TITLE
Comorbidities and sociodemographic factors on COVID-19 fatalities

AUTHORS
Jacob Gerken, M.Sc., Demi Zapata, M.Sc., Daniel Kuivinen, B.Sc. & Isain Zapata, Ph.D.

AFFILIATIONS
Department of Biomedical Sciences, Rocky Vista University, Parker, CO 80134, USA
Corresponding Author: Isain Zapata, PhD Address: 8401 S. Chambers Rd. Parker, CO 80134, USA. Phone: +1(720)795-9736 Email: izapata@rvu.edu

ABSTRACT
Introduction: Previous studies have evaluated comorbidities and sociodemographic factors individually or by type but not comprehensively. This study aims to analyze the influence of a wide variety of factors in a single study to better understand the big picture of their effects on case-fatalities.

Methods: County-level comorbidities, social determinants of health such as income and race, measures of preventive healthcare, age, education level, average household size, population density, and political voting patterns were all evaluated on a national and regional basis. Analysis was performed through Generalized Additive Models and adjusted by CCVI.

Results: Factors associated with reducing COVID-19 case fatality rates were mostly sociodemographic factors such as age, education and income, and preventive health measures. Obesity, minimal leisurely activity, binge drinking, and higher rates of individuals taking high blood pressure medication were associated with increased case fatality rate in a county. Political leaning influences case fatality rates. Regional trends showed contrasting effects where larger household size was protective in the Midwest, yet harmful in Northeast. Notably, higher rates of respiratory comorbidities such as asthma and COPD diagnosis were associated with reduced case-fatality rates in the Northeast. Increased rates of CKD within counties were often the strongest predictor of increased case-fatality rates for several regions.

Conclusion: Our findings highlight the importance of considering the full context when evaluating contributing factors to case-fatality rates. The spectrum of factors identified in this study must be analyzed in the context of one another and not in isolation.

KEYWORDS: Sociodemographic, CCVI, politics, county-level
INTRODUCTION

The SARS-CoV-2 Virus, also known as the COVID-19 virus, has currently led to over 990,000 (April, 2022) deaths in the United States. The virus has ravaged not only the United States financially and put a stop to every in-person activity for the last two years but has highlighted how embedded community traits affect the outcome of a whole community. In the end, it is characteristics like comorbidities and sociodemographic factors that play in a defining role in determining a community’s fatality outcome.

While it has been shown that a wide range of comorbidities has an impact COVID-19 outcomes; there has been a great effort to define the specific contributions of comorbidities in their impact on COVID-19 morbidity and mortality rates. Comorbidities such as hypertension, diabetes mellitus, chronic kidney disease, chronic obstructive pulmonary disease, and cardiovascular disease among others, have important repercussions on COVID-19 outcomes. However, there are also a variety of reasons for which to consider some sociodemographic factors as deleterious, such as lacking healthcare insurance, being of a specific racial background, or even voting patterns. All these factors combined would point in a direction that suggests that differences in COVID-19 fatality rates are a potential outcome of embedded community characteristics.

Even though previous studies have examined factors influencing COVID-19 fatalities, no study has performed a comprehensive analysis of all the aforementioned factors together to the same extent as our study. In our study, a multitude of key community indicators such as comorbidities, sociodemographic factors (including voting patterns), and determinants of health have been examined to reflect trends and potential associations that can be compared against each other. Therefore, the objective of this study was to perform a comprehensive evaluation of comorbidities, sociodemographic factors, and determinants of health at a national level using county aggregated to define their association to COVID-19 case-fatalities. These patterns may allow us to alter the way communities handle public health crises, utilize public health interventions that could deflect harmful outcomes, and provide resources to communities in a timely manner based on their community characteristics.

