

Letters

RESEARCH LETTER

Discriminant Accuracy of the SOFA Score for Determining the Probable Mortality of Patients With COVID-19 Pneumonia Requiring Mechanical Ventilation

The coronavirus disease 2019 (COVID-19) pandemic has raised concern regarding the capacity to provide care for a surge of critically ill patients that might require excluding patients with a low probability of short-term survival from receiving mechanical ventilation.¹ A survey identified 26 unique COVID-19 triage policies, of which 20 used some form of the Sequential Organ Failure Assessment (SOFA) score.²

However, studies performed in 2016 and 2017 have shown only moderate discriminant accuracy of the SOFA score for predicting survival in intensive care unit (ICU) patients with sepsis and an area under the receiver operating characteristic curve (AUROC) of 0.74 to 0.75.^{3,4} We hypothesized that the SOFA score might be less accurate in patients requiring mechanical ventilation for COVID-19 pneumonia because such patients generally have severe single-organ dysfunction and less variation in SOFA scores.

Methods | This retrospective study was approved and exempted from the requirement for informed consent by the University of Arizona institutional review board. Data were from patients treated at 18 ICUs in the southwestern US between March 1, 2020, and August 31, 2020. We included consecutive patients aged 18 years or older with a diagnosis of COVID-19 pneumonia and receiving oxygen therapy for 4 hours or longer before undergoing endotracheal intubation. We calculated that a sample size of 640 patients would provide a 95% CI for an AUROC of $\pm 5\%$, assuming an AUROC of 70% and mortality of 25%.

The main outcome variable was hospital mortality or hospice discharge. The SOFA score comprises 0 to 4 points assigned to each of 6 organ systems based on ratio of PaO₂ to fraction inspired oxygen, Glasgow Coma Scale score, mean arterial pressure, serum creatinine level, bilirubin level, and platelet count. The SOFA score ranges from 0 to 24 points and higher scores indicate worse organ function. The variables used to calculate SOFA score and describe the patients were collected by electronic medical record review.

The SOFA score was calculated using the worst values observed within 48 hours prior to intubation, which is the point in time when ventilator triage for a patient with COVID-19 pneumonia would theoretically occur. The AUROCs for SOFA score and age (as single variables) were calculated and compared using the χ^2 statistic with a 2-sided $P \leq .05$ as the threshold for significance. Stata version 15 (StataCorp) was used.

Results | Between March 1, 2020, and August 31, 2020, 2546 patients with COVID-19 were admitted to study ICUs. Of these, 972 were intubated 4 hours or longer after receiving

Table. Clinical Characteristics of 675 Study Patients

Characteristic	No. (%) ^a
Age, median (IQR), y	63 (53-72)
Age group, y	
18-44	92 (14)
45-64	267 (40)
65-74	199 (29)
75-84	102 (15)
≥ 85	15 (2)
Sex	
Female	270 (40)
Male	405 (60)
Race/ethnicity ^b	
Non-Hispanic White	259 (38)
Hispanic	286 (42)
Native American	68 (10)
Black	28 (4)
Body mass index, median (IQR) ^c	33 (29-39)
Medications	
Dexamethasone	255 (36)
Remdesivir	326 (48)
Anticoagulants	607 (90)
Norepinephrine	33 (5)
Comorbidities	
Diabetes	413 (61)
Hypertension	500 (74)
Coronary artery disease	134 (20)
Chronic obstructive lung disease	112 (17)
Cancer	67 (10)
Laboratory values, median (IQR)	
C-reactive protein, mg/L	168 (96-260)
D-dimer, $\mu\text{g/mL}$	1.76 (0.93-4.66)
Brain-type natriuretic peptide, pg/mL	461 (180-1328)
Creatinine, mg/dL ^d	0.9 (0.65-1.30)
Bilirubin, mg/dL ^d	0.6 (0.4-0.9)
Ratio of PaO ₂ to FiO ₂ ^d	56 (50-68)
Platelets, $\times 10^3/\mu\text{L}$ ^d	225 (164-299)
Arterial pressure, mean (IQR), mm Hg ^d	74 (64-84)
Glasgow Coma Scale score, median (IQR) ^d	15 (13-15)

Abbreviations: IQR, interquartile range; FiO₂, fraction of inspired oxygen.

SI conversion factors: To convert bilirubin to $\mu\text{mol/L}$, multiply by 17.104; creatinine to $\mu\text{mol/L}$, multiply by 88.4; D-dimer to nmol/L, multiply by 5.476.

^a Unless otherwise indicated.

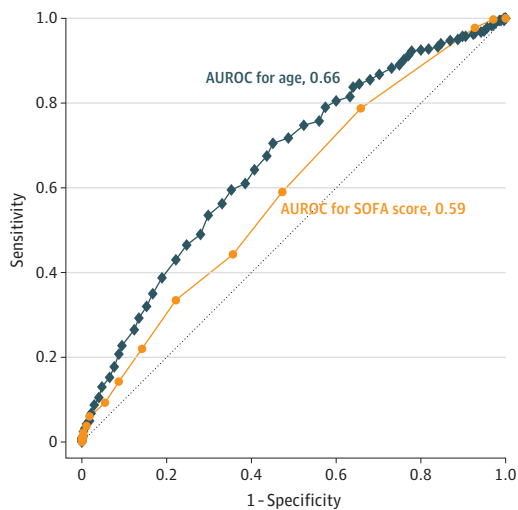
^b Self-reported at hospital admission and was assessed to aid consideration of the generalizability of the study findings.

