Original Article

Poor performance of the modified early warning score for predicting mortality in critically ill patients presenting to an emergency department

Le Onn Ho[1], Huihua Li[2], Nur Shahidah[1], Zhi Xiong Koh[1], Papia Sultana[1], Marcus Eng Hock Ong[1,3]

[1] Department of Emergency Medicine, Singapore General Hospital, Singapore

[2] Department of Clinical Research, Singapore General Hospital, Singapore

[3] Office of Clinical Science, Duke-NUS Graduate Medical School, Singapore

Corresponding Author: Le Onn Ho, Email: leonn.ho@gmail.com

BACKGROUND: This study was undertaken to validate the use of the modified early warning score (MEWS) as a predictor of patient mortality and intensive care unit (ICU)/ high dependency (HD) admission in an Asian population.

METHODS: The MEWS was applied to a retrospective cohort of 1 024 critically ill patients presenting to a large Asian tertiary emergency department (ED) between November 2006 and December 2007. Individual MEWS was calculated based on vital signs parameters on arrival at ED. Outcomes of mortality and ICU/HD admission were obtained from hospital records. The ability of the composite MEWS and its individual components to predict mortality within 30 days from ED visit was assessed. Sensitivity, specificity, positive and negative predictive values were derived and compared with values from other cohorts. A MEWS of ≥4 was chosen as the cut-off value for poor prognosis based on previous studies.

RESULTS: A total of 311 (30.4%) critically ill patients were presented with a MEWS ≥4. Their mean age was 61.4 years (SD 18.1) with a male to female ratio of 1.10. Of the 311 patients, 53 (17%) died within 30 days, 64 (20.6%) were admitted to ICU and 86 (27.7%) were admitted to HD. The area under the receiver operating characteristic curve was 0.71 with a sensitivity of 53.0% and a specificity of 72.1% in addition to a positive predictive value (PPV) of 17.0% and a negative predictive value (NPV) of 93.4% (MEWS cut-off of ≥4) for predicting mortality.

CONCLUSION: The composite MEWS did not perform well in predicting poor patient outcomes for critically ill patients presenting to an ED.

KEY WORDS: Modified early warning score; Emergency department; Outcomes; Triage

INTRODUCTION

In the emergency department, triage is used to assess the severity of patients' conditions and to assign appropriate treatment priorities. This clinical process is especially important for critically ill patients. Triage is a necessity as medical resources are never sufficient for all patients to be attended instantaneously in busy emergency departments and hospitals, with limited numbers of doctors, nurses, monitored beds, resuscitation facilities, intensive care beds, operating theatres, etc.

Thus triage at the emergency department (ED) becomes crucial to matching patient acuity to hospital resources.

As patient care processes are analyzed, one identified risk area involves incidents where patients are not recognized by healthcare staff as being potentially ill, despite presenting with deranged physiological parameters. This results in a delay in recognizing critically ill patients, instituting appropriate clinical measures and leading to poor patient outcomes. In some general ward settings, abnormal parameters have gone

on for hours or even days before the patient is finally recognized as being critically ill. The health grades quality study has listed such incidents of 'Failure to Rescue' as one of most common patient safety incidents which attributes to inpatient mortality.[1]

One way of identifying at-risk patients on the ward is through physiologically based early warning scores, where an ordinal score is assigned to increasingly derange physiological values. One such scoring system is the modified early warning score (MEWS) which was studied in the Addenbrooke's Hospital and the West Suffolk Hospital, UK, where it was used to trigger emergency outreach programs and escalate the level of care for critically ill general ward surgical patients. It is based on physiological parameters: systolic blood pressure, pulse rate, respiratory rate, temperature and AVPU score (A for 'alert', V for 'reacting to vocal stimuli', P for 'reacting to pain', U for 'unconsciousness').[2] MEWS has a score range of 0 (lowest) to 3 (highest) for each of its 5 parameters, with the composite score from all 5 parameters representing the MEWS. Lower MEWS predicts better outcomes and a cutoff composite score of 4 or higher is considered predictive of poor outcomes. Other studies done in the UK have confirmed the utility of the MEWS in predicting poor outcomes.[3–5] However, these studies have been conducted in primarily homogenous western populations and the utility of the MEWS in an Asian population has not been previously reported.

