Utility of commonly captured data from an EHR to identify hospitalized patients at risk for clinical deterioration.

Abel Kho, MD MS[1,2,3], David Rotz RN[2], Kinan Alrahi[2], Wendy Cárdenas[2]
Kristin Ramsey RN[2], David Liebovitz MD[1,2], Gary Noskin MD[1,2], Chuck Watts MD[1,2]
[1]Northwestern University, Chicago, IL.
[2]Northwestern Memorial Hospital, Chicago, IL.
[3]Regenstrief Institute, Inc, Indianapolis, IN

Abstract

Rapid Response Teams (RRTs) respond to critically ill patients in the hospital. Activation of RRTs is highly subjective and misses a proportion of at-risk patients. We created an automated scoring system for non-ICU inpatients based on readily available electronic vital signs data, age, and body mass index. Over two weeks, we recorded scores on 1,878 patient with a range of scores from 0 to 10. Fifty patients reached the primary outcome of code call, cardiopulmonary arrest, or transfer to an ICU. Using a cutoff score of 4 or greater would result in identification of an additional 20 patients over the 7 patients identified by the current method of RRT activation. The area under the Receiver Operating Curve for the prediction model was 0.72 which compared favorably to other scoring systems. An electronic scoring system using readily captured EMR data may improve identification of patients at risk for clinical deterioration.

Background

Failure to quickly identify and respond to deteriorating patients can lead to increased morbidity and mortality in hospitalized patients. In response to this need, hospitals have created rapid response teams (RRT) for prompt assessment and intervention on this at-risk population.[1,2,3] Despite initial enthusiasm, controversy surrounds the evidence supporting implementation of RRTs.[4]

RRT activation is highly subjective. The Modified Early Warning Score has been proposed as a simple bedside scoring system to identify patients at risk for subsequent deterioration.[5] The MEWS takes into account five physiological parameters: systolic blood pressure, pulse rate, respiratory rate, temperature, and mental status. At our own institution, Northwestern Memorial Hospital (NMH), we piloted bedside nurse paper collection of a MEWS over two weeks, but abandoned this approach due to excessive burden on nursing staff.

Automated monitoring of patient data may provide for earlier recognition of a patient's impending deterioration and minimize additional work on nursing.[6,7] We hypothesized that an automatically generated score based on readily available data from an electronic medical record can accurately detect patients at risk for cardiopulmonary collapse, death, or transfer to an intensive care unit.

Methods

Setting:
Northwestern Memorial Hospital is a 725 bed academic medical center in Chicago, Illinois. NMH uses Cerner's Powerchart electronic medical record, and currently all vital signs and virtually 100% of all orders are entered electronically.

The RRT at NMH consists of five nurses with a collective 79 years of ICU experience and can be activated by any patient care provider (nurses or physicians) concerned about the state of any patient admitted to NMH. In place since 2006, the RRT at NMH responds to an average of three to five calls per day. The RRT is a quality improvement initiative at NMH in line with the Institute for Healthcare Improvement's 100,000 Lives Campaign.[8]

Study Design:
We based our scoring system on the previously validated MEWS. Within our EMR, patient mental status is infrequently recorded, and we removed the AVPU ("**A**lert", "Reacting to **V**ocal Stimuli", "Reacting to **P**ain" or "**U**nconscious") component of the MEWS. We performed a retrospective analysis of prior RRT calls to determine the common data elements that triggered a call to the RRT in our population. Based on our observations, and a review of the literature, we added age, which in a number of previous prediction models, including the original MEWS added significantly to discriminatory power.[9,10] We calculated and included a body mass index (weight in kilograms divided by the square of height in meters).[11] Our resulting scoring system consisted of the MEWS, minus the AVPU score, plus age and BMI score (Table 1).

