Causal Inference and COVID-19 Nursing Home Patients: Identifying Factors That Reduced Mortality Risk

Amina Ahmed MD¹, Robert Goldberg PhD², Joseph Swiader BS², Zachary A.P. Wintrob MS MA², Margaret Yilmaz MS²

Abstract

Less than 1% of the US population lives in long-term care facilities, yet this subset of the population accounts for 22% of total COVID-19 related deaths.¹ Because of a lack of experimental evidence to treat COVID-19, analysis of real-world data to identify causal relationships between treatments/policies to mortality and morbidity among high-risk individuals is critical. We applied causal inference (CI) analysis to longitudinal patient-level health data of 4,091 long-term care high-risk patients with COVID-19 to determine if any actions or therapies delivered from January to August of 2020 reduced COVID-19 patient mortality rates during this period.

Causal inference findings determined that certain supportive care interventions caused reduced mortality rates for nursing home residents regardless of severity of disease (as measured by oxygen saturation level, presence of pneumonia and organ failure), comorbidities or social determinants of health such as race, age, and weight.² While we do not address the biological mechanisms associated with specific medical interventions and their impact on mortality, this analysis suggests methods to validate and optimize treatment protocols using domain knowledge and causal inference analysis of real-world data across patient populations and care settings.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Introduction

Throughout the COVID-19 pandemic, nursing home residents have died of COVID-19 at substantially higher rates than people in hospitals or living at home. As the pandemic began in early 2020, the death rate for nursing home patients infected with the SARS-CoV2 virus was over 30 percent. By August 30, 2020, the COVID-19 death rate in nursing homes had dropped to 18.2 percent, and the number of new COVID-19 cases declined as well. COVID-19 mortality rates declined during the first wave of the pandemic (from January through October) among nursing home residents and the general population. As of September 2020, “people in long-term care facilities were 8 percent of coronavirus cases, but 45 percent of all COVID-19 death.”

There is a limited amount of COVID-19 research seeking to identify treatments or detection steps associated with reduced mortality. While randomized controlled clinical trials can be used to establish whether a treatment causes a reduction in mortality, such studies are often difficult to conduct or may be unethical. A recent study of mortality rates among nursing home residents concluded that mortality risk “declined for residents with both symptomatic and asymptomatic infections and for residents with both high and low clinical complexity.” The authors surmised that a change in the transmissibility and severity of the virus may have caused this shift and did not find any interventions that were associated with reduced mortality. Other studies have attempted to identify specific treatment regimens or detect strategies that might have conveyed such a benefit. A Cochrane Collaboration meta-analysis found that “based on currently available data, neither absence nor presence of signs or symptoms are accurate enough to rule in or rule out COVID-19. The presence of anosmia or ageusia may be useful as a red flag for COVID-19. The presence of fever or cough, given their high sensitivities, may also be useful to identify people for further testing. Prospective studies in an unselected population presenting to primary care or hospital outpatient settings, examining
combinations of signs and symptoms to evaluate the syndromic presentation of COVID-19, are still urgently needed.”

In the absence of randomized controlled trials, the analysis of observational longitudinal data should be considered for evaluating the impact of specific treatments and interventions on COVID-19 mortality. However, as stated by Judea Pearl PhD, pioneer of Bayesian statistics and causal inference, “Interventional questions cannot be answered from purely observational information (i.e., from statistical data alone).”

A recent article on causal inference and clinical decision-making observes, “Prediction calculates a future event in the absence of any action or change; intervention presumes an enacted choice that may influence the future, which requires consideration of the underlying causal structure.”

Predictions alone do not provide clinicians with the ability to determine whether certain interventions or policies have improved outcomes. This analysis sought to identify the causal effects of actions taken to treat high-risk COVID-19 patient with regard to the reduction of mortality risk. Specifically, we asked if various treatments of COVID-19 patients caused a lower rate of mortality.