This study aims to validate the ability of the MEWS to predict mortality and poor outcomes for critically ill patients in a Singapore context by evaluating the utility of MEWS in the emergency department setting. We hypothesize that the sensitivity, specificity, positive and negative predictive value of the MEWS in the local population would be similar to that reported for non-Asian populations.

METHODS
Study setting
This was a retrospective cohort study of patients who presented to the emergency department of Singapore General Hospital (SGH), a large tertiary hospital from November 23, 2006 to December 12, 2007. Singapore is an independent urban island nation in Southeast Asia with a population of 4.5 million.[6] SGH is one of the 7 tertiary hospitals on the island and is the oldest and largest acute tertiary hospital, with an ED volume of 300 to 500 patients daily as well as being a level 1 trauma center. SGH accounts for about one third of all acute care public sector beds and about a quarter of acute beds nationwide. Annually, about 60 000 patients are admitted and another

600 000 are attended to specialist outpatient clinics.

Patient selection
In this opportunistic study we used data collected on critically ill patients for a study on heart rate variability. Our study included patients at age of 18 years or older presenting to the ED who required continuous ECG monitoring, with patient acuity category scale (PACS) of 1 or 2 (see below), recruited during the office hours of 8:00 to 18:00. The patients who presented with cardiac arrest without any 'vital signs' available were excluded. We also excluded PACS 3 and 4 patients to focus on critically ill patients because of low expected event rate (death) in PACS 3 and 4 patients.

Data collection
We collected initial vital signs obtained during triage in the ED. This data collection was part of an ongoing study of critically ill patients in the ED. Patient demographic data were also collected and included age, gender, race and medical history.

All public hospitals in Singapore use a national patient acuity category scale (PACS) for triage at the ED. PACS 1 patients are the most critically ill and attended to without delay. These patients would have been triaged by a nurse as having airway, breathing and circulation problems or thought to be possibly unstable and to need close monitoring. PACS 2 patients are non-ambulant and would appear to be in a stable state on initial cardiovascular examination, and not at risk of imminent collapse. PACS 3 patients are ambulant and PACS 4 patients are considered non-emergencies. At triage, blood pressure (BP), heart rate (HR), respiratory rate (RR), temperature and general patient condition were recorded and a composite MEWS for each patient using pre-specified criteria was calculated. Each individual component of the patient's MEWS was summed for the final composite MEWS.

Follow up
We followed all patients for inpatient outcomes, subsequent transfers and other dispositions. Outcomes such as high dependency or intermediary care area (HD/ICA) admissions, transfers to intensive care units (ICU) and death were obtained from the review of hospital charts and inpatient electronic record systems. Follow up occurred until the patient was discharged from hospital, died or 30 days after being seen in the ED.

Outcome measures
The primary outcome was patient mortality during

the inpatient period following admission from the ED up to 30 days. Secondary outcome measures included direct admission from the ED to the high dependency unit, intermediary care area or the intensive care unit.

Statistical analysis

Sample size was opportunistic, based on the maximum data collection possible during office hours, over a 1-year period. Entry of data was performed with MS Excel (Microsoft Corp., Redmond WA) and data analysis was made with SPSS (SPSS Inc., Chicago, IL). A MEWS ≥4 was used as a cut-off for validation of outcomes.[2] Area under the receiver operating curve (ROC), sensitivities and specificities were calculated using STATA 11.0 (StataCorp, Texas) and logistic regression was used to evaluate the predictive capability of each individual MEWS component. Statistical significance was set at <0.05.

RESULTS

A total of 1 024 patients who presented to the ED during the hours of 8:00 to 18:00 from November 2006 to December 2007 were included in the analysis. Patient demographics are shown in Table 1. There were 713 patients with a MEWS score of <4 and 311 patients with a MEWS score of ≥4. The two groups had a similar mean age of 62.3 (15.4) years and 61.4 (18.1) years respectively. In both groups, there was a propensity of the patients being male with a male to female ratio of 1.8 in the MEWS <4 group and 1.1 in the MEWS ≥4 group. There were no significant differences in race distributions between the two groups. This was also the case for the distributions of co-morbidities. Table 2 shows the outcomes between the groups MEWS <4 and MEWS ≥4.

