SAPS-3 Performance for Hospital Mortality Prediction in 30,571 Patients With COVID-19 Admitted to ICUs in Brazil

Electronic Supplementary Material

Kurtz et al.

Supplemental Methods

Data source and study population

We included all adult patients (>16 y.o.) with RT-PCR-confirmed SARS-CoV-2 infection admitted to 188 ICUs of 45 hospitals (*Rede D'Or São Luiz)* from February 26th, 2020, to April 30th, 2021. Local and National Ethics Committee (CAAE: 17079119.7.0000.5249) approved the study without the need for informed consent. Anonymized information was obtained from an electronic system, which contains prospectively collected structured data for all ICU admissions (Epimed Monitor®, Rio de Janeiro, Brazil). We used descriptive analysis to report baseline and outcome data as appropriate.

Statistical Analysis

We estimated SAPS-3 mortality probabilities using the SAPS-3 standard equation (SAPS3-SE). In addition, we obtained recalibrated probabilities for COVID-19 patients after performing a first-level customization of the SAPS-3 equation, using a logistic regression model with hospital mortality as the dependent variable and SAPS-3 score's predicted probabilities from the SAPS-3-SE as the sole predictor.

We assessed the discrimination for hospital mortality using the area under the receiver operating curve (AUROC) and Brier's Score. Calibration was evaluated using the Hosmer-Lemeshow goodness-of-fit (GOF) test and the calibration belt method[5]. We estimated 95% confidence intervals (95%CI) using 2,000 resamples.

R 4.1 was used for all analyses.

Supplementary Table 1 – Characteristics of COVID-19 ICU admissions

Characteristic	N = 30,571
Age, median (IQR)	55 (42, 69)
16-39, N (%)	5,935 (19%)
40-49	6,267 (20%)
50-59	6,077 (20%)
60-69	5,058 (17%)
70-79	3,827 (13%)
>=80	3,407 (11%)
Female, N (%)	12,301 (40.3%)
Modified Frailty Index (MFI), median (IQR)	1.00 (0.00, 2.00)
Non-frail (MFI = 0)	13,400 (44%)
Pre-frail (MFI = 1-2)	13,429 (44%)
Frail (MFI >= 3)	3,742 (12%)
Charlson Comorbidity Index, median (IQR) [n = 30,244]	0.00 (0.00, 1.00)
Any comorbidities, N (%)	20,890 (68%)
Hypertension	13,167 (43%)
Diabetes	7,368 (24%)
Cardiovascular disease	3,667 (12%)
Imunossupression	4,117 (13%)
Obesity	2,882 (9.4%)
COPD or Asthma	2,209 (7.2%)
Cerebrovascular disease	1,725 (5.6%)
Malignancy	1,707 (5.6%)
Chronic kidney disease	1,282 (4.2%)
Tobacco	894 (2.9%)
Liver Cirrhosis	157 (0.5%)
Other comorbidities	6,138 (20%)
SOFA score, median (IQR)	1 (0, 4)
SAPS-3, median (IQR)	43 (39, 51)
Initial respiratory support (First 24h)	
Oxygen	17,868 (58%)
NIRS first	9,383 (31%)
IMV first	3,320 (11%)
PaO_2/FiO_2, median (IQR) [n = 10,800]	200 (98, 317)
Normal (> 300)	2,960 (27%)
Mild (201–300)	2,405 (22%)
Moderate (101–200)	2,639 (24%)
Severe (≤ 100)	2,796 (26%)
In-hospital mortality, N (%)	4,581 (15%)
SAPS-3 predicted probability, median (IQR)	0.09 (0.06, 0.19)
SAPS-3 recalibrated probabilities, median (IQR)	0.08 (0.05, 0.18)

IQR: Interquartile range (1^{st} quartile – 3^{rd} quartile); COPD: Chronic obstructive pulmonary disease; NIRS: Noninvasive respiratory support; IMV – invasive respiratory support; SOFA: Sequential Organ Failure Assessment; SAPS-3: Simplified acute physiology score 3

Supplementary Table 2 – Goodness-of-fit, calibration and discrimination of SAPS-3 models

Hosmer-Lemeshow test

Model	C statistic	H statistic
SAPS-3 standard equation	172.8 ($p < 0.001$)	259.4 ($p < 0.001$)
SAPS-3 recalibrated	127.4 ($p < 0.001$)	245.1 ($p < 0.001$)