**Methods**

*Data Source and Population.*

The source data for the analysis was provided by CareOne and consisted of clinical records for 4,091 COVID-19 patients diagnosed and treated from January 2020 through August 2020. Data were de-identified and determined by an independent institutional review board as nonhuman subject research. The cohort was composed of both inpatient and outpatient residents at 55 nursing and assisted living facilities in the northeastern United States. All continuous variables were standardized prior to model fitting.
We analyzed data with several hundred variables for every patient including demographics (age, gender, race, region, insurance status), preexisting conditions such as diabetes, obesity, BMI, heart disease, cancer, as well up to twenty-six different lab results, oxygen, incentive spirometry, etc. Patient data also included vital signs, medications and whether the patient was being provided hospice or palliative care or were in an assisted living, post-acute care, or long-term care setting.

Of the 4,091 residents with confirmed COVID-19 infections in the study population, 2,277 (55.7%) were female. Most COVID-19 positive cases identified as Caucasian, 2,630 (64.3%) with 375 (9.2%) identifying as African American. Of the total COVID-19 patients evaluated during the study time frame there 1,204 new admissions and re-admissions that tested positive for COVID-19 after being admitted from hospitals. All COVID-19 infections in this subset occurred within 14 days of admission. Baseline characteristics of the study population and survivors subgroups are shown in Tables 1 and 2, respectively. Table 3 shows the distribution of comorbidities among the study population.

**Causal Inference.**

As Pearl notes, “Graphical models such as Directed Acyclic Graphs (DAGs) can be used for encoding as well as portraying conditional independencies and causal relations.” Accordingly, we constructed directed acyclic graphs using the patient health record variables to differentiate causal relationships from associations in the data. A DAG clarifies study questions and identifies and illustrates assumptions (in the form of lines or edges) about the relationships between variables (nodes) and reduces bias by separating individual effects and establishing appropriate covariates. The DAG is *directed* in that its lines point from cause to effect (causal effects cannot be bidirectional); and *acyclic* in that no directed path can form a closed loop in a DAG.

DAGs make assumptions and potential causal relationships transparent. Changes to the models are made by holding every other variable constant that might account for the outcome to observe the
effect of changing or removing a factor or action.\textsuperscript{10} We generated a DAG for every patient by using lab values as instrumental variables (i.e., variables more strongly associated with the treatment than with the outcome) to eliminate confounding and to identify latent (unobserved) variables to eliminate residual bias carried by unmeasured confounders.\textsuperscript{11}

To isolate specific independent effect of medicines or treatment steps, we generated thousands of possible of heterogeneous treatment effects of concurrent pharmacotherapy using Orthogonal Random Forests to control for confounders. For each patient, treatment specific models were trained to predict constant marginal average treatment effects and constant marginals for mortality with a minimum leaf size of 50 and a maximum depth of two.\textsuperscript{12} We then developed and calculated a propensity score for causal relationships for each patient to establish the conditional probability of observed variables. This was accomplished by using treatment specific logistic-regression models to predict constant average treatment effects and constant marginals for specific therapies across the entire patient population. The internal treatment predictions were fit by logistic regression with 3-fold cross-validation. The internal outcome model also implemented 3-fold cross-validation with a ridge classifier with 2 exceptions, the morphine and sennosides outcome models were fit with weighted LASSO.\textsuperscript{13} Analysis was conducted using Python 3.9.0 and the EconML and Scikit-Learn packages.\textsuperscript{14} We measure survival as a Boolean feature representing whether a COVID-19 patient or former COVID-19 patient “died” or “survived”. If a patient died within 120 days of a covid diagnosis the mortality outcome target was “1”, otherwise it was “0”.\textsuperscript{15}
Results

Mortality.

Table 4 provides mortality rates for the entire population (17.3%). The COVID-19 mortality rate of non-hospice COVID-19 patients between January and August 2020 was 9.9% percent. During this period, CareOne took in several hundred medically complex COVID-19 patients from hospitals and other facilities at the request of the New Jersey State Department of Health, including those who were receiving hospice and palliative care in other settings.\(^{16}\) By comparison, a recent study found that during the period of study the COVID-19 mortality nationally was 21 percent.\(^{17}\)

In the early days of the pandemic, CareOne isolated all patients transferred from a hospital to a CareOne facility for two weeks to stop the spread of SARS-CoV2. We found that this CareOne containment strategy and the two-week isolation period was effective in reducing transmission hospital acquired infection.

