ORIGINAL ARTICLE

Hospital-Level Variation in Death for Critically Ill Patients with COVID-19

Matthew M. Churpek[1], Shruti Gupta[2], Alexandra B. Spicer[1], William F. Parker[3], John Fahrenbach[3], Samantha K. Brenner[4,5], and David E. Leaf[2]; for the STOP-COVID Investigators

[1]Division of Allergy, Pulmonary and Critical Care Medicine, School of Medicine and Public Health, University of Wisconsin, Madison, Wisconsin; [2]Division of Renal Medicine, Brigham and Women's Hospital, Boston, Massachusetts; [3]Section of Pulmonary and Critical Care, Department of Medicine, University of Chicago, Chicago, Illinois; [4]Department of Internal Medicine, Hackensack Meridian School of Medicine, Seton Hall, New Jersey; and [5]Heart and Vascular Hospital, Hackensack University Medical Center, Hackensack Meridian Health, Hackensack, New Jersey

ORCID ID: 0000-0002-5747-2151 (S.G.).

Abstract

Rationale: Variation in hospital mortality has been described for coronavirus disease (COVID-19), but the factors that explain these differences remain unclear.

Objective: Our objective was to use a large, nationally representative data set of critically ill adults with COVID-19 to determine which factors explain mortality variability.

Methods: In this multicenter cohort study, we examined adults hospitalized in ICUs with COVID-19 at 70 U.S. hospitals between March and June 2020. The primary outcome was 28-day mortality. We examined patient-level and hospital-level variables. Mixed-effect logistic regression was used to identify factors associated with interhospital variation. The median odds ratio was calculated to compare outcomes in higher- versus lower-mortality hospitals. A gradient-boosted machine algorithm was developed for individual-level mortality models.

Measurements and Main Results: A total of 4,019 patients were included, 1,537 (38%) of whom died by 28 days. Mortality varied considerably across hospitals (0–82%). After adjustment for patient- and hospital-level domains, interhospital variation was attenuated (odds ratio decline from 2.06 [95% confidence interval (CI), 1.73–2.37] to 1.22 [95% CI, 1.00–1.38]), with the greatest changes occurring with adjustment for acute physiology, socioeconomic status, and strain. For individual patients, the relative contribution of each domain to mortality risk was as follows: acute physiology (49%), demographics and comorbidities (20%), socioeconomic status (12%), strain (9%), hospital quality (8%), and treatments (3%).

Conclusions: There is considerable interhospital variation in mortality for critically ill patients with COVID-19, which is mostly explained by hospital-level socioeconomic status, strain, and acute physiologic differences. Individual mortality is driven mostly by patient-level factors.

Keywords: COVID-19; critical care; ICU; health disparities

A complete list of STOP-COVID Investigators may be found in the online supplement.

At a Glance Commentary

Scientific Knowledge on the Subject: Considerable variation in hospital mortality has been described for patients admitted to the ICU with coronavirus disease (COVID-19). However, the factors that explain these differences remain unclear.

What This Study Adds to the Field: In this study of 4,019 patients in 70 hospitals, we found significant interhospital variation in mortality for critically ill patients with COVID-19. This hospital-level variation was mostly explained by hospital-level socioeconomic status, strain, and physiologic differences, although individual mortality was driven mostly by patient-level factors.

As of April 2021, coronavirus disease (COVID-19) has killed more than 500,000 people in the United States (1). When patients develop severe disease, they are typically transferred to the ICU, which provides more intensive monitoring together with potentially life-saving critical care therapies such as mechanical ventilation, vasoactive agents, and extracorporeal membrane oxygenation (2, 3). Studies conducted before the pandemic demonstrated that outcomes of critically ill patients vary across hospitals, which may relate to differences in patient characteristics and the quality of care provided at different hospitals (4). Emerging data suggest similar variability in outcomes across hospitals for critically ill patients admitted with COVID-19 (5–8). The causes of this variability are unclear and could include differences in demographics, comorbidities, the physiologic severity of illness, socioeconomic status, resource strain, hospital quality, and treatments provided. It is also unknown how each of these domains impacts mortality risk for individual patients. A better understanding of the patient- and hospital-level factors impacting death could lead to insights into the reasons for the wide variation in reported outcomes, the determinants of individual patient outcomes, and improved healthcare delivery.

Our objective was to use a large, nationally representative data set of critically ill adults with COVID-19 to determine which factors explain the variability in mortality at both the hospital and the patient level. To do this, we linked detailed patient information with hospital-level data and then explored how different domains explained variations in 28-day mortality.

Methods

Study Design, Setting, and Population

We used the database of the multicenter STOP-COVID (Study of the Treatment and Outcomes in Critically Ill Patients with COVID-19), a cohort study of 5,154 patients with COVID-19 admitted to ICUs across the United States (*see* Table E1 in the online supplement for the sites included in this study) (5). We included consecutive adults (age ⩾ 18 yr) admitted to the ICU with laboratory-confirmed COVID-19 admitted between March 4 and June 29, 2020. Patients were followed until the first of hospital discharge, death, or at least 28 days after ICU admission. Patients transferred to the ICU from other hospitals, admitted to a hospital not linked to the Medicare Hospital Compare ratings, or admitted to a hospital with fewer than 10 COVID-19 ICU admissions in the data set were excluded. A sensitivity analysis was performed by including patients transferred from outside hospitals. The study was approved by institutional review boards at each site, with a waiver of informed consent being given.

Data Collection and Outcome

Manual chart review was performed at each site by using a standardized case report form, as previously described (5). Patient-level data collected included admission day, demographic information, comorbidities, vital signs at ICU admission, laboratory values, medications, nonmedication treatments, organ support in the first 2 weeks of ICU admission, and outcomes, including in-hospital mortality. The STOP-COVID data set also included what type of ICU bed the patient was admitted to (e.g., medical–surgical), whether the patient was admitted to a COVID-19–specific ICU or surge unit, and the number of ICU beds at each hospital before the COVID-19 pandemic.

