Risk Stratification of Hospitalized Patients on the Wards

Matthew M. Churpek, MD, MPH; Trevor C. Yuen, BA; and Dana P. Edelson, MD

Patients who suffer adverse events on the wards, such as cardiac arrest and death, often have vital sign abnormalities hours before the event. Early warning scores have been developed with the aim of identifying clinical deterioration early and have been recommended by the National Institute for Health and Clinical Excellence. In this review, we discuss recently developed and validated risk scores for use on the general inpatient wards. In addition, we compare newly developed systems with more established risk scores such as the Modified Early Warning Score and the criteria used in the Medical Early Response Intervention and Therapy (MERIT) trial in our database of > 59,000 ward admissions. In general we found the single-parameter systems, such as the MERIT criteria, to have the lowest predictive accuracy for adverse events, whereas the aggregate weighted scoring systems had the highest. The Cardiac Arrest Risk Triage (CART) score was best for predicting cardiac arrest, ICU transfer, and a composite outcome (area under the receiver operating characteristic curve [AUC], 0.83, 0.77, and 0.78, respectively), whereas the Standardized Early Warning Score, VitalPAC Early Warning Score, and CART score were similar for predicting mortality (AUC, 0.88). Selection of a risk score for a hospital or health-care system should be guided by available variables, calculation method, and system resources. Once implemented, ensuring high levels of adherence and tying them to specific levels of interventions, such as activation of a rapid response team, are necessary to allow for the greatest potential to improve patient outcomes.

CHEST 2013; 143(6):1758–1765

Abbreviations: AUC = area under the receiver operating characteristic curve; CART = Cardiac Arrest Risk Triage; DNR = do not resuscitate; MERIT = Medical Early Response Intervention and Therapy; MEWS = Modified Early Warning Score; SEWS = Standardized Early Warning Score; ViEWS = VitalPAC Early Warning Score

Adverse events on the wards, such as cardiac arrest and death, are rarely sudden and are often heralded by abnormal vital signs hours before the event.^{1–5} However, these signs are often missed or not acted on appropriately, even in hospitals with mature rapid response systems.⁶ In 2007, the National Institute for Health and Clinical Excellence recommended that physiologic track and trigger systems should be used to monitor all adult patients in acute hospital settings.⁷ These systems, also known as early warning scores, typically

use vital sign thresholds to identify at-risk patients. Currently there are > 100 different published track and trigger systems,^{8–10} most of which are hospital-specific modifications of the original Early Warning Score,¹¹ developed using expert opinion, and have demonstrated variable levels of reliability, validity, and usefulness.¹² In the past few years, there has been a move toward scientifically derived risk scores and unifying the criteria across hospitals in some countries.^{13–15}

In this review, we discuss risk scores for use on the general inpatient wards, with a focus on those that were developed recently. In addition, we compare newly developed systems with more established risk scores, such as the Modified Early Warning Score (MEWS)¹⁶ and the criteria used in the Medical Early Response Intervention and Therapy (MERIT) trial¹⁷ in our database of > 59,000 ward admissions. We limit our discussion to objective vital sign-based risk scores designed for adult patients on the general hospital wards, but it should be noted that much progress has been made

Manuscript received June 26, 2012; revision accepted October 15, 2012.

Affiliations: From the Section of Pulmonary and Critical Care (Dr Churpek), the Health Studies Department (Dr Churpek), and the Section of Hospital Medicine (Mr Yuen and Dr Edelson), University of Chicago, Chicago, IL.

Correspondence to: Dana Edelson, MD, 5841 S Maryland Ave, MC 5000, Chicago, IL 60637; e-mail: dperes@uchicago.edu

© 2013 American College of Chest Physicians. Reproduction of this article is prohibited without written permission from the American College of Chest Physicians. See online for more details.
DOI: 10.1378/chest.12-1605

in pediatric risk scores and in the development of systems for specific patient populations.¹⁸⁻²¹

RECENTLY DEVELOPED RISK
SCORES FOR WARD PATIENTS

Physiologic track and trigger systems are often divided into single-parameter, multiple-parameter, and aggregate weighted systems. There have been recent developments in each of these categories.⁸⁻¹⁰

