

# SARS2 simplified scores to estimate risk of hospitalization and death among patients with COVID-19

Hesam Dashti<sup>1</sup>, Elise C. Roche<sup>1</sup>, David William Bates<sup>1</sup>, Samia Mora<sup>1,\*</sup>, Olga Demler<sup>1</sup>

<sup>1</sup>Center for Lipid Metabolomics, Division of Preventive Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA

\* smora@bwh.harvard.edu

## **Supplements.**

Supplementary Figure 1. SARS2 scores heatmap.

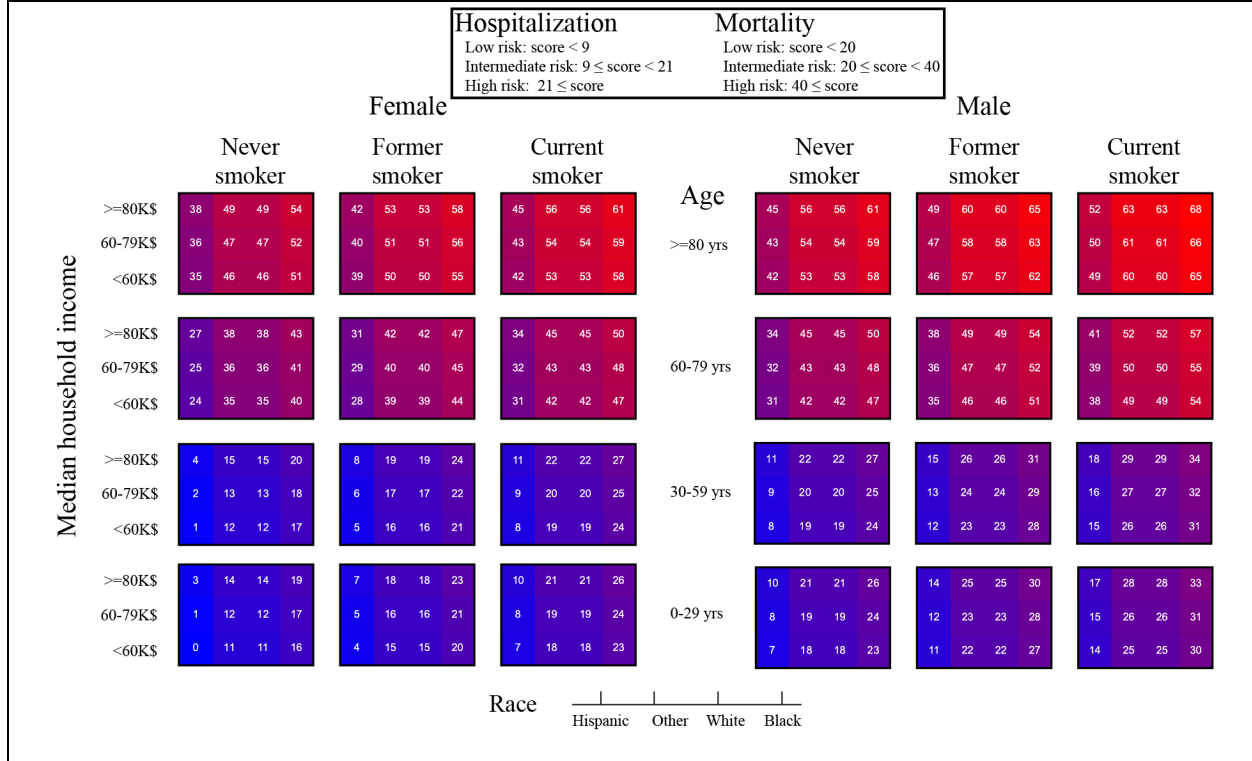
Supplementary Figure 2. Receiver operating characteristic and calibration plots.

Supplementary Table 1. Characteristics of MGB employees and non-employees.