	3	2	1	0	1	2	3
Systolic blood pressure (mmHg)	< 70	71-80	81-100	101-199		≥ 200	
Heart rate (bpm)		< 40	41-50	51-100	101-110	110-129	≥ 130
Respiratory rate (bpm)		< 9		9-14	15-20	21-29	≥ 30
Temperature (°C)		< 35		35-38.4		≥ 38.5	
Age (y)					65-74	75-84	≥ 85
BMI (kg/m²)			< 18.5		25.1-34.9	> 35	

Table 1. Scoring system combining original MEWS, minus AVPU score, and addition of age score and body mass index score.

We defined our primary outcome as a code call for cardiopulmonary arrest, in-hospital death, or transfer to an intensive care unit (ICU). We recorded data on all RRT calls during the study period. We recorded all codes and cardiopulmonary arrests and identified all patients transferred from general care wards to ICUs by review of electronic admission records for all ICUs. The Institutional Review Board of Northwestern University approved this study.

Automated Reports:
We created a standardized report using Cerner's Command Language (CCL) tool within Cerner's PowerChart application. In instances where some vital signs data were missing, the query retrieved the most recent vital sign prior to the missing value for the same admission. Inadequate height or weight data resulted in a default BMI score of zero. At the start of each shift, the RRT nurse generated a report (twice a day). The RRT nurses were instructed to not act on the electronic score sheets during this preliminary study, but only to store them in a secure file folder on the RRT computer. Over two weeks, we collected the vital signs, age and BMI on all adult patients admitted to NMH. The automated CCL reports generated a summary score that added the MEWS score (minus AVPU score), age score, and BMI score. For analysis we used the maximum MEWS score recorded on the ward for each patient.

Statistical Analysis:

We used SAS software 9.1.3 for analysis (SAS Institute, Inc., Cary, NC). We determined the sensitivity and specificity of various score cutoffs to predict our primary outcome and compared this with the current practice of provider-initiated RRT activation. We created a Receiver Operating Curve as an estimate of the discriminatory power of our automated electronic score.

Results

During the study period, we collected data on 1,878 patients, with an average age of 50.3 years, and 63% female predominance. The 1,277 patients with both a measured height and weight, had an average BMI of 28.3. Fifty patients were either transferred to the ICU, or coded, and none died. During the same time period, the RRT responded to 32 patient calls initiated by concerned ward personnel. Six of these patients subsequently required transfer to the ICU and one patient coded.

The composite scores of MEWS minus AVPU score, plus age score, plus BMI score ranged from 0-10, with a mean of 1.7 (Figure 1).
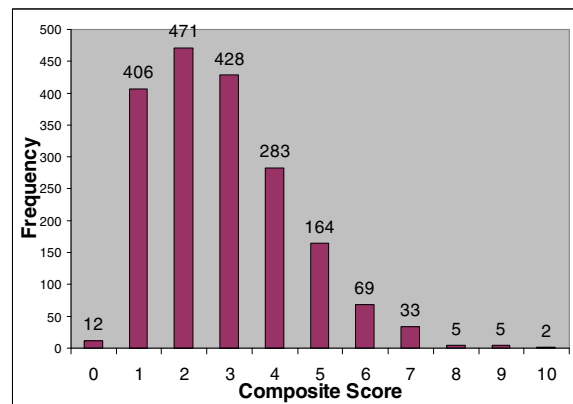

Figure 1. Distribution of composite scores from 0 to 10.

We calculated the sensitivity and specificity of the automated composite scores, and for the current system of staff activation of RRT to detect patients progressing to our primary outcome (Table 2).

Total Score	Sensitivity	Specificity
1	99	1
2	98	23
3	88	48
4	54	71
5	34	86
6	18	94
7	6	98
8	1	99
RRT Activated By Ward Personnel	22	98

Table 2. Sensitivity and specificity for the composite score and the current method of RRT activation to detect the primary study outcome.

Using a cutoff score of 4 or higher, would capture an additional 20 patients who would go on to code, or require elevation of care to an ICU, but result in an additional 229 RRT activations. Notably however, a cutoff score of 4 would have missed the one patient seen by the RRT who progressed to code (composite score of only 3).

