Factors associated with hospitalization and critical illness among 4,594 patients with COVID-19 disease in New York City: A prospective cohort study

Journal:	*BMJ*
Manuscript ID	BMJ-2020-057265
Article Type:	Research
BMJ Journal:	BMJ
Date Submitted by the Author:	20-Apr-2020
Complete List of Authors:	Petrilli, Christopher; NYU School of Medicine, Medicine Jones, Simon; NYU School of Medicine, Population Health Yang, Jie; NYU School of Medicine, Population Health Rajagopalan, Harish; NYU Langone Health O'Donnell, Luke; NYU School of Medicine, Medicine Chernyak, Yelena; NYU Langone Health Tobin, Katie; NYU Langone Health Cerfolio, Robert; NYU School of Medicine, Cardiothoracic surgery Francois, Fritz; NYU School of Medicine, Medicine Horwitz, Leora; NYU School of Medicine, Population Health
Keywords:	SARS-Cov-2, COVID-19, hospital medicine, mortality

SCHOLARONE™
Manuscripts

Title: Factors associated with hospitalization and critical illness among 4,594 patients with COVID-19 disease in New York City: A prospective cohort study

Running title: COVID-19

Authors:

Christopher M. Petrilli, MD[1,2]
Simon A. Jones, PhD[3,4]
Jie Yang, MPH, MS[4]
Harish Rajagopalan[2]
Luke O'Donnell, MD[1]
Yelena Chernyak[2]
Katie A. Tobin, MPA[2]
Robert J. Cerfolio, MD[2,5]
Fritz Francois, MD[2,6]
Leora I. Horwitz, MD, MHS[1,3,4]

Affiliations:

[1]Division of General Internal Medicine and Clinical Innovation, Department of Medicine, NYU Grossman School of Medicine, New York, NY

[2]NYU Langone Health, New York, NY

[3]Division of Healthcare Delivery Science, Department of Population Health, NYU Grossman School of Medicine, New York, NY

[4]Center for Healthcare Innovation and Delivery Science, NYU Langone Health, New York, NY

[5]Department of Cardiothoracic Surgery, NYU Grossman School of Medicine, New York, NY

[6]Division of Gastroenterology, Department of Medicine, NYU Grossman School of Medicine, New York, NY

Corresponding author:

Leora Horwitz, MD, MHS
227 E. 30th St. #633
New York, NY 10016
Leora.horwitz@nyulangone.org
(646) 501-2685

Word count: 3,826

Tables: 3

Figures: 2

References: 36

Confidential: For Review Only

Key words: SARS-Cov-2, COVID-19, hospital medicine, mortality

Acknowledgements: The authors thank Andrew Admon, MD, MPH; Brian Bosworth, MD; Steven Chatfield; Thomas Doonan; Robert Grossman, MD; Juan Peralta, and Daniel Widawsky for operational and technical support. We also thank the thousands of NYU Langone Health employees who have cared for these patients.

Funding: This work was funded in part by the Kenneth C. Griffin Charitable Fund, which had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

Contributorship statement: CMP, LO'D, KAT, RJC, FF and LIH conceived of the project. CMP, YC and HR obtained, validated and cleaned the data. JY and SAJ performed the statistical analyses. RJC and FF provided administrative and operational support. LIH supervised the project and drafted the manuscript. All authors discussed the results and contributed to the final manuscript. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Data sharing: Identifiable patient level data from this project is not available to the public.

Confidential: For Review Only

Abstract

Objective: To describe outcomes of hospitalized patients with COVID-19 in the United States, and the clinical and laboratory characteristics associated with severity of illness. This information may help assist in patient triage and may improve the accuracy of modeling for anticipated resource needs and mortality rates.

Design, Setting and Participants: Cross-sectional analysis of all patients with laboratory-confirmed SARS-Cov-2 infection at a single academic medical center in New York City between March 1, 2020 and April 5, 2020. The final date of follow up was April 16, 2020.

Main outcomes and measures: Primary outcomes were hospitalization and critical illness (intensive care, mechanical ventilation, discharge to hospice and/or death). Predictors included patient demographics, medical history, vital signs and laboratory results. We conducted multivariable logistic regression to identify risk factors for adverse outcomes, and maximum information gain decision tree classifications to identify key splitters.

Results: Of 9,205 patients tested for COVID-19, 4,734 (51.4%) were positive; 4,594 were included. A total of 2,390 (52.0%) were hospitalized, of whom 1,504/2,390 (62.9%) have been discharged, and 486/2,390 (20.3%) have died or been discharged to hospice. Of 580 (24.3%) patients requiring mechanical ventilation, 278/580 (47.9%) have died and 91 (15.7%) have been discharged or extubated. Among all COVID-19 patients, the strongest risk for hospitalization was age, with odds ratio (OR)>4 for all age groups above 54 years and OR 48.5 for age ≥75 years. Heart failure (OR 3.6, 95% CI 1.8-7.9), chronic kidney disease (OR 3.2, 95% CI, 2.0-5.3), male sex (OR 2.9, 95% CI 2.5-3.4), and body mass index (BMI) > 40 (OR 2.7, 95% CI, 2.0-3.8) were other risks. For development of critical illness, the strongest risks besides age were BMI>40 (OR 2.5, 95% CI 1.8-3.6), male sex (OR 2.1, 95% CI 1.8-2.5), heart failure (OR 1.8, 95% CI, 1.3-2.7) and diabetes (OR 1.5, 95% CI 1.2-1.9). However, among hospitalized patients, admission oxygen saturation <88% (OR 4.3, 95% CI 3.2-5.9), C-reactive protein (CRP) >200 (OR 4.9, 95% CI, 2.6-9.6), and d-dimer>2500 (OR 3.9, 95% CI, 2.3-6.6) were more strongly associated with critical illness than age or comorbidities. In the classification tree for hospitalization, the most important features identified were age >65, male sex and diabetes; for critical illness with lab results, the most important was SpO2<88, followed by procalcitonin >0.5, troponin <0.1 (protective), and age >59.

Conclusions: Age and comorbidities are powerful predictors of hospitalization and to a lesser extent of critical illness; however, when added, admission oxygen impairment and markers of inflammation are most strongly associated with critical illness.

Confidential: For Review Only

Background

The first announcement of a cluster of novel pneumonia-like illness was made on December 31, 2019 by China. Since then, the causative organism, SARS-Cov-2, has produced a global pandemic that to date has infected over 2.2 million people and directly resulted in over 150,000 known deaths.

While several reports from China,[1,2] Italy,[3,4] and most recently the United States Centers for Disease Control and Prevention[5,6] have described some characteristics of patients with COVID-19, the disease caused by SARS-Cov-2, little is understood about factors associated with hospital admission and with severe disease. Studies to date have included few patients with severe outcomes,[1,7-11] or have not compared those to patients with less virulent disease,[12-14] making it difficult to assess characteristics associated with poor outcomes. No large studies have conducted multivariable regression to help identify the strongest risk factors. Moreover, very few studies to date describe the experience in the United States. Differences in population demographics and behaviors may limit generalizability of studies from China to patients in the United States, which currently has the most COVID-19 cases in the world.

Understanding which patients are most at risk for hospitalization and the distribution of outcomes among hospitalized patients is crucial for many reasons. It can assist emergency providers in making triage decisions and ambulatory clinicians in identifying patients who would most benefit from early treatment once available. It can help inform policymakers about highest risk populations, who may need particular protection in policy determinations. Finally, it can help epidemiologists to improve the accuracy of projections about likely need for hospital beds and staffing needs in a region given its demographic characteristics.[15]

Confidential: For Review Only

For similar reasons, is also important to understand the risk of critical illness among those hospitalized. Clinicians need this information both to identify patients at greatest risk of deterioration, and to inform decision-making about hospital discharge. Policymakers and epidemiologists need this information to project likely need for intensive care unit capacity, ventilators and associated staffing. It would also help improve the reliability of future mortality rate estimation.