METHODS

Datasets

The focus of the study was to evaluate regional trends of COVID-19 case-fatality rate compared to comorbidities and sociodemographic factors. This study was vetted and categorized as exempt by the Institutional Research Board. Our study utilized countywide data for each county in the entire continental United States and Hawaii; Alaska and Puerto Rico were excluded from the analysis due to differences in their county data reporting. COVID-19 case-fatality rates were gathered from the COVID-19 Community Profile Report for the January 2-8, 2021 week cutoff, this report included the CCVI. This cutoff week was selected because
it allowed for the evaluation of the COVID-19 case fatality rate, without the influence of the vaccines or newer, more infectious strains. Rates of various comorbidities such as chronic obstructive pulmonary disease (COPD), hypertension, cancer, asthma, chronic heart disease (CHD), cholesterol, diabetes, chronic kidney disease (CKD), smoking, stroke, and obesity were obtained on a per-county basis from the CDC 2020 Population Level Analysis and Community Estimates (PLACES) project. Rates of poor mental health, binge drinking, lack of health insurance, time allocated to leisurely activity, and preventive care consisting of cervical cancer screening, routine doctor visits, dental visits, cholesterol screening, and routine mammography were obtained from the 2020 CDC PLACES Project, as well. Other variables such as average household size and population density for each county were acquired from the United States Census Bureau COVID-19 website. Latitude of each county was also included in the analysis and obtained from the United States Census Bureau. The 2020 Presidential voting records of each county were obtained from the Harvard Dataverse. Racial makeup in each county was obtained from the 2020 decennial United States census, while income, age, and education level were retrieved from the 2019 American Community Survey 5-year estimates.

**Statistical analysis**

Data was evaluated for associations using a Generalized Additive Models (GAMs) approach. COVID-19 case fatality rates per 100k people were set as the dependent variable while each comorbidity, sociodemographic and health determinant factor was set as an independent variable. All models were adjusted using COVID-19 Community Vulnerability Indexes (CCVI) which normalized the data for inherent inequity. This study uses individual counties as the experimental unit. All analyses were evaluated on the national sample and in a per region samples, as to determine regional patterns. Independent variables were introduced into the model using smoothing splines starting with three degrees of freedom. Models assumed Gaussian residual distributions. All analyses were performed using PROC GAM in SAS/STAT v.9.4 (SAS Inc., Cary, NC). Risk ratios were estimated with 99% confidence intervals and the coefficients sign determined effect directions. Negative coefficients indicated a reduction in COVID-19 case-fatality rates while positive coefficients indicated an increase in case-fatality. Coefficient standardization was done with a normally distributed $Z$-score transformation. All associations presented were tested using two-tailed tests. Regional pattern models were performed independently to identify the top contributors - negatively and positively associated. Even with 99% confidence intervals, all tests were declared significant at a Bonferroni threshold. Regional pattern top contributors that did not reach the Bonferroni threshold are indicated in the figures.

**RESULTS**

**National level trends**

Data from 3140 counties was included in our analysis. Our GAM approach examined the data for associations to case-fatality rates, all these values were adjusted to CCVI to normalize the inherent differences a raw
risk ratio would estimate, and the risk ratio standardized estimates are presented in Figure 1. Raw estimates showcase the extent of the association without considering their spread while standardized estimates adjust the extent of the association to the spread. Raw estimate findings (Figure 1A) revealed that comorbidities above sociodemographic factors have the largest effects associated with case-fatalities; however, these associations can go in both positive and negative directions. A diagnosis of cancer provided the largest effect decreasing COVID-19 case-fatalities while CKD and stroke had the largest effects increasing them. Similarly, asthma (decreased risk), CHD (increased risk) and diabetes (increased risk) displayed moderate effects. Household size was the largest significant sociodemographic factor in positive outcomes of COVID-19 with an effect in a range comparable to relevant comorbidities. Other demographic effects such as age and education displayed significant associations that reduced or increased case-fatality risk. Populations with higher educational achievements displayed significantly reduced case-fatality rates. Increased income always displayed a protective effect. Political preference was significantly associated with case-fatalities such as voting for Biden reduced case-fatalities while voting for Trump had the opposite effect. Racial and ethnic backgrounds were only associated to COVID-19 case–fatalities for Pacific Islanders, Asian and Black groups. Determinants of health such as cervical cancer screening and people using high blood pressure medication also showed mixed direction associations. Cervical cancer screening had the largest case-fatality reducing effect from this category while the people using high blood pressure medication had the largest opposite effect. Standardized estimates (Figure 1B) show a different perspective and allow for comparisons across factors as they are standardized. In this case, a routine colonoscopy procedure was found to be the largest protecting effect against COVID-19 case-fatality followed by a combination of sociodemographic factors such as age, education, and income. On the other side of the spectrum, obesity had the largest negative impact deleterious outcome in COVID-19 patients followed by having no leisure physical activity, binge drinking and higher proportions of people taking high blood pressure medication in a specific county.

