

# Identification of Phenotypes Among COVID-19 Patients in the United States Using Latent Class Analysis

Catherine Teng<sup>1</sup>  
 Unnikrishna Thampy<sup>1</sup>  
 Ju Young Bae<sup>1</sup>  
 Peng Cai<sup>2</sup>  
 Richard AF Dixon<sup>3</sup>  
 Qi Liu<sup>3</sup>  
 Pengyang Li<sup>4</sup>

<sup>1</sup>Department of Medicine, Yale New Haven Health – Greenwich Hospital, Greenwich, CT, USA; <sup>2</sup>Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, MA, USA; <sup>3</sup>Molecular Cardiology Research, Texas Heart Institute, Houston, TX, USA; <sup>4</sup>Division of Cardiology, Pauley Heart Center, Virginia Commonwealth University, Richmond, VA, USA

Correspondence: Pengyang Li  
 Pauley Heart Center, Virginia Commonwealth University, 1200 E Marshall St, Richmond, VA, 23219, USA  
 Tel +1 626-420-5811  
 Fax +1 832-355-9692  
 Email leelpy0109@gmail.com

Qi Liu  
 Texas Heart Institute, 6770 Bertner Avenue, MC 2-255, Houston, TX, 77030, USA  
 Tel +1 832-355-8006  
 Fax +1 832-355-9692  
 Email QLiu@texasheart.org

**Background:** Coronavirus disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2 or COVID-19) is a heterogeneous disorder with a complex pathogenesis. Recent studies from Spain and France have indicated that underlying phenotypes may exist among patients admitted to the hospital with COVID-19. Whether those same phenotypes exist in the United States (US) remains unclear. Using latent class analysis (LCA), we sought to determine whether clinical phenotypes exist among patients admitted for COVID-19.

**Methods:** We reviewed the charts of adult patients who were hospitalized primarily for COVID-19 at Greenwich Hospital and performed LCA using variables based on patient demographics and comorbidities. To further examine the reliability and replicability of the clustering results, we repeated LCA on the cohort of patients who died during hospitalization for COVID-19.

**Results:** Two phenotypes were identified in patients admitted for COVID-19 (N = 483). According to phenotype, patients were designated as cluster 1 (C1) or cluster 2 (C2). C1 (n = 193) consisted of older individuals with more comorbidities and a higher mortality rate (25.4% vs 8.97%, p < 0.001) than patients in C2. C2 (n = 290) consisted of younger individuals who were more likely to be obese, male, and nonwhite, with higher levels of the inflammatory markers C-reactive protein and alanine aminotransferase. When we performed LCA on the cohort of patients who died during hospitalization for COVID-19 (n = 75), we found that the distribution of patient baseline characteristics and comorbidities was similar to that of the entire cohort of patients admitted for COVID-19.

**Conclusion:** Using LCA, we identified two clinical phenotypes of patients who were admitted to our hospital for COVID-19. These findings may reflect different pathophysiologic processes that lead to moderate to severe COVID-19 and may be useful for identifying treatment targets and selecting patients with severe COVID-19 disease for future clinical trials.

**Keywords:** COVID-19 (United States), hospitalization, death, phenotypes, latent class analysis, comorbidities

## Introduction

Since January 2020, considerable emphasis has been placed on identifying distinct risk factors in patients with COVID-19. Certain risk factors,<sup>1–3</sup> such as age, obesity, sex, and other underlying comorbidities, have been extensively described in patients with COVID-19,<sup>1–3</sup> and a recent study has indicated that patients can be grouped on the basis of those risk factors.<sup>4</sup> A retrospective study conducted in Spain showed that patients who were admitted with COVID-19 could be grouped according to three distinct phenotypes that correlate with mortality.<sup>4</sup> In the first

group, patients were younger, were less frequently male, and had normal levels of inflammatory parameters; in the second group, patients were obese with moderately elevated levels of inflammatory markers; in the third group, patients were older with more comorbidities.<sup>4</sup> Another retrospective study of 85 patients with COVID-19 in Paris, France reported similar phenotypes in patients who were admitted to the intensive care unit.<sup>5</sup> Although focus is currently directed toward the relevance of phenotypes, it remains unclear whether those same phenotypes exist in the United States (US), given the significant differences in mortality rates between the US and major European countries.<sup>6–9</sup> In addition, patient populations are genetically and ethnically heterogeneous across these continents, and the virus strains are potentially different. In this study, we conducted latent class analysis (LCA) of patients who were admitted to Greenwich Hospital, Yale New Haven Health with COVID-19 and sought to identify whether clinical phenotypes exist among this US patient population.