Additional hospital-level variables were collected by linking each study hospital to data from the following sources: the American Hospital Association Annual Survey 2020 database for hospital strain and capacity variables, the 2017 Medicare Hospital Compare ratings for hospital quality ratings, the Healthcare Cost Report Information System, and the 2015 American Community Survey socioeconomic status data, which incorporates information from communities surrounding each hospital by using a previously described methodology (Table E2) (9–11). Furthermore, time-varying variables describing hospital-level strain were collected from the STOP-COVID data set (i.e., number of other patients with COVID-19 currently in the ICU at a given hospital when a patient was admitted) and from publicly available data on the number of new COVID-19 cases from the past 30 days for the county where each hospital was located (1).

The primary outcome of the study was in-hospital death within 28 days of ICU admission. If a patient was discharged alive before Day 28, they were assumed to be alive at Day 28. This assumption was confirmed in a sample of patients in a previous study (5). A sensitivity analysis was performed using in-hospital mortality as the outcome.

Statistical Analysis

Explanatory variables were categorized into six domains, including three patient-level domains and three hospital-level domains. The individual variables and domains were chosen *a priori* on the basis of prior literature and availability. Patient-level domains included acute physiology and severity of illness in the first 48 hours of ICU admission (e.g., vital signs, laboratory values, ventilatory support, number of vasopressors, and renal replacement therapy); demographics and comorbidities (e.g., age, sex, race, body mass index, smoking status, and preexisting conditions); and treatments provided in the first 48 hours of ICU admission (e.g., corticosteroids, remdesivir, tocilizumab, prone position ventilation). Hospital-level treatment intensity was also included as a variable in the treatment domain by calculating the percentage of mechanically ventilated patients with a Pa_{O_2}/Fi_{O_2} ratio <150 who were treated with extracorporeal membrane oxygenation, inhaled pulmonary vasodilators, tocilizumab, prone positioning, or neuromuscular blockade. Hospital-level domains included socioeconomic factors at the hospital level (e.g., percentage with a high school diploma, percentage unemployed, percentage who were English speaking, percentage with a travel time to work >45 min), hospital strain (e.g., number of ICU beds before COVID-19, time-varying number of

ICU beds filled with patients with COVID-19, whether the patient was admitted to a COVID-19–specific ICU or surge unit, total number of medical–surgical beds, prepandemic ICU occupancy rate, number of hospital beds in the county, number of COVID-19 cases in the county from the prior 30 d), and hospital quality scores (mortality, readmission, safety, timeliness, patient experience, and effectiveness). The ICU admission day was used to create a variable that denotes the "days since study start" that a patient was admitted to the ICU, which was assigned to each patient to account for possible longitudinal changes in hospital quality (12). The full variable list for each domain, together with additional descriptions, is provided in Table E2. Missing values were imputed by using bagged forests from the caret package in R, which builds ensembles of decision trees, with each tree being fit to a randomly selected, bootstrapped sample of the data set by using nonmissing variables to impute missing variables (see Table 1 for the amount of missing data for each variable). This approach has the advantage of automatically modeling nonlinearities and interactions that may be important for accurate variable imputation (13). Comparisons between patients who survived and those who died within 28 days were made for all study variables by using Wilcoxon rank sum tests and chi-square tests.

Next, mixed-effect logistic regression models were fit, first with an empty model with a random effect for each hospital and then by sequentially adjusting for variables from each domain in the order described above, which moves from patient-level to hospital-level factors. This ordering allowed for the separation of patient- and hospital-level variables to determine their contributions to interhospital variation in mortality. The change in the adjusted variation of 28-day mortality was calculated, moving from one model to the next, by examining the median odds ratio for each model. The median odds ratio can be interpreted as the difference in odds between a randomly selected lower-risk hospital and a randomly selected higher-risk hospital. It can be conceptualized as the increased risk that a subject would have if he or she were admitted to a higher-risk hospital (14, 15). Pseudo-R^2 values were also calculated for each individual domain by using Efron's R^2, which is calculated by taking the sum of the squared model residuals divided by the total variability in the dependent variable.

Finally, to calculate the contribution of the domains to an individual's risk of mortality, a gradient-boosted tree machine learning model was fit by using all of the variables from each domain (16). Tenfold cross-validation was used to optimize the model's area under the receiver operating characteristic curve. Shapley values, which estimate the contribution of each variable for that individual patient's risk of 28-day mortality (17), were then calculated for each individual patient. The individual Shapley values were then combined across all patients in the data set by using the mean of their absolute value to determine the percent mortality risk explained by each domain. All analyses were performed by using Stata version 16.1 (StataCorp) and R version 4.2 (R Foundation for Statistical Computing) with the caret, XGBoost, and iml packages. A two-sided P value of <0.05 denoted statistical significance.

Results

Patient Characteristics

A total of 4,019 patients (median age [interquartile range (IQR)], 63 [53–72]; 63% male [$n = 2,532$]) from 70 hospitals were included in the analysis after exclusion criteria were applied (Figure E1 and Table E1), and 1,537 patients (38%) died by 28 days. The median number of patients at a given hospital was 34 (IQR, 20–79; Figure E2). Patients who died were older (median [IQR], 68 [59–76] yr vs. 60 [49–68] yr), more likely to be male (66% vs. 61%), and more likely to be current or former smokers (30% vs. 23%) and had higher frequencies of most comorbidities than those who survived at 28 days (Table 1). Most vital signs and laboratory results were significantly different during the first 48 hours of ICU admission between those who died and those who survived (Table 1). Patients who died were also more likely to have received invasive mechanical ventilation (80% vs. 58%) and renal replacement therapy (9% vs. 5%) during the first 48 hours of ICU admission. Finally, certain medications were more often provided to those who died, such as neuromuscular blocking agents (25% vs. 17%), hydroxychloroquine (63% vs. 59%), and corticosteroids (35% vs. 21%) (Table 1).