Regional trends

The main analysis was also replicated independently within each of the ten US Health and Human Services defined regions. These models were also adjusted by CCVI. Risk ratio effect estimates for the ten regions are displayed in Figure 2. These analyses detected a wide array of effects that in some cases go in opposite directions across all regions. No single factor was consistently associated for all regions suggesting that regional associations are not generalizable. These regional assessments all have different sample sizes that are based on the number of counties within each state. These can range from 67 (Region 1) to 736 (Region 4); however, this discrepancy did not affect the capacity of each regional analysis to detect associations at a Bonferroni level (adjusted for 470 tests across all sets). The top variables reducing and increasing COVID-19 case-fatalities for each region are presented in Figure 3. The map in Figure 3A shows that the strongest protective...
regional effects were observed towards the east of the country where Stroke and Cancer Diagnosis were highly protective in the Northeast regions (Regions 1 and 2) and being of Pacific Islander descent was protective in the Southeast United States (Region 4). The Midwest displayed some moderate protective effects where household size was the top reducing factor in two regions (Region 5 and 7). Western regions displayed smaller COVID-19 case-fatality protective effects. Regions in the Western United States displayed smaller effect sizes in comparisons to regions in the South or Midwest. Lastly, Figure 3B shows variables that most significantly contributed to increased COVID-19 case-fatalities. Chronic kidney disease (CKD) was the most prevalent comorbidity across several regions (Regions 4, 6, 8 and 9). Among other findings, household size had a negative impact on COVID-19 outcomes specifically in the Northeast regions (Regions 1 and 2). As previously mentioned, some of the top variables displayed opposing effects which suggested that the interpretation must be done in context with the specific characteristics of that region, and interpretations cannot be generalized to others.

**DISCUSSION**

The objective of this paper was to evaluate COVID-19 case-fatality rate in conjunction with sociodemographic factors, comorbidities, and determinants of health. Previous studies have evaluated the influence of various socioeconomic factors 7,14,24 and comorbidities 4-6 on COVID-19 case-fatality rate; however, these analyses do not pair together their findings to be comparable with each other. Our study evaluates COVID-19 fatality rates from a wide timeframe, without potential influence from vaccines reducing case-fatality rate and the addition of major COVID-19 variants. Our study builds on the efforts of previous studies by presenting together a wide array of variables that describe community characteristics. It is necessary to emphasize that the associations between these factors and case-fatality is not necessarily or entirely causative. All comorbidities, sociodemographic and determinants of health variables presented describe characteristics of the population that are not isolated or independent. Therefore, the causality that could be attributed to each factor evaluated must always be provided with context as a community indicator as they are all dependent or interconnected on each other, examples of this are binge drinking, mammography and visits to the dentist rates, which are likely indirectly describing a characteristic of the community. In general, all these variables must be interpreted in a continuum of causality that can vary across regions depending on the context.

Chronic kidney disease rates were the strongest predictor of increase COVID-19 case-fatality in several US regions. This relationship is likely to be predominantly causal due to people with this condition being medically vulnerable 25-28. Other comorbidities followed a similar trend such as higher rates of hypertension, diabetes, stroke, and smoking being associated with an increased case-fatality rate in a specific county. This finding aligns with other studies 5,29,30 linking increased rates of comorbidities to poorer COVID-19 outcomes. Even though comorbidities were most often associated with worse COVID-19 outcomes, stroke and cancer
diagnosis were linked to reduced case-fatality rates in the Northeast region. We speculate that a potential
explanation for this relationship is possible more frequent mask usage and more precautions taken by this
group of people. Northeastern states had mask adherence rates greater than 75% during the latter part of
2020 with usage potentially diminishing the influence of comorbidities such as asthma and COPD on COVID-19
case-fatality rate. Those with asthma and COPD in these communities were more likely to wear a mask, further reducing their chance of acquiring and succumbing to COVID-19.

