Development and validation of a decision tree early warning score based on routine laboratory test results for the discrimination of hospital mortality in emergency medical admissions

Dr. Stuart W Jarvis, PhD, Research Assistant, Centre for Healthcare Modelling and Informatics, University of Portsmouth, Portsmouth, UK

Mrs. Caroline Kovacs, BSc, Research Assistant, Centre for Healthcare Modelling and Informatics, University of Portsmouth, Portsmouth, UK

Mrs. Tessy Badriyah, BSc. MSc., Research Student, Centre for Healthcare Modelling and Informatics, University of Portsmouth, Portsmouth, UK

Dr. Jim Briggs, BA, DPhil, Principal Lecturer & Director, Centre for Healthcare Modelling and Informatics, University of Portsmouth, Portsmouth, UK

Dr. Mohammed A Mohammed, PhD, Visiting Fellow, Centre for Healthcare Modelling and Informatics, University of Portsmouth, Portsmouth, UK and Senior Research Fellow, Primary Care Clinical Sciences, University of Birmingham, Birmingham, UK

Dr. Paul Meredith, PhD, Data Analyst, TEAMS centre, Portsmouth Hospitals NHS Trust, Portsmouth, UK

Dr. Paul E Schmidt, MRCP, B.Med.Sc, Consultant In Acute Medicine, Portsmouth Hospitals NHS Trust, Portsmouth UK

Dr. Peter I Featherstone, FRCP, Consultant In Acute Medicine, Portsmouth Hospitals NHS Trust, Portsmouth UK

Professor David R Prytherch, PhD, MIPEM, CSci, Clinical Scientist, TEAMS centre, Portsmouth Hospitals NHS Trust & Visiting Professor, Centre for Healthcare Modelling and Informatics, University of Portsmouth, Portsmouth, UK

Professor Gary B Smith, FRCA, FRCP, Visiting Professor, School of Health & Social Care, University of Bournemouth, Bournemouth, UK

Correspondence to:

Professor G B Smith, FRCA, FRCP,
Centre of Postgraduate Medical Research & Education (CoPMRE),
The School of Health & Social Care,
Bournemouth University, Royal London House,
Christchurch Road, Bournemouth,
Dorset BH1 3LT, United Kingdom

Tel: +44 (0) 1202 962782

Fax: +44 (0) 1202 962218

Email: garybsmith3@virginmedia.com

Word count = 2959

Number of references = 42

Figures = 5; Tables =1

Abstract

Aim of study: To build an early warning score (EWS) based exclusively on routinely undertaken laboratory tests that might provide early discrimination of in-hospital death and could be easily implemented on paper.

Materials and Methods: Using a database of combined haematology and biochemistry results for 86472 discharged adult patients for whom the admission specialty was Medicine, we used decision tree (DT) analysis to generate a laboratory decision tree early warning score (LDT-EWS) for each gender. LDT-EWS was developed for a single set ($n = 3496$) (Q_1) and validated in 22 other discrete sets each of three months long ($Q_2, Q_3, \dots, Q_{23}$) (total $n = 82976$; range of $n = 3428$ to 4093) by testing its ability to discriminate in-hospital death using the area under the receiver-operating characteristic (AUROC) curve.

Results: The data generated slightly different models for male and female patients. The ranges of AUROC values (95% CI) for LDT-EWS with in-hospital death as the outcome for the validation sets Q_2 - Q_{23} were: 0.755 (0.727 to 0.783) (Q_{16}) to 0.801 (0.776 to 0.826) [all patients combined, $n = 82976$]; 0.744 (0.704-0.784, Q_{16}) to 0.824 (0.792-0.856, Q_2) [39591 males]; and 0.742 (0.707-0.777, Q_{10}) to 0.826 (0.796-0.856, Q_{12}) [43385 females].