Single-Parameter Systems

A single-parameter system is composed of a list of individual physiologic criteria that, if reached by a particular patient, trigger a response. Since any one abnormality on the list will trigger the response, these systems are the easiest to implement, requiring no score calculation. The first such system was developed in the early 1990s in Liverpool, Australia and included vital signs, laboratory values, and specific conditions such as new arrhythmia and amniotic fluid embolism.²² The MERIT trial used a variation of these criteria, and it remains the most commonly described of the single-parameter tools today.¹⁷ Both of these early systems were developed using expert opinion. However, recently Cretikos and colleagues²³ used a case-control design to develop an evidence-based modification to the MERIT criteria using data from the control arm of the trial. This model (respiratory rate ≥ 28 breaths/min, heart rate ≥ 140 beats/min, systolic BP ≤ 85 mm Hg, or a decrease in Glasgow Coma Scale score of > 2 points), had a sensitivity of 59.6% and specificity of 93.7% for predicting the composite outcome of cardiac arrest, death, or ICU transfer, compared with 50.4% sensitivity and 93.3% specificity for the original criteria.

Multiple-Parameter Systems

Multiple-parameter systems use combinations of different physiologic criteria, without calculation of a score, to activate the rapid response system. These systems are the least commonly described but have the advantage of allowing for risk stratification and a graded response, without requiring a complex calculation. A recent example of this type of system was developed by Bleyer and colleagues²⁴ using vital sign cutoffs that

were individually associated with at least 5% in-hospital mortality. The critical values they identified were a systolic BP < 85 mm Hg, heart rate > 120 beats/min, temperature $< 35^{\circ}\text{C}$ or $> 38.9^{\circ}\text{C}$, oxygen saturation $< 91\%$, respiratory rate < 13 or > 23 breaths/min, and level of consciousness recorded as anything but “alert.” They assigned one point for each critical value and found that having three simultaneous critical values was associated with 23.6% in-hospital mortality. In addition, the authors compared their score to two aggregate weighted risk scores, the MEWS (Table 1)¹⁶ and the VitalPAC Early Warning Score (ViEWS) (Table 2),²⁵ and found similar accuracy for detecting in-hospital mortality (areas under the receiver operating characteristic curves [AUCs] of 0.85, 0.87, and 0.86, respectively).

Aggregate Weighted Systems

Aggregate weighted scoring systems are the most complex of the early warning scores. They categorize vital signs and other variables into different degrees of physiologic abnormality and then assign point values for each category. They have the advantage of allowing for risk stratification of patients and responses based on severity level but can be error prone when calculated manually.^{26,27} Most published systems are expert opinion-based variations of the original Early Warning Score,¹² including the well-described MEWS and the Standardized Early Warning Score (SEWS) (Table 3).^{16,28}

One of the most recently developed aggregate weighted risk scores is the ViEWS.²⁵ The ViEWS is similar to previously published early warning scores and includes heart rate, respiratory rate, temperature, and systolic BP but adds oxygen saturation and the use of supplemental oxygen (Table 2). In addition, the weighting for different vital sign abnormalities was adjusted based on the investigators’ analyses, prior literature, and trial and error.²⁵ A study in the dataset from which the ViEWS was derived found that it outperformed 33 other risk scores in a cohort of 35,585 acute medical patients for the outcome of death within 24 h of a ward vital sign set. In addition, an abbreviated ViEWS, without mental status, was externally validated in a separate study in a Canadian hospital, where it was found to have an AUC of 0.93 for mortality and

Table 1—Modified Early Warning Score¹⁶

Score	3	2	1	0	1	2	3
Respiratory rate, breaths/min	...	< 9	...	9-14	15-20	21-29	> 29
Heart rate, beats/min	...	< 40	41-50	51-100	101-110	111-129	> 129
Systolic BP, mm Hg	< 70	71-80	81-100	101-199	...	> 199	...
Temperature, $^{\circ}\text{C}$	...	< 35	...	35-38.4	...	> 38.4	...
Neurologic	...	...	...	Alert	Voice	Pain	Unresp

Unresp = unresponsive.

Table 2—VitalPAC Early Warning Score²⁵

Score	3	2	1	0	1	2	3
Respiratory rate, breaths/min	<9	...	9-11	12-20	...	21-24	> 24
Oxygen saturation, %	<92	92-93	94-95	96-100	...	...	...
Supplemental oxygen use	...	...	...	No	...	...	Yes (any oxygen)
Heart rate, beats/min	...	<41	41-50	51-90	91-110	111-130	> 130
Systolic BP, mm Hg	<91	91-100	101-110	111-249	> 249	...	...
Temperature, °C	<35.1	...	35.1-36	36.1-38	38.1-39	> 39	...
Neurologic	...	...	...	Alert	...	...	Voice
	...	...	...	...	...	...	Pain
	...	...	...	...	...	...	Unresp

See Table 1 legend for expansion of abbreviation.