Household size was identified early on to be a risk factor for COVID-19 transmission. Our results have shown a contrasting effect that supports this notion in the Northeast but has the opposite effect in the Midwest. The difference between these regions is likely related to the specific context of living conditions. Although the mean household size for regions 1 and 2 is not far from the mean household size for regions 5 and 7, with 2.48 and 2.41 respectively, this difference may capture differences in housing quality and composition. This may also be indirectly identifying behavioral factors that are not obvious but can be implied such as house proximity (higher in cities) which can affect the capacity of self-isolation. Population density could be partially be
influenced by household size, which has shown to have an impact on transmission. The Northeast states have higher population densities when compared to midwestern states. In summary, living conditions, housing quality and composition, and population density all be important component that define the impact of household size on case fatality rate. Generally, higher income was also associated with decreased COVID-19 case fatality rates.

There could be a multitude of reasons why a higher income may be beneficial. This could include not being classified as an essential worker leading to being more likely to take time off or work from home, or even being able to live outside high-density population and compact housing areas. Income is often reflective and associated with racial discrepancies in COVID-19 outcomes. In a study examining neighborhood median income and COVID-19, when examining the neighborhoods, Black populations more often lived in neighborhoods with a significantly lower median income ($35,000) whereas White populations more often lived in wealthier neighborhoods ($63,000). Communities with lower median incomes are more often Medicaid patients and have COVID-19 complications that require invasive tactics such as mechanical ventilation. Income inequalities have been strong predictors of higher case numbers and fatalities throughout this pandemic.

The context that defines the influence of social dynamics on COVID-19 is complicated. Political affiliation has been repeatedly evaluated as a potential factor influencing the pandemic’s mortality. This pattern is not new, as Republican party affiliation was associated with decreased influenza vaccinations and the Democrat party was associated with increased compliance. The politicization of pandemic response has continued into the COVID-19 pandemic, with behaviors such as masking, social distancing, and vaccination being often divided along party lines. Some studies have shown a decrease in pandemic preventive health measures among Republicans while there has been increased adherence to public health recommendations among Democrats.
Similarly, Republicans have been shown to have lower COVID-19 risk perception, when compared to other political parties which may influence their likelihood of contracting COVID-19. The behaviors exhibited by each political party may influence the results of this study. Our study showed that voting for Joseph Biden in the 2020 presidential election was mildly associated with decreased case-fatality rate while voting for Donald Trump was associated with increased rate. With Democrats being shown to be more likely to adhere to public health guidelines, they may be less likely to acquire and perish from COVID-19 while the inverse is true for Republicans. In December 2020, states with Republican governors had higher rates of cases, deaths, and positive tests than states with Democrat governors. This trend is evident in a similar approach using national data presented by NPR in collaboration with researchers at John’s Hopkins University where it was shown that voting Republican also had a deleterious effect. Rural and urban differences have been shown to play a major role in case-fatality rate as well with rural counties having a higher case-fatality ratio than urban counties. Rural voters are more likely to vote Republican and therefore, the influence of politics in our findings may also be capturing geographical differences. Rural areas tend to have worse health outcomes in general and have significantly less access to care compared to their urban counterparts. These disparities add to the likelihood of developing comorbidities and ultimately, poorer COVID-19 outcomes. The link between political affiliation and COVID-19 case fatality rate is far more complex than the individual candidates that people of a county voted for. Political affiliation in our study is an indicator of underlying sociodemographic, health, and psychological trends that are more causative rather than associative.

Limitations and future directions

Our study utilized aggregate data on a per-county basis instead of individual patient data; therefore, it is not possible to evaluate factors that contribute to COVID-19 case-fatality on a per case fashion which could help avoid any erroneous generalizations of specific regions. Another limitation of using county level data is that there is significant variability in the size and number of counties across the United States. Some counties may have only a few hundred people, while other counties may have a few million and this may lead, to some extent, representation bias. Future directions of this study include using these results to guide the evaluation of individuals’ factors that contribute directly to illness outcomes.