Hospital-Level Analysis

Compared with patients who survived, patients who died were admitted to hospitals with a higher percentage of ICU beds occupied by patients with COVID-19 (48% vs. 31%), a higher percentage of the population traveling >45 minutes to work (23% vs. 18%), a lower prepandemic ICU occupancy rate (69% vs. 74%), a lower number of pre-COVID-19 ICU beds (median [IQR], 53 [47–98] vs. 98 [54–115]), a higher number of COVID-19 cases in the county in the prior 30 days (median [IQR], 2,279 [640–7,268] vs. 1,416 [398–4,585]), and a lower hospital quality score (Table 2).

Twenty-eight–day mortality varied widely across hospitals, from 0% at the lowest-risk hospital to 82% at the highest-risk hospital. In the mixed-effect regression model, the median odds ratio decreased from 2.06 (95% confidence interval [CI], 1.73–2.37) in the unadjusted model to 1.22 (95% CI, 1.00–1.38) in the fully adjusted model (Figure 1). This was associated with a change in the range of mortality across hospitals from 12–91% (random effects only) to 32–44% (fully adjusted model). Model adjustment with variables from the physiology, socioeconomic status, and strain domains were associated with the greatest change in the median odds ratio (all with a >0.20 change in the point estimate). The fully adjusted model explained nearly all the variability across hospitals (P value for random effect term = 0.73; see Table E3 for model coefficients). Pseudo-R^2 values for each individual domain demonstrated similar results, with physiology (0.2), demographics (0.11), socioeconomic status (0.10), and strain (0.09) having the highest values, followed by quality (0.06) and treatments (0.04).

Patient-Level Analysis

The Shapley values calculated from the XGBoost model using variables from all the domains found that physiology (49%), demographics and comorbidities (20%), hospital socioeconomic status (12%), strain (9%), hospital quality (8%), and treatments (3%) all contributed to mortality risk (Figure 2). The mean contributions of the individual variables in each domain are shown in Figures E3–E8. Thus, for patients in the data set, on average, their presenting physiology explained half of their quantifiable individual risk of mortality, whereas external factors such as hospital socioeconomic status, hospital capacity and strain, hospital quality, and the treatments clinicians provided explained over one-quarter (31%) of their mortality risk. Among patient demographics, age had the highest contribution, explaining 12% of the mortality risk, whereas comorbidities explained 4% of a patient's mortality risk. Temporal trends captured by the days since study start variable only explained a small percentage of a patient's mortality risk (1%).

Table 1. Patient Characteristics at Baseline

Variable	All Patients (N = 4,019)	28-Day Survivors (n = 2,482)	28-Day Nonsurvivors (n = 1,537)
Demographics and preexisting comorbidities			
Demographics			
Age, yr, median (IQR)	63 (53–72)*	60 (49–68)	68 (59–76)
Sex, M, n (%)	2,532 (63.0)*	1,520 (61.2)	1,012 (65.8)
Race, n (%)			
White	1,527 (38.0%)	950 (38.3%)	577 (37.5)
Black	1,238 (30.8%)	782 (31.5%)	456 (29.7%)
Other	328 (8.2%)	213 (8.6%)	115 (7.5%)
Unknown/not reported	926 (23.0%)	537 (21.6%)	389 (25.3%)
Ethnicity, n (%)			
Hispanic	954 (23.7%)	601 (24.2%)	353 (23.0%)
Non-Hispanic	2,600 (64.7%)	1,604 (64.6%)	996 (64.8%)
Unknown/not reported	465 (11.6%)	277 (11.2%)	188 (12.2%)
Current or former smoker, n (%)	1,039 (25.9)*	581 (23.4)	458 (29.8)
BMI, kg/m^2, median (IQR)	30.2 (26.3–35.5)*	30.6 (26.5–35.9)	29.7 (26.0–34.9)
Preexisting comorbidities, n (%)			
Active cancer	190 (4.7)*	80 (3.2)	110 (7.2)
Congestive heart failure	425 (10.6)*	224 (9.0)	201 (13.1)
Chronic obstructive pulmonary disease	356 (8.9)*	175 (7.1)	181 (11.8)
Coronary artery disease	567 (14.1)*	277 (11.2)	290 (18.9)
Diabetes	1,713 (42.6)*	972 (39.2)	741 (48.2)
End-stage renal disease	153 (3.8)*	79 (3.2)	74 (4.8)
Hypertension	2,476 (61.6)*	1,398 (56.3)	1,078 (70.1)
Physiology			
Vital signs[†]			
Altered mental status, n (%)	997 (24.8)*	420 (16.9)	577 (37.5)
Heart rate, beats/min, median (IQR)	105 (91–120)*	103 (90–118)	109 (93–125)
Respiratory rate, breaths/min, median (IQR)	32 (26–39)*	32 (26–39)	31 (26–38)
Systolic blood pressure, mm Hg, median (IQR)	97 (85–111)*	99 (88–112)	94 (82–109)
Temperature, °C, median (IQR)	37.9 (37.2–38.8)*	38.0 (37.2–38.8)	37.8 (37.1–38.7)
Laboratory results[‡]			
Arterial pH, median (IQR)	7.3 (7.3–7.4)*	7.4 (7.3–7.4)	7.3 (7.2–7.4)
Aspartate aminotransferase, U/L, median (IQR)	60 (39–86)*	56 (37–79)	67 (42–105)
Creatinine, mg/dl, median (IQR)	1.2 (0.9–2.1)*	1.1 (0.8–1.6)	1.6 (1.1–2.9)
C-reactive protein, mg/L, median (IQR)	173 (115–238)*	168 (108–229)	185 (127–250)
D-dimer, ng/ml, median (IQR)	2,340 (1,015–6,135)*	2,024 (825–4,340)	3,841 (1,593–9,305)
Ferritin, ng/ml, median (IQR)	1,291 (661–2,214)*	1,177 (622–1,933)	1,588 (776–2,682)
High troponin indicator,[§] n (%)	1,769 (44.0)*	820 (33.0)	949 (61.7)
Lactate, mmol/L, median (IQR)	1.7 (1.3–2.4)*	1.6 (1.3–2.1)	2.0 (1.4–2.9)
Lymphocytes, %, median (IQR)	8.9 (5.4–13.3)*	10.0 (6.4–14.6)	7.1 (4.0–11.1)
Procalcitonin,[†] ng/ml, median (IQR) (1)	1.3 (0.2–4.6)*	0.8 (0.2–2.3)	1.4 (0.5–8.0)
Sodium,[†] mEq/L, median (IQR) (1)	137 (134–140)*	136 (134–139)	137 (134–141)
Urine output, ml, median (IQR)	716 (436–1,000)*	792 (550–1,050)	579 (300–875)
WBC count per mm^3, median (IQR)	9.6 (6.8–13.3)*	9.0 (6.5–12.1)	10.6 (7.6–15.1)
Severity of Illness[‡]			
P/F ratio,[‖] mm Hg, median (IQR)	131 (102–158)*	135 (112–159)	123 (86–155)
Invasive mechanical ventilation, n (%)	2,681 (66.7)*	1,449 (58.4)	1,232 (80.2)
Renal replacement therapy,[¶] n (%)	258 (6.4)*	118 (4.8)	140 (9.1)
Vasopressors,[**] n (%)			
One	1,367 (34.0)*	790 (31.8)	577 (37.5)
Two or more	648 (16.1)*	282 (11.4)	366 (23.8)
Treatments,[‡] n (%)			
Aspirin	696 (17.3)*	389 (15.7)	307 (20.0)
Azithromycin	2,003 (49.8)*	1,270 (51.2)	733 (47.7)
Hydroxychloroquine	2,423 (60.3)*	1,457 (58.7)	966 (62.8)
Neuromuscular blockade	812 (20.2)*	428 (17.2)	384 (25.0)
Prone positioning	1,087 (27.0)	663 (26.7)	424 (27.6)

