JAMA | Original Investigation | CARING FOR THE CRITICALLY ILL PATIENT

Development and Validation of the Sequential Organ Failure Assessment (SOFA)-2 Score

Otavio T. Ranzani, MD, MSc, PhD; Mervyn Singer, MD; Jorge I. F. Salluh, MD, PhD; Manu Shankar-Hari, MD, PhD; David Pilcher, MBBS; Joana Berger-Estilita, MD, PhD; Craig M. Coopersmith, MD; Nicole P. Juffermans, MD, PhD; John Laffey, MD, DSc; 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, PhD; Diptesh Aryal, MD, PhD; Elie Azoulay, MD, PhD; Leo Anthony Celi, MD, MPH, MSc; 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; Teresa Engelbrecht, BSc; Fathima Fazla, MSc; Ricard Ferrer, MD, PhD; Stefano Finazzi, PhD; Tomoko Fujii, MD, PhD; Hayley B. Gershengorn, MD; John D. Greene, MA; Rashan Haniffa, MD, PhD; Sicheng Hao, MSc; Mohd Shahnaz Hasan, MBBS, MAnaes; Steve Hollenberg, MD; Mariachiara Ippolito, MD; Christian Jung, MD, PhD; Mikhail Kirov, MD, PhD; Shigetaka Kobari, BS; Inès Lakbar, MD, PhD; Jeffrey Lipman, MBBCH, MD, DMed(Res); Vincent Liu, MD; Xiaoli Liu, PhD; Suzana M. Lobo, MD, PhD; Demetrio Magatti, MSc; Greg S. Martin, MD, MSc; Barbara Metnitz, PhD; Philipp Metnitz, MD, PhD; Sheila N. Myatra, MD; Simon Oczkowski, MD, MSc, MHSc; José-Artur Paiva, MD, PhD; Fathima Paruk, MD, PhD; Pirkka T. Pekkarinen, MD, PhD; Lise Piquilloud, MD, PhD; Anssi Pölkki, MD, PhD; Hallie C. Prescott, MD, MSc; Annika Reintam Blaser, MD, PhD; Ederlon Rezende, MD; Chiara Robba, MD, PhD; Bram Rochwerg, MD, MSc; Stephane Ruckly, MSc; Rasoul Samei, BSci; Edward J. Schenck, MD, MS; Paul Secombe, BMBS, MSc; 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; Shiro Tanaka, PhD; Giovanni Tricella, PhD; Jean-Louis Vincent, MD, PhD; Julia Wendon, MBChB; Fernando G. Zampieri, MD, PhD; Andrew Rhodes, MB, BS, MD(Res); Rui Moreno, MD, PhD

IMPORTANCE Acute dysfunction of vital organs is the hallmark of critical illness. The Sequential Organ Failure Assessment (SOFA) score, the most widely adopted approach to describe organ dysfunction, has not been updated in 30 years and therefore may not appropriately capture current clinical practice and outcomes.

OBJECTIVES To inform the data-driven component of an updated score (SOFA-2) in varied geographical and resource settings (stages 6-8) after expert input via a modified Delphi process (stages 1-5).

DESIGN, SETTING, AND PARTICIPANTS A federated analysis was performed on data collected from adult patients admitted to 1319 intensive care units (ICUs) in 9 countries (Australia, Austria, Brazil, France, Italy, Japan, Nepal, New Zealand, United States) between 2014 and 2023. Four representative multicenter cohorts containing data from 2 098 356 patients were used for data-driven score development and internal validation. External validation was performed on 6 cohorts containing data from 1 241 114 patients.

MAIN OUTCOMES AND MEASURES Content validity for organ dysfunction identified through the modified Delphi process should be reflected by predictive validity using the area under the receiver operating characteristic (AUROC) curve of the score measured on the first ICU day (higher scores indicate worse organ dysfunction).

RESULTS Of 3.34 million patient encounters, 270 108 (8.1%) died in the ICU (range, 4.5% to 20.5% across the 10 cohorts). SOFA-2 modified the 6 organ systems of the original SOFA score (brain, respiratory, cardiovascular, liver, kidney, hemostasis), including new variables and revised thresholds that better describe the organ dysfunction distribution from 0 to 4 points and their associated mortality (SOFA-2 AUROC, 0.79; 95% CI, 0.76-0.81; SOFA-1 AUROC, 0.77; 95% CI, 0.74-0.81). Evaluation of sequential SOFA-2 data from ICU day 1 to day 7 maintained its predictive validity. Insufficient data and lack of content validity precluded incorporation of gastrointestinal and immune dysfunction scores into SOFA-2.

CONCLUSIONS AND RELEVANCE The SOFA-2 score, updated to include contemporary organ support treatments and new score thresholds, describes organ dysfunction in a large, geographically and socioeconomically diverse population of critically ill adults.

Editorial page 2075

Multimedia

Supplemental content

Related article at jamanetworkopen.com

Author Affiliations: Author affiliations are listed at the end of this article.

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

JAMA. 2025;334(23):2090-2103. doi:10.1001/jama.2025.20516
Published online October 29, 2025.

The provision of intensive care has undergone many changes in the past 30 years. New methods for advanced life support, avoidance of iatrogenic harm, and closer monitoring have impacted the understanding and treatment of organ dysfunction in critically ill patients.[1] However, the measurement of organ dysfunction with the Sequential Organ Failure Assessment (SOFA) score (now SOFA-1) has not changed since 1996.[2,3]

The SOFA-1 score describes 6 organ systems—neurological, cardiovascular, respiratory, hepatic, renal, and coagulation—using clinical and biochemical variables in routine clinical use during the 1990s, with a total score ranging from 0 to 24 (higher scores indicate worse organ dysfunction). It fails to capture contemporary interventions (drugs and devices) that provide support for failing organs.[4-6] Thresholds for organ support may also be impacted by trends toward less invasive treatment, initiation earlier in the disease trajectory, and ICU case mix.[4,5,7-9] On this backdrop, a new update to organ dysfunction measurement in critically ill patients is needed, particularly one that is generalizable to both high-income countries and low- and middle-income countries.[10]

An accompanying manuscript describes a modified Delphi (mDelphi) procedure to generate the conceptual framework and proposed score.[11] This article describes the data-driven development and validation of the final SOFA-2 score.

Key Points

Question Does an updated Sequential Organ Failure Assessment (SOFA)-2 score describe organ dysfunction in critically ill patients and its association with intensive care unit (ICU) mortality?

Findings The SOFA-2 score was developed and validated in 10 international multicenter cohorts of 3.3 million adult ICU patients. SOFA-2 includes the original 6 organ systems with a total score ranging from 0 to 24 (higher scores indicate worse organ dysfunction). Possible inclusion of immune and gastrointestinal systems was investigated but not added. The updated score now incorporates commonly used drugs and mechanical organ supports that were rarely or not used when the original version was published in 1996. Some thresholds were modified to improve predictive validity against ICU mortality.

Meaning The SOFA-2 score, updated to include contemporary organ support treatments and new score thresholds, describes organ dysfunction, supported by good predictive validity, in a large, geographically and socioeconomically diverse population of critically ill adults.

Methods

Overview of SOFA-2 Process

The update to the SOFA score occurred in 8 stages (**Figure 1**). These stages included expert selection for mDelphi rounds, systematic reviews, and internal and external data validation. The work included the assessment of 6 domains (reliability; content; construct; criterion; predictive validity; and clarity, measurement burden, and timeliness) aligned to the appropriate stages. The first 5 stages were completed and described in the accompanying manuscript.[11] The results of the data-driven stages (6-8) are reported herein. Findings are presented in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines.

Summary of Findings of Stages 1 Through 5

A panel of 60 intensive care experts participated in 2 mDelphi exercises to develop updated definitions of organ dysfunction and to propose routinely measured or clinical and laboratory variables that reflect contemporary identification and management of organ dysfunction (stages 1-3). The diverse expert panel addressed evidence gaps and ensured clinical relevance. These discussions were based on the theoretical framework developed for Sepsis-3.[12,13] Organ dysfunction was considered a scoreable construct defined by operational criteria rather than a precise biological truth. Following the principles underpinning SOFA,[2,4] the update prioritized simplicity, clinical usability, widespread applicability, and content validity. This ensured that the score reflects meaningful categories of organ dysfunction severity with a stepwise increase in mortality risk (Supplement 1).[12] In stage 4, the systematic reviews were used to match the ratio of arterial oxygen tension (Pa_{O_2}) to fraction of inspired oxygen (F_{IO_2}) and arterial oxygen saturation (Sp_{O_2}) to F_{IO_2} thresholds for when arterial blood gas measurements are unavailable, to specify indication criteria for commencing renal replacement therapy and to characterize the associations of norepinephrine drenaline dosage, total white cell count, lymphocyte count, and intra-abdominal pressure with mortality risk. Stage 5 comprised a second mDelphi round to secure agreement on proxy and feasibility for a draft SOFA-2 proposal. Eight organ systems (brain, respiratory, cardiovascular, liver, kidney, hemostasis, gastrointestinal, immune) were proposed for evaluation in subsequent internal and external validation stages.

