medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

COVID-19 Outcomes and Genomic characterization of SARS-CoV-2 isolated from Veterans in New England States

Megan Lee, BS[1*]; Ya Haddy Sallah, MPH[1*]; Mary E. Petrone, BS[2*]; Matthew Ringer, MD[3]; Danielle Cosentino, BS[4]; Chantal B.F. Vogels, PhD[2]; Joseph R. Fauver, PhD[2]; Tara Alpert, PhD[2]; Nathan D. Grubaugh, PhD[2]; Shaili Gupta, MBBS[3,5]

[1]Yale School of Medicine, 333 Cedar St., New Haven, CT 06511

[2] Department of Epidemiology of Microbial Diseases, Yale School of Public Health, 60 College St, New Haven, CT 06510

[3] Department of Internal Medicine, Yale School of Medicine, 333 Cedar St, New Haven, CT 06510

[4] Department of Clinical Informatics, AV Connecticut Healthcare System, 950 Campbell Ave, West Haven, CT 06516

[5]Department of Medicine, VA Connecticut Healthcare System, 950 Campbell Ave, West Haven, CT 06516

* M.L., Y.H.S. and M.E.P. contributed equally to this work and are co-first authors.

Running Title: COVID-19 Outcomes and SARS-CoV-2 lineages in New England Veterans

Corresponding Author:

Shaili Gupta, MBBS

Department of Medicine, Yale School of Medicine

VA Connecticut Healthcare System

950 Campbell Ave, Bldg 1, 5th floor, Mailstop 111

West Haven, CT 06516

E-mail: shaili.gupta@yale.edu; phone: 203-932-5711, ext 4412; fax: 203-932-3748

Word count: Abstract: 250. Main text: 2,037.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract:

Background: Clinical and virologic characteristics of COVID-19 infections in veterans in New England have not been described.

Objectives: To evaluate clinical and virologic factors impacting COVID-19 outcomes.

Study Design: We reviewed charts and sequenced virus from nasopharyngeal specimens with confirmed SARS-CoV-2 from 426 veterans in six New England states between April and September, 2020. Peak disease severity, hospitalization, and mortality were correlated to clinical, demographic, and virologic factors.

Results: Of 426 veterans, 274 had complete and accessible charts. 92.7% were men, 83.2% White, with mean age 63 years. On multivariate regression, significant predictors of hospitalization were age (OR: 1.05) and non-white race (OR: 2.39). Mortality and peak disease severity varied by age (OR 1.06 and 1.07 respectively) and oxygen requirement on admission (OR6.74 and 45.7). Dementia (OR: 3.44) was also associated with mortality. Most (97.3%) of our samples were dominated by the spike protein D614G substitution, and were from SARS-CoV-2 B.1 lineage or one of 37 different B.1 sub-lineages, with none representing more than 8.7% of the cases.

Conclusions: In an older cohort of veterans from the six New England states with a high comorbidity burden, age was the largest predictor of hospitalization, peak disease severity, and mortality. Non-white veterans were more likely to be hospitalized, and patients who required oxygen on admission were more likely to have severe disease and higher rates of mortality. Multiple SARS-CoV-2 lineages were distributed in patients in New England early in the COVID-19 era, mostly related to viruses from New York with D614G mutation.

Keywords: COVID-19, Outcomes, Veterans, Age, SARS-CoV-2, D614G

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Background

The Coronavirus Disease 2019 (COVID-19) caused by Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) has infected >30 million people in the United States and caused nearly 600,000 deaths [1]. Disease severities and outcomes vary among individuals who become infected. Throughout the course of the pandemic, researchers have tried to identify patient and viral characteristics that predispose patients to having worse outcomes. Some host-factors known to impact disease severity and mortality include Charlson comorbidity index score, age, and body mass index (BMI) [2-6].

Virologic characteristics have also been suggested to impact the severity of disease. A dynamic nomenclature system (Pangolin) was developed to classify SARS-CoV-2 and identify lineages and mutations that could impact could impact infectivity and virulence [7]. Early during the pandemic course in the United States, viral isolates on the East Coast of the US were predominantly of the B.1 lineage that contained the Spike protein D614G substitution [8]. This substitution has been suggested to have higher transmissibility [9], supported by the higher upper airway viral loads and higher viral replication in lung epithelial cells [10]. At the same time, viruses with the D614G substitution has also been found to be more susceptible to neutralization by antibodies generated by infection with wild-type virus [11,12]. With the increasing focus on emerging SARS-CoV-2 genetic variants, such as the lethality of B.1.1.7 [13], more research is required on how mutations affect patient outcomes. Furthermore, understanding viral epidemiology and a regional evaluation of viral variants, with their respective clinical correlations, is important to provide a full understanding of the disease.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Objectives

Given the high variability and conflicting data in predicting who will have poor outcomes, assessment of specific populations is necessary to give providers the best clinical picture on their patients. Clinical outcomes of veterans with COVID-19 in New England and respective SARS-CoV-2 genomics have not been described. We aimed to (1) describe patient characteristics, comorbidities, and disease factors that impact patient outcomes and (2) present data on the genomic composition of the SARS-CoV-2 infecting these patients. Factors impacting outcomes help inform triage and management algorithms for COVID-19.

