

SAS® GLOBAL FORUM 2021

Paper 1191-2021

Disproportional Impact of COVID-19 on Marginalized Communities

David J Corliss, Peace-Work

ABSTRACT

The impact of COVID-19 pandemic has been especially severe among marginalized communities, including persons of color, Native Americans, LGBTQ+, the homeless, and prison populations. This study leverages county-level COVID mortality data from the New York Times database and socio-economic and demographic data from the American Community Survey to identify risk factors and calculate odds ratios for COVID-19 mortality among various marginalized communities. Disproportionate impact is assessed using only the first wave of the pandemic in order to measure intrinsic risk unaffected by politically-driven behaviors and associated mortality developing in later waves of the pandemic. Per capita mortality is calculated for sets of counties with 60%+ population of at-risk groups and compared using odds ratios to the set of counties with less than 5%. These risks are compared to other COVID-19 mortality risk factors, e.g., chronic lung disease. This study finds the two factors with highest risk, with an odds ratio of about 10, are persons of color and cardiac patients. Implications of these results are discussed to advise potential mitigation strategies.

INTRODUCTION: DISPROPORTIONATE IMPACT

Disasters of all kinds often have a particularly severe impact on minorities and other marginalized communities, and COVID-19 has been no exception. There have been reports of high death rates among African American, Native American, immigrants, prison populations, and other marginalized groups. This analysis seeks to quantify the disproportionate impact of the COVID-19 pandemic among different groups. Investigating the relevant data, a number of issues were identified. Metrics need to be established to track the pandemic. Incomplete race / ethnicity data at the case level leads to taking a propensity approach using complete data by geography and Census Bureau demographics. The analysis also separates the effect of intrinsically higher risk among traditionally marginalized population from the politicization of this pandemic in the US that has resulted in behavior-driven morbidity and mortality for certain groups.

Early evidence of disproportionate impact is seen in a case study in the Detroit metropolitan area. Wayne County, Michigan, containing the city of Detroit, had the highest COVID mortality rate outside of the New York City area among counties with at least 100,000 residents. This geography is one of the most racially separated areas of the United States, with the city of Detroit having 14.7% white population in comparison to 80.2% in the surrounding suburbs, according to US Census Bureau data (2019). The prevalence of poverty versus affluence is markedly different also, the suburban areas of the Detroit area consist of Wayne County outside the city of Detroit, together with Oakland and Macomb counties. Analysis is facilitated by the State of Michigan reporting COVID-19 statistics separately for the city of Detroit and Western Wayne county, the only county in the state reporting by geographic subdivision.

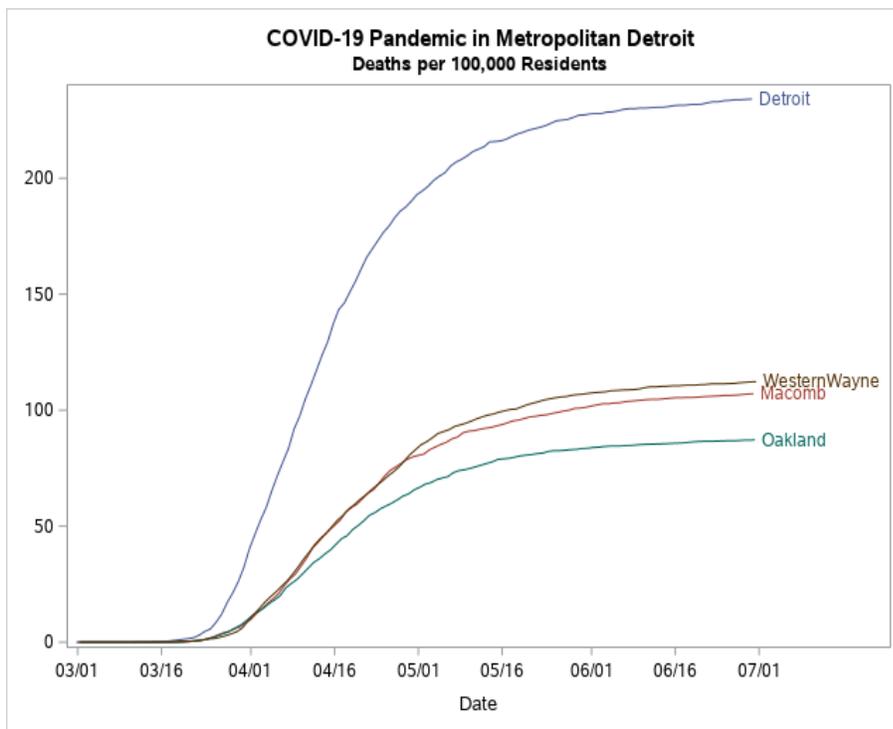


Figure 1. COVID-19 First Wave Cumulative Deaths per Capita in Metropolitan Detroit

