



Consensus Statement | Critical Care Medicine

Rationale and Methodological Approach Underlying the Development of the Sequential Organ Failure Assessment (SOFA)-2 Score

A Consensus Statement

Rui Moreno, MD, PhD; Andrew Rhodes, MD(Res); Otavio Ranzani, MD, MSc (Epi), PhD; Jorge I. F. Salluh, MD, PhD; Joana Berger-Estilita, MD, PhD; Craig M. Coopersmith, MD; Nicole P. Juffermans, MD, PhD; John Laffey, MD, DSc, MRJA; Matti Reinikainen, MD, PhD; Ary Serpa Neto, MD, MSc, PhD; Miguel Tavares, MD; Jean-François Timsit, MD, PhD; Maria Del Pilar Arias Lopez, MD; Nish Arulkumaran, MBBS, PhD; Elie Azoulay, MD, PhD; Dipayan Chaudhuri, MD, MSc; Dylan De Lange, MD, PhD; Jan De Waele, MD, PhD; Claudia C. Dos Santos, MD, MSc; Bin Du, MD; Sharon Einav, MD, MSc; Ricard Ferrer, MD, PhD; Tomoko Fujii, MD, PhD; Hayley B. Gershengorn, MD; Rashan Haniiffa, MD, PhD; Mohd Shahnaz Hasan, MBBS, MAnaes; Steve Hollenberg, MD; Mariachiara Ippolito, MD; Christian Jung, MD, PhD; Mikhail Kirov, MD, PhD; Inès Lakbar, MD, PhD; Jeffrey Lipman, MBBS, MD, DMed (Res); Vincent Liu, MD; Suzana Margareth Lobo, MD, PhD; Greg S. Martin, MD, MSc; Philipp Metnitz, MD, PhD; Sheila N. Myatra, MD; Simon Oczkowski, MD, MSc, MHS; José-Artur Paiva, MD, PhD; Fathima Paruk, MD, Crit Care(SA), PhD; Pirkka T. Pekkarinen, MD, PhD; David Pilcher, MBBS; Lise Piquilloud, MD, PhD; Anssi Pölkki, MD, PhD; Halle C. Prescott, MD, MSc; Annika Reintam Blaser, MD, PhD; Ederlon Rezende, MD; Chiara Robba, MD, PhD; Bram Rochweg, MD, MSc; Edward J. Schenck, MD, MS; Cornelius Sendagire, MD; Moses Siaw-Frimpong, MD; Andrew J. Simpkin, MD; Márcio Soares, MD, PhD; Charlotte Summers, BM, PhD; Wojciech Szczeklik, MD, PhD; Jukka Takala, MD, PhD; Jean-Louis Vincent, MD, PhD; Julia Wendon, MBChB; Fernando G. Zampieri, MD, PhD; Manu Shankar-Hari, MD, PhD; Mervyn Singer, MD

Abstract

IMPORTANCE The Sequential Organ Failure Assessment (SOFA) score was published in 1996 to describe organ dysfunction in critically ill adult patients in a readily quantifiable and sequential manner. Considerable changes have occurred over the last 3 decades in the use of organ support drugs and devices and in patient outcomes, necessitating revision of the score.

OBJECTIVES To develop definitions of organ dysfunction that reflect current understanding and to identify representative variables to generate a revised SOFA score (SOFA-2) of individual organ dysfunction.

EVIDENCE REVIEW A task force of experts in intensive care medicine and epidemiology generated definitions of organ dysfunction, identified relevant variables (physiological and laboratory data specific to the organ system, pharmacological and mechanical organ support), and proposed a 0 to 4-point grading of dysfunction severity through meetings, Delphi processes, and explicit rules, informed by data synthesis, including systematic reviews and meta-analysis. Variables were tested in 2 validation exercises using separate datasets totaling 3.34 million patients within 10 representative databases from diverse geographical and socioeconomic settings to assess distribution and predictive validity (mortality at intensive care unit discharge).

FINDINGS A total of 60 experts participated, with 18 (30%) female participants. Overall, 65 countries were represented, with 33 (51%) from Europe and Central Asia, 13 (20%) from North America; and 8 (12%) from Latin America and the Caribbean. The physiological variables within the 6 organ systems used in the original SOFA score were retained, although some categories were renamed (ie, central nervous system was changed to brain, renal to kidney, coagulation to hemostasis, and hepatic to liver). Revisions of organ support drug and device variables were made to reflect current practice. Alternative variables were added for instances when laboratory data and/or organ support interventions would be inaccessible (eg, in some low-resource settings) or not indicated (eg, ceiling of treatment). Some point cutoff thresholds were modified based on evidence from systematic reviews and data analyses. Scores could not be developed for 2 additional organ systems (gastrointestinal and immune) due to insufficient data, complexity, or lack of content and

(continued)

Key Points

Question Can an update to the Sequential Organ Failure Assessment (SOFA) score that reflects current clinical practice demonstrate content, construct, and predictive validity?

Findings In this consensus statement, a modified Delphi process, systematic reviews, and analysis of physiological, laboratory and organ support data conducted by 60 experts and using data from 3 339 470 patients in adult intensive care units in 9 countries generated an updated SOFA score (SOFA-2). This update provided new definitions, new variables, and revised thresholds to categorize the severity of organ dysfunction.

Meaning The SOFA-2 score captures contemporaneous clinical practice of organ support and organ dysfunction-associated outcomes in a large, geographically and socioeconomically diverse population of critically ill adults receiving critical care.

+ [Multimedia](#)

+ [Supplemental content](#)

+ [Related article at jama.com](#)

Author affiliations and article information are listed at the end of this article.

Open Access. This is an open access article distributed under the terms of the CC-BY License.

Abstract (continued)

predictive validity for the variables assessed. Explicit rules were developed to facilitate scoring consistency.

CONCLUSIONS AND RELEVANCE Through a methodologically robust development process, the SOFA-2 score offers updated definitions to describe organ dysfunction in adult patients requiring critical care and readily quantifiable criteria to grade the degree of dysfunction in individual organ systems. This score considers contemporaneous changes in patient management and outcomes.

JAMA Network Open. 2025;8(10):e2545040. doi:10.1001/jamanetworkopen.2025.45040

Introduction

The Sequential Organ Failure Assessment (SOFA) score, published in 1996, was designed to describe organ dysfunction in patients with sepsis¹ and subsequently expanded to include all critically ill patients.²⁻⁵ The original score (hereafter, SOFA-1) consisted of 6 organ systems: respiratory, cardiovascular, central nervous system, renal, coagulation, and hepatic. Each organ had a 5-point score (0-4), ranging from normal functioning through increasing degrees of organ impairment, from dysfunction to failure. The categories comprised commonly measured physiological and laboratory variables and, in some organ systems, the intensity of pharmacological or mechanical organ support.

SOFA-1 has since become widely integrated into clinical research and is used for multiple purposes, including epidemiological studies, trial entry criteria, evaluation of treatment response, and as the clinical criteria for the latest definition of sepsis (Sepsis-3).⁶ Here, sepsis is characterized by a likely infection triggering an increase in SOFA score of 2 or more points greater than the patient's normal baseline. Some clinicians use the score and temporal change (ie, delta SOFA) within their daily clinical practice.

Importantly, SOFA-1 was designed to describe organ dysfunction in a readily quantifiable and repeatable manner. This score is distinct from more complex prognostication scores, such as the Acute Physiology and Chronic Health disease Classification System (APACHE) and Simplified Acute Physiology Score (SAPS), that incorporate weighted physiological variables; individual patient characteristics, such as age and severe chronic disease; diagnosis; and type of intensive care unit (ICU) admission.^{7,8} Accordingly, the SOFA-1 scores for each individual organ were not weighted on their relative prognostic impact.⁹ The variables and point score thresholds were empirically selected, and the association with ICU outcomes (as a measure of predictive validity) was subsequently confirmed in multiple prospective and retrospective studies.²⁻⁵

Nonetheless, after nearly 30 years, the original SOFA-1 score requires a major update. Critical care management has evolved significantly, with expanded access to noninvasive respiratory supports and renal replacement therapy, widened use of cardiovascular drugs and devices that were either unavailable or not considered in SOFA-1, and less use of drugs and devices that are now largely discarded (eg, dopamine), among other changes.¹⁰ Patient outcomes have also improved over time,^{11,12} potentially affecting the strength and patterns of association with SOFA-1 score categories. SOFA-1 also lacked explicit rules on handling some situations, eg, chronic (and acute on chronic) organ dysfunction; transient phenomena, such as a brief episode of cardiac arrest or endotracheal suction-induced hypoxemia; and the impact of sedation on the central nervous system domain. Furthermore, the score assumes availability of resources, such as arterial blood gas analysis, organ support drugs, and specialized equipment, that may not be available in resource-limited settings.

This article details the methodological approach and the rationale underlying the development of an updated version of the SOFA score, SOFA-2. A companion article¹³ provides an in-depth evaluation of the findings from the data analyses conducted in 3.34 million patients that underpinned and validated the recommendations made by the working groups.

