EDITORIAL

MSOFA: An Important Step Forward, but Are We Spending Too Much Time on the SOFA?

Lewis Rubinson, MD, PhD, FCCP; Ann Knebel, RN, DNSc, FAAN; John L. Hick, MD

In this issue of *Disaster Medicine and Public Health Preparedness*, Grissom and colleagues[1] present a comparison of the Sequential Organ Failure Assessment (SOFA)[2] physiologic score vs a newly derived score, based on SOFA but less dependent on laboratory measurements, which they term the Modified Sequential Organ Failure Assessment (MSOFA). The impetus for the study was that SOFA, which has been used frequently to compare patients in critical care research studies, has been recommended as a means to assist in the objective prioritization of patients during mass critical care events,[3-6] but it requires multiple laboratory measurements that may not be available during catastrophic events. Grissom et al empirically developed MSOFA and piloted the use of the score to predict mortality from a retrospective cohort of intensive care unit (ICU) patients at a tertiary care academic medical center. They then prospectively compared SOFA to MSOFA for their ability to predict the need for mechanical ventilation as well as mortality; this prospective comparison is reported in this issue.

THE ORIGINAL CASE FOR SOFA

Several years ago, the Working Group on Emergency Mass Critical Care considered the use of physiologic scoring systems for prioritizing patients to guide the allocation of scarce clinical resources during a mass critical care event.[7] Working Group members were drawn to the potential "objective" assignment of priority. However, precision and logistical concerns were raised, including the concern that many scores required multiple laboratory diagnostic results, so the group did not achieve consensus regarding use of scoring systems for triage and allocation during severe epidemics. Later, Hick and O'Laughlin described the initial attempt of the state of Minnesota to define a process and criteria for definitive triage during a serious epidemic.[8] Their work showed that objective criteria applied within a standardized decision-making process across hospitals was conceptually palatable to response agencies and incited a number of communities to address allocation of clinical resources during mass critical care. A group from Ontario made the next major contribution, which included use of SOFA,[1] and their model was incorporated with minor modifications into subsequent efforts by the New York State Task Force on Life and the Law and the Task Force for Mass Critical Care.[5,6]

SOFA has appeal because it can predict outcomes early in critical illness, has relatively few variables, includes variables that are frequently collected as part of routine clinical care, offers predictive value for both medical and trauma patients, and has been validated for serial measurements. Its shortfalls include lack of validation in children, uncertainty that it will predict outcomes for novel diseases because it was derived and validated on a mix of critically ill patients, and some of the variables may not accurately reflect the level of organ dysfunction for all populations (eg, creatinine in elderly adults). Finally, 4 of the 6 SOFA variables require laboratory measurements. Grissom et al point out that such data may not be available during a medical catastrophe to optimize the allocation of clinical resources.

IS MSOFA AN ALTERNATIVE TO SOFA?

The proposal to use SOFA was an improvement over haphazard resource allocation without a standard method for triage; however, it is far from perfect. Grissom and colleagues recognized the shortcomings of the laboratory-dependent variables of SOFA and consequently derived the MSOFA. The fact that only 1 laboratory measurement is necessary to calculate the MSOFA and that the score performs as well as SOFA is highly appealing. We believe the authors have performed a laudable service by establishing a credible alternative to SOFA.

Grissom and colleagues' study addressed whether 2 laboratory-derived variables could be exchanged for clinical or noninvasive measures of dysfunction of the same organ system—and do without a measure of hematologic dysfunction—while still having the same or a better ability to predict mortality. The substitution of SpO_2/FIO_2 for PaO_2/FIO_2 seems reasonable because it had been investigated for patients with acute lung injury,[9] and the information lost by not including thrombocytopenia appears to be limited for general ICU conditions.

We are concerned, however, that the nurses likely had knowledge of the patients' bilirubin results when recording their clinical examinations because 86% of patients received bilirubin measurements by ICU day 1. Although the MSOFA score was calculated from the clinical determination of liver function, the examination results may have been influenced by laboratory measurements, and thus MSOFA may not perform as well in settings in which bilirubin is truly not available. The published literature suggests that the reliability of bedside examination may be insufficient to identify scleral icterus or jaundice when clinicians are truly blinded to the bilirubin measurement[10,11]; this must be investigated further to confirm the utility of MSOFA.