Study Design, Setting, and Population for Stages 6 to 8

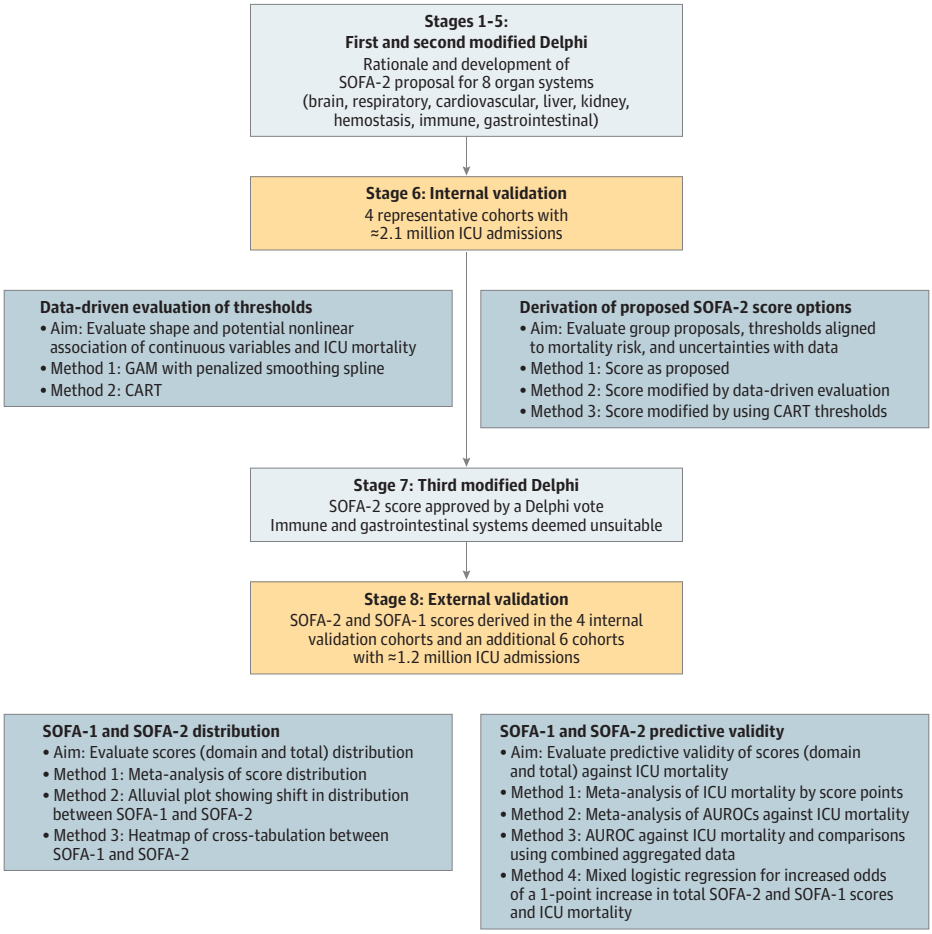
A federated analysis was performed using data from 10 multicenter, international cohorts, comprising encounters from 1319 intensive care units (ICUs) in 9 countries (Australia, Austria, Brazil, France, Italy, Japan, Nepal, New Zealand, United States).

Data were extracted from 8 national ICU registries[10,14] and 2 multicenter electronic health record–based datasets that fulfilled predefined criteria for data completeness, feasibility, and protocol adherence (eMethods in Supplement 2). Single-center datasets were excluded.

We included critically ill adults aged 18 years or older admitted to ICUs between January 1, 2014, and December 31, 2023. We excluded ICU readmissions within the same acute hospital stay, patients with missing ICU discharge status from their index ICU admission, and admissions exclusively for organ donation.

In the internal validation phase, 4 cohorts (Australian and New Zealand Intensive Care Society [ANZICS],[15] Austrian Center for Documentation and Quality Assurance in Intensive Care [ASDI],[16] Kaiser Permanente Northern California [KPNC],[17] and Organizational Characteristics in Critical Care [ORCHESTRA][18]) were analyzed. External validation used an additional 6

Figure 1. SOFA-2 Update Stages and Methods



Stages 1-8 were aligned with Sequential Organ Failure Assessment (SOFA) principles and with 6 score utility domains: stages 1-5 and 7, reliability; content; construct; criterion validity: and clarity measurement, burden, and timeliness. Stages 6 and 8: construct and predictive validity.

Further details on rationale and Delphi methodology are provided in a companion article.[11]

AUROC indicates area under the receiver operating characteristic curve; CART, classification and regression tree; GAM, generalized additive model; ICU, intensive care unit.

cohorts (eICU [electronic Intensive Care Unit Collaborative Research Database],[19] GiViTI-PROSAFE [Gruppo Italiano per la Valutazione degli Interventi in Terapia Intensiva-Product Safety Forum of Europe],[20] JIPAD [Japanese Intensive Care Patient Database],[21] GiViTI-MargheritaTre,[22] Nepal Intensive Care Registry Foundation [NICRF],[23] and OutcomeRea [OutcomeRéanimation][24]). Databases were required to have key variables for each organ system (ie, Glasgow Coma Scale [GCS]; Pao_2:FIO_2 ratio; mean arterial pressure and vasopressor dosage; and bilirubin, creatinine, and platelet values, eTable 1 in Supplement 2). Values outside the plausible range for continuous variables, defined by group consensus (eTable 2 in Supplement 2), were considered missing.

Methods for Stages 6 to 8

First, the internal validation of the proposed SOFA-2 domains was conducted, evaluating the distributions of candidate variables, threshold values for SOFA cut points, and predictive validity for ICU mortality. ICU mortality was chosen by consensus as the primary outcome because it was consistently available across all registries.[14,25] Continuous distributions were evaluated with generalized additive models (GAMs) using the mgcv R-package with penalized smoothing splines.[26] Threshold values proposed for continuous variables (ie, GCS; Pao_2:FIO_2 ratio; and bilirubin, creatinine, and platelet levels) were compared against those generated using a classification and regression tree (CART) model.[27] Four cutoffs for each continuous variable were obtained in the CART models when possible, and 10-fold cross validation was used (eMethods in Supplement 2).[27]

Next, the third mDelphi reviewed results in stage 7. When data-driven results conflicted with expert consensus, decisions were made through structured committee discussions, guided by the SOFA protocol rules (eg, content validity taking precedence over predictive validity; Supplement 1). Finally, external validation of the final SOFA-2 score was performed (stage 8). These analyses focused on predictive validity for ICU mortality, longitudinal measurement in the first 7 days of intensive care, and sensitivity analyses (see Supplement 1 for more detail).

Statistical Analyses

Descriptive statistics were generated using mean (SDs), median (IQRs), and proportions. The proportion of ICU deaths were illustrated for each SOFA-1 and SOFA-2 scores, as well as for each SOFA domain. Data were pooled using a multilevel

meta-analysis model, applying a logit transformation, a random intercept for each database, and categorical points as covariates.[28,29] Pooled proportions and 95% CIs estimating marginal means were derived from the meta-analytic model for each point category. Cohorts were entered into the models for each organ system when at least 2 organ systems with 2 consecutive point categories were available for that cohort (Supplement 1). For total SOFA, only cohorts with fully available scores were included.

To assess predictive validity, the area under the receiver operating characteristic (AUROC) curves were estimated (1) in a single stage combining the number of patients with a given score and associated deaths in mixed-effects logistic models, with a random intercept per cohort, and (2) in 2 stages, pooling estimates extracted from each cohort using random-effects meta-analysis models, with logit transformation for the AUROC and its standard error.[30] Standard errors for the area under the curve in each cohort were estimated using the DeLong method.[31] Restricted maximum likelihood was used for meta-analyses. Generalized mixed models with a logit link evaluated the association between a 1-point increase in total SOFA score and ICU mortality, including random intercepts for each database.

Multiple sensitivity analyses were conducted. First, we tested whether findings using normal value imputation were mirrored by analyses limited to complete-case data only. Second, we tested assumptions about the time window to profile organ dysfunction on day 1, performing an analysis comparing cohorts with the worst values recorded within the first 24 hours of ICU admission vs cohorts with both first-hour and first 24-hour worst values. Third, we tested whether partial or full data availability for each system impacted the pooled results; meta-analyses were then performed leveraging data from the 10 cohorts. Fourth, we assessed whether the magnitude of association between a 1-point increase in each organ dysfunction score and ICU mortality could vary between cohorts; this included a random slope for score points in the meta-analyses and mixed models.

To understand how the SOFA-2 score reclassified patients into different score categories compared with SOFA-1, a reclassification heat map was used with patient-level data from eICU, an open access database.[19]

In the longitudinal analysis, the daily, domain-specific and total SOFA-2 scores were measured on ICU days 1 to 7, reporting both mean and maximum values, when available.

Missing data were primarily handled using normal value imputation[3,4,32-34] and alternative approaches[35-37] described in Supplement 1 and the eMethods section and eTable 3 in Supplement 2.

All analyses were performed using R version 4.2.1 (R Foundation).

Results

Patients

In internal validation data (4 cohorts, total n = 2 098 356 patients; mean age, 63.1 years [SD, 18], 44.6% female) most ICU admissions were for medical diagnoses (n = 1 129 428, 53.9%), and ICU mortality ranged from 4.5% to 10.1% (**Figure 2** and **Table 1**). External validation data (6 cohorts, n = 1 241 114 patients) were similar, noting that patients were older (mean age, 65.1 years [SD, 16], 40.7% female) and ICU mortality ranged from 4.0% to 20.5%. Distributions varied between databases (n = 7) with respect to illness severity within the first day of ICU admission (range, 12.4%-31.2% of predicted in-hospital mortality using varied scoring systems; eTable 4 in Supplement 2).

Stage 6, Internal Validation

The distributions for candidate variables for the SOFA-2 score were consistent across cohorts (eFigures 1-6 in Supplement 2). When evaluated using GAM and CART models, candidate variable thresholds ratified those proposed from the second mDelphi process (eTable 5 in Supplement 2). For example, for the respiratory system, the new PaO_2:FIO_2 ratio thresholds are 300, 225, 150, and 75. For the liver system, thresholds based on bilirubin levels were adjusted to 1.2, 3, 6, and 12 mg/dL (to convert bilirubin from mg/dL, multiply by 17.104). Changes in cutoffs in other organ systems are shown in eTable 6 in Supplement 2.

The proposed gastrointestinal system score was evaluated in one cohort (ASDI, n = 406 469), and no association was observed with ICU mortality (eFigure 7 in Supplement 2). For the immune system, there was a U-shaped association between both total white blood cell and lymphocyte count with ICU mortality (eFigure 8 in Supplement 2).