## Materials and Methods

### Database and Study Population

We reviewed the charts of adult patients who were hospitalized for COVID-19 from February 1 to May 30, 2020 at Greenwich Hospital, Yale New Haven Health. Diagnoses were made on the basis of SARS-CoV-2 PCR test results. Patients were excluded who were younger than 18, were pregnant at the time of diagnosis, or had incomplete hospitalization records. This study received ethics approval by the institutional review board of Greenwich Hospital, Yale New Haven Health. Data acquired from electronic medical records were de-identified, encrypted, and stored securely. The requirement to obtain patient informed consent to review their medical records was waived because of de-identification. This study was conducted in accordance with the Declaration of Helsinki.

### Study Variables

We retrieved data on well-documented prognostic factors,<sup>1–3,10</sup> including patient demographics (age, sex, race, and body mass index [BMI]), comorbidities (hypertension, diabetes mellitus, coronary artery disease [CAD], chronic heart failure [CHF], chronic kidney disease [CKD], neurological disease, and pre-existing respiratory disease), in-hospital laboratory results, and treatment information (medications received during hospitalization,

hospital length of stay, maximized oxygen requirement, and requirement for intubation). During chart review, patient information was de-identified for statistical analysis.

## Statistical Analysis

We first analyzed the characteristics of each cohort by using a *t*-test, Fisher's exact test, or a  $\chi^2$  test, depending on the nature of the variable. A  $p < 0.05$  result was considered statistically significant. LCA was first applied to the entire cohort of patients who were admitted to the hospital for COVID-19. To further examine the reliability and replicability of the clustering results, we repeated LCA on the cohort of patients who died during hospitalization for COVID-19.

LCA is a statistical method used to identify classes of individuals on the basis of known categorical variables.<sup>11</sup> In traditional LCA, class membership probabilities and item-response probabilities conditional on class membership are estimated.<sup>11</sup>

We examined our data by using LCA to determine the number of clusters that best fit the data.<sup>12</sup> We calculated the item response probabilities (IRP) in the datasets to define each cluster's characteristics. Starting with a model that had two classes, we increased the number of classes up to seven to determine whether the set of available diagnostic methods suggested an optimal number of classes. By examining the number of parameters, log-likelihood, Bayesian information criterion (BIC), sample-size-adjusted BIC, Akaike's information criterion, Pearson  $\chi^2$  goodness of fit, and likelihood ratio  $\chi^2$  (G2) statistic and entropy, we found that having two classes presented the most parsimonious solution, considering goodness-of-fit measures and the interpretability of the model.

## Results

### Population Characteristics

Our study included 483 patients admitted to Greenwich Hospital, Yale New Haven Health from February 1 to May 20, 2020. Overall, 46.8% of patients admitted for COVID-19 were 65 or older (average age=64 years), 63.1% were men, and 46.8% were white. The average BMI was 28.2 kg/m<sup>2</sup>. Patient comorbidities included HTN (49.1%), obesity (35.4%), DM (25.9%), previous gastrointestinal disease (17.8%), pre-existing respiratory disease (16.6%), CAD (12.8%), CKD (11.8%), and CHF (8.3%). Patient characteristics are shown in detail in [Table 1](#).

**Table 1** Patient Characteristics in Risk Factor Clusters Identified by Performing Latent Class Analysis of Patients Who Were Admitted for COVID-19