New York City is now the epicenter of the COVID-19 outbreak in the United States, with over 122,000 known cases in the city and over 7,800 confirmed deaths as of April 17: more than anywhere else in the country.[16] In this report, the largest case series from the United States to date, we describe characteristics of COVID-19 patients treated at a large quaternary academic health system in New York City and Long Island, and the association of these characteristics with adverse outcomes.

Methods

Study setting

The study was conducted at NYU Langone Health, which includes over 260 outpatient office sites and four acute care hospitals (two in Manhattan, one in Brooklyn, one in Long Island), ranging from a quaternary care hospital to a safety net institution. As the epidemic evolved, the health system added intensive care unit beds and inpatient capacity, resulting in approximately 394 ICU beds and 1,357 non-ICU beds at time of writing.

Study cohort

We began with all patients tested for SARS-Cov-2 between March 1, 2020 and April 5, 2020. We then created a cohort of confirmed COVID-19 cases, defined as a positive result on

https://mc.manuscriptcentral.com/bmj

Confidential: For Review Only

real-time reverse-transcriptase-polymerase-chain-reaction (RT-PCR) assay of nasopharyngeal or oropharyngeal swab specimens. Initial tests were conducted by the New York City Department of Health and Mental Hygiene; as of March 16, tests were conducted in our clinical laboratory using the Roche SARS-CoV2 assay in the Cobas 6800 instruments through emergency use authorization (EUA) granted by the FDA. On March 31 we added testing using the SARS-CoV2 Xpert Xpress assay in the Cepheid GeneXpert instruments also under EUA by FDA. The targets amplified by these assays are the ORF1/a and E in the Roche Cobas assay and N2 and E genes in the Cepheid XpertXpress. After March 16, only nasopharyngeal samples were collected and tested.

Testing was performed for patients presenting to the emergency department with any complaint consistent with COVID-19, including fever, cough, shortness of breath, fatigue, gastrointestinal complaints, syncope, known exposure to a COVID-19 positive patient, or clinician concern. In addition, ambulatory testing was available by appointment with clinician's referral until March 26, 2020, when New York State recommended restricting testing of patients with mild or moderate illness. Outpatient testing of symptomatic or concerned employees has remained available throughout the study period. Repeat testing of negative specimens was conducted at clinician discretion. If testing was repeated and discordant (i.e. negative test followed by a positive test), we used the positive result.

We excluded from the COVID-19 cohort 140 patients who were not hospitalized and were missing all data besides age and sex. We obtained complete follow up on the COVID-19 cohort through April 16, 2020.

Data sources and patient and public involvement

https://mc.manuscriptcentral.com/bmj

Confidential: For Review Only

All study data was obtained from the electronic health record (Epic Systems, Verona, WI), which is an integrated electronic health record (EHR) including all inpatient and outpatient visits in the health system. For data on tobacco use, body mass index (BMI) and comorbidities we included any data in the EHR, including data entered during prior inpatient or outpatient visits.

Patients and the public were not a priori involved in the design and conduct of the study, in the choice of outcomes, in recruitment, or in planned dissemination. However, we incorporated many comments from the public on an earlier preprint version of the paper into the final analysis.

Main outcomes

We assessed two primary outcomes: inpatient hospitalization and critical illness, defined as a composite of care in the intensive care unit, use of mechanical ventilation, discharge to hospice, or death. We assessed outcomes longitudinally over the entire study period, not just at the time of the initial testing event. For patients with multiple visits, the most severe outcome was assigned. For instance, patients who did not need hospitalization at time of initial testing but were later hospitalized were assigned to the hospitalization group. Similarly, patients who were initially admitted and discharged and then readmitted requiring invasive ventilation were assigned to the critical illness group.

Predictors

We obtained from the electronic health record the following variables: age at time of testing, sex, race/ethnicity as reported by the patient (aggregated into non-Hispanic white, non-Hispanic African American, Asian, Hispanic, other/multiracial and unknown), and history of

https://mc.manuscriptcentral.com/bmj

Confidential: For Review Only

hypertension, hyperlipidemia, coronary artery disease, heart failure, pulmonary disease (defined by chronic obstructive pulmonary disease or asthma), malignancy (excluding non-metastatic non-melanoma skin cancer), diabetes, and obesity (defined by most recent body mass index). We also obtained vital signs and first set of laboratory results where available. For multivariable modeling, we bucketed vital sign and laboratory results into categories by degree of abnormality based on clinical judgment because of non-linear associations with outcome. We selected these predictors based on prior published literature[1,5] and our clinical experience with COVID-19 patients.

Statistical analysis

We used descriptive statistics to characterize each cohort of patients: those not hospitalized, all those hospitalized, those hospitalized without critical illness, and those with critical illness (care in intensive care unit, mechanical ventilation, discharge to hospice, or death). We then fitted multivariable logistic regression models with admission and with critical illness as the outcomes to identify factors associated with those outcomes. In analyses using hospitalization as the outcome, we included only patient demographics and comorbidities, since 83% of the patients who were not admitted were evaluated in ambulatory testing centers and did not have vitals or laboratory studies collected. For the critical illness analyses, we included the above predictors and for one of the models added temperature and oxygen saturation on presentation, as well as the first result of c-reactive protein, d-dimer, ferritin, procalcitonin, and troponin when obtained. We included all selected predictors based on *a priori* clinical significance after testing for collinearity using the variance inflation factor (VIF) and ensuring none had VIF>2.[17]

Confidential: For Review Only

For the admission model, we included all patients testing positive (excluding the 140 patients with no data besides age and sex). We constructed three models for association with critical illness. First, we constructed a model using the entire positive cohort and including all demographic and comorbidity predictors, to assess associations with critical illness among all positive patients. In this case, the comparison group included both hospitalized and non-hospitalized patients. Second, we constructed a model restricted to hospitalized patients, excluding those still hospitalized without critical illness (4.8% of total), to assess associations with critical illness among hospitalized patients with a known outcome. This model also included all demographic and comorbidity predictors. Third, we added to that model vital signs and the first set of laboratory results, to assess clinical associations with critical illness among hospitalized patients. We excluded from the third model four patients who expired in the emergency department before vital signs or laboratory results could be collected. We obtained odds ratios from the models and profiled confidence intervals for the odds ratios using the approach of Venables & Ripley,[18] since assuming normality of the maximum likelihood estimate to estimate Wald-type confidence intervals can lead to biased estimates.[19]

Finally, we constructed maximum information gain decision tree classifications for both hospital admission and severe complication to identify the variables that best classified patients into different outcome cohorts. For a given population, the decision tree classification method splits the population into two groups using one feature at a time, starting with the feature that maximizes the split between groups relative to the outcome in question.[20,21] Subsequent splits reevaluate each split subgroup for the next best feature. The final population in each end node has similar characteristics and outcomes. We used the decision tree classifier from Python 3.7.4 scikit-learn library. We chose to maximize information gain (which minimizes entropy) for each

https://mc.manuscriptcentral.com/bmj

Confidential: For Review Only

branch split in the classification tasks. We also pruned the trees to prevent overfitting by limiting the maximum depth, minimum samples in a leaf, and minimum sample splits. To avoid splitting the tree by meaningless characteristics like whether information is present or not, we dropped from these classifications all patients with missing data for BMI and tobacco use. For all decision tree classifications, we split the data into a training set (80%) and hold-out set (20%) and calculated area under the curve (AUC) on the hold-out set. For the admission model, we ran 48 iterations to achieve optimal parameters. For the critical illness model, we ran 64 iterations.