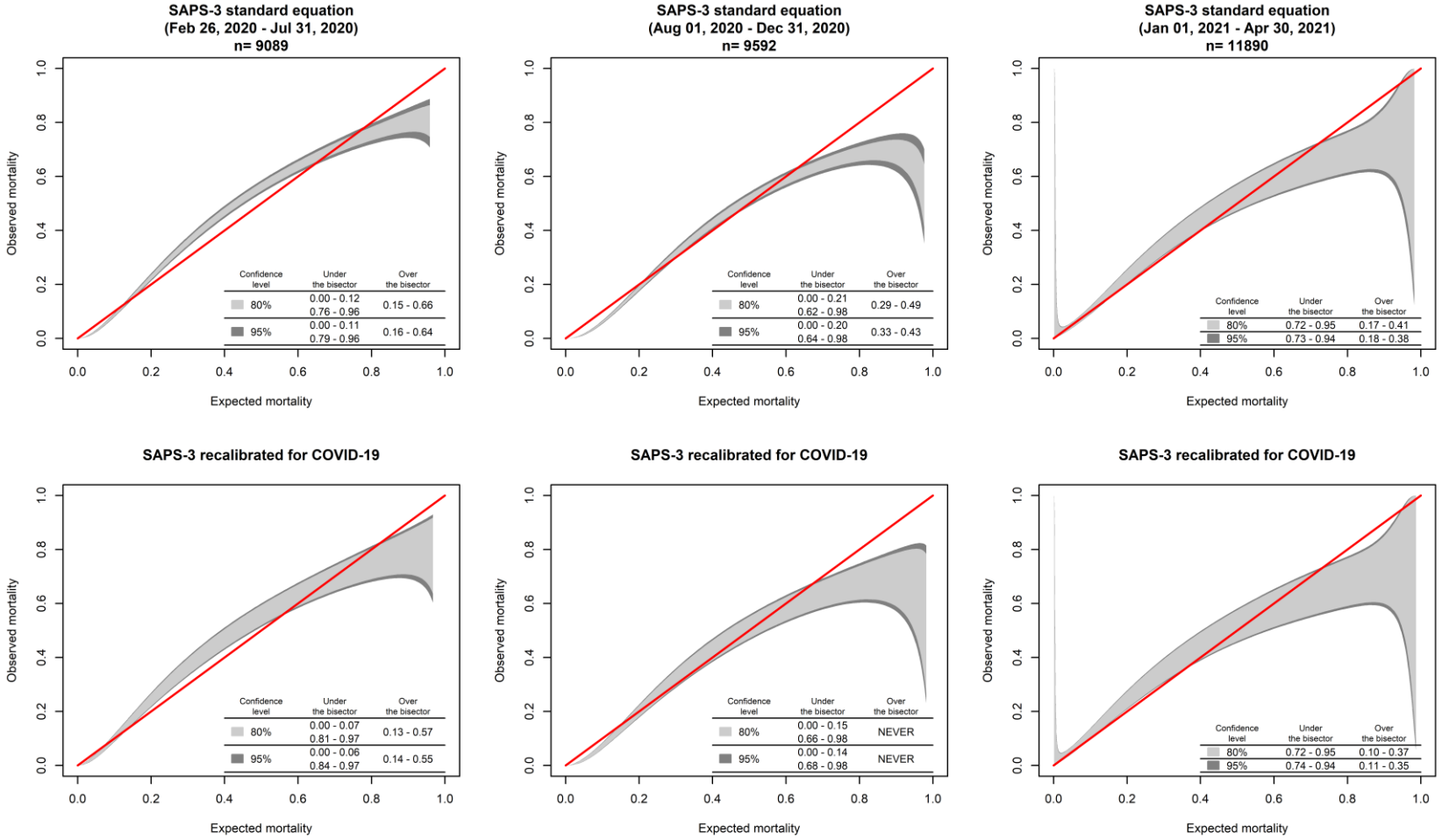
SAPS-3 discrimination and calibration

Model	AUROC	Brier's Score
SAPS-3 standard equation	0.83 (0.82-0.84)	0.097 (0.095-0.01)
SAPS-3 recalibrated	0.83 (0.82-0.84)	0.097 (0.095-0.01)

95% CI – Bootstrap (2000 resamples)

AUROC – Area Under the Receiving Operating Characteristics (ROC) curve

Supplementary Figure 1 - Calibration belts for SAPS-3 predicted probabilities (standard equation and COVID-19 customized equation) stratifying patients in three consecutive periods.