(Continued)

Table 1. (*Continued*)

Variable	All Patients (*N* = 4,019)	28-Day Survivors (*n* = 2,482)	28-Day Nonsurvivors (*n* = 1,537)
Remdesivir	238 (5.9)*	168 (6.8)	70 (4.6)
Statin	913 (22.7)	553 (22.3)	360 (23.4)
Corticosteroid	1,057 (26.3)*	522 (21.0)	535 (34.8)
Tocilizumab	497 (12.4)*	331 (13.3)	166 (10.8)
Vitamin C	281 (7.0)*	148 (6.0)	133 (8.7)

Definition of abbreviations: BMI = body mass index; IQR = interquartile range; PEEP = positive end-expiratory pressure; P/F = Pa$_{O_2}$/Fi$_{O_2}$; WBC = white blood cell.

Data regarding troponin were missing for 2,544 (63%), data regarding P/F were missing for 1,576 (39%), data regarding PEEP Day 1 were missing for 1,513 (38%), data regarding procalcitonin were missing for 1,455 (36%), data regarding D-dimer were missing for 1,233 (31%), data regarding urine output were missing for 1,210 (30%), data regarding lactate were missing for 1,198 (30%), data regarding ferritin were missing for 1,054 (26%), data regarding CRP were missing for 926 (23%), data regarding arterial pH were missing for 902 (22%), data regarding smoking status were missing for 745 (19%), data regarding lymphocytes were missing for 36 (11%), data regarding aspartate aminotransferase were missing for 353 (9%), data regarding mental status were missing for 220 (5%), data regarding PEEP Day 2 were missing for 163 (4%), data for BMI were missing for 152 (4%), data regarding WBC counts were missing for 77 (2%), data regarding creatinine were missing for 70 (2%), and data regarding sodium were missing for 24 (<1%). Missing data were imputed by using bagImpute and are included in the table.
*P value of <0.05 for difference between survivors and nonsurvivors (Wilcoxon rank sum test for continuous variables and chi-square test for categorical variables).
†Collected on ICU admission.
‡Worst value or if occurred anytime during Days 1–2 in the ICU.
§Troponin T or I value greater than the 99th percentile upper reference limit of normal for that laboratory test.
‖Refers to the P/F ratio and was only recorded in patients receiving invasive mechanical ventilation. Other values were imputed.
¶Received renal replacement therapy for acute or chronic renal failure.
**Included phenylephrine hydrochloride, epinephrine, norepinephrine bitartrate, vasopressin, dopamine hydrochloride, dobutamine, and milrinone.

Sensitivity Analysis

Performing the analyses using in-hospital mortality (*n* = 3,904 [97.1%] with complete hospital follow-up) demonstrated results similar to those of the main analysis that used 28-day mortality (Figures E9 and E10). For example, the median odds ratio decreased from 2.10 (95% CI, 1.76–2.41) in the unadjusted model to 1.18 (95% CI, 1.00–1.36) in the fully adjusted model, and adjustment with variables from the physiology, socioeconomic status, and strain domains were associated with the greatest change in the median odds ratio. The ordering and magnitude of the domains regarding their contribution to individual risk were also similar. Adding outside hospital transfers back into the cohort also demonstrated results similar to those of the primary analysis (Figures E11 and E12).

Discussion

In this multicenter cohort study of 4,019 critically ill adults with COVID-19 admitted to ICUs at 70 geographically diverse hospitals across the United States, we found wide variation in 28-day mortality across hospitals. This hospital-level variability was mostly explained by differences in socioeconomic status of the hospital population, hospital capacity and strain, and presenting ICU physiology. Furthermore, the mortality risk for individual patients was largely explained by demographic characteristics and comorbidities as well as acute physiology. To our knowledge, this is the first manuscript of its kind to investigate both hospital- and individual-level contributors to variation in mortality from a large, nationally representative cohort of critically ill patients with COVID-19. Our results help explain the wide variation in published mortality rates for critically ill patients with COVID-19 and quantify how different factors contribute to an individual patient's mortality.

Published reports on the outcomes of critically ill patients with COVID-19 have shown wide variations in mortality. For example, an early report by Arentz and colleagues (18) reported an in-hospital mortality rate of 67% for patients admitted to the ICU at one hospital in Washington State. In contrast, a study by Cummings and colleagues (19) reported a mortality of 39% in a study from two hospitals in New York City. This variability was summarized in a recent systematic review by Serafim and colleagues (8), which reported an in-hospital mortality range of 1–62%. The cause of this variation has been hypothesized to be related to various factors, such as hospital strain, patient characteristics, and variability in treatment practices (5, 20–23).