Methods

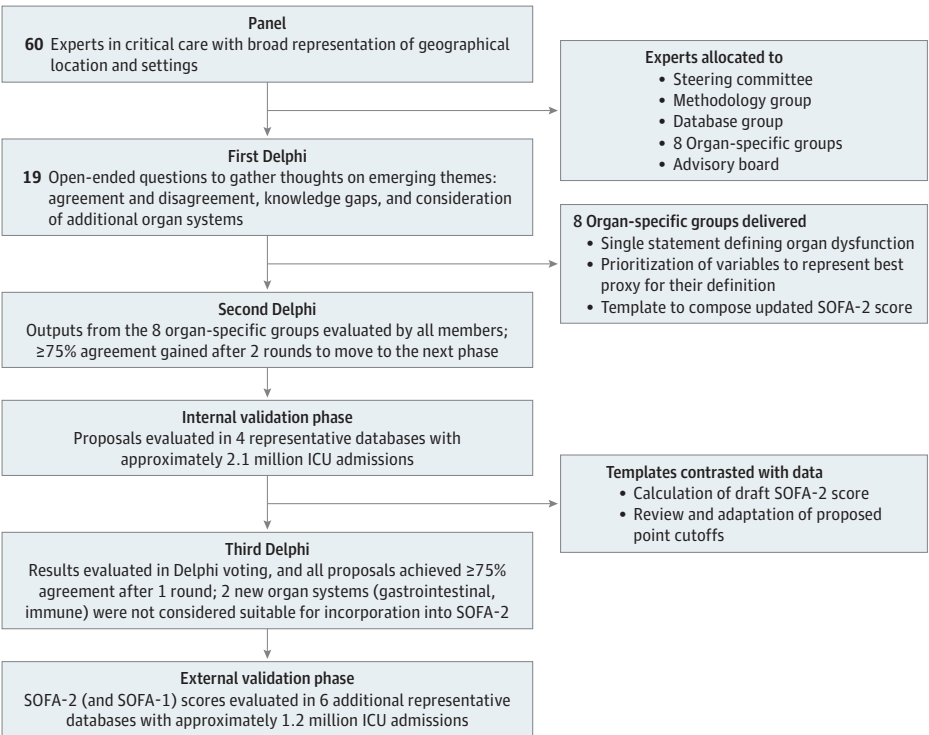
The process of updating and validating SOFA-2 is summarized in the **Figure**. This new iteration involved a large, diverse group of expert intensive care clinicians and methodologists, with strong representation from low-resource settings and different health care systems. A robust methodological approach was adopted that included a modified Delphi (mDelphi) method, generating definitions of organ dysfunction and informing variable selection for testing; systematic reviews including meta-analyses; and 2 data validation phases using multicenter registry and electronic health care record (EHR) databases from upper, upper-middle and low-middle income countries. Two new organ systems (gastrointestinal and immune) not scored in SOFA-1 were also considered for inclusion.

Stage 1: Panel of Experts

The steering committee invited 60 experts in critical care medicine with strong publication records and diversity in geographical location, socioeconomic settings, and gender to participate in the development of SOFA-2 (eTable 1 in the [Supplement](#)). In addition, independent methodology and data validation groups were established with the aims of providing guidance and expertise to the design, analysis, and interpretation of results. An advisory group composed of long-serving intensivists opinion leaders was also established.

An mDelphi approach informed all stages of development of the SOFA-2 score. The Delphi process started with briefing meetings conducted online. The guiding principles were that chosen variables should be (1) clearly indicative of the degree of organ dysfunction, independent of illness type but as specific as possible for that organ; (2) routinely measured, objective, reproducible, and globally feasible, including within low-resource settings; (3) quantitative (ideally a continuous variable or a predefined category); and (4) associated with short-term mortality (ie, predictive validity), supported by the literature. No specific hierarchy was used. The Delphi questions were

Figure. Study Flowchart



ICU indicates intensive care unit; SOFA, Sequential Organ Failure Assessment.

drafted by 2 of us (O.R. and M.S.-H.); responses were distilled by 1 of us (O.R.) and then assessed by the steering and methodology committees and domain group heads.

Stage 2: First mDelphi

The first mDelphi used a questionnaire approach, with 19 open-ended questions (eTable 2 in the [Supplement](#)) circulated to all expert members via online surveys and grouped into 3 domains. The first domain explored the current state of knowledge of the purpose of SOFA-1, including assumptions made. The second domain explored the validity, reliability, and feasibility of SOFA-1. The third domain focused on informing the development of SOFA-2, including representation of organ systems, guidance on operationalization, and additional areas of uncertainty that needed addressing. Responses provided the basis for further discussion and decisions. This first mDelphi had only 1 round.

Stage 3: Group Proposals

After the first mDelphi, experts were allocated to 8 groups, representing the 6 organ systems represented in SOFA-1 (in SOFA-1: respiratory, cardiovascular, central nervous system, renal, coagulation, and hepatic; in SOFA-2: brain, respiratory, cardiovascular, liver, kidney, and hemostasis) and 2 new systems (immune and gastrointestinal). Each organ-system group had a chair and an expert in systematic reviews, plus at least 4 geographically spread intensivists who were not expert in that specific domain to reduce academic bias and firmly held convictions.

At this stage, the organ system groups met individually to perform 3 tasks: (1) to generate a descriptive definition of organ dysfunction; (2) to agree on whether each of the 6 organ dysfunctions from SOFA-1 should be retained within SOFA-2 and whether gastrointestinal and immune organ dysfunction should be potentially incorporated; and (3) to generate a set of variables to represent the proposed organ dysfunction definition and a hierarchy of the chosen variables. These could include variables already included in SOFA-1 or new variables, including biochemical, pharmacological, or mechanical support, not represented within SOFA-1. Only organ dysfunction in adults (age ≥ 18 years) was considered.

Each organ system group met with members of the methodology committee to refine the descriptive definition of organ dysfunction, to prioritize proposed variables to mirror the refined definition, to select variables to take forward to systematic reviews and analyses within databases, and to propose organ dysfunction categories (within a 0-4-point score) to test within databases. Groups were asked to consider alternative variables for use in low-resource settings where diagnostic tests, therapies, or mechanical organ support devices may be unavailable. Rules and exceptions were agreed on (eTable 3 in the [Supplement](#)).¹⁴⁻²¹

Key principles underpinning the selection of variables representing impaired organ function or acute illness were:

- Measured routinely and frequently in general critical care practice;
- Objective and specific for that organ;
- Globally feasible, including in low-resource settings;
- Readily available from EHRs or paper medical records; and
- Had a well-established association with mortality.

Predictive validity was framed as a measure of discrimination against an outcome (ie, whether the score correctly ranks individuals by risk), such as mortality.²² ICU mortality was chosen by consensus, as in SOFA-1. In this framing we were explicit that predictive accuracy was not the primary goal of the analyses used to derive SOFA-2; nonetheless, increasing organ dysfunction should reflect mortality risk. The quantitative variable could be a continuous variable or a predefined category but, ideally, not a marker of comorbidity. All groups received a 9-step guide to aid discussions and development of outputs (eTable 4 in the [Supplement](#)).

Stage 4: Systematic Reviews

To support group decisions, systematic reviews were conducted, if needed, to ensure an association between a variable and short-term mortality; to gather knowledge to support new definitions to be used in SOFA-2; and to evaluate changing or adapting SOFA-1 variables and respective cutoffs. To promote standardization, the methodology group provided a template with minimal requirements for conduct of the systematic reviews (eTable 5 in the [Supplement](#)). Meta-analysis or quantitative analysis was encouraged, but if heterogeneity was too great, a narrative synthesis was performed. Five systematic reviews were performed, covering (1) the equivalence between the ratios of arterial oxygen tension (PaO_2) to fraction of inspired oxygen (FiO_2) and arterial oxygen saturation (SpO_2) to FiO_2 ; (2) the association of total white blood cell count and lymphocyte count with short-term mortality (eg, ICU, hospital, 28-day mortality); (3) criteria for initiation of renal replacement therapy; (4) potential thresholds for the association between norepinephrine (noradrenaline) dose and short-term mortality; and (5) associations between intra-abdominal pressure, gastrointestinal bleeding, and short-term mortality. These systematic reviews will be published in detail in future publications. During stages 3 and 4, in addition to frequent feedback from the methodology team, each group received feedback during regular meetings between group chairs or their representatives and the steering, methodology, and data validation committees.

Stage 5: Second mDelphi

The second mDelphi evaluated expert panel agreement on the 3 outputs from stage 3. Each member voted on questions for all organ systems. Examples are shown in eTable 6 in the [Supplement](#). Scales, thresholds, the number of rounds, and actions were prespecified, and the iterative processes was conducted through online surveys. Although a maximum of 3 rounds were allowed, only 2 were needed. The 8 organ systems were presented in random order. Each output was followed by 5-point Likert scale questions (strongly agree, agree, neither agree nor disagree, disagree, or strongly disagree). The second Delphi used a 3-point Likert scale (strongly agree, agree, or disagree). Consensus at this stage was defined as at least 75% of answers agreeing or strongly agreeing.²³ When evaluating agreement on proxy and feasibility of a proposed variable, both features needed to meet the threshold of at least 75% agreement.

