Resuscitation 85 (2014) 549–552

Contents lists available at ScienceDirect

Resuscitation

journal homepage: www.elsevier.com/locate/resuscitation

Short communication

Widely used track and trigger scores: Are they ready for automation in practice?☆,☆☆

Santiago Romero-Brufau[a], Jeanne M. Huddleston[a,b,c,*], James M. Naessens[a,c], Matthew G. Johnson[a,c], Joel Hickman[a,c], Bruce W. Morlan[a,c], Jeffrey B. Jensen[d], Sean M. Caples[e], Jennifer L. Elmer[f], Julie A. Schmidt[f], Timothy I. Morgenthaler[e], Paula J. Santrach[g]

[a] Mayo Clinic Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, 200 First Street SW, Rochester, MN 55905, United States
[b] Division of Hospital Internal Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, United States
[c] Division of Health Care Policy and Research, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, United States
[d] Department of Anesthesiology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, United States
[e] Division of Pulmonary Medicine and Critical Care, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, United States
[f] Department of Nursing, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, United States
[g] Department of Laboratory Medicine and Pathology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, United States

ARTICLE INFO

Article history:
Received 6 August 2013
Received in revised form
10 December 2013
Accepted 11 December 2013

Keywords:
Early warning score
Track-and-trigger
Rapid response system
Automation
Patient safety
Cardiorespiratory arrests

ABSTRACT

Introduction: Early Warning Scores (EWS) are widely used for early recognition of patient deterioration. Automated alarm/alerts have been recommended as a desirable characteristic for detection systems of patient deterioration. We undertook a comparative analysis of performance characteristics of common EWS methods to assess how they would function if automated.
Methods: We evaluated the most widely used EWS systems (MEWS, SEWS, GMEWS, Worthing, ViEWS and NEWS) and the Rapid Response Team (RRT) activation criteria in use in our institution. We compared their ability to predict the composite outcome of Resuscitation call, RRS activation or unplanned transfer to the ICU, in a time-dependent manner (3, 8, 12, 24 and 36 h after the observation) by determining the sensitivity, specificity and positive predictive values (PPV). We used a large vital signs database (6,948,689 unique time points) from 34,898 unique consecutive hospitalized patients.
Results: PPVs ranged from less than 0.01 (Worthing, 3 h) to 0.21 (GMEWS, 36 h). Sensitivity ranged from 0.07 (GMEWS, 3 h) to 0.75 (ViEWS, 36 h). Used in an automated fashion, these would correspond to 1040–215,020 false positive alerts per year.
Conclusions: When the evaluation is performed in a time-sensitive manner, the most widely used weighted track-and-trigger scores do not offer good predictive capabilities for use as criteria for an automated alarm system. For the implementation of an automated alarm system, better criteria need to be developed and validated before implementation.

1. Introduction

Multiple studies demonstrate that abnormal vital signs precede critical events such as death, cardiorespiratory arrest (CRA) or unplanned transfer to the intensive care unit (ICU).[1–4] The concept of a Rapid Response Team was developed to act upon early signs of deterioration before a critical event developed.[5] The concept evolved to the idea of implementing a system of care (the Rapid Response System – RRS) rather than merely a team.[6] These *systems* are comprised of an *afferent limb* (which can be thought of as the 'sensing' structure, responsible for detecting deteriorating patients and activating assistance), an *efferent limb* (usually a team that responds to the deterioration events), and *administrative and data analysis limb* that ensures continuous assessment and improvement of the system.[6]

Several previous reviews focus on early warning scores that could be used in afferent limbs, both single-parameter[7] and weighted track-and-trigger systems.[8] These papers analyzed the

☆ A Spanish translated version of the summary of this article appears as Appendix in the final online version at http://dx.doi.org/10.1016/j.resuscitation.2013.12.017.

* Corresponding author at: Center for the Science of Health Care Delivery, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, United States.
E-mail address: huddleston.jeanne@mayo.edu (J.M. Huddleston).

0300-9572/$ – see front matter © 2014 The Authors. Published by Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.resuscitation.2013.12.017

capacity of the criteria to predict death during hospitalization. In a recent consensus conference about the afferent limb of Rapid Response Systems, “automated alert/alarm” was identified as one of the desirable characteristics of electronic monitoring systems.[9] Use of early warning scores to implement automated alarm systems requires understanding of the time-dependent nature of the alerts and the resulting automatically triggered workload.

Our aim was to compare the relative performance of the more commonly used[9–11] physiologically based early warning scores (MEWS,[12] SEWS,[13] GMEWS,[4] Worthing,[14] ViEWS[15] and NEWS[16]), with the RRT single parameter calling criteria currently in use in our institution to assess the feasibility of using them in a real-time automated alert setting.