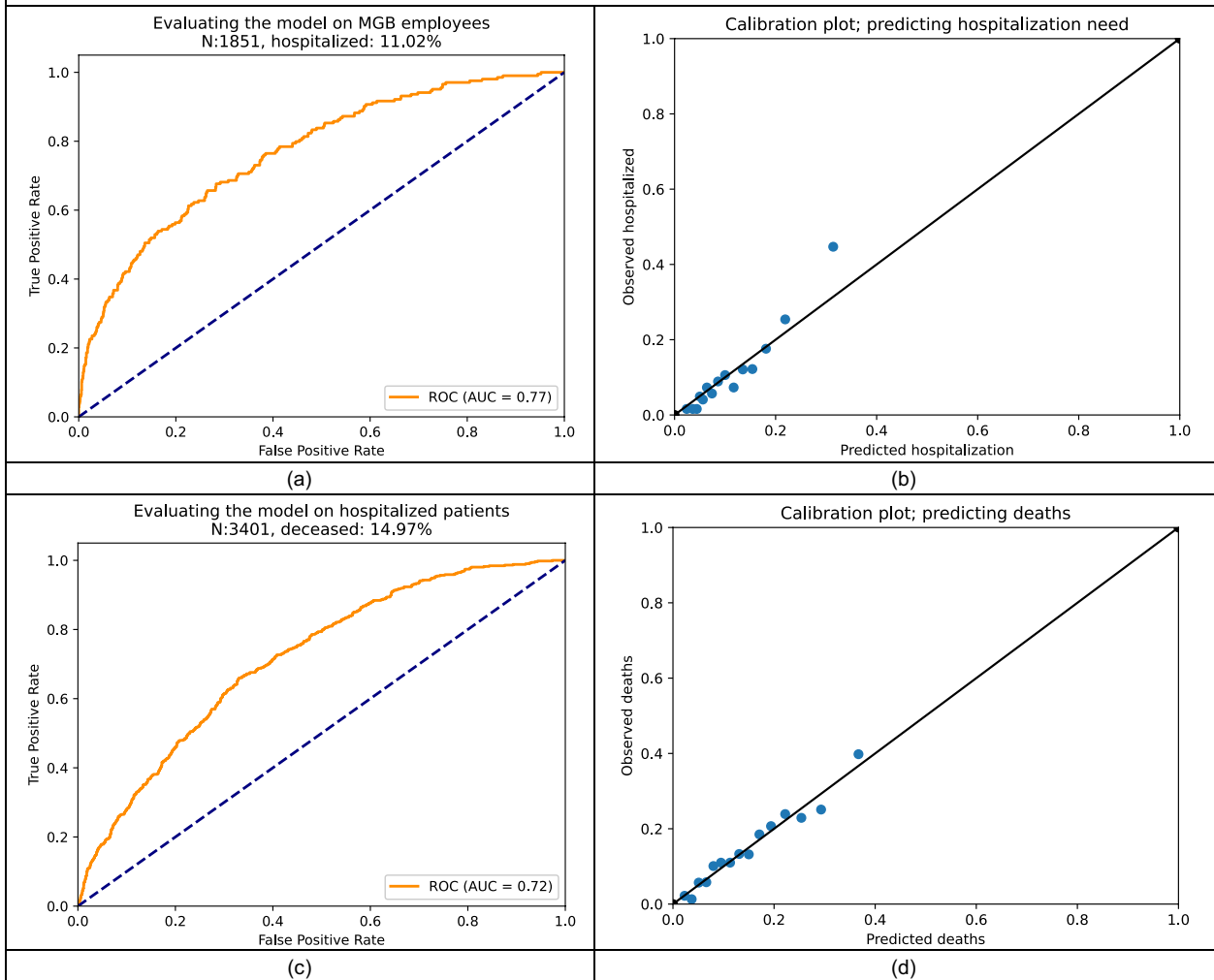
Supplementary Table 2. Odds ratios of the GLM for predicting hospitalization of the patients when age and median household income are categorized

Supplementary Table 3. Results of the sensitivity analysis on COVID-19 test date.

**Supplementary Figure 1. SARS2 scores heatmap.** Scaled risk scores are divided to female patients with lower risks (left panel) and male patients (right panel). The columns in each panel represent different smoking status. Every row indicates an age category. Each matrix shows the three median household incomes in rows, and columns of these matrices indicate four race groups. The cells are colored from light-blue (lowest score: 0) to dark red (highest score: 68). The top panel indicates score ranges corresponding to low, intermediate, and high-risk groups for hospitalizations and mortality among hospitalized patients.



**Supplementary Figure 2.** (a) Receiver operating characteristic (ROC) curve of the model for predicting hospitalization when the model applied on the validation cohort of MGB employees. (b) The Hosmer-Lemeshow GOF plot for the model predicting hospitalization on the validation population (number of bins: 15, p-value: 0.11). The higher rate of observed hospitalizations compared to the mean predicted risks could be due to the use of MGB employees as the validation group which have a better access to medical facilities. (c) ROC of the model for predicting mortality among hospitalized patients. (d) The Hosmer-Lemeshow GOF plot for the model predicting mortality on the hospitalized patients (number of bins: 15, p-value: 0.62).



**Supplementary Table 1. Characteristics of MGB employees and non-employees.**

<b>Characteristics</b>	<b>Non-employees N=10496</b>	<b>Employees N=1851</b>
Age (years), Median (IQR)	48.0 (32.0-64.0)	41.0 (30.0-54.0)
Women, N (%)	5251 (50.03)	1334 (71.88)
Race, N (%)		
White	4444 (42.34)	837 (45.10)
Black	1472 (14.02)	492 (26.51)
Hispanic	974 (9.28)	143 (7.70)
Other/Not recorded	3606 (34.36)	379 (20.42)
Median household income (\$1000), Median (IQR)	60.4 (53.3-86.2)	70.2 (55.0-96.8)
Smoking, N (%)		
Current	489 (4.66)	63 (3.39)
Former	1772 (16.88)	253 (13.63)
Never	5715 (54.44)	1223 (65.89)
Unknown	2520 (24.01)	312 (16.81)

IQR: interquartile range

**Supplementary Table 2.** Odds ratios of the GLM for predicting hospitalization of the patients when age and median household income are categorized (N=10,496).

