



Published in final edited form as:

*Crit Care Med.* 2016 December ; 44(12): 2283–2285. doi:10.1097/CCM.0000000000002105.

## Moving Beyond Single Parameter Early Warning Scores For Rapid Response System Activation

Matthew M Churpek, MD, MPH, PhD<sup>1,\*</sup> and Dana P Edelson, MD, MS<sup>1</sup>

<sup>1</sup>Department of Medicine, University of Chicago, Chicago, IL

### Keywords

Hospital Rapid Response Team; Monitoring; Physiologic; Quality improvement; Vital Signs; Medical Emergency Team

Early warning systems provide an assessment of a patient's likelihood of developing critical illness and thus requiring additional critical care resources. The groundwork for these systems was laid millennia ago, with the Hippocratic "Book of Prognostics." The statement attributed to Hippocrates that "it is bad if he has dyspnoea, and urine that is thin and acrid, and if sweats come out about the neck and head" includes clinical variables (respiratory rate and urine output) still used in early warning systems today (1). These systems now form the foundation for activating Rapid Response and Medical Emergency Teams.

Traditionally, early warning systems have come in two primary configurations: single parameter criteria and aggregated weighted scores (2). The former originated in Australia over two decades ago as a set of equally weighted abnormal physiologic thresholds (e.g., respiratory rate >36), the presence of any of which would trigger the system (3). In contrast, aggregated weighted scoring systems, such as the Modified Early Warning Score (MEWS), which arose in the United Kingdom around the same time, involve summing up points from multiple parameters based on the degree of derangement (e.g., 2 points for a respiratory rate of 21–29 and 3 points for >30) (4, 5).

In the current issue of the *Journal*, Professor Smith and colleagues provide important evidence regarding the comparative accuracy of the National Early Warning Score (NEWS), an aggregated weighted score similar to the MEWS, which was developed by the Royal College of Physicians as a uniform method of identifying clinical deterioration in patients across the National Health Service (NHS) in the United Kingdom (6). Using data from an NHS District General Hospital, the authors compared NEWS to 44 distinct single parameter tools and found it to be superior for predicting death, cardiac arrest, and/or unanticipated intensive care unit transfer. This study is limited by the fact that this is a single center study using a population arising from the same hospital in which the VitalPAC™ Early Warning Score (ViEWS), its immediate precursor, was originally derived (7). However, these

\*Corresponding author and requests for reprints (Matthew M Churpek), Mail: University of Chicago Medical Center, Section of Pulmonary and Critical Care Medicine, 5841 South Maryland Avenue, MC 6076, Chicago, IL 60637, matthew.churpek@uchospitals.edu; Office: (773) 702-1092; Fax: (773) 702-6500.

concerns are largely mitigated by the fact that there is no overlap between the ViEWS derivation cohort and the current study population, and that the findings are consistent with independent studies demonstrating the superiority of aggregated weighted scoring systems over single parameter criteria (2, 8).

From a statistical modeling perspective, the finding that an aggregate weighted scoring system is more accurate than single parameter criteria is not surprising. Single parameter tools are generally based on single cut-points of continuous variables, which result in the loss of valuable information. For example, respiratory rates of 18 and 30 count similarly if they are both below the activation threshold. Furthermore, these criteria will miss subtle abnormalities in multiple vital signs, which have been shown to be more important for predicting outcomes than more dramatic elevations in a single vital sign (9). Aggregate weighted scores, which include several gradations of derangement and allow high scores to occur from both individual and combinations of vital sign abnormalities, do not suffer from these limitations. The NEWS has the added benefit of being informed by the dataset used to derive the ViEWS, rather than having been developed solely on the basis of expert opinion, upon which the vast majority of single parameter and many commonly used aggregated weighted scores were, including the MEWS. This is evident in the heavier weighting of subtler respiratory rate derangements, for example, which has been shown to be the vital sign with the strongest correlation to clinical deterioration (10, 11). In fact, the use of patient data in its development is the likely rationale for the superiority of NEWS to MEWS in prior head-on comparisons.

However, the improvements in accuracy need not stop there. Additional variables like laboratory data can be added and the full range of values can be utilized with logistic regression models and other similar models (12, 13). The use of vital sign trends can also increase accuracy, although accounting for these is more complicated than initially thought (14). Furthermore, the advent of machine learning tools, such as random forests, enable even more accurate models for predicting clinical deterioration (11).

If one believes that accuracy matters, and any hospital that has ever struggled with false alarms or missed opportunities would be hard pressed to argue that it doesn't, each hospital system owes it to its providers and patients to implement the most accurate activation tool it can. For those hospitals still using paper charts, that should be one of the aggregated weighted scores, of which the NEWS appears to be one of the stronger contenders. However, for those hospitals that have transitioned to the computer age, it's time to start thinking beyond paper based screening tools and make our expensive computers and electronic health records (EHRs) do the work they were designed to do. Retrofitting them with less accurate paper-based tools makes little sense.

