Diagnosing sepsis: a step forward, and possibly a step back

Steven Q. Simpson

Division of Pulmonary and Critical Care Medicine, University of Kansas, Kansas City, KS, USA

Correspondence to: Steven Q. Simpson, MD, FCCP, FACP. Division of Pulmonary and Critical Care Medicine, University of Kansas, 3901 Rainbow Boulevard, Mail Stop #3007, Kansas City, KS 66160, USA. Email: ssimpson3@kumc.edu.

Sepsis/severe sepsis continues to be a worldwide public health problem, ending some lives prematurely, maiming others, and requiring inordinate amounts of healthcare expenditures (1-4). One of the keys to effective treatment of the condition is early recognition and initiation of therapy, and it has long been a goal of physicians in general, and intensivists specifically, to use diagnostic approaches that allow us to predict adverse outcomes, such as death, in patients who are sick but who are not so critically ill that our interventions lack effectiveness. The political writer Niccolo Machiavelli, writing in the 16th century, quoted the physicians of his time, "*As the physicians say of hectic fever, that in the beginning of the malady it is difficult to detect but easy to treat, but in the course of time, having been neither detected nor treated in the beginning, it becomes easy to detect but difficult to treat.*" (5). We, as physicians, have attempted to move forward the state of the art in sepsis diagnosis and treatment, but in many ways we are no more advanced in our approach than those physicians in the early 1500's.

In February of 2016 a group of physicians appointed by the European Society of Intensive Care Medicine and the Society of Critical Care Medicine published a revised definition and set of proposed diagnostic criteria for sepsis (6). The authors have dubbed their consensus conference and its output "Sepsis-3" and previous consensus conferences "Sepsis-1" and "Sepsis-2". The new definition of sepsis as life threatening organ dysfunction due to a dysregulated host response to infection seems reasonable, at least for a global description of the entity of sepsis, and was actually encapsulated in the pre-existing criteria for

diagnosis of severe sepsis and septic shock. It is worth noting, as did the "consensus" definition's authors, that there is a difference between a definition and diagnostic criteria. It may be fairly stated, in fact, that what were previously called definitions of sepsis, severe sepsis, and septic shock, were, in reality, diagnostic criteria for the conditions. It is obviously difficult to recognize when a host response has become "dysregulated", so diagnostic criteria for sepsis that are based on patient features that can be objectively recognized and recorded are necessary.

The Sepsis-3 authors have stated that their diagnostic criteria are the first to be based in clinical evidence or derived from clinical data, although it is more precise to say that they are the first to be formally derived from a specific clinical data set. It is quite clear that for millennia such features as fever, tachycardia, and tachypnea (commonly known as SIRS criteria) have been associated with serious infection, and that their presence in the setting of infection imputes an associated risk of mortality (7-9). Moreover, scores of papers have demonstrated that the associated mortality can be reduced by a variety of means, including early diagnosis with organized, standard treatment approaches (1,10).

The Sepsis-3 authors retrospectively evaluated a large database of patients from the 12 Hospital University of Pittsburgh Medical Center to propose alternatives to the current Sepsis-1 and Sepsis-2 diagnostic criteria (11). They compared sepsis according to the Sepsis-1 and Sepsis-2 criteria of known or suspected infection associated with ≥2 SIRS criteria with the newly minted Sepsis-3 criteria

of known or suspected infection with organ dysfunction, as defined by Sequential Organ Dysfunction Assessment score. Additionally, they compared with a screening test for non-ICU patients that they derived from their data set and called qSOFA. Finally, they compared with another organ dysfunction scoring system, the Logistic Organ Dysfunction Score (LODS) (12). The outcomes of interest for the Sepsis-3 authors were mortality or ICU stay ≥3 days. They compared the predictive ability of the diagnostic parameters for these outcomes by assessing the area under the receiver operator characteristic (ROC) curve for various values of the parameter. Interestingly, the LODS performed as well as or better than the ultimately chosen SOFA and qSOFA. Nevertheless, the Sepsis-3 authors proposed that increases of ≥2 SOFA points or ≥2 qSOFA features represent diagnostic criteria for ICU patients and non-ICU patients, respectively.

