

UCSF

UC San Francisco Previously Published Works

Title

Clinical Criteria to Identify Patients With Sepsis

Permalink

<https://escholarship.org/uc/item/0c46g4gf>

Journal

JAMA, 316(4)

ISSN

0098-7484

Authors

Makam, Anil N
Nguyen, Oanh Kieu

Publication Date

2016-07-26

DOI

10.1001/jama.2016.6407

Peer reviewed

Letters

COMMENT & RESPONSE

Clinical Criteria to Identify Patients With Sepsis

To the Editor Dr Seymour and colleagues¹ assessed the predictive validity of various clinical criteria to identify patients with sepsis. However, their conclusion that the Sequential [Sepsis-related] Organ Failure Assessment (SOFA) and quick SOFA (qSOFA) scores are more clinically useful than the systemic inflammatory response syndrome (SIRS) criteria was primarily based on differences in the area under the receiver operating curve (AUROC), which has several limitations.²

First, differences in the AUROC may be of minimal clinical relevance. For example, although the American College of Cardiology/American Heart Association (ACC/AHA) and the Adult Treatment Panel III cardiovascular risk prediction models have similar AUROCs, the ACC/AHA model recommends statin therapy for more adults and better reclassifies patients according to their true risk (net reclassification index [NRI], 0.332).^{3,4} To better assess for clinically meaningful differences in the utility of different sepsis diagnostic criteria, the authors should report the overall categorical NRI (for ≥ 2 vs < 2 points), event and nonevent NRIs, and risk reclassification tables.

Second, Seymour and colleagues proposed implementing both the SOFA and qSOFA scores in practice by defining patients with scores of 2 points or greater as having “sepsis” and those with less than 2 points as not having sepsis. However, the reported AUROCs were calculated using continuous scores. Assessing the performance of each set of clinical criteria using the dichotomous categorization and reporting the positive and negative predictive values and corresponding likelihood ratios compared with the SIRS criteria would be more clinically relevant than reporting the differences in the AUROC or the relative difference in outcomes across deciles of baseline risk.

Most important, it is unclear that the lower AUROC reflects inferior predictive validity of the SIRS criteria. Rather, the lower AUROC (and relative fold difference in outcomes) may reflect the influence of incorporation bias. The current standard of care for clinicians during the period of this study was to initiate early goal-directed therapy for patients with suspected sepsis based on SIRS criteria. Thus, the lower performance of SIRS for predicting mortality may simply be an artifact—patients diagnosed as having sepsis by 2 or more SIRS criteria were more likely to have received early goal-directed therapy and less likely to die as a consequence.⁵

Before the current diagnostic criteria for sepsis are replaced, we believe that more robust analyses are needed, including a prospective study of the utility of these various clinical criteria that would eliminate the influence of incorporation bias.

Anil N. Makam, MD, MAS
Oanh Kieu Nguyen, MD, MAS

Author Affiliations: Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas.

Corresponding Author: Anil N. Makam, MD, MAS, Division of General Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390-9169 (anil.makam@utsouthwestern.edu).

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

1. Seymour CW, Liu VX, Iwashyna TJ, et al. Assessment of clinical criteria for sepsis. *JAMA*. 2016;315(8):762-774.

2. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007;115(7):928-935.

3. Kavousi M, Leening MJ, Nanchen D, et al. Comparison of application of the ACC/AHA guidelines, Adult Treatment Panel III guidelines, and European Society of Cardiology guidelines for cardiovascular disease prevention in a European cohort. *JAMA*. 2014;311(14):1416-1423.

4. Paixao AR, Ayers CR, Berry JD, de Lemos JA, Khera A. Atherosclerotic cardiovascular disease prevention. *Circ Cardiovasc Qual Outcomes*. 2014;7(5):778-779.

5. Mouncey PR, Osborn TM, Power GS, et al. Trial of early, goal-directed resuscitation for septic shock. *N Engl J Med*. 2015;372(14):1301-1311.

To the Editor In their study, Dr Seymour and colleagues developed a simple model for identifying ward patients at risk of sepsis.¹ Their qSOFA score was compared with previously published criteria. We agree with the authors that there is a need for early recognition of patients with sepsis and that such systems should be based on measures of organ dysfunction, such as easily recorded vital signs, and feasible for health staff to use. The authors stated that the greater “predictive validity for in-hospital mortality” of the qSOFA supports its use.

However, we are not sure that the qSOFA is ready to be used. First, the qSOFA was derived through a data-driven approach in a large sample from 1 country. Standard significance levels combined with many tests increase the risk of false-positive results, an issue neither accounted for nor discussed by the authors.² The data were largely from 1 country, and transferring vital sign-based models to other settings can adversely affect their accuracy.³ Second, although the authors claimed to have favored biological plausibility, we find it hard to believe that a respiratory rate cutoff of 22 is more plausible than, for example, the more standard cutoff of 30.

Third, the clinical usefulness of the qSOFA is compromised by its design for use only in patients with suspected infection; it is difficult to distinguish infected patients from other critically ill patients. Fourth, most systems using vital signs show an association with mortality. A number of predictive systems have been developed claiming good predictive value, some simpler⁴ and some more complex.⁵ Seymour and colleagues' study does not prove that the qSOFA is the best available system. Fifth, the call for use of the qSOFA may be premature because it has the weakness found in many proposed predictive clinical criteria: a lack of evidence that use leads to improved outcomes.