The logistic regression models were conducted with R, version 3.6.3, and the decision trees with Python, version 3.7.4. All analyses used 2-sided statistical tests and we considered a p value < 0.05 to be statistically significant without adjustment for multiple testing.

This study was approved by the NYU Grossman School of Medicine Institutional Review Board, which granted both a waiver of informed consent, and a waiver of the Health Information Portability and Privacy Act.

Results

During the study period, the health system tested 9,205 patients for COVID-19, of whom 4,734 (51.4%) were positive and 140 were then excluded for lack of data. Of the remaining 4,594 patients testing positive, 2,204 (48.0%) were treated only as outpatients through the end of the study period, and 2,390 (52.0%) required admission to the hospital. Among those admitted to the hospital, 2,275 (95.2%) have experienced a study outcome, among which 1,416/2,275 (62.2%) were discharged without critical illness and 859/2,275 (37.8%) experienced critical illness, including 486/2,275 (21.4%) who have been discharged to hospice or died.

Confidential: For Review Only

Among the 859 patients with critical illness, 580/859 (67.5%) required mechanical ventilation, 94/859 (10.9%) required intensive care without mechanical ventilation, and 185/859 (21.5%) were discharged to hospice or died without either intensive care or mechanical ventilation. Final outcomes to date for each subgroup are shown in **Figure 1**.

Characteristics of study population

The median age of the COVID-19 positive study population was 52 years (interquartile range, 37 to 66), and 2,317 (50.4%) were male. A total of 749 (16.3%) had diabetes, 1,569 (34.2%) obesity, and 1,455 (31.7%) any form of cardiovascular disease. Among hospitalized patients, the median length of stay among those with final discharge disposition (discharged alive or died) was 6.3 days (interquartile range, 3.7 to 9.7). Median days of follow up for those still hospitalized with critical illness (N=285) was 17 (IQR 13 to 21); for those still hospitalized without critical illness (N=115) was 14 (IQR 12-16).

Hospitalized patients were more likely to be male (62.3% vs 37.6%) and were substantially more likely to have comorbidities than non-hospitalized patients (any comorbidity, 57.0% vs 27.4%), particularly with regard to cardiovascular disease (45.3% vs. 17.1%), diabetes (26.0% vs 5.8%), obesity (38.9% vs 29.0%), and chronic kidney disease (10.6% vs. 1.1%) (**Table 1**). Differences in sex and comorbidities between hospitalized patients experiencing critical illness and those who did not were much smaller. Among these patients, differences in clinical presentation and laboratory results were more prominent. Patients with critical illness more often presented with hypoxia (initial O2 saturation 25th percentile 86% versus 92%), and had higher initial levels of c-reactive protein (median 140 vs 88), d-dimer (median 508 vs 315), ferritin (median 931 vs 628), procalcitonin (0.28 vs 0.10) and troponin (0.07 vs 0.02) (**Table 2**).

Confidential: For Review Only

Predictors of hospitalization

In multivariable analysis of the full COVID-19 positive cohort, the factors most associated with hospitalization were age, including 75 years or older (OR 48.5, 95% CI, 33.7-71.6), age 65-74 (OR 10.1, 95% CI, 7.9-13.1), heart failure (OR 3.6, 95% CI 1.8-7.9), chronic kidney disease (OR 3.2, 95% CI, 2.0-5.3), male sex (OR 2.9, 95% CI 2.5-3.4), and BMI>40 (OR 2.7, 95% CI, 2.0-3.8). Also significant were overweight, BMI 30-40, hypertension, and pulmonary disease. Hyperlipidemia was associated with lower hospitalization risk (OR 0.6, 95% CI, 0.5-0.8) (**Table 1**).

Among all COVID-19 positive patients, the factors most associated with critical illness were age, including 75 years or older (OR 14.2, 95% CI 10.5-19.4) and age 65-74 (OR 7.7, 95% CI, 5.7-10.3); BMI > 40 (OR 2.5, 95% CI 1.8-3.6); male sex (OR 2.1, 95% CI 1.8-2.5) and heart failure (OR 1.8, 95% CI, 1.3-2.7); with diabetes and BMI 30-40 also significant. Hyperlipidemia was associated with lower risk of critical illness among all COVID-19 patients (OR 0.70, 95% CI, 0.6-0.9) (**Table 3**).

In our cohort, being self-reported Hispanic was associated with increased risk of admission (OR 1.6, 95% CI, 1.3-1.9) but not critical illness; being self-reported Asian was associated with increased risk both of hospitalization (OR 1.4, 95% CI, 1.0-1.8) and critical illness (OR 1.6, 95% CI, 1.2-2.3). African American patients had similar admission risk as white patients and lower risk of critical illness (OR 0.8, 95% CI, 0.6-1.0).

By contrast, among hospitalized patients, the only factors associated with critical illness were age, including age 75 years or older (OR 3.5, 95% CI, 2.5-5.0) and age 65-74 (OR 2.7, 95% CI 2.0-3.8); BMI>40 (OR 1.8, 95% CI, 1.2-2.7); history of heart failure (OR 1.5, 95% CI, 1.0-

Confidential: For Review Only

2.3); and male sex (OR 1.5, 95% CI, 1.2-1.8). Hospitalized African American patients had lower risk for critical illness (OR 0.74, 95% CI 0.6-0.99) (**Table 3**).

After adding admission vitals and first set of laboratory results, only age and BMI remained significant risks; in this model, the factors most associated with critical illness were admission oxygen saturation <88% (OR 4.3, 95% CI 3.2-5.9), troponin >1 (OR 5.5, 95% CI, 2.4-14.2), C-reactive protein (CRP) >200 (OR 4.9, 95% CI, 2.6-9.6), and d-dimer>2500 (OR 3.9, 95% CI, 2.3-6.6). Age 0-18 had a high OR of 4.2 (95% CI, 1.7 to 9.8), but this age group included only 32 patients as children are rarely tested; most of the critically ill were > 16 years. There were no deaths in this age group (**Table 3**).

In a maximum information gain decision tree classification for hospitalization, the most important feature at the top-level branch point was age >65, followed by age >38, male sex and diabetes. The AUC of the model was 0.82. For critical illness without vital signs and laboratory results, the top branch point was age>63, followed by age>87 and heart failure. The AUC of the model was 0.65. Finally, for critical illness with vital signs and laboratory results, the top branch point was SpO2<88, followed by procalcitonin >0.5, troponin <0.1 (protective) and age >59. The additional information increased the AUC to 0.72 (**Figure 2**).

Discussion

In this report, we describe characteristics of 4,594 patients with laboratory-confirmed COVID-19 disease in New York City, of whom 2,390 required hospital admission and 859 required intensive care, mechanical ventilation, were discharged to hospice and/or died. We find particularly strong associations of older age, male sex, heart failure, chronic kidney disease and obesity with hospitalization and critical illness risk among all patients with COVID-19, with less

Confidential: For Review Only

influence of chronic pulmonary disease and other forms of heart disease. By contrast, we found comorbidities to be less strongly associated with critical illness once patients were hospitalized. Among those patients, we noted the importance of hypoxia despite supplemental oxygen and early elevations in inflammatory markers (especially d-dimer and c-reactive protein) in distinguishing among patients who go on to develop critical illness and those who do not. In the hospitalized population, measures of inflammation were much more important than demographic characteristics and comorbidities.

The largest detailed case series published to date included 1,099 hospitalized patients with laboratory-confirmed COVID-19 infection in China, of whom only 25 (2.3%) underwent invasive ventilation and 15 (1.4%) died.[1] By contrast, 24% of hospitalized patients with definitive outcomes in this case series required invasive ventilation and 20% have died. Given the very high prevalence of disease in New York City and the relative paucity of baseline hospital beds per capita (1.5-2.7 beds per 1,000 in all boroughs except Manhattan), admission thresholds may be higher in New York City than in China (4.2 beds per 1,000).[22,23] Moreover, in the series by Guan et al, only a quarter of the hospitalized patients had any chronic comorbidity, whereas in our series 54% of hospitalized patients had at least one of eight major chronic diseases.

