Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Original article

Coronavirus disease 2019 (COVID-19) mortality and neighborhood characteristics in Chicago

Molly Scannell Bryan, PhD a, *, Jiehuan Sun, PhD b, Jyotsna Jagai, PhD b, Daniel E. Horton, PhD c, Anastasia Montgomery, BA c, Robert Sargis, MD, PhD d, Maria Argos, PhD b

a Department of Medicine, Institute for Minority Health Research, University of Illinois at Chicago, Chicago
b School of Public Health, University of Illinois at Chicago, Chicago
c Department of Earth and Planetary Sciences and Institute for Sustainability and Energy, Northwestern University, Chicago, IL
d Department of Medicine, University of Illinois at Chicago, Chicago

Article info

Article history:
Received 24 July 2020
Accepted 28 October 2020
Available online 10 November 2020

Keywords:
COVID-19
Prevention
Social determinants of health
Built environment
Health disparities

A B S T R A C T

Purpose: To describe coronavirus disease 2019 (COVID-19) mortality in Chicago during the spring of 2020 and identify at the census-tract level neighborhood characteristics that were associated with higher COVID-19 mortality rates.

Methods: Using Poisson regression and regularized linear regression (elastic net), we evaluated the association between neighborhood characteristics and COVID-19 mortality rates in Chicago through July 22 (2514 deaths across 795 populated census tracts).

Results: Black residents (31% of the population) accounted for 42% of COVID-19 deaths. Deaths among Hispanic/Latino residents occurred at a younger age (63 years, compared with 71 for white residents). Regarding residential setting, 52% of deaths among white residents occurred inside nursing homes, compared with 35% of deaths among black residents and 17% among Hispanic/Latino residents. Higher COVID-19 mortality was seen in neighborhoods with heightened barriers to social distancing and low health insurance coverage. Neighborhoods with a higher percentage of white and Asian residents had lower COVID-19 mortality. The associations differed by race, suggesting that neighborhood context may be most tightly linked to COVID-19 mortality among white residents.

Conclusions: We describe communities that may benefit from supportive services and identify traits of communities that may benefit from targeted campaigns for prevention and testing to prevent future deaths from COVID-19.

© 2020 Elsevier Inc. All rights reserved.

Introduction

Coronavirus disease 2019 (COVID-19), the illness caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in the fall of 2019 and was declared a pandemic by the World Health Organization in March 2020 [1]. No natural immunity exists, and while severely ill patients can benefit from supportive care [2,3], no medical cure is available. In the United States, more than 200,000 Americans have died of the disease before the middle of September 2020 [4]. The risk of COVID-19—related morbidity and mortality increases with age [5,6], and is higher in individuals with pre-existing conditions, including respiratory disease, cardiovascular disease, recent cancer history, kidney disease, liver disease, vascular disease, immune disorders, and diabetes [7].

As the pandemic has progressed in the United States, it has become clear that the impact of COVID-19 has been felt more acutely in some communities [8,9], most clearly among black Americans, who acquire the disease and die at disproportionate rates [10–12]. Social determinants of health may drive some of this disparity, and neighborhood traits may be particularly relevant, given that infectious disease spread is often influenced by the built environment [13]. Determining neighborhood factors that are associated with COVID-19 could identify modifiable risk factors that may directly influence COVID-19 risk, and recent research has postulated associations with air quality [14,15], population density [16], and public transit use [17]. Further, identifying neighborhood traits that are associated with COVID-19 risk, even if they are not...
causally related, can identify communities with high need for support and testing and prevention services to mitigate the continued impact of COVID-19.

In the present analysis, we describe COVID-19 mortality in the spring and early summer of 2020 in Chicago, the third most populous city in the United States [18]. Further, with a focus on deaths that occurred outside of nursing homes, we identify at the census-tract level neighborhood characteristics that are associated with COVID-19 mortality in Chicago.

**Methods**

**Deaths related to COVID-19 in Chicago**

The Cook County Medical Examiner provides a public record of deaths in Cook County [19]. Beginning on March 16, 2020, there was a rapid increase in daily deaths from COVID-19; this count peaked in early May, at which point deaths declined through July (eFig. 1). Between March 16 and July 22, 2020, 4834 deaths were recorded with “novel corona virus” listed as a primary or secondary cause of death. Along with the age and race of the decedent, each record includes an incident address (populated in 97% of COVID-19 deaths); when the cause of death is an infectious disease, this almost always reflects home address. Addresses were standardized and geocoded [20], and Chicago residents were included in the subsequent analysis.