Stage 6: Internal Validation

Informed by the Delphi results, prospective validation of individual organ-level variables was performed to derive SOFA-2 criteria using 3 large multicenter databases: the Australian and New Zealand Intensive Care Society (ANZICS) Adult Patient Database (Australia/New Zealand),²⁴ Austrian Center for Documentation and Quality Assurance in Intensive Care (ASDI; Austria),²⁵ and Organizational Characteristics in Critical Care (ORCHESTRA; Brazil)²⁶ as well as from the EHRs of ICU patients admitted within the Kaiser Permanente North California (KPNC) health care system (eAppendix in the [Supplement](#)).²⁷ These analyses were standardized and centrally coordinated by the methodology and validation committees to ensure consistency. To explore the point cutoffs proposed by the experts, a generalized additive model was applied with penalized smoothing splines and classification and regression tree models with cross-validation. Further details are provided in a companion article.¹³

Findings from this validation process were collated and finalized with the group leads before presentation to the entire SOFA-2 group. Data included frequency of occurrence, mortality rates at ICU discharge within each point threshold, and a comparison between SOFA-1 and the proposed SOFA-2 draft score. A new Delphi process was then performed to confirm the overall agreement on this latest SOFA-2 proposal.

Stage 7: Third mDelphi

Using data from stage 6, a revised table for SOFA-2 was proposed after discussion between the steering, data validation, and methods committees and the chairs of each organ-system group. A

third mDelphi was then conducted to evaluate whether the expert panel agreed with this proposal. eTable 7 in the [Supplement](#) provides examples of the research questions. Rules and principles used in the second mDelphi were followed. The third mDelphi was completed after 1 round.

Stage 8: External Validation

An external data validation was conducted on 4 new registries (GiViTI-PROSAFE [Italy],²⁸ Japanese Intensive care Patient Database [JIPAD; Japan],²⁹ Nepal Intensive Care Research Foundation [NICRF; Nepal],³⁰ and OUTCOMEREA [France]³¹) and 2 new multicenter EHRs (electronic Intensive Care Unit Collaborative Research Database [eICU; US]³² and GiViTI-MargheritaTre [Italy]³³) (eAppendix in the [Supplement](#)). A longitudinal analysis of SOFA-2 was performed for score distribution and predictive validity on 3 datasets with available data (eICU, KPNC, and OUTCOMEREA). Additional detail on methods, comprehensive data analyses, other approaches for handling missing data, including longitudinal assessments, and sensitivity analyses are described in the companion article.¹³ As a final step, feedback and approval of the final SOFA-2 score were sought from the entire SOFA-2 group.

Results

The timeline and process for developing and validating the SOFA-2 score is shown in **Table 1**. A total of 60 experts participated, with 18 (30%) female and 42 (70%) male participants. Overall, 65 countries were represented, with 33 (51%) from Europe and Central Asia; 13 (20%), North America; 8 (12%), Latin America and the Caribbean; 5 (8%), East Asia and the Pacific; 3 (5%), sub-Saharan Africa; 2 (3%), South Asia; and 1 (2%), Middle East and North Africa.

First and Second mDelphi

The emerging themes from the first mDelphi emphasized (1) the need for SOFA-2 to mirror current clinical practice, while maintaining the simplicity of SOFA-1; (2) the provision of a clear description of how to define and operationalize scoring in future work; and (3) broad support for development of scores for gastrointestinal and immune dysfunction (eTable 8 in the [Supplement](#)). Potential variables were suggested (eTable 9 in the [Supplement](#)), and after detailed discussion, prioritized variables were voted on in the first round of the second mDelphi (eTable 10 in the [Supplement](#)). Fifty-nine complete responses were received from 60 invitations (98% response rate). Except for the immune domain, the proposed organ definitions achieved high agreement, greater than the 75% threshold (eFigure 1 in the [Supplement](#)). Regarding the proxy and feasibility of each variable, 12 of 21 (57%) achieved the threshold for both features, while some achieved consensus for proxy but not feasibility or vice versa. At this stage, gastrointestinal and immune domains did not achieve agreement in both features for any variable. Agreement on the combined SOFA-2 proposal was 69% (41 of 59).

In the second round of the second mDelphi, updated variables (eTable 10 in the [Supplement](#)) were voted on. Based on comments from the previous round, the central nervous system and coagulation domains were relabeled as brain and hemostasis. Fifty-four complete responses were obtained from 60 invitations (90% response rate). All domains achieved agreement greater than the 75% threshold, while agreement on the combined SOFA-2 proposal was 93% (50 of 54) (eFigure 2 in the [Supplement](#)). **Table 2** shows the agreed definitions for each organ dysfunction and the proposed variables for data analyses and score construction.

Third mDelphi

Before the third mDelphi, the internal validation analyzed 2 098 356 patients from the 4 cohorts.¹³ After collating results from the data validation exercise, a proposal for SOFA-2 with 6 organ systems was presented to all experts in a third mDelphi. In the first round, 57 complete responses were received (95% response rate). All 6 remaining organ systems achieved agreement greater than the threshold, ranging from 84% (48 of 57) for liver to 93% (53 of 57) for respiratory. Overall agreement on the final SOFA-2 proposal was 95% (54 of 57) (eFigure 3 in the [Supplement](#)).

Table 1. Timeline for Development and Validation of the SOFA-2 Score

Step	Period
Steering group constitution	Mar 2022
Invitation to experts to join groups and committees	Mar 2022 to Feb 2023
Background discussions	Mar 2022 to Feb 2023
First mDelphi	Mar 2023
Definitions, prioritization of variables, systematic reviews	May 2023 to Feb 2024
Second mDelphi	
First round	Mar 2024
Second round	Apr 2024
Internal validation	May to Sept 2024
Third mDelphi	Sept 2024
External validation	Oct 2024 to Feb 2025

Abbreviations: mDelphi, modified Delphi; SOFA, Sequence Organ Failure Assessment.

Gastrointestinal and immune dysfunctions were not included within SOFA-2 for several reasons. There were insufficient data available within any of the databases and EHRs for gastrointestinal dysfunction to validate the proposed score. Data that could be extracted also showed a lack of predictive validity. For immune dysfunction, both total lymphocyte and total white cell counts on day 1 showed good predictive validity for ICU mortality with a U-shaped curve. However, the content validity criterion could not be met, as neutrophilia often represents an appropriate response to an inflammatory insult such as sepsis, trauma, or surgery, rather than immune dysfunction. Second, immune dysfunction can exist despite normal leukocyte counts; functional tests (eg, phagocytic capacity) and markers of immune suppression (eg, monocyte HLA-DR) are not performed in routine clinical practice, even in high-income countries.

Second Data Validation and Endorsement

The external data validation included 1 241 114 patients from the 6 further cohorts.¹³ Findings were then distributed to the entire SOFA-2 group who endorsed the final SOFA-2 score. The score is shown in **Table 3**, with footnotes providing clarifications to enhance consistency in scoring. **Table 4** displays differences between SOFA-2 and the SOFA-1 score. The main changes made were (1) incorporation of new drugs and devices to support organ systems to reflect current standard management practices, (2) new score cutoffs that align better with mortality risk, and (3) alternative variables to be used when primary options are not being used or unavailable (eg, arterial blood gas analysis, renal replacement therapy) or not indicated (eg, ceiling of treatment). Explicit rules were also provided, eg, neurological scoring when sedation is being administered, inclusion of chronic organ dysfunction, handling of data missingness.

Table 2. Definitions for Each Organ Dysfunction

Organ domain	Definition of dysfunction	Proposed variables to take forward
Brain	Disruption or impairment in the sensory, processing, or motor functioning of the brain.	Glasgow Coma Scale; use of drugs for delirium according to international guidelines ¹⁵
Respiratory	Functional impairment due to loss of gas exchange capacity due to primary respiratory illness or a systemic illness causing secondary respiratory failure. This results in reduction of the oxygenation capacity of respiratory system (ie, impaired oxygenation), causing hypoxemia.	PaO ₂ :FiO ₂ ratio; use of ventilatory support; ECMO
Liver	Loss of synthetic and catabolic functions, impairing the ability for normal blood clotting, protein synthesis, and liver-mediated toxin clearance. Liver dysfunction commonly manifests as elevated serum bilirubin levels and hepatic encephalopathy.	Bilirubin level
Hemostasis	Hemostasis failure caused by an acute eliciting factor and characterized by consumption of platelets with concomitant consumption of coagulation factors and anticoagulant proteins. Severity can range from a mild decrease in platelet count to overt disseminated intravascular coagulation. Clinically manifests as an increased risk for both bleeding and thromboembolic events. Hemostasis dysfunction can also be acute-on chronic.	Platelet levels
Cardiovascular	Insufficient perfusion leading to tissue hypoxia caused by low blood pressure and/or low cardiac output.	Blood pressure; doses of norepinephrine and epinephrine; use of other vasopressor and inotropic drugs; use of any type of mechanical cardiac support
Kidney	Kidney impairment manifests as a loss of function and/or structural damage, related to multiple coexisting etiologies, including renal and/or extrarenal.	Creatinine level; urine output; use of renal replacement therapy for renal dysfunction indications
Gastrointestinal	Functional impairment that may include disturbances in motility, absorption, mucosal integrity, and mesenteric perfusion. These impairments may contribute to patient morbidity, aggravate multi-organ failure, and further deteriorate to life-threatening emergencies, such as abdominal compartment syndrome, acute mesenteric ischemia, or major hemorrhage.	Oral, enteral, and parenteral nutrition; grade III-IV intraabdominal hypertension; and life-threatening acute gastrointestinal condition
Immune	Dysregulation of the host immune response that involves complex pathophysiological mechanisms, such as release of cytokines and adhesion molecules and activation of complement and coagulation pathways, leading to inappropriate and potentially life-threatening inflammation. This ultimately results in a state of immune paralysis that increases the risk of nosocomial infection, prolonged organ dysfunction, and death.	Total white cell count, lymphocyte count

Abbreviations: ECMO, extracorporeal membrane oxygenation; FiO₂, fraction of inspired oxygen; PaO₂, partial pressure of oxygen.

