Appraisal of the literature

The majority of the studies evaluated reported some ability to predict mortality from physiological scores within the ED. The most effective scores were those derived from physiological observations in ED patients. This was particularly true of the Rapid Emergency Medicine Score (REMS) which performed well in the authors' initial validation studies (10, 11) and also in two other independent studies. (12, 13) This score is an extension of the Rapid Acute Physiology Score (RAPS) which was originally assessed in the ED by Olsson *et al* to determine its performance against the well recognised intensive care scoring system APACHE (acute physiology and chronic health evaluation) from which it was derived. (10) REMS is a 26-point scale which uses blood pressure, respiratory and pulse rate, oxygen saturations, age and GCS. Over the four studies which have investigated this score, over 20,000 patients from Sweden, the UK and US have now been studied. These include general non-surgical ED patients and also more specific groups, such as those with suspected infection and those with presumed life threatening illness. Using area under receiver operating characteristic (AUROC) curves to assess the performance of REMS, high values (0.74 - 0.91) for the prediction of death were reported over the four studies. (10-13) In the first study by Olsson *et al*, the physiological observations of 885 patients attending the ED and 165 patients admitted directly to ICU were collected prospectively. (10) This study demonstrated that REMS was superior to RAPS and equally as good as the more involved APACHE score. This study did not include a power calculation and may be under powered to predict death which occurred in 7.5% of the study group. Moreover, in this study the sample was supplemented with patients admitted

directly to ICU, thereby reducing the applicability of the findings to a general ED population. In the subsequent study by Olsson *et al*, 12,006 consecutive ED attendees were prospectively studied and information regarding all the physiological parameters of interest was available for 98% of the patients studied. (11) In addition to this high quality data, the findings were corroborated using a split sample technique; a statistical method of validation. It is likely that REMS performed better than RAPS because of the addition of oxygen saturations and age which were both found to be independent predictors of death in the multivariate analysis. Interestingly, blood pressure was not found to be an independent predictor of death, but the authors advocated its inclusion in the REMS score because it is routinely measured in all patients. The lack of a statistically significant association between blood pressure and death in this study may reflect the loss of accuracy in a one off manual measurement in comparison to the invasive BP monitoring used in the APACHE score. Importantly, this study also demonstrated that all patients with a score <3 survived, thereby providing a clinically useful cut-off.

The study by Goodacre *et al* was retrospective and was only able to calculate RAPS for 65% and REMS for 39% of the 5,583 patients included. (13) As this study was a secondary analysis, the study group only included patients admitted via ambulance with life-threatening conditions; this obviously skews the observations and may make comparisons to a general ED population invalid. On the basis of their multivariate analysis the authors derived a new score using age, oxygen saturations and GCS. Using this score, they found the AUROC value to be 0.81 which is very similar to the values obtained for REMS in the studies by Olsson. (10, 11)The most recent study to assess REMS

included 2,152 patients who attended the ED with signs of an infective illness. This was also a secondary analysis of a cohort of patients selected for another study in which the mortality rate was 3.9%. In this study the authors did not include GCS in the REMS score as this information was not available for all patients, however despite this omission, the score had an AUROC of 0.8. This group also compared this modified REMS to their previously validated score MEDS (Mortality in Emergency Department Sepsis Score) which was derived using a population of patients with sepsis. MEDS performed well with an AUROC of 0.85, although this is likely to reflect the suitability of this score to this particular cohort of patients, who were all admitted with infection, rather than a general ED population. Importantly this score also incorporates laboratory indices such as white cell count and other non physiological data.

Buising *et al* performed a novel study assessing a modified version of the CURB65 score incorporating blood pressure, level of confusion, respiratory rate and oxygen saturations (rather than blood urea) in patients presenting to the ED with clinical signs consistent with pneumonia. (14) Observations from 392 patients were used to derive the score, which was then validated prospectively in 332 patients. The mortality rate in this group was 9.4% and all four variables contributing to the score were found to be independent predictors of death and/or requirement for mechanical ventilation or inotropic support. Using simple cut-off values for each variable, the presence of two parameters above the cut-off threshold was associated with a sensitivity of 72% and a specificity of 70%. This study demonstrates the use of a simple physiological score in identifying high and low risk patients with a specific clinical diagnosis. It is logical that a score derived from a tight clinical phenotype will perform

well; the challenge for any score which will be clinically useful in the ED however, is one which performs well across all admissions.

In a retrospective cohort of 452 patients, who were subsequently admitted to ICU, Etter *et al* used the Medical Emergency Team (MET) calling criteria to create a MET score. (15) These MET criteria include the clinical assessment of a compromised airway, haemodynamic and respiratory instability, neurological impairment and overall concern about the patient. The AUROC was 0.82 for the prediction of hospital survival and multivariate analysis demonstrated that respiratory rate and GCS were the most significant independent outcome predictors. Seizures, threatened airway and general concern for the patient were not statistically significant predictors of mortality. Although in a small cohort, despite a mortality rate of 16%, this study may be under powered.

