The upcoming SOFA 2.0 score: a roadmap for future developments in critical care?

Bruno Adler Maccagnan Pinheiro Besen[1,2], Andre C Kalil[3], Elisa Estenssoro[4,5], Pedro Póvoa[6,7,8]

1 Instituto D'Or de Pesquisa e Ensino - São Paulo (SP), Brazil.
2 Postgraduate Program in Medical Sciences, Faculdade de Medicina, Universidade de São Paulo - São Paulo (SP), Brazil.
3 Division of Infectious Diseases, University of Nebraska Medical Center - Omaha, Nebraska, United States.
4 Servicio de Terapia Intensiva, Hospital Interzonal de Agudos San Martin de La Plata - Buenos Aires, Argentina.
5 Dirección de Investigación; Ministerio de Salud - Provincia de Buenos Aires, Argentina.
6 NOVA Medical School, NOVA University of Lisboa - Lisboa, Portugal.
7 Research Unit of Clinical Epidemiology, Department of Clinical Research, Odense University Hospital - Odense, Denmark.
8 Department of Intensive Care, Hospital São Francisco Xavier - Lisboa, Portugal.

WHY DO WE NEED ORGAN DYSFUNCTION SCORES?

Intensive care medicine is characterized by providing organ support to dysfunctional or failing organs during acute or acute-on-chronic illness. Organ dysfunction may present in isolation or combination and occur at different times during critical illness in each patient. Although intensivists are trained to recognize and respond promptly to organ dysfunction to avoid further deterioration and, when this occurs, to provide multiple organ support, the evaluation of organ dysfunctions might differ between intensivists. Therefore, organ dysfunction scores that standardize how organ dysfunction is measured are especially warranted. They may both enhance and standardize communication by intensivists in their day-to-day interactions and aid in research purposes.

In 1996, Vincent et al. introduced the Sequential Organ Failure Assessment (SOFA) score based on a consensus definition.[1] The SOFA score allowed the description of organ dysfunction dynamics in patients with multiple organ failure and its prognostic association with clinical outcomes. A few are worth remembering: (i) non-recovery within a few days of organ failure onset was associated with similar outcomes to organ failure worsening;[2] (ii) the maximum intensity of organ dysfunction in each domain during intensive care unit (ICU) stay (maximum total SOFA) was perhaps the strongest predictor of mortality;[3] (iii) the timing of organ dysfunction was variable during acute illness, with neurological dysfunction presenting earlier, while liver and coagulation abnormalities have lagged presentations;[3,4] and (iv) the prognostic role of non-neurological organ dysfunction in neurocritical care patients.[5]

Significantly, the SOFA score was not developed to be a prognostic score. Neither could it be, as we know that the clinical outcome of critically ill patients is only partially explained by the intensity and broadness of organ dysfunctions, with additional contribution from both patient antecedent characteristics and contextual variables reflecting ICU organizational characteristics.[6,7] The SOFA score describes organ dysfunction dynamics through time, domains, and intensity.

WHY SHOULD THE SOFA SCORE BE UPDATED?

The SOFA score, initially published in 1996,[1] has been successfully incorporated into critical care medicine research.[8,9] It allowed descriptive and prognostic research and has been used to determine surrogate outcomes in critical care trials.[10] It has also become part of the latest sepsis definitions.[8] Reasons for SOFA score success perhaps involve its underlying principles: (i) inclusion of low-cost, simple, routinely collected data aiming at a low number of variables; (ii) ability to be serially measured, as organ dysfunction is dynamic throughout critical illness; and (iii) stratification of organ dysfunction in clear-cut, objectively defined, discrete categories that reflect increased organ dysfunction burden and severity. Nevertheless, although its simplicity has certainly passed the test of time and facilitated its adoption even in low-resource settings, new additions to critical care daily practice have brought the need to update the score.[11-13]

A few of these advancements include: (1) new technologies, either more invasive – such as venous-venous (VV-) and veno-arterial (VA-) extracorporeal membrane oxygenation (ECMO) – or less invasive – as high flow nasal oxygen (HFNO) and noninvasive ventilation (NIV) –; (2) new drugs, such as vasopressin and angiotensin II; (3) the capacity to better assess neurological dysfunction with light sedation approaches; (4) less invasive approaches to the assessment of critically ill patients, with oxygen saturation/fraction of inspired oxygen (SpO_2/FiO_2) ratios instead of partial pressure of oxygen/fraction of inspired oxygen (PaO_2/FiO_2) ratios as an example;[14] (5) the (virtual) abandonment of some treatment strategies (such as dopamine); and (6) identification of other possible domains to be included in the score, such as immunological dysfunction.[15,16] Development and validation of an updated SOFA score is therefore the consequence of advances in critical care medicine over the last decades.