0.72 for ICU transfer within 48 h using admission vital sign data.²⁹

Another aggregate weighted scoring system was developed by Tarassenko and colleagues³⁰ using continuous vital sign data in high-risk patients on the wards or step-down units. They derived a scoring system based on the distributions of respiratory rate, heart rate, systolic BP, and oxygen saturation in their dataset. Although they did not evaluate the accuracy of their system for detecting adverse outcomes, they are currently conducting a clinical trial using their system on the trauma wards at a teaching hospital.³⁰

Our group published an aggregate weighted system known as the Cardiac Arrest Risk Triage (CART) score (Table 4).³¹ The CART score was derived using logistic regression to detect in-hospital cardiac arrest and was validated for detecting ward-to-ICU transfers. In that study, the CART score outperformed the MEWS for detecting both cardiac arrest (AUC, 0.84 vs 0.76; $P = .001$) and ICU transfer (AUC, 0.71 vs 0.67; $P < .001$). However, it was not compared with other early warning scores and has not yet been validated using external data.

There are several other recent additions to the literature. These include the Worthing physiologic scoring system, which was statistically derived using admission data.³² When applied to a validation dataset, the system had an AUC of 0.72 for the outcome of in-hospital mortality. A disadvantage of this system is that it requires the measurement of oxygen saturation on room air, which is not always collected. In addition, a form of the

MEWS was recently introduced in The Netherlands that includes nurse worry and urine output and was validated in a population of surgical patients using the highest score achieved for each patient during that hospitalization.¹⁵ A score ≥ 3 had a sensitivity of 74% and specificity of 82% for a composite outcome that included mortality, ICU transfer, and severe surgical complications. In another study, Kho and colleagues³³ developed a risk score by altering the vital sign weightings in the MEWS, adding BMI and age and removing mental status, based on prior literature and a review of calls to their hospital's rapid response team. Using the maximum score for each patient on the wards, their model had an AUC of 0.72 for the combined outcome of code team activation, cardiac arrest, or ICU transfer. In addition, Cuthbertson and colleagues^{34,35} used discriminant function analysis in two case-control studies, one in medical and another in surgical patients. They used the highest, lowest, and median vital sign values for patients in the 48-h period before transfer to the ICU (cases) or to a lower-acuity unit (control subjects). The resulting models were at least as accurate as the other published risk scores they compared their models to in both studies. Finally, using electronic medical record data, investigators from Kaiser Permanente developed a 24-variable model that included vital signs, laboratory values, severity of illness scores, and longitudinal chronic illness burden scores.³⁶ Their model had an AUC of 0.78 in a validation dataset for detecting a combined end point of ICU transfer or death outside the ICU.

Table 3—Standardized Early Warning Score²⁸

Score	3	2	1	0	1	2	3
Respiratory rate, breaths/min	< 9	...	...	9-20	21-30	31-35	> 35
Oxygen saturation, %	< 85	85-89	90-92	93-100	...	...	...
Heart rate, beats/min	< 30	30-39	40-49	50-99	100-109	110-129	> 129
Systolic BP, mm Hg	< 70	70-79	80-99	100-199	...	> 199	...
Temperature, °C	< 34	34-34.9	35-35.9	36-37.9	38-38.9	> 38.9	...
Neurologic	...	...	...	Alert	Voice	Pain	Unresp

See Table 1 legend for expansion of abbreviation.

Table 4—Cardiac Arrest Risk Triage Score³¹

Vital Sign	Score
Respiratory rate, breaths/min	
< 21	0
21-23	8
24-25	12
26-29	15
> 29	22
Heart rate, beats/min	
< 110	0
110-139	4
> 139	13
Diastolic BP, mm Hg	
> 49	0
40-49	4
35-39	6
< 35	13
Age, y	
< 55	0
55-69	4
> 69	9

CHOOSING THE OPTIMAL SCORING SYSTEM

Introduction

The accuracy of the system has strong implications for long-term success, both in terms of identifying more cases (ie, sensitivity) to minimize adverse events and preventing false alarms (ie, specificity) to minimize resource expenditure and alarm fatigue. For example, consider a hospital that has 20,000 admissions per year and 1,000 events (ICU transfers, deaths, and ward cardiac arrests). An improvement in sensitivity of 5%, at the same specificity level, would result in the detection of an additional 50 adverse events per year (1,000 events multiplied by the difference in sensitivity of 5%). In addition, an improvement in specificity of 5%, at the same level of sensitivity, would result in 950 fewer “false alarms” per year (19,000 admissions who did not experience the event multiplied by the difference in specificity of 5%). Multiplying these results over many hospitals across the country would result in a considerable public health benefit and illustrates the importance of efforts to improve the accuracy of early warning scores.