Note the use of CURVELABEL in the SAS code used to create this plot is:

```

title1 'COVID-19 Pandemic in Metropolitan Detroit';
title2 'Deaths per 100,000 Residents';
proc sgplot data=c19.detroit;
  where date le mdy(6,30,2020);
  series x=date y=deaths_100k / group=county groupdisplay=cluster curvelabel;
  xaxis label="Date" type=time valuesformat=mmddy5.;
  yaxis display=(nolabel);
run;
title1 ''; title2 '';

```

While Detroit was an early hotspot in the COVID-19 pandemic, Figure 1 shows this was confined to a small area with a very high population percentage of BIPOC (Black, Indigenous, Persons of Color). The three suburban geographies present very similar time series. Despite their close proximity, the first wave of the COVID-19 pandemic was far more severe in the city of Detroit than its suburbs. Distinctly higher mortality was apparent by early April. In the First Wave, through the end of June, the mortality odds ratio of the city versus the suburbs was 7.12.

STUDY METHODOLOGY

Official US COVID-19 statistics from the Centers for Disease Control (Figure 2, updated March 6, 2021) includes confirmed cases, hospitalizations, and deaths from COVID-19. While these reported risk ratios appear to indicate greater risk for American Indian / Alaskan Native, Black / African American, and Hispanic / Latinx populations, there are concerns with the methodology. It is noted that that the data reported to the CDC is very incomplete: only 53% of case reports include race or ethnicity. This leads the possibility of not only complete data, but also sample and response biases in the cases that are reported.

The CDC Race / Ethnicity cannot be recommended for the analysis of the impact of the pandemic on marginalized communities.

Risk for COVID-19 Infection, Hospitalization, and Death By Race/Ethnicity				
Rate ratios compared to White, Non-Hispanic persons	American Indian or Alaska Native, Non-Hispanic persons	Asian, Non-Hispanic persons	Black or African American, Non-Hispanic persons	Hispanic or Latino persons
Cases ¹	1.7x	0.7x	1.1x	1.3x
Hospitalization ²	3.7x	1.0x	2.9x	3.1x
Death ³	2.4x	1.0x	1.9x	2.3x

Race and ethnicity are risk markers for other underlying conditions that affect health, including socioeconomic status, access to health care, and exposure to the virus related to occupation, e.g., among frontline, essential, and critical infrastructure workers.

Figure 2. Reporting by Race/Ethnicity: Incomplete Data Reported Over All Time

In addition, the point-in-time data provided by the CDC averages the time series across all waves of the pandemic. In the United States, response to the pandemic after the initial wave has become highly politicized, leading politically-driven behaviors such as disregard of recommended safety precautions such as wearing masks and social distancing.

As a result, the CDC report understates the intrinsic risk, which is most apparent at the very beginning of the pandemic, *before* politically-driven risk behaviors emerged. This happens because of politicization, which has been highly present in the case of the COVID 19 pandemic level cold responses have created a group of degree of behavior driven more bit busy a mortality. Factor driven more by behaviors and by intrinsic underlying risk, and it tends to mask the under this underlying risk. Point-in-time data, looking at the entire pandemic up to a certain date, masks the actual underlying risk. If the pandemic in the US is instead analyzed by interval as it evolves, the intrinsic risk can be ascertained from the very beginning of the pandemic. Once the intrinsic risks by race / ethnicity is measured, it can be compared to subsequent intervals to estimate the risk associated with politically-driven behaviors that have played an important role in the COVID-19 pandemic in the United States.

COVID-19 cases often go unreported due to being asymptomatic or so mild that they are not reported to local health authorities. Deaths, while not perfect, are much more likely to be reported. To support comparison between different geographic areas, mortality counts must be normalized to population size. This established deaths per capita as the best metric to analyze the pandemic.