Conclusions: This study provides evidence that the results of commonly measured laboratory tests collected soon after hospital admission can be represented in a simple, paper-based EWS (LDT-EWS) to discriminate in-hospital mortality. We hypothesise that, with appropriate modification, it might be possible to extend the use of LDT-EWS throughout the patient's hospital stay.

Keywords

Early warning scores; Risk prediction; Biochemistry; Haematology; Illness severity score

Introduction

It is now common for hospitals to use predetermined 'calling criteria' as indicators of the need to increase the physiological monitoring of acutely ill patients or to deliver expert help to their bedside. In the UK, most hospitals use one of the large number of available early warning scores (EWS).¹ EWSs use measurements of vital signs (e.g., pulse rate, blood pressure, breathing rate and conscious level) as their basis. However, their usefulness is affected by the quality of the process of measurement of vital signs,² which relies upon (a) staff training,³ (b) the accuracy of clinical monitors,⁴ and (c) a correct measurement technique.^{5,6} Although measured less often than vital signs, the results of laboratory tests are subject to strict quality control and have independently been identified as risk factors for poor patient outcome.⁷⁻²⁸ Therefore, prediction models or EWSs based exclusively on laboratory test results may offer an additional opportunity to identify sick or 'at risk' patients, either at hospital admission or, perhaps, throughout the hospital stay. Whilst physiological EWS have currency and utility in acute hospital care, there is no equivalent for common blood tests.

Our group has previously shown that the risk of in-hospital death can be modelled in general surgical and unselected medical patients using the results of the seven most commonly used laboratory tests in our hospital (i.e., haemoglobin (Hb); white cell count (WCC); serum urea (U); serum albumin (Alb); serum creatinine (Cr); serum sodium (Na); and serum potassium (K)), and several administrative items (i.e., patient age at admission, patient gender, and mode of admission (elective or emergency)).^{7,8} All of these are routinely available on or near hospital admission. We termed this algorithmic predictive model the 'Biochemistry and Haematology Outcome Model' (BHOM).^{7,8} However, the binary logistic regression method used to develop BHOM^{7,8} produces a complex algorithm that has little utility as a paper-based EWS for general hospital areas because of the nature of calculations involving logarithms. Therefore, we decided to investigate if it was feasible to use a computerised, statistical method – Decision Tree (DT) analysis²⁹ – to build a EWS that might predict patients at risk of in-hospital death early in their hospital stay using only laboratory tests and that could be easily implemented on paper.

Method

Local research ethics committee approval was obtained for this study from the Isle of Wight, Portsmouth and South East Hampshire Research Ethics Committee.

Laboratory test results database and its development

We searched the computerised hospital records of all discharged adult patients admitted to Portsmouth Hospitals NHS Trust during the period 01/07/2006 to 31/03/2012, for whom the admission specialty was Medicine. Data from patients aged <16 years at hospital admission; patients admitted as elective admissions; and patients discharged alive on the day of admission were excluded. We extracted the first routinely collected values taken during the hospital stay of a range of common haematology and biochemistry blood tests (i.e., Hb; WCC; U; Alb; Cr; Na; and K), which were within the acceptable measurement ranges of the hospital laboratory service analysers. There were no unacceptable upper and lower limits for Hb and WCC. The acceptable ranges for the biochemistry parameters were U (0.4-107.1 mmol/L); Alb (10-70 g/L); Cr (8.8-2210 μ mol/L); Na (100-200 mmol/L); and K (1-15 mmol/L). Patient gender and patient outcome (alive or dead at hospital discharge) were also extracted. We constrained the data for analysis to those records where there was a value for each of the seven parameters from specimens taken on the day of admission or the following day. There was no requirement that the results came from the same venesection.

The dataset of routine blood test results was derived from episodes of patient care recorded over 69 consecutive months (23 quarters Q_1 - Q_{23}). This dataset was split into a single development set (data from the three month period 01/07/2006 to 30/09/2006; Q_1) and 22 validation sets, each of three months long (Q_2 , Q_3 ,..... Q_{23}). The decision to select Q_1 as the development set was an arbitrary one.