Discussion

The model of a physiologically-derived score has been investigated in a number of ED settings and patient cohorts. The overall conclusion from the studies which have been evaluated is that physiological scores, applied to patients in the ED, can predict hospital mortality. The strongest evidence to support this comes from Olsson *et al*'s REM score (10, 11) which has subsequently been tested in two other cohorts with promising results. (12, 13) Other scoring systems (CORB, MET) were also found to be predictive, with respiratory rate, neurological status and age being the most consistent parameters to be associated with subsequent mortality. The ability of a physiological score to predict mortality ultimately relies on the nature of the physiology being monitored. Physiological observations are frequently

unpredictable and often a dynamic assessment is necessary to identify deterioration in a patient's condition. Although a combination of extremely deranged physiological parameters, which contribute to a high EWS, have been demonstrated to be significantly associated with subsequent death, it does not necessarily follow that the same parameters will be as reliably predictive of other outcomes, such as admission to ICU or prolonged hospital stay. For example, Olsson *et al* included an analysis of the association between the ED REM score and the length of hospital stay and found only a modest correlation. (11) To allow direct comparisons, only studies which described mortality as their primary outcome were selected for this review. This has the advantage of being an easily definable outcome, however the inclusion of other clinical endpoints may be beneficial in determining the clinical usefulness of an EWS.

All the studies evaluated have demonstrated that when a cumulative physiological score is analysed as a continuous variable it has a strong correlation with subsequent mortality. This is not surprising, however the clinical value of an EWS system will be determined by its discriminatory power. As a measure of sensitivity and specificity the AUROC has been universally reported, but only one of the studies included the positive and negative predictive value of a cut-off threshold score. To be effective in clinical practice, an EWS needs to have both a high specificity and a reasonable sensitivity. A high specificity will minimise the number of patients whose score is falsely reassuring, whereas a reasonable sensitivity will ensure that only a manageable number of patients trigger unnecessarily therefore not significantly affecting the efficiency of the department. An EWS also needs to have a high degree of reproducibility in order to perform well. In related

studies which have investigated the use of EWS on medical wards, poor intra and inter-observer reproducibility has been reported, (16) however, this variability may be improved by the use of digital recording systems. (17) Reproducibility has not been rigorously tested in the studies described here and would need testing at a local level prior to the introduction of any such system into clinical practice. It is also vital to assess the quality of the observations documented and accuracy of the calculated physiological score before a new trigger system could be relied upon; an original audit performed in my unit is described in appendix B.

Although the data for some of the studies evaluated in this review has been collected prospectively, to date there has been no prospective evaluation of the performance of an EWS in a clinical setting. This would be necessary to ascertain whether the use of such a scoring system would result in improved management and therefore subsequently reduce mortality. It is surprising that the scores based on Morgan's EWS (6) have such widespread use amongst ward patients given the minimal evidence in the literature to support their use. (18) This lack of evidence is cause for concern as EDs may be attracted to the use of these scores if they are already in place in their hospital. It is possible in the absence of stringent scientific assessment of both their performance and applicability to the ED that many such scores will be inadequate.

The Future

There is sufficient evidence to demonstrate that physiological scores applied in the ED can predict mortality. This is only the first step towards the introduction of a system which reduces mortality as there is not yet sufficient evidence to

suggest that EWS have the utility to be introduced into the ED. Furthermore, in isolation a scoring system has limited scope to change or improve practice. Subsequent to the implementation of a physiological score calculation, the next stage would be the introduction of a protocol-guided trigger system (example included in appendix B). This process has been demonstrated to improve outcomes for patients on medical and surgical wards by outreach teams or a MET. (8, 19)

An EWS system could benefit the modern ED in the following ways:

- Help in identifying 'sick' patients and those at risk of adverse events (cardiac arrest, death and ITU admission). This would allow intensive treatment to be commenced earlier in the patient's journey which may have a positive effect on outcome. (20)
- Allow a graded response in the ED in relation to score. For example, a score >2 and <3 would lead to nurse led intervention and a score ≥3 would require ED senior doctor review prompting more timely intervention.
- Aid disposition both within ED and on transfer from the ED. For example, score >3 on transfer would prompt a medical review prior to transfer and/or disposition to a higher dependency area and a score >5 would result in transfer to the resuscitation room if resources allow. In patients who continue to trigger despite ED-based interventions, early referral to the critical care team would be mandatory.
- Allow assessments of response to treatment and the physiological wellbeing of the patient to be tracked from the beginning of their hospital stay. This would allow health professionals involved in ongoing care to

communicate in a common language.