©2010 American Medical Association. All rights reserved.
https://doi.org/10.1001/dmp.2010.41 Published online by Cambridge University Press

MSOFA AND SOFA HAVE SIMILAR LIMITATIONS

For the ICU population in Grissom and colleagues' study, MSOFA performs as well as SOFA and with fewer laboratory requirements. Still, it is important to remember that the study neither validates MSOFA for use in children nor determines that MSOFA variables are the most appropriate for predicting outcomes for unusual conditions such as critical illness associated with novel strains of influenza.

Physiologic scoring systems provide the appearance of objectivity, but their use as a stand-alone tool for guiding decisions of life and death, even during a catastrophe, is fraught with limitations. Grissom et al found that the fixed cutoff of a SOFA score >11 would have excluded only a small fraction of patients and would not have freed up sufficient resources; at the same time, a number of individuals who would have been excluded were found ultimately to have significant rates of survival. The evaluation by Guest et al of SOFA performance during the 2009 H1N1 influenza pandemic offers similar conclusions.[12] These limitations are not specific to SOFA but are germane to the application of fixed cutoffs of existing scoring systems, and they highlight how the circumstances in which scoring systems are applied may influence the score's performance.

For the sake of demonstration, assume that the sensitivity of a SOFA score >11 is 90% (ie, 90% of patients who die have a SOFA score >11 at the appropriate time point) and the specificity is 90% (ie, 90% of those who survive have a SOFA score <11). These hypothetical levels of sensitivity and specificity are considered generally to be good. If SOFA were used for a cohort whose mortality is 50%, then half of the patients would be excluded as being "too ill to benefit" and the mortality of the excluded group would be 90%. SOFA would be a useful tool in this situation. However, if SOFA were used for a cohort whose mortality is 10%, such as in the Grissom et al cohort or for pediatric critical illness associated with 2009 H1N1, then only 18% of the cohort would have a SOFA score >11, and of these, half would survive if they received care in the ICU. The actual performance of the scoring system also would be influenced by other circumstances of score implementation, such as when along the disease's time course the score is measured and how organ dysfunction is linked to the mortality of a specific event.

Thus, although MSOFA may be an improvement over SOFA, it still does not address some of the key limitations inherent in any scoring system, particularly when using a system designed to predict cohort outcome to prospectively predict individual outcomes. We are not advocating against the use of scoring systems as an element of a comprehensive triage process, because it is useful to have a common, objective system agreed upon before a mass critical care event that may be used for comparison among patients. We are, however, advocating that triage systems allow flexibility for disease-specific factors and ideally develop the capability for rapid assessment of score performance (to determine whether revised scores or modified cutoffs are necessary) during a response. Given the anticipated duress that medical staff will undergo during such events, data collection is rarely considered. However, in the case of the SOFA in the 2009 H1N1 influenza pandemic, this could have been catastrophic because many people with survivable conditions may have been triaged to end-of-life care had the pandemic been more severe. The evaluations of SOFA made by Grissom et al and Guest et al point out that even the best-intentioned triage protocols contain a profound requirement for early data collection and rapid analyses to ensure that optimal criteria are being applied to allocation decisions that are likely to have profound long-term effects on providers and communities.

Development and evaluation of improved scoring systems are important activities. Consideration also should be given to how to more nimbly and intelligently use these scores within a rational decision-making framework. Scoring systems are unlikely to perform well enough to be used in isolation. Instead, triage personnel can use sequential screening criteria to first identify individuals at high risk for mortality or poor outcomes and then use a physiologic score to better determine who in this high-mortality group is likely to be too ill to benefit from life-sustaining interventions. These sequential processes should be designed to minimize the time to complete the triage process. Still, if a majority of people need resources but they are not screened out because of intermediate likelihood of poor outcomes, then criteria other than physiologic scores may be necessary to guide fair and just allocation of scarce resources (eg, consideration of duration of use, duration of benefit, lottery, first-come, first-served prioritization categories determined by community).

Although important as objective, predictive comparators of outcome, MSOFA, SOFA, and similar scoring systems are only a component of the decision-making tools used when making resource-allocation decisions in a crisis setting. These decision-making tools themselves require adaptation to the specifics of the event; although it is important to examine and refine them, they are a small part of wider preparedness for making such allocation decisions. As described in a recent Institute of Medicine report,[13] ensuring institutional preparedness for use of such tools is the critical factor and includes components such as the following:

- Incident management and authorities within the institution
- Assessment and prioritization of clinical resources
- Use of triage teams, their function, and oversight
- Regional health care resource coordination.

Mass critical care events are likely to be dynamic. Scoring systems can assist in patient prioritization, but the decision about which resources to allocate to patients requires up-to-date knowledge of the available resources. Many hospitals and health care regions do not have in place systems to best match available and anticipated resources with patient needs. This must become a priority for their preparedness capabilities.