Stage 7, Third mDelphi

In review of the internal validation, there was consensus that the gastrointestinal score lacked predictive validity whereas the immune score did not fulfil content validity. Consequently, both were excluded from the final SOFA-2 score.[11]

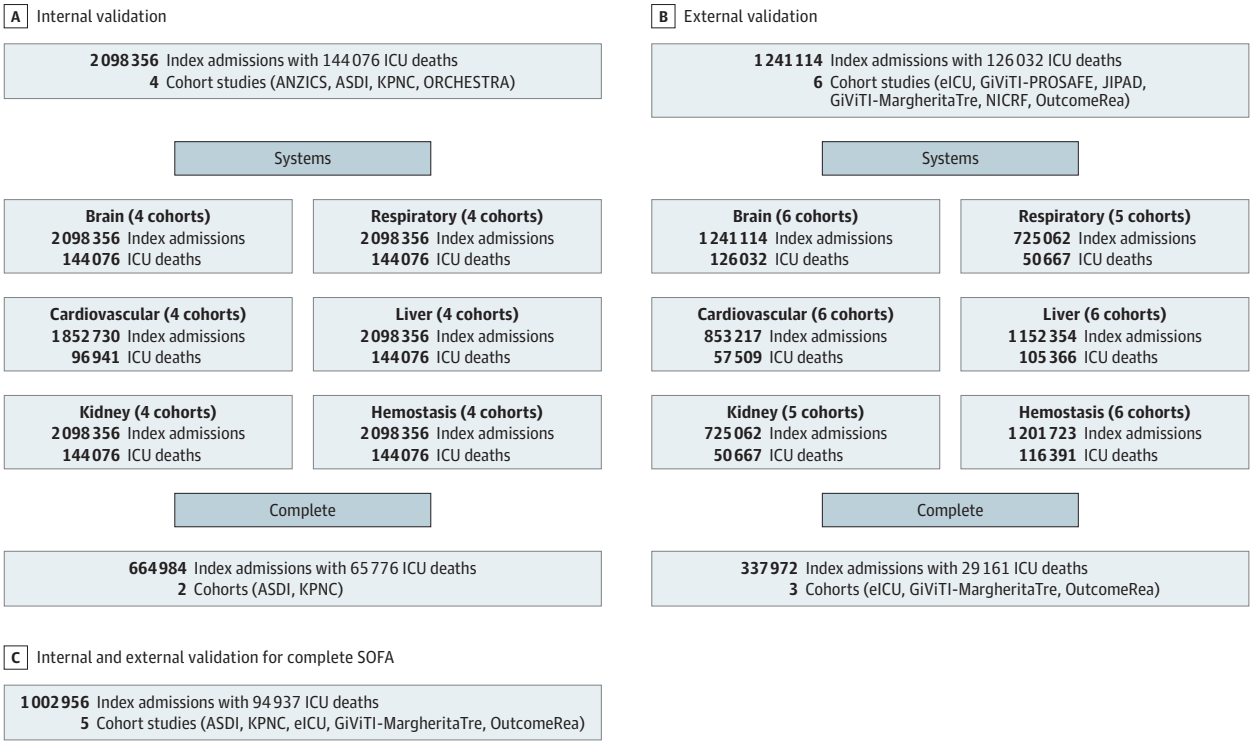
Stage 8, External Validation

External validation in 6 separate cohorts confirmed findings from the internal validation, including complete-case and other sensitivity analyses (**Figure 3**; eFigures 9-14 in Supplement 2). Each individual organ system was associated with an incremental increase in ICU mortality as the SOFA-2 score increased from 0 to 4 points (eFigures 15-22 in Supplement 2).

Pooled results from both internal and external validation from either 9 (respiratory, kidney) or 10 (brain, cardiovascular, liver, hemostasis) cohorts totaled a minimum of 2.5 million encounters in each system (Figure 2; eTable 7 in Supplement 2). Total SOFA analyses pooled data from 5 cohorts (2 from internal and 3 from external validation) totaling 1 002 956 patients and 94 937 deaths (9.5%). The same observed pattern for score distribution and associated mortality was observed in meta-analyses of all available data (eFigures 23-26 in Supplement 2).

Complete-case data for total SOFA included 116 481 patients and 22 476 deaths (19.3%). The distributions shifted in the complete case analyses toward higher scores (eFigures 27-30 in Supplement 2), although a similar pattern for ICU mortality risk was retained. Assuming a linear association between total SOFA score and ICU mortality, there was an increase in the odds of ICU mortality (odds ratio, 1.378; 95% CI, 1.375-1.381) for each 1-point increase in the SOFA-2 score.

Figure 2. Study Flowchart for Internal and External Validations



A and B, The total number of index admissions, intensive care unit (ICU) deaths, and number of cohorts for the internal (stage 6) and external validation (stage 8), stratified by each organ system stand alone and for complete Sequential Organ Failure Assessment (SOFA; ie, when the 6 systems were available to estimate the total score).

C, The total number of index admissions, ICU deaths, and number of cohorts for complete SOFA, combining the internal (stage 6) and external validation (stage 8; data are shown in Figures 4B and 5B).

The number for each domain can be smaller than the total because of missing data in the cohorts, including an inability to calculate the full score (eg, norepinephrine dosage not known). ANZICS indicates Australian and New Zealand Intensive Care Society; ASDI, Austrian Center for Documentation and Quality Assurance in Intensive Care; eICU, electronic Intensive Care Unit Collaborative Research Database; GiViTI, Gruppo Italiano per la Valutazione degli Interventi in Terapia Intensiva; KPNC, Kaiser Permanente Northern California; MargheritaTre, Petal Three; ORCHESTRA, Organizational Characteristics in Critical Care; OutcomeRea, OutcomeRéanimation.

Similar results were found in sensitivity analyses (eTable 8 in Supplement 2).

The final SOFA-2 score is shown in **Table 2** with the footnotes describing important rules for consistent scoring.

Comparison Between SOFA-1 and SOFA-2

Overall, the total SOFA-1 score (median, 3; IQR, 1-6]) was similar to total SOFA-2 (median, 3; IQR, 1-5, eTable 9 in Supplement 2) with normal value imputation. In complete-case data, total scores were higher (median SOFA-1 score, 8; IQR, 5-11; median SOFA-2 score, 7; IQR, 4-10). The distribution of patients within each organ system differed between SOFA-1 and SOFA-2 (**Figure 4**). For example, 2 points in the cardiovascular system for SOFA-1 only contained 0.9% of patients compared with 8.9% in SOFA-2 (Figure 4A and eFigures 31-34 in Supplement 2). For the total SOFA score, more patients had lower values in SOFA-2 (Figure 4B).

For the eICU database, cardiovascular, total SOFA and longitudinal evaluations were conducted in a subset of 289 000 patients (72%), excluding hospitals that did not report vasopressor or inotrope usage. Reclassification analyses found that 49% of patients had the same total SOFA-1 and SOFA-2 score, whereas SOFA-2 was greater in 11% (median difference, 2; IQR, 1 to 3 points) and lower in 40% (median difference, −3; IQR, −4 to −1] points; **Figure 5**A and eFigures 35-41 in Supplement 2). ICU mortality was 4.7% when scores were equal, 13.5% when SOFA-2 was higher, and 8.6% when SOFA-2 was lower than SOFA-1.

Predictive Validity

The predictive validity of SOFA-2 and SOFA-1 for ICU mortality was similar (SOFA-2 AUROC, 0.81; 95% CI, 0.81-0.81; SOFA-1 AUROC, 0.80; 95% CI, 0.79-0.80 combined, single-stage estimate). The 2-stage, meta-analyses estimates were similar (SOFA-2 AUROC, 0.79; 95% CI, 0.76-0.81; SOFA-1 AUROC, 0.77; 95% CI, 0.74-0.81; Figure 5B, eFigures 42-43, and eTable 10 in Supplement 2). These data were consistent for individual cohorts, complete-case analysis, sensitivity analyses, and organ system (eFigures 44-46 and eTable 11 in Supplement 2).

Longitudinal Data

Daily SOFA-2 scores were measured in 553 901 patients (eICU n = 289 000, KPNC n = 258 515, OutcomeRea n = 6386), totaling 2 072 285 patient-days. Approximately 80% of patients