Characteristics	All	C1	C2	P value
	N=483	n=193	n=290	
<b>Risk factors</b>				
Age	64.0 ± 17.9	79.5 ± 9.88	53.7 ± 14.3	<0.001
Age ≥65 years	226 (46.8%)	180 (93.3%)	46 (15.9%)	<0.001
Male	305 (63.1%)	101 (52.3%)	204 (70.3%)	<0.001
Nonwhite	257 (53.2%)	55 (28.5%)	202 (69.7%)	<0.001
BMI ≥30 kg/m <sup>2</sup>	171 (35.4%)	54 (28.0%)	117 (40.3%)	0.007
BMI, kg/m <sup>2</sup>	28.2 ± 5.88	26.8 ± 5.88	29.2 ± 5.70	<0.001
CAD	62 (12.8%)	62 (32.1%)	0 (0.0%)	<0.001
HTN	237 (49.1%)	171 (88.6%)	66 (22.8%)	<0.001
CHF	40 (8.28%)	40 (20.7%)	0 (0.0%)	<0.001
DM	125 (25.9%)	74 (38.3%)	51 (17.6%)	<0.001
CKD	57 (11.8%)	49 (25.4%)	8 (2.76%)	<0.001
Neurological disease	72 (14.9%)	62 (32.1%)	10 (3.45%)	<0.001
Gastrointestinal disease	86 (17.8%)	49 (25.4%)	37 (12.8%)	0.001
Pre-existing respiratory disease	80 (16.6%)	48 (24.9%)	32 (11.0%)	<0.001
<b>Treatments</b>				
Tocilizumab	147 (30.4%)	43 (22.3%)	104 (35.9%)	0.002
Protease Inhibitors	176 (36.4%)	55 (28.5%)	121 (41.7%)	0.004
Convalescent plasma	14 (2.90%)	0 (0.00%)	14 (4.83%)	0.005
Remdesivir	18 (3.73%)	3 (1.55%)	15 (5.17%)	0.07
Anticoagulation	162 (33.5%)	60 (31.1%)	102 (35.2%)	0.41
Hemodialysis	20 (4.14%)	6 (3.11%)	14 (4.83%)	0.34
Intubation	73 (15.1%)	28 (14.5%)	45 (15.5%)	0.61
Oxygen	0.43 ± 0.26	0.45 ± 0.27	0.42 ± 0.25	0.24
<b>Laboratory data</b>				
Creatinine, mg/dL	1.52 ± 1.42	1.86 ± 1.57	1.29 ± 1.27	<0.001
Troponin, ng/mL	1.08 ± 8.93	1.71 ± 11.5	0.64 ± 6.63	0.25
Pro-BNP, pg/mL	2720 ± 6131	3541 ± 6035	1959 ± 6142	0.04
Ferritin, ng/mL	1949 ± 5471	1644 ± 5235	2152 ± 5622	0.33
CRP, mg/L	15.3 ± 35.0	10.7 ± 13.7	18.3 ± 43.4	0.006
D-dimer, mg/L	5.95 ± 9.17	5.82 ± 8.52	6.03 ± 9.58	0.82
ALT, U/L	164 ± 421	96.2 ± 204	209 ± 512	0.001
AST, U/L	203 ± 805	127 ± 436	253 ± 970	0.05
<b>Outcomes</b>				
Length of stay, days	10.5 ± 13.7	9.57 ± 10.1	11.2 ± 15.7	0.18
Death	75 (15.5%)	49 (25.4%)	26 (8.97%)	<0.001

**Note:** Data are presented as the mean ± standard deviation or as the number (%).

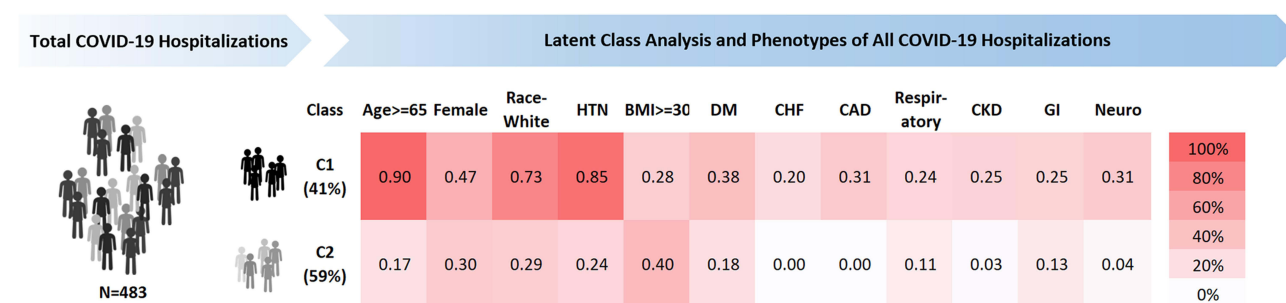
**Abbreviations:** ALT, alanine transaminase; AST, aspartate aminotransferase; BMI, body mass index; CAD, coronary artery disease; CHF, chronic heart failure; CKD, chronic kidney disease; CRP, C-reactive protein; DM, diabetes mellitus; HTN, hypertension; pro-BNP, pro-B-type natriuretic peptide.