In fact, outcomes in the majority of reports are similar to ours. A commentary by the Chinese Center for Disease Control and Prevention described outcomes for 72,314 cases, of which 14% were severe (similar to hospitalized patients in our series) and 5% critical with respiratory or multi-organ failure (similar to those with critical illness in our series).[2] Among critical cases, mortality was 49%; ours is 48% to date among ventilated patients. This is also similar to the typical mortality rate from acute respiratory distress syndrome (ARDS) of about

Confidential: For Review Only

35-45%.[24,25] Finally, our results are also consistent with a recent national case series reported by the US CDC that found that 457 of 1,037 (44%) hospitalized patients required ICU admission, and that three quarters had at least one chronic condition.[6]

The comorbidities we identified as associated with hospitalization in COVID-19 are largely similar to those associated with any type of severe infectious disease requiring hospitalization or ICU level care,[26] though we were surprised that cancer and chronic pulmonary disease did not feature more prominently. Others have also noted the surprising absence of asthma and COPD as risks for severity of illness in patients with Covid-19.[27] The pathophysiologic reasons for this are unknown. The demographic distribution of hospitalized patients is also similar to other acute respiratory infections. For instance, while advanced age was by far the most important predictor of hospitalization and an important predictor of severe outcomes (as it is for most illnesses), 54% of hospitalized patients were younger than 65 years. This is typical of the hospitalization pattern in viral respiratory disease. Studies of influenza hospitalizations in the United States have found that people younger than 65 years account for 53-57% of influenza-related hospitalizations.[28,29]

Surprisingly, though some have speculated that high rates of smoking in China explained some of the morbidity in those patients, we did not find tobacco use to be associated with increased risk of hospitalization or critical illness; in fact, it even appeared protective. However, this may be artifactual: patients with unknown smoking status had significantly higher risks of admission and of critical illness. It is possible that data are disproportionately missing for current or former smokers who might not care to answer that question; if so that would attenuate the apparent benefit of smoking. Very few (<5%) of patients had a recorded history of vaping; separate analyses could not be conducted for those patients.

Confidential: For Review Only

Our findings about the importance of inflammatory markers in distinguishing future critical from non-critical illness among hospitalized patients were striking. Among these, early elevations in c-reactive protein and d-dimer had the strongest association with mechanical ventilation or mortality. Hyperinflammatory states are well described in severe sepsis;[30] however, the degree to which COVID-19 related inflammation is similar to or different than that typically found in sepsis is unknown. Some emerging case reports suggest that patients with critical COVID-19 disease are developing complications from hypercoagulability,[9] including both pulmonary emboli[31] and microscopic thrombi.[32] In this regard it is notable that one of the chronic conditions strongly associated with critical illness was obesity. Obesity is well-recognized to be a pro-inflammatory condition.[33,34] In addition, this might explain why hyperlipidemia appeared to be protective: statin therapy is anti-inflammatory and has been shown to reduce cytokine levels.[35] Some studies suggest that elevated LDL itself may be beneficial in reducing mortality from respiratory diseases through direct anti-infectious properties, though mean LDL levels were low in our population.[36] Finally, we noted that early elevation in procalcitonin was a powerful splitter in the classification tree, although COVID-19 appears to be characterized by low procalcitonin levels in general. While many patients with elevated procalcitonin and critical illness were treated with antibiotics, it remains unclear whether these patients actually had bacterial disease or whether the elevation in procalcitonin was another manifestation of a general hyperinflammatory state.

This study includes several limitations. Most important, data on non-hospitalized patients was more limited because many came only for testing and did not have history, vital signs or blood samples collected. We may therefore be overestimating the importance of chronic disease in hospitalization risk. This limitation was not present for the analyses limited to the

https://mc.manuscriptcentral.com/bmj

Confidential: For Review Only

hospitalization cohort. We did not have access to symptom duration, which is an important predictor of hospitalization: patients rarely require hospitalization with less than a week of symptoms. However, this limitation should not affect the demographic and clinical characteristics of those requiring admission and having severe deterioration. Our patients were all from a single geographic region, treated within a single health system; factors associated with poor outcomes may differ elsewhere, though our patient population is very diverse. We did not have inflammatory markers available for non-hospitalized patients; it is possible that these would have been strong predictors for hospitalization risk, not just critical illness, if available. Finally, a standardized admission laboratory protocol was only established about two weeks into the epidemic, resulting in missing laboratory data for earlier patients, especially those who were less acutely ill.

Overall, we find that age and comorbidities are powerful predictors of requiring hospitalization rather than outpatient care; however, degree of oxygen impairment and markers of inflammation are strongest predictors of poor outcomes during hospitalization. Clinicians should consider routinely obtaining inflammatory markers during hospitalizations for COVID-19.

Confidential: For Review Only

References

1. Guan WJ, Ni ZY, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med.* 2020.
2. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA : the journal of the American Medical Association.* 2020.
3. Livingston E, Bucher K. Coronavirus Disease 2019 (COVID-19) in Italy. *JAMA : the journal of the American Medical Association.* 2020.
4. Onder G, Rezza G, Brusaferro S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *JAMA : the journal of the American Medical Association.* 2020.
5. Cdc Covid- Response Team. Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) - United States, February 12-March 16, 2020. *MMWR Morb Mortal Wkly Rep.* 2020;69(12):343-346.
6. Cdc Covid- Response Team. Preliminary Estimates of the Prevalence of Selected Underlying Health Conditions Among Patients with Coronavirus Disease 2019 - United States, February 12-March 28, 2020. *MMWR Morb Mortal Wkly Rep.* 2020;69(13):382-386.
7. Wu C, Chen X, Cai Y, et al. Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. *JAMA Intern Med.* 2020.
8. Wang Z, Yang B, Li Q, Wen L, Zhang R. Clinical Features of 69 Cases with Coronavirus Disease 2019 in Wuhan, China. *Clin Infect Dis.* 2020.
9. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet.* 2020;395(10229):1054-1062.
10. Guo T, Fan Y, Chen M, et al. Cardiovascular Implications of Fatal Outcomes of Patients With Coronavirus Disease 2019 (COVID-19). *JAMA Cardiol.* 2020.
11. Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA : the journal of the American Medical Association.* 2020.
12. Korean Society of Infectious D, Korea Centers for Disease C, Prevention. Analysis on 54 Mortality Cases of Coronavirus Disease 2019 in the Republic of Korea from January 19 to March 10, 2020. *J Korean Med Sci.* 2020;35(12):e132.
13. Bhatraju PK, Ghassemieh BJ, Nichols M, et al. Covid-19 in Critically Ill Patients in the Seattle Region - Case Series. *N Engl J Med.* 2020.
14. Grasselli G, Zangrillo A, Zanella A, et al. Baseline Characteristics and Outcomes of 1591 Patients Infected With SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy. *JAMA : the journal of the American Medical Association.* 2020.
15. Lipsitch M, Swerdlow DL, Finelli L. Defining the Epidemiology of Covid-19 - Studies Needed. *N Engl J Med.* 2020;382(13):1194-1196.
16. NYC Health. Cases, Hospitalizations and Deaths. https://www1.nyc.gov/site/doh/covid/covid-19-data.page#download. Published 2020. Accessed 4 April, 2020.