THE SOFA SCORE IS NOW 2.0: WHAT ARE ITS MERITS?

Figure 1 presents a few of the new development's strengths and limitations. SOFA 2.0 was developed following state-of-the-art methodology, which will allow internal validity and substantial generalizability. First, the authors designed a panel composition that is more representative of the global critical care community, multinational, with representatives from all continents and from different country-level incomes. The panel also balanced content and methodological expertise, a significant strength to avoid expert opinion-based round table discussions. Second, the authors did not abandon the underlying theoretical principles to update the score. Not only has the score been based on the content expertise, which has driven Delphi rounds, but the working groups have systematically reviewed published evidence to identify possible targets to characterize the severity of organ dysfunction across each domain. Third, after considering underlying theory and published evidence, the score used a modern data-driven validation that leverages currently available real-world ICU databases. Therefore, the association of organ dysfunction categories with outcomes is based not only on theory or consensus but also on predictive validity. Finally, SOFA 2.0 is likely to be generalizable to many critical care scenarios, given the diversity in the databases to validate the cut-off points of the score.
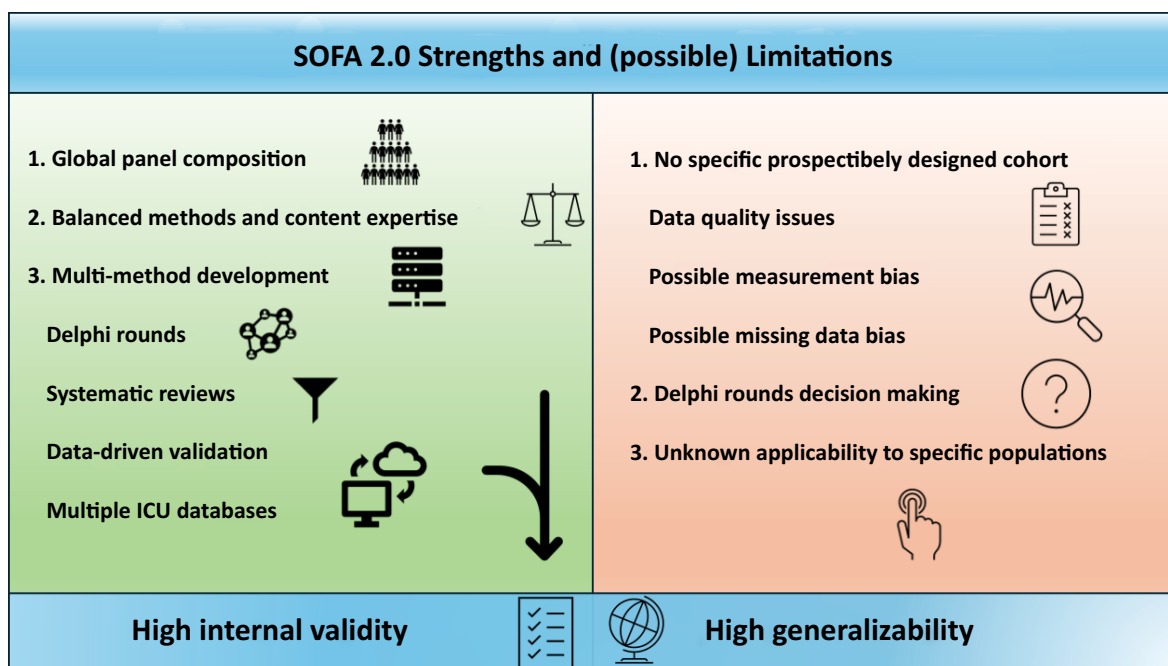

Figure 1 - SOFA 2.0 score strengths and (possible) limitations.*
ICU - intensive care unit.
* These strengths and limitations were considered based on past European Society of Intensive Care Medicine (ESICM) LIVES 2024 and *Congresso Brasileiro de Medicina Intensiva* (CBMI) 2024 presentations by the principal authors of the revised version.

ARE THERE ANY LIMITATIONS?

SOFA 2.0 will undoubtedly come with some limitations. First, it was not based on prospective cohort studies, which may reduce data quality and completeness.[17] Although large databases designed to capture clinically relevant data might partly overcome this limitation, large databases cannot necessarily prevent some sorts of bias, such as measurement bias.[13] Other limitations stem from decisions made during Delphi rounds – a reflection of the choice of panelists of the working group –,[18] that might be subject to bias and criticisms. Another limitation is that a single score derivation and validation may hamper generalizability to frequent and specific scenarios and subgroups, such as critical illness in pregnancy, or to underrepresented conditions from low-income countries, such as malaria and other high-burden infectious diseases of interest.