2. Methods

We used the following definitions:

Trigger: instances in which a patient met the conditions of a specific rule (score/vital sign exceeded the published/defined threshold).
Event: incidence of one of the following in a general care setting: unplanned transfer to the ICU, Resuscitation call, or RRT activation. Any movement directly from a general care bed to an intensive care unit was considered unplanned. Transfers to the ICU directly from the Emergency Department or Operating Room were not considered as an unplanned transfer.
Resuscitation call: call for a cardiopulmonary resuscitation when a patient has a cardiorespiratory arrest. Other institutions use the term “Code blue” or “Code45”.
Current clinical practice: in our institution, RRT activations require health provider action. Our RRT calling criteria are a single variable trigger. Staff should call an RRT assessment any time a patient has an acute and persistent change in any one or more of the following: oxygen saturations <90%, heart rate <40 or >130, systolic blood pressure <90, respiratory rate <10 or >28. This is in contrast to the comparison scores which are aggregate composites of multiple parameters.
Coverage time: the period after a rule triggers that is observed for events. For example, if a coverage time of 3 h is considered, a trigger is counted as a true positive if an event occurred during the following 3 h.
Episode: continuous time on the general care floor within a hospitalization, excluding times when a patient was in the operating room or intensive care unit (ICU). For example, if a patient was admitted to a general bed on a surgery floor, subsequently went to the operating room, and then returned to the surgery floor, two separate episodes were considered.

We selected a retrospective cohort comprised of all adult acute care in-patients discharged from two academic hospitals in 2011. The two hospitals had a total of 1300 eligible general care beds approximately, with approximately 250 ICU beds. We excluded Psychiatric or Rehabilitation inpatients, and patients admitted for research purposes.

We developed a longitudinal database that included patients’ data (vital signs, frailty measures, laboratory test results, demographics, urinary output, administrative data, comorbidities, code status and hospital floor) at the minute level throughout each patient’s hospital stay. Vital signs are manually collected and entered into the electronic medical record by a nurse. Physiologically impossible values (e.g. Blood Pressure of 1200 mmHg) were considered entered by mistake and eliminated from the database. Time spent in OR or ICU, or time patients were in “comfort care” status was excluded as RRT activation would not be applied.

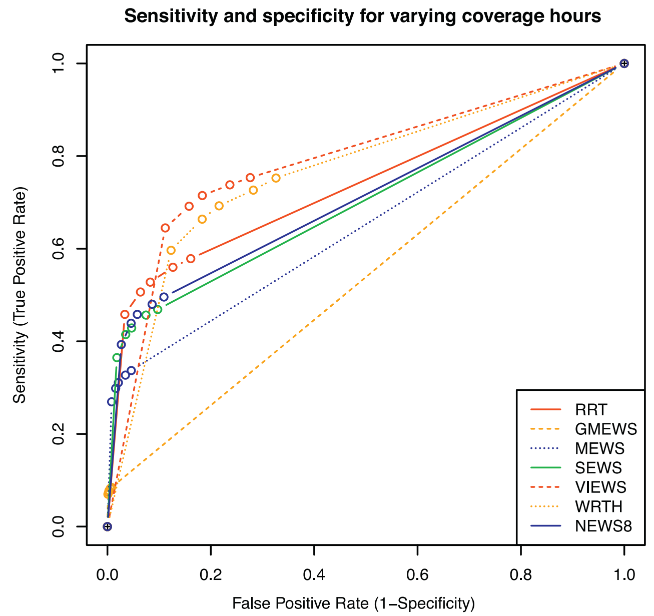

Fig. 1. RRT0: study institution’s Rapid Response Team calling criteria. WRTH, Worthing.

For our analysis, we determined that “events” (see above) were our outcome variables, and that triggers were our independent variables.

We calculated a variety of triggers using the MEWS, ViEWS, SEWS, GMEWS, NEWS and Worthing scores and our own RRT criteria in a rolling fashion through episodes of care. The score was updated every time a new parameter was entered into the analytical electronic medical record, and last values were carried forward to complete the rest of the required parameters to calculate the score.

We used the published thresholds to create rule triggers. For the RRT activation criteria in use in our institution, we used the criteria mentioned above. We simulated the alerts that would have been triggered using each of the different criteria, and compared it to the actual outcomes that occurred.

We defined different periods of coverage time (3, 8, 12, 24 and 36 h) after a rule trigger. If an event occurred within that time frame the trigger was considered a true positive. Analysis was performed for time periods after each rule triggered. Each one of those 5 analyses for each rule represents a point in Figs. 1 and 2.