Copyright Draft For Review Only

17. Fox J, Monette G. Generalized collinearity diagnostics. *J American Statistical Association.* 1992;87:178-183.
18. Venables WN, Ripley BD. *Modern Applied Statistics.* 4 ed: Springer; 2002.
19. Venzon DJ, Moolgavkar SH. A method for computing profile-likelihood based confidence intervals. *Applied Statistics.* 1988;37:87-94.
20. Kotsiantis SB. Decision trees: a recent overview. *Artificial Intelligence Review.* 2013;39:261-283.
21. Song YY, Lu Y. Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry.* 2015;27(2):130-135.
22. Melby C, Gu J, Rojanasakul M. Mapping New York City Hospital Beds as Coronavirus Cases Surge. https://www.bloomberg.com/graphics/2020-new-york-coronavirus-outbreak-how-many-hospital-beds/. Published 2020. Accessed 7 Apr, 2020.
23. World Health Organization. Hospital beds (per 10 000 population). https://www.who.int/data/gho/data/indicators/indicator-details/GHO/hospital-beds-(per-10-000-population). Published 2020. Accessed 7 Apr, 2020.
24. Rubenfeld GD, Caldwell E, Peabody E, et al. Incidence and outcomes of acute lung injury. *N Engl J Med.* 2005;353(16):1685-1693.
25. Bellani G, Laffey JG, Pham T, et al. Epidemiology, Patterns of Care, and Mortality for Patients With Acute Respiratory Distress Syndrome in Intensive Care Units in 50 Countries. *JAMA : the journal of the American Medical Association.* 2016;315(8):788-800.
26. Cecconi M, Evans L, Levy M, Rhodes A. Sepsis and septic shock. *Lancet.* 2018;392(10141):75-87.
27. Halpin DMG, Faner R, Sibila O, Badia JR, Agusti A. Do chronic respiratory diseases or their treatment affect the risk of SARS-CoV-2 infection? *Lancet Respir Med.* 2020.
28. Simonsen L, Fukuda K, Schonberger LB, Cox NJ. The impact of influenza epidemics on hospitalizations. *J Infect Dis.* 2000;181(3):831-837.
29. Zhou H, Thompson WW, Viboud CG, et al. Hospitalizations associated with influenza and respiratory syncytial virus in the United States, 1993-2008. *Clin Infect Dis.* 2012;54(10):1427-1436.
30. Hotchkiss RS, Karl IE. The pathophysiology and treatment of sepsis. *N Engl J Med.* 2003;348(2):138-150.
31. Danzi GB, Loffi M, Galeazzi G, Gherbesi E. Acute pulmonary embolism and COVID-19 pneumonia: a random association? *European heart journal.* 2020.
32. Tang N, Bai H, Chen X, Gong J, Li D, Sun Z. Anticoagulant treatment is associated with decreased mortality in severe coronavirus disease 2019 patients with coagulopathy. *J Thromb Haemost.* 2020.
33. Schmidt FM, Weschenfelder J, Sander C, et al. Inflammatory cytokines in general and central obesity and modulating effects of physical activity. *PLoS One.* 2015;10(3):e0121971.
34. Caer C, Rouault C, Le Roy T, et al. Immune cell-derived cytokines contribute to obesity-related inflammation, fibrogenesis and metabolic deregulation in human adipose tissue. *Sci Rep.* 2017;7(1):3000.
35. Geovanini GR, Libby P. Atherosclerosis and inflammation: overview and updates. *Clin Sci (Lond).* 2018;132(12):1243-1252.

Copyright: For review only

36. Jacobs D, Blackburn H, Higgins M, et al. Report of the Conference on Low Blood Cholesterol: Mortality Associations. *Circulation.* 1992;86(3):1046-1060.

Confidential: For Review Only

Figure legends

Figure 1: Flow diagram of included patients

Figure 2A: Maximum likelihood decision tree classification for hospitalization

Figure 2B: Maximum likelihood decision tree classification for critical illness, without vital signs and laboratory results

Figure 2C: Maximum likelihood decision tree classification for critical illness, with vital signs and laboratory results

Confidential: For Review Only

Table 1: Characteristics of Covid-19 patients, by hospitalization status, and multivariable risk of hospitalization

Characteristic	Overall, N=4,594 N (%) or median (IQR)	Not hospitalized, N=2,204 N (%) or median (IQR)	Hospitalized, N=2,390 N (%) or median (IQR)	Risk of hospitalization Odds ratio (95% CI)	p
Age, median	52 (37, 66)	41 (32, 54)	63 (50.25, 74)		
Age, years					
0-18	58 (1.3)	25 (1.1)	33 (1.4)	4.39 (2.46-7.94)	<0.001
19-44	1588 (34.6)	1213 (55.0)	375 (15.7)	1.00 (0.00-0.00)	<0.001
45-54	778 (16.9)	417 (18.9)	361 (15.1)	2.53 (2.07-3.09)	<0.001
55-64	885 (19.3)	356 (16.2)	529 (22.1)	4.18 (3.42-5.13)	<0.001
65-74	678 (14.8)	152 (6.9)	526 (22.0)	10.12 (7.89-13.05)	<0.001
≥75	607 (13.2)	41 (1.9)	566 (23.7)	48.53 (33.65-71.59)	<0.001
Male	2317 (50.4)	828 (37.6)	1,489 (62.3)	2.89 (2.48-3.36)	<0.001
Race/ethnicity					
Non-Hispanic white	1824 (39.7)	837 (38.0)	987 (41.3)	Reference	
Non-Hispanic African American	710 (15.5)	391 (17.7)	319 (13.3)	0.83 (0.66-1.04)	0.114
Asian	338 (7.4)	181 (8.2)	157 (6.6)	1.37 (1.02-1.83)	0.037
Hispanic	1128 (24.6)	506 (23.0)	622 (26.0)	1.57 (1.29-1.90)	<0.001
Other/ Multiracial	296 (6.4)	121 (5.5)	175 (7.3)	1.87 (1.39-2.54)	<0.001
Unknown	298 (6.5)	168 (7.6)	130 (5.4)	1.26 (0.92-1.73)	0.146
Tobacco use					
Never, including 10 passive exposure	2776 (60.4)	1386 (59.1)	1390 (58.2)	Reference	
Former	764 (16.6)	278 (11.9)	486 (20.3)	0.67 (0.54-0.83)	<0.001
Current	238 (5.2)	121 (5.2)	117 (4.9)	0.60 (0.42-0.84)	0.003
Unknown	816 (17.8)	419 (19.0)	397 (16.6)	1.31 (1.07-1.62)	0.011
Obesity					
BMI <25 kg/m^2	1227 (26.7)	643 (29.2)	584 (24.4)	Reference	
BMI 25 to <30 kg/m^2	1489 (32.4)	676 (30.7)	813 (34.0)	1.33 (1.09-1.62)	0.005
BMI 30 to <40 kg/m^2	1304 (28.4)	537 (24.4)	767 (32.1)	1.85 (1.51-2.28)	<0.001
BMI ≥ 40 kg/m^2	265 (5.8)	102 (4.6)	163 (6.8)	2.73 (1.96-3.81)	<0.001
Unknown	309 (6.7)	246 (11.2)	63 (2.6)	0.30 (0.21-0.43)	<0.001
Any chronic condition*	1876 (40.8)	513 (27.4)	1363 (57.0)		
Any cardiovascular condition*	1455 (31.7)	377 (17.1)	1078 (45.1)		
Coronary artery disease	254 (5.5)	37 (1.7)	217 (9.1)	0.86 (0.56-1.34)	0.504
Heart failure	168 (3.7)	10 (0.5)	158 (6.6)	3.61 (1.81-7.94)	<0.001
Hyperlipidemia	866 (18.9)	241 (10.3)	625 (26.2)	0.63 (0.50-0.80)	<0.001
Hypertension	1165 (25.4)	267 (12.1)	898 (37.6)	1.40 (1.12-1.74)	0.003