## Identification of Phenotypes by Clustering

By performing cluster analysis of the variables using patient demographics and comorbidities, we identified two clusters that provided the best fit among patients who were admitted for COVID-19: cluster 1 (C1; n=193) and cluster 2 (C2; n=290) (Table 1 and Figure 1).

## C1 Cohort: Older Patients with More Comorbidities

In C1, patients were older than in C2 (79.5 vs 53.7 years,  $p<0.001$ ), with 93.3% of patients 65 or older in C1 and only 15.9% 65 or older in C2. C1 also had a lower proportion of men (52.3% vs 70.3%,  $p<0.001$ ) and nonwhite individuals (28.5% vs 69.7%,  $p<0.001$ ) than did C2. The



**Figure 1** Probability of individual risk factors in the two clusters among total COVID-19 hospitalizations.

**Abbreviations:** BMI, body mass index (kg/m<sup>2</sup>); CAD, coronary artery disease; CHF, chronic heart failure; CKD, chronic kidney disease; DM, diabetes mellitus; GI, gastrointestinal disease; HTN, hypertension; neuro, neurological disease; respiratory, pre-existing respiratory disease.

overall mortality rate in C1 was higher than that in C2 (25.4% vs 8.97%,  $p<0.001$ ).

C1 had a significantly higher proportion of comorbidities than did C2, including HTN (88.6% vs 22.8%,  $p<0.001$ ), CAD (32.1% vs 0.0%,  $p<0.001$ ), CHF (20.7% vs 0%,  $p<0.001$ ), DM (38.3 vs 17.6%,  $p<0.001$ ), CKD (25.4% vs 2.8%,  $p<0.001$ ), pre-existing respiratory disease (24.9% vs 11.0%,  $p=0.004$ ), and pre-existing neurological disease (32.1% vs 3.5%,  $p<0.001$ ).

C1 also had higher peak levels of creatinine ( $1.86 \pm 1.57$  vs  $1.29 \pm 1.27$  mg/dL,  $p<0.001$ ) and pro-natriuretic peptide (pro-BNP) ( $3541 \pm 6035$  vs  $1959 \pm 6142$  pg/mL,  $p=0.036$ ) but lower levels of the inflammatory markers C-reactive protein (CRP) ( $10.7 \pm 13.7$  mg/L vs  $18.3 \pm 43.4$ ,  $p=0.006$ ) and alanine aminotransferase (ALT) ( $96.2 \pm 204$  vs  $209 \pm 512$  U/L,  $p=0.001$ ) than did C2.

### C2 Cohort: Younger Patients with Obesity

C2 patients were younger than C1 patients (53.7 vs 79.5 years,  $p<0.001$ ) and were more likely to be obese (BMI  $\geq 30$  kg/m<sup>2</sup>; 40.3% vs 28.0%,  $p=0.007$ ), male (70.3% vs 52.3%,  $p<0.001$ ), and nonwhite (69.7% vs 28.5%,  $p<0.001$ ). The IRPs of each risk factor in C1 and C2 are shown in [Figure 1](#).

Overall, C2 had higher levels of the inflammatory markers CRP and ALT than did C2. Levels of other markers were also higher in C2 than in C1, including ferritin ( $2152 \pm 5622$  vs  $1644 \pm 5235$  ng/L,  $p=0.33$ ), D-dimer ( $6.03 \pm 9.58$  vs  $5.82 \pm 8.52$  mg/L,  $p=0.815$ ), and aspartate aminotransferase ( $253 \pm 970$  vs  $127 \pm 436$  U/L,  $p=0.054$ ), but the differences were not statistically significant.