Study Design

Our cohort was composed of all veterans receiving care at Veterans Administration (VA) healthcare centers who tested positive in six New England states (Connecticut, Massachusetts, Maine, New Hampshire, Rhode Island, Vermont) from April 8, 2020, to September 16, 2020. Inclusion criteria included patients with accessible chart records and a positive COVID-19 test. We manually reviewed charts and recorded demographics [age, gender, race, body-mass index (BMI), long-term care facility (LTC) status, and state of residence when diagnosed with COVID-19]. Comorbidities recorded were immunosuppression, dementia, diabetes mellitus, chronic kidney disease (CKD) stage ≥ 3, chronic liver disease, coronary artery disease (CAD), heart failure, atrial fibrillation (aFib), chronic obstructive pulmonary disease (COPD), asthma, and active tobacco use. A clinical database repository and virologic repository was created for all veterans diagnosed with COVID-19 in the VA healthcare system of New England, and was approved by VA Connecticut Institutional Review Board. Whole genome sequencing was

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

conducted on SARS-CoV-2 isolates with a cycle threshold (Ct) value <36, and provided near-complete or complete genome results where Ct value was <30.

We recorded hospitalization status of patients, mortality, and whether patients required oxygen (O2) within 24 hours of admission. Lastly, we divided patients based on their peak disease severity into five categories depending on oxygenation requirements: 1) no O2 requirement, 2) 1-3 liters (L) by nasal cannula (NC), 3) 4-6 L NC, 4) >6 L O2 or non-invasive positive pressure ventilation (NIPPV), and 5) mechanical ventilation. We used STATA v16 for univariate and multivariate logistic regressions to predict the outcomes of interest.

For the genetic data, we sequenced whole virus genomes (≥20x coverage depth across ≥70% of the genome) using Illumina (n = 238) and Nanopore (n = 61) platforms. Using BWA-MEM version 0.7.15, we aligned reads to the Wuhan-Hu-1 reference genomes (GenBank MN908937.3). With iVar v1.2.1, we trimmed sequencing adaptors and primer sequences and called bases by simple majority (>50% frequency) at each site to generate consensus genomes. An ambiguous N was used when fewer than 10 reads were present at a site. We aligned these consensus genomes with MAFFT [14], and masked problematic sites [15]. We then built a phylogenetic tree with IQTree [16] using an HKY substitution model and 1000 bootstraps. We visualized this tree using the Python module baltic v0.1.5. We used Pangolin to assign lineages to our genomes [7].

Results

Of 476 veterans in six New England states with confirmed SARS-CoV-2 during the study period, 274 had complete and accessible charts. Of 274 veterans, 92.7% were men, 83.2% White, and mean age was 63 years (standard deviation: 17.6 years) (Table 1). Over a third resided in

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

LTC (n=92), of which 76% were at the VA's community living center. Over half the residents were from Massachusetts, and one third from Connecticut. The most common comorbidities were CAD (27%), diabetes (25%), and tobacco use (23%).

Notably, 11.7% patients required O2 above their baseline home oxygen requirements within 24 hours of admission, and 20.8% of all patients required oxygen support at some point during hospitalization. In terms of peak severity, 78.5% required only room air, 10.9% required 1-3 L O2 via NC, 4.0% required 4-6 L O2 via NC, 3.6% required >6L O2 or NIPPV, and 2.2% were intubated. The hospitalization rate was 28.8% (Figure 1), with the highest rates in people with COPD (50%), CAD (45.9%), and those over age 80 years (44.4%). The overall mortality rate was 10.6% (Figure 2), with the highest rates in people with dementia (33.3%), age >80 (28.9%), and those from LTC (23.9%).

In univariate regressions, significant predictors of hospitalization were age (Odds Ratio (OR), 95% confidence interval (CI): 1.06), non-White race (OR: 1.97), LTC residence (OR: 1.78), COPD (OR: 3.04), CAD (OR: 2.93), and aFib (OR: 2.05). Peak severity increased with age (OR: 1.06), LTC residence (OR: 2.8), CAD (OR: 2.1), aFib (OR: 2.97), and O2 requirement on admission (OR: 46.2). Higher mortality was seen with age (OR: 1.10), LTC residence (OR: 7.86), BMI<30 (OR: 5.80), dementia (OR: 7.23), aFib (OR: 3.11), hospitalization (OR: 2.58), and O2 requirement on admission (OR: 4.34) (Table 2).