The Sepsis-3 authors called for additional research to inform their findings, and it has not taken long for a response. Although a series of editorials had called into question the usefulness and validity of the new diagnostic criteria, the first subsequent study of the proposed Sepsis-3 criteria was published online by Drs. Churpek *et al*. in September of 2016 (13-16). These authors studied patients with suspected infection at the University of Chicago healthcare system. They used techniques identical to those used by Seymour *et al*. in the investigations supporting the Sepsis-3 criteria, but limited themselves to patients diagnosed outside the ICU. They compared infection plus either SIRS or qSOFA, but added comparisons with the National Early Warning System (NEWS) and the Modified Early Warning System (MEWS) (17,18). They evaluated the same endpoints of mortality or ICU stay ≥3 days that the Sepsis-3 investigators used.

The University of Chicago (U of C) investigators found 30,677 evaluable patients in the period from November 2008 through January 2016. Of these patients, 60% met the clinical definition for inclusion in the emergency department, while 40% met the definition on the wards. Area under the ROC curve with mortality as the outcome was greatest for NEWS (0.77), followed by MEWS (0.73), qSOFA (0.69), and SIRS (0.65). Areas under the curve were qualitatively similar, i.e., showed the same order of discriminative ability, but were slightly smaller for the composite outcome of mortality or prolonged ICU stay. NEWS and MEWS were shown to be more efficient predictors of both mortality and of the composite outcome than SIRS or qSOFA, meaning that for any given level of

sensitivity, the proportion of patients with a score above the threshold value was lower. However, there was a rapid convergence of efficiency at sensitivities higher than 90%.

There are two key messages from the University of Chicago study. First, while the Sepsis-3 investigators did use their own data to determine the utility of SOFA and qSOFA and even to define qSOFA, other scoring systems can provide equal or better discriminatory ability for prediction of mortality or prolonged ICU stay in patients diagnosed outside the ICU. Second, NEWS and SIRS identify patients at risk for mortality or prolonged ICU stay substantially earlier than qSOFA in non-ICU patients. The first message is important mostly because many hospitals are currently equipped with EMRs that calculate the NEWS and/or MEWS scores in the background and display them to providers. The latter message is of key importance to patients and to clinicians, because adequate treatment of sepsis, severe sepsis, or septic shock necessitates early recognition and requires time for important additional testing, for ordering and administration of antimicrobials and fluids, and for ICU transfer of those patients who require it. The best outcomes are likely to be achieved by identifying patients with a time cushion that allows for the foibles and delays of human activity. In other words, all human activities, and especially those involved in the care of sick patients, take time and will be affected by weaknesses in our delivery systems. We should take advantage of early warning signs whenever we can.

One important weakness of both the U of C and the Sepsis-3 investigations is that they compare diagnostic criteria for sepsis that include organ dysfunctions (infection with qSOFA, NEWS, or MEWS) with diagnostic criteria that specifically exclude organ dysfunction (infection with SIRS). It should not be surprising that infection with organ dysfunction has a higher predictive ability for death or prolonged ICU stay than infection without organ dysfunction. There is a flavor of sleight of hand to assigning a new meaning to the word sepsis, then showing that the word now has different implications than it previously did. Both studies seem to forget that the diagnosis of principal concern in our current Sepsis-1, 2 diagnostic framework is severe sepsis, meaning infection with SIRS and organ dysfunction. Guidelines for the care of patients with serious infection are aimed at patients with severe sepsis or septic shock, as are nearly all major investigations of therapies for sepsis (19-23). Both studies would be more germane to clinical practice if they compared the predictive ability of the proposed new sepsis diagnostic criteria with the

Table 1 Sensitivity and specificity for mortality of diagnostic criteria for sepsis/severe sepsis in non-ICU settings

Criteria	Source	Sensitivity (%)	Specificity (%)
SIRS ≥2	Sepsis-3	64.0	65.0
	U of C	93.8	12.3
qSOFA ≥2	Sepsis-3	55.0	84.0
	U of C	68.7	63.5
NEWS ≥7	U of C	86.6	47.5
MEWS ≥5	U of C	71.4	65.0
Severe sepsis	Kaukonen *et al.*	92.0	84.0

predictive ability of the current severe sepsis criteria.