## LCA of Patients Who Died from COVID-19

Of the 483 patients who were hospitalized for COVID-19, 75 died during hospitalization. For the cohort of 75

deceased patients, we independently applied the latent class model. The analysis identified two clusters that we designated as cluster 1' (C1') and cluster 2' (C2'), which were analogues to those identified in the LCA for the entire cohort of 483 patients admitted for COVID-19 ([Supplemental Table 1](#)).

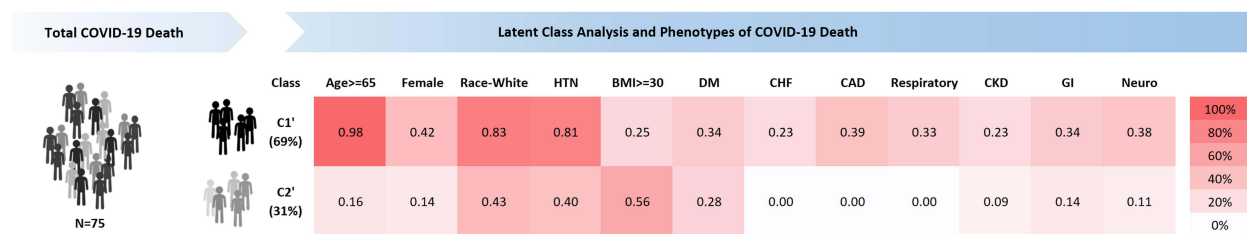
### Overall Characteristics

The distribution of baseline characteristics and comorbidities in the cohort of patients who died during hospitalization for COVID-19 was similar to that observed for the entire cohort of patients who were admitted for COVID-19. The patient characteristics of the 75 deceased patients are shown in [Supplemental Table 1](#).

### Identification of C1' and C2' by Clustering

Patients in C1' exhibited characteristics similar to those in C1. C1' patients were older than patients in C2' (83.4 vs 54.9 years,  $p<0.001$ ) and had a higher proportion of comorbidities such as HTN (80.8% vs 39.1%,  $p=0.001$ ), CAD (38.5% vs 0%,  $p<0.001$ ), CHF (23.1% vs 0%,  $p=0.01$ ), and pre-existing respiratory disease (32.7% vs 0%,  $p<0.001$ ). No statistically significant difference was observed between C1' and C2' in the proportion of patients with DM (32.7% vs 30.4%,  $p=1.00$ ), CKD (23.1% vs 8.7%,  $p=0.20$ ), neurological disease (36.5% vs 13.0%,  $p=0.05$ ), or gastrointestinal disease (34.6% vs 13.0%,  $p=0.09$ ).

C2' was comparable to C2, consisting of patients who were younger and more likely to be obese (56.5% vs 25.0%,  $p=0.02$ ), male (87.0% vs 57.7%,  $p=0.02$ ), and nonwhite (60.9% vs 15.4%,  $p<0.001$ ) than were patients in C1'. The IRPs of each risk factor in C1' and C2' are shown in [Figure 2](#). Compared with patients in C1', patients in C2' showed a pattern of increased expression of the inflammation markers CRP ( $84.4 \pm 117$  vs  $12.6 \pm 10.8$  mg/L,  $p=0.008$ ) and D-dimer ( $17.8 \pm 13.6$  vs  $8.46 \pm 10.3$  mg/L,



**Figure 2** Probability of individual risk factors in the two clusters among total COVID-19 deaths during hospitalization.

**Abbreviations:** BMI, body mass index ( $\text{kg}/\text{m}^2$ ); CAD, coronary artery disease; CHF, chronic heart failure; CKD, chronic kidney disease; DM, diabetes mellitus; GI, gastrointestinal disease; HTN, hypertension; neuro, neurological disease; respiratory, pre-existing respiratory disease.

$p=0.007$ ). No significant differences were observed between clusters in the levels of ferritin, aspartate aminotransferase, ALT, creatinine, troponin, or pro-BNP.

This subgroup analysis confirms the existence of two latent classes within our group of patients. Furthermore, these findings reveal that two different patterns may contribute to not only hospitalization in the vulnerable population of patients admitted with COVID-19 infection, but also death from COVID-19.