Causal Factors Reducing Mortality Risk.

The causal inference analysis asked counterfactual questions to determine if any treatment provided to COVID-19 patients had a direct causal effect on reducing mortality risk on a case-by-case basis. Such questions posed by the analysis asked whether if patients receiving specific treatments earlier would have had a lower risk of death even if they had not received such care, and conversely, whether people who did not receive treatment would have had a lower risk of death if they had received the treatment.

DAGs were used to eliminate potential confounders such as co-morbidities, severity of illness as measured by oxygen saturation levels, clinical or radiographic presence of pneumonia and multi-organ failure as measured by albumin, alanine transaminase (ALT) and creatine levels.
Figure 1 shows the comprehensive DAG used to assess the causal relationships inferred from our analysis. Prosperdtx developed a complete DAG for every patient in the population to identify a minimal set of covariates (as seen in Chart 1 below) followed by a structural approach to develop a simplified DAG to eliminate confounding, collider bias, selection bias, etc. The simplified DAG of the CareOne patient data set is found in Chart 2 below.

Table 5 displays the results of counterfactual analysis conducted using DAGs to determine what, if any, treatment provided to all COVID-19 positive residents caused a reduction in mortality risk. We also estimated the average treatment effect (ATE) for each intervention. The ATE measures the expected difference in the outcomes if every patient diagnosed with COVID-19 received a treatment to which we posed the counterfactual question: What if every patient diagnosed with COVID-19 had not received the medication? Using DAGs to eliminate confounders allowed us to estimate that the ATE was a function of the observed data distribution.\textsuperscript{18} Counterfactual analysis of the data was conducted to identify interventions with the greatest average treatment effect (ATE) and predict patient level response. We found that the administration of dextromethorphan, prednisone, apixaban, cholecalciferol and omeprazole, at any level of dosing reported in the EHRs reduced the risk of mortality. Table 5 shows that the use of dextromethorphan, prednisone, apixaban, cholecalciferol and omeprazole had average treatment effect estimates across all COVID-19 positive residents of 23.2%, 15.5%, 11.5%, 10.5% and 10.4% and reductions in risk of mortality, respectively. We found that the administration of 800 mg of hydroxychloroquine sulfate (604 unique patients received an administration) did not increase mortality risk as some studies have claimed\textsuperscript{19} nor did it reduce mortality risk.

We conducted an independent component analysis (ICA) of every lab test value to eliminate the possibility of latent variables (Table 6). Using ICA, we were able to determine whether an individual test at a specific measurement influenced the causal model. We ran a 2-way analysis of every lab value
against all others and found that no lab value was more predictive of mortality than any other. While some lab values were associated with higher or lower risk than others, they did not influence mortality risk in a causal manner. Further, our component analysis showed that no lab value, independently or in combination, was confounding the impact of specific treatment on mortality rates.

**Discussion**

Causal inference analysis showed that specific medications independent of other medications reduced COVID-19 mortality risk among nursing home patients regardless of their comorbidities, demonstrating the utility of observational, longitudinal patient-level data to identify causal relationships between patients, treatments, and outcomes. We can more effectively control for the confounders (a variable, often hidden, that influences the relationship between a dependent variable and independent variable) and collider bias (a third variable is caused by the relationship between an intervention and outcome) using causal inference. Electronic medical records (EHRs) often miss data for one or more variables as was the case in this study. The use of DAGs allowed us to identify when ‘missingness’ might be hiding variables that could bias estimate. However, when we were unable to fully account for missing data in every patient record, for which conditional independence evaluation identified the existence of latent variables, causal relationships were not assessed. However, we showed that causal inference, combined with existing machine leaning methods is an efficient and reliable tool for establishing the causal relationships needed to optimize treatment and personalize medicine. It is possible to do so at scale using automated causal structure learning and to eliminate confounders and selection bias more rapidly and effectively.20