Conclusion

Our study evaluated a multitude of factors that may affect COVID-19 case-fatality rate. Unlike previous studies that evaluated these factors separately, we performed a comprehensive analysis of all these variables together where they interact amongst each other. We identified several unique regionally dependent and independent relationships that highlighted the various factors that might influence COVID-19. Like other studies, we determined that comorbidities and demographic factors together are strong drivers of COVID-19 case fatalities. However, our study presents an assessment that puts them side to side for direct comparison. Our
study highlights how any association is often dependent on the regional context. For example, household size in the Northeastern region of the United States was associated with more fatalities, while larger household size in the Midwest regions had a protective effect. Political voting patterns were also indicative of underlying healthcare patterns, with overall reduced case fatality rates in Democratic voting counties compared to increased fatality rates in Republican voting counties. The trends we identified in our study emphasize the importance of interpreting each factor in the context of other variables instead of in isolation.

CONFLICT OF INTEREST DISCLOSURES

The authors have no conflict of interests to declare.

FUNDING

No special funding was received for this study.

DATA AVAILABILITY

All data used in the study is publicly available. Curated datasets can be made available with a reasonable request to the corresponding author.

AUTHOR CONTRIBUTIONS

Jacob Gerken: Conceptualization, Methodology, Data Curation, Investigation, Writing of Original Draft. Demi Zapata and Daniel Kuivinen: Conceptualization, Methodology, Investigation, Writing of Original Draft. Isain Zapata: Conceptualization, Methodology, Investigation, Formal Analysis, Visualization, Supervision, Writing and Editing of Final Draft. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

REFERENCES

4. Iyanda AE, Boakye KA, Lu Y, Oppong JR. Racial/Ethnic Heterogeneity and Rural-Urban Disparity of COVID-


32. Faria N, Costa MI, Gomes J, Sucena M. Reduction of Severe Exacerbations of COPD during COVID-19


358    46. Milligan MA, Hoyt DL, Gold AK, Hiserodt M, Otto MW. COVID-19 vaccine acceptance: influential roles of