An additional important limitation of both studies, but especially of the Sepsis-3 study, is the switch from area under the ROC curve to the choice of specific diagnostic criteria for sepsis. The ROC curve plots the true positive rate, or sensitivity, as a function of the false positive rate, also expressed as 1—specificity, for an array of chosen diagnostic cutoffs of a test. In this case, the tests are SIRS, qSOFA, NEWS, and MEWS, with the diagnostic cutoffs being various levels of change in vital signs or organ dysfunction. The area under the ROC curve is of interest, because it demonstrates which test is, in a general sense, a more discriminative test. The U of C and Sepsis-3 studies clearly demonstrate that in the presence of suspected infection NEWS, MEWS, and qSOFA have better overall discriminative ability than SIRS for the outcomes of mortality and prolonged ICU stay.

One does not, however, use an overall test to diagnose patients, but a specific value of the test, e.g., SIRS ≥2 or qSOFA ≥2. Once we limit ourselves to a specific diagnostic cutoff value, then sensitivity and specificity of that test at that value become the comparison points of interest. *Table 1* shows a comparison of the sensitivity and specificity of SIRS, SIRS + organ dysfunction, qSOFA, and NEWS as predictors of mortality in patients with suspected infection, all displayed at their normally used cutoff values. The data for SIRS, qSOFA, and NEWS are taken from the U of C study, while the data for SIRS + organ dysfunction are taken from the study of Kaukonen *et al.*, with organ dysfunction as defined by the Sepsis-2 criteria (3). This table suggests that at currently used cutoffs severe sepsis is a far more accurate predictor of mortality than any of the other proposed diagnostic criteria. Unfortunately, no currently published study evaluates the utility of severe sepsis as a predictor of prolonged ICU stay. Nevertheless, it is readily seen that at the cutoff points for severe sepsis that have previously been studied, severe sepsis substantially outperforms any of the other tested diagnostic criteria for predicting mortality.

If one accepts that "life threatening organ dysfunction" is the hallmark feature of sepsis or severe sepsis, then mortality or prolonged ICU stay seem to be appropriate endpoints for study. However, if we view the proposition from the viewpoint of patients who are ill and needing intervention, then it might seem that those two outcomes are rather unpalatable, and we should ask if it would be better to identify sepsis early enough so that intervention could avert either of those endpoints. In other words, patients are not really interested in our prognostic ability to predict that they will suffer or die in spite of intervention; that is an ability that physicians find useful, not patients. What may be more useful is some indicator that in the presence of infection with signs of physiological stress or distress (SIRS), there is some indicator of "dysregulated" host response. Such findings could be immunological in nature, as proposed by Bermejo-Martin *et al.* or they could be more physiological biomarkers (24-26). However, even with a biomarker or biomarkers to indicate when the host response has become dysregulated, we will need to recognize which patients require testing, and that recognition will be based on some clinical syndrome. Efforts to define clinical criteria for the early diagnosis of sepsis, such as those by the U of C and Sepsis-3 investigators, are therefore very germane to clinical practice and to outcomes that matter to our patients.

In the absence of a clear diagnostic pathophysiology or a defining biomarker, the Sepsis-3 investigators attempted to use a large data set to define, for better or worse, a new syndrome more predictive of mortality or prolonged ICU stay. In reality, they have managed an attempt to move

from one less than perfect syndrome to another less than perfect syndrome. The U of C data suggest that, even if we accept the study endpoints of mortality and prolonged ICU stay, other syndromes could potentially be implemented to greater effect. In one way, the outcomes of these investigations reveal that we are the proverbial blind men attempting to describe an elephant; each of us perceives something different, based on our reasons for diagnosing the condition. Patients need for physicians to detect sepsis early and intervene when the simple treatments of IV fluids and antibiotics are sufficient to halt or markedly slow a downward spiral. Clinical investigators who desire to develop new therapeutic agents need a syndrome that is most closely associated with mortality, so that the impact of the agent can be readily demonstrated in clinical trials. Both of these endeavors are ultimately of benefit to patients. It seems that the medical community faces a choice. We can use separate diagnostic criteria for studying sepsis and treating sepsis. We can work together to develop criteria that are broad-based enough to recognize and accommodate the needs of both patients and researchers. Or, we can take another look at our longstanding criteria and recognize that they have actually encompassed the desired concepts all along.