Deaths that occurred in nursing homes or assisted living facilities were identified using three methods: address match with Medicare’s Nursing Home Compare database [21] (836 deaths), address match with Medicaid-managed care facilities contracted with Cook County Health and Hospital System (33 additional deaths identified), and a manual review of address in which more than four deaths occurred (21 additional deaths identified).

**Neighborhood characteristics**

Following the framework laid out by others [22], we curated a list of characteristics of neighborhoods that may be associated with (1) increased chance of COVID-19 infection or (2) increased risk of mortality if infected. In total, we considered 33 neighborhood characteristics, detailed in Table 1. Unless otherwise noted, characteristics were estimated at the census-tract level from the American Community Survey 2018 five-year estimates [23]. Characteristics putatively associated with infection risk included the following: crowded living conditions, transportation habits, dense housing, and sociodemographic characteristics related to heightened barriers to social distancing. Characteristics putatively associated with risk of mortality included the following: health care access, comorbid conditions and demographic traits associated with more severe COVID-19 infection, indicators of poverty, and chronic exposure to poor air quality.

To assess air quality at the census tract scale, high-resolution (1.3 km²) simulations from the two-way coupled community multiscale air quality-weather research and forecasting ([26]) (CMAQ-WRF) model were performed. Our air quality exposure analysis included nitrogen dioxide (NO₂) and particulate matter less than 2.5 microns (PM₂.5) and ozone (O₃) pollutant data. While air quality varies seasonally, and air quality was impacted by the widespread stay-at-home directives during the spring of 2020 ([27–29]), our aim was to model the influence of long-term exposure to air quality, rather than acute exposure as the pandemic was ongoing. As such, for PM₂.5 and NO₂, we averaged the exposure from two seasonally representative pre-COVID-19 months (August 2018 and January 2019). For ozone, given its high dependency on temperature, we averaged 8-hour daytime max values during a relevant representative pre-COVID summertime month (August 2018).

We also included indicators of the racial and ethnic makeup of the census tract as a proxy for the influence of structural racism that may not be captured by the other characteristics [22].

**Statistical approach**

In institutionalized settings, the community in which the facility is located may have a more distant relationship with COVID-19 mortality, as the traits of the facility itself most directly influence the risk of COVID-19 [30]. As such, we next focused on deaths that occurred outside of nursing homes (also excluding the five deaths that occurred at the Cook County jail).

In Chicago, many neighborhood characteristics are highly correlated with one another (eFig. 2). In recognition of this, two analytical approaches were taken. The aim of the first was to

---

**Table 1**

Neighborhood characteristics evaluated with COVID-19 infection and mortality

| Infection risk | Crowded living conditions | Residences without complete kitchens, residences with more than one occupant per room, grandparents living with children under 18 years. Residences without a car available, commuting primarily by public transit, commuting primarily by carpool. |
|----------------|----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                | Transportation habits | Housing units in buildings with more than 20 units, population density. SNAP use, broadband Internet at home, educational attainment, ability to work from home. |
|                | Dense housing | Health insurance status (American Community Survey), access to a primary care provider (Chicago Health Atlas estimate at the community area level 2016–2018 ([24])). |
|                | Sociodemographic characteristics that might be associated with heightened barriers to social distancing | Rate of heart disease deaths per 100,000, rate of diabetes-related deaths per 100,000, rate of nephrotic disease deaths per 100,000, and rate of tobacco-related deaths per 100,000 (all from Chicago Health Atlas ([24]) estimate at the community area level 2013–2017). |
| Mortality Risk | Health care access | Male sex, population aged 65–74 years, population aged 75+ years. Poverty rate, unemployment rate, households spending more than 35% of their income on rent, historical redlining of the neighborhood (University of Richmond Mapping Inequality ([25])). |
|                | Presence of comorbid conditions suspected to be associated with more severe disease | Concentration of nitrogen dioxide (NO₂), ozone (O₃), and particulate matter smaller than 2.5 microns (PM₂.5). |
|                | Age and biological sex Indicators of poverty | Percent population that is non-Hispanic black, percent of population that is non-Hispanic white, percent of population that is Hispanic/Latino, percent of population that is non-Hispanic Asian. |
|                | Air quality | | |
describe Chicago census tracts with high COVID-19 mortality rates. The count of COVID-19 deaths was modeled as a Poisson distribution, and regressed upon each of the 33 neighborhood characteristics sequentially, with the log of the population of the census tract as an offset. The \( P \)-values from these bivariate regressions were corrected using the Benjamini-Hochberg procedure. In each regression, no other tract-level descriptors were included as covariates. As such, while this analysis well-described communities that were heavily impacted by COVID-19 deaths, any given association may be explained by the presence of other correlated characteristics.