Development of laboratory data DT EWS (LDT-EWS) using DT analysis

To take account of any variation arising from the different reference ranges for males and females for haemoglobin and creatinine, data from Q_1 were split on the basis of gender. For each gender, we used DT analysis to generate a LDT-EWS from the laboratory data. DT analysis is a data mining classification technique for building decision trees by recursively splitting or partitioning of datasets into homogenous groups.²⁹ This partitioning is based on derived associations between the outcome – in our case, in-hospital death - and one or more covariates. Our tree modelling strategy assessed the

following covariates individually: Hb; WCC; U; Alb; Cr; Na; and K. Each of the terminal leaves of the DT produced by the software (*R* statistical programming language³⁰ version 2.15.2, using the *rpart* package³¹) provides the number of cases in it and the outcome, so that a risk of the outcome can be generated. As an example, the decision tree and the nodes obtained for serum albumin for male patients is shown in Figure 1. The *rpart* package provides various settings to constrain the trees, including a maximum number of levels, maximum number of leaf nodes and constraints on the minimum number of episodes that must be present in each leaf node. To prevent over-fitting, we constrained the trees so that there were no leaf nodes with fewer than 5% of the total cases.

To replicate the classic approach used by most EWS^{1,32,33}, we chose to develop LDT-EWS with a 0, 1, 2 and 3 weighting system for the risk bands. The risk bands were set as follows: where the risk generated by the DT analysis for any given parameter was $<$ mean risk of in-hospital death (relative risk <1), a value of 0 was ascribed; if the risk was $\geq$ mean risk and $<$ 2 times mean risk, a value of 1 was ascribed; if the risk was $\geq$ 2 times mean risk and $<$ 3 times mean risk, a value of 2 was ascribed; and if the risk was $\geq$ 3 times mean risk, a value of 3 was ascribed. DTs for some of the laboratory parameters presented more than seven distinct bands of measurement values, each with different risks. In these cases, once weightings had been assigned to the risks, leaves that have the same or similar risks were merged. This process was repeated for each of the seven laboratory parameters to produce a LDT-EWS for each gender (Table 1).

Table 1: Laboratory data decision tree EWS (LDT-EWS) for male and female patients**Males**

	3	2	1	0	1	2	3
Hb		≤11.1	11.2-12.8	≥12.9			
WCC				≤9.3	9.4-16.6	≥16.7	
U				≤9.4	9.5-13.7		≥13.8
Cr				≤114	115-179	≥180	
Na		≤132		133-140	≥141		
K			≤3.7	3.8-4.4	4.5-4.7	≥4.8	
Alb		≤30	31-34	≥35			

Females

	3	2	1	0	1	2	3
Hb			≤12.0	12.1-14.8	≥14.9		
WCC				≤12.6	12.7-14.8	≥14.9	
U				≤8.4	8.5-13.8		≥13.9
Cr				≤91	92-157	≥158	
Na			≤134	135-140	≥141		
K			≤3.3	3.4-4.5	≥4.6		
Alb		≤28	29-34	≥35			

Where Hb = haemoglobin; WCC = white cell count; U = serum urea; Alb = serum albumin; Cr = serum creatinine; Na = serum sodium; K = serum potassium

Evaluation of LDT-EWS

For each patient admission in each of the additional 22 sets (Q_2 - Q_{23}), we converted raw laboratory test results into aggregate LDT-EWS values using the appropriate scoring system for the patient's gender. For each of the 22 validation sets (Q_2 - Q_{23}), we combined all LDT-EWS values and used the area under the receiver-operating characteristics (AUROC) curve³⁴ to evaluate the ability of LDT-EWS to discriminate patients at risk of hospital death for each set. Therefore, the LDT-EWS generated from Q_1 was validated against 22 other sets (Q_2 - Q_{23}). AUROC analysis was performed using the *pROC* package³⁵ in *R*. Additionally, as a measure of the relative number of “triggers” that would be generated at different values of LDT-EWS, we also produced an “EWS efficiency curve”³¹ for in-hospital death for LDT-EWS using the amalgamated set of Q_2 - Q_{23} .