Considering the emergence of new COVID-19 variants and the fact that anti-viral medications and vaccines may not be effective in preventing or controlling the disease in all cases, identifying
treatment choices that can improve outcomes in all patients, regardless of demography, other diseases, care settings, is critically important. Interventional models derived from causal inference can be used to extrapolate “findings across domains (i.e., settings, populations, environments) that differ both in their distributions and in their inherent causal characteristics.”

Our study also highlights the value of supportive care and its role in achieving optimal health of critically and chronically ill patients. As more care is delivered at home or remotely outside of a healthcare facility, applying causal inference to supportive care in these settings for patients may lead to improved outcomes.

At the same time, we look forward to the development of causal inference methods to generate personalized care pathways that can be dynamically updated to optimize health as patient personal and clinical conditions change. At the outset, we need tools to identify changes in the type, timing, and frequency of care that may have a causal effect on well-being. For example, Marie Davidian pioneered the development of individualized dynamic treatment regimens and created software to allow the use of average treatment effects to estimate optimized treatment outcomes at an individual level. Finally, there is a need for the deployment of technologies that allow people to create fully integrated health records that can be updated over time. Such data could be collected from electronic medical records, genomic information, wearables, remote monitoring devices and patient surveys. Longitudinal real-time information, combined with causal analysis could be used to make supportive care both prospective and personalized.

Conclusions

Nursing homes not only dealt with higher rates of Covid-19 infection and death among residents, in several states, patients with COVID-19 from other care settings were taken in by long term care facilities. Diagnosis and treatment decisions were often made without sufficient resources,
without an existing body of clinical knowledge for how to best are for a heterogenous group of patients who were also critically ill and without the time and resources to conduct randomized prospective trials to measure treatment effect.

We used causal inference to establish the certain causal and effect relationships of patient treatments and admission policies that providers could use to guide care under such circumstances. This study shows that causal inference can be used to identify important clinical practices for optimizing patient outcomes in various care settings and in people with other conditions. As Pearl has observed, “the entire enterprise known as “personalized medicine” and, more generally, any enterprise requiring inference from populations to individuals, rests on counterfactual analysis and AI now holds the key theoretical tools for operationalizing this analysis.”

Machine learning not only enables the collection and organization of patient level data from many sources, but it can also be used to automate the generation of individual treatment decisions over time as a patient’s condition and circumstances change. When faced with variations of a novel virus, the capacity to quickly determine what supportive care can reduce the risk of death may be as valuable as predicting who is likely to die at a population level.
Acknowledgments

We are thankful for the insightful writings of Dr. Judea Pearl.
Table 1: Select COVID-19 Patient Demographics

