A warning against using static US county-level community data to guide equity in COVID-19 vaccine distribution: Temporal and spatial correlations of community characteristics with COVID-19 cases and deaths vary enormously and are increasingly uninformative

Nancy Krieger, PhD,¹ Christian Testa,¹ Jarvis T. Chen, ScD,¹ Pamela D. Waterman, MPH¹, William P. Hanage, PhD²

December 1, 2020

The views expressed in this paper are those of the author(s) and do not necessarily reflect those of the Harvard Center for Population and Development Studies.
Affiliations

1. Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA
2. Department of Epidemiology and Center for Communicable Disease Dynamics, Harvard T.H. Chan School of Public Health, Boston, MA

Corresponding Author

Nancy Krieger, PhD
Professor of Social Epidemiology
American Cancer Society Clinical Research Professor
Department of Social and Behavioral Sciences
Harvard T.H. Chan School of Public Health
677 Huntington Avenue
Boston, MA 02115
Email: nkrieger@hsph.harvard.edu
Abstract

Discussions in the United States about criteria for ensuring equitable vaccine distribution for COVID-19 have focused on both individual and community characteristics. Is it reasonable, however, to assume that US community characteristics have a clear-cut relationship to risk of COVID-19 in the context of a dynamic and increasingly out-of-control pandemic? To test this hypothesis, we examined temporal and regional variation in the correlations of a range of county-level social and economic metrics, including the CDC’s social vulnerability index, with COVID-19 case and death rates per capita, spanning from February 1, 2020 through November 10, 2020: for the US as a whole, and for four regions – Northeast, Midwest, South, and West. Consistent with our hypothesis, the observed correlations for each metric varied considerably by time and by region, as did the relative strength of the correlations for the economic and racial/ethnic variables – and as of early November 2020, all correlations were under 0.3 and most hovered close to 0. Our findings offer an important warning against using static US county-level community characteristics to guide equitable allocation of COVID-19 vaccines.
Title: A warning against using static US county-level community data to guide equity in COVID-19 vaccine distribution: temporal and spatial correlations of community characteristics with COVID-19 cases and deaths vary enormously and are increasingly uninformative

Authors: Nancy Krieger, PhD, Christian Testa, Jarvis T. Chen, ScD, Pamela D. Waterman, MPH, William P. Hanage, PhD

1 Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA
2 Department of Epidemiology and Center for Communicable Disease Dynamics, Harvard T.H. Chan School of Public Health, Boston, MA

Word Count: 662 words

Figures: Figure 1 (panels 1a and 1b)

References: 12 (all hyperlinked, with list of references also provided)

IRB approval: not required (use solely public access de-identified county-level data)

Funding: none to report

Conflicts of interest: none to report

Submitted: December 1, 2020
Discussions in the United States about criteria for ensuring equitable vaccine distribution for COVID-19 have focused on both individual and community characteristics. The latter are not simply a proxy for individual-level data, since the prevalence of virus (or immunity) within a community affects its individuals’ risk of exposure. For example, the National Academy of Sciences, Engineering, and Medicine’s October 2020 report, *Framework for Equitable Allocation of COVID-19 Vaccine*, focuses on individual characteristics (pertaining to occupation, co-morbidities, age, disability, and residence in congregate or crowded residential facilities), but further notes: “Equity is a crosscutting consideration. In each population group, vaccine access should be prioritized through CDC’s Social Vulnerability Index or another more specific index,” with such indices based on community-level data.

The first three types of individual characteristics listed in the National Academy report are likewise prioritized by the November 2020 Advisory Committee on Immunization Practices (ACIP) ethical principles for allocating the initial supplies of COVID-19 vaccine, but this report also states that equity considerations require attention to the higher risks faced by communities of color. Concomitantly, some discussions have urged caution using racial/ethnic data, pointing to the risks of potentially stigmatizing communities of color and wrongly implying their “race,” rather than racism, underlies their elevated COVID-19 risks. Others have suggested using instead only socioeconomic indicators or else community-level indices that include “race” as only one of many variables, in part to deflect conservative opposition to race-conscious programs. One example mentioned is the CDC’s Social Vulnerability Index (SVI), developed to guide preparedness for physical disasters, which contains 15 items, of which only one involves race/ethnicity, and is typically used at the county level.