Data analysis

All data manipulation was performed using Microsoft® Visual FoxPro 9.0. All analyses were undertaken in *R* version 2.15.2.

Results

In the study period, there were 97585 discharges of patients admitted during the period 01/07/2006 to 31/03/2012, where the admission specialty was Medicine, the mode of admission was non-elective, the patient was aged ≥ 16 and the patient was not discharged alive on the day of admission. 91242 episodes (93.5%) had a full and valid set of the necessary haematology and biochemistry blood tests for the analysis. Of the 91242 episodes with valid tests, 86472 had the tests taken on either the day of admission or the following day.

Therefore, our final dataset comprised 86472 episodes from an initial dataset of 97585 episodes (88.6%). The mortality (95% CI) in our initial dataset ($n=97585$) was 8.64% (8.47% to 8.82%). Our requirements for valid haematology and biochemistry blood tests to be taken on the day of admission or the next day excluded 11113 episodes and 976 deaths from our analysis. In these episodes, the mortality (95% CI) was 8.77% (8.27% to 9.32%). The mortality (95% CI) in our final dataset ($n=86472$) was 8.63% (8.44% to 8.81%).

The final dataset ($n = 86472$) was split into a single development set (Q_1 : $n = 3496$) and 22 validation sets ($Q_2, Q_3, \dots, Q_{23}$: $n = 82976$). Each validation set contained between 3428 and 4093 patient episodes. The development set (Q_1) included 47.71% males. The mean incidence of death in the development set was 8.27% for males and 10.07% for females.

The distribution of LDT-EWS values and the relationship to in-hospital death in Q_2 - Q_{23} combined is shown in Figure 2. The AUROC values (95% CI) for the discrimination of in-hospital death for validation sets Q_2 - Q_{23} , for males and female patients combined, are shown in Figure 3. The AUROC values (95% CI) ranged from 0.755 (0.727 to 0.783) (Q_{16}) to 0.801 (0.776 to 0.826) (Q_9). The CIs for all AUROC values overlap, suggesting that there is no significant difference between the AUROC values of any of the datasets. If male and female patients are studied separately, the AUROC values range from 0.744 (0.704-0.784, Q_{16}) to 0.824 (0.792-0.856, Q_2) for males and 0.742 (0.707-0.777, Q_{10}) to 0.826 (0.796-0.856, Q_{12}) for females.

Figure 4 shows the “EWS efficiency curve”³² for LDT-EWS and NEWS for in-hospital death for

Q_2 - Q_{23} combined. This provides a relative measure of the number of “triggers” that would be generated at different values of LDT-EWS. Choosing a LDT-EWS value of 4 as the trigger for calling for additional assistance would mean that 40.7% of all laboratory test result datasets would trigger and that 79.7% of all patients subsequently dying would be visited.

Figure 5 shows separate EWS efficiency curves for LDT-EWS for 39591 male and 43385 female patients in Q_2 - Q_{23} . Whilst the curves are very similar in shape, the position of LDT-EWS scores differs for males and females - a LDT-EWS score = 4 for a female (35.2% calls triggered; 75.3% deaths visited) is in an almost identical position on its curve to a LDT-EWS score = 5 for a male (36.7% calls triggered; 75.8% deaths visited).