Prior to rolling out the qSOFA, we would like it to be subjected to 2 challenges: (1) a comparison of its discrimination, calibration, and clinical usefulness in various settings with other models derived using subject matter knowledge or based on single vital signs and (2) a prospective trial of the effect, including patient outcomes, time burden, and costs, of using the qSOFA in clinical practice. If the qSOFA overcomes these challenges, then we too will be as optimistic as Seymour and colleagues.

Martin Gerdin, MD, PhD
Tim Baker, MB ChB, PhD

Author Affiliations: Department of Public Health Sciences, Karolinska Institutet, Stockholm, Sweden.

Corresponding Author: Martin Gerdin, MD, PhD, Department of Public Health Sciences, Karolinska Institutet, Stockholm, 17177, Sweden (martin.gerdin@ki.se).

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

1. Seymour CW, Liu VX, Iwashyna TJ, et al. Assessment of clinical criteria for sepsis. *JAMA*. 2016;315(8):762-774.
2. Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124.
3. Gerdin M, Roy N, Felländer-Tsai L, et al. Traumatic transfers. *Clin Epidemiol*. 2016;74:177-186.
4. Baker T, Blixt J, Lugazia E, et al. Single deranged physiologic parameters are associated with mortality in a low-income country. *Crit Care Med*. 2015;43(10):2171-2179.
5. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. 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(4):465-470.

In Reply Both letters suggest misunderstandings about the scientific goal of the task force, which was to explore the predictive validity of diagnostic criteria for sepsis. There is no gold standard for sepsis, which precludes simple measures of validity based on the presence of true positives (cases with sepsis) and true negatives (controls without sepsis). Predictive validity permits assessment of the extent to which potential criteria, applied in a population at risk of the unmeasurable condition (sepsis), predict outcomes more common in the condition.

Outcomes such as hospital mortality were chosen because sepsis is life threatening, implying that death is more common in infected patients who have sepsis. However, mortality is not necessarily caused by sepsis. Therefore, not only was the AUROC used but also the fold change within each decile of baseline risk of death. This approach explores the consistency with which death occurs more frequently than expected among patients with potential criteria for sepsis. We agree with Drs Makam and Nguyen regarding the advantages of the NRI in the outcome prediction examples they cite, but the situation is not analogous.

The task force did not propose the qSOFA as a stand-alone criterion for sepsis but rather as a prompt among clinicians of patients with infection to identify those who might fare badly. This decision reflects consideration of more validity domains than just predictive validity. Makam and Nguyen's comment about SIRS and early goal-directed

therapy is also not relevant because more than half of the patients were outside the emergency department at the onset of infection, and no early goal-directed therapy protocols were uniformly adopted across all hospitals. Also, the proportion of patients with 2 or more SIRS criteria and signs of hypoperfusion was low (<5%).¹

Counter to the claims of Drs Gerdin and Baker, the task force had no a priori hypotheses regarding which criteria would have the greatest predictive validity. We do contend, however, that altered mentation, hypotension, and tachypnea are biologically and clinically plausible as criteria associated with increased odds of poor outcome. The threshold for respiratory rate was simply the cut point associated with the greatest explanatory power in the model.

The article encouraged prospective validation in other data sets, ideally in broader settings. Gerdin and Baker suggest that the reason for a need for validation is because the predictive score may not calibrate well. However, calibration is not a priority for this exercise: the reasons to test externally are to understand if the fundamental relationship endures, regardless of calibration, and then if prospective deployment can be integrated to improve care and outcomes.

A separate question is how to help clinicians manage patients in whom infection is not suspected. However, this blends 2 tasks: diagnosis of infection (beyond the remit of the task force) and a severity of illness assessment, regardless of cause. There are countless severity instruments, and it was not the goal to compare the qSOFA with all possible scores. Nonetheless, it is reassuring that the elements in the qSOFA were similar to those of other scores such as the CURB-65,² yet dissimilar from SIRS.

Christopher W. Seymour, MD, MSc
Derek C. Angus, MD, MPH

Author Affiliations: Department of Critical Care Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania.

Corresponding Author: Christopher W. Seymour, MD, MSc, Departments of Critical Care Medicine and Emergency Medicine, University of Pittsburgh School of Medicine, Clinical Research, Investigation, and Systems Modeling of Acute Illness (CRISMA) Center, 3550 Terrace St, Scaife Hall, Ste 639, Pittsburgh, PA 15261 (seymourcw@upmc.edu).

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Seymour reports receiving grants from the National Institutes of Health and personal fees from Beckman Coulter, Cytovale, and Edwards. Dr Angus is Associate Editor, *JAMA*. No other disclosures were reported.

1. Seymour CW, Coopersmith CM, Deutschman CS, et al. Application of a framework to assess the usefulness of alternative sepsis criteria. *Crit Care Med*. 2016;44(3):e122-e130.
2. Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital. *Thorax*. 2003;58(5):377-382.

Defining Septic Shock

To the Editor The proposed new definition of septic shock, part of the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3), requires the simultaneous presence of hypotension and hyperlactatemia for making the diagnosis, instead of hypotension or hyperlactatemia.¹ In our opinion, this is a step backward compared with previous definitions.

First, including both hypotension and hyperlactatemia conflicts with the pathophysiology of shock. Shock is a