Variables	Median (IQR), N (%)	OR (95% CI)	P
Age (year)			<0.001*
0-29	2149 (20.47)	0.39 (0.33-0.45)	<.001
30-59	5040 (48.02)	ref	ref
60-79	2245 (21.39)	3.13 (2.80-3.50)	<.001
≥80	1062 (10.12)	5.61 (4.83-6.54)	<.001
Sex			
Female	5251 (50.03)	ref	ref
Male	5245 (49.97)	1.39 (1.27-1.53)	<.001
Race			<0.001*
White	4444 (42.34)	ref	ref
Black	1472 (14.02)	1.30 (1.13-1.49)	<.001
Hispanic	974 (9.28)	0.58 (0.48-0.69)	<.001
Other/Not recorded	3606 (34.36)	1.00 (0.89-1.12)	0.95
Median household income (\$1000)			<0.001*
<60	4412 (42.04)	ref	ref
60-79	2936 (27.97)	0.94 (0.83-1.05)	0.25
≥80	3148 (29.99)	0.87 (0.78-0.98)	0.02
Smoking			<.001*
Current	489 (4.66)	1.42 (1.16-1.74)	<.001
Former	1772 (16.88)	1.25 (1.10-1.41)	<.001
Never	5715 (54.45)	ref	ref
Unknown	2520 (24.01)	0.52 (0.46-0.60)	<.001

Each variable shown was mutually adjusted for the other variables in the table. IQR: interquartile range. Medians, interquartile ratios, and percentages are reported on the derivation population. \* Test of trend p-value.

Supplementary Table 3. Results of the sensitivity analysis on COVID-19 test date.									
Variables	COVID-19 tested before April 29, 2020 (MGB non-employees N=6,624)			COVID-19 tested after April 29, 2020 (MGB non-employees N=3,872)			Main model (MGB non-employees N=10,496)		
	Median (IQR), N (%)	OR (95% CI)	P	Median (IQR), N (%)	OR (95% CI)	P	Median (IQR), N (%)	OR (95% CI)	P
Age (years)	51.0 (35.0-65.0)	1.58 (1.53-1.63)	<.001	43.0 (28.0-61.0)	1.44 (1.38-1.50)	<.001	48.0 (32.0-64.0)	1.53 (1.49-1.57)	<.001
Sex									
Female	3285 (49.6)	ref	ref	1966 (50.8)	ref	ref	5251 (50.03)	ref	ref
Male	3339 (50.4)	1.46 (1.31-1.64)	<.001	1906 (49.2)	1.30 (1.10-1.52)	0.002	5245 (49.97)	1.40 (1.28-1.54)	<.001
Race			<.001*			<.001*			<.001*
White	2741 (41.4)	ref	Ref	1703 (44.0)	ref	ref	4444 (42.3)	ref	ref
Black	957 (14.4)	1.21 (1.02-1.43)	0.03	515 (13.3)	1.48 (1.17-1.87)	<.001	1472 (14.0)	1.30 (1.13-1.49)	<.001
Hispanic	627 (9.5)	0.64 (0.51-0.79)	<.001	347 (9.0)	0.40 (0.27-0.58)	<.001	974 (9.3)	0.58 (0.48-0.70)	<.001
Other/Not recorded	2299 (34.7)	1.03 (0.90-1.19)	0.66	1307 (33.8)	0.94 (0.76-1.15)	0.54	3606 (34.4)	1.02 (0.91-1.15)	0.74
Median household income (\$1000)	60.4 (53.3-88.7)	0.95 (0.93-0.97)	<.001	60.4 (53.3-85.3)	1.04 (1.01-1.07)	0.005	60.4 (53.3-86.2)	0.98 (0.96-0.99)	0.007
Smoking			<.001*			<.001*			<.001*
Current	296 (4.5)	1.40 (1.08-1.81)	0.01	193 (5.0)	1.60 (1.15-2.22)	0.005	489 (4.7)	1.44 (1.17-1.76)	<.001
Former	1213 (18.3)	1.15 (0.99-1.33)	0.08	559 (14.4)	1.33 (1.07-1.65)	0.009	1772 (16.9)	1.22 (1.08-1.38)	0.002
Never	3666 (55.3)	ref	ref	2049 (52.9)	ref	ref	5715 (54.4)	ref	ref
Unknown	1449	0.51 (0.44-0.60)	<.001	1071 (27.7)	0.58 (0.47-0.73)	<.001	2520 (24.0)	0.53 (0.46-0.60)	<.001

Each variable shown was mutually adjusted for the other variables in the table. IQR: interquartile range. Medians, interquartile ratios, and percentages are reported on the derivation population. Odd ratios (OR) and the corresponding 95% confidence interval (CI) for age is reported per 10 years increment, and these values for median household income are shown per 10,000\$. \* Test of trend p-value.