## Discussion

To the best of our knowledge, this is the first study in the US to report the underlying phenotypes of patients with COVID-19 admitted to the hospital. These phenotypes, determined on the basis of patient demographics and underlying comorbidities, were identified in two separate analyses: patients admitted to the hospital with COVID-19 and patients who died during hospitalization for COVID-19. Given the variations in clinical presentation<sup>10,13</sup> and potential genetic polymorphisms<sup>14,15</sup> of COVID-19 (common gene clusters among patients with severe COVID-19 pneumonia), identifying distinct underlying phenotypes is vital to establishing a targeted therapy for each phenotype and facilitating further study on the mechanism of COVID-19.

Overall, our study shows that patients who were admitted for COVID-19 can be categorized into two cohorts, C1 and C2. C1 consists of older individuals with more comorbidities and higher mortality rates than C2, whereas C2 consists of younger individuals who are more often male, nonwhite, obese, and have increased expression of CRP. This partially correlates with the recent retrospective studies conducted in Spain<sup>4</sup> and France,<sup>5</sup> where three phenotypes instead of two were identified. In the study of the COVID-19 @ Spain registry by Gutierrez et al,<sup>4</sup> the phenotype A group consisted of younger patients who were

less often male and had mild viral symptoms and normal inflammatory parameters; the phenotype B group included patients who were obese with moderately elevated inflammatory markers; and the phenotype C group included older patients with more comorbidities. Overall, patients in the phenotype A group were significantly younger and were less frequently male than were patients in phenotype B and phenotype C groups. Furthermore, the mortality rate was the lowest in the phenotype A group. Overall, the phenotype C group, which had more elderly patients with more comorbidities, had the highest mortality rate. Similar results were noted in the retrospective study conducted in France.<sup>5</sup>

The C1 group in our study correlates to the phenotype C group in the COVID-19 @ Spain study, which included older patients with more comorbidities and had the highest mortality rate. Similar to patients in the phenotype C group, patients in C1 were much older and had significantly higher chances of having HTN, CHF, CAD, and pre-existing respiratory disease, which may help explain the higher mortality rate in this cohort. Given that age is an independent predictor of do not resuscitate (DNR) status,<sup>16</sup> elderly patients may have a higher rate of DNR status<sup>16</sup> and, thus, a higher mortality rate and lower likelihood of undergoing invasive procedures (e.g., intubation).<sup>17</sup> Finally, at the time of this study, certain medications were experimental drugs that adhered to strict inclusion criteria. Elderly patients with more comorbidities may have been more likely to have been excluded in the clinical trials because of underlying conditions. For example, at the time of the study, remdesivir was given to patients who were intubated and who did not require dialysis as part of the clinical trial inclusion criteria in the Yale New Haven Health System; having dementia as a comorbidity is still a relative contraindication for remdesivir administration.



The C2 group in our study correlates to the phenotype B group, which included younger patients who were more likely to be obese with moderately elevated levels of inflammatory markers. In C2, only 15.9% of patients were 65 or older, and none of them had CHF or CAD. C2 also had a much smaller percentage of patients with HTN, DM, CKD, or pre-existing neurological, gastrointestinal, or respiratory disease. Obesity has been associated with elevated inflammation status,<sup>18</sup> and young adults with morbid obesity have been shown to experience substantial rates of adverse outcomes such as requiring intensive care or mechanical ventilation, and even death.<sup>19</sup> In addition, patients in C2 also had increased CRP expression, which has been suggested to be an underlying mechanism of severe COVID-19 pneumonia and acute respiratory distress syndrome (ARDS).<sup>20,21</sup> The higher likelihood of being obese and having an elevated CRP level in this phenotype cluster may suggest a different pathophysiology and possible differences in response to a specific COVID-19 treatment.<sup>22</sup>

Unlike the studies conducted in Spain and France, we did not identify the phenotype A group, which consisted of young patients who were less frequently male and had minor symptoms with normal levels of inflammatory markers. The following reasons may have contributed to the difference in phenotype allocation. First, in both studies conducted in Spain and France, the admission criteria were not published, which may have contributed to the slight difference in phenotype attribution, particularly for patients of the phenotype A group, who presented with minor symptoms and normal levels of inflammatory markers. In general, the patients who required hospitalization in our study must have required oxygen supplementation, and a large portion of them manifested at least moderate symptoms. Second, the patient populations were significantly different in some regards, particularly with respect to demographic differences between the US and Spain and France.<sup>23</sup> For example, the US has higher rates of obesity, diabetes, and chronic lung disease at baseline than do peer countries such as Spain and France.<sup>24</sup> Third, given the geographical differences, it is unclear whether the study populations among the different countries were infected with the same strain of virus.