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>State</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>79</td>
<td>23</td>
<td>109</td>
<td>New Jersey</td>
<td>2,188</td>
<td>53.5%</td>
</tr>
<tr>
<td>BMI</td>
<td>26.6</td>
<td>11.3</td>
<td>83.1</td>
<td>Massachusetts</td>
<td>1,471</td>
<td>36.0%</td>
</tr>
<tr>
<td>Height</td>
<td>67.5</td>
<td>45.0</td>
<td>-</td>
<td>New York</td>
<td>120</td>
<td>2.9%</td>
</tr>
<tr>
<td>IBW</td>
<td>133.4</td>
<td>65.2</td>
<td>214.0</td>
<td>Maryland</td>
<td>54</td>
<td>1.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pennsylvania</td>
<td>20</td>
<td>0.5%</td>
</tr>
<tr>
<td>Category</td>
<td>N</td>
<td>%</td>
<td></td>
<td>Connecticut</td>
<td>15</td>
<td>0.4%</td>
</tr>
<tr>
<td>Females</td>
<td>2,275</td>
<td>55.6%</td>
<td></td>
<td>Rhode Island</td>
<td>13</td>
<td>0.3%</td>
</tr>
<tr>
<td>Males</td>
<td>1,816</td>
<td>44.4%</td>
<td></td>
<td>New Hampshire</td>
<td>12</td>
<td>0.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Florida</td>
<td>11</td>
<td>0.3%</td>
</tr>
<tr>
<td>Race</td>
<td>N</td>
<td>%</td>
<td></td>
<td>Virginia</td>
<td>5</td>
<td>0.1%</td>
</tr>
<tr>
<td>Asian</td>
<td>89</td>
<td>2.2%</td>
<td></td>
<td>Maine</td>
<td>4</td>
<td>0.1%</td>
</tr>
<tr>
<td>Black or African</td>
<td>374</td>
<td>9.1%</td>
<td></td>
<td>Vermont</td>
<td>4</td>
<td>0.1%</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>229</td>
<td>5.6%</td>
<td></td>
<td>Delaware</td>
<td>2</td>
<td>&gt;0.1%</td>
</tr>
<tr>
<td>Others</td>
<td>769</td>
<td>18.8%</td>
<td></td>
<td>Ohio</td>
<td>2</td>
<td>&gt;0.1%</td>
</tr>
<tr>
<td>White</td>
<td>2,630</td>
<td>64.3%</td>
<td></td>
<td>Texas</td>
<td>2</td>
<td>&gt;0.1%</td>
</tr>
<tr>
<td>Total</td>
<td>4,091</td>
<td>100.0%</td>
<td></td>
<td>Washington DC</td>
<td>1</td>
<td>&gt;0.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Illinois</td>
<td>1</td>
<td>&gt;0.1%</td>
</tr>
<tr>
<td>Education</td>
<td>N</td>
<td>%</td>
<td></td>
<td>Minnesota</td>
<td>1</td>
<td>&gt;0.1%</td>
</tr>
<tr>
<td>Graduate degree</td>
<td>51</td>
<td>1.2%</td>
<td></td>
<td>New Mexico</td>
<td>1</td>
<td>&gt;0.1%</td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td>112</td>
<td>2.7%</td>
<td></td>
<td>Nevada</td>
<td>1</td>
<td>&gt;0.1%</td>
</tr>
<tr>
<td>Some college</td>
<td>164</td>
<td>4.0%</td>
<td></td>
<td>Unknown</td>
<td>163</td>
<td>4.0%</td>
</tr>
<tr>
<td>High School</td>
<td>511</td>
<td>12.5%</td>
<td></td>
<td>Total</td>
<td>3,928</td>
<td>100.0%</td>
</tr>
<tr>
<td>Less than High</td>
<td>177</td>
<td>4.3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>3076</td>
<td>75.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100.0%</td>
</tr>
</tbody>
</table>
Table 2: COVID-19 Survivors

<table>
<thead>
<tr>
<th>Category</th>
<th>Sub-category</th>
<th>Value</th>
<th>SD or %</th>
</tr>
</thead>
<tbody>
<tr>
<td>N = 3,395</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, mean (sd)</td>
<td></td>
<td>78.2 years</td>
<td>12.6 years</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td>Female</td>
<td>1,898</td>
<td>55.9%</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td>Male</td>
<td>1,497</td>
<td>44.1%</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td>Asian</td>
<td>69</td>
<td>2.0%</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td>Black of African</td>
<td>343</td>
<td>10.1%</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td>Hispanic or Latino</td>
<td>206</td>
<td>6.1%</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td>Others</td>
<td>621</td>
<td>18.3%</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td>White</td>
<td>2,156</td>
<td>63.5%</td>
</tr>
<tr>
<td>Temperature to onset*</td>
<td>Days</td>
<td>2.7 days</td>
<td>13.7 days</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Mild</td>
<td>100</td>
<td>2.9%</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Moderate</td>
<td>72</td>
<td>2.1%</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Severe</td>
<td>3,223</td>
<td>94.9%</td>
</tr>
</tbody>
</table>