Methods

Because the systems described previously represent recent advances, many have not been directly compared with one another in the same dataset. Therefore, we used our database of ward admissions from November 2008 until August 2011 consisting of both medical and surgical patients to compare the different early warning scores described previously. Our patient population has been previously described; in brief, it consists of all patients on the ward from a single urban academic center in the United States.³¹ Our hospital

has had a rapid response team in place since 2008, which is led by a critical care nurse and is separate from the team that responds to cardiac arrests. Respiratory therapy also responds to team activations and a hospitalist attending physician and/or pharmacist are available upon request.

We compared recently developed or validated single-parameter (ie, Cretikos et al²³), multiple-parameter (ie, Bleyer et al²⁴), and aggregate weighted (ie, Tarassenko et al,³⁰ VIEWS, SEWS, and the CART score^{25,28,31}) risk scores to previously validated systems that are commonly used (ie, the MERIT criteria and MEWS).^{16,17} We were unable to compare all the recently developed systems because room air oxygen saturation, BMI, urine output, and accurate determination of patients admitted to the surgical service were not available. In addition, some aspects of the MERIT criteria were not captured in our dataset, such as a drop in Glasgow Coma Scale score, the presence of seizures, and airway emergencies.

Ward vital signs were extracted from our electronic health record (EPIC), and each of the previously mentioned early warning scores was calculated for every simultaneous ward vital sign set in the entire dataset. If a vital sign necessary for score calculation was missing, then the most recent value was carried forward. In addition, if there were no previous values, then a median value was imputed. Cardiac arrest was determined using a prospectively validated quality improvement database, and ICU transfer and mortality were determined using administrative databases.

Accuracy was calculated using the AUC, sensitivity, and specificity for detecting cardiac arrest, ICU transfer, mortality, and a composite outcome of any of these events using each patient’s highest score prior to the event or during their entire admission for those who did not experience an event. Ward patients transferred to the ICU from the operating room were not counted as an ICU transfer event. Of note, the CART score was developed to detect cardiac arrest using an older version of these data that account for approximately 80% of the patients in this updated dataset.

Results

During the study period, there were 59,643 admissions with ward vital signs, including 109 ward cardiac arrests, 291 deaths within 24 h of a ward vital sign, and 2,655 ward-to-ICU transfers. The included patients had a mean age of 55 ± 18 years; 56% were women, 43% were black, 36% were white, and 34% underwent surgery during the hospitalization. Results from the early warning score comparisons are shown in Table 5, separated by outcome. We found a wide range of accuracy, both across outcomes for a given system and across systems. In general, mortality resulted in the

Table 5—Accuracy of Track and Trigger Systems for Different Outcomes

Track and Trigger System	Cardiac Arrest	ICU Transfer	Mortality	Composite
MERIT	0.63 (0.59-0.68)	0.64 (0.63-0.65)	0.74 (0.71-0.76)	0.64 (0.64-0.65)
Modified MERIT	0.69 (0.65-0.74)	0.69 (0.68-0.70)	0.79 (0.76-0.81)	0.70 (0.69-0.70)
Multiple parameter, from Bleyer et al ²⁴	0.73 (0.68-0.78)	0.72 (0.71-0.73)	0.84 (0.82-0.87)	0.73 (0.72-0.74)
Centile-based, from Tarassenko et al ³⁰	0.70 (0.65-0.76)	0.71 (0.69-0.72)	0.83 (0.80-0.86)	0.72 (0.70-0.73)
MEWS	0.76 (0.71-0.81)	0.74 (0.73-0.75)	0.87 (0.84-0.89)	0.75 (0.74-0.76)
SEWS	0.76 (0.71-0.81)	0.75 (0.74-0.76)	0.88 (0.86-0.90)	0.76 (0.75-0.77)
ViEWS	0.77 (0.72-0.82)	0.73 (0.72-0.75)	0.88 (0.86-0.91)	0.75 (0.74-0.76)
CART score	0.83 (0.79-0.86)	0.77 (0.76-0.78)	0.88 (0.86-0.90)	0.78 (0.77-0.79)

Data are shown as area under the receiver operating characteristic curve (95% CI). CART = Cardiac Arrest Risk Triage; MERIT = Medical Early Response Intervention and Therapy; MEWS = Modified Early Warning Score; SEWS = Standardized Early Warning Score; ViEWS = VitalPAC Early Warning Score.

