

Supplementary Information for:

Analysis of the Mount Sinai Health System electronic medical records reveals racial disparities and identifies prognostic factors in the COVID-19 hospitalized patient population

Supplementary tables

Table S1. Estimated coefficients from age-adjusted logistic regression model showing the racial disparities in the infection rates.

	ESTIMATED COEFFICIENT	STD ERR	Z	P> Z	2.5% CI	97.5% CI
(INTERCEPT)	-7.818	0.046	-171.245	0	-7.908	-7.729
AGE	0.0348	0.001	53.567	0	0.034	0.036
AFRICAN- AMERICAN/BLACK	0.7264	0.034	21.292	0	0.659	0.793
ASIAN	0.0788	0.061	1.303	0.193	-0.04	0.197
HISPANIC/LATINO	1.2644	0.033	38.593	0	1.2	1.329
OTHER	0.5699	0.04	14.382	0	0.492	0.648

Table S2. Descriptive statistics of the age and baseline O2 saturation at admission of hospitalized COVID-19 patients by hospitals.

FACILITY	AGE MEAN	AGE STD	O2 SAT BASELINE MEAN	O2 SAT BASELINE STD
MOUNT SINAI BI BROOKLYN	69.73	13.82	94.41	7.78
MOUNT SINAI QUEENS HOSPITAL	65.16	15.39	91.26	8.89
MOUNT SINAI ST. LUKE'S	66.58	16.41	93.57	6.94
MOUNT SINAI WEST	62.07	18.68	93.59	5.87
THE MOUNT SINAI HOSPITAL	58.88	17.77	94.20	5.80

Table S3. Summary of models and metrics used across this study.

MODEL	LOCATION	DEGREE OF FREEDOM	AUROC	CONCORDANCE
MULTIVARIATE LOGISTIC REGRESSION	Fig 2	31	0.934[se = 0.011]	NA
ETIOLOGIC COMPETING RISKS FOR DEATH	Fig 4A	29	NA	0.788[se = 0.01]
ETIOLOGIC COMPETING RISKS FOR CURED	Fig 4B	29	NA	0.791[se = 0.006]
PROGNOSTIC COMPETING RISKS FOR DEATH	Fig 5A	26	NA	0.768[se = 0.011]
PROGNOSTIC COMPETING RISKS FOR CURED	Fig 5B	26	NA	0.685[se = 0.009]

Table S4. Generalized variance-inflation factors (GVIFs) for covariates used in the multivariate etiologic models.

COVARIATES	GVIF	DEGREE OF FREEDOM
DURATION (DAYS)	1.22	1
AGE	1.51	1
SEX	1.20	1
RACE	1.23	5
SMOKING_STATUS	1.31	3
BMI	1.27	1
TEMPERATURE	1.32	1
TEMP_MAX	1.48	1
O2_SAT	1.16	1
O2SAT_MIN	1.24	1
HEART_RATE	1.19	1
RESPIRATORY RATE > 25	1.08	1
HIGH BP	1.04	1
ASTHMA	1.06	1
COPD	1.09	1
HTN	1.42	1
OBESITY	1.24	1
DIABETES	1.32	1
HIV	1.04	1
CANCER	1.08	1

IN_ICU	1.19	1
WBC	1.11	1
CREATININE > 1.2	1.17	1
ALT	1.07	1

Table S5. Generalized variance-inflation factors (GVIFs) for covariates used in the multivariate prognostic models.

COVARIATES	GVIF	DEGREE OF FREEDOM
DURATION (DAYS)	1.05	1
AGE	1.40	1
SEX	1.17	1
RACE	1.20	5
SMOKING_STATUS	1.28	3
BMI	1.31	1
TEMPERATURE	1.10	1
O2_SAT	1.08	1
HEART_RATE	1.20	1
RESPIRATORY RATE > 25	1.10	1
HIGH BP	1.05	1
ASTHMA	1.08	1
COPD	1.08	1
HTN	1.47	1
OBESITY	1.29	1
DIABETES	1.34	1
HIV	1.04	1
CANCER	1.08	1
WBC	1.09	1
CREATININE > 1.2	1.15	1
ALT	1.06	1

Table S6. p-values and odd ratios (OR) for less common baseline laboratory tests in multivariate association analyses between COVID-19 mortality and discharge

LABORATORY TESTS	P-VALUE	LOG ODDS	ODDS RATIO	2.5% OR	97.5% OR
LDH	7.11E-32	0.0030	1.0030	1.0025	1.0035
CRP	7.15E-19	0.0080	1.0080	1.0063	1.0098
D-DIMER	1.05E-15	0.1524	1.1646	1.1228	1.2097
BUN	1.14E-15	0.0235	1.0238	1.0181	1.0299

PROCALCITONIN	1.08E-10	0.1153	1.1222	1.0858	1.1652
FERRITIN	2.07E-09	0.0001	1.0001	1.0001	1.0002
FIBRINOGEN	2.89E-03	0.0012	1.0012	1.0004	1.0020
HGB	4.08E-02	-0.0683	0.9340	0.8748	0.9973
EGFR NON-AFRICAN AM	1.07E-01	-0.0080	0.9920	0.9823	1.0017
EGFR AFRICAN AM	5.21E-01	0.0035	1.0035	0.9927	1.0145

Supplementary figure legends

Figure S1. Age distributions across racial groups of patients from the MSHS EMR (A) and Covid-19 patients (B).

Figure S2. Scatter plots showing various vital signs with respect to age of the hospitalized Covid-19 cohort.

Figure S3. Scatter plots showing related vital signs against each other.

Figure S4. Overall CIFs for the hospitalized Covid-19 cohorts for deceased and discharged events, respectively.

Figure S5. Coefficients from logistic regression models analyzing covariates associated with final outcomes (cured versus deceased) for COVID-19 patients with and without ICU stays (A, B). The estimated coefficients from the logistic regression model, also known as log odds, are plotted for the covariates. An intercept term was included in the model but excluded from the plot. Error bars indicate 95% CI.