- Physiological scoring during the triage assessment would facilitate an ongoing dynamic review in contrast to the single snap shot triage assessment which is currently in place in many units. The Cape Triage Score (CTS) (21) has the potential for this, combining clinical discriminators (such as chest pain) with physiological scoring

Summary

This review demonstrates there is evidence linking abnormal physiological scores in the ED and mortality. Further evidence is needed to change our current practice, ideally this would involve a large multi-centre trial using a validated ED-based physiological early warning score to drive an intervention-based response. The introduction of a physiological score which was able to discriminate between high and low risk patients and trigger prompt and appropriate intervention could not only improve the effectiveness of the ED but potentially save lives. This would help emergency departments who are striving to improve their efficiency and the quality of care that they provide.

References

1. Kause J, Smith G, Prytherch D, Parr M, Flabouris A, Hillman K. 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;62:275-82.

2. Goldhill DR, McNarry AF. Physiological abnormalities in early warning scores are related to mortality in adult inpatients. *Br J Anaesth* 2004;92:882-4.

3. Buist M, Bernard S, Nguyen TV, Moore G, Anderson J. Association between clinically abnormal observations and subsequent in-hospital mortality: a prospective study. *Resuscitation* 2004;62:137-41.

4. McGloin H, Adam SK, Singer M. Unexpected deaths and referrals to intensive care of patients on general wards. Are some cases potentially avoidable? *J R Coll Physicians Lond* 1999;33:255-9.

5. McQuillan P, Pilkington S, Allan A, Taylor B, Short A, Morgan G, Nielsen M, Barrett D, Smith G, Collins CH. Confidential inquiry into quality of care before admission to intensive care. *BMJ* 1998;316:1853-8.

6. Morgan R. An early warning score for the early detection of patients with impending illness. *Clinical Intensive Care* 1997;8:100-2.

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

8. Priestley G, Watson W, Rashidian A, Mozley C, Russell D, Wilson J, Cope J, Hart D, Kay D, Cowley K, Pateraki J. Introducing Critical Care Outreach: a ward-randomised trial of phased introduction in a general hospital. *Intensive Care Med* 2004;30:1398-404.

9. Simpson HK, Clancy M, Goldfrad C, Rowan K. Admissions to intensive care units from emergency departments: a descriptive study. *Emerg Med J* 2005;22:423-8.

10. Olsson T, Lind L. Comparison of the rapid emergency medicine score and APACHE II in nonsurgical emergency department patients. *Acad Emerg Med* 2003;10:1040-8.

11. Olsson T, Terent A, Lind L. Rapid Emergency Medicine Score can predict long-term mortality in nonsurgical emergency department patients. *Acad Emerg Med* 2004;11:1008-13.

12. Howell MD, Donnino MW, Talmor D, Clardy P, Ngo L, Shapiro NI. Performance of severity of illness scoring systems in emergency department patients with infection. *Acad Emerg Med* 2007;14:709-14.

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

14. Buising KL, Thursky KA, Black JF, MacGregor L, Street AC, Kennedy MP, Brown GV. Identifying severe community-acquired pneumonia in the emergency department: a simple clinical prediction tool. *Emerg Med Australas* 2007;19:418-26.

15. Etter R, Ludwig R, Lersch F, Takala J, Merz TM. Early prognostic value of the medical emergency team calling criteria in patients admitted to intensive care from the emergency department. *Crit Care Med* 2008;36:775-81.

16. Subbe CP, Gao H, Harrison DA. Reproducibility of physiological track-and-trigger warning systems for identifying at-risk patients on the ward. *Intensive Care Med* 2007;33:619-24.

17. Prytherch DR, Smith GB, Schmidt P, Featherstone PI, Stewart K, Knight D, Higgins B. Calculating early warning scores--a classroom comparison of pen and paper and hand-held computer methods. *Resuscitation* 2006;70:173-8.

18. Gao H, McDonnell A, Harrison DA, Moore T, Adam S, Daly K, Esmonde L, Goldhill DR, Parry GJ, Rashidian A, Subbe CP, Harvey S. 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.

19. Buist MD, Moore GE, Bernard SA, Waxman BP, Anderson JN, Nguyen TV. Effects of a medical emergency team on reduction of incidence of and mortality from unexpected cardiac arrests in hospital: preliminary study. *BMJ* 2002;324:387-90.

20. Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B, Peterson E, Tomlanovich M. Early goal-directed therapy in the treatment of severe sepsis and septic shock. *N Engl J Med* 2001;345:1368-77.

21. Bruijns SR, Wallis LA, Burch VC. A prospective evaluation of the Cape triage score in the emergency department of an urban public hospital in South Africa. *Emerg Med J* 2008;25:398-402.