Our findings provide important insights into the reasons for this wide variation in hospital-level mortality. We found that hospital socioeconomic status, physiology, and hospital strain were the most important factors explaining this variability, whereas treatments provided to patients contributed least. To our knowledge, we are the first to show that the socioeconomic status of the community surrounding a hospital is an important contributor to hospital-level variability in outcomes in a geographically representative sample of critically ill patients with COVID-19. This finding could be due to factors related to either the impact of socioeconomic status on the health status of individual patients in the study or the unobserved variability in the quality of care that hospitals provide for a population with a lower socioeconomic status (22, 24). Interestingly, the most important individual variable from the socioeconomic status domain was the percentage of patients at the hospital who traveled >45 minutes to work. This variable has been previously used to capture the spatial mismatch hypothesis theory (25, 26), which relates to discrepancies between the location of low-income neighborhoods and the locations of employment opportunities. This variable was also found to be one of the most
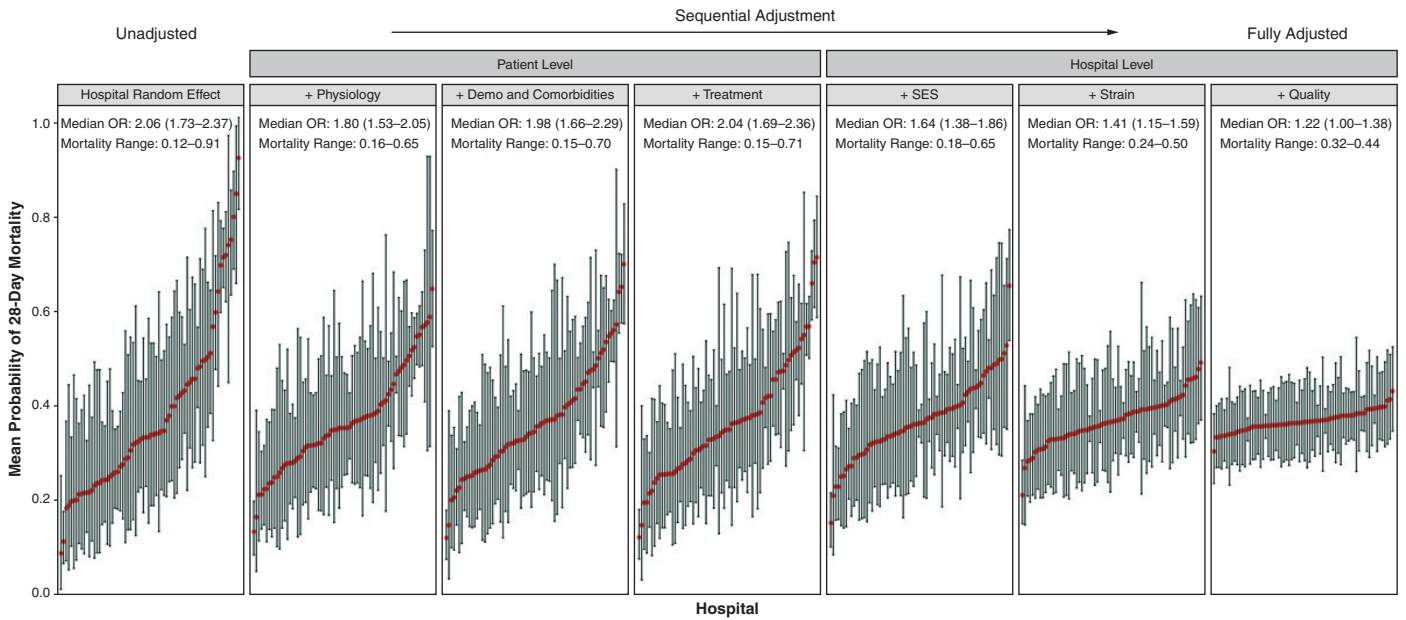
Figure 1. Case mix–adjusted probabilities of 28-day mortality. The graphs illustrate the change in interhospital variation in death as each domain is added to the unadjusted mixed-effect model (leftmost panel) and end with the fully adjusted model (rightmost panel), which shows that most of the variation in mortality across hospitals can be explained by the domains included. The x-axis is hospital ranked by increasing probability of death in 28 days, and the y-axis shows the case mix–adjusted probability of death in the mixed-effect regression model, with the red dots denoting the point estimates and the whiskers denoting the 95% confidence intervals. The median OR and range in mortality are presented for each model. Demo = demographics; OR = odds ratio; SES = socioeconomic status.

important metrics of social risk in a study investigating hospital ratings and neighborhood disadvantage (9). Our findings of increased mortality related to hospital population socioeconomic status suggest that COVID-19 may be exacerbating existing healthcare disparities in the United States (27).

The majority of an individual's risk of mortality was related to the presenting physiology, demographics, and preexisting conditions. Only one-quarter of a patient's quantifiable mortality risk was related to other factors such as hospital capacity and strain, hospital socioeconomic status, hospital quality, and treatments. Prior work suggested that the number of preexisting ICU beds is an important predictor of mortality among critically ill patients with COVID-19 (5), suggesting a correlation between ICU capacity and outcomes.