The aim of the second approach was to account for the correlation between the neighborhood traits and identify those that continued to be robustly associated with COVID-19 mortality. In the second approach, all 33 neighborhood characteristics were included as predictors in a regularized generalized linear regression model with elastic net penalty. We used the R package glmnet \([31],[32]\) with the elastic net mixing parameter \( \lambda \) set at 0.5 and 20-fold cross-validation to select the tuning parameter. COVID-19 deaths were again modeled as a Poisson distribution with the log of the population of the census tract as an offset. We used 200 bootstrap replications to obtain 95% confidence intervals and to obtain the frequency selection of each trait. Empirically, all penalization penalties \( \lambda > 0 \) produced substantively similar results (i.e., a ridge penalty was found to be inappropriate, and a least absolute shrinkage and selection operator (LASSO) penalty produced substantively similar results), and a tuning parameter of 0.5 was chosen to make use of the flexibility of the elastic net.

**Results**

Between March 16 and July 22 2020, 2514 COVID-19 deaths were recorded within Chicago. Although non-Hispanic black residents comprise 31% of Chicago’s population, they accounted for 42% of the COVID-19 deaths. All other racial and ethnic groups were under-represented: non-Hispanic white residents comprise 32% of Chicago’s population, Hispanic/Latino residents 29%, and Asian residents 6%.

We identified 895 deaths (36%) among people living in an institutionalized setting. Selected demographic characteristics of individuals who died are shown in Table 2. The percentage of deaths that occurred in nursing homes differed by race and ethnicity: 52% of those who died and were white lived in a nursing home, compared with 18% of those who died and were Hispanic/Latino. Outside of nursing homes, the proportion of deaths by race and ethnicity changed throughout the course of the spring of 2020 (eFig. 3). Before mid-May, more than 50% of the deaths occurred among black individuals. While initially few deaths occurred among Hispanic/Latino residents (12% of deaths before April 1), the proportion of deaths among this demographic steadily rose (32% of deaths by July 22).

The mean age at death remained relatively stable over the same time (eFig. 4). For those who died outside of nursing homes, the age at death differed by a decade by race and ethnicity. The mean age at death among Hispanic/Latino residents was 63 years, 69 years for black residents, 71 years for white residents, and 74 years for Asian residents.

Figure 1 maps the rate of COVID-19 deaths that occurred among individuals in noninstitutionalized residential settings in each of Chicago’s 795 populated census tracts per thousand residents \([33]\). Panel A displays this death rate overall, and panels B–D are restricted to deaths that occurred among residents who were black (B), white (C), or Hispanic/Latino (D) (as recorded by the medical examiner), each as a fraction of the number of residents in that tract of the given race/ethnicity. Chicago is a city with high levels of residential segregation (eFig. 5) \([34]\). In the overall Chicago population, areas with higher death rates were located on the west and south sides of Chicago, with the highest death rates occurring in majority-minority census-tracts. This pattern was particularly pronounced when focusing on deaths that occurred to white residents; in neighborhoods that were majority-minority, the proportion of white residents who died from COVID-19 (as a proportion of the white population in that neighborhood) was generally higher than it was in neighborhoods who were majority white. In contrast, high rates of COVID-19 mortality among black and Hispanic/Latino residents were found throughout the city.

**Bivariate Poisson regression associations**

Table 3 (column A) summarizes the bivariate associations of neighborhood characteristics with COVID-19 death rate. Among neighborhood characteristics putatively associated with increased risk of infection Table 3A, top), higher death rates were seen in neighborhoods with heightened barriers to social distancing (high proportion of supplemental nutrition assistance program (SNAP) recipients, fewer households with broadband Internet, lower educational attainment, and fewer workers able to work from home). Crowded living conditions were similarly associated with higher COVID-19 mortality, whereas neighborhoods with dense housing had lower rates. Higher COVID-19 death rates were also seen in neighborhoods where fewer residents had cars or where more residents carpooled; no association was seen with public transportation.