Discussion

The SOFA-2 score was developed to provide a common language to describe organ dysfunction in critically ill adult patients, reflecting contemporary clinical practice. SOFA-2 incorporates both commonly measured and readily collectable physiological and laboratory variables as well as the degree of organ support required to maintain the physio-biochemical variables at normal, or acceptably abnormal, levels. Organ support has been expanded to include present-day management

Table 3. The SOFA-2 Score^{a,b}

Organ system	Score				
	0	1	2	3	4
Brain ^{c,d}	GCS 15 (or thumbs-up, fist, or peace sign)	GCS 13-14 (or localizing to pain) ^d or need for drugs to treat delirium ^e	GCS 9-12 (or withdrawal to pain)	GCS 6-8 (or flexion to pain)	GCS 3-5 (or extension to pain, no response to pain, generalized myoclonus)
Respiratory ^f	PaO ₂ :FiO ₂ ratio >300 mm Hg (>40 kPa)	PaO ₂ :FiO ₂ ratio ≤300 mm Hg (≤40 kPa)	PaO ₂ :FiO ₂ ratio ≤225 mm Hg (≤30 kPa)	PaO ₂ :FiO ₂ ratio ≤150 mm Hg (≤20 kPa) and advanced ventilatory support ^{g,h}	PaO ₂ :FiO ₂ ratio ≤75 mm Hg (≤10 kPa) and advanced ventilatory support ^{g,h} or ECMO ⁱ
Cardiovascular ^{j,k,l,m}	MAP ≥70 mm Hg, no vasopressor or inotrope use	MAP <70 mm Hg, no vasopressor or inotrope use	Low-dose vasopressor (sum of norepinephrine and epinephrine ≤0.2 µg/kg/min) or any dose of other vasopressor or inotrope	Medium-dose vasopressor (sum of norepinephrine and epinephrine >0.2 to ≤0.4 µg/kg/min) or low-dose vasopressor (sum norepinephrine and epinephrine ≤0.2 µg/kg/min) with any other vasopressor or inotrope	High-dose vasopressor (sum of norepinephrine and epinephrine >0.4 µg/kg/min) or medium-dose vasopressor (sum of norepinephrine and epinephrine >0.2 to ≤0.4 µg/kg/min) with any other vasopressor or inotrope or mechanical support ^{l,n}
Liver	Total bilirubin ≤1.20 mg/dL (≤20.6 µmol/L)	Total bilirubin ≤3.0 mg/dL (≤51.3 µmol/L)	Total bilirubin ≤6.0 mg/dL (≤102.6 µmol/L)	Total bilirubin ≤12.0 mg/dL (≤205 µmol/L)	Total bilirubin >12.0 mg/dL (>205 µmol/L)
Kidney	Creatinine ≤1.20 mg/dL (≤110 µmol/L)	Creatinine ≤2.0 mg/dL (≤170 µmol/L) or urine output <0.5 mL/kg/h for 6-12 h	Creatinine ≤3.50 mg/dL (≤300 µmol/L) or urine output <0.5 mL/kg/h for ≥12 h	Creatinine >3.50 mg/dL (>300 µmol/L) or urine output <0.3 mL/kg/h for ≥24 h or anuria (0 mL) for ≥12 h	Receiving or fulfils criteria for RRT ^{o,p,q} (includes chronic use)
Hemostasis	Platelets >150 × 10 ³ /µL	Platelets ≤150 × 10 ³ /µL	Platelets ≤100 × 10 ³ /µL	Platelets ≤80 × 10 ³ /µL	Platelets ≤50 × 10 ³ /µL

Abbreviations: ECMO, extracorporeal membrane oxygenation; GCS, Glasgow Coma Scale; MAP, mean arterial pressure; PaO₂:FiO₂, ratio of partial pressure of oxygen to fraction of inspired oxygen; RRT, renal replacement therapy; SOFA, Sequential Organ Failure Assessment.

^a The final score is obtained by summing the maximum points from each of the 6 organ systems individually within a 24-hour period, with the final score ranging from 0 to 24.

^b For missing values at day 1, the general recommendation is to score these as 0 points. This may vary for specific purposes, eg, for research, where multiple imputation is often used. For sequential scoring, if there are missing data after day 1, it is recommended to carry forward the last observation, the rationale being that nonmeasurement suggests stability.

^c For sedated patients, use the last recorded GCS before sedation. If the previous GCS is unknown, score 0.

^d When not possible to evaluate the 3 domains of GCS, use the best achieved score in the motor-scale domain.

^e If receiving drug treatment for delirium (short- or long-term), score 1 point even if GCS is 15. For relevant drugs, see International PADIS Guidelines.¹⁵

^f Use arterial oxygen saturation (SpO₂) to FiO₂ ratio only when PaO₂:FiO₂ ratio is unavailable and where SpO₂ is less than 98%. Cutoffs: 0 points, greater than 300 mm Hg; 1 point, less than or equal to 300 mm Hg; 2 points, less than or equal to 250 mm Hg; 3 points, less than or equal to 200 mm Hg with ventilatory support; and 4 points, less than or equal to 120 mm Hg with ventilatory support or ECMO.

^g Advanced ventilatory support is defined as receipt of high-flow nasal cannula, continuous positive airflow pressure, bilevel positive airway pressure, noninvasive ventilation, invasive mechanical ventilation, or long-term home ventilation. This is required to score 3 to 4 points, in addition to the PaO₂:FiO₂ or SpO₂:FiO₂ ratio being within the specified range. Changes in PaO₂:FiO₂ or SpO₂:FiO₂ within less than 1 hour (eg, after suctioning) should not be considered.

^h Patients not receiving advanced respiratory support can score a maximum 2 points unless ventilatory support is (1) not available or (2) precluded due to the ceiling of treatment; if so, severity is scored by PaO₂:FiO₂ or SpO₂:FiO₂ ratio.

ⁱ If used for respiratory therapy, ECMO (all forms) should be scored 4 in the respiratory component (regardless of PaO₂:FiO₂ ratio), but not in the cardiovascular component. If used for cardiovascular indications (all forms), it should be automatically scored in both the cardiovascular and the respiratory systems.

^j Vasopressor medication is only scored if given by continuous intravenous infusion for at least 1 hour.

^k Norepinephrine is usually dispensed as the salt (eg, hemitartrate or bitartrate).¹⁷ Dose should be expressed as the base. One mg of norepinephrine base is equivalent to 2 mg of norepinephrine bitartrate monohydrate, 1.89 mg of the anhydrous bitartrate (also called hydrogen tartrate, acid tartrate, or tartrate), and 1.22 mg of the hydrochloride.

^l If dopamine is used as a single vasopressor, scoring is based on the following cutoffs: 2 points (≤20 µg/kg/min); 3 points (>20 to ≤40 µg/kg/min); 4 points (>40 µg/kg/min). These cutoffs are based on norepinephrine equipotency studies.¹⁸⁻²⁰

^m Where vasoactive drugs are unavailable or precluded due to a ceiling of treatment, use the following MAP cut-offs for scoring: 0 points, 70 mm Hg or greater; 1 point, 60 to 69 mm Hg; 2 points, 50 to 59 mm Hg; 3 points, 40 to 49 mm Hg; and 4 points, less than 40 mm Hg.

ⁿ Any type of mechanical cardiovascular support: eg, veno-arterial ECMO, intra-aortic balloon pump, left ventricular assist device, microaxial flow pump.

^o Excludes patients receiving RRT exclusively for nonrenal causes (eg, removal of toxic products, bacterial toxins, cytokines).

^p For patients not receiving RRT (eg, ceiling of treatment, machine unavailability, or decision to delay commencement), score 4 points if they otherwise meet criteria for RRT, ie, creatinine level greater than 1.2 mg/dL (>110 µmol/L) or oliguria (<0.3 mL/kg/h) for more than 6 hours plus at least 1 of either serum potassium of 6.0 mmol/L or greater or metabolic acidosis with pH of 7.20 or less and serum bicarbonate of 12 mmol/L or less.

^q For patients receiving intermittent RRT, score 4 points on days not receiving RRT until RRT use is terminated.

practices, such as noninvasive ventilation modalities, renal replacement therapy, and cardioactive drugs and devices, use of which was far more limited when the SOFA-1 score was published in 1996.¹ The major changes are the inclusion of definitions, new variables, and revision of categories in respiratory, kidney, brain, and cardiovascular organ system scores, as well as altered point score cutoffs.

A robust methodological approach was adopted with explicit rules set prior to any analysis. This included a Delphi process, conducted among a large group of intensivists representing each of the 6 populated continents, that generated definitions of organ dysfunction and the proposal of variables that offered both content validity and specificity to describe dysfunction for that organ system. After systematic reviews and meta-analyses, the predictive validity of SOFA-2, both for each organ system using a 0 to 4–point scale and the total summated score, was confirmed in large multicenter databases with 3.4 million patients admitted to ICUs across diverse geographic and socioeconomic settings.