However, additional factors such as the baseline occupancy rate before the pandemic and the number of patients with COVID-19 currently admitted to the ICU are important to consider when determining the strain on critical care resources. By including these variables and other related factors in one domain, we were able to show that strain and capacity contribute to both hospital-level

variability and individual mortality. This contribution to mortality risk may be related to rationing, more aggressive goal-of-care discussions, and treatment of critically ill

patients outside the normal ICU or by less experienced providers. Hospital quality scores also had some explanatory power, albeit they had less than hospital socioeconomic status or
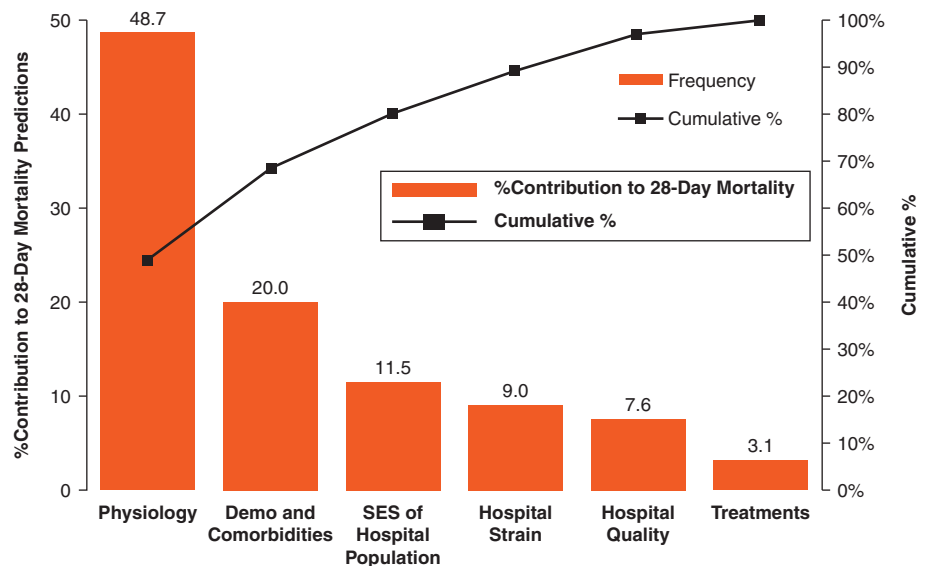

Figure 2. Contributions to 28-day mortality risk based on Shapley values. The figure illustrates the relative contribution of all variables in each domain based on Shapley values calculated from the XGBoost machine learning model (red bars; left y-axis). The cumulative contribution of the domains, moving from left to right in the figure, is shown with the line plot (right y-axis). Demo = demographics; SES = socioeconomic status.

Table 2. Hospital Characteristics for Patients Included in the Study

Variable	All Patients (N = 4,019)	28-Day Survivors (n = 2,482)	28-Day Nonsurvivors (n = 1,537)
Socioeconomics of hospital pop			
Commute to work takes >45 min, % of pop, median (IQR)	18.3 (12.9 to 25.5)*	17.8 (11.8 to 21.0)	23.4 (15.9 to 29.3)
In households speaking English only, % of pop, median (IQR)	72.6 (62.7 to 80.9)*	72.9 (64.3 to 82.4)	70.9 (61.8 to 78.9)
Uninsured, % of pop, median (IQR)	9.1 (5.7 to 13.0)	9.0 (5.7 to 13.0)	10.0 (5.6 to 13.0)
Of Black race, % of pop, median (IQR)	15.2 (8.9 to 27.0)	16.2 (9.7 to 27.0)	15.2 (8.3 to 28.0)
Dual eligible, % of pop, median (IQR)	2.5 (1.3 to 3.3)	2.6 (1.5 to 3.2)	2.3 (1.3 to 3.3)
High school diploma, % of pop, median (IQR)	88.0 (83.3 to 93.5)	88.1 (83.2 to 92.3)	87.4 (83.3 to 94.5)
Unemployed, % of pop, median (IQR)	7.3 (5.6 to 9.7)	7.3 (5.6 to 9.2)	7.3 (5.3 to 9.7)
In single-parent households, % of pop, median (IQR)	17.4 (12.8 to 22.1)*	17.0 (12.8 to 22.1)	17.5 (13.1 to 22.2)
Mean household size, median (IQR)	2.5 (2.3 to 2.8)*	2.5 (2.2 to 2.7)	2.7 (2.4 to 2.9)
Mean median home value, $, median (IQR)	305,629 (212,487 to 519,842)*	275,506 (206,403 to 519,842)	392,285 (229,155 to 519,842)
Mean median income, $, median (IQR)	63,078 (50,637 to 87,753)*	62,915 (51,208 to 87,754)	64,957 (49,753 to 96,175)
Metro area, n (%)	3,534 (87.9)	2,124 (85.6)	1,410 (91.7)
Hospital strain, median (IQR)			
Hospital ICU beds w/ STOP-COVID patients,† %	37.5 (14.4 to 69.6)*	30.6 (12.3 to 53.8)	48.2 (19.0 to 104.3)
County pop	932,202 (798,975 to 1,628,706)	945,726 (593,490 to 1,628,706)	932,202 (798,975 to 1,628,706)
ICU occupancy rate	75.0 (58.7 to 83.2)*	76.4 (63.2 to 84.2)	69.3 (54.2 to 82.1)
Number of hospital medical–surgical beds	510 (329 to 718)*	555 (358 to 733)	437 (266 to 691)
Hospital total occupancy rate	77.3 (69.5 to 84.6)*	79.5 (69.6 to 84.6)	74.6 (66.8 to 84.5)
Number of ICU beds before COVID-19	88 (48 to 112)*	98 (54 to 115)	53 (47 to 98)
Total number of county COVID-19 cases in the 30 d before admission†	1,743 (475 to 5,845)*	1,416 (398 to 4,585)	2,279 (640 to 7,268)
Total number of hospital beds	682 (448 to 1,006)*	794 (522 to 1,006)	610 (355 to 937)
Total number of hospital beds in the county	3,411 (2,286 to 5,326)*	3,657 (2,156 to 5,344)	2,768 (2,310 to 5,069)
In COVID-19 ICU or surge, n (%)	3,047 (76)	1,879 (76)	1,168 (76)
Hospital quality, median (IQR)			
Standardized outcomes mortality score	0.9 (0.4 to 1.9)*	1.0 (0.4 to 2.0)	0.5 (0.4 to 1.4)
Standardized outcomes readmission score	−0.9 (−2.0 to 0.4)*	−0.9 (−1.7 to 0.4)	−0.9 (−2.2 to 0.4)
Standardized outcomes safety score	−0.1 (−1.1 to 0.6)*	−0.1 (−1.0 to 0.7)	−0.3 (−1.7 to 0.2)
Standardized patient experience score	−0.3 (−0.7 to 0.4)*	0.0 (−0.6 to 0.4)	−0.6 (−0.9 to 0.2)
Standardized process effect score	−0.1 (−0.8 to 0.6)*	0.1 (−0.5 to 0.6)	−0.3 (−1.4 to 0.5)
Standardized process time score	−1.6 (−2.8 to 0.8)	−1.6 (−2.7 to 0.9)	−1.6 (−2.8 to 0.4)
Study day	30 (23 to 40)	30 (23 to 41)	30 (23 to 39)
Hospital treatment intensity, n (%)			
Patients vented w/ a P/F ratio <150 receiving more intense therapies, %	60 (47 to 67)*	57 (47 to 67)	63 (43 to 67)