Using multivariate regression, significant predictors of hospitalization were age (OR: 1.05) and non-White race (OR: 2.39) (Table 3). Peak severity also varied by age (OR: 1.07) and O2 requirement on admission (OR: 45.7). Mortality was predicted by age (OR: 1.06), dementia (OR:3.44), and O2 requirement on admission (OR: 6.74). In other words, for every year increase

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

in age, the odds of hospitalization increased by 5%, peak severity increased by 7%, and mortality increased by 6%.

For the genomic characteristics, amongst the 426 patients, 299 patients genomes had adequate coverage for analysis. We found the majority of our specimens (154/299, 51.5%) were from SARS-CoV-2 lineage B.1, or a sub-lineage of B.1 (e.g. B.1.302, B.1.303, B.1.356; 137/299, 45.8%), all of which are defined by D614G substitution (Figure 3, Appendix Table 1). Only 2.4% (7/299) were from the lineage A that lack the D614G mutation. Since there were 41 SARS-CoV-2 lineages detected in our cohort, with none of the sub-lineages consisting of more than 26 COVID-19 cases, we did not have the power to test for clinical correlates. Our sequencing data does inform us that the outcomes presented in this VA cohort are dominated by the impacts, whatever they may be, of the B lineage D614G mutation.

Discussion

Our study found that in an older cohort of veterans with a high comorbidity burden, age significantly associated with of hospitalization, peak disease severity, and mortality. Our study is unique in that it is the first to look at several COVID-19 outcomes in veterans from the six New England states. The CDC currently has a list of chronic medical conditions that predispose individuals to severe illness from SARS-CoV-2 infection [17]. This comprehensive list of comorbidities can be difficult to interpret, however, and one study found that >75% of United States adults would fall under a high-risk category [18]. Veterans are a unique cohort because they suffer from high rates of comorbid disease and advanced age. Understanding clinical factors that impact outcomes in veterans will help clinicians risk-stratify patients with similar demographic profiles.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Many COVID-19 studies have found age to be a predictor of worse outcomes [4,19-22]. In our study, age was a significant predictor for all of our outcomes and was a confounder for other variables. Interestingly, LTC status predicted all three of our outcomes on univariate analysis, but not on multivariate analyses. Because age is associated with both having COVID-19 and living in an LTC, it confounded the relationship between the outcomes and LTC status. Earlier in the COVID-19 pandemic, residents of nursing homes had higher rates of infection as well as severe illness and mortality [23]. Our study shows that among veterans in LTC facility, disease outcomes were not impacted by their residence status.

Concurrent work in our lab suggests that oxygen requirement on admission predicts poor outcomes in veterans (Ringer et al, unpublished work). This has helped inform the triage guidelines at our healthcare system. This is an important finding because other ways of determining oxygenation status, such as the arterial oxygen saturation/ fraction of inspired oxygen (SaO2/FiO2) ratio [24], can frequently change, and thus become difficult for clinicians to use in practice.

Our study supports data from previous trials that non-White patients are at increased risk of hospitalization but have similar peak severity and mortality outcomes [25,26]. One study of 5,902 patients with COVID-19 in New York found that non-Hispanic Black patients were more likely to test positive for coronavirus but had similar outcomes [27]. Similar to our study, Azar et al. found that in 1052 cases of COVID-19 in California, non-Hispanic African American patients had 2.7 increase in odds of being hospitalized [28]. Cardemil et al. also found a 4.6 and 4.2-fold increase in rates of hospitalization for Hispanic and Black veterans, respectively.[25] Many studies have shown that minorities often have delays in care due to mistrust of the medical system because of providers' implicit biases [29,30], which might explain the outcomes in our

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

study. It is critical to continue ongoing efforts to combat medical inequities and target prevention efforts and education to communities and racial groups most affected by COVID-19.

After adjusting for age and other comorbidities, we found that patients with dementia had a higher risk of death. This is similar to other studies on patients with COVID-19 and dementia [20,31,32]. In a meta-analysis of 46,391 patients, Hariyanto et al. found that dementia was associated with a higher risk of severe infection (relative risk: 2.63) and death (relative risk: 2.62) [33]. This may be explained by a host of biological factors, but it emphasizes the importance of extra care and monitoring required when approaching a patient with dementia. In this vulnerable population, providers and caregivers must ensure they receive sufficient support throughout the course of their illness.