Confidential: For Review Only

Characteristic	Overall, N=4,594 N (%) or median (IQR)	Not hospitalized, N=2,204 N (%) or median (IQR)	Hospitalized, N=2,390 N (%) or median (IQR)	Risk of hospitalization Odds ratio (95% CI)	p
Diabetes	749 (16.3)	128 (5.8)	621 (26.0)	2.96 (2.32-3.81)	<0.001
Asthma or chronic obstructive pulmonary disorder	363 (7.9)	117 (5.3)	246 (10.3)	1.45 (1.08-1.95)	0.014
Chronic kidney disease	278 (6.1)	24 (1.1)	254 (10.6)	3.18 (1.97-5.32)	<0.001
Cancer	213 (4.6)	45 (2.0)	168 (7.0)	1.29 (0.87-1.94)	0.216
Temperature at presentation, degrees Celsius*	37.4 (36.9 – 38.2)	37.3 (36.9 – 37.8)**	37.5 (36.9 – 38.2)		
Temperature ≥ 38^{o} C at presentation*	892 (19.4)	114 (5.2)**	778 (32.6)		
Oxygen saturation at presentation*	95 (91 – 97)	97 (96-99) †	94 (90-96) ‡		
Oxygen saturation < 88% at presentation*	357 (7.8)	2 (0.1)	355 (14.9)		

*Not included in multivariable model

**Missing for 74% of non-hospitalized patients

†Measured on supplemental oxygen for 23% of patients

‡Measured on supplemental oxygen for 99% of patients

Confidential: For Review Only

Table 2: Characteristics of admitted patients, by complication status, among those with definitive outcomes

Characteristic	All hospitalized with outcomes, N=2,275 N (%) or median (IQR)	Discharged, no critical illness, N=1,416 N (%) or median (IQR)	Critical illness, N=859 N (%) or median (IQR)
Age, median years	62 (50 – 74)	60 (47-71)	67 (57-78)
Age, years			
0-18	32 (1.4)	21 (1.5)	11 (1.3)
19-44	361 (15.9)	277 (19.6)	84 (9.8)
45-54	346 (15.2)	259 (18.3)	87 (10.1)
55-64	500 (22.0)	315 (22.3)	185 (21.5)
65-74	494 (21.7)	275 (19.4)	219 (25.5)
≥75	542 (23.8)	269 (19.0)	273 (31.8)
Male sex	1423 (62.5)	852 (60.2)	571 (66.5)
Race/ethnicity			
Non-Hispanic white	940 (41.3)	558 (39.4)	382 (44.5)
Non-Hispanic African American	306 (13.5)	211 (14.9)	95 (11.1)
Asian	149 (6.5)	82 (5.8)	67 (7.8)
Hispanic	588 (25.8)	395 (27.9)	193 (22.5)
Other/Multiracial	167 (7.3)	97 (6.9)	70 (8.1)
Unknown	125 (5.5)	73 (5.2)	52 (6.1)
Tobacco use			
Never, including 3 passive exposure	1326 (58.3)	866 (61.2)	460 (53.6)
Former	455 (20.0)	264 (18.6)	191 (22.2)
Current	117 (5.1)	84 (5.9)	33 (3.8)
Unknown	377 (16.6)	202 (14.3)	175 (20.4)
Obesity			
BMI <25 kg/m^2	552 (24.3)	329 (23.3)	223 (26.0)
BMI 25.0-29.9 kg/m^2	770 (33.8)	497 (35.1)	273 (31.8)
BMI 30.0-39.9 kg/m^2	733 (32.2)	464 (32.8)	269 (31.3)
BMI >=40 kg/m^2	159 (7.0)	93 (6.6)	66 (7.7)
Unknown	61 (2.7)	33 (2.3)	28 (3.3)
Any chronic condition*	1291 (56.8)	783 (55.3)	508 (59.1)
Any cardiovascular condition*	1020 (44.8)	604 (42.7)	416 (48.4)
Coronary artery disease	207 (9.1)	99 (7.0)	108 (12.6)
Heart failure	152 (6.7)	70 (4.9)	82 (9.5)
Hyperlipidemia	588 (25.8)	355 (25.1)	233 (27.1)
LDL, median	82 (58 – 110)	85 (62 – 116)	76 (54 – 101)
Hypertension	850 (37.4)	500 (35.3)	350 (40.7)
Diabetes	584 (25.7)	339 (23.9)	245 (28.5)
Cancer	160 (7.0)	88 (6.2)	72 (8.4)

Confidential: For Review Only

Characteristic	All hospitalized with outcomes, N=2,275 N (%) or median (IQR)	Discharged, no critical illness, N=1,416 N (%) or median (IQR)	Critical illness, N=859 N (%) or median (IQR)
Asthma or chronic obstructive pulmonary disorder	232 (10.2)	138 (9.8)	94 (10.9)
Chronic kidney disease	235 (10.3)	127 (9.0)	108 (12.6)
Temperature at presentation, degrees Celsius	37.5 (36.9 – 38.3)	37.5 (37.0 – 38.2)	37.5 (36.9 – 38.3)
Temperature ≥ 38^{o} C at presentation	1227 (53.9)	456 (32.2)	291 (33.9)
Oxygen saturation at presentation	94 (91 – 97)	95 (92-97) †	92 (86-95) ‡
Oxygen saturation < 88 percent at presentation	326 (14.3)	86 (6.1) ‡	240 (27.9) ‡
First absolute lymphocyte count, 10^3/µl	0.8 (0.6 – 1.2)	0.9 (0.6-1.2)	0.8 (0.5-1.1)
Missing lymphocyte count	82 (3.6)	54 (3.8)	28 (3.3)
First creatinine, mg/dL	0.98 (0.80 – 1.32)	0.93 (.77-1.21)	1.07 (.83-1.54)
Missing creatinine	148 (6.5)	108 (7.6)	40 (4.7)
First alanine aminotransferase, units/L*	34 (23 – 54)	33 (23-54)	35 (24-54)
Missing alanine aminotransferase	214 (9.4)	166 (11.7)	48 (5.6)
First aspartate aminotransferase, units/L*	46 (31 – 69)	42 (29.0-61.0)	53 (35.3-79.8)
Missing aspartate aminotransferase	246 (10.8)	181 (12.8)	65 (7.8)
First C-reactive protein, mg/L	110.0 (56.0 – 170.2)	87.5 (42.1-145.2)	139.7 (86.5-205.0)
Missing C-reactive protein	277 (12.2)	199 (14.1)	78 (9.1)
First d-dimer, ng/mL	375 (235 – 673)	315 (204 – 529)	508 (307 - 1038)
Missing d-dimer	462 (20.3)	340 (24.0)	122 (14.2)
First ferritin, ng/mL	730.8 (362.0 – 1487.0)	628.0 (316.0-1320.3)	931.0 (472.3-1706.4)
Missing ferritin	329 (14.5)	237 (16.7)	92 (10.7)
First procalcitonin, ng/mL	0.14 (0.07 – 0.40)	0.10 (0.05-0.23)	0.28 (0.10-0.80)
Missing procalcitonin	306 (13.5)	232 (16.4)	74 (8.6)
First troponin-I, ng/mL	0.03 (0.01 – 0.10)	0.02 (0.01-0.10)	0.07 (0.01-0.10)
Missing troponin-I	309 (13.6)	234 (16.5)	75 (8.7)
Length of stay, days*	6.9 (3.8 – 12.0)	5.7 (3.6 – 8.8)	[Discharged or Died] 7.6 (4.6-11.8) [Still Hospitalized] 17 (13-21)