Although results like the paper by Smith and colleagues suggest that this could and should be the beginning of the end for single parameter tools, it is becoming clear that sometime in the future we will be saying the same thing about simple aggregated weighted scores, like the NEWS, at least in their current form. EHRs are already ubiquitous in the United States, and are becoming more common in Europe, Australia, and other parts of the world as well. The EHR can harness the promise of "big data," with countless variables and high power

computing to automatically calculate complex and accurate algorithms in real-time. The future will belong to comprehensive and complex scores that are more accurate than NEWS, examples of which are already up and running in several hospitals today (13, 15). For hospitals that have already fully transitioned to using EHRs, it's time to make this future a reality. At a minimum, it's time to retire the single-parameter activation criteria once and for all.

## Acknowledgments

Conflicts of Interest and Source of Funding: Drs. Churpek and Edelson have a patent pending (ARCD. P0535US.P2) for risk stratification algorithms for hospitalized patients. Dr. Churpek is supported by a career development award from the National Heart, Lung, and Blood Institute (K08 HL121080) and an ATS Foundation Recognition Award for Early Career Investigators. In addition, he has received honoraria from Chest for invited speaking engagements. Dr. Edelson has received research support from Philips Healthcare (Andover, MA) and Early Sense (Tel Aviv, Israel). She has ownership interest in Quant HC (Chicago, IL), which is developing products for risk stratification of hospitalized patients.

Copyright form disclosures:

Dr. Churpek received funding (honoraria from Chest for invited speaking engagements), disclosed other support (Drs. Churpek and Edelson have a patent pending [ARCD. P0535US.P2] for risk stratification algorithms for hospitalized patients), and received support for article research from the National Institutes of Health (NIH). Dr. Edelson disclosed other support (A patent pending [ARCD. P0535US.P2] for risk stratification algorithms for hospitalized patients) and received support from ownership interest in QuantHC. Her institution received funding from grant from Philips Healthcare and grant from EarlySense.

## References

1. Hippocrates. *On Regimen in Acute Diseases*. Kessinger Publishing Co; 2004. Printed
2. Churpek MM, Yuen TC, Edelson DP. Risk stratification of hospitalized patients on the wards. *Chest*. 2013; 143(6):1758–1765. [PubMed: 23732586]
3. Lee A, Bishop G, Hillman KM, et al. The Medical Emergency Team. *Anaesth Intensive Care*. 1995; 23(2):183–186. [PubMed: 7793590]
4. Morgan RM, Williams F, Wright MM. An early warning scoring system for detecting developing critical illness. *Clin Intensive Care*. 1997; 8:100.
5. Subbe CP, Kruger M, Rutherford P, et al. Validation of a modified Early Warning Score in medical admissions. *QJM : monthly journal of the Association of Physicians*. 2001; 94(10):521–526. [PubMed: 11588210]
6. Smith GB, Prytherch D, Jarvis S, et al. A comparison of the ability of the physiological components of Medical Emergency Team criteria and the UK National Early Warning Score (NEWS) to discriminate patients at risk of a range of adverse clinical outcomes. *Crit Care Med*. 2016 in press.
7. Prytherch DR, Smith GB, Schmidt PE, et al. ViEWS--Towards a national early warning score for detecting adult inpatient deterioration. *Resuscitation*. 2010; 81(8):932–937. [PubMed: 20637974]
8. Tirkkonen J, Olkkola KT, Huhtala H, et al. Medical emergency team activation: performance of conventional dichotomised criteria versus national early warning score. *Acta anaesthesiologica Scandinavica*. 2014; 58(4):411–419. [PubMed: 24571384]
9. Jarvis S, Kovacs C, Briggs J, et al. Aggregate National Early Warning Score (NEWS) values are more important than high scores for a single vital signs parameter for discriminating the risk of adverse outcomes. *Resuscitation*. 2015; 87:75–80. [PubMed: 25433295]
10. Cuthbertson BH, Boroujerdi M, McKie L, et al. Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? *Crit Care Med*. 2007; 35(2):402–409. [PubMed: 17205002]
11. Churpek MM, Yuen TC, Winslow C, et al. Multicenter Comparison of Machine Learning Methods and Conventional Regression for Predicting Clinical Deterioration on the Wards. *Crit Care Med*. 2016; 44(2):368–374.

12. Churpek MM, Yuen TC, Winslow C, et al. Multicenter development and validation of a risk stratification tool for ward patients. *Am J Respir Crit Care Med.* 2014; 190(6):649–655. [PubMed: 25089847]
13. Escobar GJ, LaGuardia JC, Turk BJ, et al. 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. [PubMed: 22447632]
14. Churpek MM, Adhikari R, Edelson DP. The value of vital sign trends for detecting clinical deterioration on the wards. *Resuscitation.* 2016; 102:1–5. [PubMed: 26898412]
15. Kang MA, Churpek MM, Zdravetz FJ, et al. Real-Time Risk Prediction on the Wards: A Feasibility Study. *Crit Care Med.* 2016