highest AUCs, whereas ICU transfer resulted in the lowest. Overall, the aggregate weighted scoring systems outperformed the other systems for most outcomes, with the SEWS, MEWS, ViEWS, and CART score being the most accurate. In addition, the modified MERIT criteria described by Cretikos and colleagues²³ were more accurate than the original MERIT criteria for all outcomes. Although the ViEWS, CART, MEWS, and SEWS were similar in performance across the outcomes, the CART score had the highest AUC for cardiac arrest (0.83), ICU transfer (0.77), and the composite outcome (0.78), whereas the CART score, ViEWS, and SEWS all had the same AUC for mortality (0.88). The ViEWS was the second most accurate system for detecting cardiac arrest (0.77), and the SEWS was the second most accurate for ICU transfer (AUC 0.75) and the composite outcome (AUC, 0.76). Since the CART score was derived using many of the patients from this dataset, we repeated the analysis for the CART score using only those patients not in the original study (ie, prospective validation) and found similar results (AUCs of 0.86, 0.76, 0.87, and 0.77 for cardiac arrest, ICU transfer, mortality, and the composite outcome, respectively). Sensitivity and specificity values at cut-points closest to 85%, 90%, and 95% specificity for the four most accurate systems for detecting cardiac arrest are shown in Table 6. At a specificity threshold of approximately 90%, the CART score had a sensitivity of 49%, compared with the ViEWS (41%), MEWS (39%), and the centile-based system (35%, data not shown). The SEWS and the multiple-parameter system by Bleyer and colleagues²⁴ did not have cut-offs near this level of specificity. The MERIT criteria had a sensitivity and specificity of 45% and 82% for detecting cardiac arrest compared with the modified criteria proposed by Cretikos et al,²³ which had a sensitivity of 54% and specificity of 84%.

IMPLEMENTATION OF EARLY WARNING SCORES

An important part of the process of implementing a physiologic track and trigger system, especially an

aggregate weighted system, is determining how the score will be calculated. Options include calculation by hand, using a calculator or handheld device developed specifically for the scoring system, and using the electronic medical record. Manual calculation, with or without a calculator, is the most commonly used method and the least expensive to implement. However, studies suggest that calculation errors are not uncommon.^{26,27} Preprogrammed applications decrease these errors but still require manual entering of the data, which can be redundant to workflow and error prone in its own right.²⁷ Completely automated systems, such as those integrated into electronic medical records, are likely to be the least labor intensive from the clinician standpoint and least error prone. Moreover, they have the potential to incorporate other patient data, such as demographic characteristics, location, and laboratory values and even be tied into automated notification systems.^{33,36} However, these systems require institutional resources to implement and may not be an option for most hospitals, especially those with paper-based medical records.

CURRENT CONTROVERSIES IN THE FIELD

Outcome Selection for Development and Validation Studies

As demonstrated in the results of our comparison study, the reported accuracy of a scoring system will depend in large part on the outcome chosen for validation. Therefore, comparison of systems requires a consistency in outcome selection. Although mortality is the easiest to predict, as evidenced by the higher AUCs, it may not be the most useful, since many deaths in the hospital are fully expected, and detecting those events is neither necessary nor helpful. However, most studies of early warning scores have not omitted do not resuscitate (DNR) patients from their analyses. Reasons for this may include difficulty in identifying such patients in large datasets, the inability to determine the exact time when the goals in patient care transition from life-saving (when early warning scores

Table 6—Sensitivity and Specificity at Selected Cut Points for the Most Accurate Track and Trigger Systems for Identifying Patients With Cardiac Arrest

Track and Trigger System Cut Point	Sensitivity, %	Specificity, %
SEWS		
> 3	55	85
> 4	38	94
> 5	19	97
MEWS		
> 3	67	80
> 4	39	91
> 5	20	96
ViEWS		
> 8	60	83
> 9	41	91
> 10	29	95
CART		
> 16	61	84
> 20	49	90
> 24	35	95

See Table 5 legend for expansion of abbreviations.

might be beneficial) to comfort (when risk scores would not be useful), the belief that some DNR patients still desire other life-saving interventions, and previous studies suggesting that rapid response teams can improve some aspects of end-of-life care.³⁷ ICU transfer represents the most common outcome and thus results in the highest statistical power but is the least generalizable, given the heterogeneous criteria for admission, resulting in the lowest AUCs. Moreover, it is, by definition, an event already recognized by hospital staff, albeit late on some occasions. In addition, criteria for ICU admission in some hospitals may include vital sign cutoffs, and so scores derived or validated in such hospitals would be affected by these criteria. Similar to mortality, cardiac arrest has the benefit of being objectively defined. However, unlike mortality, it is always worth identifying and preventing, if only to institute a DNR order in some cases. As such, a cardiac arrest on the floor always represents a failure of the current system and thus may be the most clinically relevant of the outcomes to use for derivation and validation. However, reporting all four outcomes in future studies would allow readers to draw their own conclusions about relevant outcomes for their own practice.