Definition of abbreviations: COVID-19 = coronavirus disease; IQR = interquartile range; P/F = Pa_{O_2}/Fi_{O_2}; pop = population; STOP-COVID = Study of the Treatment and Outcomes in Critically Ill Patients w/ COVID-19; w/ = with.
Data regarding ICU occupancy rate were missing for 97 (2%). Missing data were imputed by using bagImpute and are included in the table.
*$P < 0.05$ for difference between survivors and nonsurvivors (Wilcoxon rank sum test for continuous variables and chi-square test for categorical variables).
†Time varying based on the date of patient admission.

strain. This suggests that the quality of the hospital a patient with COVID-19 goes to has a small but measurable effect on their outcome, which is consistent with prior work in all hospitalized patients (28).

Of all the domains studied, the treatments provided to patients had the least impact on hospital-level variability and individual-level mortality risk. This may be explained by the fact that few treatments have shown a mortality benefit for critically ill patients with COVID-19 (29–32). Notably,

the three treatments that contributed the most to improved mortality— neuromuscular blockade, aspirin, and tocilizumab—are all therapies that have previously been shown to improve outcomes for patients with COVID-19 or acute respiratory distress syndrome (33–35).

This study has several strengths. Our cohort consisted of a geographically diverse sample of critically ill adults with COVID-19. We had access to detailed patient

characteristics, physiology, interventions, and medications during their ICU stay. In addition, we were able to link the hospitals where these patients were admitted to quality scores and hospital-level socioeconomic status. Furthermore, by linking patients to the American Hospital Annual Survey data and time-varying county-level COVID-19 data, we were able to better quantify capacity and strain. Finally, in addition to standard mixed-effect regression models, we also used a state-of-the-art machine learning approach to determine

the contribution of individual variables to patient mortality (17).

This study also has several limitations. First, although we were able to identify variables associated with mortality, our study design does not lend itself to inferring causality. In addition, our findings only apply to patients admitted to the ICU, as we did not have data on patients who were critically ill but were not admitted to an ICU (e.g., because of bed rationing or goals of care). Furthermore, there may be additional variables that contribute to mortality risk that we did not account for in our study. For example, best practices and supportive care interventions, such as low-V_T ventilation for patients with acute respiratory distress syndrome, were not collected, nor were other hospital-level factors (e.g., teaching status, intensivist coverage, and nurse-to-patient ratios) or the duration of treatments. Similarly, the Shapley values measure only quantifiable mortality that is explained by the variables in the model. It is also possible that some patients were discharged alive before Day 28 only to die at home soon thereafter (e.g., patients discharged to home hospice). Although we verified in 50 patients at six participating hospitals that all patients discharged alive before 28 days were still alive at Day 28, this might not be true at all centers. In addition, the hospital quality data were collected in 2017, which may not reflect quality of care during the present-day pandemic. Finally, we only had hospital-level socioeconomic status available as opposed to individual socioeconomic status, so we could not determine whether the impact of this domain was related to the socioeconomic status of individual patients or the resources and quality that might be associated with hospitals that provide care for patients with varying socioeconomic status characteristics.

In conclusion, we found considerable interhospital variation in death among critically ill patients with COVID-19. This variability is explained by several domains, including hospital socioeconomic status, presenting physiology, and hospital capacity and strain. Similar factors contribute to an individual patient's risk of mortality, with patient-level factors (e.g., physiology, demographics, and comorbidities) explaining most of their mortality risk. ∎