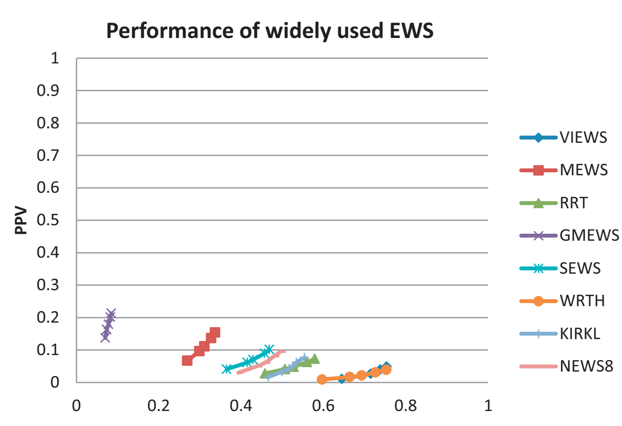

Fig. 2. RRT: study institution’s Rapid Response Team calling criteria. PPV, positive predictive value.

Table 1
Early warning systems.

Patients	
Total	34,898
Male (%)	17,001 (48.7%)
Female (%)	17,897 (51.3%)
Hospitalizations	
Total	46,366
Mean age (median)	
Male	60.9 (62.0%)
Female	56.9 (59.0%)
Mean LOS (quartiles)	
Male	5.2 (2; 3; 6)
Female	4.7 (2; 3; 5)
Episodes	
Total	75,240
Mean LOS (quartiles)	
Male	2.7 (0; 2; 4)
Female	2.5 (0; 2; 3)
Vitals: *N* (mean ± s.d.)	
Heart rate (beats/min)	5,794,425 (79.3 ± 16.2)
Respiration rate (breaths/min)	1,792,210 (17.9 ± 4.5)
Systolic BP (mmHg)	1,443,072 (123.4 ± 21.7)
Temperature (°C)	866,269 (36.8 ± 0.4)
SpO2 (%)	3,704,682 (95.2 ± 3.6)
Events	
Total	4747
RRS activations (%)	1888 (39.8%)
Resuscitation calls (%)	203 (4.3%)
Unsched Xfer to ICU (%)	2656 (55.9%)
Triggers by rule	
Current RRS criteria	256,762
GMEWS	5311
MEWS	50,446
SEWS	115,479
ViEWS	1,322,042
Worthing	1,129,067
NEWS	215,628

LOS, length of stay; BP, blood pressure; RRT, rapid response team. The following units were used: Age, years; LOS, days; Heart Rate, beats/minute; Respiratory Rate, breaths/minute; Temperature, degrees Celsius; SpO2, percentage.

3. Results

Table 1 includes the patients' demographics, number of hospitalizations and mean length of stay (LOS). During the year of study, there were 203 Resuscitation calls (0.128 cardiorespiratory arrests per hospital bed per year, or 4.37 per 1000 hospital admissions).

Fig. 1 shows the sensitivity and specificity curve for periods of time of 3, 8, 12, 24 and 36 h following each of the rules triggers. Fig. 2 shows a similar graph, but with positive predictive value (PPV) in the vertical axis and sensitivity in the horizontal axis. The latter demonstrates the tradeoffs an automated alarm system must weigh including the proportion of false alarms (1-PPV) generated. The number of false alarms increases significantly the further away (in hours) from the hypothetical trigger time for both the ViEWS and the Worthington early warning scores. The MEWS, SEWS and our institution's RRT criteria show less of a decay in specificity; however, do not reach the same degree of sensitivity through time.

4. Discussion

Most of the physiologically-based early warning scoring systems were dominated by the performance of a few approaches with VIEWS, MEWS and our own RRS activation criteria having the best performance among certain portions of the ROC curve. MEWS had the best specificity, but missed many events; VIEWS detected more events, but identified many false positive alerts.

GMEWS had the highest PPV, with 0.21 but a sensitivity of less than 0.08. That means that, if used in an automated alarm system, for every 10 alarms about 2 would predict a deterioration event, while about 92% of events would not be predicted by this index.