There were 47 deaths (6.6%) in the MEWS <4 group as compared with 53 (17.0%) deaths in the MEWS ≥4 group. Two hundred and sixty-seven patients were admitted to HD/ICA (37.4%) in the MEWS <4 group as compared with 86 (27.7%) patients in the MEWS ≥4 group. The average length of stay (ALOS) for the MEWS <4 group was 6.97 days and for the MEWS <4 group was 7.75 days.

Tables 3 and 4 describe the sensitivity and

specificity of MEWS in predicting mortality and hospital admissions. A MEWS score of <4 had a sensitivity of 47% and a specificity of 27.92% in predicting patient mortality. This corresponded to a 6.66% positive predictive value (PPV), 82.96% negative predictive value (NPV) and an area under the receiver operating characteristic (ROC) curve of 0.68. A MEWS score of <5 had a sensitivity of 66% and a specificity of 12.34% with a corresponding PPV of 7.53%, a NPV of 77.03% and an area under ROC curve of 0.662. When predicting admission to ICU, HD or the ICA, a MEWS score of <4 had a sensitivity of 74.16% and a specificity of 33.91% with a corresponding PPV of 46.7%, a NPV of 62.7% and an area under the ROC curve of 0.49. The MEWS score of <5 had a sensitivity of 87.75% and a specificity of 16.17% with a corresponding PPV of 44.98%, a NPV of 62.84% and an area under the ROC curve of 0.47.

Table 1. Characteristics of study population

Variables	MEWS <4 (*n*=713)	MEWS ≥4 (*n*=311)	*P* value
Age, mean (SD)	62.3 (15.4)	61.4 (18.1)	0.44
Gender, ratio (M:F)	1.80	1.10	<0.001
Race (%)			
Chinese	468 (65.6)	221 (71.1)	
Malay	111 (15.6)	42 (13.5)	0.39
Indian	92 (12.9)	34 (10.9)	
Others	42 (5.9)	14 (4.5)	
Co-morbidities (%)			
Nil	60 (8.4)	22 (7.1)	0.47
Hypertension	404 (56.7)	154 (49.5)	0.03
Diabetes	240 (33.7)	98 (31.5)	0.50
Heart disease	264 (37.0)	98 (31.5)	0.09
Cancer	48 (6.7)	31 (10.0)	0.07
Stroke	48 (6.7)	24 (7.7)	0.57
Respiratory disease	74 (10.4)	49 (15.8)	0.01
Renal disease	99 (13.9)	39 (12.5)	0.56

Table 2. Comparison of outcomes between the MEWS <4 and MEWS ≥4 groups

Variables	MEWS <4	MEWS≥4	*P* value
Outcomes (%)			
Died	47 (6.6)	53 (17.0)	<0.001
Admitted to ICU	137 (19.2)	64 (20.6)	0.61
Admitted to HD/ICA	267 (37.4)	86 (27.7)	0.00
Average LOS (SD)	6.97 (12.83)	7.75 (12.04)	0.36
LOS (median/Interquartile range)	4 (2, 7)	4 (1, 9)	0.99

LOS: Length of stay.

Table 3. Sensitivity and specificity of MEWS in predicting patient mortality

Variables	MEWS <2 (*n*=11)	MEWS <3 (*n*=23)	MEWS <4 (*n*=47)	MEWS <5 (*n*=66)	MEWS <6 (*n*=81)
Sensitivity (%)	11.00	23.00	47.00	66.00	81.00
Specificity (%)	61.47	44.80	27.92	12.34	5.84
PPV (%)	3.00	4.31	6.66	7.53	8.52
NPV (%)	86.45	84.32	82.96	77.03	73.97
ROC area	0.52	0.62	0.68	0.66	0.68

PPV: positive predictive value; NPV: negative predictive value; ROC: receiver operating characteristics.