Some authors have stated that the original early warning scores were not developed to be highly accurate predictors of any specific outcome due to the many confounding events that occur during a hospitalization.³⁸ In addition, as described previously, investigators have also used the distribution of vital signs to determine cutoffs for risk scores, arguing that deriving models based on outcomes, when used prospectively, will disadvantage those patients who would have been previously “salvaged” by the vital sign monitoring sys-

tem that had been previously in use in the development dataset.³⁰

Inclusion of Age in Risk Scores

Older age is a known risk factor for cardiac arrest and death and is often included in risk scores used in the ICU. However, age is only rarely included in current risk scores for ward patients.³⁹ Several studies have shown that the increased risk of adverse outcomes associated with increased age is independent of vital sign derangements, and the inclusion of age has been shown to improve the accuracy of risk scores, although to varying degrees.^{16,24,31,32,39} Some concerns raised about including age in early warning scores are ethical in nature,³⁹ and including age could make it less likely for younger patients with vital sign abnormalities to be identified. It is unknown to what degree this would occur, given that cardiac arrest and death are much more common in older age groups. In addition, it is unknown whether vital signs prior to these events differ between younger and older patients, given the increased use of β -blockers in older patients and the physiologic changes that occur with aging.

Impact on Clinical Outcomes

The most important unanswered question is whether early warning scores improve outcomes. To definitively answer this question, a large randomized trial would be needed that used a well-validated risk score, and even then separating the effects of the specific risk score used in the study from the intervention would be difficult. Many before-and-after studies have been published highlighting the usefulness of early warning scores (usually concurrently implemented with rapid response systems), including improved vital sign documentation^{40,41} and improved patient outcomes.^{37,42–44} However, these findings have not been universal,^{45,46} and it is currently unclear whether early warning scores improve important patient outcomes, such as hospital-wide cardiac arrest and mortality rates, or decrease costs. Importantly, delayed response has been identified as one of the strongest predictors of mortality and unexpected ICU transfer in patients evaluated by the rapid response team.^{47,48} Future efforts to improve the evidence base for early warning scores are needed, as are methods to improve adherence to vital sign documentation and rapid response system activation after they are implemented. Automated implementation with built-in notification, as described previously, is likely to help, but this remains to be demonstrated, and the cost-effectiveness needs to be studied. Finally, pairing the different risk strata to specific levels of interventions, such as increased monitoring, consultation by the ICU team, and automatic calls to the rapid response team, is essential, and different workflows have been proposed.⁸

CONCLUSIONS

The number of published risk scores for ward patients has grown rapidly over the past decade, in large part due to the popularity of rapid response systems. We found that aggregate weighted early warning scores, specifically the ViEWS, SEWS, MEWS, and CART score, were more accurate than other types of scoring systems for detecting cardiac arrest, mortality, ICU transfer, and a composite outcome in our dataset. Hospitals seeking to implement an aggregate weighted scoring system should consider the available variables, possible calculation methods, and system resources when selecting the appropriate tool for their setting.

ACKNOWLEDGMENTS

Financial/nonfinancial disclosures: The authors have reported to *CHEST* the following conflicts of interest: Dr Edelson is supported by a career development award from the National Heart, Lung, and Blood Institute [K23 HL097157-01] and receives research support from Koninklijke Philips Electronics N.V. (Andover, MA). Dr Churpek is supported by a National Institutes of Health grant [T32 HL 07605]. Mr Yuen has reported that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

Other contributions: We thank Jesse Hall, MD; David Meltzer, MD, PhD; Nicole Babusko, BS; Krista Scheffey, BA; Don Saner, MS; and Contessa Tsu, AAS, for their support and guidance.