*Not included in multivariable model

†Measured on supplemental oxygen for 100% of patients

‡Measured on supplemental oxygen for 99% of patients

Confidential For Review Only

Table 3: Multivariable logistic regression results for risk of critical illness

Characteristic	Among all patients (N=4,594)		Among hospitalized patients, excluding vitals and laboratory results (N=2,275)		Among hospitalized patients, including vitals and laboratory results (N=2,275)	
	Odds ratio (95% CI)	p	Odds ratio (95% CI)	p	Odds ratio (95% CI)	p
Age, years						
0-18	3.46 (1.62-6.82)	<0.001	1.49 (0.66-3.23)	0.322	4.15 (1.70-9.81)	0.001
19-44	1.00 (0.00-0.00)	<0.001	Reference		Reference	
45-54	2.03 (1.47-2.79)	<0.001	1.06 (0.75-1.51)	0.742	0.80 (0.53-1.19)	0.264
55-64	4.25 (3.21-5.68)	<0.001	1.93 (1.41-2.66)	<0.001	1.35 (0.94-1.95)	0.107
65-74	7.66 (5.72-10.33)	<0.001	2.72 (1.97-3.77)	<0.001	1.88 (1.29-2.75)	0.001
≥75	14.17 (10.45-19.37)	<0.001	3.54 (2.53-4.98)	<0.001	2.73 (1.85-4.05)	<0.001
Male	2.08 (1.75-2.48)	<0.001	1.45 (1.19-1.75)	<0.001	1.08 (0.86-1.36)	0.516
Race/ethnicity						
Non-Hispanic white	Reference)		Reference		Reference	
Non-Hispanic African American	0.76 (0.58-1.00)	0.050	0.74 (0.55-0.99)	0.043	0.68 (0.48-0.95)	0.027
Asian	1.62 (1.15-2.25)	0.005	1.44 (0.99-2.09)	0.057	1.40 (0.90-2.15)	0.130
Hispanic	1.13 (0.91-1.41)	0.272	0.92 (0.72-1.16)	0.460	0.83 (0.63-1.08)	0.167
Other/ Multiracial	1.66 (1.20-2.29)	0.002	1.31 (0.91-1.86)	0.141	1.51 (1.00-2.26)	0.047
Unknown	1.29 (0.88-1.85)	0.179	1.09 (0.73-1.63)	0.673	1.11 (0.69-1.76)	0.671
Tobacco use						
Never	Reference)		Reference		Reference	
Former	0.85 (0.68-1.05)	0.141	1.01 (0.79-1.28)	0.951	0.96 (0.73-1.26)	0.763
Current	0.58 (0.37-0.87)	0.012	0.61 (0.39-0.95)	0.032	0.62 (0.37-1.02)	0.064
Unknown	1.65 (1.31-2.08)	<0.001	1.55 (1.21-1.99)	<0.001	1.33 (0.99-1.78)	0.055
Obesity						
BMI <25 kg/m^2	Reference)		Reference		Reference	
BMI 25.0-29.9 kg/m^2	1.02 (0.81-1.27)	0.885	0.94 (0.74-1.20)	0.613	1.01 (0.76-1.33)	0.964
BMI 30.0-39.9 kg/m^2	1.45 (1.15-1.83)	0.002	1.19 (0.93-1.53)	0.175	1.34 (1.00-1.79)	0.051
BMI >=40 kg/m^2	2.52 (1.75-3.62)	<0.001	1.78 (1.20-2.65)	0.004	1.90 (1.20-3.00)	0.006
Unknown	0.54 (0.33-0.84)	0.009	1.13 (0.64-2.00)	0.674	1.33 (0.67-2.59)	0.410
Coronary artery disease	1.22 (0.89-1.67)	0.220	1.31 (0.94-1.82)	0.113	1.44 (0.99-2.10)	0.058
Heart failure	1.82 (1.25-2.65)	0.002	1.53 (1.04-2.26)	0.029	1.28 (0.82-2.00)	0.272
Hyperlipidemia	0.70 (0.56-0.89)	0.003	0.82 (0.64-1.04)	0.108	0.85 (0.64-1.12)	0.255