When the individual derivatives of the MEWS score were analyzed for their predictive value for patient mortality, a systolic BP score of ≥1 had an odds ratio of 1.88, heart rate score of ≥2 had an odds ratio of 2.10, respiratory rate score of score ≥2 had an odds ratio of 3.44, and an AVPU score of score ≥1 had an odds ratio of 8.35. A temperature score of 2 had an odds ratio of 0.80 (Table 5).

The physiological parameters were used to calculate MEW score and its corresponding scoring criteria (Table 6).

DISCUSSION

Our study failed to meet our hypothesis that the MEWS would perform as well in our Asian population as in other non-Asian studies. In our study, we found the selected MEWS score of ≥4 predicted a higher rate of mortality compared to patients who presented with a MEWS score of <4 (*P* value <0.001). However, the MEWS had fairly low sensitivity and specificity compared with western populations in whom the sensitivity and specificity were 75% and 83% respectively.[2] For example, using a cut-off of MEWS <4 for risk stratification in our critically ill patients only had a sensitivity of 47.0% and a specificity of 27.9% for death. This indicates that there would be a number of 'missed' cases that might require ICU care but were assessed as low risk, and similarly a number of 'false positives' that might not need ICU care.

Risk stratification for mortality and other adverse outcomes play an essential role in the management of patients at the emergency department (ED).[7] Medical decisions for disposition as well as the level of intensive monitoring required rest on this perceived risk.[8] However any triage score used at the ED needs to balance the risks of over- and under-triage. The over-triage can be a problem

Table 4. Sensitivity and specificity of MEWS in predicting ICU/HD/ICA admission

Variables	MEWS <2 (*n*=174)	MEWS <3 (*n*=243)	MEWS <4 (*n*=333)	MEWS <5 (*n*=394)	MEWS <6 (*n*=420)
Sensitivity (%)	38.75	54.12	74.16	87.75	93.54
Specificity (%)	66.43	49.57	33.91	16.17	7.65
PPV (%)	47.41	45.59	46.70	44.98	44.16
NPV (%)	58.14	58.04	62.70	62.84	60.27
ROC area	0.47	0.46	0.49	0.47	0.46

PPV: positive predictive value; NPV: negative predictive value; ROC: receiver operating characteristics.

Table 5. Predictive values of individual MEWS score derivatives

Variables	Mortality (*N*=100)			ICU/HD/ICA Admission (*N*=449)		
	n	OR (95% *CI*)	*P* value	*n*	OR (95% *CI*)	*P* value
Systolic BP						
Score=3	3	1.21 (0.36–4.11)	0.76	12	1.10 (0.50–2.40)	0.81
Score≥2	18	1.56 (0.90–2.69)	0.11	62	1.16 (0.80–1.67)	0.44
Score≥1	37	1.88 (1.22–2.90)	0.00	102	0.80 (0.60–1.06)	0.12
Heart rate						
Score=3	15	1.42 (0.79–2.56)	0.24	35	0.51 (0.33–0.77)	0.00
Score≥2	43	2.10 (1.38–3.21)	0.00	108	0.70 (0.53–0.93)	0.01
Respiratory rate						
Score=3	6	2.39 (0.95–6.00)	0.06	12	0.85 (0.40–1.78)	0.67
Score≥2	37	3.49 (2.24–5.45)	<0.001	66	0.78 (0.56–1.09)	0.15
Temperature						
Score=2	5	0.80 (0.31–2.05)	0.64	22	0.69 (0.40–1.18)	0.17
Score=0	95	1.25 (0.49–3.19)	0.64	427	1.45 (0.85–2.48)	0.17
AVPU						
Score=3	14	8.68 (4.14–18.22)	<0.001	21	2.77 (1.29–5.95)	0.01
Score≥2	28	8.37 (4.89–14.33)	<0.001	41	1.96 (1.19–3.23)	0.01
Score≥1	35	8.35 (5.10–13.65)	<0.001	52	1.80 (1.16–2.78)	0.01

Table 6. Modified early warning score (MEWS)

Variables	MEWS						
	3	2	1	0	1	2	3
Systolic BP (mmHg)	≤70	71–80	81–100	101–199		≥200	
Heart rate (bpm)		≤40	41–50	51–100	101–110	111–129	≥130
Respiratory rate (bpm)	<9		9–14	15–20	21–29		≥30
Temperature (°C)	<35			35–38.4		38.5	
AVPU Score				Alert	Reacting to voice	Reacting to pain	Unresponsiveness

if patients who do not require intensive monitoring are allowed to consume such precious resources in a limited supply situation. In order to be useful at the ED, a high degree of accuracy is needed from such scores.

