BH, Sprung CL. Clinical antecedents to in-hospital cardiac arrest. *Chest* 1990; **98**: 1388-1392.
2. Franklin C, Mathew J. Developing strategies to prevent in-hospital cardiac arrest: analysing responses of physicians and nurses in the hours before the event. *Crit Care Med* 1994; **22**: 244-7.
3. 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**: 547–553
4. Kause J, Smith G, Prytherch D, Parr M, Flabouris A, Hillman K, for the Intensive Care Society (UK) & Australian and New Zealand Intensive Care Society Clinical Trials Group ACADEMIA Study investigators. A comparison of Antecedents to Cardiac Arrests, Deaths and Emergency Intensive care Admissions in Australia and New Zealand and the United Kingdom—the ACADEMIA study. *Resuscitation* 2004; **62**: 275–282
5. Jacques T, Harrison GA, McLaws ML, Kilborn G. Signs of critical conditions and emergency responses (SOCCER): A model for predicting adverse events in the inpatient setting. *Resuscitation* 2006; **69**, 175-183
6. Lee A, Bishop G, Hillman KM, Daffurn K. The Medical Emergency Team. *Anaesth Intens Care* 1995; **23**: 183-186
7. 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**: 2463-2478
8. Gao H, McDonnell A, Harrison DA, Moore T, Adam S, Daly K *et al*. Systematic review and evaluation of physiological track and trigger warning systems for identifying at-risk patients on the ward. *Intensive Care Med* 2007; **33**: 667-79.

9. Cretikos M, Chen J, Hillman K, Bellomo R, Finfer S, Flabouris A, the MERIT study investigators. The objective medical emergency team activation criteria: A case—control study. *Resuscitation* 2007; **73**: 62—72
10. Cuthbertson BH, Bourjerdi M, Aucott L, Prescott G. The use of physiological parameters and scoring systems to recognise the deteriorating surgical patients. Accepted for publication. *Critical Care Med* 2007; 35: 402-409
11. Kellet J, Deane B. The Simple Clinical Score predicts mortality for 30 days after admission to an acute medical unit. *Q J Med* 2006; **99**: 771–781
12. Olsen et al. Comparison of the rapid emergency medicine score and APACHE II in non-surgical emergency department patients. *Acad Emerg Med* 2003; **10**: 1040-8
13. Olsson et al. Rapid Emergency Medicine score: a new prognostic tool for in-hospital mortality in non-surgical emergency department patients. *J Int Med* 2004; 255: 579-87.
14. Duckit, R, Buxton-Thomas R, Walker J, Cheek E, Bewick V, Venn RM, Forni LG. The Worthing Physiological Scoring System: Derivation and Validation of a physiological early warning system for medical admissions. Observational population based single-centre study. *Brit J Anaesth* (details to be entered)
15. Smith GB, Prytherch DR, Schmidt P, Featherstone PI, Knight D, Clements G, Mohammed MA. Hospital-wide physiological surveillance—A new approach to the early identification and management of the sick patient. *Resuscitation* 2006; **71**: 19-28
16. Watkinson PJ, Barber VS, Price JD, Hann A, Tarassenko L, Young JD. A randomised controlled trial of the effect of continuous electronic physiological monitoring on the adverse event rate in high risk medical and surgical patients. Anaesthesia 2006; **61**: 1031–1039