Among neighborhood characteristics associated with populations at higher risk of severe disease (Table 3A, middle), most were associated with higher rates of COVID-19 mortality. Mortality was higher in neighborhoods with worse access to health care, more comorbid conditions, older age, higher rates of poverty, and

---

**Table 2**

Demographic characteristics of COVID-19 deaths in Chicago through July 22 by residential setting

| Race and ethnicity | Overall n – 2514 | Institutionalized n – 885 | Noninstitutionalized n – 1619 |
|-------------------|----------------|--------------------------|-----------------------------|
| Mean age (SD)     |                |                          |                             |
| Black, non-Hispanic/Latino | 71 (15)        | 78 (12)                  | 67 (15)                     |
| White, non-Hispanic/Latino  | 663 (26%)      | 343 (38%)                | 531 (33%)                   |
| Hispanic/Latino  | 644 (26%)      | 13 (13%)                 | 47 (3%)                     |
| Asian, non-Hispanic/Latino | 85 (3%)        | 38 (4%)                  | 17 (2%)                     |
| Other, non-Hispanic/Latino | 35 (1%)        | 10 (1%)                  | 24 (1%)                     |
| Gender            |                |                          |                             |
| Female            | 1002 (40%)     | 393 (44%)                | 609 (38%)                   |
| Male              | 1507 (60%)     | 501 (56%)                | 1006 (62%)                  |
| Unknown           | 5 (0%)         | 1 (0%)                   | 4 (0%)                      |
neighborhoods that had historically been redlined. Summertime concentrations of maximum $O_3$ were associated with higher death rates, and average concentrations of $NO_2$ were inversely associated with death rates.

Our analysis confirms that neighborhoods with a greater percentage of black and Hispanic/Latino residents saw higher COVID-19 death rates during the spring of 2020 (Table 3A, bottom). For each additional percentage point of the population that was black, there was a 32% increase in the COVID-19 death rate, and for each additional percentage point of the population that was Hispanic/Latino, there was an 19% increase in the COVID-19 death rate. Conversely, neighborhoods with a higher percentage of Asian or white residents saw lower death rates.

**Regularized regression associations**

The results from the regularized regression model are shown on the right of Table 3 (Table 3B). After implementing the regularized regression to further account for the correlation between the neighborhood traits, COVID-19 death rate remained robustly associated with four neighborhood traits (here, “robustly” includes traits selected by more than more than 90% of the bootstrap replications). The death rate was lower in neighborhoods with higher penetration of broadband internet service and higher percentages of white or Asian residents, and the death rate was higher in neighborhoods with more residents without health insurance. Additional characteristics were suggestively associated with COVID-19 death rate: lower education, higher rates of severe kidney disease, and a higher percentage of the population above age 65 years (here, “suggestively” includes traits selected by between 50% and 90% of the bootstrap replications).

Table 4 displays the results of the regularized regression for the race and ethnicity-specific death rates. As shown in Table 4 (left), the COVID-19 death rate among black residents was only weakly associated with neighborhood characteristics, as none were selected by more than 50% of the bootstrap replications, and the deviance ratio was essentially zero.

Among Hispanic/Latino residents (Table 4, right), no robust associations were found, and the deviance ratio was low (0.0198). Suggestive associations were found with public transit, high school
Table 3
Bivariate and regularized linear regression associations between neighborhood characteristics and COVID-19 death rates in the noninstitutionalized population of Chicago (n = 1619 deaths)