Limitations of this work include the smaller sample size. Furthermore, our study is specific to veterans, which is a largely male and older cohort, and results may not therefore be generalizable. The time period of this study was prior to established medical therapies for COVID-19 and our outcomes reported are likely worse than you would expect today. Since this was during the first wave of the pandemic when standard guidelines were not yet solidified, many patients not requiring oxygen were admitted which is not standard of practice today. Strengths of our study include its comprehensive scope, wide geographic range, manual chart review allowing for the capturing of all comorbidities and oxygenation parameters that may not be available otherwise in a database, and multivariate analysis of many potential risk factors.

Conclusion

Our study found that in an older cohort of veterans from the six New England states with a high comorbidity burden, age was the single strongest predictor of hospitalization, peak severity, and

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

mortality. Non-white veterans were more likely to be hospitalized, and patients who required oxygen on admission were more likely to have severe disease and higher rates of mortality. Furthermore, patients with dementia were more likely to die. Multiple genomic variants of SARS-CoV-2 were distributed in patients in New England early in the COVID-19 era, mostly from a B.1 sub-lineage with the spike D614G mutation.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Acknowledgements: The authors report no conflicts of interest.

Funding: This work was supported by Centers for Disease Control and Prevention [BAA 75D301-20-R-68024 to N.D.G. and S.G.] and the Yale Center for Clinical Investigation TL1 TR001864 (M.E.P.).

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Literature Cited

1. U.S. Department of Health & Human Services. Published COVID Data Tracker. Accessed March 25, 2021.
2. Simonnet A, Chetboun M, Poissy J, et al. High Prevalence of Obesity in Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) Requiring Invasive Mechanical Ventilation. *Obesity (Silver Spring).* 2020;28(7):1195-1199.
3. Garg S, Kim L, Whitaker M, et al. Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 - COVID-NET, 14 States, March 1-30, 2020. *MMWR Morb Mortal Wkly Rep.* 2020;69(15):458-464.
4. Ioannou GN, Locke E, Green P, et al. Risk Factors for Hospitalization, Mechanical Ventilation, or Death Among 10 131 US Veterans With SARS-CoV-2 Infection. *JAMA Netw Open.* 2020;3(9):e2022310.
5. 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(20):2052-2059.
6. 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;180(7):934-943.
7. Rambaut A, Holmes EC, O'Toole Á, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nature Microbiology.* 2020;5(11):1403-1407.
8. Brufsky A. Distinct viral clades of SARS-CoV-2: Implications for modeling of viral spread. *Journal of Medical Virology.* 2020;92(9):1386-1390.
9. Korber B, Fischer WM, Gnanakaran S, et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell.* 2020;182(4):812-827.e819.
10. Lorenzo-Redondo R, Nam HH, Roberts SC, et al. A clade of SARS-CoV-2 viruses associated with lower viral loads in patient upper airways. *EBioMedicine.* 2020;62.
11. Plante JA, Liu Y, Liu J, et al. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature.* 2021;592(7852):116-121.
12. Weissman D, Alameh MG, de Silva T, et al. D614G Spike Mutation Increases SARS CoV-2 Susceptibility to Neutralization. *Cell Host Microbe.* 2021;29(1):23-31.e24.
13. Davies NG, Jarvis CI, Edmunds WJ, Jewell NP, Diaz-Ordaz K, Keogh RH. Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. *Nature.* 2021.
14. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution.* 2013;30(4):772-780.
15. Nicola De Maio* CW, Rui Borges, Lukas Weilguny, Greg Slodkowicz, Nick Goldman. Masking strategies for SARS-CoV-2 alignments. https://virological.org/t/masking-strategies-for-sars-cov-2-alignments/480. Published 2020. Accessed April 9, 2021.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