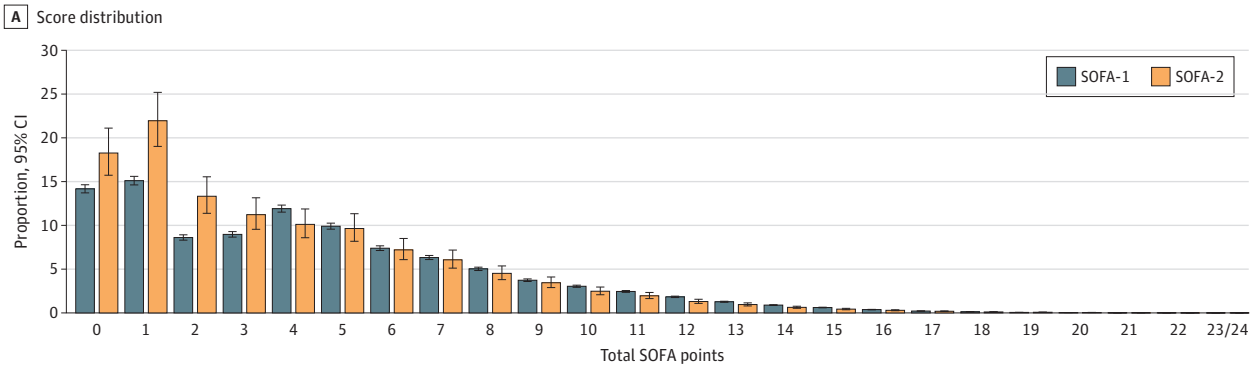
Table 1. Summary of Cohorts

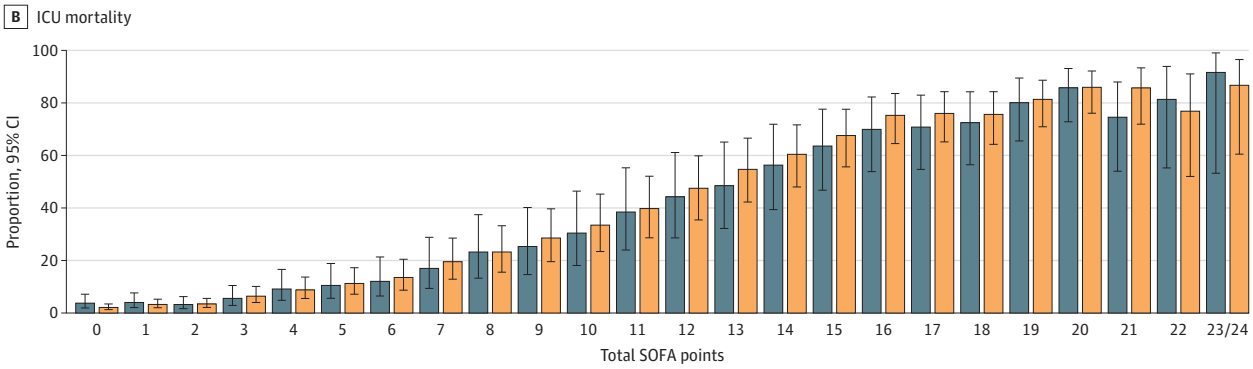
	Internal validation				External validation					
Variable	ANZICS	ASDI	KPNC	ORCHESTRA	eICU	GiViTI-PROSAFE	JIPAD	GiViTI-MargheritaTre	NICRF	OutcomeRea
Period	2018-2023	2014-2023	2014-2023	2022-2023	2014-2015 and 2021-2022	2014-2023	2018-2022	2014-2023	2019-2023	2014-2023
Countries	Australia and New Zealand	Austria	US	Brazil	US	Italy	Japan	Italy	Nepal	France
Type	Registry	Registry	Electronic health record	Registry	Electronic health record	Registry	Registry	Electronic health record	Registry	Registry
No. of ICUs	181 Australia; 21 New Zealand	146	21	218	339	220	100	25	24	24
No. of hospitals	181 Australia; 20 New Zealand	82	21	111	339	>159	NA	22	15	24
No. of patients	1 091 034	406 469	258 515	342 338	401 613	516 052	245 250	42 586	29 227	6386
Age, mean (SD), y	62 (17)	66 (16)	65 (17)	62 (22)	63 (17)	66 (16)	68 (15)	65 (16)	56 (19)	61 (16)
Sex, No./total (%)										
Female	475 582/ 1 090 119 (43.6)	170 520/406 439 (42.0)	113 826 (44.0)	176 059 (51.4)	180 530/401 434 (45.0)	199 214 (39)	94 393 (38.5)	16 692 (39)	12 184 (41.7)	2333 (37)
Male	614 537/ 1 090 119 (56.3)	235 919/406 439 (58.0)	144 689 (56.0)	166 279 (48.6)	220 904/401 434 (55.0)	315 887 (61)	150 851 (61.5)	25 894 (61)	17 023 (58.2)	4042 (63)
ICU admission type, No. (%)										
Medical	496 151 (45.5)	204 784/402 157 (50.9)	173 854 (67.3)	254 639 (74.4)	339 684 (84.6)	215 763 (42)	75 401 (30.7)	21 295 (50)	23 920 (81.8)	4924 (87)
Elective surgical	419 149 (38.4)	130 674/402 157 (32.5)	39 772 (15.4)	62 866 (18.4)	58 050 (14.5)	184 028 (36)	138 127 (56.3)	11 528 (27)	3717 (12.7)	180 (3)
Emergency surgical	175 734 (16.1)	66 699/402 157 (16.6)	44 886 (17.4)	24 833 (7.3)	3879 (1.0)	116 252 (23)	31 722 (12.9)	9586 (23)	1590 (5.4)	548 (10)
Invasive mechanical ventilation at ICU admission	410 456 (37.6)	139 209/406 400 (34.3)	54 399 (21.0)	40 792 (11.9)	54 105 (13.5)	352 604 (70)	90 486 (36.9)	29 480 (69)	5516 (18.9)	2179 (34)
ICU LOS, median (IQR), d	2 (1-3)	3 (2-6)	2 (1-3)	3 (2-5)	2 (1-3)	2 (1-6)	1 (1-4)	2 (1-7)	4 (2-6)	4 (2-9)
ICU mortality	49 483 (4.5)	39 538 (9.7)	26 238 (10.1)	28 817 (8.4)	27 977 (7.0)	75 365 (14.6)	9896 (4.0)	6924 (16.0)	4560 (15.6)	1310 (20.5)
Domain contribution to the pooled results	Brain (0-4), respiratory (0-4), cardiovascular (0-1),[a] liver (0-4), kidney (0-4), hemostasis (0-4)	Brain (0-4), respiratory (0-4), cardiovascular (0-4), liver (0-4), kidney (0-4), hemostasis (0-4)	Brain (0-4), respiratory (0-4), cardiovascular (0-4), liver (0-4), kidney (0-4), hemostasis (0-4)	Brain (0-4), respiratory (0-4), cardiovascular (0-1),[a] liver (0-4), kidney (0-4), hemostasis (0-4)	Brain (0-4), respiratory (0-4), cardiovascular (0-4), liver (0-4), kidney (0-4), hemostasis (0-4)	Brain (0-4), cardiovascular (0-1),[a] liver (0,3,4),[a] hemostasis (0,1,4),[a]	Brain (0-4), respiratory (0-4), cardiovascular (0-1),[a] liver (0-4), kidney (0-4), hemostasis (0-4)	Brain (0-4), respiratory (0-4), cardiovascular (0-4), liver (0-4), kidney (0-4), hemostasis (0-4)	Brain (0-4), respiratory (0-4), cardiovascular (0-1),[a] liver (0-4), kidney (0-4), hemostasis (0-4)	Brain (0-4), respiratory (0-4), cardiovascular (0-4), liver (0-4), kidney (0-4), hemostasis (0-4)

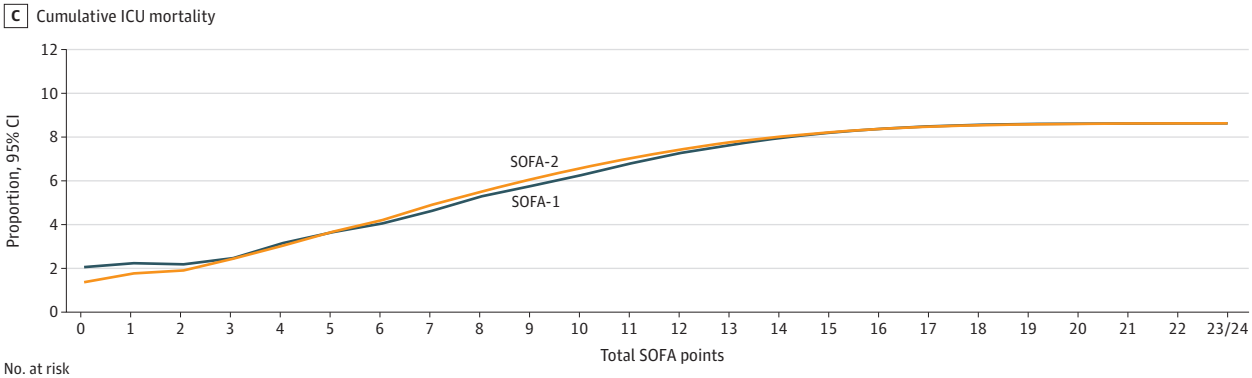
Abbreviations: ANZICS, Australian and New Zealand Intensive Care Society; ASDI, Austrian Center for Documentation and Quality Assurance in Intensive Care; eICU, electronic Intensive Care Unit Collaborative Research Database; ICU, intensive care unit; GiViTi, Gruppo Italiano per la Valutazione degli Interventi in Terapia Intensiva; JIPAD, Japanese Intensive Care Patient Database; KPNC, Kaiser Permanente Northern California; LOS, length of stay; MargheritaTre, Petal Three; NICRF, Nepal Intensive Care Research Foundation; ORCHESTRA, Organizational Characteristics in Critical Care; OutcomeRea, OutcomeRéanimation.

[a] Domains that were derived partially (ie, not the complete 0-4 point score). At least 2 domains with at least 2 consecutive points were the criteria for inclusion of a cohort.