As shown in Figure 3, the time series of this pandemic has been complex, presenting as a series of waves.

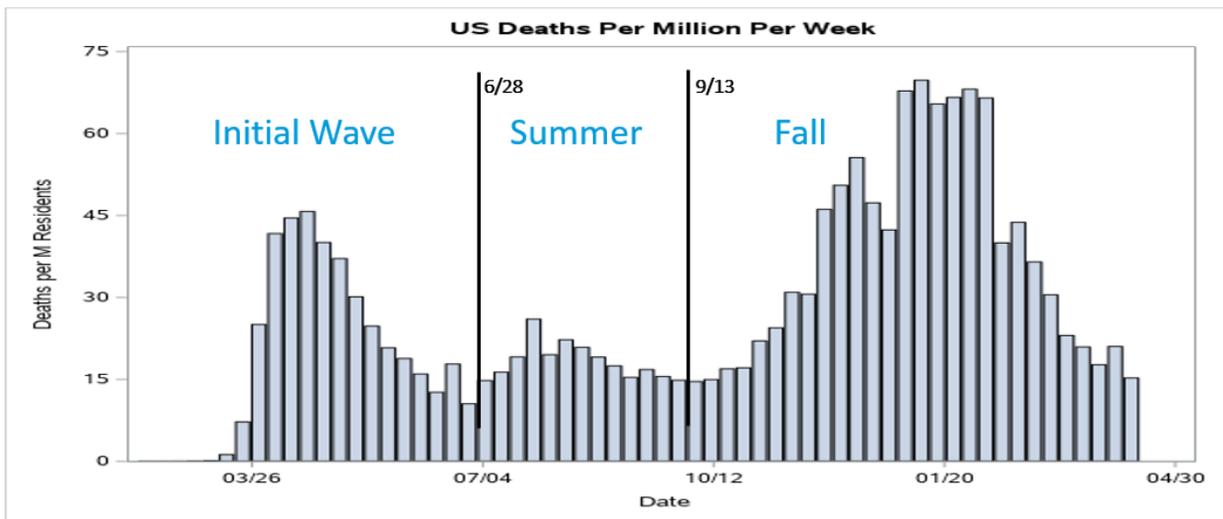


Figure 3. Pandemic US Mortality Rate by Week

In the evolution of the pandemic in the US, substantial variation by geographic area is seen (Figure 4). As seen in the Metro Detroit case study, some areas were much more impacted in the first wave – especially those places with a high percentage BIPOC populations. However, the second, Summer wave, was only found in some geographic areas. This wave is present in the time series in the US, not other countries, and only in a subset of US states.