16. Minh BQ, Schmidt HA, Chernomor O, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution.* 2020;37(5):1530-1534.
17. People with Certain Medical Conditions. In: CDC, ed2021.
18. Ajufo E, Rao S, Navar AM, Pandey A, Ayers CR, Khera A. U.S. population at increased risk of severe illness from COVID-19. *Am J Prev Cardiol.* 2021;6:100156-100156.
19. Docherty AB, Harrison EM, Green CA, et al. Features of 20□133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ.* 2020;369:m1985-m1985.
20. Giorgi Rossi P, Marino M, Formisano D, et al. Characteristics and outcomes of a cohort of COVID-19 patients in the Province of Reggio Emilia, Italy. *PLoS One.* 2020;15(8):e0238281-e0238281.
21. 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:ciaa1012.
22. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature.* 2020;584(7821):430-436.
23. Bagchi S MJ, Li Q, et al. *MMWR Morb Mortal Wkly Rep 2021*. Rates of COVID-19 Among Residents and Staff Members in Nursing Homes — United States, May 25–November 22, 2020;70:52–55.
24. Garibaldi BT, Fiksel J, Muschelli J, et al. Patient Trajectories Among Persons Hospitalized for COVID-19 : A Cohort Study. *Ann Intern Med.* 2021;174(1):33-41.
25. Cardemil CV, Dahl R, Prill MM, et al. COVID-19-Related Hospitalization Rates and Severe Outcomes Among Veterans From 5 Veterans Affairs Medical Centers: Hospital-Based Surveillance Study. *JMIR Public Health Surveill.* 2021;7(1):e24502-e24502.
26. Gold JAW, Wong KK, Szablewski CM, et al. Characteristics and Clinical Outcomes of Adult Patients Hospitalized with COVID-19 - Georgia, March 2020. *MMWR Morb Mortal Wkly Rep.* 2020;69(18):545-550.
27. Kabarriti R, Brodin NP, Maron MI, et al. Association of Race and Ethnicity With Comorbidities and Survival Among Patients With COVID-19 at an Urban Medical Center in New York. *JAMA network open.* 2020;3(9):e2019795-e2019795.
28. Azar KMJ, Shen Z, Romanelli RJ, et al. Disparities In Outcomes Among COVID-19 Patients In A Large Health Care System In California. *Health Aff (Millwood).* 2020;39(7):1253-1262.
29. Hall WJ, Chapman MV, Lee KM, et al. Implicit Racial/Ethnic Bias Among Health Care Professionals and Its Influence on Health Care Outcomes: A Systematic Review. *Am J Public Health.* 2015;105(12):e60-e76.
30. Nelson A. Unequal treatment: confronting racial and ethnic disparities in health care. *J Natl Med Assoc.* 2002;94(8):666-668.
31. Harrison SL, Fazio-Eynullayeva E, Lane DA, Underhill P, Lip GYH. Comorbidities associated with mortality in 31,461 adults with COVID-19 in the United States: A federated electronic medical record analysis. *PLoS Med.* 2020;17(9):e1003321-e1003321.
32. Miyashita S, Yamada T, Mikami T, Miyashita H, Chopra N, Rizk D. Impact of dementia on clinical outcomes in elderly patients with coronavirus 2019 (COVID-19): an experience in New York. *Geriatr Gerontol Int.* 2020;20(7):732-734.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

33. Hariyanto TI, Putri C, Arisa J, Situmeang RFV, Kurniawan A. Dementia and outcomes from coronavirus disease 2019 (COVID-19) pneumonia: A systematic review and meta-analysis. *Arch Gerontol Geriatr.* 2021;93:104299-104299.

Demographics		**Comorbidities**	
Age (years)	62.9+17.6	Immunosuppressed	10 (4)
Gender		Dementia	42 (15)
Male	254 (93)	Diabetes	68 (25)
Female	20 (7)	CKD 3	18 (7)
Race		Chronic liver disease	32 (12)
White	228 (83)	Chronic heart disease	107 (39)
Non-white	46 (17)	CAD	74 (27)
BMI > 30	110 (40)	Heart failure	29 (11)
From LTC	92 (34)	Atrial fibrillation	40 (15)
State		Chronic Lung disease	97 (35)
Connecticut	89 (32)	COPD	44 (16)
Massachusetts	150 (55)	Asthma	19 (7)
Maine	4 (1)	OSA	55 (20)
New Hampshire	9 (3)	Tobacco Use	62 (23)
Rhode Island	20 (7)		
Vermont	2 (1)		

Results are presented as numbers (percentages).
Abbreviations: LTC - long term care, BMI – body mass index, CKD – chronic kidney disease, CAD – coronary artery disease, COPD – chronic obstructive pulmonary disease, OSA – obstructive sleep apnea

Table 1. Patient Characteristics (n=274)