doi:10.1177/19485506211055295


doi:10.1016/j.amepre.2021.01.034


**Figure 1. Comorbidity, sociodemographic and determinants of health associations to COVID-19 case-fatalities.**

All models are CCVI adjusted. A) Risk ratio raw estimates B) Risk ratio standardized estimates (ordered). Blue diamonds indicate case-fatality reduction, red diamonds indicate case-fatality increase, and Black diamonds indicate no association. A total of 3140 counties were included in the study. Even when 99% CI are presented, association are declared significant at a Bonferroni adjusted threshold (47 tests $P_{adj} \leq 1.06E-03$).
<table>
<thead>
<tr>
<th>Variable</th>
<th>US Health and Human Services Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Region 1</td>
</tr>
<tr>
<td>Comorbidities - Arthritis (%)</td>
<td>-14.66</td>
</tr>
<tr>
<td>Comorbidities - Asthma (%)</td>
<td>-40.97</td>
</tr>
<tr>
<td>Comorbidities - Bad Mental Health (%)</td>
<td>-29.56</td>
</tr>
<tr>
<td>Comorbidities - Bad Physical Health (%)</td>
<td>-25.44</td>
</tr>
<tr>
<td>Comorbidities - Binge Drinking (%)</td>
<td>-32.49</td>
</tr>
<tr>
<td>Comorbidities - Cancer diagnosis</td>
<td>-114.77</td>
</tr>
<tr>
<td>Comorbidities - CHD (%)</td>
<td>-57.94</td>
</tr>
<tr>
<td>Comorbidities - CKD (%)</td>
<td>-132.53</td>
</tr>
<tr>
<td>Comorbidities - COPD (%)</td>
<td>-39.76</td>
</tr>
<tr>
<td>Comorbidities - Diabetes (%)</td>
<td>4.87</td>
</tr>
<tr>
<td>Comorbidities - High Blood Pressure (%)</td>
<td>7.92</td>
</tr>
<tr>
<td>Comorbidities - High Cholesterol (%)</td>
<td>29.67</td>
</tr>
<tr>
<td>Comorbidities - No Leisure Physical Activity (%)</td>
<td>-11.68</td>
</tr>
<tr>
<td>Comorbidities - Obesity (%)</td>
<td>-6.94</td>
</tr>
<tr>
<td>Comorbidities - Poor Sleeping (%)</td>
<td>7.61</td>
</tr>
<tr>
<td>Comorbidities - Smoking (%)</td>
<td>-14.24</td>
</tr>
<tr>
<td>Comorbidities - Stroke (%)</td>
<td>-125.62</td>
</tr>
<tr>
<td>Demographics - 25 to 64 years of age (%)</td>
<td>8.86</td>
</tr>
<tr>
<td>Demographics - Household Size (Persons)</td>
<td>96.97</td>
</tr>
<tr>
<td>Demographics - No Insurance (%)</td>
<td>-15.97</td>
</tr>
<tr>
<td>Demographics - Over 65 years of age (%)</td>
<td>9.63</td>
</tr>
<tr>
<td>Demographics - Under 25 years of age (%)</td>
<td>9.99</td>
</tr>
<tr>
<td>Education - Bachelor's or Higher (%)</td>
<td>3.58</td>
</tr>
<tr>
<td>Education - High School (%)</td>
<td>-7.08</td>
</tr>
<tr>
<td>Education - No High School (%)</td>
<td>12.78</td>
</tr>
<tr>
<td>Education - Some College or Associate's (%)</td>
<td>-7.38</td>
</tr>
<tr>
<td>Geography - Latitude (degrees)</td>
<td>-36.75</td>
</tr>
<tr>
<td>Healthcare - Carcinoma Screening (%)</td>
<td>17.67</td>
</tr>
<tr>
<td>Healthcare - Cholesterol Screening (%)</td>
<td>21.38</td>
</tr>
<tr>
<td>Healthcare - Colonoscopy (%)</td>
<td>9.57</td>
</tr>
<tr>
<td>Healthcare - High Blood Pressure Medication (%)</td>
<td>17.86</td>
</tr>
<tr>
<td>Healthcare - Mammography Use (%)</td>
<td>11.68</td>
</tr>
<tr>
<td>Healthcare - Routine Checkup (%)</td>
<td>13.81</td>
</tr>
<tr>
<td>Healthcare - Visited Dentist (%)</td>
<td>7.42</td>
</tr>
<tr>
<td>Income - Families Median (USD$)</td>
<td>2.91</td>
</tr>
<tr>
<td>Income - Household Median (USD$)</td>
<td>2.96</td>
</tr>
<tr>
<td>Income - Married Median (USD$)</td>
<td>2.46</td>
</tr>
<tr>
<td>Income - Nonfamily Median (USD$)</td>
<td>4.32</td>
</tr>
<tr>
<td>Politics - Voted for Biden (%)</td>
<td>2.21</td>
</tr>
<tr>
<td>Politics - Voted for Trump (%)</td>
<td>-1.74</td>
</tr>
<tr>
<td>Race - Asian (%)</td>
<td>15.58</td>
</tr>
<tr>
<td>Race - Black (%)</td>
<td>12.45</td>
</tr>
<tr>
<td>Race - Hispanic (%)</td>
<td>7.43</td>
</tr>
<tr>
<td>Race - Native American (%)</td>
<td>47.13</td>
</tr>
<tr>
<td>Race - Other (%)</td>
<td>-12.96</td>
</tr>
<tr>
<td>Race - White (%)</td>
<td>-4.75</td>
</tr>
<tr>
<td>Race - Pacific Islander (%)</td>
<td>-275.10</td>
</tr>
</tbody>
</table>


**Figure 2. Regional risk ratio estimates for comorbidity, sociodemographic and determinants of health in association to COVID-19 case-fatalities.** All models were performed independently by region and are adjusted to CCVI. Sample sizes correspond to the number of counties in the analysis for each region. Estimates signs indicate effect direction. Red boxes are significant to a Bonferroni level ($P_{adj}$≤0.04). Orange boxes are significant to a 95% confidence level.
Figure 3. Top county level factors associated to COVID-19 case-fatalities for each United States Department of Health and Human Services regions. A) Top factors associated with reduced case fatalities B) Top factors associated with increased case-fatalities. Shading indicates the effect size across regions (adjusted by CCVI). Alaska was excluded from the analysis because of differences in their county level reporting.