References

1. COVID data tracker. Atlanta, GA: CDC; 2021 [accessed 2021 Apr 2]. Available from: https://www.cdc.gov/covid-data-tracker/#cases.
2. Severe outcomes among patients with coronavirus disease 2019 (COVID-19): United States, February 12–March 16, 2020. Atlanta, GA: CDC; 2020 [accessed 2021 Apr 2]. Available from: https://www.cdc.gov/mmwr/volumes/69/wr/mm6912e2.htm.
3. Berlin DA, Gulick RM, Martinez FJ. Severe COVID-19. *N Engl J Med* 2020; 383:2451–2460.
4. Kahn JM, Goss CH, Heagerty PJ, Kramer AA, O'Brien CR, Rubenfeld GD. Hospital volume and the outcomes of mechanical ventilation. *N Engl J Med* 2006;355:41–50.
5. Gupta S, Hayek SS, Wang W, Chan L, Mathews KS, Melamed ML, *et al.*; STOP-COVID Investigators. Factors associated with death in critically ill patients with coronavirus disease 2019 in the US. *JAMA Intern Med* 2020; 180:1436–1447.
6. Gupta S, Coca SG, Chan L, Melamed ML, Brenner SK, Hayek SS, *et al.*; STOP-COVID Investigators. AKI treated with renal replacement therapy in critically ill patients with COVID-19. *J Am Soc Nephrol* 2021;32:161–176.
7. Hayek SS, Brenner SK, Azam TU, Shadid HR, Anderson E, Berlin H, *et al.*; STOP-COVID Investigators. In-hospital cardiac arrest in critically ill patients with COVID-19: multicenter cohort study. *BMJ* 2020;371:m3513.
8. Serafim RB, Póvoa P, Souza-Dantas V, Kalil AC, Salluh JIF. Clinical course and outcomes of critically ill patients with COVID-19 infection: a systematic review. *Clin Microbiol Infect* 2021;27:47–54.
9. Fahrenbach J, Chin MH, Huang ES, Springman MK, Weber SG, Tung EL. Neighborhood disadvantage and hospital quality ratings in the Medicare hospital compare program. *Med Care* 2020;58:376–383.
10. Overall hospital quality star ratings overview. Baltimore, MD: Centers for Medicare and Medicaid Services; 2017 [accessed 2019 Mar 5]. Available from: www.qualitynet.org/dcs/ContentServer?c=Pagepagename=Qnet Public%2FPage%2FQnetTier2cid=1228775183434.
11. 2015 American Community Survey socioeconomic status data. Suitland, Suitland-Silver Hill, MD: U.S. Census Bureau; 2015 [accessed 2021 Mar 22]. Available from: https://data.census.gov.
12. Doidge JC, Gould DW, Ferrando-Vivas P, Mouncey PR, Thomas K, Shankar-Hari M, *et al.* Trends in intensive care for patients with COVID-19 in England, Wales, and Northern Ireland. *Am J Respir Crit Care Med* 2021;203:565–574.
13. Tang F, Ishwaran H. Random forest missing data algorithms. *Stat Anal Data Min* 2017;10:363–377.
14. Larsen K, Merlo J. Appropriate assessment of neighborhood effects on individual health: integrating random and fixed effects in multilevel logistic regression. *Am J Epidemiol* 2005;161:81–88.
15. Sanagou M, Wolfe R, Forbes A, Reid CM. Hospital-level associations with 30-day patient mortality after cardiac surgery: a tutorial on the application and interpretation of marginal and multilevel logistic regression. *BMC Med Res Methodol* 2012;12:28.
16. Hastie T, Tibshirani R, Friedman JH. The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York, NY: Springer; 2009.
17. Molnar C. Interpretable machine learning. Victoria, BC: Leanpub; 2020.
18. Arentz M, Yim E, Klaff L, Lokhandwala S, Riedo FX, Chong M, *et al.* Characteristics and outcomes of 21 critically ill patients with COVID-19 in Washington State. *JAMA* 2020;323:1612–1614.
19. Cummings MJ, Baldwin MR, Abrams D, Jacobson SD, Meyer BJ, Balough EM, *et al.* Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study. *Lancet* 2020;395:1763–1770.
20. Townsend MJ, Kyle TK, Stanford FC. Outcomes of COVID-19: disparities in obesity and by ethnicity/race. *Int J Obes* 2020;44:1807–1809.
21. Khunti K, Singh AK, Pareek M, Hanif W. Is ethnicity linked to incidence or outcomes of COVID-19? *BMJ* 2020;369:m1548.
22. Muñoz-Price LS, Nattinger AB, Rivera F, Hanson R, Gmehlin CG, Perez A, *et al.* Racial disparities in incidence and outcomes among patients with COVID-19. *JAMA Netw Open* 2020;3:e2021892.
23. Lim ZJ, Subramaniam A, Ponnapa Reddy M, Blecher G, Kadam U, Afroz A, *et al.* Case fatality rates for patients with COVID-19 requiring invasive mechanical ventilation: a meta-analysis. *Am J Respir Crit Care Med* 2021;203:54–66.
24. Soto GJ, Martin GS, Gong MN. Healthcare disparities in critical illness. *Crit Care Med* 2013;41:2784–2793.
25. Fernandez RM. Race, spatial mismatch, and job accessibility: evidence from a plant relocation. *Soc Sci Res* 2008;37:953–975.
26. Casciano R, Massey DS. Neighborhoods, employment, and welfare use: assessing the influence of neighborhood socioeconomic composition. *Soc Sci Res* 2008;37:544–558.
27. Thakur N, Lovinsky-Desir S, Bime C, Wisnivesky JP, Celedón JC. The structural and social determinants of the racial/ethnic disparities in the U.S. COVID-19 pandemic: what's our role? *Am J Respir Crit Care Med* 2020;202:943–949.
28. Werner RM, Bradlow ET. Relationship between Medicare's Hospital Compare performance measures and mortality rates. *JAMA* 2006;296:2694–2702.

29. Cao B, Wang Y, Wen D, Liu W, Wang J, Fan G, et al. A trial of lopinavir-ritonavir in adults hospitalized with severe COVID-19. N Engl J Med 2020;382:1787–1799.

30. Beigel JH, Tomashek KM, Dodd LE, Mehta AK, Zingman BS, Kalil AC, et al.; ACTT-1 Study Group. Remdesivir for the treatment of COVID-19: final report. N Engl J Med 2020;383:1813–1826.

31. Cavalcanti AB, Zampieri FG, Rosa RG, Azevedo LCP, Veiga VC, Avezum A, et al.; Coalition COVID-19 Brazil I Investigators. Hydroxychloroquine with or without Azithromycin in Mild-to-Moderate COVID-19. N Engl J Med 2020;383:2041–2052.

32. Stone JH, Frigault MJ, Serling-Boyd NJ, Fernandes AD, Harvey L, Foulkes AS, et al.; BACC Bay Tocilizumab Trial Investigators. Efficacy of tocilizumab in patients hospitalized with COVID-19. N Engl J Med 2020; 383:2333–2344.

33. Papazian L, Forel JM, Gacouin A, Penot-Ragon C, Perrin G, Loundou A, et al.; ACURASYS Study Investigators. Neuromuscular blockers in early acute respiratory distress syndrome. N Engl J Med 2010;363: 1107–1116.

34. Horby P, Lim WS, Emberson JR, Mafham M, Bell JL, Linsell L, et al.; RECOVERY Collaborative Group. Dexamethasone in hospitalized patients with COVID-19. N Engl J Med 2021; 384:693–704.

35. Chow JH, Khanna AK, Kethireddy S, Yamane D, Levine A, Jackson AM, et al. Aspirin use is associated with decreased mechanical ventilation, intensive care unit admission, and in-hospital mortality in hospitalized patients with coronavirus disease 2019. Anesth Analg 2021;132: 930–941.