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Demographics	**Hospitalization**		**Peak Severity**		**Death**	
	Unadjusted OR (95% CI)	P-value	Unadjusted OR (95% CI)	P-value	Unadjusted OR (95% C.)	P-value
Age	1.06 (1.03-1.08)	<0.001*	1.06 (1.04 - 1.09)	<0.001*	1.1 (1.06 - 1.14)	<0.001*
Gender						
Male	Ref		Ref		Ref	
Female	0.41 (0.12-1.45)	0.17	0.42 (0.95 - 1.9)	0.26	n/a	n/a
Race						
White	Ref		Ref		Ref	
Non-White	1.97 (1.02 - 3.8)	0.04*	1.3 (0.6 - 2.6)	0.55	0.54 (0.16 - 1.9)	0.33
From Long Term Facility	1.78 (1.04 - 3.07)	0.04*	2.8 (1.5 - 5.1)	0.001*	7.9 (3.2 - 19.2)	<0.001*
Clinical Characteristics						
BMI < 30	1.52 (0.9 - 2.7)	0.14	1.1 (0.6 - 2)	0.77	5.8 (2.1 - 15.9)	0.001*
Dementia	0.63 (0.29 - 1.4)	0.25	1.6 (0.7 - 3.3)	0.26	7.2 (3.2 - 16.6)	<0.001*
COPD	3.04 (1.6 - 5.9)	0.001*	1.5 (0.74 - 3.2)	0.25	1.8 (0.7 - 4.5)	0.22
Heart Failure	1.6 (0.7 - 3.5)	0.25	1.8 (0.8 - 4.0)	0.18	1.9 (0.67 - 5.5)	0.23
CAD	2.93 (1.66 - 5.2)	<0.001*	2.1 (1.2 - 3.9)	0.016*	2.1 (0.9 - 4.6)	0.07
Atrial Fibrillation	2.05 (1.03 - 4.1)	0.04*	3 (1.4 - 5.9)	0.002*	3.1 (1.3 - 7.4)	0.01*
Hospitalization	n/a		n/a		2.58 (1.18-5.65)	0.017*
O2 on admission	n/a		46.2 (19.9, 107.3)	<0.001*	4.34 (1.77 - 10.6)	0.001*

*p<0.05

Table 2. Univariate regression analysis of factors that predict hospitalization, peak severity, and death

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Demographics	**Hospitalization**		**Peak Severity**		**Death**	
	OR (95% C.I)	P-value	OR (95% C.I)	P-value	OR (95% C.I)	P-value
Age	1.05 (1.03, 1.08)	<.001*	1.07 (1.03 - 1.11)	<0.001*	1.06 (1.01, 1.11)	0.019*
Race						
White	Ref		n/a		n/a	
Non-White	2.39 (1.13, 5.01)	0.021*				
From long term facility	0.70 (0.36, 1.38)	0.307	1.18 (0.55, 2.55)	0.668	2.0 (0.58, 6.88)	0.273
Clinical Characteristics						
BMI < 30	n/a		n/a		2.95 (0.81, 10.75)	0.102
Dementia	n/a		n/a		3.44 (1.07, 11.1)	0.038*
COPD	1.73 (0.84, 3.35)	0.136	n/a		n/a	
CAD	1.44 (0.75, 2.81)	0.273	0.71 (0.32, 1.55)	0.390	0.57 (0.209, 1.57)	0.283
Atrial Fibrillation	0.984 (0.45, 2.16)	0.969	1.11 (0.46, 2.64)	0.819	1.25 (0.42, 3.69)	0.686
Hospitalization	n/a		n/a		1.36 (0.40, 4.65)	0.631
O2 on admission	n/a		45.7 (18.79, 111)	<0.001*	6.74 (1.74, 26.1)	0.006*

*p<0.05

Table 3. Multivariate regression analysis of factors that predict hospitalization, peak severity, and death