The number of arrests in our cohort is 0.128 per hospital bed per year, or 4.37 per 1000 hospital admissions, similar to other reported numbers in the literature.[17–19] Recent reports have shown rates of <1 cardiac arrest per 1000 admissions after the implementation of a Medical Emergency Team.[20,21] In the recent paper presenting the most recent of the evaluated scores, NEWS,[16] both vitals and outcomes were based on only the first 24 h of hospitalization, whereas our analysis was time-dependent, with the score updating every time a new set of vitals was recorded. The PPV for the proposed EWS was near 30%, but only 20% of events would be detected (80% false negative rate). The main weaknesses of this study include: (1) all study patients were hospitalized in a single tertiary care institution limiting the generalizability of the results; (2) analysis was performed on the aggregate hospital population without evaluation of patient subgroups, thereby making the assumption that all patients demonstrate similar physiologic changes prior to deterioration; (3) only the published early warning score thresholds were analyzed. It is possible that altering score thresholds could improve the PPVs, although it would be at the cost of lower sensitivity. Finally, in our analysis we defined an event by whether or not the decline in patient condition was ultimately recognized and thought serious enough to merit an escalation of care. Obviously such decisions depend not only on the physiological condition of the patient, but also on clinical judgments and comfort of the care team, anticipated needs for enhanced monitoring or nursing care, etc., so that the calculated positive predictive value does not only depend on the evaluated score, but on the recognition of the physiological deterioration. However, developing tools to help make such decisions earlier are the main purpose of EWSs, and so even though there are subjective elements that enter into the decision to request a higher level of acute care, we feel our definition of "event" is not unreasonable. Future, prospective studies could define an event in a way that is less affected by the clinicians' different judgments, and explore the possibility of calculating the score only in those instances when all variables were collected. The primary strength of this study is the complete enumeration of the periods of time patients were in the general care setting. This allowed for use of time-dependent statistical analyses to assess risk of an outcome event following the scores' triggers.

5. Conclusions

The most widely used weighted track-and-trigger scores do not offer good predictive capabilities for use as criteria for a time-sensitive automated alarm system. For the implementation of an automated alarm system, better criteria need to be developed and validated before implementation.

Conflict of interest statement

No conflicts of interest to declare.

Acknowledgements

This study was funded with Mayo Clinic internal funding. The authors graciously thank Megan Reinalda for her assistance with accessing the data sources and creation of the analytic data set necessary for these analyses; and the Morbidity and Mortality Council for their critical review of analytic approach and results.

References

1. Buist MD, et al. Recognising clinical instability in hospital patients before cardiac arrest or unplanned admission to intensive care. A pilot study in a tertiary-care hospital. Med J Aust 1999;171:22–5.
2. Schein RM, et al. Clinical antecedents to in-hospital cardiopulmonary arrest. Chest 1990;98:1388–92.
3. Hillman KM, et al. Antecedents to hospital deaths. Intern Med J 2001;31:343–8.
4. Harrison GA, et al. Combinations of early signs of critical illness predict in-hospital death-the SOCCER study (signs of critical conditions and emergency responses). Resuscitation 2006;71:327–34.
5. Hillman K, et al. Redefining in-hospital resuscitation: the concept of the medical emergency team. Resuscitation 2001;48:105–10.
6. Devita MA, et al. Findings of the first consensus conference on medical emergency teams. Crit Care Med 2006;34:2463–78.
7. Smith GB, et al. A review, and performance evaluation, of single-parameter track and trigger systems. Resuscitation 2008;79:11–21.
8. Smith GB, et al. Review and performance evaluation of aggregate weighted 'track and trigger' systems. Resuscitation 2008;77:170–9.
9. DeVita MA, et al. Identifying the hospitalised patient in crisis – a consensus conference on the afferent limb of rapid response systems. Resuscitation 2010;81:375–82.
10. Bellomo R, et al. Prospective controlled trial of effect of medical emergency team on postoperative morbidity and mortality rates. Crit Care Med 2004;32:916–21.
11. Griffiths JR, Kidney EM. Current use of early warning scores in UK emergency departments. Emerg Med J 2012;29:65–6.
12. Subbe CP, et al. Validation of a modified Early Warning Score in medical admissions. QJM 2001;94:521–6.
13. Paterson R, et al. Prediction of in-hospital mortality and length of stay using an early warning scoring system: clinical audit. Clin Med 2006;6: 281–4.
14. Duckitt RW, 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;98:769–74.
15. Prytherch DR, et al. ViEWS – towards a national early warning score for detecting adult inpatient deterioration. Resuscitation 2010;81:932–7.
16. Smith GB, et al. 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;84:465–70.
17. DeVita MA, et al. Use of medical emergency team responses to reduce hospital cardiopulmonary arrests. Qual Saf Health Care 2004;13:251–4.
18. Goncales PD, et al. Reduced frequency of cardiopulmonary arrests by rapid response teams. Einstein 2012;10:442–8.
19. Peberdy MA, et al. Cardiopulmonary resuscitation of adults in the hospital: a report of 14720 cardiac arrests from the National Registry of Cardiopulmonary Resuscitation. Resuscitation 2003;58:297–308.
20. Beckett DJ, et al. Reducing cardiac arrests in the acute admissions unit: a quality improvement journey. BMJ Qual Saf 2013;22:1025–31.
21. Konrad D, et al. Reducing in-hospital cardiac arrests and hospital mortality by introducing a medical emergency team. Intensive Care Med 2010;36: 100–6.