Acknowledgements

Footnote

References

1. Rhodes A, Phillips G, Beale R, et al. The Surviving Sepsis Campaign bundles and outcome: results from the International Multicentre Prevalence Study on Sepsis (the IMPreSS study). Intensive Care Med 2015;41:1620-8.
2. Iwashyna TJ, Ely EW, Smith DM, et al. Long-term cognitive impairment and functional disability among survivors of severe sepsis. JAMA 2010;304:1787-94.
3. Kaukonen KM, Bailey M, Suzuki S, et al. Mortality related to severe sepsis and septic shock among critically ill patients in Australia and New Zealand, 2000-2012. JAMA 2014;311:1308-16.
4. Kumar G, Kumar N, Taneja A, et al. Nationwide trends of severe sepsis in the 21st century (2000-2007). Chest 2011;140:1223-31.
5. Machiavelli N, Skinner Q, Price R. Machiavelli: The Prince. Cambridge University Press; 1988.
6. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA 2016;315:801-10.
7. Funk DJ, Parrillo JE, Kumar A. Sepsis and septic shock: a history. Crit Care Clin 2009;25:83-101, viii.
8. Bone RC, Fisher CJ Jr, Clemmer TP, et al. Sepsis syndrome: a valid clinical entity. Methylprednisolone Severe Sepsis Study Group. Crit Care Med 1989;17:389-93.
9. Rangel-Frausto MS, Pittet D, Costigan M, et al. The natural history of the systemic inflammatory response syndrome (SIRS). A prospective study. JAMA 1995;273:117-23.
10. Rivers EP. Rebuttal From Dr Rivers. Chest 2010;138:483–4.
11. Seymour CW, Liu VX, Iwashyna TJ, et al. Assessment of Clinical Criteria for Sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA 2016;315:762-74.
12. Le Gall JR, Klar J, Lemeshow S, et al. The Logistic Organ Dysfunction system. A new way to assess organ dysfunction in the intensive care unit. ICU Scoring Group. JAMA 1996;276:802-10.
13. Simpson SQ. New Sepsis Criteria: A Change We Should Not Make. Chest 2016;149:1117-8.
14. Marshall JC. Sepsis-3: What is the Meaning of a Definition? Crit Care Med 2016;44:1459-60.
15. Churpek MM, Snyder A, Han X, et al. qSOFA, SIRS, and Early Warning Scores for Detecting Clinical Deterioration in Infected Patients Outside the ICU. Am J Respir Crit Care Med 2016. [Epub ahead of print].
16. Cortés-Puch I, Hartog CS. Opening the Debate on the New Sepsis Definition Change Is Not Necessarily Progress: Revision of the Sepsis Definition Should Be Based on New Scientific Insights. Am J Respir Crit Care Med 2016;194:16-8.
17. Smith GB, Prytherch DR, Meredith P, et al. 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;84:465-70.
18. Gardner-Thorpe J, Love N, Wrightson J, et al. The value of Modified Early Warning Score (MEWS) in surgical in-patients: a prospective observational study. Ann R Coll Surg Engl 2006;88:571-5.
19. Dellinger RP, Levy MM, Rhodes A, et al. Surviving sepsis

campaign: international guidelines for management of severe sepsis and septic shock: 2012. Crit Care Med 2013;41:580-637.

20. Levy MM, Rhodes A, Phillips GS, et al. Surviving Sepsis Campaign: association between performance metrics and outcomes in a 7.5-year study. Crit Care Med 2015;43:3-12.

21. ARISE Investigators; ANZICS Clinical Trials Group, Peake SL, et al. Goal-directed resuscitation for patients with early septic shock. N Engl J Med 2014;371:1496-506.

22. ProCESS Investigators, Yealy DM, Kellum JA, et al. A randomized trial of protocol-based care for early septic shock. N Engl J Med 2014;370:1683-93.

23. Mouncey PR, Osborn TM, Power GS, et al. Trial of early, goal-directed resuscitation for septic shock. N Engl J Med 2015;372:1301-11.

24. Bermejo-Martin JF, Andaluz-Ojeda D, Almansa R, et al. Defining immunological dysfunction in sepsis: A requisite tool for precision medicine. J Infect 2016;72:525-36.

25. Ventetuolo CE, Levy MM. Biomarkers: diagnosis and risk assessment in sepsis. Clin Chest Med 2008;29:591-603, vii.

26. Kibe S, Adams K, Barlow G. Diagnostic and prognostic biomarkers of sepsis in critical care. J Antimicrob Chemother 2011;66 Suppl 2:ii33-40.