FURTHER RESEARCH AGENDA

The expected publication of SOFA 2.0 in 2025 will likely come with a substantial new research agenda. Should we start using SOFA 2.0 to define sepsis instead of SOFA? What will the predictive validity of SOFA 2.0 be in defining sepsis? Will the predictive validity of SOFA 2.0 to operationally define sepsis still be based on a two-point variation of baseline SOFA? Is SOFA 2.0 applicable to specific subgroups of patients, and how are subgroups different? If new domains were not included, was there any issue in the score development approach that may be prevented in the future?

CONCLUSIONS

The authors of SOFA 2.0 should be commended for their concerted efforts to update the SOFA score. It will hopefully be an example of a balanced panel of content experts and methodologists experienced in critical care research. Upon its publication, we expect further external prospective validation, which will be necessary to determine its accuracy and generalizability.

Publisher's note

Conflicts of interest: None.

Corresponding author:
Bruno Adler Maccagnan Pinheiro Besen
Faculdade de Medicina, Universidade de São Paulo
Av. Dr. Enéas Carvalho de Aguiar, 255, room 11.083
Zip code: 05403-000 - São Paulo (SP), Brazil
E-mail: brunobesen@yahoo.com.br

Responsible editor: Jorge Ibrain Figueira Salluh ⓘ

REFERENCES

1. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, 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):707-10.
2. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. Serial evaluation of the SOFA score to predict outcome in critically ill patients. JAMA. 2001;286(14):1754-8.
3. Moreno R, Vincent JL, Matos R, Mendonça A, Cantraine F, Thijs L, et al. The use of maximum SOFA score to quantify organ dysfunction/failure in intensive care. Results of a prospective, multicentre study. Working Group on Sepsis related Problems of the ESICM. Intensive Care Med. 1999;25(7):686-96.
4. Vincent JL, de Mendonça A, Cantraine F, Moreno R, Takala J, Suter PM, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Working group on "sepsis-related problems" of the European Society of Intensive Care Medicine. Crit Care Med. 1998;26(11):1793-800.
5. Kurtz P, Taccone FS, Bozza FA, Bastos LS, Righy C, Gonçalves B, et al. Systemic severity and organ dysfunction in subarachnoid hemorrhage: a large retrospective multicenter cohort study. Neurocrit Care. 2021;35(1):56-61.
6. Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, et al.; SAPS 3 Investigators. 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(10):1345-55.
7. Soares M, Bozza FA, Angus DC, Japiassú AM, Viana WN, Costa R, et al. Organizational characteristics, outcomes, and resource use in 78 Brazilian intensive care units: the ORCHESTRA study. Intensive Care Med. 2015;41(12):2149-60.
8. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, et al. Assessment of Clinical Criteria for Sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016;315(8):762-74.
9. Lambden S, Laterre PF, Levy MM, Francois B. The SOFA score-development, utility and challenges of accurate assessment in clinical trials. Crit Care. 2019;23(1):374.
10. de Grooth HJ, Geenen IL, Girbes AR, Vincent JL, Parienti JJ, Oudemans-van Straaten HM. SOFA and mortality endpoints in randomized controlled trials: a systematic review and meta-regression analysis. Crit Care. 2017;21(1):38.
11. Moreno R, Rhodes A, Piquilloud L, Hernandez G, Takala J, Gershengorn HB, et al. The Sequential Organ Failure Assessment (SOFA) Score: has the time come for an update? Crit Care. 2023;27(1):15.
12. Moreno R, Singer M, Rhodes A. Why the Sequential Organ Failure Assessment score needs updating? Crit Care Sci. 2024;36:e20240296en..

13. Lam BD, Struja T, Li Y, Matos J, Chen Z, Liu X, et al. Analyzing how the components of the SOFA score change over time in their contribution to mortality. Crit Care Sci. 2024;36:e20240030en.

14. Carvalho EB, Pinheiro BV, Silva PL. Peripheral arterial oxygen saturation to fraction of inspired oxygen ratio: a versatile parameter for critically ill patients. Crit Care Sci. 2025;37:e20250156

15. Cidade JP, Souza-Dantas VC, Mamfrim RB, Miranda RC, Caroli HT, Oliveira NA, et al. Advancing insights in critical COVID-19: unraveling lymphopenia through propensity score matching - Findings from the Multicenter LYMPH-COVID Study. Crit Care Sci. 2024;36:e20240236en.

16. Cidade JP, Guerreiro G, Povoa P. A clinical guide to assess the immune response to sepsis: from bench to bedside. Crit Care Sci. 2024;36:e20240179en.

17. Roepke RM, Sendagire C, Pilcher D. Challenges in using the dynamic components of the SOFA score in health care databases. Crit Care Sci. 2024;36:e20240224en.

18. Campbell SM, Hann M, Roland MO, Quayle JA, Shekelle PG. The effect of panel membership and feedback on ratings in a two-round Delphi survey: results of a randomized controlled trial. Med Care. 1999;37(9):964-8.