*Mean time in days from first 99°F temperature to onset of COVID-19
Table 3: COVID-19 Patient Comorbidities

<table>
<thead>
<tr>
<th>Comorbidity</th>
<th>Counts</th>
<th>% Patients with comorbidity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension</td>
<td>1,806</td>
<td>44.6%</td>
</tr>
<tr>
<td>Dementia</td>
<td>1,362</td>
<td>33.3%</td>
</tr>
<tr>
<td>Behavioral</td>
<td>1,081</td>
<td>26.4%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>995</td>
<td>24.3%</td>
</tr>
<tr>
<td>Depression</td>
<td>943</td>
<td>23.1%</td>
</tr>
<tr>
<td>Obesity</td>
<td>906</td>
<td>22.2%</td>
</tr>
<tr>
<td>Anemia</td>
<td>836</td>
<td>20.4%</td>
</tr>
<tr>
<td>Heart disease</td>
<td>726</td>
<td>17.8%</td>
</tr>
<tr>
<td>Psychosis</td>
<td>686</td>
<td>16.8%</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>612</td>
<td>15.0%</td>
</tr>
<tr>
<td>COPD or asthma</td>
<td>604</td>
<td>14.8%</td>
</tr>
<tr>
<td>Chronic kidney disease</td>
<td>541</td>
<td>13.2%</td>
</tr>
<tr>
<td>Anxiety</td>
<td>538</td>
<td>13.2%</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>431</td>
<td>10.5%</td>
</tr>
<tr>
<td>Cerebrovascular</td>
<td>421</td>
<td>10.3%</td>
</tr>
<tr>
<td>Morbid obesity</td>
<td>348</td>
<td>8.5%</td>
</tr>
<tr>
<td>Cancer</td>
<td>346</td>
<td>8.5%</td>
</tr>
<tr>
<td>Peripheral vascular</td>
<td>346</td>
<td>8.5%</td>
</tr>
<tr>
<td>Kidney failure</td>
<td>278</td>
<td>6.8%</td>
</tr>
<tr>
<td>Vitamin D deficiency</td>
<td>175</td>
<td>4.3%</td>
</tr>
<tr>
<td>Parkinson's</td>
<td>167</td>
<td>4.1%</td>
</tr>
<tr>
<td>Hyponatremia</td>
<td>135</td>
<td>3.3%</td>
</tr>
<tr>
<td>Hypokalemia</td>
<td>115</td>
<td>2.8%</td>
</tr>
<tr>
<td>Weight loss &gt; 5 pounds</td>
<td>63</td>
<td>1.6%</td>
</tr>
<tr>
<td>Other autoimmune</td>
<td>63</td>
<td>1.6%</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>58</td>
<td>1.4%</td>
</tr>
<tr>
<td>Dehydration</td>
<td>52</td>
<td>1.3%</td>
</tr>
<tr>
<td>Hyperkalemia</td>
<td>50</td>
<td>1.2%</td>
</tr>
<tr>
<td>Heart bypass</td>
<td>41</td>
<td>10.0%</td>
</tr>
<tr>
<td>Lupus</td>
<td>8</td>
<td>0.0%</td>
</tr>
</tbody>
</table>
Table 4: COVID-19 Mortality Rates

<table>
<thead>
<tr>
<th>COVID-19 Mortality Rates</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality rate in the COVID-19 palliative care population</td>
<td>33.6%</td>
</tr>
<tr>
<td>Mortality rate in the non-COVID-19 palliative care population</td>
<td>28.3%</td>
</tr>
<tr>
<td>Excess mortality</td>
<td>5.3%</td>
</tr>
<tr>
<td>Mortality rate in the COVID-19 non-palliative care population</td>
<td>9.9%</td>
</tr>
<tr>
<td>Mortality in the non-COVID-19 non-palliative care population</td>
<td>5.7%</td>
</tr>
<tr>
<td>Excess mortality</td>
<td>4.2%</td>
</tr>
</tbody>
</table>
Table 5. Predicted Average Treatment Effect (ATE) on Mortality