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

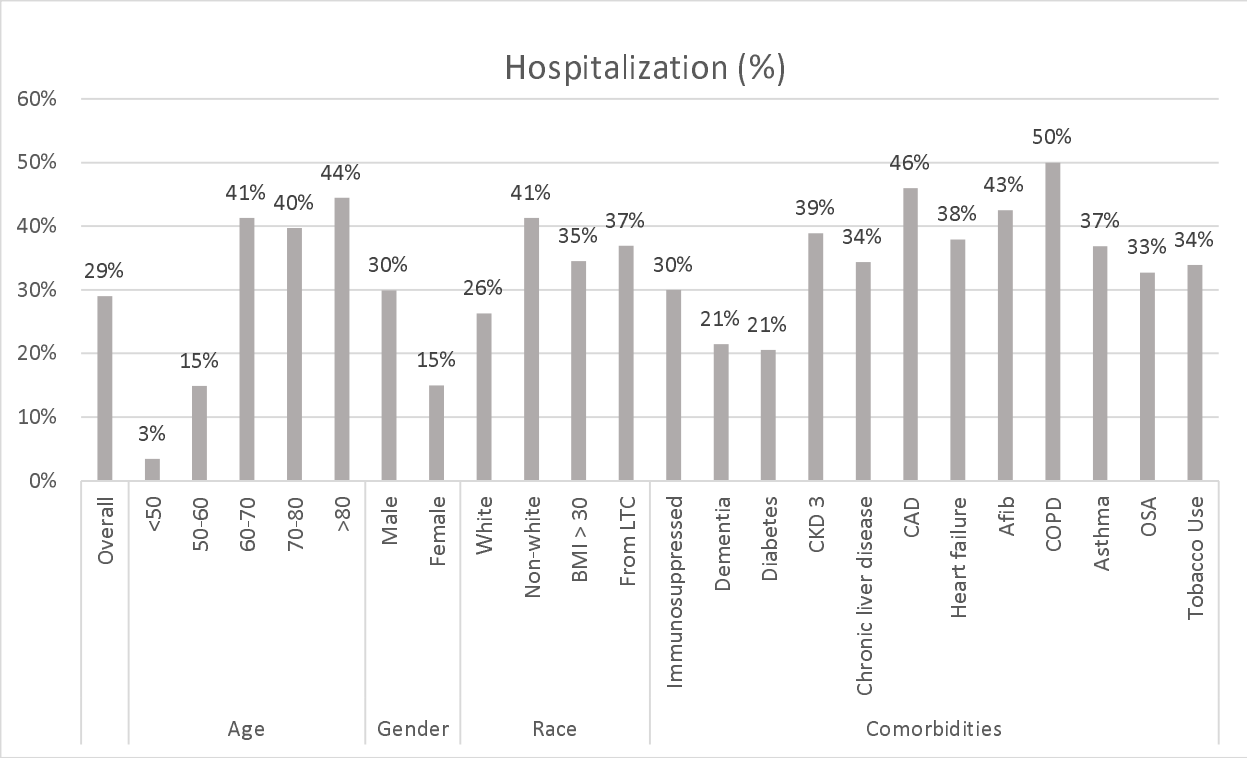

Figure 1. Percent of patients hospitalized, based on patient demographics and comorbidities

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

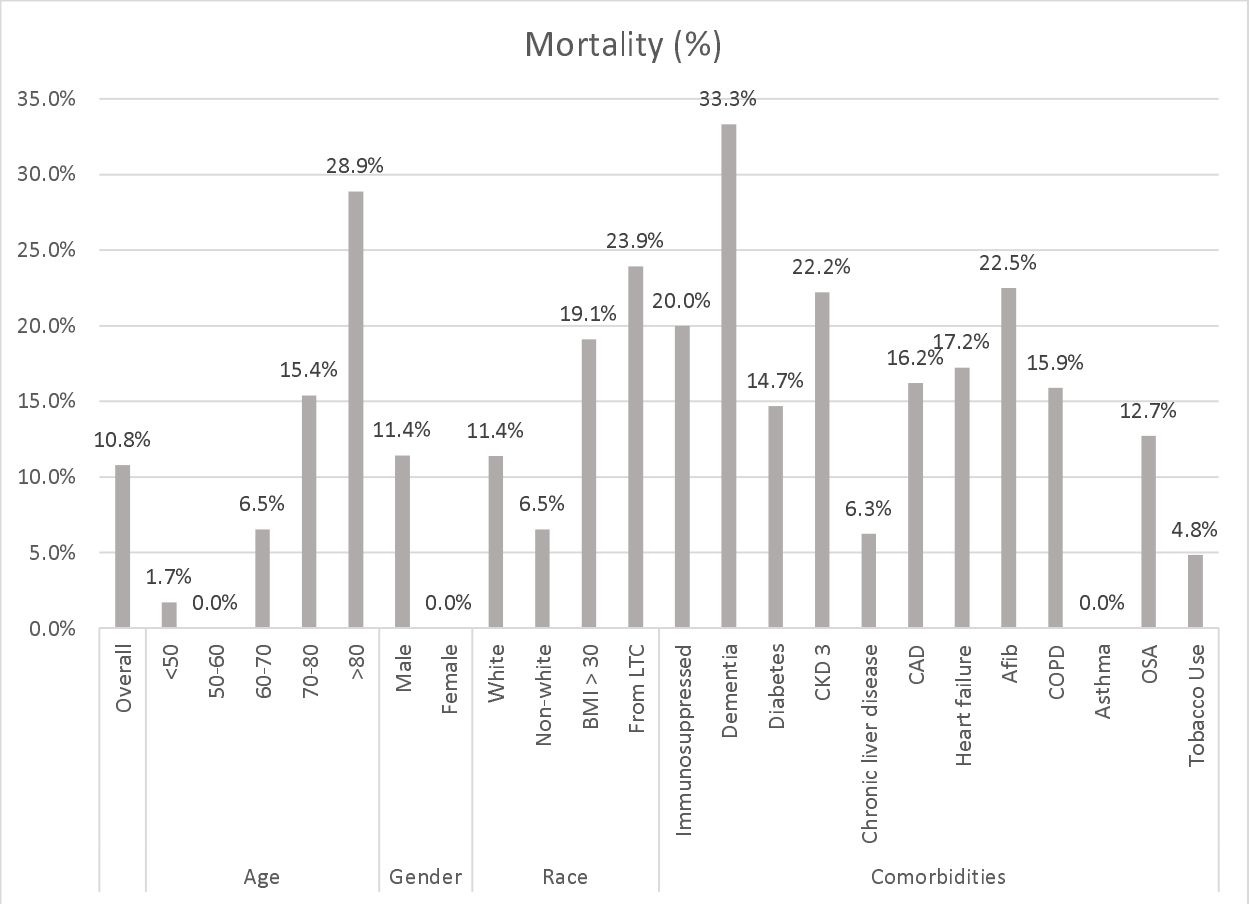

Figure 2. Percent of patients who died, based on patient demographics and comorbidities