While increasing organ dysfunction is clearly associated with mortality risk, it should be stressed that SOFA is primarily intended to describe organ dysfunction consistently across different settings and longitudinally, if required, rather than competing with mortality risk prediction scores for which more complex models already exist.^{7,8} Intentionally, therefore, the point score and underpinning variables within each organ system were not weighted nor optimized to increase discriminative performance (ie, the area under the receiver operating characteristic curve [AUROC]). The provision of explicit rules as to how variables should be handled, clarifying areas of uncertainty in SOFA-1, will facilitate consistency of scoring. Table 4 and eTable 3 in the [Supplement](#) provide the rationale for the decisions taken and the recommendations made to promote consistency. Another important addition in SOFA-2 is the provision of alternatives where the chosen variables and/or organ support (respiratory, cardiovascular, kidney) are either unavailable (eg, in low-resource settings, nonuse of an arterial line) or not indicated in patients with ceilings of treatment; this facilitates global use of the score and comparisons to be drawn.

There was a strong intent to incorporate scores for immune and gastrointestinal dysfunction into SOFA-2 given their significance in critical illness. Consensus was achieved on establishing definitions of dysfunction. However, the complexity and lack of necessary data to validate a gastrointestinal score and the lack of suitable markers to identify immune dysfunction on a routinely measured and widely available basis precluded their inclusion within SOFA-2. While increasingly

Table 4. Differences Between SOFA-2 and SOFA-1 Scores

Organ system	Change from SOFA-1
Brain	<ul style="list-style-type: none">• Incorporation of delirium (using specific drug therapy as a surrogate)• Modified point score cutoffs for Glasgow Coma Score• Renamed from central nervous system
Respiratory	<ul style="list-style-type: none">• Incorporation of noninvasive respiratory support modalities and ECMO• Modified point score cutoffs for PaO₂:FiO₂ ratio
Cardiovascular	<ul style="list-style-type: none">• Summation of doses of catecholaminergic agents• Simple incorporation of all other vasopressor and inotrope agents as well as mechanical support devices
Liver	<ul style="list-style-type: none">• Modified point score cutoff for bilirubin level• Renamed from hepatic
Kidney	<ul style="list-style-type: none">• Modified point score cutoffs for creatinine level and urine output• Incorporation of renal replacement therapy (or fulfilment of criteria)• Renamed from kidney
Hemostasis	<ul style="list-style-type: none">• Modified point score cutoff for platelet count• Renamed from coagulation/hematological
General	<ul style="list-style-type: none">• Provision of definitions for dysfunction in each organ system• Provision of alternatives if variables are either not available (eg, arterial blood gas analysis, renal replacement therapy) or not indicated (eg, ceiling of treatment)• Explicit rules for handling scoring of, eg, sedation, chronic organ dysfunction• For missing data on day 1, normal value imputation (rationale: if not measured, likely to be within normal range); for missing data after day 1, carry forward last observation (rationale: if not measured, likely to be stable or irrelevant in a clinical context)• For more detailed discussion of missing data handling, please see the accompanying SOFA-2 article¹³

Abbreviations: ECMO, extracorporeal membrane oxygenation; FiO₂, fraction of inspired oxygen; PaO₂, partial pressure of oxygen; SOFA, Sequential Organ Failure Assessment.

abnormal white blood cell and lymphocyte counts (both high and low) have good predictive ability for ICU mortality,¹³ a high count often represents an appropriate response to a severe infection rather than dysfunction. Conversely, normal immune cell counts do not exclude immune dysfunction. Of note, values falling within the extremes of the reference (ie, healthy) range for lymphocytes (1000-4000 cells/ μ L [to convert to cells $\times 10^9$ per liter, multiply by 0.001]) were associated with a 2- to 3-fold increase in ICU mortality risk compared with the nadir, seen at 2300 cells/ μ L.¹³ These markers thus fail our content validity criterion. Simple, accessible, and reliable scores for gastrointestinal and immune dysfunction nonetheless represent important research questions for future iterations of the SOFA score.

Some important research questions need to be addressed. First, it would be useful to develop and validate an updated SOFA version suitable for pediatric populations, such as the equivalents developed for pediatric sepsis.^{34,35} In the current study, we arbitrarily set an age cutoff of 18 years, recognizing that normal ranges of blood pressure and creatinine are age dependent in children and adolescents, and the use of invasive monitoring may be less frequent. Second, the validity of SOFA-2 in patients outside the ICU, eg, in the emergency department or general wards, should be assessed. We were only able to access the necessary high granularity data to compute SOFA from multicenter databases and registries of patients admitted to ICUs. Single-center EHRs were excluded for generalizability concerns. Third, the value of dynamic measures, such as delta SOFA as well as maximum SOFA and discharge SOFA,^{3,4,36} should be explored in future SOFA-2 studies to track the evolution of the pathological process organ by organ³⁷ using harmonized criteria.³⁸ Finally, SOFA was proposed as a triage tool during the COVID-19 pandemic³⁹; in our opinion any such use should be validated before implementation. The sponsoring bodies of the International Sepsis-3 Definitions should determine whether SOFA-2 should supersede SOFA-1 to better reflect current practices.

Limitations

There are several limitations to our work. Although we analyzed data from 10 large cohorts from 9 varied countries, we would strongly encourage prospective validation of the SOFA-2 score in different geographical and resource settings. We would welcome identification of any further scoring inconsistencies that can be addressed. Other than OUTCOMEREA, the registry databases lacked sequential (ie, daily) data; these data were only available in the eICU and KPNC electronic health care record databases that mainly included US patients. We also acknowledge that SOFA-2 will underestimate organ dysfunction in a minority of patients, such as liver dysfunction with a normal bilirubin level or hypercapnic respiratory failure with relatively normal oxygenation, and does not include organ systems such as immune, gastrointestinal, and neuromuscular. However, we believe the pragmatic simplicity of the score outweighs the need for all-encompassing complexity. For predictive validity, only ICU mortality was evaluated, as this was the only outcome available across all cohorts. Notably, the SOFA-1 score at ICU discharge is predictive for longer-term mortality.^{40,41} As mentioned previously, an increasing number and severity of organ dysfunctions should, for content validity, show an association with mortality risk. However, the main emphasis of SOFA is to describe organ dysfunction in a critically ill population during their ICU stay rather than aiming to compete with weighted predictive scores such as APACHE and SAPS.

Conclusions

In summary, we developed a new SOFA score to describe organ dysfunction in critical illness, accounting for changes in clinical practice and improvements in patient outcomes over the last 3 decades. A robust methodological approach was adopted with validation across 9 large databases in geographical and socioeconomically diverse settings.

ARTICLE INFORMATION

Accepted for Publication: September 30, 2025.

Published: October 29, 2025. doi:10.1001/jamanetworkopen.2025.45040

Open Access: This is an open access article distributed under the terms of the [CC-BY License](#). © 2025 Moreno R et al. *JAMA Network Open*.

Corresponding Author: Rui Moreno, MD, PhD, Hospital de São José, ULS de São José, Rua José António Serrano, Lisbon 1150-199 Lisboa, Portugal (r.moreno@mail.telepac.pt).

