medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089698; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

# Social determinants of COVID-19 mortality at the county level

Rebecca Fielding-Miller, MSPH PhD[1]*
Maria Sundaram, MSPH PhD[2]
Kimberly Brouwer, PhD[1]

[1] University of California, San Diego
[2] Emory University
*Corresponding author: email at rfieldingmiller@health.ucsd.edu

## Abstract

**Objective:** To assess the associations between COVID-19 mortality and immigrant and farm worker population at the county level.

**Methods:** We used publicly accessible datasets to build a series of spatial autoregressive models assessing county level associations between COVID-19 mortality and (1) Percentage of Non-English speaking households, (2) percentage of individuals engaged in hired farm work, (3) percentage of uninsured individuals under the age of 65, and (3) percentage of individuals living at or below the poverty line.

**Results:** In urban counties (n=114), only population density was significantly associated with COVID19 mortality ($b = 0.21$, $p < 0.001$). In non-urban counties (n=2,629), all hypothesized social determinants were significantly associated with higher levels of mortality. Percentage of uninsured individuals was associated with lower reported COVID-19 mortality ($b = -0.36$, $p = 0.001$).

**Conclusions:** Individuals who do not speak English, individuals engaged in farm work, and individuals living in poverty may be at heightened risk for COVID-19 mortality in non-urban counties. Mortality among the uninsured may be being systematically undercounted in county and national level surveillance.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089698; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

## Introduction

A novel coronavirus responsible for COVID-19 respiratory disease is causing a global pandemic which has already resulted in over 3 million cases and 240,000 deaths since early January[1]. The United States currently has more cases than any other nation in the world, with just over 1 million cases and 61,000 deaths as of May 1, 2020[1]. Preliminary data indicates that existing health inequities in the United States are likely linked to COVID-19 morbidity and mortality[2].

## Methods

We built a series of spatial autoregressive models to assess county-level associations between COVID-19 mortality and: (1) Percentage of Non-English speaking households (defined as households in which no one 14 years or older reports speaking English at least "very well"), (2) percentage of individuals engaged in hired farm work [3] in the county as of 2018, (3) percentage of uninsured individuals under the age of 65, (3) percentage of individuals living at or below the poverty line, (4) percentage of residents age 65 or older, and (5) county density, measured as number of residents per square mile.

COVID-19 mortality data was sourced from county public health agencies, aggregated and made publicly available by the New York Times[4]. The proportion of households with limited English speaking ability was drawn from the American Community Survey's (ACS) 2014 5-year estimate, percentages of individuals living below poverty and over the age of 65 were from 2017 ACS data, and percentage of farmworkers was taken from the US Bureau of Economic Analysis. Percent uninsured was based on the US Census Small Area Health Insurance Estimates (SAHIE) program's 2018 estimates. Density was measured as number of individuals per square mile, based on US census data.

In addition to hypothesized predictors, we adjusted our models to account for the stage of the local epidemic by including a variable for the number of days since a county reported its first case of COVID-19, and the number of days between the

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089698; this version posted May 8, 2020. 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. It is made available under a CC-BY-NC-ND 4.0 International license .

100th case in a state and the declaration of a state-wide shelter in place (SIP) order. Arkansas, Iowa, Nebraska, North Dakota, and Wyoming were assigned a '0', denoting that they had not yet implemented an SIP order at the time of these analyses.

We first built a series of simple linear regression models to assess the bivariate association between number of deaths within a county and our hypothesized predictors, adjusting for days since 1st case and SIP order. We then constructed a spatial contiguity matrix, and checked the assumption that residuals were distributed spatially using a Moran's I test.

We next built three separate spatial autoregressive models to assess the association between number of deaths and our hypothesized social determinants, adjusting for potential confounders, and fitted the model with a spatial lag of the dependent variable based on our contiguity matrix. Our first model assessed relationships across all counties. We then stratified our analyses by population density to model (1) counties with population greater than 1000 individuals per square mile, and (2) a non-urban model with population density less than or equal to 1000 individuals per square mile.

**Results**

This analysis encompassed 2,743 counties across all 50 states. As of April 26, 2020, the number of deaths reported in the NY Times aggregated dataset ranged from 0 to 11,648 per county, with a median of 0 an interquartile range (IQR) of 0-3[1]. We classified 114 counties as urban and 2629 counties as non-urban. Deaths in urban counties ranged from 0 – 11648, with a median of 73 and IQR = 25-291. Deaths in rural counties ranged from 0-410 with a median of 0 and IQR = 0-2.

[1] Within this dataset, the 5 boroughs/counties of New York are treated as a single entity. We have done the same in these analyses, assigning all 5 counties the values associated with New York County.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089698; this version posted May 8, 2020. 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. It is made available under a CC-BY-NC-ND 4.0 International license .

In each simple (non-spatial) linear regression, the Moran's I test was statistically significant at $p < 0.01$, indicating a significant spatial pattern to associations between predictors and mortality.