Discussion

There is increasingly interest in the role of laboratory tests as predictors of patient outcome in recent years.⁷⁻²⁸ Many of these authors have described the development of complex prediction models based on logistic regression techniques, but all use one or more variables, such as combinations of age, symptoms and physiology, in addition to routinely measured laboratory test results.⁷⁻²⁸ DT analysis has been studied previously in healthcare,³⁶⁻³⁹ partly because tree models are visual, easy to interpret and reflect aspects of human decision making processes. To our knowledge, DT techniques have not been used to describe a simple tabular display of weighted values for the incorporated covariates based on risk. Therefore, we believe that LDT-EWS is the first example of a statistically developed EWS, based exclusively on the results of commonly undertaken laboratory test results, which appears to be suitable for use as a simple, paper-based scoring system for evaluating the risk of in-hospital mortality. Its simplicity mimics that of the widely used physiologically-based EWS.^{1,32,33}

Importantly, all of the necessary data to calculate a LDT-EWS value are readily obtainable within the first hours of hospital admission via a single venesection. Additionally, all data are collected routinely as part of the clinical process such that their collection presents no added burden to the staff providing care. Further, all components of LDT-EWS are subject to strict quality control by the laboratory (unlike physiological variables measured by ward staff which can be more significantly affected by human error).²⁻⁶

We have previously shown that an initial set of vital signs measured after admission to hospital can be used to produce EWS values that discriminate final hospital outcome well using certain EWS systems.¹ The AUROC ($\pm 95\%$ CI) for 33 unique EWSs ranged from 0.657 (0.636-0.678) to 0.782 (0.767-0.797) using in-hospital death as the outcome.¹ Hence the range of AUROC values ($\pm 95\%$ CI) obtained from the 22 test sets in our study using laboratory data only (from 0.755 (0.727 to 0.783) to 0.801 (0.776 to 0.826)) highlights the predictive power that laboratory based EWS systems may possess. Not surprisingly, the AUROC values for vital sign-based EWSs are increased as the time between the vital signs measurement and the outcome is decreased (i.e., the certainty of outcome becomes greater as one moves closer to the outcome).³² By extrapolation, it is therefore possible that EWS systems using only laboratory data might also have an important role to play in the early determination of adverse outcomes.

The efficiency curves (Figures 4 and 5) provide a measure of the number of “triggers” that would be generated at different values of LDT-EWS, thereby providing an estimate of the likely clinical workload. The finding that the position of LDT-EWS scores differs for males and females, even though the shapes of their respective curves are similar, implies that, in clinical practice, it might be appropriate to have a different trigger point for male and female patients.

One drawback to our current approach is that whilst many, if not all, non-elective admissions to hospital have a full set of biochemistry and haematology tests done during the first hours of hospital admission, often via a single admission venesection, subsequent tests are done in a ‘piecemeal’ fashion. However, a recent study by Loekito et al used batches of laboratory tests, where each batch contained a variably sized subset of 30 common laboratory measurements, such as ‘serum levels of Na, K, Cl, U, Cr, Alb, total bicarbonate, bilirubin, alkaline phosphatase, alanine aminotransferase, and gamma-glutamyl transferase’²⁷ They were able to show that multivariate logistic modelling achieved an AUROC ($\pm 95\%$ CI) of 0.87 (0.85–0.89) for the prediction of imminent death, defined as the patient’s death either on the same day or during the following calendar day. A similar AUROC ($\pm 95\%$ CI) of 0.88 (0.85–0.90) was obtained by employing their method in a second hospital.²⁷ The same authors have also shown good discrimination in detecting imminent ICU admission and imminent death, using the same technique.²⁸ These findings provide optimism that an LDT-EWS might be predictive of outcome even when a full set of biochemistry and haematology tests is not available for use in the calculation of an LDT-EWS value. This hypothesis clearly needs further testing.

Additionally, the ability to discriminate the risk of mortality using precisely measurable and objective data, which are subject to routine quality control processes, could be an adjunct to the physiological-based early warning systems commonly used in acute care. Potentially, the use of an EWS based upon vital signs and laboratory data may be useful. Alternatively, laboratory based EWS values may be predictive of long-term outcome, whilst values from vital sign-based EWS may provide ‘minute-by-minute’ prediction.