| Domain                      | Neighborhood characteristic | (A) Bivariate association | (B) Regularized linear regression |
|-----------------------------|----------------------------|---------------------------|-----------------------------------|
| Infection                   | % population 30–living w/ grandparents ≤ 18 | 1.27 (1.22–1.33); P ≤ .0001* | 1 (122; 0.989 to 1.000) |
|                             | % occupied units w/out complete kitchen       | 1.10 (1.05–1.14); P ≤ .0001* | 1 (133; 1.000 to 1.022) |
|                             | % occupied units >1 resident per room         | 1.21 (1.16–1.27); P ≤ .0001* | 1 (322; 1.000 to 1.107) |
|                             | % occupied units w/no car available           | 1.09 (1.04–1.15); P ≤ .0004* | 1 (1; 1.000 to 1.000) |
|                             | % workers >16 commuting public transit        | 0.97 (0.93–1.02); P = 3157 | 1 (296; 1.000 to 1.077) |
|                             | % workers >16 commuting carpool               | 1.23 (1.18–1.28); P ≤ .0001* | 1 (181; 1.000 to 1.005) |
| Dense housing               | % units in buildings w/20+ units              | 0.76 (0.71–0.80); P ≤ .0001* | 1 (363; 0.997 to 1.000) |
| Barriers to social distancing| % households w/broadband internet             | 0.63 (0.60–0.67); P ≤ .0001* | 0.992 (99%; 0.986 to 0.999) |
|                             | % households w/SNAP in last year             | 1.50 (1.43–1.57); P ≤ .0001* | 1 (444; 1.000 to 1.006) |
|                             | % adults ≥25 w/h high school diploma          | 1.39 (1.34–1.45); P ≤ .0001* | 1 (822; 1.000 to 1.111) |
|                             | % adults ≥25 w/bachelor's degree or higher    | 0.57 (0.54–0.61); P ≤ .0001* | 0.999 (68%; 0.995 to 1.000) |
|                             | % workers >16 who work from home              | 0.79 (0.74–0.83); P ≤ .0001* | 1 (8; 1.000 to 1.006) |
| Mortality                   | % population without health insurance         | 1.35 (1.30–1.41); P ≤ .0001* | 1.013 (98%; 1.000 to 1.025) |
| Comorbid conditions         | % population w/primary care provider          | 0.81 (0.77–0.85); P ≤ .0001* | 1 (144; 0.996 to 1.000) |
|                             | Rate of heart disease deaths                  | 1.19 (1.13–1.25); P ≤ .0001* | 1 (600; 1.000 to 1.000) |
|                             | Rate of diabetes-related deaths               | 1.31 (1.24–1.37); P ≤ .0001* | 1 (300; 1.000 to 1.002) |
|                             | Rate of nephritic disease deaths              | 1.26 (1.20–1.31); P ≤ .0001* | 1 (624; 1.000 to 1.009) |
| Age and gender              | % population male                             | 0.95 (0.90–1.00); P = 0.035* | 1 (222; 1.000 to 1.104) |
|                             | % population between 65 and 74               | 1.11 (1.05–1.17); P ≤ .0001* | 1.005 (622; 1.000 to 1.033) |
|                             | % population 75+                              | 1.11 (1.06–1.16); P ≤ .0001* | 1.005 (683; 1.000 to 1.024) |
| Indicators of poverty       | % population below poverty                    | 1.38 (1.32–1.45); P ≤ .0001* | 1 (14; 1.000 to 1.004) |
|                             | % population ≥16 unemployed                   | 1.36 (1.30–1.42); P ≤ .0001* | 1 (188; 0.996 to 1.004) |
|                             | % occupied units w/rent > 35% income          | 1.35 (1.28–1.42); P ≤ .0001* | 1 (133; 1.000 to 1.002) |
| Air quality                 | Rated C/D by the HOLC                        | 1.13 (1.08–1.19); P ≤ .0001* | 1 (420; 1.000 to 1.194) |
| Structural racism           | Average NO2 in tract (ppbv)                  | 0.92 (0.88–0.97); P = 0.0012* | 1 (19; 0.983 to 1.000) |
| Structural                   | Average summitte O2 in tract (ppbV)           | 1.13 (1.08–1.18); P ≤ .0001* | 1 (8; 1.000 to 1.007) |
| Structural racism           | Average PM2.5 in tract (ug/m3)               | 1.01 (0.97–1.06); P = 0.017 | 1 (28; 0.991 to 1.000) |
| Model metrics               | % population Black                           | 1.32 (1.25–1.38); P ≤ .0001* | 1 (4; 1.000 to 1.000) |
|                             | % population Hispanic/Latino                  | 1.19 (1.14–1.24); P ≤ .0001* | 1 (181; 1.000 to 1.003) |
|                             | % population Asian                           | 0.67 (0.62–0.73); P ≤ .0001* | 0.995 (95%; 0.987 to 1.000) |
|                             | % population white                           | 0.57 (0.54–0.61; P ≤ .0001* | 0.993 (1000; 0.990 to 0.997) |