In order to estimate the ability of our scoring system to discriminate between true and false positives for our primary outcome, we constructed a Receiver Operating Curve (Figure 2) and estimated the area under the curve (AUC). An AUC of 1 defines a perfect test with 100% sensitivity and 100% specificity whereas an AUC of 0.5 defines a test with no discriminatory power. Our AUC of 0.722 was almost identical to the AUC for the original MEWS (0.72), and suggests that despite significant modification, our scoring system discriminated between true and false primary outcomes as well as the original validated MEWS.
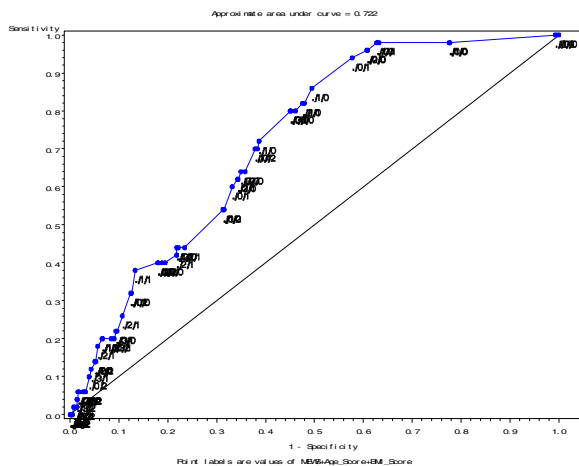

Figure 2. Receiver Operating Curve for MEWS Score + Age Score + BMI Score.

Discussion

Automated warnings based on continuous physiological monitoring may detect at-risk patients for earlier intervention.[12] Our results support such efforts, and suggest that simplified algorithms including readily available patient data may provide similar benefit.

However, health care, and patient deterioration cannot be predicted with 100% certainty. Experienced clinicians have excellent ability to identify a deteriorating patient based on subtle signs and symptoms not easily captured electronically.[13] Our scoring system could detect a greater number of at-risk patients (54% sensitivity compared with 22% for standard ward initiated RRT calls), at the tradeoff of numerous false positives. And one patient who progressed to cardiopulmonary arrest did so without preceding vital sign abnormalities. No detection system is perfect, although a combination of the two systems, automated surveillance, with human adjudication of suspected at-risk patients may ideally balance sensitivity and specificity better than either system alone.

Innovations in the airline industry have been held as a model for improving the safety of the healthcare industry.[14,15] Central to this model is the concept of multiple layers of checks and balances to prevent errors. We believe that automated central monitoring can add an additional layer of patient safety analogous to an "air traffic controller" for a hospital. This additional layer of data review may serve as a check against "group think" of a treating clinical team, who may not notice the incremental downward changes of a patient seen often throughout a day.

We hypothesize that electronic notification of RRT members to concerning clinical trends will both decrease the time to recognition of patients, and increase the accuracy of the RRT to recognize patients in danger of clinical deterioration. We believe that a hybridized model which combines the speed of electronic notification, with the accuracy of human review, will lead to a decrease in the overall rate of preventable code calls. Such an electronic dashboard can be utilized by a wide range of clinical services to prospectively monitor patient populations for concerning trends.

There were a number of limitations. Either height or weight were absent for approximately 1/3 of our sample. We conducted this interim analysis in preparation for a future trial, and additional data

collection may strengthen our model. We based our scoring on absolute data values although changes in values may add additional predictive power. Mis-identification of patients, may create additional burden on the RRT personnel and lead to decreased attention, analogous to "alert fatigue" with CPOE. Our scoring system utilized a minimum of data points and will almost certainly be improved by the addition of additional patient data, such as laboratory results.[16] In this initial study, we generated reports twice daily, which introduced a significant lag time for patients whose vital signs changed within the hours between reports. We are redesigning our reports to dynamically update reports as vital signs are entered into the EMR.