Figure 3. Distribution and ICU Mortality for Total SOFA-1 and SOFA-2 at ICU Admission From the Meta-Analyses of 3 Cohorts at the External Validation (Stage 8)







No. at risk

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23/24
SOFA-1																								
Encounters	46 977	49 868	28 570	29 689	39 525	32 797	24 514	20 907	16 576	12 359	10 083	8115	6067	4208	2937	2017	1252	703	432	189	119	45	18	5
ICU mortality	967	1199	575	1089	2314	2167	1904	2420	2689	2086	2116	2229	1961	1493	1259	1019	721	413	266	133	93	29	14	5
SOFA-2																								
Encounters	51 096	62 527	39 203	34 111	30 926	29 687	22 085	18 333	13 774	10 459	7528	5882	3975	2912	1920	1312	896	564	335	221	134	60	18	14
ICU mortality	699	1313	901	1656	2052	2474	2253	2785	2492	2369	2016	1921	1555	1337	992	773	607	386	230	166	108	50	13	13

The 3 cohorts pooled are the electronic Intensive Care Unit Collaborative Research Database, Gruppo Italiano per la Valutazione degli Interventi in Terapia Intensiva–MargheritaTre, and OutcomeRéanimation. A and B, The bars represents the proportion of the outcome for each point and the whiskers, 95% CIs. Estimates were retrieved from multilevel meta-analysis models pooling data from the 3 cohorts that contain data allowing calculation of the Sequential Organ Failure Assessment (SOFA) score in its totality. C, Cumulative intensive care unit (ICU) mortality was retrieved from the combined aggregated raw data from the same 3 cohorts. Only 4 patients (all of whom died) scored 24 points in SOFA-2. They are thus grouped with patients scoring 23 points as 23/24 points. All results considered missing values on the specific domains as 0; ie, without dysfunction. The analysis for the internal validation as well as for complete-case data are shown in eFigures 9 to 11 in Supplement 2.

contributed to day 2, and 60% to day 3 cross-sectional daily estimates (eFigure 47 in Supplement 2). The median SOFA-2 score was 6 (IQR, 4-9) on ICU day 1; 5 (IQR, 3-8) on ICU day 2; and 5 (IQR, 3-8) on ICU day 3. SOFA-2 values, on average, were

Table 2. 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	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 >.02 to ≤0.4 µg/kg/min) with any other vasopressor or inotrope or mechanical support[i,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 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 (includes chronic use)[o,p,q]
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; PaO_2:FIO_2, 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, 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, bedside use, research, etc). For sequential scoring, for missing data after day 1, it is 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 the International Management of Pain, Agitation, and Delirium in Adult Patients in the ICU Guidelines.[3]

[f] Use the arterial oxygen saturation (SpO_2) to FIO_2 ratio only when the PaO_2:FIO_2 ratio is unavailable and when the SpO_2 is less than 98%. Cutoffs: 0 points, greater than 300 mm Hg; 1 point, 300 mm Hg or less; 2 points, 250 mm Hg or less; 3 points, 200 mm Hg or less with ventilatory support; 4 points, 120 mm Hg or less 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 PaO_2:FIO_2 or SpO_2:FIO_2 ratio being within the specified range. Changes in PaO_2/FIO_2 or SpO_2:FIO_2 within an 1 hour (eg, after suctioning) should not be considered.

[h] Patients not receiving advanced respiratory support can score a maximum of 2 points unless ventilatory support is (1) not available or (2) precluded due to the ceiling of treatment; if so, severity is scored by the PaO_2:FIO_2 or SpO_2:FIO_2 ratio.

[i] If used for respiratory failure, ECMO (all forms) should be scored 4 in the respiratory component (regardless of PaO_2:FIO_2 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).[39] 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.[40-42]

[m] When vasoactive drugs are unavailable or precluded due to a ceiling of treatment, use the following MAP cutoffs for scoring: 0 point, 70 mm Hg or higher; 1 point, 60 to 69 mm Hg; 2 points, 50 to 59 mm Hg; 3 points, 40 to 49 mm Hg; 4 points, less than 40 mm Hg.

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

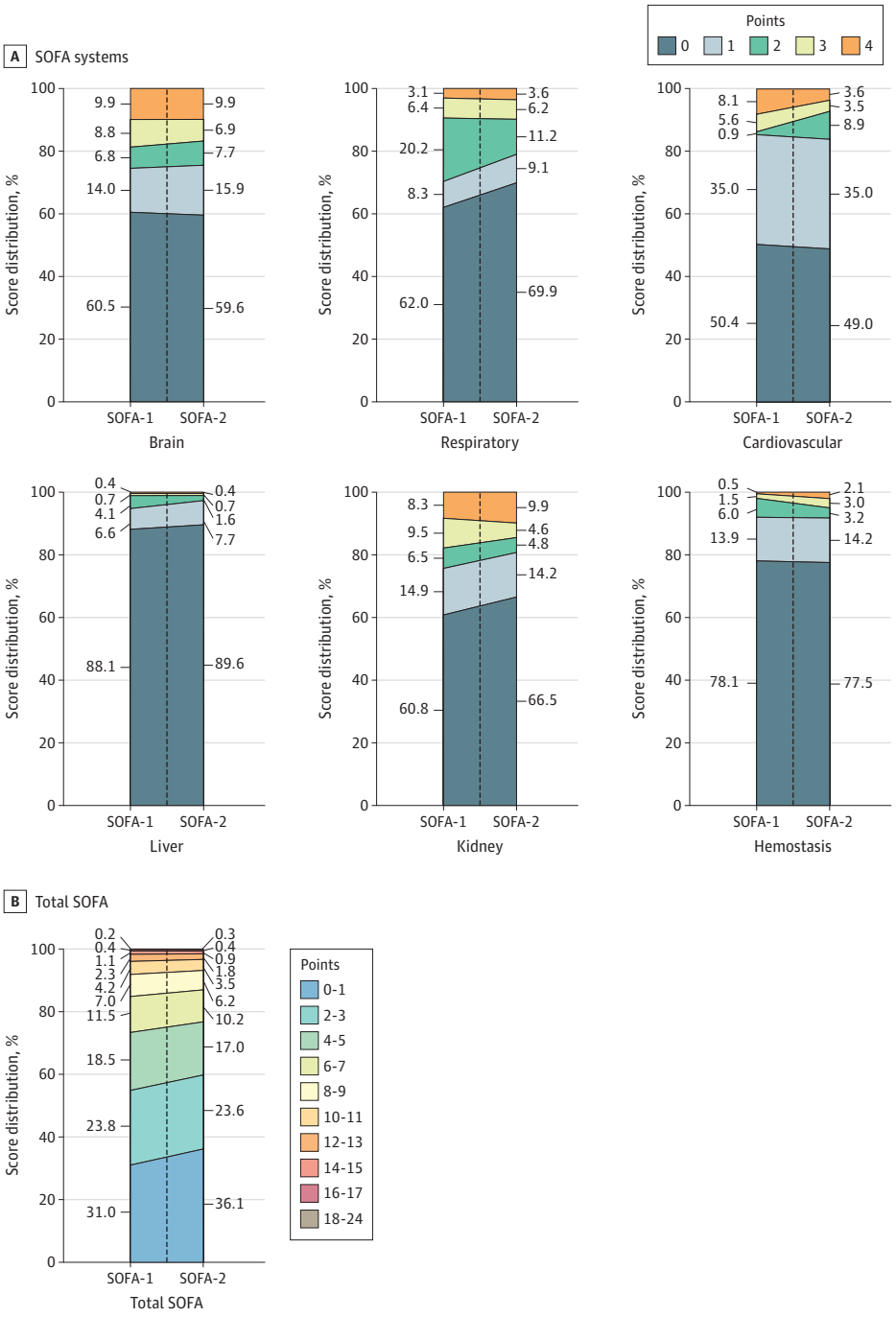
[o] Excludes patients receiving RRT exclusively for nonrenal causes (eg, removal of toxic products, bacterial toxins, cytokines).

[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.

higher among patients who died in the ICU (median SOFA-2 score, 9; IQR, 6-13] vs 5; IQR, 3-7; $P < .001$; eFigure 48 in Supplement 2). The pattern over time differed for each organ system (eFigure 49 in Supplement 2). Predictive validity of the SOFA-2 score for ICU mortality was highest using the mean score across the ICU stay (AUROC, 0.87; 95% CI, 0.80-0.92) compared with the maximum score achieved on any day (AUROC, 0.84, 95% CI, 0.79-0.87; eFigure 50 in Supplement 2). Alternative approaches to handling longitudinal missing data did not meaningfully change results (eFigure 51 in Supplement 2).

Figure 4. Distribution Change From SOFA-1 to SOFA-2 Systems and Total Scores at ICU Admission Combining Data From 2 Internal Cohorts and 3 External Cohorts Validation



The proportion of each score point was retrieved from the 5 cohorts (n = 1 002 956) for which the complete Sequential Organ Failure Assessment (SOFA) score could be estimated, comprising 2 from the internal validation (Austrian Center for Documentation and Quality Assurance in Intensive Care and Kaiser Permanente Northern California) and 3 from the external validation (the electronic Intensive Care Unit Collaborative Research Database, Gruppo Italiano per la Valutazione degli Interventi in Terapia Intensiva–MargheritaTre, and OutcomeRéanimation).

A, The distribution change from SOFA-1 to SOFA-2 for each system is shown. B, The distribution change from total SOFA-1 to SOFA-2 is shown. The shift between SOFA-1 and SOFA-2 represents the shift on the aggregated distribution and not individual-level change. No difference in the aggregated distribution shift was expected for certain domain points. Nevertheless, in some cases, small differences were seen between SOFA-1 and SOFA-2 due to changes in the exact threshold definition and in the differences in scores (eg, extracorporeal membrane oxygenation and mechanical devices for cardiovascular domain). All results with missing values on the specific domains were considered as 0; ie, without dysfunction. For SOFA-2 calculation instructions, see Table 2 footnotes.