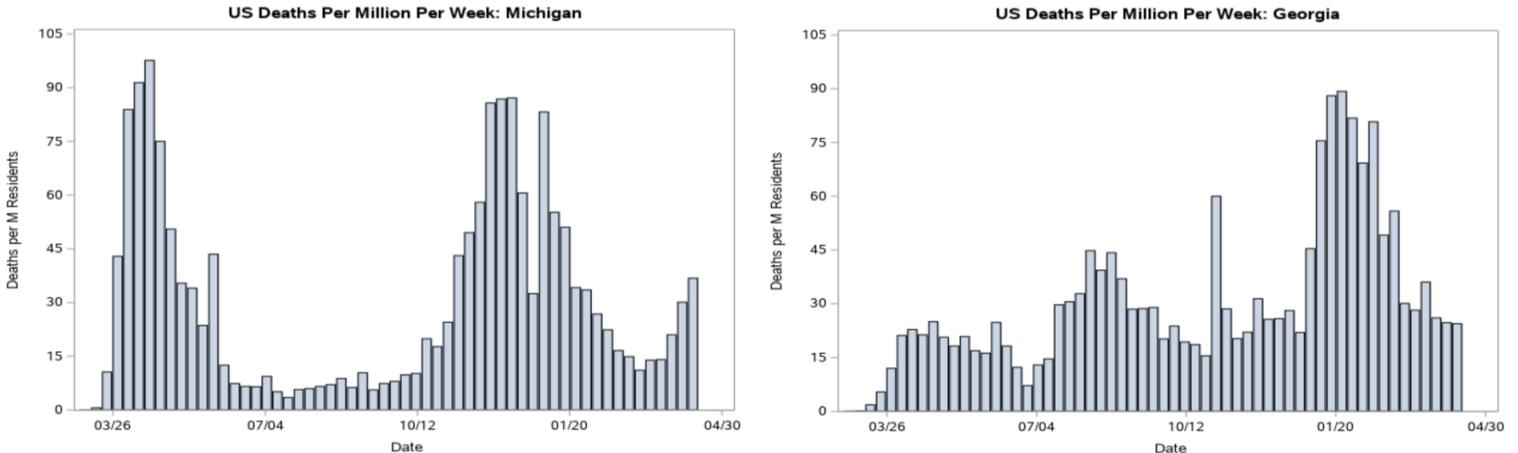


Figure 4. Variation by Geographic Area

The areas where the Summer wave was prominent is the geographic subset of the United States marked by politically driven denial of the pandemic and resistance to safety precautions.

The First Wave, reflecting intrinsic risk, was most prominent in areas with high percentage BIPOC populations. As the pandemic evolved, behavior-driven risks emerged: rejection of and even large protests against safety precautions such as wearing masks and social distancing. In terms of methodology, the impact of the first wave indicates intrinsic risk - before full understanding of the virus and how to address it. The start of the summer wave,

where the increase is driven by behavioral factors, marks the end of the period when intrinsic risk can be measured separately. Accordingly, the key metric for this study is the cumulative COVID-19 deaths per capita during the First Wave from January 1 through June 30, 2020.

COUNTY-LEVEL ANALYSIS

With the outcome key metric established, demographic data was obtained from the US Census Bureau American Community Survey. Data were compiled at a US county level. Quintile subsets of counties by percent BIPOC were created using PROC RANK:

```
Proc rank data=c19.Phase1_DeathsByCounty groups=5
  out=quint(keep=state county pop_2019 deaths_p1_per_m bipoc_q);
  var bipoc_pct;
  ranks bipoc_q;
  where bipoc_pct not = .;
run;
```

PROC RANK assigns values beginning with 0, so a value of 1 was added to each quintile to number them 1 through 5. PROC MEANS was used to find the median mortality rate for the counties in each quintile:

```
data work.quint;
  set work.quint;
  bipoc_q = bipoc_q + 1;
run;

proc sort data=work.quint;
  by bipoc_q;
run;

proc means data=work.quint;
  var deaths_p1_per_m;
  by bipoc_q;
  weight pop_2019;
  output median=deaths_p1_per_m out=work.plot_data;
run;
```

These data were plotted using a Vertical Bar Chart with SGPLOT:

```
title 'BIPOC Mortality Rate by Population Quintile - Phase 1';
proc sgplot data=work.plot_data;
  xaxis label="BIPOC % of Population - US Counties by
  Quintile(*ESC*) {unicode '000a'x} (Increasing BIPOC % --->);
  yaxis label="Deaths per M Residents" values=(0 to 300 by 50);
  vbar bipoc_q / response=deaths_p1_per_m datalabel;
run;
title '';
```