Rather than using specific cut-off values of the MEWS score as a predictor of poor outcome, it may be more prudent to monitor the MEWS score sequentially. Studying the MEWS score for a dynamic deterioration may be a better predictor of poor outcome. Conversely, monitoring the MEWS score for an improvement following any specific intervention may assist in judging the efficacy of the intervention and help guide any subsequent management.

Current triage systems are based on clinical judgment and traditional vital signs including heart rate, respiratory rate, blood pressure, temperature and pulse oximetry.[9] Unfortunately, vital signs have not been shown to correlate well with short or long-term clinical outcomes.[10]

One reason why the MEWS performed poorly in our population may be related to underlying disease condition. The diverse types of conditions in the patients recruited may have different effects on the MEWS scores. The main diagnosis groupings are cardiovascular, respiratory, neurological, gastrointestinal, renal, endocrine, infectious disease, vascular, trauma, cancer and others. In future studies, the different subgroups could possibly be analyzed separately.

Another limitation of this study is that we only included patients with a higher acuity of presentation, rather than the full spectrum of less acute patients who usually present to the ED. Nevertheless it is of concern that MEWS did not perform well in patients with a higher acuity. Also, this study was carried out in a single-center study at a tertiary teaching hospital in Singapore and the results may not be generalizable to other settings.

In the future, we are aiming to prospectively validate the MEWS score in a more representative sample of ED patients, including those of lesser acuity. We also plan to look at the effect of ongoing treatment on the serial MEWS score and how this relates to outcomes. In addition, we are looking to using new technology such as heart rate variability and machine learning techniques to improve prediction of adverse outcomes for critically ill patients. We believe this technology has the potential to become a useful bedside triage and monitoring tool for EDs, wards, ambulances, disaster situations and even for home monitoring. Clinical validation by comparing these new techniques with existing scores such as MEWS will be helpful.

In conclusion, the composite MEWS did not perform well in predicting poor patient outcomes for critically ill patients presenting to an ED.

REFERENCES

1 HealthGrades Quality Study. Patient Safety in American Hospitals: HealthGrades Inc; 2004 [cited July 2011]. Available from: http://www.healthgrades.com/media/english/pdf/hg_patient_safety_study_final.pdf.
2 Gardner-Thorpe J, Love N, Wrightson J, Walsh S, Keeling N. The value of Modified Early Warning Score (MEWS) in surgical in-patients: a prospective observational study. Ann Roy Coll Surg Engl 2006; 88: 571–575.
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 Paterson R, MacLeod DC, Thetford D, Beattie A, Graham C, Lam S, et al. Prediction of in-hospital mortality and length of stay using an early warning scoring system: clinical audit. Clin Med (London, England) 2006; 6: 281–284.
5 Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning Score in medical admissions. QJM 2001; 94: 521–526.
6 Time Series Population. Department of Statistics Singapore [cited July 2011]. Available from: http://www.singstat.gov.sg/stats/themes/people/hist/popn.html.
7 Goldman L. Using prediction models and cost-effectiveness analysis to improve clinical decisions: emergency department patients with acute chest pain. Proc Assoc Am Physicians 1995; 107: 329–333.
8 Peacock WF, Soto-Ruiz KM. Risk stratification for suspected acute coronary syndromes and heart failure in the emergency department. Acute Card Care 2009; 11: 138–145.
9 Gottschalk SB, Wood D, DeVries S, Wallis LA, Bruijns S. The Cape Triage Score: a new triage system South Africa. Proposal from the Cape Triage Group. Emerg Med J 2006; 23: 149–153.
10 Subbe CP, Slater A, Menon D, Gemmell L. Validation of physiological scoring systems in the accident and emergency department. Emerg Med J 2006; 23: 841–845.