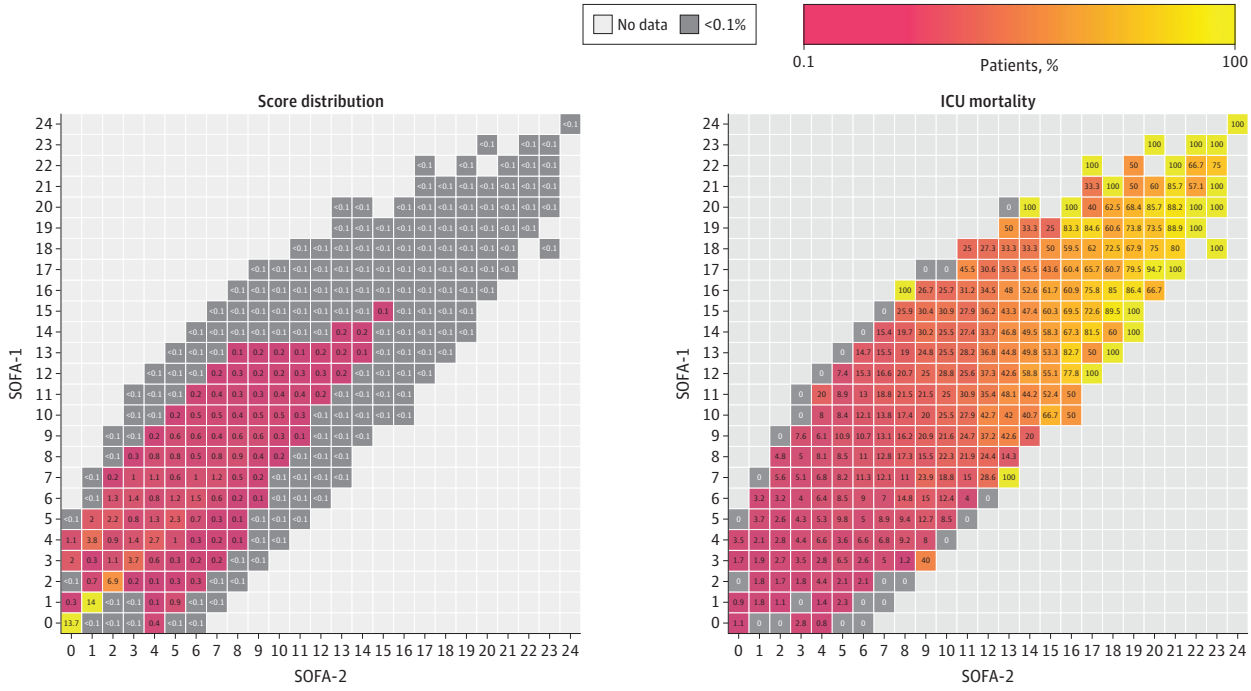
Discussion

The SOFA score, updated to match contemporary practice for organ support of the critically ill patient through an mDelphi and data-driven analyses, maintained 6 organ systems with a score ranging from 0 to 24. In data from 9 countries and more than 3 million ICU admissions in varied geographical and economic areas, the SOFA-2 score was associated with ICU mortality.

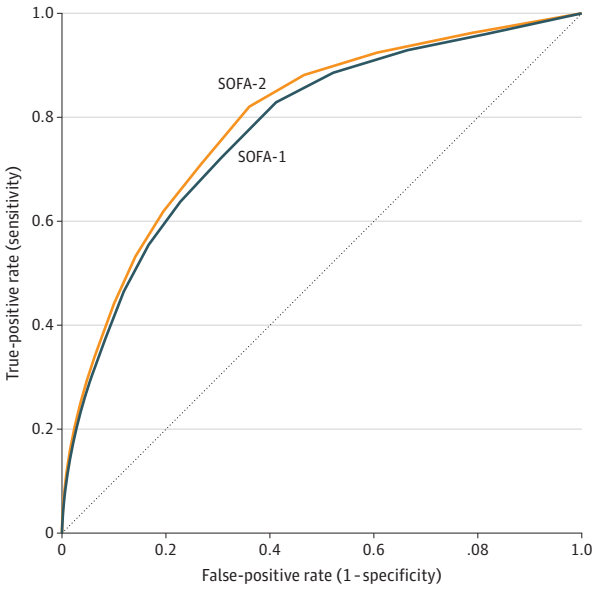
The need to update SOFA-1 has been long recognized by intensive care physicians and clinical researchers.[4,6,7] SOFA-1 has become outdated for some organ systems, because of advances in drugs and organ support devices. As we also observed, this limitation is particularly evident within cardiovascular, respiratory, and kidney systems.[5,8] Additionally, ambiguities in interpretation have led to inconsistent scoring.[6] SOFA-2 addresses these deficiencies by incorporating currently used drugs and devices, providing explicit instructions,

Figure 5. Reclassification and AUROC for Total SOFA-1 and SOFA-2 at ICU Admission

A Reclassification between SOFA-1 and SOFA-2 in eICU cohort



B AUROCs for ICU mortality for 5 cohorts



A, The reclassification heat map used individual-level data from the electronic Intensive Care Unit Collaborative Research Database (eICU) cohort comprising 289 000 patients. This represents results from the eICU cohort because it was the cohort with accessible patient-level data publicly available. In the score distribution, the percentages inside each cell represent the proportion of patients in that cell compared with the total. The percentages inside each cell represent ICU mortality.

B, The area under the receiver operating characteristic (AUROC) of total SOFA-1 and SOFA-2 for ICU mortality was retrieved from the 5 cohorts (n = 1 002 956) for which the complete SOFA score could be estimated: 2 from the internal validation (Austrian Center for Documentation and Quality Assurance in Intensive Care and Kaiser Permanente Northern California); 3 from the external validation (eICU, Gruppo Italiano per la Valutazione degli Interventi in Terapia Intensiva–MargheritaTre, and OutcomeRéanimation). The AUROC was built expanding the aggregated data from the five cohorts.

All results considered missing values on the specific domains as 0; ie, without dysfunction.

and extending applicability to ceilings of treatment and resource-limited settings, with the goals of improving standardization and generalizability.

The most notable differences between SOFA-1 and SOFA-2 were observed in respiratory, cardiovascular, and kidney systems. These changes result in a more plausible distribution of intermediate scores (eg, 0-2 points for respiratory dysfunction and 2 points for cardiovascular dysfunction). Reclassification between total SOFA-1 and SOFA-2 occurred for nearly half of the patients. The associated ICU mortality gradients (ie, 13.5% when SOFA-2 was higher and 8.6% when SOFA-2 was lower than SOFA-1) indicate that SOFA-2 better aligns with organ dysfunction. By redistributing points in the key systems, SOFA-2 improves content validity and enhances interpretability with contemporary clinical practice, satisfying 2 priorities underpinning the SOFA update.

Normal value imputation for day-1 data provides a more realistic representation of score distribution among all ICU patients.[11] Observed ICU mortality was higher in complete-case data (19.3%) compared with the imputed version (9.5%). This may relate to more comprehensive data collection in the most severely ill patients. For longitudinal missing values, predictive validity was similar across imputation methods up to 7 days. For bedside use, we recommend the last-observation-carried-forward method as the best trade-off. Methods such as multiple imputation may be preferable for research purposes, including syndromic criteria or trial outcomes.[8,35,38]

SOFA-2 demonstrates good predictive validity for ICU mortality using data from the first ICU day and from longitudinal analyses. These data are consistent with a prior systematic review of 18 studies (1999-2008, ≈30 000 patients) of SOFA-1 at ICU admission, where the AUROC for short-term mortality ranged from 0.61 to 0.88.[43] Most importantly, predictive validity was consistent across countries and cohorts, suggesting generalizability to variations in geoeconomics, case-mix, and management strategies.

The study has notable strengths. First, the diverse working group comprised intensive care, epidemiology, and data science experts from multiple regions, including low-resource settings. Second, the real-world analyses, including low- and middle-income countries with a wide geographic distribution, from 1319 ICUs in 9 countries support generalizability. Prior steps to validate SOFA-1 relied on far fewer patients.[2,44] Third, the current process was a collaboration between multiple rounds of mDelphi and data analyses, enabling the development of a score that is both evidence-based and applicable at the bedside across different ICUs.

Limitations

The study has several limitations. First, only ICU mortality was used for predictive validity assessment because ICU outcome alone was collected in all contributing cohorts. Importantly, we did not necessarily seek to improve the predictive ability of SOFA-1 nor to compete with existing prognostic scores such as the Acute Physiology and Chronic Health Evaluation but rather to enhance score distribution to better describe organ dysfunction in the general ICU population. ICU mortality is strongly associated with other critical illness outcomes, such as ICU stay, cost, staff burnout, family satisfaction, and hospital mortality.[45-48] Second, although recommended by the early mDelphi discussions, gastrointestinal and immune system dysfunction were not included in the final SOFA-2 score. The inability of candidate variables to satisfy both content and predictive validity precluded inclusion.[11] Third, the decisions for SOFA-2 thresholds were based on data from the first day of ICU admission. Alternative thresholds may be optimal later in the course of critical illness. Fourth, the SOFA-2 score was developed and validated only in intensive care patients; generalizability to patients located outside the ICU, eg, emergency department patients or pediatric care requires future investigation.

Conclusions

The SOFA-2 score, updated to include contemporary organ support treatments and new score thresholds, describes organ dysfunction in a large, geographically and socioeconomically diverse population of critically ill adults, and is supported by good predictive validity.

ARTICLE INFORMATION

Accepted for Publication: October 9, 2025.

Published Online: October 29, 2025. doi:10.1001/jama.2025.20516

Author Affiliations: DataHealth Lab, Institut de Recerca Sant Pau (SANT PAU), Barcelona, Spain (Ranzani); ISGlobal, Barcelona, Spain (Ranzani); Pulmonary Division, Faculty of Medicine, Heart Institute, Hospital das Clinicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil (Ranzani); Bloomsbury Institute of Intensive Care Medicine, Division of Medicine, University College London, London, United Kingdom (Singer, Arulkumaran); Postgraduate Program, Instituto D'Or de Pesquisa e Ensino Rio de Janeiro, Brazil (Salluh, Aryal, Sendagire, Soares); Institute for Regeneration and Repair, University of Edinburgh, Edinburgh, Scotland (Shankar-Hari); The Australian and New Zealand Intensive Care Society (ANZICS) Centre for Outcomes and Resources Evaluation, Australian (Pilcher, Neto, Secombe); Department of Intensive Care, Alfred Health, Prahran, Victoria, Australia (Pilcher); New Zealand Intensive Care Research Centre (ANZIC-RC), School of Public Health and Preventive Medicine, Monash University (Pilcher); 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, Ireland (Laffey); Department of Anaesthesia, Galway University Hospitals, Galway, Ireland (Laffey); University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland (Reinikainen); 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, Ruckly); Medical and Infectious Diseases Intensive Care Unit, APHP, Hôpital Bichat-Claude Bernard, Paris, France (Timsit, Ruckly); Management, Quality and Data Committee, SATI-Q Program, Argentine Society of Intensive Care (SATI), Buenos Aires, Argentina (Arias Lopez); PICU, Hospital de Niños Ricardo Gutierrez, Buenos Aires, Argentina (Arias Lopez); Nepal Intensive Care Research Foundation, Nepal (Aryal); APHP, Paris City University, Saint-Louis Hospital, Paris, France (Azoulay); Laboratory for Computational Physiology, Massachusetts Institute of Technology, Cambridge,