Is it reasonable, however, to assume that US community characteristics have a clear-cut relationship to risk of COVID-19 in the context of a dynamic and increasingly out-of-control pandemic? We suspect not. Instead, the growing dispersion of COVID-19 – starting in the US East and West Coasts, then flaring up in the US South, and then spreading to the US Mid-West, with cases in November now rising rapidly across the nation – implies that the predictive value of community characteristics may vary by time and region. At issue are both substantial heterogeneity in community characteristics across these regions and the impact of temporally rising infection rates and immunity within and across communities.

We accordingly examined temporal and regional variation in the correlations of a range of county-level social and economic metrics, including the CDC’s SVI, with COVID-19 case and death rates per capita, spanning from February 1, 2020 through November 10, 2020: for the US as a whole, and for four regions – Northeast, Midwest, South, and West. We obtained COVID-19 case, death, and county population data from USA Facts and used county-level metrics derived from the American Community Survey 2014-2018 5-year estimates and the CDC SVI.
Exhibits 1a and 1b present our results, for COVID-19 cases and deaths respectively. Consistent with our hypothesis, the observed correlations for each metric varied considerably by time and by region, as did the relative strength of the correlations for the economic and racial/ethnic variables. For all metrics, the national aggregate correlations obscured regional heterogeneity. Temporally, correlations rose and then fell, especially for cases, and sometimes changed signs (from positive to negative). As of early November 2020, all correlations were under 0.3 and most hovered close to 0.

Our findings offer an important warning against using static US county-level community characteristics to guide equitable allocation of COVID-19 vaccines. While more granular data, with community input and guidance, might potentially be informative, even correlations based on such data can vary by time and place. Our results further indicate that community-level racial/ethnic and socioeconomic data cannot be treated as stand-ins for each other. If, indeed, equity is the goal, then perhaps attention should turn to assessing the structural drivers of inequitable government and private sector policies and practices that perpetuate – versus prevent – viral transmission via unsafe workplaces, inadequate sick pay and family leave, precarious employment, unaffordable housing, and crowded prisons.

**Exhibit 1a. County-level correlations over time, for the US and sub-regions, between community characteristics and COVID-19 cases per capita, February 1, 2020 through November 10, 2020.**

**Exhibit 1b. County-level correlations over time, for the US and sub-regions, between community characteristics and COVID-19 deaths per capita, February 1, 2020 through November 10, 2020.**

**Figure legend:**

<table>
<thead>
<tr>
<th>Label</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Percent below poverty</td>
</tr>
<tr>
<td>S2</td>
<td>Percent in crowded housing</td>
</tr>
<tr>
<td>S3</td>
<td>CDC Social Vulnerability Index – Socioeconomic component</td>
</tr>
<tr>
<td>S4</td>
<td>CDC Social Vulnerability Index – Household Composition and Vulnerability component</td>
</tr>
<tr>
<td>S5</td>
<td>CDC Social Vulnerability Index – Housing Type and Transportation component</td>
</tr>
<tr>
<td>R1</td>
<td>Percent people of color&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>R2</td>
<td>CDC Social Vulnerability Index – Minority Status and Language component</td>
</tr>
<tr>
<td>SR1</td>
<td>Racialized Economic Segregation – White high-income households vs. People of Color in low-income households&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>SR2</td>
<td>CDC Social Vulnerability Index – Overall</td>
</tr>
<tr>
<td>D1</td>
<td>Population density</td>
</tr>
</tbody>
</table>

<sup>a</sup> Defined as all persons in US other than White Non-Hispanic
<sup>b</sup> This metrics is computed as an Index of Concentration at the Extremes; for explication, see: [https://www.hsph.harvard.edu/thegeocodingproject/covid-19-resources/](https://www.hsph.harvard.edu/thegeocodingproject/covid-19-resources/)
References cited (with links included via hypertext):


Figure 1a. US County Level Correlations Between Area Based Measures and COVID−19 Case Rates per Capita
Figure 1b. US County Level Correlations Between Area Based Measures and COVID−19 Death Rates per Capita