^c Calculated as weight in kilograms divided by height in meters squared.

^d Indicates the variables that were incorporated into the Sequential Organ Failure Assessment score.

oxygen, but 297 lacked sufficient data to calculate the SOFA score. The characteristics of the remaining 675 patients appear in the **Table**.

Figure. Area Under the Receiver Operating Characteristic Curve (AUROC) for Sequential Organ Failure Assessment (SOFA) Score and Age



The AUROC is for the outcome of hospital mortality or hospice discharge.

The median SOFA score was 6 (interquartile range, 4-8). Respiratory SOFA subscores were 3 to 4 in 83.5% of patients. The other SOFA subscores were 0 to 1 in 72.1% of patients for the renal system, 78.5% for the central nervous system, 94.2% for coagulation, 95.1% for the cardiovascular system, and 96.5% for the hepatobiliary system. Four hundred patients (59.3%) died or were discharged to hospice. The AUROC for SOFA score was 0.59 (95% CI, 0.55-0.63) and for age was 0.66 (95% CI, 0.62-0.70) ($P = .02$; **Figure**).

Discussion | The discriminant accuracy of the SOFA score for mortality prediction in patients prior to intubation for COVID-19 pneumonia was poor and significantly inferior to simply using age. This finding has several potential explanations. The SOFA score was designed for patients with sepsis and only 3 of the 6 equally weighted organ system subscores (respiratory, renal, and hepatobiliary) are associated with mortality in COVID-19.⁵ Compared with previous studies,^{4,6} this study population had higher, less variable SOFA scores with a lower proportion of patients (9/675) with cumulative scores of 0 to 2; such patients do not exhibit severe organ system dysfunction and can be relatively accurately predicted to survive. All patients in this study had respiratory failure requiring mechanical ventilation, which is the major cause of death in patients with COVID-19 pneumonia. Limitations of this study include missing data for 297 of 972 patients and restricted generalizability due to the distinctive patient population.

The SOFA score possesses inadequate discriminant accuracy to be used for ventilator triage of COVID-19 patients. A better option is needed that incorporates variables specifically related to mortality in patients with COVID-19 pneumonia requiring mechanical ventilation.

Robert A. Raschke, MD
Sumit Agarwal, MBBS, MBA
Pooja Rangan, MBBS, MPH
C. William Heise, MD
Steven C. Curry, MD

Author Affiliations: Division of Clinical Data Analytics and Decision Support, University of Arizona College of Medicine, Phoenix (Raschke, Heise, Curry); College of Medicine, University of Arizona, Phoenix (Agarwal); Department of Internal Medicine, Banner-University Medical Center Phoenix, Phoenix, Arizona (Rangan).

Accepted for Publication: February 1, 2021.

Published Online: February 17, 2021. doi:10.1001/jama.2021.1545

Corresponding Author: Robert A. Raschke, MD, 475 N Fifth St, Phoenix, AZ 85004 (raschkebob@gmail.com).

Author Contributions: Dr Raschke 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: Raschke, Rangan, Heise, Curry.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Raschke, Rangan, Heise.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Raschke, Agarwal, Rangan, Heise.

Administrative, technical, or material support: Agarwal, Heise, Curry.

Supervision: Heise.

Conflict of Interest Disclosures: None reported.

Funding/Support: Funded in part by grant 2196 from the Flinn Foundation.

Role of the Funder/Sponsor: The Flinn Foundation had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

1. Truog RD, Mitchell C, Daley GQ. The toughest triage—allocating ventilators in a pandemic. *N Engl J Med*. 2020;382(21):1973-1975. doi:10.1056/NEJMp2005689
2. Antommaria AHM, Gibb TS, McGuire AL, et al. Ventilator triage policies during the COVID-19 pandemic at US hospitals associated with members of the Association of Bioethics Program Directors. *Ann Intern Med*. 2020;173(3):188-194. doi:10.7326/M20-1738
3. 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
4. Raith EP, Udy AA, Bailey M, et al. Prognostic accuracy of the SOFA score, SIRS criteria, and qSOFA score for in-hospital mortality among adults with suspected infection admitted to the intensive care unit. *JAMA*. 2017;317(3):290-300. doi:10.1001/jama.2016.20328
5. Gupta S, Hayek SS, Wang W, et al; STOP-COVID Investigators. Factors associated with death in critically ill patients with coronavirus disease 2019 in the US. *JAMA Intern Med*. 2020;180:1436-1446. doi:10.1001/jamainternmed.2020.3596
6. Kovach CP, Fletcher GS, Rudd KE, Grant RM, Carlborn DJ. Comparative prognostic accuracy of sepsis scores for hospital mortality in adults with suspected infection in non-ICU and ICU at an academic public hospital. *PLoS One*. 2019;14(9):e0222563. doi:10.1371/journal.pone.0222563