Author Affiliations: Hospital de São José, ULS de São José, Lisboa, Portugal (Moreno); Faculdade de Ciências Médicas de Lisboa, Nova Medical School, Lisboa, Portugal (Moreno); Faculdade de Ciências da Saúde, Universidade da Beira Interior, Covilhã, Portugal (Moreno); Adult Critical Care, St George's University Hospitals NHS Foundation Trust, London, United Kingdom (Rhodes); DataHealth Lab, Institut de Recerca Sant Pau (IR SANT PAU), Barcelona, Spain (Ranzani); ISGlobal, Barcelona, Spain (Ranzani); Pulmonary Division, Faculty of Medicine, Heart Institute, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil (Ranzani); Postgraduate Program, Instituto D'Or de Pesquisa e Ensino, Rio de Janeiro, Brazil (Salluh); Institute for Medical Education, University of Bern, Bern, Switzerland (Berger-Estilita); Department of Surgery and Emory Critical Care Center, Emory University, Atlanta, Georgia (Coopersmith); Department of Intensive Care and Laboratory of Translational Intensive Care, Erasmus Medical Center, Rotterdam, the Netherlands (Juffermans); Anaesthesia and Intensive Care Medicine, School of Medicine, University of Galway, Galway, Ireland (Laffey); Department of Anaesthesia, Galway University Hospitals, Galway, Ireland (Laffey); University of Eastern Finland, Kuopio, Finland (Reinikainen); Kuopio University Hospital, Kuopio, Finland (Reinikainen); Australian and New Zealand Intensive Care Research Centre (ANZIC-RC), School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia (Neto); Department of Intensive Care, Austin Hospital, Melbourne, Australia (Neto); Department of Critical Care Medicine, Hospital Israelita Albert Einstein, São Paulo, Brazil (Neto); Unidade Local de Saúde Santo António, Porto, Portugal (Tavares); ICBAS-Instituto de Ciências Biomédicas Abel Salazar, Porto, Portugal (Tavares); IAME (Infection Antimicrobials Modelling Evolution), UMR 1137, University Paris-Cité, Paris, France (Timsit); Medical and Infectious Diseases Intensive Care Unit, APHP, Hôpital Bichat-Claude Bernard, Paris, France (Timsit); Management, Quality and Data Committee, SATI-Q Program, Argentine Society of Intensive Care (SATI), Buenos Aires, Argentina (Arias Lopez); PICU, Hospital de Niños Ricardo Gutiérrez, Buenos Aires, Argentina (Arias Lopez); Bloomsbury Institute of Intensive Care Medicine, Division of Medicine, University College London, London, United Kingdom (Arulkumaran); APHP, Paris City University, Saint-Louis Hospital, Paris, France (Azoulay); Division of Critical Care, McMaster University and Department of Critical Care, St Joseph's Healthcare, Hamilton, Ontario, Canada (Chaudhuri); Intensive Care Unit and Dutch Poison Information Center, University Medical Center Utrecht, the Netherlands (De Lange); Department of Intensive Care Medicine, Ghent University Hospital, Ghent, Belgium (De Waele); Interdepartmental Division of Critical Care, University of Toronto, Ontario, Canada (Dos Santos); Keenan Research Centre for Biomedical Science, Unity Health Toronto, Toronto, Ontario, Canada (Dos Santos); Medical Intensive Care Unit, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China (Du); Maccabi Healthcare Services, Sharon Region, Hebrew University Faculty of Medicine, Tel Aviv, Israel (Einav); Department Intensive Care and SODIR-VHIR Research Group, Vall d'Hebron University Hospital, Barcelona, Spain (Ferrer); Department of Intensive Care, Jikei University Hospital, Tokyo, Japan (Fujii); Division of Pulmonary, Critical Care and Sleep Medicine, University of Miami Hospital and Clinics, Miami, Florida (Gershengorn); Centre for Inflammation Research, University of Edinburgh, Edinburgh, Scotland (Haniffa); Department of Intensive Care, University College London Hospitals NHS Foundation Trust, London, United Kingdom (Haniffa); Department of Anaesthesiology, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia (Hasan); Department of Medicine, Emory University School of Medicine, Atlanta, Georgia (Hollenberg); Department of Precision Medicine in Medical, Surgical and Critical Care (Me.Pre.C.C.), University of Palermo, Palermo, Italy (Ippolito); Department of Cardiology, Pulmonology, and Vascular Medicine, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany (Jung); Cardiovascular Research Institute Dusseldorf (CARID), Heinrich-Heine-University Dusseldorf, Dusseldorf, Germany (Jung); Department of Anesthesiology and Intensive Care Medicine, Northern State Medical University, Arkhangelsk, Russia (Kirov); Anesthesia and Critical Care Department B, Saint Eloi Teaching Hospital, PhyMedExp, University of Montpellier, INSERM U1046, Montpellier, France (Lakbar); Jamieson Trauma Institute, Royal Brisbane and Women's Hospital, University of Queensland, Brisbane, Australia (Lipman); Nîmes University Hospital, Nîmes, France (Lipman); Kaiser Permanente Northern California Division of Research and Department of Critical Care, Kaiser Permanente Medical Center, Santa Clara (Liu); Division of Intensive Care, Hospital de Base-São Jose do Rio Preto and Faculdade de Medicina de São Jose do Rio Preto, Brazil (Lobo); Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, Emory University School of Medicine, and Grady Memorial Hospital, Atlanta, Georgia (Martin); Department of General Anaesthesiology and Intensive Care, Medical University of Graz, Graz, Austria (Metnitz); Department of Anesthesiology, Critical Care and Pain, Tata Memorial Hospital, Homi Bhabha National Institute, Mumbai, India (Myatra); Department of Medicine and Department of Health Research Methods, Evidence, and Impact, McMaster

University, Hamilton, Ontario, Canada (Oczkowski); Serviço de Medicina Intensiva, Unidade Local de Saúde São João, Faculty of Medicine, University of Porto and Grupo de Investigação e Desenvolvimento em Infecção e Sepsis, Porto, Portugal (Paiva); Department of Critical Care, University of Pretoria and Department of Critical Care, Steve Biko Academic Hospital, Pretoria, South Africa (Paruk); Division of Intensive Care, Department of Anaesthesiology and Intensive Care, University of Helsinki and Helsinki University Hospital, Helsinki, Finland (Pekkarinen); The Australian and New Zealand Intensive Care Society (ANZICS) Centre for Outcomes and Resources Evaluation, Australian and New Zealand Intensive Care Research Centre (ANZIC-RC), School of Public Health and Preventive Medicine, Monash University, and Department of Intensive Care, Alfred Health, Prahran, Victoria, Australia (Pilcher); Adult Intensive Care Unit, University Hospital and University of Lausanne, Lausanne, Switzerland (Piquilloud); Department of Anaesthesiology and Intensive Care, Kuopio University Hospital and University of Eastern Finland, Kuopio, Finland (Pölkki); Department of Internal Medicine, University of Michigan, Ann Arbor (Prescott); VA Center for Clinical Management Research, Ann Arbor, Michigan (Prescott); University of Tartu, Estonia, and Luzerner Kantonsspital, Lucerne, Switzerland (Blaser); Serviço de Terapia Intensiva na IAMSPE-Instituto de Assistência Médica ao Servidor Público Estadual, São Paulo, Brazil (Rezende); IRCCS, Policlinico San Martino, Genova, Italy (Robba); Dipartimento di Scienza Chirurgica Diagnostica Integrate, University di Genova, Genova, Italy (Robba); Department of Medicine, and Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, Ontario, Canada (Rochweg); Division of Pulmonary and Critical Care Medicine, Department of Medicine, New York-Presbyterian Hospital, Weill Cornell Medical Center, Weill Cornell Medicine, New York, New York (Schenck); Instituto D'Or de Pesquisa e Ensino, Rio de Janeiro, Brazil (Sendagire); Makerere University, Kampala, Uganda (Sendagire); Directorate of Anaesthesia and Intensive Care, Komfo Anokye Teaching Hospital, Kumasi, Ghana (Siaw-Frimpong); School of Mathematical and Statistical Sciences, University of Galway, Galway, Ireland (Simpkin); SFI Insight Centre for Data Analytics, Galway, Ireland (Simpkin); Department of Critical Care and Postgraduate Program, D'Or Institute for Research and Education, Rio de Janeiro, Brazil (Soares); Victor Phillip Dahdaleh Heart & Lung Research Institute, University of Cambridge, Cambridge, United Kingdom (Summers); Jagiellonian University Medical College, Centre for Intensive Care and Perioperative Medicine, Krakow, Poland (Szczeklik); Faculty of Medicine, University of Bern, Bern, Switzerland (Takala); Department of Intensive Care, Erasme University Hospital, Université Libre de Bruxelles, Brussels, Belgium (Vincent); Liver Intensive Care Unit, King's College Hospital, London, United Kingdom (Wendon); Department of Critical Care Medicine, University of Alberta, Edmonton, Alberta, Canada (Zampieri); Institute for Regeneration and Repair, University of Edinburgh, Edinburgh, Scotland (Shankar-Hari); Bloomsbury Institute of Intensive Care Medicine, Division of Medicine, University College London, London, United Kingdom (Singer).

Author Contributions: Dr Ranzani had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Moreno, Rhodes, Ranzani, Salluh, Juffermans, Laffey, Reinikainen, Serpa Neto, Einav, Gershengorn, Haniffa, Jung, Kirov, Metnitz, Myatra, Paiva, Pekkarinen, Pilcher, Piquilloud, Rezende, Robba, Sendagire, Siaw-Frimpong, Soares, Szczeklik, Vincent, Wendon, Zampieri, Shankar-Hari, Singer.

Acquisition, analysis, or interpretation of data: Moreno, Rhodes, Ranzani, Salluh, Berger-Estilita, Coopersmith, Juffermans, Laffey, Reinikainen, Serpa Neto, Tavares, Timsit, Arias López, Arulkumaran, Azoulay, Chaudhuri, de Lange, De Waele, dos Santos, Du, Ferrer, Fujii, Gershengorn, Hasan, Hollenberg, Ippolito, Jung, Lakbar, Lipman, Liu, Lobo, Martin, Metnitz, Myatra, Oczkowski, Paruk, Pilcher, Pölkki, Prescott, Reintam Blaser, Rezende, Rochweg, Schenck, Simpkin, Soares, Summers, Szczeklik, Takala, Shankar-Hari, Singer.

Drafting of the manuscript: Moreno, Rhodes, Ranzani, Salluh, Juffermans, Serpa Neto, de Lange, Du, Kirov, Myatra, Paiva, Pilcher, Robba, Sendagire, Szczeklik, Shankar-Hari, Singer.

Critical review of the manuscript for important intellectual content: Moreno, Rhodes, Salluh, Berger-Estilita, Coopersmith, Juffermans, Laffey, Reinikainen, Serpa Neto, Tavares, Timsit, Arias López, Arulkumaran, Azoulay, Chaudhuri, De Waele, dos Santos, Du, Einav, Ferrer, Fujii, Gershengorn, Haniffa, Hasan, Hollenberg, Ippolito, Jung, Kirov, Lakbar, Lipman, Liu, Lobo, Martin, Metnitz, Myatra, Oczkowski, Paiva, Paruk, Pekkarinen, Pilcher, Piquilloud, Pölkki, Prescott, Reintam Blaser, Rezende, Robba, Rochweg, Schenck, Sendagire, Siaw-Frimpong, Simpkin, Soares, Summers, Szczeklik, Takala, Vincent, Wendon, Zampieri, Shankar-Hari, Singer.