Nonetheless, similar to our study, both studies in Spain and France showed distinct phenotypes of patients who were hospitalized with COVID-19. Moreover, our findings indicate that younger patients with obesity and moderately elevated levels of inflammatory markers may have

a different underlying pathophysiology from elderly patients with more comorbidities.

After analyzing patients who died during hospitalization for COVID-19, we found that the distribution of patient baseline characteristics and comorbidities was similar to that of the entire cohort admitted for COVID-19. The C1' cluster was older with more comorbidities and had a lower intubation rate and a shorter admission-to-death timeframe. C2' patients were younger with few comorbidities, but they were more likely to be obese, be male, and have increased expression of certain inflammatory markers. This additional analysis further validated the existence of two latent classes within a large heterogeneous group of patients with COVID-19 and revealed that two different patterns of risk factors may contribute to not only requiring hospitalization, but also death from a severe case of COVID-19.

## Limitations

The clusters identified in this study were predominantly differentiated by age; because of this, it is possible that different levels of care were received between classes. Nonetheless, age alone cannot fully explain the differences in sex, race, obesity, and biomarkers between these two groups.

Although the small sample size of this single-center retrospective study is a limitation, our statistically significant findings indicate that risk factors are clustered together in two distinct groups in US patients who were hospitalized for COVID-19. These results suggest that different pathophysiologic processes lead to moderate to severe COVID-19 and death and may be useful in identifying treatment targets and selecting patients with severe COVID-19 disease for future clinical trials. Further studies in a larger cohort should be conducted to test the generalizability of our findings.

## Conclusion

We identified two distinct phenotypes in a sample of US patients admitted for COVID-19. These phenotypes were based on patient demographics and underlying comorbidities. C1 featured older individuals with more comorbidities and a higher mortality rate, whereas C2 consisted of individuals who were younger and more likely to be male, with increased expression levels of CRP. A similar pattern was also identified in patients who died from COVID-19. Our findings support the pathophysiological and clinical heterogeneity of COVID-19, which may be central to the early identification of high-risk patients and treatment optimization on the basis

of each phenotype's underlying mechanism. Our study suggests that future research efforts that provide a comprehensive understanding of each phenotype may be worthwhile and that, ultimately, targeted clinical trials may lead to specific treatments for each cohort.

## Acknowledgments

The authors thank Nicole Stancel, PhD, ELS(D) of the Department of Scientific Publications at the Texas Heart Institute, for providing editorial support.

## Funding

There is no funding to report.

## Disclosure

The authors report no conflicts of interest in this work.