We would emphasise that mathematical systems such as LDT-EWS should not be regarded as the sole solution to detecting patient deterioration. Their use should complement the use of other triggers,

such as symptoms (e.g., chest pain); signs (e.g., sweating); other assessment scores (e.g., the Glasgow Coma Scale)⁴⁰; and nurse⁴¹ or family concern.⁴²

The strengths of the study are that it uses a large database from almost six years of completed inpatient admissions. Of most importance is that the data used for the analysis has been subjected to the highest quality control measures. We were also able to confirm the repeatability and stability of the EWS over time by demonstrating stable AUROC values in the 22 test sets, which also ensured that the EWS was not tested against the data used for its derivation. The weaknesses of the study are that it is a single centre study and the extent to which the findings are generalisable to other hospitals remains to be seen. Another potential weakness is that laboratory values may be subject to the influence of treatment strategies, such as potassium during beta-agonist therapy. Our study analysed data from the first routinely collected laboratory test result taken during the hospital stay, which were as close to admission as possible and, therefore, least likely have been influenced by prior therapy. However, the possibility of some modulation of clinical intervention cannot be completely discounted. Nevertheless, and in spite of any such potential impact, we were still able to show that LDTEWS is clearly capable of discriminating hospital death.

Finally, other authors have questioned the ability to develop a laboratory test based EWS similar to the numerous vital signs-based early warning scores in existence, because of the complexity of the necessary calculations, and the need for appropriate automated software to calculate and transmit the information to clinicians.²⁷ We have clearly shown that it is indeed possible to develop a simple, paper-based laboratory test result-based EWS that currently provides good discrimination of in-hospital death. The challenge is now to demonstrate that such a system can be safely implemented into the process of care in a hospital without unintended consequences (e.g., unnecessary cost) and to investigate if such systems can be used throughout the patient's hospital stay.

Conclusions

This study provides evidence that the results of commonly measured laboratory tests collected soon after hospital admission can be used in a simple, paper-based scoring EWS (LDT-EWS) to discriminate in-hospital mortality. When compared with published data regarding the discriminating ability of a large number of vital sign-based EWSs for in-hospital mortality, LDT-EWS performed

similarly. We hypothesise that, with appropriate modification, it might be possible to extend the use of LDT-EWS for use on an ongoing basis throughout the patient's hospital stay.

Conflicts of interest

None

Funding

None.

References

1. Smith GB, Prytherch DR, Schmidt P, Featherstone PI. Review and performance evaluation of aggregate weighted 'track and trigger' systems. *Resuscitation* 2008;77:170-179.
2. Smith GB, Prytherch DR. The Afferent Limb of Rapid Response Systems. In DeVita MA, Hillman K, Bellomo R. *Textbook of Rapid Response Systems: Concept and Implementation*, Springer, 2011.
3. Department of Health. *Competencies for Recognising and Responding to Acutely Ill Patients in Hospital*. London; 2009.
4. Milner QJ, Mathews GR. An assessment of the accuracy of pulse oximeters. *Anaesthesia* 2012;67:396-401.
5. Smith GB. Measuring pulse and breathing rates - Simple, yet complex. *Resuscitation* 2011;82:1367-1368
6. Kellett J, Li M, Rasool S, Green GC, Seely A. Comparison of the heart and breathing rate of acutely ill medical patients recorded by nursing staff with those measured over 5 minutes by a piezoelectric belt and ECG monitor at the time of admission to hospital. *Resuscitation* 2011;82:1381-6.
7. Prytherch DR, Sirl JS, Weaver PC, Schmidt P, Higgins B, Sutton GL. Towards a national clinical minimum data set for general surgery. *Br J Surg* 2003;90:1300-5.
8. Prytherch D, Sirl J, Schmidt P, Featherstone P, Weaver PC, Smith GB. The use of routine laboratory data to predict in-hospital death in medical admission. *Resuscitation* 2005;66:203-207.
9. Mohammed MA, Rudge G, Wood G et al. Which is more useful in predicting hospital mortality - dichotomised blood test results or actual test values? A retrospective study in two hospitals. *Plos One* in press 2012.
10. Asimwe A, Brims F, Andrews N, Prytherch D, Higgins B, Kilburn S and Chauhan A. Routine laboratory tests can predict in-hospital mortality in acute exacerbations of COPD. *Lung* 2011;189:225-232.
11. Vroonhof K, van Solinge WW, Rovers MM, Huisman A. Differences in mortality on the basis of complete blood count in an unselected population at the emergency department. *Lab Hematol* 2006; 12:134-8.