REFERENCES

- Berlot G, Pangher A, Petrucci L, Bussani R, Lucangelo U. Anticipating events of in-hospital cardiac arrest. *Eur J Emerg Med*. 2004;11(1):24-28.
- Hodgetts TJ, Kenward G, Vlackonikolis I, et al. Incidence, location and reasons for avoidable in-hospital cardiac arrest in a district general hospital. *Resuscitation*. 2002;54(2):115-123.
- Schein RM, Hazday N, Pena M, Ruben BH, Sprung CL. Clinical antecedents to in-hospital cardiopulmonary arrest. *Chest*. 1990;98(6):1388-1392.
- Kause J, Smith G, Prytherch D, Parr M, Flabouris A, Hillman K; Intensive Care Society (UK); Australian and New Zealand Intensive Care Society Clinical Trials Group. 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(3):275-282.
- Churpek MM, Yuen TC, Huber MT, Park SY, Hall JB, Edelson DP. Predicting cardiac arrest on the wards: a nested case-control study. *Chest*. 2012;141(5):1170-1176.
- Gallotra S, DeVita MA, Simmons RL, Dew MA; Members of the Medical Emergency Response Improvement Team (MERIT) Committee. Mature rapid response system and potentially avoidable cardiopulmonary arrests in hospital. *Qual Saf Health Care*. 2007;16(4):260-265.
- National Institute of Health and Clinical Excellence. *Acutely Ill Patients in Hospital: Recognition of and Response to Acute Illness in Adults in Hospital*. London, England: National Institute of Health and Clinical Excellence; 2007. NICE Clinical Guideline No. 50.
- Jansen JO, Cuthbertson BH. Detecting critical illness outside the ICU: the role of track and trigger systems. *Curr Opin Crit Care*. 2010;16(3):184-190.

- Smith GB, Prytherch DR, Schmidt PE, Featherstone PI. Review and performance evaluation of aggregate weighted 'track and trigger' systems. *Resuscitation*. 2008;77(2):170-179.
- Smith GB, Prytherch DR, Schmidt PE, Featherstone PI, Higgins B. A review, and performance evaluation, of single-parameter "track and trigger" systems. *Resuscitation*. 2008;79(1):11-21.
- Morgan RM, Williams F, Wright MM. An early warning scoring system for detecting developing critical illness. *Clin Intensive Care*. 1997;8(2):100.
- Gao H, McDonnell A, Harrison DA, 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(4):667-679.
- Subbe CP. Better ViEWS ahead? It is high time to improve patient safety by standardizing early warning scores. *Resuscitation*. 2010;81(8):923-924.
- Jenkins PF, Thompson CH, Barton LL. Clinical deterioration in the condition of patients with acute medical illness in Australian hospitals: improving detection and response. *Med J Aust*. 2011;194(11):596-598.
- Smith T, Den Hartog D, Moerman T, Patka P, Van Lieshout EM, Schep NW. Accuracy of an expanded early warning score for patients in general and trauma surgery wards. *Br J Surg*. 2012;99(2):192-197.
- Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning Score in medical admissions. *QJM*. 2001;94(10):521-526.
- Hillman K, Chen J, Cretikos M, et al; MERIT study investigators. Introduction of the medical emergency team (MET) system: a cluster-randomised controlled trial. *Lancet*. 2005;365(9477):2091-2097.
- Taenzer AH, Pyke JB, McGrath SP, Blike GT. Impact of pulse oximetry surveillance on rescue events and intensive care unit transfers: a before-and-after concurrence study. *Anesthesiology*. 2010;112(2):282-287.
- Akre M, Finkelstein M, Erickson M, Liu M, Vanderbilt L, Billman G. Sensitivity of the pediatric early warning score to identify patient deterioration. *Pediatrics*. 2010;125(4):e763-e769.
- Skaletsky SM, Raszynski A, Totapally BR. Validation of a modified pediatric early warning system score: a retrospective case-control study. *Clin Pediatr (Phila)*. 2012;51(5):431-435.
- Hravnak M, Devita MA, Clontz A, Edwards L, Valenta C, Pinsky MR. Cardiorespiratory instability before and after implementing an integrated monitoring system. *Crit Care Med*. 2011;39(1):65-72.
- Lee A, Bishop G, Hillman KM, Daffurn K. The Medical Emergency Team. *Anaesth Intensive Care*. 1995;23(2):183-186.
- Cretikos M, Chen J, Hillman K, Bellomo R, Finfer S, Flabouris A; MERIT study investigators. The objective medical emergency team activation criteria: a case-control study. *Resuscitation*. 2007;73(1):62-72.
- Bleyer AJ, Vidya S, Russell GB, et al. Longitudinal analysis of one million vital signs in patients in an academic medical center. *Resuscitation*. 2011;82(11):1387-1392.
- Prytherch DR, Smith GB, Schmidt PE, Featherstone PI. ViEWS—Towards a national early warning score for detecting adult inpatient deterioration. *Resuscitation*. 2010;81(8):932-937.
- 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(4):619-624.
- Prytherch DR, Smith GB, Schmidt P, et al. Calculating early warning scores—a classroom comparison of pen and paper and hand-held computer methods. *Resuscitation*. 2006;70(2):173-178.