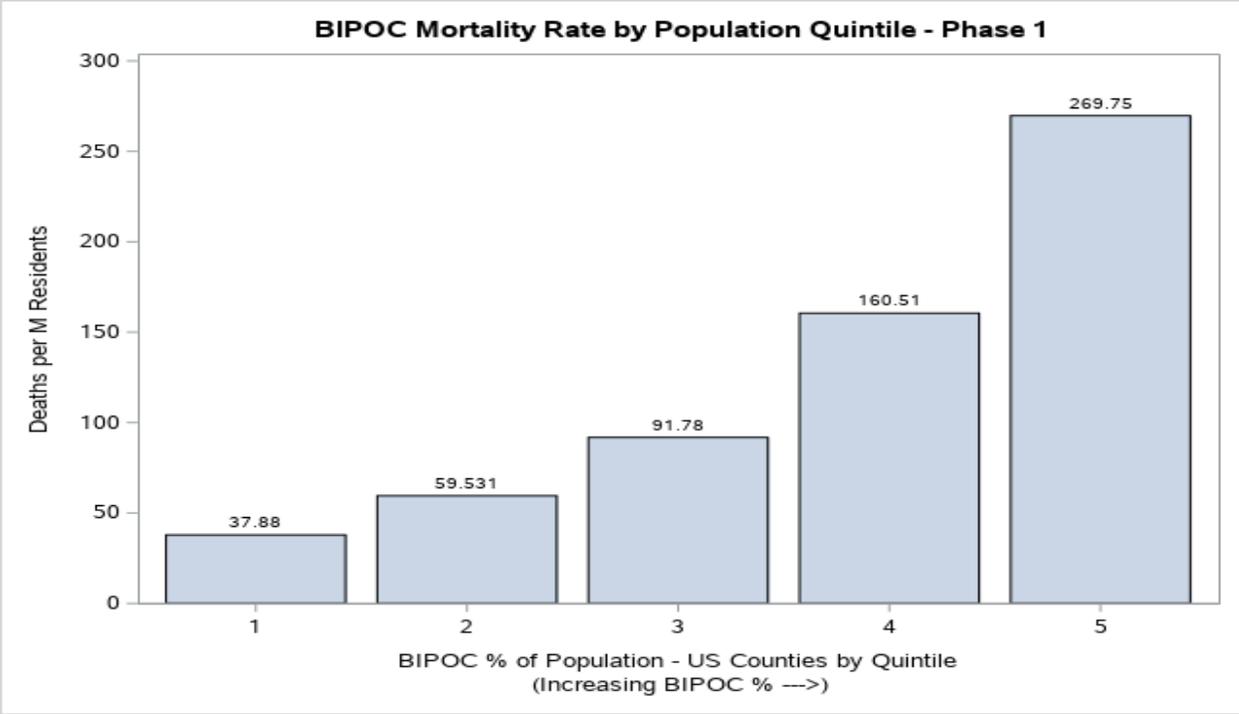


Figure 5. COVID-19 Mortality Rate by % BIPOC Quintile

Dividing all US counties into quintiles from lowest to highest BIPOC % of total population, each successive quintile has higher COVID mortality (Figure 5).

Comparing this to the summer wave (Figure 6), a demographic shift is seen with death rates increasing in largely white areas quintiles while decreasing in communities with the largest percent BIPOC population.