Statistical analysis: Moreno, Rhodes, Ranzani, Salluh, Juffermans, Chaudhuri, Haniffa, Liu, Metnitz, Pilcher, Simpkin, Shankar-Hari.

Obtained funding: Rhodes, Azoulay, Shankar-Hari.

Administrative, technical, or material support: Moreno, Rhodes, Salluh, Laffey, Serpa Neto, Timsit, Azoulay, de Lange, Du, Fujii, Haniffa, Jung, Metnitz, Myatra, Pilcher, Rochweg, Sendagire, Siaw-Frimpong, Soares, Vincent, Shankar-Hari, Singer.

Supervision: Moreno, Rhodes, Ranzani, Salluh, Berger-Estilita, Arias López, Du, Ferrer, Jung, Lipman, Metnitz, Myatra, Robba, Szczeklik, Vincent, Wendon, Zampieri, Shankar-Hari, Singer.

Conflict of Interest Disclosures: Dr Ranzani reported receiving a fellowship from the Spanish Ministry of Science and Innovation, Ramon y Cajal Program (RYC2023-002923-C) awarded by the Spanish Ministry of Science, Innovation and Universities (MICIU/AEI/10.13039/501100011033) and by the European Social Fund Plus (ESF+); receiving grants from the National Council for Scientific and Technological Development (CNPq), Brazil (2023-2025), to Pontifical Catholic University; receiving grants from the Brazilian Ministry of Health, PROADI Program (2022-2026) to Hospital Albert Einstein; receiving a fellowship from the Instituto de Salud Carlos III, Spanish government (2020-2023); and receiving grants from Health Effects Institute (2021-2023) to ISGlobal outside the submitted work. Dr Berger-Estilita reported receiving personal fees paid to the institution from Medtronic outside the submitted work. Dr Coopersmith reported receiving grants from National Institutes of Health outside the submitted work. Dr Azoulay reported receiving personal fees for lectures from Alexion, Baxter, Pfizer, and Gilead outside the submitted work. Dr De Waele reported receiving honoraria paid to the employer from Biomerieux, Grifols, Menarini, Pfizer, Roche Diagnostics, and Viatris outside the submitted work as well as receiving funding from the Research Foundation Flanders. Dr dos Santos reported receiving grants from the Canadian Institutes of Health Research during the conduct of the study; having a patent pending; and being cofounder and chief scientific officer for NorthMiRs. Dr Ferrer reported receiving personal fees from Cytosorbents, Viatris, AOP, and Pfizer outside the submitted work. Dr Fujii reported receiving grants from the Japan Society for the Promotion of Science and Terumo Life Science Foundation as well as receiving personal fees from Japan Council for Quality Health Care and Wolters Kluwers Health outside the submitted work. Dr Gershengorn reported receiving salary support from the University of Miami Hospital and Clinics Data Analytics Research Team during the conduct of the study as well as serving as editor-in-chief for *CHEST Critical Care*. Dr Haniffa reported receiving grants to the institution from Wellcome during the conduct of the study as well as receiving travel and accommodation from the World Health Organization Guideline development group for influenza outside the submitted work. Dr Lakbar reported receiving personal fees from AOP Health and Viatris outside the submitted work. Dr Liu reported receiving grants from the National Institutes of Health (NIH) during the conduct of the study. Dr Lobo reported receiving personal fees from Edwards and Baxter outside the submitted work. Dr Martin reported receiving personal fees for serving on the research advisory board of Grifols, receiving personal for serving on the data safety monitoring board from Eagle Pharmaceuticals, receiving personal fees for serving as a section editor UpToDate, and receiving personal fees from Elsevier as editor for Critical Care Clinics outside the submitted work. Dr Oczkowski reported receiving travel support from Fisher & Paykel and receiving consulting fees from Vitalaire outside the submitted work. Dr Paruk reported receiving advisory fees from Baxter and Novo Nordisk and speaker fees from Thermofisher Scientific and Pfizer outside the submitted work. Dr Piquilloud reported receiving lecture fees from Getinge, GE Healthcare, Fisher & Paykel, and Air Liquid System outside the submitted work. Dr Prescott reported receiving grants from NIH, Veterans Affairs Health Systems Research, and the Agency for Healthcare Research and Quality and receiving consulting fees from Aurobac Therapeutics outside the submitted work as well as serving as cochair of Surviving Sepsis Campaign Guidelines and being a government employee. Dr Reintam Blaser reported receiving grants from Estonian Research Council outside the submitted work. Dr Robba reported receiving personal fees from BD outside the submitted work. Dr Soares reported being founder of and holding equity in Epimed Solutions and receiving grants from CNPq and FAPERJ outside the submitted work. Dr Summers reported being an employee of the University of Cambridge and chair of the UKRI-Medical Research Council Experimental Medicine Panel; receiving research funding to the University of Cambridge from the National Institute for Health and Care Research, UKRI-Medical Research Council, and the Wellcome Trust. Dr Vincent reported being codeveloper of the original SOFA score; he has no financial interest related to its use or dissemination. Dr Singer reported receiving grants from Gentian; receiving speaker fees to the institution from Biotest, AOP, and Pfizer; receiving speaker fees from Biomerieux; receiving advisory board fees to the institution from Safeguard, Matisse, Volition, and Radiometer; receiving advisory board fees from Aptarion, Roche Diagnostics, Sanofi, deePull, and Bayer; receiving royalties to the institution from Deltex outside the submitted work; and serving as a council member for the International Sepsis Forum and a sepsis topic advisor for National Institute for Health and Care Excellence. No other disclosures were reported.

Additional Contributions: We acknowledge the major contribution made to this project by the analysts and data collectors in participating hospitals for the following databases: ANZICS Centre for Outcome and Resources Evaluation; Austrian Center for Documentation and Quality Assurance in Intensive Care; Organizational Characteristics in Critical Care (Brazil); Kaiser Permanente North California; Gruppo Italiano per la Valutazione degli Interventi in Terapia Intensiva; Japanese Intensive Care Patient Database; Nepal Intensive Care Research Foundation; OutcomeRea; and eICU (eAppendix in the [Supplement](#)).