## References

- Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature*. 2020;584:430–436. doi:10.1038/s41586-020-2521-4
- Parohan M, Yaghoubi S, Seraji A, et al. Risk factors for mortality in patients with Coronavirus disease 2019 (COVID-19) infection: a systematic review and meta-analysis of observational studies. *Aging Male*. 2020;23:1416–1424. doi:10.1080/13685538.2020.1774748.
- Kim L, Garg S, O'Halloran A, et al. Risk factors for intensive care unit admission and in-hospital mortality among hospitalized adults identified through the U.S. coronavirus disease 2019 (COVID-19)-associated hospitalization surveillance network (COVID-NET). *Clin Infect Dis*. 2020. doi:10.1093/cid/ciaa1012
- Gutierrez-Gutierrez B, Del Toro MD, Borobia AM, et al. Identification and validation of clinical phenotypes with prognostic implications in patients admitted to hospital with COVID-19: a multicentre cohort study. *Lancet Infect Dis*. 2021. doi:10.1016/S1473-3099(21)00019-0
- Azoulay E, Zafrani L, Mirouse A, et al. Clinical phenotypes of critically ill COVID-19 patients. *Intensive Care Med*. 2020;46:1651–1652. doi:10.1007/s00134-020-06120-4
- Stokes AC, Lundberg DJ, Elo IT, et al. COVID-19 and excess mortality in the United States: a county-level analysis. *PLoS Med*. 2021;18:e1003571. doi:10.1371/journal.pmed.1003571
- Preston SH, Vierboom YC. Excess mortality in the United States in the 21st century. *Proc Natl Acad Sci U S A*. 2021;118:e2024850118. doi:10.1073/pnas.2024850118
- Centers for Disease Control and Prevention. Covid-19 forecasts: cumulative deaths. Available from: <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html>. Accessed June 7.
- Institut National d'Etudes Démographiques. The demographics of Covid-19 deaths. Available from: <https://dc-covid.site.ined.fr/en/>. Accessed June 7.
- Richardson S, Hirsch JS, Narasimhan M, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City Area. *JAMA*. 2020;323:2052–2059. doi:10.1001/jama.2020.6775
- Lanza ST, Collins LM, Lemmon DR, et al. PROC LCA: a SAS procedure for latent class analysis. *Struct Equ Modeling*. 2007;14:671–694. doi:10.1080/10705510701575602
- Li P, Dai Q, Cai P, et al. Identifying different phenotypes in takotsubo cardiomyopathy by latent class analysis. *ESC Heart Failure*. 2021;8:555–565. doi:10.1002/ehf2.13117
- Sudre CH, Lee KA, Lochlainn MN, et al. Symptom clusters in Covid19: a potential clinical prediction tool from the COVID Symptom study app. *Sci Adv*. 2021;7(12). doi:10.1101/2020.06.12.20129056
- Severe Covid-19 GWAS Group. Genomewide association study of severe Covid-19 with respiratory failure. *N Engl J Med*. 2020;383:1522–1534. doi:10.1056/NEJMoa2020283
- Kaser A. Genetic risk of severe Covid-19. *N Engl J Med*. 2020;383:1590–1591. doi:10.1056/NEJMe2025501
- Messinger-Rapport BJ, Kamel HK. Predictors of do not resuscitate orders in the nursing home. *J Am Med Dir Assoc*. 2005;6:18–21. doi:10.1016/j.jamda.2004.12.006
- Boonmee P, Ruangsomboon O, Limsuwat C, et al. Predictors of mortality in elderly and very elderly emergency patients with sepsis: a retrospective study. *West J Emerg Med*. 2020;21:210–218. doi:10.5811/westjem.2020.7.47405
- Samad F, Ruf W. Inflammation, obesity, and thrombosis. *Blood*. 2013;122:3415–3422. doi:10.1182/blood-2013-05-427708
- Cunningham JW, Vaduganathan M, Claggett BL, et al. Clinical outcomes in young US adults hospitalized with COVID-19. *JAMA Intern Med*. 2020. doi:10.1001/jamainternmed.2020.5313
- Sinha P, Calfee CS, Cherian S, et al. Prevalence of phenotypes of acute respiratory distress syndrome in critically ill patients with COVID-19: a prospective observational study. *Lancet Respir Med*. 2020;8:1209–1218. doi:10.1016/S2213-2600(20)30366-0
- Manson JJ, Crooks C, Naja M, et al. COVID-19-associated hyperinflammation and escalation of patient care: a retrospective longitudinal cohort study. *Lancet Rheumatol*. 2020;2:e594–e602. doi:10.1016/S2665-9913(20)30275-7
- Sanders JM, Monogue ML, Jodlowski TZ, et al. Pharmacologic treatments for coronavirus disease 2019 (COVID-19): a review. *JAMA*. 2020;323:1824–1836. doi:10.1001/jama.2020.6019
- Perez AD, Hirschman C. The changing racial and ethnic composition of the us population: emerging American identities. *Popul Dev Rev*. 2009;35:1–51. doi:10.1111/j.1728-4457.2009.00260.x
- National Research Council, Committee on Population. U.S. Health In International Perspective: Shorter Lives, Poorer Health. Washington (DC): National Academies Press (US); 2013.

### Infection and Drug Resistance

### Publish your work in this journal

Infection and Drug Resistance is an international, peer-reviewed open-access journal that focuses on the optimal treatment of infection (bacterial, fungal and viral) and the development and institution of preventive strategies to minimize the development and spread of resistance. The journal is specifically concerned with the epidemiology of

antibiotic resistance and the mechanisms of resistance development and diffusion in both hospitals and the community. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/infection-and-drug-resistance-journal>

Dovepress