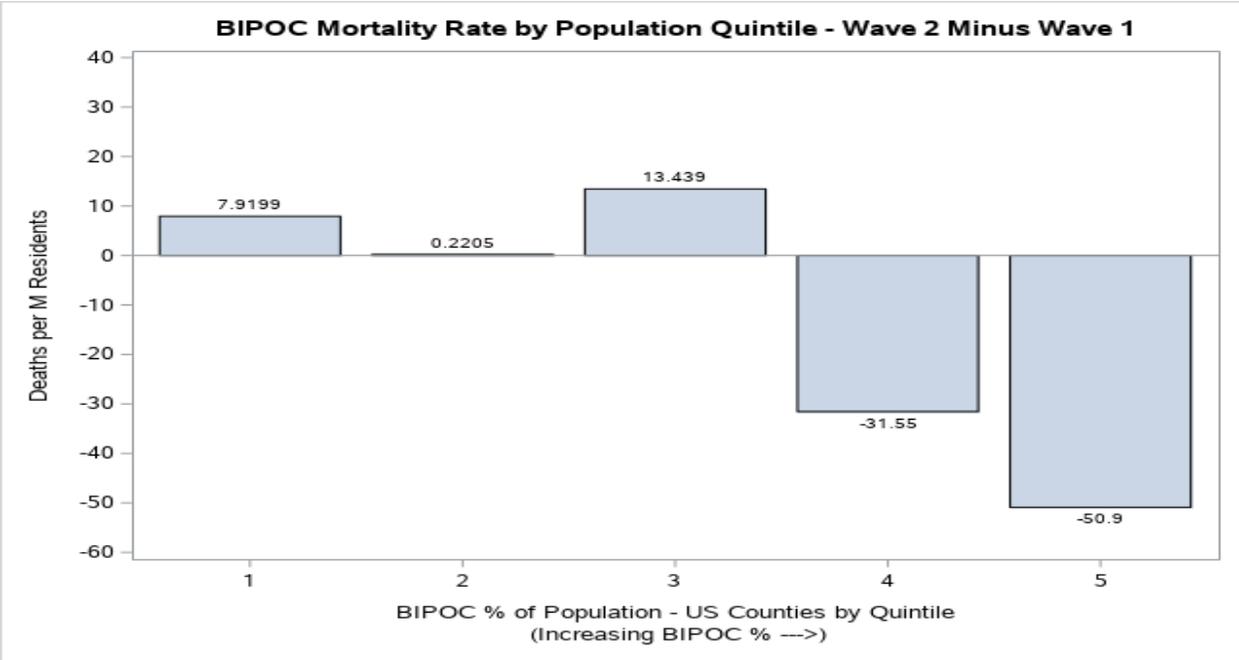


Figure 6. Change in COVID-19 Mortality Rate During the Second Wave

This shift reflects the behavior-driven nature of the second wave. Keeping in mind that no countries outside the US experienced a COVID-19 wave at this time, this analysis finds the Summer Wave was predominately a characteristic of a subset of US areas with a small percentage of BIPOC which analysis of the First Wave found had the highest intrinsic COVID-19 risk.

This analysis finds disproportionate impact on of the COVID-19 pandemic on BIPOC communities. An important dimension of this result is government management of this challenge. This is especially important due to politicization of pandemic management. This effect can be investigated by dividing states into Red and Blue based on the political party of the Governor (Mayor in the case of Washington, DC) to indicate the affiliation of the administration. Identifying Red and Blue states in this matter, summing the deaths by month and calculating the average mortality rates for each provides the time series plotted in Figure 7:

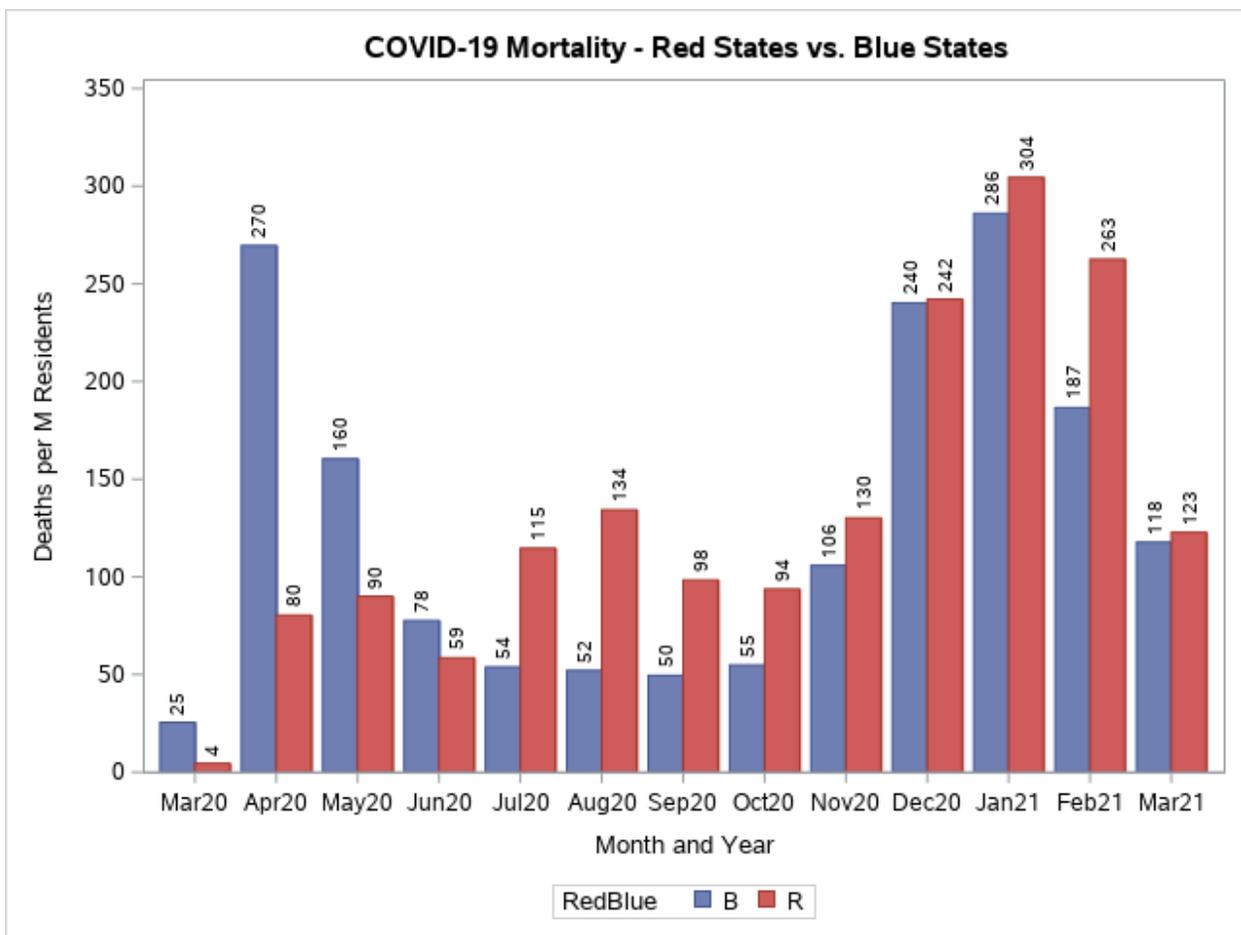


Figure 7. Red State vs. Blue States: Different Evolution of the Pandemic

The SAS code used to create this plot is:

```

title 'COVID-19 Mortality - Red States vs. Blue States';
proc sgplot data=work.plot_data;
  xaxis label="Month and Year";
  yaxis label="Deaths per M Residents" values=(0 to 350 by 50);

```

```

vbar date / response=deaths_per_m group=redblue groupdisplay=cluster
datalabel;
format date monyy5. deaths_per_m comma8.0;
where date between mdy(3,1,2020) and mdy(3,31,2021);
run;
title '';

```

Examining the results by wave, we find

- The First Wave, through June 2020, shows the initial impact much was more severe in Blue States, dominated by poor urban centers with a large BIPOC population.
- The Summer wave, in July – October, only occurs in Red States: the mortality rate in Blue states stayed flat. This also allows us to estimate the excess deaths driven by Red State policies. With a difference over the period of 235 deaths per million and a combined population of 151 million, the additional Red State deaths are estimated to be about 35,000 in the Summer Wave.
- While severe in all areas, the Fall wave rose faster in Red states on average and declined faster in Blue states. Estimated 23,800 addition Red State deaths.

OTHER DISPROPORTIONATE IMPACTS

COVID-19 has had many wide ranging effects. Economic dislocation, isolation, and other consequences of the pandemic the have aggravated social pathologies known to have disproportional impacts on marginalized communities, including addiction, domestic violence, suicide, human trafficking, and others. Changes in the prevalence of these concerns due to the pandemic can be estimated using the following process:

1. Using statistical analysis, identify risk factors for the particular concern and develop a predictive model. Many of these have been well-studied, with propensity models found in the literature on the subject.
2. Using the model, calculate prevalence using Pre-COVID values for the independent variables for each geography or other subset of interest
3. Identify the factors impacted by the pandemic
4. Determine Post-COVID values for the values that have changed
5. Estimate post-COVID prevalence using the updated values
6. Calculate the pre / post difference in prevalence and % change

As an example, we will consider human trafficking in the United States. A propensity model was presented by the author at SAS Global in 2017 with the following predictive variables:

- Poverty: % Below Poverty Line
- Affluence: Per Capita Income
- Gini Index: % Change
- New Homelessness: Foreclosure % Increase
- Race: % Persons of Color
- History of Legal Slavery: % in slavery in 1860
- Approach to Business Regulation: Right to Work States

Of these risk factors, poverty percent, per capita income, income Gini Index, and Foreclosures have been impacted by the pandemic, while the others were not. Finding new values for the affected predictors allows estimation of the changes caused by the pandemic (Figure 8).