12. ten Boekel E, Vroonhof K, Huisman A, van Kampen C, de Kieviet W. Clinical laboratory findings associated with in-hospital mortality. *Clinica Chimica Acta* 2006; 372: 1-13.
13. Asadollahi K, Hastings IM, Beeching NJ, Gill GV. Laboratory risk factors for hospital mortality in acutely admitted patients. *Q J Med* 2007; 100: 501–507.
14. Novack V, Pencina M, Zahger D, et al. Routine Laboratory Results and Thirty Day and One-Year Mortality Risk Following Hospitalization with Acute Decompensated Heart Failure. *PLoS ONE* 2010; 5: e12184. doi:10.1371/journal.pone.0012184
15. Hodgetts TJ, Kenward G, Vlachonikolis IG, Payne S, Castle N. The identification of risk factors for cardiac arrest and formulation of activation criteria to alert a medical emergency team. *Resuscitation* 2002;54:125-131.
16. Silke B, Kellett J, Rooney T, Bennett K, O’Riordan D. An improved medical admissions risk system using multivariable fractional polynomial logistic regression modelling. *QJM*. 2010;103:23-32.
17. O’Sullivan E, Calley E, O’Riordan D, Bennett K, Silke B. Predicting outcomes in emergency medical admissions – role of laboratory data and co-morbidity. *Acute Medicine* 2012; 11(2): 59-65.
18. Romero-Ortuno R, Silke B. Use of a laboratory only score system to define trajectories and outcomes of older people admitted to the acute hospital as medical emergencies. *Geriatr Gerontol Int* 2012. DOI: 10.1111/j.1447-0594.2012.00917.x
19. Prytherch DR, Briggs JS, Weaver PC, Schmidt P, Smith GB. Measuring clinical performance using routinely collected clinical data. *Medical Informatics and the Internet in Medicine* 2005;30:151-156
20. Tang T, Prytherch D, Walsh S et al. The development of a VBHOM-based outcome model for lower limb amputation performed for critical ischaemia. *Eur J Vasc and Endovasc Surg* 2009;37:62-66.
21. Farooq N, Patterson A, Walsh S, Prytherch D, Justin T, Tang T. Predicting outcome following colorectal cancer surgery using a colorectal biochemical and haematological outcome model (Colorectal BHOM). *Colorectal Disease* 2010;13:1237-1241.
22. Nichols K, Prytherch D, Fancourt M, Gilkison W, Kyle S; Mosquera D. Risk-adjusted general surgical audit in octogenarians. *ANZ J Surg* 2008;78:990-994.