©2010 American Medical Association. All rights reserved.

https://doi.org/10.1001/dmp.2010.41 Published online by Cambridge University Press

Most institutions have no written crisis standards of care plan and even fewer have exercised these processes. With so much work to be done to establish operational plans for crisis conditions, we offer a proviso against "spending too much time on the SOFA" while recognizing the efforts of Grissom et al and others to advance our understanding of the predictive abilities of these systems.

Author Affiliations: Dr Rubinson is Deputy Chief Medical Officer, National Disaster Medical System, and Dr Knebel is Deputy Director, Preparedness Planning, Office of the Assistant Secretary for Preparedness and Response, Department of Health and Human Services; Dr Hick is Associate Professor of Emergency Medicine, University of Minnesota.

Address correspondence and reprint requests to Dr Lewis Rubinson, National Disaster Medical System, Department of Health and Human Services, 409 3rd St SW, Suite 330, Washington, DC 20024 (e-mail: lewis.rubinson@hhs.gov).

The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the Department of Health and Human Services or its components.

Received and accepted for publication October 6, 2010.

Author Disclosures: The authors report no conflicts of interest.

REFERENCES

1. Grissom CK, Brown SM, Kuttler KG, et al. A modified sequential organ failure assessment score for critical care triage. *Disaster Med Public Health Prep*. 2010;4(4):277-284.
2. Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med*. 1996;22(7):707-710.
3. Christian MD, Hawryluck L, Wax RS, et al. Development of a triage protocol for critical care during an influenza pandemic. *CMAJ*. 2006;175(11):1377-1381.
4. Devereaux AV, Dichter JR, Christian MD, et al; Task Force for Mass Critical Care. Definitive care for the critically ill during a disaster: a framework for allocation of scarce resources in mass critical care: from a Task Force for Mass Critical Care summit meeting, January 26-27, 2007, Chicago, IL. *Chest*. 2008;133(5)(Suppl):51S-66S.
5. Powell T, Christ KC, Birkhead GS. Allocation of ventilators in a public health disaster. *Disaster Med Public Health Prep*. 2008;2(1):20-26.
6. Christian MD, Joynt GM, Hick JL, Colvin J, Danis M, Sprung CL; European Society of Intensive Care Medicine's Task Force for intensive care unit triage during an influenza epidemic or mass disaster. Chapter 7. Critical care triage. Recommendations and standard operating procedures for intensive care unit and hospital preparations for an influenza epidemic or mass disaster. *Intensive Care Med*. 2010;36(Suppl 1):S55-S64.
7. Rubinson L, Nuzzo JB, Talmor DS, O'Toole T, Kramer BR, Inglesby TV. Augmentation of hospital critical care capacity after bioterrorist attacks or epidemics: recommendations of the Working Group on Emergency Mass Critical Care. *Crit Care Med*. 2005;33(10):2393-2403.
8. Hick JL, O'Laughlin DT. Concept of operations for triage of mechanical ventilation in an epidemic. *Acad Emerg Med*. 2006;13(2):223-229.
9. Rice TW, Wheeler AP, Bernard GR, Hayden DL, Schoenfeld DA, Ware LB; for the National Institutes of Health, National Heart, Lung, and Blood Institute ARDS Network. Comparison of the SpO2/FIO2 ratio and the PaO2/FIO2 ratio in patients with acute lung injury or ARDS. *Chest*. 2007;132(2):410-417.
10. Ruiz MA, Saab S, Rickman LS. The clinical detection of scleral icterus: observations of multiple examiners. *Mil Med*. 1997;162(8):560-563.
11. Hung OL, Kwon NS, Cole AE, et al. Evaluation of the physician's ability to recognize the presence or absence of anemia, fever, and jaundice. *Acad Emerg Med*. 2000;7(2):146-156.
12. Guest T, Tantam G, Donlin N, Tantam K, McMillan H, Tillyard A. An observational cohort study of triage for critical care provision during pandemic influenza: 'clipboard physicians' or 'evidenced based medicine'? *Anaesthesia*. 2009;64(11):1199-1206.
13. Altevogt BM, Stroud C, Hanson SL, Hanfling D, Gostin LO, eds. *Guidance for Establishing Crisis Standards of Care for Use in Disaster Situations: A Letter Report*. Washington, DC: The National Academies Press; 2009.

©2010 American Medical Association. All rights reserved.

https://doi.org/10.1001/dmp.2010.41 Published online by Cambridge University Press