Published in final edited form as:

Crit Care Med. 2016 March ; 44(3): 639–640. doi:10.1097/CCM.0000000000001501.

Keeping score of severity scores: taking the next step

Vincent Liu, MD, MS[1]

[1]Kaiser Permanente Division of Research; Oakland, CA

Keywords

Severity of illness index; Sepsis; critical care; risk adjustment; informtics

When it comes to severity scoring, critical care is fortunate to have an abundance of excellent options. Thanks to four decades of investment by diverse groups, data drawn from >300,000 critically ill patients around the world have been used to develop intensive care unit (ICU) scoring systems.(1) The investments have produced well-validated and contemporary iterations of severity scores including the Acute Physiology and Chronic Health Evaluation (APACHE), the Simplified Acute Physiology Score (SAPS), and the Mortality Probability Model (MPM), among several others.(2–5) In addition, numerous follow-on studies have compared how these scores perform in new patient subgroups or care settings.(1,6) And, the verdict is in: critical care is fortunate to have an abundance of excellent options.

In light of what could be considered an embarrassment of riches, how can we thus maximize the benefits of new investments in scoring system development? A natural inclination would aim to further improve the precision of outcome model predictions, a goal that seems well aligned with leveraging individual variability to direct precision medicine.(7) However, the benefits of such an investment may be limited for several reasons. First, despite the improved performance of contemporary iterations of ICU severity scores, their older versions seem good enough—scores developed in the 1980's persist even in the highest-profile studies.(8) Second, outcome prediction models were designed to be applied to populations of patients and may have limited utility when applied to individual patients.(1) Thus, despite our lengthy history with ICU scoring systems, we still lack any examples where these severity scores are used to direct treatments. Finally, greatly improved precision —for example, the ability to reliably discriminate between a 42% and 48% risk of developing shock—may have limited incremental benefit when there are only a few treatment options.

The study by Williams and others in this issue of Critical Care Medicine offers one potential way forward: choose parsimony.(9) Clinical data was prospectively collected from 8,871 Australian patients treated in the emergency department with presumed infection. The corresponding infection severity was categorized as infection without SIRS, sepsis, severe

Author information: Vincent Liu, MD MS, Kaiser Permanente Division of Research, 2000 Broadway, Oakland, CA 94612, (408) 769-8497, Vincent.x.liu@kp.org.

sepsis, and septic shock. The discrimination of severity score models with respect to mortality were then compared, including the Mortality in Emergency Department Sepsis (MEDS), the Sequential Organ Failure Assessment (SOFA), the Severe Sepsis Score (SSS), the APACHE II, and the SAPS 2 scores. In the overall cohort, the MEDS, SAPS, and APACHE scores all displayed exquisite discrimination for predicting mortality (c-statistics 0.90 to 0.92 with overlapping confidence intervals), likely owing to the relative health of the sample—53% of patients did not meet SIRS criteria. In all severity comparisons, the MEDS score displayed the highest discrimination. And, as expected, calibration was poor for the APACHE and SAPS—scores initially developed for critically ill patients in the 1980s.

While the study firmly supports the use of simpler scoring systems, it highlights relevant challenges as well. To its credit, the MEDS reduces the number of variables needed to produce a score from 16 (APACHE II) or 17 (SAPS2) to 9. However, many of the left-out variables in other scoring systems are routinely available for most patients (e.g., heart rate, sodium). Thus, if used at the point of care, it is arguable whether the MEDS would actually translate into a decrease in a clinician's workload. And, for the purposes of population comparisons for research or risk adjustment, the MEDS introduces new variables with heterogeneous availability and reliability (e.g., rapidly terminal comorbid illness, nursing home resident). What the MEDS does offer is a greatly simplified point assignment system where only one value is assigned if an abnormality is present (e.g., 3 points for tachypnea or hypoxia).