The reported bivariate associations (Column A) are the rate ratios from a Poisson regression with counts of the noninstitutionalized deaths predicted by the given neighborhood characteristic, with the offset equal to the population of the census tract. 95% confidence intervals are reported in parentheses. Benjamini-Hochberg p-values are reported, and those BH adjusted p-values that are below 0.05 are marked with an asterisk (*). The reported regularized linear regression associations (Column B) are the rate ratios from a regularized (elastic net) Poisson regression with counts of the noninstitutionalized deaths of Chicago residents assigned the given race predicted by the given neighborhood characteristic, with the offset equal to the log of the population of the census-tract. In parentheses is the percentage of bootstrap replications in which the variable was selected, along with bootstrap 95% confidence intervals for the estimate. Variables that were selected by more than 50% of the bootstrap replications are highlighted with a caret (^), and variables that were selected by more than 90% of the bootstrap replications are highlighted with a dagger (|). The reported R2-like statistic is: 1 - \( \sum (\hat{y}_i - y_i)^2 / \sum (y_i - \bar{y})^2 \), where \( \hat{y}_i \) = population_i, \( \sum \) population_i, 0.2518 deviation ratio: 0.2638

diplomas, health insurance coverage, primary care providers, gender, and exposure to NO2.

In contrast, among white residents (Table 4, middle), there were robust associations between COVID-19 death rate and four neighborhood characteristics. The death rate among white residents was higher in neighborhoods with more residents without a high school diploma, and a greater percentage of residents who were Hispanic or Latino, and the death rate was lower in neighborhoods with higher percentages of white or Asian residents. There were further suggestive associations between higher death rates and intergenerational households, crowded housing, carpool use, population density, broadband Internet access, SNAP use, bachelor degree attainment, health insurance, age, rent-to-income ratios, and summertime ozone concentration.

To assess whether these race-specific findings were due to more uniformity in neighborhood traits among nonwhite residents, for the traits that were robustly associated with white death rate but not with black and Hispanic/Latino death rate, we compared the variability in neighborhood traits by the race of the decedent (eFig. 6). While neighborhoods of black residents who died had less variability terms of school completion, there was no such difference in variability in the neighborhoods of Hispanic/Latino residents who died, and further, there was broad variability in the neighborhood-level race/ethnicity makeup. As such, homogeneity in neighborhood traits could not fully explain the lack of associations among nonwhite deaths.

During the observed study period, the city of Chicago implemented a phased reopening policy, in which residents and businesses were able to participate in increasingly less restrictive activities. On June 3, and June 26, Chicago entered “phase 3” and “phase 4,” respectively ([35],[36]). To assess whether the results were sensitive to this timing, we repeated the overall regularized regression analyses twice: first limiting to deaths that occurred before June 3 (2118 deaths; 1324 among the noninstitutionalized population), and those that occurred before June 26 (2417 and 1544 deaths, respectively). The results were substantively similar (eTable 1).

Discussion

During the spring of 2020, COVID-19 impacted the residents of the city of Chicago unevenly. Black residents were at the highest
The associations between neighborhood characteristics and COVID-19 mortality largely mirrors that for white residents, COVID-19 mortality may be more tightly linked to neighborhood context than for black residents.

Our analysis suggests that controlling for the correlated nature of the neighborhood characteristics is necessary, as bivariate associations suggested inverse associations between COVID-19 mortality rates and air quality, density, and public transit use. These associations that generally did not persist in the regularized regression models. These paradoxical bivariate associations may be due to Chicago-specific desire of affluent residents to live close to the lake (with dense housing) or highways (which are near public transit hubs and often have poor air quality). Although additional study designs are necessary to confirm that these associations are not due to residual confounding, the regularized regression approach statistically controls for more aspects of this confounding than a standard Poisson regression.

This work is part of a growing body of literature that explores social determinants of COVID-19 severity. A study from New York City also found that for white residents, COVID-19 mortality may be more tightly linked to neighborhood context than for black residents.