Massachusetts (Celi, Hao, X. Liu); Division of Pulmonary, Critical Care and Sleep Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts (Celi); Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts (Celi); Division of Critical Care, McMaster University, Hamilton, Ontario, Canada (Chaudhuri); 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, Gent, Belgium (De Waele); Interdepartmental Division of Critical Care, University of Toronto, 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, Beijing, China (Du); Chinese Academy of Medical Sciences, Beijing, China (Du); Maccabi Healthcare Services, Sharon Region, Hebrew University Faculty of Medicine, Tel Aviv, Israel (Einav); Center for Medical Data Science, Medical University of Vienna, Vienna, Austria (Engelbrecht); Nat-Intensive Care Surveillance M.O.R.U, Colombo, Sri Lanka (Fazla); Department Intensive Care and SODIR-VHIR Research Group, Vall d'Hebron University Hospital, Barcelona, Spain (Ferrer); Laboratory of Clinical Data Science, Department of Medical Epidemiology, Mario Negri Institute for Pharmacological Research IRCCS, Milan, Italy (Finazzi, Magatti, Samei, Tricella); Deptarment 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); Kaiser Permanente Division of Research, Pleasanton, California (Greene); 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); Emory University School of Medicine, Atlanta, Georgia (Hollenberg); Department of Precision Medicine in Medical, Surgical and Critical Care, University of Palermo, Palermo, Italy (Ippolito); Department of Cardiology, Pulmonology and Vascular Medicine and Cardiovascular Research Institute Dusseldorf, Heinrich-Heine-University Duesseldorf, Duesseldorf, Germany (Jung); Department of Anesthesiology and Intensive Care Medicine, Northern State Medical University, Arkhangelsk, Russia (Kirov); Department of Clinical Biostatistics, Graduate School of Medicine, Kyoto University, Kyoto, Japan (Kobari, Tanaka); 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, Department of Critical Care, Kaiser Permanente Medical Center, Santa Clara, California (V. Liu); Division of Intensive Care, Hospital de Base-Sao Jose do Rio Preto and Faculdade de Medicina de Sao 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); Austrian Center for Documentation and Quality Assurance in Intensive Care, Vienna, Austria (B. Metnitz); Department of General Anaesthesiology and Intensive Care, Medical University of Graz, Graz, Austria (P. Metnitz); Department of Anesthesiology, Critical Care and Pain, Tata Memorial Hospital, Homi Bhabha National Institute, Mumbai, India (Myatra); Departments of Medicine and 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 (Paiva); Grupo de Investigacão e Desenvolvimento em Infeção e Sépsis, Porto, Portugal (Paiva); Department of Critical Care, University of Pretoria and 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); 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 and VA Center for Clinical Management Research, Ann Arbor, Michigan (Prescott); University of Tartu, Estonia (Blaser); Luzerner Kantonsspital, Lucerne, Switzerland (Blaser); Serviço de Terapia Intensiva na IAMSPEInstituto de Assistência Médica ao Servidor Público Estadual, São Paulo, Brazil (Rezende); IRCCS, Policlinico San Martino, Genova, Italy (Robba); Dipartimento di Scienza Chirurgiche Diagnistiche Integrate, University di Genova, Genova, Italy (Robba); Department of Medicine, and Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, Ontario, Canada (Rochwerg); 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); Intensive Care Specialist, Queen Elizabeth Hospital, Woodville, SA 5011, Australia (Secombe); 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, Ireland (Simpkin); SFI Insight Centre for Data Analytics Ireland, Galway (Simpkin); 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); Deptartment 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, Canada (Zampieri); Adult Critical Care, St George's University Hospitals NHS Foundation Trust, London, United Kingdom (Rhodes); 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).

Author Contributions: Drs Ranzani and Salluh had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Ranzani, Singer, Salluh, Shankar-Hari, Pilcher, Juffermans, Laffey, Reinikainen, Serpa Neto, Arulkumaran, Aryal, Azoulay, de Lange, De Waele, dos Santos, Einav, Haniffa, Jung, Kirov, Lipman, Martin, P. Metnitz, Pekkarinen, Piquilloud, Rezende, Robba, Schenck, Sendagire, Soares, Szczeklik, Takala, Vincent, Wendon, Zampieri, Rhodes, Moreno.
Acquisition, analysis, or interpretation of data: Ranzani, Singer, Salluh, Shankar-Hari, Pilcher, Berger-Estilita, Coopersmith, Juffermans, Laffey, Reinikainen, Serpa Neto, Tavares, Timsit, Arias López, Arulkumaran, Aryal, Azoulay, Celi, Chaudhuri, de Lange, De Waele, dos Santos, Du, Einav, Engelbrecht, Fazla, Ferrer, Finazzi, Fujii, Gershengorn, Greene, Hao, Hasan, Hollenberg, Ippolito, Jung, Kobari, Lakbar, V. Liu, X. Liu, Lobo, Magatti, Martin, B. Metnitz, P. Metnitz, Myatra, Oczkowski, Paiva, Paruk, Pekkarinen, Piquilloud, Pölkki, Prescott, Reintam Blaser, Rezende, Robba, Rochwerg, Ruckly, Samei, Schenck, Secombe, Siaw-Frimpong, Simpkin, Summers, Takala, Tanaka, Tricella, Wendon, Moreno.
Drafting of the manuscript: Ranzani, Singer, Salluh, Shankar-Hari, Pilcher, Juffermans, Serpa Neto, Arulkumaran, Aryal, dos Santos, Fazla, Kirov, Lipman, Lobo, Piquilloud, Sendagire, Rhodes, Moreno.
Critical review of the manuscript for important intellectual content: Singer, Salluh, Shankar-Hari, Pilcher, Berger-Estilita, Coopersmith, Juffermans, Laffey, Reinikainen, Serpa Neto, Tavares, Timsit, Arias López, Arulkumaran, Aryal, Azoulay, Celi, Chaudhuri, de Lange, De Waele, dos Santos, Du, Einav, Engelbrecht, Ferrer, Finazzi, Fujii, Gershengorn, Greene, Haniffa, Hao, Hasan, Hollenberg, Ippolito, Jung, Kobari, Lakbar, Lipman, V. Liu, X. Liu, Magatti, Martin, B. Metnitz, P. Metnitz, Myatra, Oczkowski, Paiva, Paruk, Pekkarinen, Piquilloud, Pölkki, Prescott, Reintam Blaser, Rezende, Robba, Rochwerg, Ruckly, Samei, Schenck, Secombe, Siaw-Frimpong, Simpkin, Soares, Summers, Szczeklik, Takala, Tanaka, Tricella, Vincent, Wendon, Zampieri, Rhodes, Moreno.
Statistical analysis: Ranzani, Salluh, Shankar-Hari, Pilcher, Celi, Chaudhuri, Engelbrecht, Fazla, Greene, Hao, Kobari, V. Liu, Magatti, Martin, B. Metnitz, P. Metnitz, Ruckly, Samei, Soares, Tanaka, Tricella, Zampieri, Rhodes, Moreno.
Administrative, technical, or material support: Ranzani, Singer, Shankar-Hari, Tavares, Arulkumaran, Aryal, Azoulay, de Lange, Du, Fujii, Haniffa, Hasan, Jung, Lakbar, X. Liu, Oczkowski, Prescott, Rochwerg, Sendagire, Summers, Szczeklik, Vincent, Rhodes, Moreno.
Supervision: Ranzani, Singer, Salluh, Shankar-Hari, Rhodes, Moreno.