Conclusion

We believe that automated data delivery to a concerned expert may be a powerful tool to assist in patient care. Examples of this concept have proven value in remote critical care monitoring,[17] infection control,[18] and detection of adverse events.[19] In our example a simple electronic scoring system using readily captured EMR data identified additional patients at risk for clinical deterioration. Based on our preliminary results we are conducting a randomized trial comparing a proactive approach, based on automated alerting of our RRT to the current reactive approach to activation of the RRT.

References

[1] DeVita MA, Braithwaite RS, Mahidhara R, Stuart S, Foraida M, Simmons RL. Use of medical emergency team responses to reduce hospital cardiopulmonary arrests. Qual Saf Health Care 2004;13(4):251-4.

[2] Buist MD, Moore GE, Bernard SA, Waxman BP, Anderson JN, Nguyen TV. Effects of a medical emergency team on reduction of incidence of and mortality from unexpected cardiac arrests in hospital: preliminary study. BMJ 2002;324(7334):387-90.

[3] Braithwaite RS, DeVita MA, Mahidhara R, Simmons RL, Stuart S, Foraida M. Use of medical emergency team (MET) responses to detect medical errors. Qual Saf Health Care 2004;13(4):255-9.

[4] Winters BD, Pham J, Pronovost PJ. Rapid Response Teams – Walk, Don't Run. JAMA 2006 296(13):1645-47.

[5] Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning Score in medical admissions. QJM 2001;94(10):521-6.

[6] Tarassenko L, Hann A, Young D. Integrated monitoring and analysis for early warning of patient deterioration. Br J Anaesth 2006;97(1):64-8.

[7] Quarterman CPJ, Thomas AN, McKenna M, McNamee R. Use of a patient information system to audit the introduction of modified early warning scoring. Journal of Evaluation in Clinical Practice 2005;11(2):133-8.

[8] Institute for Healthcare Improvement website: http://www.ihi.org/ihi . Last accessed 3/15/2007.

[9] Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985;13:818-29.

[10] Kho AN, Hui S, Kesterson JG, McDonald CJ. Which observations from the complete blood count predict mortality for hospitalized patients? Journal of Hospital Medicine 2007;2:5-12.

[11] Flegal KM, Graubard BI, Williamson DF, Gail MH. Excess deaths associated with underweight, overweight, and obesity. JAMA 2005;293(15):1861-7.

[12] Tarassenko L, Hann A, Patterson A, et al. Biosign™: Multi-parameter monitoring for early warning of patient deterioration. 3rd IEE International Seminar on Medical Applications of Signal Processing. 2005:71-76.

[13] McClish DK, Powell SH. How well can physicians estimate mortality in a medical intensive care unit? Med Decision Making 1989;9(2):125-32.

[14] Helmreich RL. On error management: lessons from aviation. BMJ 2000;320(7237):781-5.

[15] Leape LL, Bates DW, Cullen DJ, et al. Systems analysis of adverse drug events. ADE Prevention Study Group. JAMA 1995;274(1):35-43.

[16] Fritsche L, Schlaefer A, Budde K, Schroeter K, Neumayer H-H. Recognition of Critical Situations from Time Series of Laboratory Results by Case-Based Reasoning. J Am Med Inform Assoc 2002;9(5):520-8.

[17] Breslow MJ, Rosenfeld BA, Doerfler M, et al. Effect of a multiple-site intensive care unit telemedicine program on clinical and economic outcomes: an alternative paradigm for intensivist staffing. Crit Care Med. 2004;32(1):31-8.

[18] Kho AN, Dexter PR, Warvel JS, Belsito AW, Commiskey M, Wilson SJ, Hui SL, McDonald CJ. An Effective Computerized Reminder for Contact Isolation of Patients Colonized or Infected with Resistant Organisms. International Journal of Medical Informatics (In press).

[19] Szekendi MK, Sullivan C, Bobb A, et al. Active surveillance using electronic triggers to detect adverse events in hospitalized patients. Qual Saf Health Care 2006;15(3):184-90.