REFERENCES

1. 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. doi:10.1007/BF01709751

2. Vincent JL, de Mendonça A, Cantraine F, 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. *Crit Care Med*. 1998;26(11):1793-1800. doi:[10.1097/00003246-199811000-00016](https://doi.org/10.1097/00003246-199811000-00016)
3. Moreno R, Vincent JL, Matos R, 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):686-696. doi:[10.1007/s001340050931](https://doi.org/10.1007/s001340050931)
4. Antonelli M, Moreno R, Vincent JL, et al. Application of SOFA score to trauma patients—Sequential Organ Failure Assessment. *Intensive Care Med*. 1999;25(4):389-394. doi:[10.1007/s001340050863](https://doi.org/10.1007/s001340050863)
5. 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):1754-1758. doi:[10.1001/jama.286.14.1754](https://doi.org/10.1001/jama.286.14.1754)
6. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA*. 2016;315(8):801-810. doi:[10.1001/jama.2016.0287](https://doi.org/10.1001/jama.2016.0287)
7. Knaus WA, Wagner DP, Draper EA, et al. The APACHE III prognostic system: risk prediction of hospital mortality for critically ill hospitalized adults. *Chest*. 1991;100(6):1619-1636. doi:[10.1378/chest.100.6.1619](https://doi.org/10.1378/chest.100.6.1619)
8. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA*. 1993;270(24):2957-2963. doi:[10.1001/jama.1993.03510240069035](https://doi.org/10.1001/jama.1993.03510240069035)
9. Pölkki A, Pekkarinen PT, Takala J, Selander T, Reinikainen M. Association of Sequential Organ Failure Assessment (SOFA) components with mortality. *Acta Anaesthesiol Scand*. 2022;66(6):731-741. doi:[10.1111/aas.14067](https://doi.org/10.1111/aas.14067)
10. Bachmann KF, Arabi YM, Regli A, Starkopf J, Reintam Blaser A. Cardiovascular SOFA score may not reflect current practice. *Intensive Care Med*. 2022;48(1):119-120. doi:[10.1007/s00134-021-06536-6](https://doi.org/10.1007/s00134-021-06536-6)
11. Kaukonen KM, Bailey M, Suzuki S, Pilcher D, Bellomo R. Mortality related to severe sepsis and septic shock among critically ill patients in Australia and New Zealand, 2000-2012. *JAMA*. 2014;311(13):1308-1316. doi:[10.1001/jama.2014.2637](https://doi.org/10.1001/jama.2014.2637)
12. Prescott HC, Harrison DA, Rowan KM, Shankar-Hari M, Wunsch H. Temporal trends in mortality of critically ill patients with sepsis in the United Kingdom, 1988-2019. *Am J Respir Crit Care Med*. 2024;209(5):507-516. doi:[10.1164/rccm.202309-1636OC](https://doi.org/10.1164/rccm.202309-1636OC)
13. Ranzani OT, Singer M, Salluh JIF, et al. Development and validation of the Sequential Organ Failure Assessment (SOFA)-2 score. *JAMA*. Published online October 29, 2025. doi:[10.1001/jama.2025.20516](https://doi.org/10.1001/jama.2025.20516)
14. Murray SA, Kendall M, Boyd K, Sheikh A. Illness trajectories and palliative care. *BMJ*. 2005;330(7498):1007-1011. doi:[10.1136/bmj.330.7498.1007](https://doi.org/10.1136/bmj.330.7498.1007)
15. Devlin JW, Skrobik Y, Gélinas C, et al. Clinical practice guidelines for the prevention and management of pain, agitation/sedation, delirium, immobility, and sleep disruption in adult patients in the ICU. *Crit Care Med*. 2018;46(9):e825-e873. doi:[10.1097/CCM.0000000000003299](https://doi.org/10.1097/CCM.0000000000003299)
16. Salluh JIF, Wang H, Schneider EB, et al. Outcome of delirium in critically ill patients: systematic review and meta-analysis. *BMJ*. 2015;350:h2538-h2538. doi:[10.1136/bmj.h2538](https://doi.org/10.1136/bmj.h2538)
17. Wierszowski PM, Leone M, Kaas-Hansen BS, et al. Position paper on the reporting of norepinephrine formulations in critical care from the Society of Critical Care Medicine and European Society of Intensive Care Medicine Joint Task Force. *Crit Care Med*. 2024;52(4):521-530. doi:[10.1097/CCM.0000000000006176](https://doi.org/10.1097/CCM.0000000000006176)
18. Goradia S, Sardaneh AA, Narayan SW, Penm J, Patanwala AE. Vasopressor dose equivalence: a scoping review and suggested formula. *J Crit Care*. 2021;61:233-240. doi:[10.1016/j.jcrc.2020.11.002](https://doi.org/10.1016/j.jcrc.2020.11.002)
19. Kotani Y, Di Gioia A, Landoni G, Belletti A, Khanna AK. An updated "norepinephrine equivalent" score in intensive care as a marker of shock severity. *Crit Care*. 2023;27(1):29. doi:[10.1186/s13054-023-04322-y](https://doi.org/10.1186/s13054-023-04322-y)
20. De Backer D, Biston P, Devriendt J, et al; SOAP II Investigators. Comparison of dopamine and norepinephrine in the treatment of shock. *N Engl J Med*. 2010;362(9):779-789. doi:[10.1056/NEJMoa0907118](https://doi.org/10.1056/NEJMoa0907118)
21. Khanna AK, Kinoshita T, Natarajan A, et al. Association of systolic, diastolic, mean, and pulse pressure with morbidity and mortality in septic ICU patients: a nationwide observational study. *Ann Intensive Care*. 2023;13(1):9. doi:[10.1186/s13613-023-01101-4](https://doi.org/10.1186/s13613-023-01101-4)
22. Ranieri VM, Rubenfeld G, Slutsky AS. Rethinking acute respiratory distress syndrome after COVID-19: if a "better" definition is the answer, what is the question? *Am J Respir Crit Care Med*. 2023;207(3):255-260. doi:[10.1164/rccm.202206-1048CP](https://doi.org/10.1164/rccm.202206-1048CP)
23. Rhodes A, Moreno RP, Azoulay E, et al; Task Force on Safety and Quality of European Society of Intensive Care Medicine (ESICM). Prospectively defined indicators to improve the safety and quality of care for critically ill patients: a report from the Task Force on Safety and Quality of the European Society of Intensive Care Medicine (ESICM). *Intensive Care Med*. 2012;38(4):598-605. doi:[10.1007/s00134-011-2462-3](https://doi.org/10.1007/s00134-011-2462-3)

24. Secombe P, Millar J, Litton E, et al. Thirty years of ANZICS CORE: a clinical quality success story. *Crit Care Resusc.* 2023;25(1):43-46. doi:10.1016/j.ccrj.2023.04.009
25. Metnitz PGH, Vesely H, Valentin A, et al. Evaluation of an interdisciplinary data set for national intensive care unit assessment. *Crit Care Med.* 1999;27(8):1486-1491. doi:10.1097/00003246-199908000-00014
26. Soares M, Salluh JIF, Zampieri FG, Bozza FA, Kurtz PMP. A decade of the ORCHESTRA study: organizational characteristics, patient outcomes, performance and efficiency in critical care. *Crit Care Sci.* Published online July 19, 2024. doi:10.62675/2965-2774.20240118-en
27. Liu V, Turk BJ, Ragins AI, Kipnis P, Escobar GJ. An electronic Simplified Acute Physiology Score-based risk adjustment score for critical illness in an integrated healthcare system. *Crit Care Med.* 2013;41(1):41-48. doi:10.1097/CCM.0b013e318267636e
28. Finazzi S, Paci G, Antiga L, et al; GiViTI-PROSAFE collaboration. PROSAFE: a European endeavor to improve quality of critical care medicine in seven countries. *Minerva Anesthesiol.* 2020;86(12):1305-1320. doi:10.23736/S0375-9393.20.14112-9
29. Irie H, Okamoto H, Uchino S, et al; JIPAD Working Group in the Japanese Society of Intensive Care Medicine. The Japanese Intensive care PATient Database (JIPAD): a national intensive care unit registry in Japan. *J Crit Care.* 2020;55:86-94. doi:10.1016/j.jcrr.2019.09.004
30. Aryal D, Thakur A, Gauli B, et al. Epidemiology of critically ill patients in intensive care units in Nepal: a retrospective observational study. *Wellcome Open Res.* 2023;8:180. doi:10.12688/wellcomeopenres.19127.1
31. Planquette B, Timsit JF, Misset BY, et al; OUTCOMEREA Study Group. Pseudomonas aeruginosa ventilator-associated pneumonia: predictive factors of treatment failure. *Am J Respir Crit Care Med.* 2013;188(1):69-76. doi:10.1164/rccm.201210-1897OC
32. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci Data.* 2018;5(1):180178. doi:10.1038/sdata.2018.178
33. Finazzi S, Mandelli G, Garbero E, et al. Data collection and research with MargheritaTre. *Physiol Meas.* 2018;39(8):084004. doi:10.1088/1361-6579/aaad10f
34. Balamuth F, Scott HF, Weiss SL, et al; Pediatric Emergency Care Applied Research Network (PECARN) PED Screen and PECARN Registry Study Groups. Validation of the Pediatric Sequential Organ Failure Assessment Score and evaluation of Third International Consensus Definitions for Sepsis and Septic Shock definitions in the pediatric emergency department. *JAMA Pediatr.* 2022;176(7):672-678. doi:10.1001/jamapediatrics.2022.1301
35. Schlapbach LJ, Watson RS, Sorce LR, et al; Society of Critical Care Medicine Pediatric Sepsis Definition Task Force. International consensus criteria for pediatric sepsis and septic shock. *JAMA.* 2024;331(8):665-674. doi:10.1001/jama.2024.0179
36. 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. doi:10.1186/s13054-017-1609-1
37. Lam BD, Struja T, Li Y, et al. Analyzing how the components of the SOFA score change over time in their contribution to mortality. *Crit Care Sci.* Published online November 22, 2024. doi:10.62675/2965-2774.20240030-en
38. 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. doi:10.1186/s13054-019-2663-7
39. Jezmir JL, Bharadwaj M, Chaitoff A, et al; STOP-COVID Investigators. Performance of crisis standards of care guidelines in a cohort of critically ill COVID-19 patients in the United States. *Cell Rep Med.* 2021;2(9):100376. doi:10.1016/j.xcrm.2021.100376
40. Matsumura Y, Nakada TA, Abe R, Oshima T, Oda S. Serum procalcitonin level and SOFA score at discharge from the intensive care unit predict post-intensive care unit mortality: a prospective study. *PLoS One.* 2014;9(12):e114007. doi:10.1371/journal.pone.0114007
41. Moreno R, Miranda DR, Matos R, Fevereiro T. Mortality after discharge from intensive care: the impact of organ system failure and nursing workload use at discharge. *Intensive Care Med.* 2001;27(6):999-1004. doi:10.1007/s001340100966

SUPPLEMENT.

eTable 1. Panel Composition of 60 Experts Involved in SOFA-2 (Voting Members)

eTable 2. Open-Ended Questions From First mDelphi Round

eTable 3. Issues Highlighted During Delphi Processes or at the Development Stages of the Analysis Plans

eTable 4. Steps Suggested to Each Group to Follow for Stage 3 Group Proposals

eTable 5. Protocol Template for Systematic Reviews in SOFA-2

eTable 6. Illustration of Questions Used in the Second mDelphi Round

eTable 7. Illustration of Questions Used in the Third mDelphi Round

eTable 8. First Round mDelphi Synthesis

eTable 9. Variables Initially Highlighted From the Group Discussions for SOFA-2

eTable 10. Variables Prioritized for Voting in the Second mDelphi Round

eFigure 1. Second mDelphi, First Iteration: Agreement on Votes on Proposed Definition and Proxy/Feasibility of Proposed Variables

eFigure 2. Second mDelphi, Second Iteration: Agreement on Votes on Proposed SOFA-2

eFigure 3. Third mDelphi, First Iteration: Agreement on Votes on Proposed SOFA-2 Domains

eAppendix. Acknowledgment of Participating Databases

eReferences.