In light of its equivalent or superior discrimination in sepsis patients, choosing parsimony with the MEDS seems a reasonable choice; however, parsimony alone is not itself a worthy goal. In an age in which microprocessors vastly outperform humans in computational speed, power, and reliability, we should harness complex informatics to simplify the prognostic and treatment decision-making of clinicians.(10,11) Novel approaches leveraging machine learning to identify highest-risk sepsis patients earlier in their clinical course show promising results.(12) Furthermore, many efforts are underway to employ computers to offload the work of iteratively computing risk scores or performing pattern recognition in thousands of hospitalized patients in real-time.(13–15) When deployed within this longitudinal framework, even relatively simplistic scores are likely to demonstrate considerably improved utility and more favorable signal-to-noise ratios.

The grand vision is a living and breathing healthcare information infrastructure which proactively identifies patients in whom targeted and timely interventions can substantially improve the quality and outcomes of care. However, this vision has yet to be realized. And, perhaps in an ironic twist, even the most cutting-edge approaches have so far produced prediction models that look nearly identical to those already standardized decades ago. Fortunately, while we wait for our turn in the data revolution, when it comes to severity scoring in critical care, we still have an abundance of excellent options.

Acknowledgments

Funding sources: Dr. Liu is supported by NIH K23GM112018 and The Permanente Medical Group.

Copyright form disclosures: Dr. Liu received support for article research from the National Institutes of Health (NIH). His institution received grant support from the NIH/NIGMS.

References

1. Vincent JL, Moreno R. Clinical review: scoring systems in the critically ill. Crit Care. 2010; 14:207. [PubMed: 20392287]
2. Zimmerman JE, Kramer AA, McNair DS, et al. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. Crit Care Med. 2006; 34:1297–1310. [PubMed: 16540951]
3. Moreno RP, Metnitz PG, Almeida E. SAPS 3. From evaluation of the patient to evaluation of the intensive care unit. Part 2: development of a prognostic model for hospital mortality at ICU admission. Intensive Care Med. 2005; 31:1345–1355. [PubMed: 16132892]
4. Higgins TL, Teres D, Copes WS, et al. Assessing contemporary intensive care unit outcome: an updated Mortality Probability Admission Model (MPM0-III). Crit Care Med. 2007; 35:827–835. [PubMed: 17255863]
5. Harrison DA, Parry GJ, Carpenter JR, et al. A new risk prediction model for critical care: the Intensive Care National Audit & Research Centre (ICNARC) model. Crit Care Med. 2007; 35:1091–1098. [PubMed: 17334248]
6. Kuzniewicz MW, Vasilevskis EE, Lane R, et al. Variation in ICU risk-adjusted mortality: impact of methods of assessment and potential confounders. Chest. 2008; 133:1319–27. [PubMed: 18403657]
7. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med. 2015; 372:793–5. [PubMed: 25635347]
8. Angus DC, Barnato AE, Bell D, et al. A systematic review and meta-analysis of early goal-directed therapy for septic shock: the ARISE, PROCESS, and ProMISE Investigators. Intensive Care Med. 2015; 41:1549–60. [PubMed: 25952825]
9. Williams JM, Greenslade JH, Chu K, et al. Severity scores in Emergency Department patients with presumed infection: A prospective validation study. Crit Care Medi. 2016 in press.
10. Bates DW, Saria S, Ohno-Machado L, et al. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. Health Aff. 2014; 33:1123–31.
11. Liu V, Turk BJ, Ragins AI, et al. An electronic Simplified Acute Physiology Score-based risk adjustment score for critical illness in an integrated healthcare system. Crit Care Med. 2013; 41:41–8. [PubMed: 23222263]
12. Henry KE, Hager DN, Pronovost PJ, et al. A targeted real-time early warning score (TREWScore) for septic shock. Sci Transl Med. 2015; 7:299.
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; 7:388–95. [PubMed: 22447632]
14. Kollef MH, Chen Y, Heard K, et al. A randomized trial of real-time automated clinical deterioration alerts sent to a rapid response team. J Hosp Med. 2014; 9:424–9. [PubMed: 24706596]
15. 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; 190:649–55. [PubMed: 25089847]