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 Singer reported receiving grants from Gentian; serving on the administrative boards of Biotest, Pfizer, Safeguard Biosystems, Aptarion, Roche Diagnostics, Sanofi, Matisse, Volition, Deepull, Bayer, and Radiometer, paid to his institution; speaker fees from Biomerieux and AOP Health to his institution; royalties from Deltex Medical to his institution; having stock options outside the submitted work; and serving as an unpaid council member for the International Sepsis Forum and as an unpaid sepsis topic advisor for the National Institute for Health and Care Excellence.
Dr Shankar-Hari reported receiving the National Institute for Health Research Clinician Scientist Award (CS-2016-16-011; 2017-2023), research grants from the Chief Scientist Office of Scotland, the UK National Institute for Health and Care Research, UK Medical Research Council, Canadian Institute for Health Research, and research grants from the Huo Foundation; and for serving as chief investigator of the Glucocorticoids in Adults With Acute Respiratory Distress Syndrome Trial (https://www.isrctn.com/ISRCTN15076735) and by the RAITS Research Programme for the Time Critical Precision Medicine in Adult Critically Ill Patients (TRAITS Program; https://traits-trial.ed.ac.uk).
Dr Berger-Estilita reported receiving nonfinancial support paid to her institution from Medtronic during the conduct of the study. Dr Laffey reported receiving consulting fees from Cellenkos outside the submitted work. Dr Serpa Neto reported receiving support from Hamilton Medical outside the submitted work. Dr Azoulay reported receiving speaker fees from Gilead, Mundipharma, Alexion, Fisher & Paykel, and Baxter outside the submitted work. Dr De Waele reported receiving honoraria from BioMerieux, Grifols, Menarini, MSD, Pfizer, Roche Diagnostics, VIatris, all paid to his employer, outside the submitted work, and being supported by a senior research grant from the Research Foundation (FWO 1881020N). Dr dos Santos reported receiving grants from the Canadian Institute of Health Research and being a tier 1 Canada research chair and a University of Toronto Pitts research chair during the conduct of the study. Dr Ferrer reported receiving personal fees from Cytosorbents, Viatris, and AOP Health outside the submitted work. Dr Fujii reported receiving fees for a manuscript and honorarium from the Japan Council for Quality Health Care, consulting fees from Wolters Kluwer Health, and research grants from the Japan Society for the Promotion of Science and Terumo Life Science Foundation for research outside the submitted work.
Dr Gershengorn reported receiving research support from the University of Miami Hospital and Clinics Data Analytics Research Team as part of her salary and serving as an editor of *CHEST Critical Care* outside the submitted work. Dr Haniffa reported receiving grants from Wellcome Trust during the conduct of the study and travel support from the World Health Organization outside the submitted work. Dr Lakbar reported receiving speaker fees from AOP Health and Viatris outside the submitted work. Dr V. Liu reported receiving grants from the National Institutes of Health during the conduct of the study. Dr Martin reported serving on the research advisory committee of Grifols, on the data and safety monitoring board for Eagle Pharmaceuticals, as a section editor for UpToDate, and as an editor for *Critical Care Clinics* outside the submitted work. Dr Oczkowski reported receiving travel support from Fisher & Paykel and consulting fees from the Brain Trauma Foundation and VitalAire outside the submitted work. Dr Paruk reported receiving speaker fees from Baxter, Thermo Fischer Scientific, Acino, and Pfizer and serving as an advisor to Baxter, Kabi, and Novo Nordisk outside the submitted work. Dr Piquilloud reported serving on the medical advisory board of Getinge Memmer and receiving fees for participating in a symposia from Löwenstein, Hamilton, Fisher & Paykel, Air Liquide Medical Systems, and GE Healthcare outside the submitted work. Dr Prescott reported serving as cochair of the Surviving Sepsis Campaign Guidelines. Dr Reintam Blaser reported receiving a grant from the Estonian Research Council outside the submitted work (PRG1255). Dr Robba reported receiving speaker fees from BD outside the submitted work. Dr Soares reported being a founding member and shareholder of Epimed Solutions outside the submitted work. Dr Summers reported receiving grants from National Institute for Health and Care Research (133788) and from UKRI-Medical Research Council (MR/S035753/1 and MR/X005070/1s) during the conduct of the study. Dr Vincent reported being a codeveloper of the original SOFA score and has no financial interest related to its use or dissemination. Dr Zampieri reported receiving support from Baxter outside the submitted work. No other disclosures were reported.

Disclaimer: The views expressed in this publication are those of the authors.

Data Sharing Statement: See Supplement 3.

Additional Contributions: We acknowledge the major contribution made to this project by the analysts and data collectors in participating hospitals for the following databases (acknowledgment in Supplement 2): 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 Northern California; Gruppo Italiano per la Valutazione degli Interventi in Terapia Intensiva (GiViTi); Japanese Intensive care Patient Database; Nepal Intensive Care Research Foundation; OutcomeRea; eICU.

REFERENCES

1. Shankar-Hari M, Wunsch H, Rowan K, Singer M, Rubenfeld GD, Angus DC. Reflections on critical care's past, present, and future. *Crit Care Med*. 2021;49(11):1855-1865. doi:10.1097/CCM.0000000000005246

2. Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure: on behalf of the working group on sepsis-related problems of the European Society of Intensive Care Medicine. *Intensive Care Med*. 1996;22(7):707-710. doi:10.1007/BF01709751

3. 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

4. Moreno R, Rhodes A, Piquilloud L, et al. The Sequential Organ Failure Assessment (SOFA) score: has the time come for an update? *Crit Care*. 2023;27(1):15. doi:10.1186/s13054-022-04290-9

5. 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

6. 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

7. Moreno R, Singer M, Rhodes A. Why the Sequential Organ Failure Assessment score needs updating? *Crit Care Sci*. 2024;36. doi:10.62675/2965-2774.20240296-en

8. Gupta T, Puskarich MA, DeVos E, et al. Sequential Organ Failure Assessment component score prediction of in-hospital mortality from sepsis. *J Intensive Care Med*. 2020;35(8):810-817. doi:10.1177/0885066618795400

9. Raschke RA, Agarwal S, Rangan P, Heise CW, Curry SC. Discriminant accuracy of the SOFA score for determining the probable mortality of patients with COVID-19 pneumonia requiring mechanical ventilation. *JAMA*. 2021;325(14):1469-1470. doi:10.1001/jama.2021.1545

10. Salluh JIF, Quintairos A, Dongelmans DA, et al; Linking of Global Intensive Care (LOGIC) and Japanese Intensive Care Patient Database (JIPAD) working group. National ICU registries as enablers of clinical research and quality improvement. *Crit Care Med*. 2024;52(1):125-135. doi:10.1097/CCM.0000000000006050

11. Moreno R, Rhodes A, Ranzani O, et al. Rationale and methodological approach underlying development of the Sequential Organ Failure Assessment (SOFA)-2 Score: a consensus statement. *JAMA Netw Open*. Published online October 29, 2025. doi:10.1001/jamanetworkopen.2025.45040

12. 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(8):762-774. doi:10.1001/jama.2016.0288

13. Angus DC, Seymour CW, Coopersmith CM, et al. A framework for the development and interpretation of different sepsis definitions and clinical criteria. *Crit Care Med*. 2016;44(3):e113-e121. doi:10.1097/CCM.0000000000001730

14. Dongelmans DA, Pilcher D, Beane A, et al. Linking of Global Intensive Care (LOGIC): an international benchmarking in critical care initiative. *J Crit Care*. 2020;60:305-310. doi:10.1016/j.jcrc.2020.08.031

15. 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

16. 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

17. 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

18. 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*. 2024;36. doi:10.62675/2965-2774.20240118-en

19. 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

20. 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 Anestesiol*. 2020;86(12):1305-1320. doi:10.23736/S0375-9393.20.14112-9

21. 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.jcrc.2019.09.004

22. 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/aad10f

23. 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

24. 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

25. McLarty J, Litton E, Beane A, et al; Linking of Global Intensive Care (LOGIC) Collaboration. Non-COVID-19 intensive care admissions during the pandemic: a multinational registry-based study. *Thorax*. 2024;79(2):120-127. doi:10.1136/thorax-2022-219592

26. Wood SN. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J R Stat Soc Series B Stat Methodol*. 2011;73(1):3-36. doi:10.1111/j.1467-9868.2010.00749.x

27. Therneau T, Atkinson B. rpart: Recursive partitioning and regression trees. Version 4.1.24. Published online April 8, 1999. Accessed February 10, 2024. https://cran.r-project.org/web/packages/rpart/rpart.pdf

28. Stijnen T, Hamza TH, Özdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Stat Med*. 2010;29(29):3046-3067. doi:10.1002/sim.4040

29. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw*. 2010;36(3):1-48. doi:10.18637/jss.v036.i03

30. Snell KI, Ensor J, Debray TP, Moons KG, Riley RD. Meta-analysis of prediction model performance across multiple studies: which scale helps ensure between-study normality for the C-statistic and calibration measures? *Stat Methods Med Res*. 2018;27(11):3505-3522. doi:10.1177/0962280217705678

31. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837-845. doi:10.2307/2531595

32. 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

33. Bosch NA, Law AC, Rucci JM, Peterson D, Walkey AJ. Predictive validity of the Sequential Organ Failure Assessment score versus claims-based scores among critically ill patients. *Ann Am Thorac Soc*. 2022;19(6):1072-1076. doi:10.1513/AnnalsATS.202111-1251RL

34. Brinton DL, Ford DW, Martin RH, Simpson KN, Goodwin AJ, Simpson AN. Missing data methods for intensive care unit SOFA scores in electronic health records studies: results from a Monte Carlo simulation. *J Comp Eff Res*. 2022;11(1):47-56. doi:10.2217/cer-2021-0079

35. Alrawashdeh M, Klompas M, Rhee C. The impact of common variations in Sequential Organ Failure Assessment score calculation on sepsis measurement using Sepsis-3 criteria: a retrospective analysis using electronic health record data. *Crit Care Med*. 2024;52(9):1380-1390. doi:10.1097/CCM.0000000000006338

36. Honaker J, King G, Blackwell M, Amelia II. A program for missing data. *J Stat Softw*. 2011;45(7):1-47. doi:10.18637/jss.v045.i07

37. Zeileis A, Grothendieck G. zoo: S3 infrastructure for regular and irregular time series. *J Stat Softw*. 2005;14(6). doi:10.18637/jss.v014.i06

38. Schenck EJ, Hoffman KL, Oromendia C, et al. A comparative analysis of the respiratory subscore of the Sequential Organ Failure Assessment scoring system. *Ann Am Thorac Soc*. 2021;18(11):1849-1860. doi:10.1513/AnnalsATS.202004-399OC

39. Wieruszewski 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

40. 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

41. 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):29-42. doi:10.1186/s13054-023-04322-y

42. 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

43. Minne L, Abu-Hanna A, de Jonge E. Evaluation of SOFA-based models for predicting mortality in the ICU: a systematic review. *Crit Care*. 2008;12(6):R161. doi:10.1186/cc7160

44. 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

45. Kramer AA, Dasta JF, Kane-Gill SL. The impact of mortality on total costs within the ICU. *Crit Care Med*. 2017;45(9):1457-1463. doi:10.1097/CCM.0000000000002563

46. Ferrando P, Gould DW, Walmsley E, et al. Family satisfaction with critical care in the UK: a multicentre cohort study. *BMJ Open*. 2019;9(8):e028956. doi:10.1136/bmjopen-2019-028956

47. Poncet MC, Toullic P, Papazian L, et al. Burnout syndrome in critical care nursing staff. *Am J Respir Crit Care Med*. 2007;175(7):698-704. doi:10.1164/rccm.200606-806OC

48. Hosein FS, Roberts DJ, Turin TC, Zygun D, Ghali WA, Stelfox HT. A meta-analysis to derive literature-based benchmarks for readmission and hospital mortality after patient discharge from intensive care. *Crit Care*. 2014;18(6):715. doi:10.1186/s13054-014-0715-6