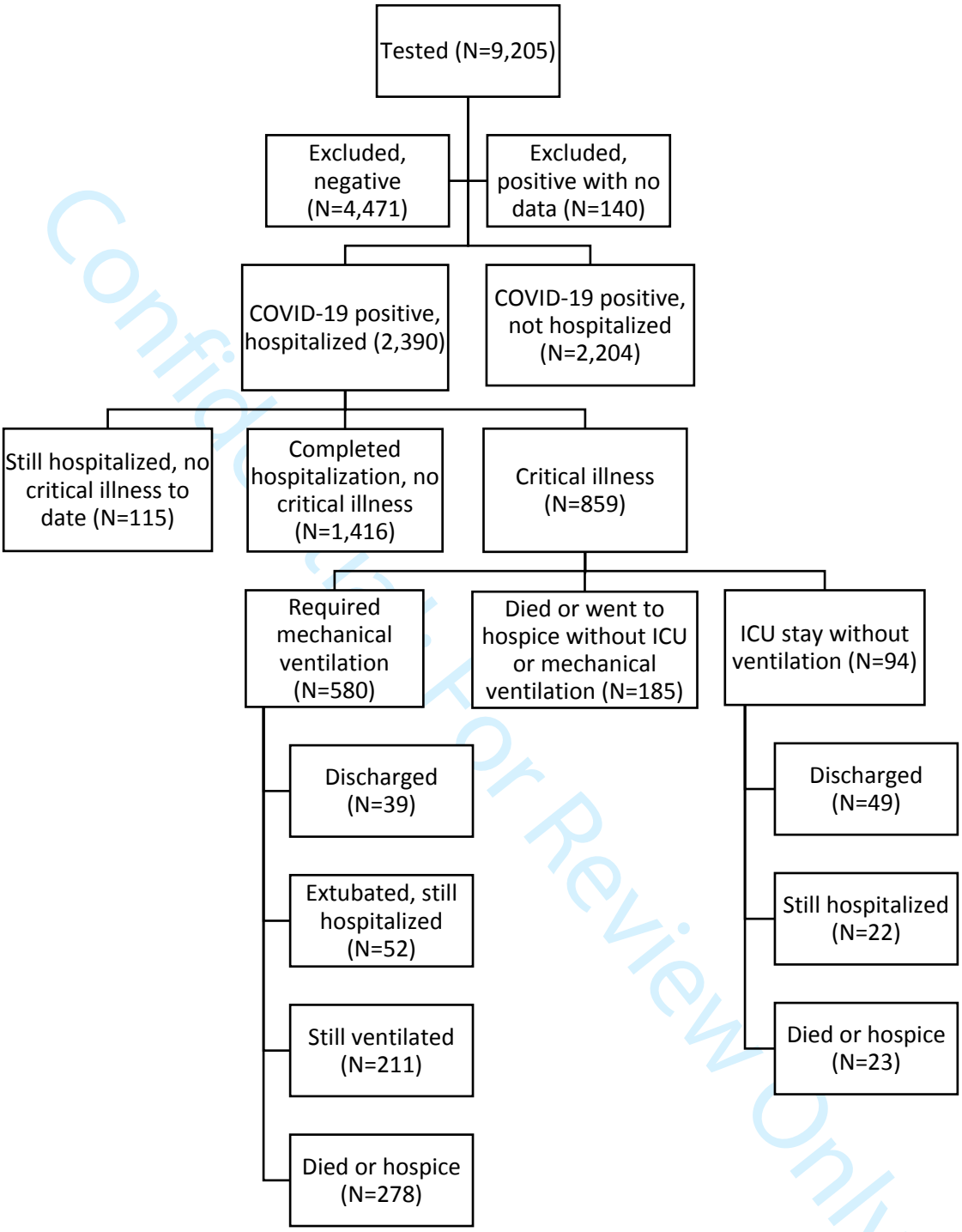
Confidential: For Review Only

Characteristic	Among all patients (N=4,594)		Among hospitalized patients, excluding vitals and laboratory results (N=2,275)		Among hospitalized patients, including vitals and laboratory results (N=2,275)	
	Odds ratio (95% CI)	p	Odds ratio (95% CI)	p	Odds ratio (95% CI)	p
Hypertension	1.10 (0.88-1.37)	0.400	0.94 (0.75-1.18)	0.593	0.95 (0.73-1.24)	0.724
Diabetes	1.48 (1.19-1.85)	<0.001	1.19 (0.95-1.49)	0.129	1.20 (0.93-1.55)	0.168
Asthma or chronic obstructive pulmonary disorder	1.12 (0.83-1.49)	0.461	1.01 (0.74-1.37)	0.961	1.17 (0.83-1.65)	0.374
Chronic kidney disease	1.23 (0.90-1.68)	0.183	1.18 (0.85-1.62)	0.318	0.76 (0.50-1.17)	0.219
Cancer	1.19 (0.85-1.64)	0.303	1.13 (0.80-1.59)	0.495	1.10 (0.74-1.63)	0.645
Temperature on presentation, degrees Celsius						
<38					Reference	
38-39					1.16 (0.90-1.49)	0.263
>39					1.29 (0.93-1.78)	0.124
Oxygen saturation on presentation, %						
SpO2 >92					Reference	
SpO2 88-92					1.71 (1.33-2.19)	<0.001
SpO2 <88					4.34 (3.19-5.94)	<0.001
First lymphocyte count, 10^3/µl						
>1.2					Reference	
>0.8-1.2					1.01 (0.74-1.38)	0.952
0.4-0.8					1.13 (0.85-1.51)	0.406
<0.4					1.93 (1.32-2.83)	<0.001
Missing					2.57 (1.11-6.21)	0.031
First creatinine, mg/dL						
0-1.2					Reference	
>1.2-2					1.14 (0.86-1.51)	0.375
>2					1.28 (0.81-2.04)	0.293
Missing					0.77 (0.32-1.80)	0.557
First C reactive protein, mg/L						
0-15					Reference	
>15-100					2.27 (1.28-4.23)	0.007
>100-200					3.71 (2.08-7.00)	<0.001
>200					4.88 (2.60-9.61)	<0.001

Confidential: For Review Only

Characteristic	Among all patients (N=4,594)		Among hospitalized patients, excluding vitals and laboratory results (N=2,275)		Among hospitalized patients, including vitals and laboratory results (N=2,275)	
	Odds ratio (95% CI)	p	Odds ratio (95% CI)	p	Odds ratio (95% CI)	p
Missing					2.74 (1.23-6.30)	0.015
First d-dimer, ng/mL						
0-250					Reference	
>250-500					1.83 (1.36-2.48)	<0.001
>500-1000					2.45 (1.74-3.46)	<0.001
>1000-2500					2.70 (1.67-4.37)	<0.001
>2500					3.88 (2.31-6.60)	<0.001
Missing					1.75 (1.18-2.60)	0.005
First ferritin, ng/mL						
0-300					Reference	
>300-500					1.12 (0.76-1.65)	0.578
>500-1000					1.11 (0.79-1.57)	0.542
>1000-2500					1.09 (0.76-1.55)	0.644
>2500					1.16 (0.74-1.82)	0.506
Missing					1.48 (0.85-2.59)	0.169
First procalcitonin, ng/mL						
0-0.5					Reference	
>0.5					1.92 (1.43-2.59)	<0.001
Missing					0.56 (0.33-0.94)	0.030
First troponin, ng/mL						
<0.1					Reference	
0.1-1					2.25 (1.56-3.27)	<0.001
>1					5.54 (2.42-14.16)	<0.001
Missing					0.84 (0.52-1.35)	0.479

Confidential: For Review Only



Confidential: For Review Only

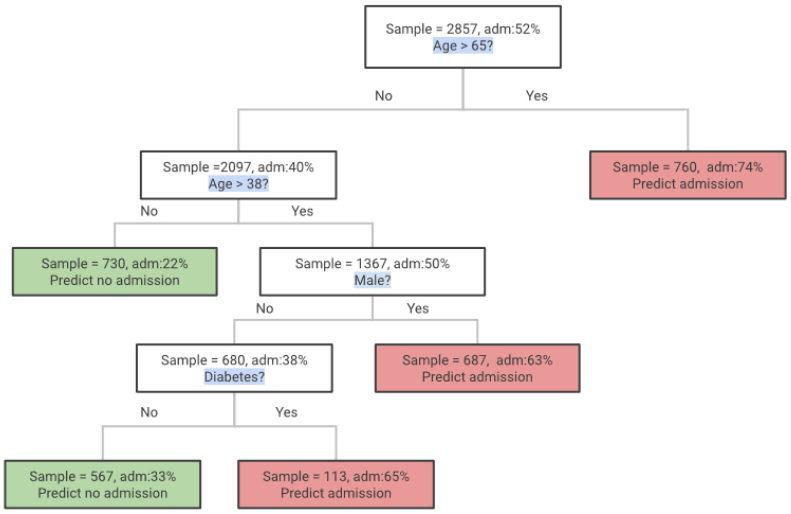


Figure 2A: Maximum likelihood decision tree classification for hospitalization

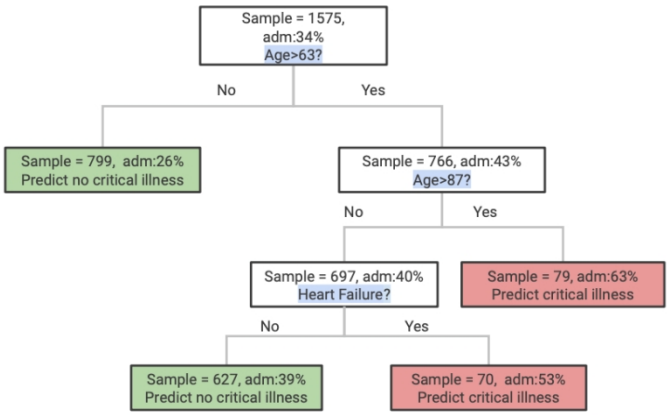


Figure 2B: Maximum likelihood decision tree classification for critical illness, without vital signs and laboratory results

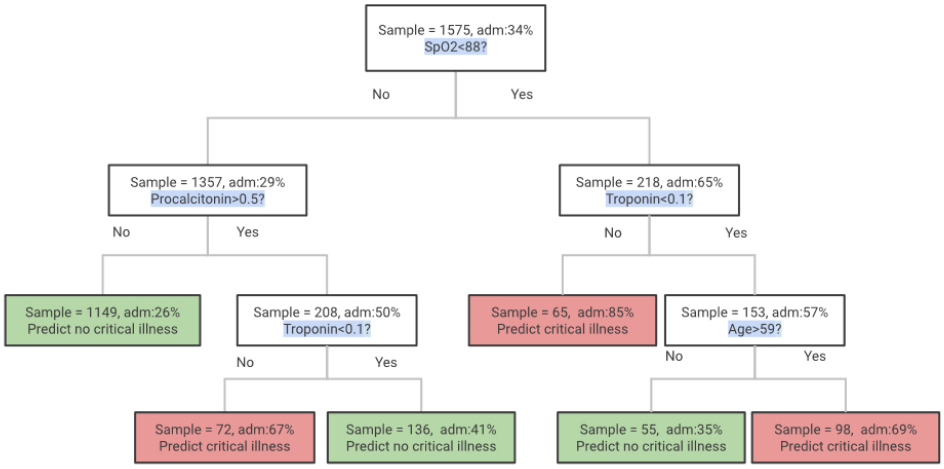


Figure 2C: Maximum likelihood decision tree classification for critical illness, with vital signs and laboratory results