23. Pine M, Jones B, Lou YB. Laboratory values improve predictions of hospital mortality. *Int J Qual Health Care* 1998; 10: 491-501.
24. Hucker TR, Mitchell GP, Blake LD, et al. Identifying the sick: can biochemical measurements be used to aid decision making on presentation to the accident and emergency department. *Br J Anaesth* 2005; 94: 735–41.
25. Froom P, Shimoni Z. Prediction of Hospital Mortality Rates by Admission Laboratory Tests. *Clinical Chemistry* 2006; 52: 325-8.
26. Asadollahi K, Hastings IM, Gill GV, Beeching NJ. Prediction of hospital mortality from admission laboratory data and patient age: A simple model. *Emergency Medicine Australasia* 2011; 23: 354-363.
27. Loekito E, Bailey J, Bellomo R et al. Common laboratory tests predict imminent death in ward patients. *Resuscitation* 2013;84:280-5.
28. Loekito E, Bailey J, Bellomo R et al. Common laboratory tests predict imminent medical emergency team calls, intensive care unit admission or death in emergency department patients. *Emergency Medicine Australasia* 2013;25:132–9.
29. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and Regression Trees*. Wadsworth 1984
30. R Core Team (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0
31. Therneau T, Atkinson B, Ripley B. *rpart: Recursive Partitioning*. R package version 4.1-0. 2012 <http://CRAN.R-project.org/package=rpart>
32. Prytherch D, Smith GB, Schmidt PE, Featherstone PI. ViEWS - towards a national Early Warning Score for detecting adult inpatient deterioration. *Resuscitation* 2010;81:932–937.
33. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. *Resuscitation* 2013 Jan 4. pii: S0300-9572(13)00002-6. doi: 10.1016/j.resuscitation.2012.12.016. [Epub ahead of print]
34. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29-36.

35. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2010; 12:77. DOI:10.1186/1471-2105-12-77
36. Ridley S, Jones S, Shahani A, Brampton W, Nielsen M, Rowan K. Classification trees. A possible method for iso-resource grouping in intensive care. *Anaesthesia*. 1998;53:833-40.
37. Abu-Hanna A, de Keizer N. Integrating classification trees with local logistic regression in Intensive Care prognosis. *Artif Intell Med*. 2003;29:5-23.
38. Javier Trujillano J, Badia M, Serviá L, March J, Rodriguez-Pozo A. Stratification of the severity of critically ill patients with classification trees. *BMC Medical Research Methodology* 2009, 9:83 doi:10.1186/1471-2288-9-83.
39. Kim S, Kim W, Park RW. A Comparison of Intensive Care Unit Mortality Prediction Models through the Use of Data Mining Techniques. *Healthc Inform Res*. 2011 Dec;17(4):232-43. doi: 10.4258/hir.2011.17.4.232. Epub 2011 Dec 31.
40. Teasdale G, Jennett B. Assessment of coma and impaired consciousness: a practical scale. *Lancet* 1974;2:81–4.
41. Santiano N, Young L, Hillman K, et al. Analysis of Medical Emergency Team calls comparing subjective to “objective” call criteria. *Resuscitation* 2009;80:44–9
42. Ray EM, Smith R, Massie S, et al. Family alert: implementing direct family activation of a pediatric rapid response team. *Jt Comm J Qual Patient Saf* 2009;35:575–80.

Figures

Figure 1

The decision tree for serum albumin for male patients and the terminal leaves generated during the development of LDT-EWS. The relative risk is determined by dividing the risk for the value/range in the terminal leaf by the mean risk of in-hospital death.

Figure 2

The distribution of LDT-EWS values in the sets Q_2 - Q_{23} and the relationship of LDT-EWS to in-hospital death.

Figure 3

AUROC values (95% CI) for LDT-EWS for each of the three-month long test sets (Q₂-Q₂₃) for male and female patients combined with in-hospital death as the outcome.

Figure 4

EWS efficiency curves for LDT-EWS. For each LDT-EWS value, this plots the percentage of the total number of laboratory test result datasets at, or above, that LDT-EWS value against the percentage of the total number of laboratory test result datasets for which the outcome – in-hospital death – was true at, or above, that LDT-EWS value. From the point at 100,100 the LDT-EWS values are 0, 1, 2, 3

Figure 5

Separate EWS efficiency curves for LDT-EWS for 39591 male and 43385 female patients in Q_2 - Q_{23} .

The position of LDT-EWS scores differs for males and females - a LDT-EWS score = 4 for a female (35.2% calls triggered; 75.3% deaths visited) is in an almost identical position on its curve to a LDT-EWS score = 5 for a male (36.7% calls triggered; 75.8% deaths visited).