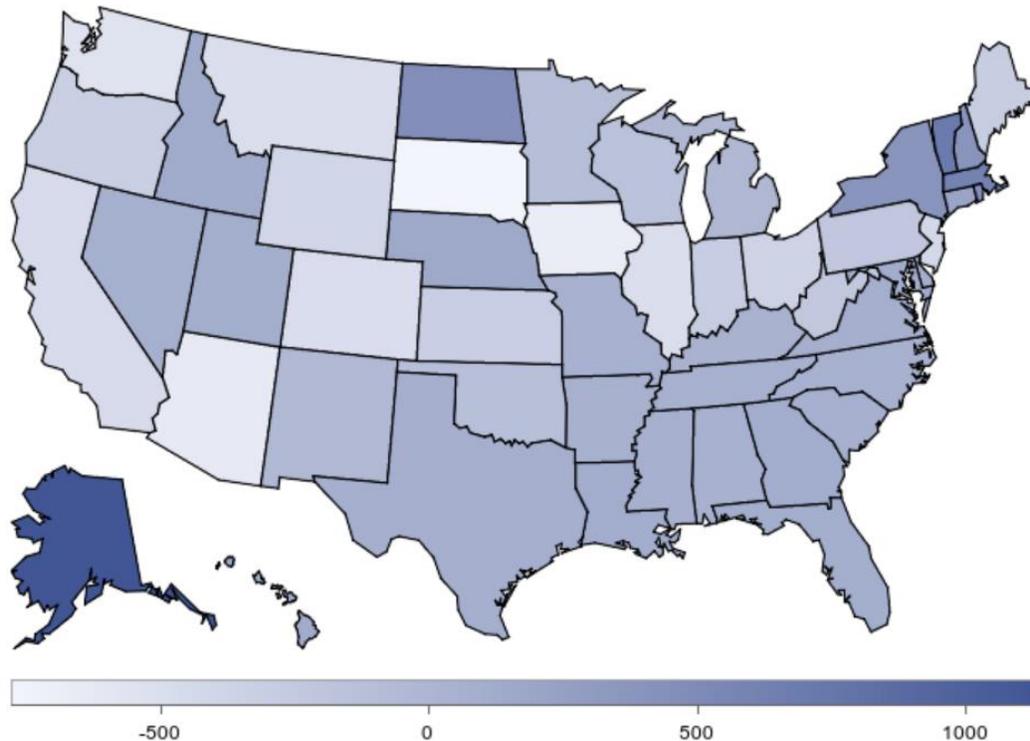


Figure 8. Change in Human Trafficking Risk Score Resulting from COVID-19

The model predicts the areas with the largest increases are Alaska, Central New England, North Dakota, New York, Nebraska, and Idaho.

CONCLUSION

In summary, disproportional impacts on marginalized populations, especially BIPOC, from the COVID-19 pandemic in the United States. Three distinct waves of the pandemic are seen:

- An initial wave, reflecting intrinsic risk, is seen from February through June. Marginalized populations were badly impacted.
- A Summer Wave from July through October, with an increase in mortality due to behavioral risks including resistance to safety precautions. This wave was only seen in in the United States and, within the US, only in Red States.
- A Fall Wave, beginning in November, was expected by epidemiologists and occurred worldwide.

The key metric for Intrinsic risk is COVID-19 mortality per capita deaths measured during the first wave, before the widespread development of behaviorally-driven risks. Mortality rates are found to increase with population percent BIPOC. The Odds Ratio for COVID mortality between the top and bottom quintiles of BIPOC population percent is 7.12. Additional deaths in Red States since July 1, 2020 due to politically-driven risk behaviors and policy decisions is estimated to be about 59,000 persons.

Connections to other social pathologies with disproportionate impact on marginalized populations as seen. These impacts can be estimated by estimating changes in prevalence model scores resulting from pandemic-driven changes in the independent variables.

REFERENCES

Banks, D. and Kyckelhahn, T. (2011), "Characteristics of Suspected Human Trafficking Incidents, 2008-2010", USDOJ

Center on Budget and Policy Priorities, Tracking the COVID-19 Recession's Effects on Food, Housing, and Employment Hardships

<https://www.cbpp.org/research/poverty-and-inequality/tracking-the-covid-19-recessions-effects-on-food-housing-and>

New York Times COVID-19 Data, accessed 4/15/2021: <https://github.com/nytimes/covid-19-data>

US Census Bureau Demographic Data <https://www.census.gov/programs-surveys/ces/data/restricted-use-data/demographic-data.html>

US Center for Disease Control and Prevention (CDC), Risk for COVID-19 Infection, Hospitalization, and Death By Race/Ethnicity, 2021 <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html#footnote03>

US Bureau of Labor Statistics Data <https://www.bls.gov/data/>

ACKNOWLEDGMENTS

Thanks are due to Rev. Dr. Donald Kreiss of the Southeast Michigan Synod of the Evangelical Lutheran Church in America, who requested the metropolitan Detroit case study leading to this paper. Initial results from this study were presented at Data Science Conference on COVID-19 from the American Statistical Association. Thanks are due to Mike Jadoo, Jun Yan, and the conference committee for an informative conference contributing to the development of this research.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

David J Corliss
Peace-Work
davidjcorliss@peace-work.org
www.peace-work.org

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.