<table>
<thead>
<tr>
<th>Medication</th>
<th>% Reduction in mortality risk</th>
<th>Lower confidence interval range 95%</th>
<th>High confidence interval range 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omeprazole</td>
<td>-0.104</td>
<td>-0.262</td>
<td>0.046</td>
</tr>
<tr>
<td>Dextromethorphan</td>
<td>-0.232</td>
<td>-0.592</td>
<td>0.116</td>
</tr>
<tr>
<td>Apixaban</td>
<td>-0.115</td>
<td>-0.278</td>
<td>0.062</td>
</tr>
<tr>
<td>Cholecalciferol</td>
<td>-0.105</td>
<td>-0.283</td>
<td>0.067</td>
</tr>
<tr>
<td>Prednisone</td>
<td>-0.155</td>
<td>-0.432</td>
<td>0.103</td>
</tr>
<tr>
<td>Benztropine Mesylate</td>
<td>-0.134</td>
<td>-0.389</td>
<td>0.107</td>
</tr>
<tr>
<td>Enoxaparin Sodium</td>
<td>-0.069</td>
<td>-0.218</td>
<td>0.065</td>
</tr>
<tr>
<td>Sennosides</td>
<td>-0.063</td>
<td>-0.303</td>
<td>0.187</td>
</tr>
<tr>
<td>Guaifenesin</td>
<td>-0.027</td>
<td>-0.187</td>
<td>0.129</td>
</tr>
</tbody>
</table>
Table 6: List of Laboratory Values for Principal Component Analysis

<table>
<thead>
<tr>
<th>Laboratory Values</th>
<th>Creatinine MAX</th>
<th>Potassium MAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin</td>
<td>Creatinine MIN</td>
<td>Potassium MIN</td>
</tr>
<tr>
<td>ALK MAX</td>
<td>CRP</td>
<td>Platelets MAX</td>
</tr>
<tr>
<td>ALK MIN</td>
<td>Hemoglobin MAX</td>
<td>Platelets MIN</td>
</tr>
<tr>
<td>ALT MAX</td>
<td>Lactate</td>
<td>Prealbumin</td>
</tr>
<tr>
<td>ALT MIN</td>
<td>Magnesium MAX</td>
<td>Respiratory rate</td>
</tr>
<tr>
<td>AST MAX</td>
<td>Magnesium MIN</td>
<td>White blood cells MAX</td>
</tr>
<tr>
<td>AST MIN</td>
<td>Max Temperature</td>
<td>White blood cells MIN</td>
</tr>
<tr>
<td>BUN MAX</td>
<td>Phosphorous</td>
<td></td>
</tr>
<tr>
<td>BUN MIN</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Comprehensive Directed Acyclic Graph (DAG) for Counterfactual Analysis
Figure 2. Simplified Directed Acyclic Graph (DAG) for Counterfactual Analysis
END NOTES

1 https://data.cms.gov/covid-19/covid-19-nursing-home-data; https://www.kff.org/other/state-indicator/number-of-nursing-facility-residents/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D

2 Covid-19 Severity Classifications and Clinical Characteristics

<table>
<thead>
<tr>
<th>Severity Level</th>
<th>Clinical characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild-</td>
<td>Clinical symptoms without clinical or radiographic evidence of pneumonia; COVID+ with SPO2&gt;94</td>
</tr>
<tr>
<td>Moderate-</td>
<td>Clinical or radiographic evidence of pneumonia with SPO2 greater than or equal to 94; COVID-19+ PNA with SPO 94 or &gt;</td>
</tr>
<tr>
<td>Severe-</td>
<td>SPO2 less than 94% or RR &gt;than or equal to 30 +/- radiographic evidence of pneumonia +/- multiorgan failure clinical evidence pneumonia (abnormal LFTs, BUN/creatinine)</td>
</tr>
</tbody>
</table>


6 Ibid.