In fully adjusted models, the percentage of residents living in poverty and residents over the age of 65, and the number of individuals per square mile were all independently and significantly associated with additional COVID19 mortality. Each percentage point increase in poverty was associated with 2.2 additional deaths across all counties ($p < 0.001$), and each increase in the percentage of residents over the age of 65 was associated with 2.3 additional COVID19 deaths ($p = 0.045$). Density was significantly associated with higher levels of mortality across all counties, with each additional person per square mile correlating to 1.2 additional deaths ($p < 0.001$).

In urban counties (n=114), only population density was significantly associated with COVID19 mortality ($b = 0.21$, $p < 0.001$). However, in non-urban counties, all hypothesized social determinants were statistically significantly associated with an increase in mortality. The percentage of individuals employed in farm work, percentage of non-English speaking households, and percentage of individuals living in poverty were all independently and significantly associated with higher mortality, as was increased density and more residents over the age of 65 (see table 1). However, contrary to our hypotheses, percentage of uninsured individuals was associated with lower reported COVID19 mortality ($b = -0.36$, $p = 0.001$).

## Discussion

Although we cannot draw conclusions about individual risk profiles, our findings do suggest that that farm work may create unique risk factors, and that farmworkers may require additional protections, such as personal protective equipment and/or targeted outreach. Immigrants provide approximately 75% of

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089698; this version posted May 8, 2020. 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. It is made available under a CC-BY-NC-ND 4.0 International license .

all farm labor in the United States[3]. Among those engaged in crop work specifically, nearly three quarters are migrants and approximately half are undocumented[3]. Undocumented status may impede an individual's willingness or ability to seek healthcare, or their ability to request additional protections from an employer if they worry doing so could result in their own deportation or that of a family member[5]. Farm labor is considered essential work, but there are reports of inadequate personal protective equipment and inadequate social distancing guidelines or enforcement[6].

The negative association we found between insured status and mortality is a point of concern. The CDC has noted higher than expected numbers of death across the United States throughout April in recent months, suggesting that COVID-19 mortality is potentially higher than what has thus far been captured by state and county level surveillance[7]. It is possible that this association represents a gap in testing and linkage to care among the uninsured, and/or a gap in ascertaining deaths due to COVID-19 among uninsured individuals.

**Public Health Implications**

COVID-19 mortality appears to be statistically significantly associated with social determinants of health at the county level, and these relationships may be more pronounced in non-urban counties. Individuals who do not speak English, individuals engaged in farm work, and individuals living in poverty may be at heightened risk for COVID-19 mortality in non-urban counties. These individuals should receive targeted and tailored outreach.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089698; this version posted May 8, 2020. 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. It is made available under a CC-BY-NC-ND 4.0 International license .

*Table 1: Spatial regression models, predictors of number of deaths across urban, non-urban, and all US counties reporting at least 1 COVID-19 case as of April 26, 2020*

| | All counties (n=2743) | | Non-urban counties (n=2629) | | Urban counties (n=114) | |
|---|---|---|---|---|---|---|
| | *b* | *p-value* | *b* | *p-value* | *b* | *p-value* |
| *% Farm workers* | 0.27 | 0.06 | 0.52 | 0.002 | 1448.56 | 0.12 |
| *% Non-English speakers* | -0.19 | 0.66 | 0.22 | <0.001 | 10.49 | 0.21 |
| *% Residents uninsured* | -1.15 | 0.25 | -0.36 | 0.001 | -39.77 | 0.13 |
| *% Residents in poverty* | 2.23 | 0.002 | 0.20 | 0.02 | 37.6 | 0.08 |
| *Residents per square mile* | 0.21 | <0.001 | 0.08 | <0.001 | 0.21 | <0.001 |
| *% Residents Over 65* | 2.27 | 0.05 | 0.35 | 0.01 | 18.96 | 0.23 |

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089698; this version posted May 8, 2020. 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. It is made available under a CC-BY-NC-ND 4.0 International license .

# REFERENCES


1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases.* 2020;20(5):533-534.
2. Dorn AV, Cooney RE, Sabin ML. COVID-19 exacerbating inequalities in the US. *The Lancet.* 2020;395(10232):1243-1244.
3. Farm Labor. *Economic Research Service* 2020; https://www.ers.usda.gov/topics/farm-economy/farm-labor/ - size. Accessed May 1, 2020, 2020.
4. Data from The New York Times, based on reports from state and local health agencies. In. *The New York Times*. New York, NY2020.
5. Philbin MM, Flake M, Hatzenbuehler ML, Hirsch JS. State-level immigration and immigrant-focused policies as drivers of Latino health disparities in the United States. *Social Science & Medicine.* 2018;199:29-38.
6. Borunda A. Farmworkers risk coronavirus infection to keep the U.S. fed. In. *National Geographic*2020.
7. Excess Deaths Associated with COVID-19. 2020; https://www.cdc.gov/nchs/nvss/vsrr/covid19/excess_deaths.htm. Accessed May 1, 2020, 2020.