Supplementary Table 3 – Goodness-of-fit, calibration and discrimination of SAPS-3 models stratifying patients in three consecutive periods.

Metrics	Feb 26, 2020 – July 31, 2020 [n = 9,089]	Aug 01, 2020 – Dec 31, 2020 [n = 9,592]	Jan 01, 2021 – Apr 30, 2021 [n = 11,890]
SAPS-3, median [IQR]	42 [37, 52]	42 [39, 50]	43 [39, 51]
Standard equation (95% CI)			
SMR	0.99 (0.95-1.03)	0.81 (0.77-0.85)	1.04 (1.002-1.08)
AUROC	0.86 (0.85-0.87)	0.86 (0.85-0.87)	0.79 (0.78-0.81)
Brier`s Score	0.095 (0.091-0.099)	0.081 (0.078-0.085)	0.111 (0.107-0.116)
H-L test	C = 95, p<0.001 H = 100.1, p<0.001	C = 163.9, p<0.001 H = 214.2, p<0.001	C = 27.7, p<0.001 H = 80.5, p<0.001
COVID-19 Recalibrated (95% CI)			
SMR	1.03 (0.99-1.08)	0.85 (0.81-0.89)	1.08 (1.05-1.12)
AUROC	0.86 (0.85-0.871)	0.86 (0.85-0.87)	0.79 (0.78-0.81)
Brier`s Score	0.095 (0.091-0.10)	0.081 (0.077-0.085)	0.112 (0.108-0.116)
H-L test	C = 80, p<0.001 H = 85, p<0.001	C = 119.2, p<0.001 H = 189.9, p<0.001	C = 56.6, p<0.001 H = 119.1, p<0.001

H-L: Hosmer-Lemeshow; SMR: Standardized Mortality Ratio; AUROC: Area Under the Receiving Operating Characteristic curve

95% CI: 95% confidence interval (2,000 resamples)

Supplementary Table 4 – Comparisons of findings from studies assessing severity scores in patients with COVID-19

	Study cohort				
Authors	Zou et al. [4]	Stephens et al. [1]	Higgins et al. [2]	Metnitz et al. [3]	Kurtz et al.
Year published	2020	2020	2021	2021	2021
ICUs (n)	01	03	285	90	188
Country	China	United Kingdom	United States	Austria	Brazil
Data collection period	Jan 10, 2020-Feb 10, 2020	Mar 10, 2020-May 22, 2020	Mar 14, 2020-June 17, 2020	Jan 01, 2020-Jan 31, 2021	Feb 26, 2020-April 30, 2021
Patients (n)	154	242	1,491	1,464	30,571
Age (years, mean or median)	60.7	59	63.4	67.0	55
Male (%)	43.5%	71.9%	58.7%	70.0%	59.7%
Score(s) evaluated	APACHE II	APACHE II, SAPS II and ICNARC	APACHE IV	SAPS 3	SAPS 3
Predicted hospital mortality	Not reported	23.5% (APACHE II) 9.7% (SAPS II) ICNARC (not reported)	16.0%	Not reported	15.7%
Observed hospital mortality	33.8%	37.6%	24.3%	34.2%	15%
Discrimination: AUROC (95%CI)	0.966 (0.942-0.990)	Not reported	Not reported	0.74 (0.72-0.77)	0.83 (0.82-0.84)
Calibration assessment	Not reported	Not reported	Graphic representation only. No specific text was used.	Hosmer-Lemeshow GOF test*; Calibration belt	Hosmer-Lemeshow GOF test*; Calibration belt
Calibration interpretation	Not reported	Not reported	APACHE IV underestimated mortality in all studied risk groups.	Standard equation had poor calibration (GOF: P<0.001) with underestimation of mortality in low and intermediate risk groups. Customized equation led to improvements in calibration (GOF=0.89)* with good fit in all risk groups.	Standard equation had poor calibration (GOF: P<0.001) with underestimation of mortality in low and intermediate risk groups and overestimation of high-risk groups. Customized equation did not improve calibration (P<0.001)

Hospital SMR (95% CI)	Not reported	Not reported	1.52 (1.35-1.68)	Standard equation, 1.20 (1.12–1.27); customized equation, not reported	Standard equation, 0.95 (0.93-0.98); customized equation, 1.00 (0.97-1.02)

ICU: intensive care unit; AUROC: area under the receiver operating characteristic curve; SMR: standardized mortality rate; CI: confidence interval

*Only C-statistics reported; GOF: goodness of fit