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

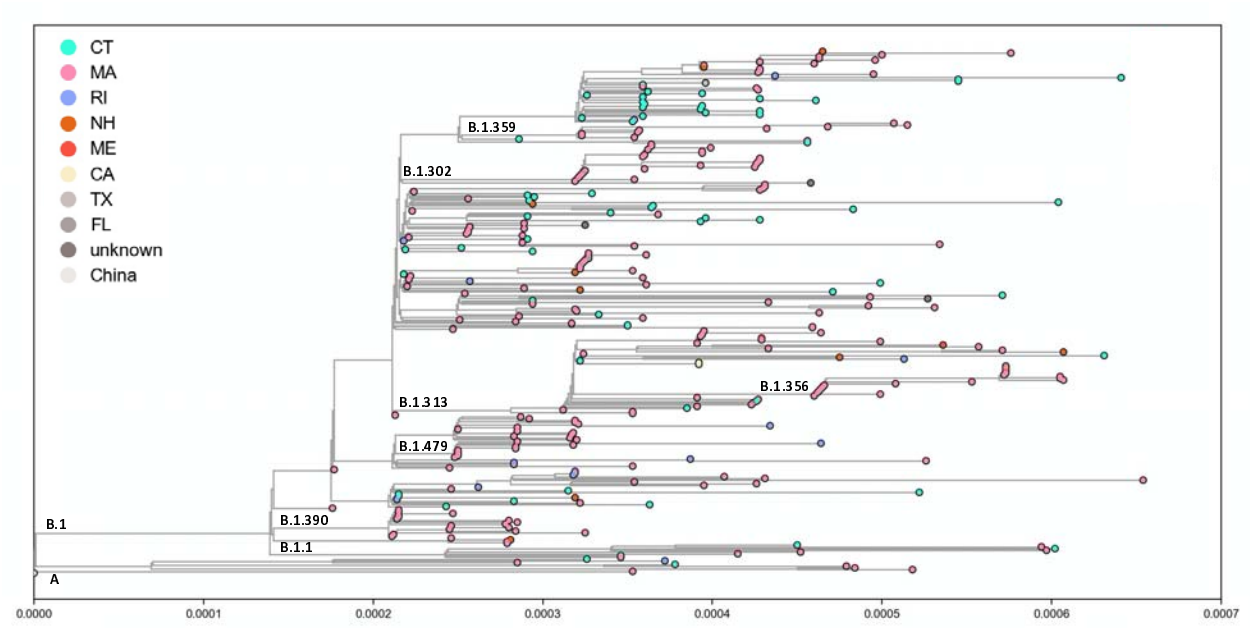

Figure 3. Maximum Likelihood tree of genomes

Note: The authors would prefer color online only

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Appendix Table. Lineages of genomes

Lineage	Count	Percent
A.1	3	1.0%
A.2	2	0.7%
A.3	2	0.7%
B	1	0.3%
B.1	154	51.5%
B.1.1.10	1	0.3%
B.1.1.113	1	0.3%
B.1.1.128	1	0.3%
B.1.1.163	1	0.3%
B.1.1.192	1	0.3%
B.1.1.225	1	0.3%
B.1.1.307	1	0.3%
B.1.104	4	1.3%
B.1.108	1	0.3%
B.1.110.3	1	0.3%
B.1.2	3	1.0%
B.1.243	1	0.3%
B.1.268	1	0.3%
B.1.280	2	0.7%
B.1.302	26	8.7%
B.1.313	19	6.4%
B.1.319	1	0.3%
B.1.320	1	0.3%
B.1.331	1	0.3%
B.1.356	18	6.0%
B.1.359	8	2.7%
B.1.36	1	0.3%
B.1.369	3	1.0%
B.1.382	1	0.3%
B.1.385	1	0.3%
B.1.390	16	5.4%
B.1.403	1	0.3%
B.1.413	1	0.3%

medRxiv preprint doi: https://doi.org/10.1101/2021.04.27.21256222; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

B.1.448	2	0.7%
B.1.448	1	0.3%
B.1.479	10	3.3%
B.1.509	2	0.7%
B.1.517	1	0.3%
B.1.521	1	0.3%
B.35	1	0.3%
N.1	1	0.3%