28. 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):281-284.
29. Kellett J, Kim A. Validation of an abbreviated Vitalpac™ Early Warning Score (ViEWS) in 75,419 consecutive admissions to a Canadian regional hospital. *Resuscitation*. 2012;83(3):297-302.
30. Tarassenko L, Clifton DA, Pinsky MR, Hravnak MT, Woods JR, Watkinson PJ. Centile-based early warning scores derived from statistical distributions of vital signs. *Resuscitation*. 2011;82(8):1013-1018.
31. Churpek MM, Yuen TC, Park SY, Meltzer DO, Hall JB, Edelson DP. Derivation of a cardiac arrest prediction model using ward vital signs. *Crit Care Med*. 2012;40(7):2102-2108.
32. 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 Anaesth*. 2007;98(6):769-774.
33. Kho A, Rotz D, Alrahi K, et al. Utility of commonly captured data from an EHR to identify hospitalized patients at risk for clinical deterioration. *AMIA Annu Symp Proc*. 2007;2007:404-408.
34. Cuthbertson BH, Boroujerdi M, McKie L, Aucott L, Prescott G. Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? *Crit Care Med*. 2007;35(2):402-409.
35. 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(1):19-25.
36. Escobar GJ, LaGuardia JC, Turk BJ, Ragins A, Kipnis P, Draper D. Early detection of impending physiologic deterioration among patients who are not in intensive care: development of predictive models using data from an automated electronic medical record. *J Hosp Med*. 2012;7(5):388-395.
37. Vazquez R, Gheorghe C, Grigoriyan A, Palvinskaya T, Amoateng-Adjepong Y, Manthous CA. Enhanced end-of-life care associated with deploying a rapid response team: a pilot study. *J Hosp Med*. 2009;4(7):449-452.
38. Morgan RJ, Wright MM. In defence of early warning scores. *Br J Anaesth*. 2007;99(5):747-748.
39. Smith GB, Prytherch DR, Schmidt PE, et al. Should age be included as a component of track and trigger systems used to identify sick adult patients? *Resuscitation*. 2008;78(2):109-115.
40. Odell M, Rechner IJ, Kapila A, et al. The effect of a critical care outreach service and an early warning scoring system on respiratory rate recording on the general wards. *Resuscitation*. 2007;74(3):470-475.
41. McBride J, Knight D, Piper J, Smith GB. Long-term effect of introducing an early warning score on respiratory rate charting on general wards. *Resuscitation*. 2005;65(1):41-44.
42. Campello G, Granja C, Carvalho F, Dias C, Azevedo LF, Costa-Pereira A. Immediate and long-term impact of medical emergency teams on cardiac arrest prevalence and mortality: a plea for periodic basic life-support training programs. *Crit Care Med*. 2009;37(12):3054-3061.
43. Mitchell IA, McKay H, Van Leuvan C, et al. A prospective controlled trial of the effect of a multi-faceted intervention on early recognition and intervention in deteriorating hospital patients. *Resuscitation*. 2010;81(6):658-666.
44. Moon A, Cosgrove JF, Lea D, Fairs A, Cressey DM. An eight year audit before and after the introduction of modified early warning score (MEWS) charts, of patients admitted to a tertiary referral intensive care unit after CPR. *Resuscitation*. 2011;82(2):150-154.
45. Subbe CP, Davies RG, Williams E, Rutherford P, Gemmell L. Effect of introducing the Modified Early Warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions. *Anaesthesia*. 2003;58(8):797-802.
46. Chan PS, Jain R, Nallmothu BK, Berg RA, Sasson C. Rapid response teams: a systematic review and meta-analysis. *Arch Intern Med*. 2010;170(1):18-26.
47. Calzavacca P, Licari E, Tee A, et al. The impact of Rapid Response System on delayed emergency team activation patient characteristics and outcomes—a follow-up study. *Resuscitation*. 2010;81(1):31-35.
48. Calzavacca P, Licari E, Tee A, et al. A prospective study of factors influencing the outcome of patients after a Medical Emergency Team review. *Intensive Care Med*. 2008;34(11):2112-2116.