### Table 4

Race and ethnicity-specific regularized regression (elastic net) associations between neighborhood characteristics and COVID-19 death rates in the noninstitutionalized population of Chicago

| Domain                        | Neighborhood characteristic                        | Non-Hispanic black | Non-Hispanic white | Hispanic/Latino |
|-------------------------------|--------------------------------------------------|--------------------|-------------------|----------------|
|                                | n = 679                                          | n = 320            | n = 531           |
| Infection crowding conditions | % population 30+ years living w/grandchildren <18|                    |                   |                |
|                               | % occupied units w/out complete kitchen           | 1 (8%; 0.988 to 1.000) | 1.006 (63%; 1.000 to 1.051) | 1 (14%; 0.990 to 1.001) |
|                               | % occupied units w/1 resident per room            | 1 (2%; 1.000 to 1.000) | 1.000 (100%; 1.000 to 1.000) | 1 (12%; 0.990 to 1.013) |
|                               | % worker >16 commuting public transit             | 1 (3%; 1.000 to 1.000) | 1 (10%; 1.000 to 1.004) | 1 (5%; 1.000 to 1.001) |
|                               | % workers >16 commuting carpool                   | 1 (0%; 1.000 to 1.000) | 1 (30%; 1.000 to 1.101) | 1.001 (60%; 1.010 to 1.105) |
|                               | % household w/SNAP in the last year               | 1 (2%; 1.000 to 1.000) | 1.003 (73%; 1.000 to 1.018) | 1 (22%; 1.000 to 1.065) |
|                               | % adults >25 w/high school diploma                | 1 (1%; 1.000 to 1.000) | 1.021 (100%; 1.005 to 1.035) | 1.000 (53%; 1.000 to 1.101) |
|                               | % adults >25 w/bachelor’s degree or higher        | 1 (0%; 1.000 to 1.000) | 0.955 (06; 0.989 to 1.000) | 1 (0%; 1.000 to 1.000) |
|                               | % workers >16 who work from home                  | 1 (1%; 1.000 to 1.000) | 1 (16%; 1.000 to 1.025) | 1 (5%; 0.994 to 1.000) |
| Mortality health care access   | % population without health insurance             | 1 (2%; 1.000 to 1.000) | 1.011 (79%; 1.000 to 1.029) | 1.003 (64%; 1.000 to 1.101) |
|                               | % population w/primary care provider              | 1 (0%; 1.000 to 1.000) | 1 (12%; 0.998 to 1.006) | 0.998 (57%; 0.984 to 1.000) |
| Comorbid conditions           | Rate of heart disease deaths                     | 1 (0%; 1.000 to 1.000) | 1 (2%; 1.000 to 1.000) | 1 (42%; 0.999 to 1.000) |
|                               | Rate of diabetes-related deaths                   | 1 (0%; 1.000 to 1.000) | 1 (1%; 1.000 to 1.001) | 1 (1%; 1.000 to 1.000) |
|                               | Rate of nephrotic disease deaths                  | 1 (0%; 1.000 to 1.000) | 1 (1%; 1.000 to 1.001) | 1 (1%; 1.000 to 1.001) |
|                               | Rate of tobacco-related deaths                    | 1 (0%; 1.000 to 1.000) | 1 (1%; 0.999 to 1.000) | 1 (30%; 0.999 to 1.000) |
| Age and gender                | % population male                                 | 1 (8%; 0.996 to 1.000) | 1 (15%; 0.990 to 1.103) | 1.002 (55%; 1.000 to 1.040) |
|                               | % population between 65 and 74                    | 1 (38%; 1.000 to 1.031) | 1 (20%; 1.000 to 1.037) | 1 (2%; 1.000 to 1.000) |
|                               | % population 75+                                 | 1 (9%; 1.000 to 1.000) | 1.002 (54%; 1.000 to 1.048) | 1 (4%; 1.000 to 1.099) |
| Indicators of poverty         | % population below poverty                        | 1 (1%; 1.000 to 1.000) | 1 (10%; 1.000 to 1.004) | 1 (3%; 1.000 to 1.099) |
|                               | % population >16 unemployed                      | 1 (14%; 1.000 to 1.013) | 1 (14%; 1.000 to 1.013) | 1 (14%; 1.000 to 1.013) |
|                               | % occupied units w/rent > 35% income              | 1 (0%; 1.000 to 1.000) | 1.003 (70%; 1.000 to 1.012) | 1 (11%; 1.000 to 1.005) |
|                               | Rated C/D by the HOLC                             | 1 (4%; 1.000 to 1.016) | 1 (23%; 0.957 to 1.197) | 1 (47%; 1.000 to 1.326) |
| Air quality                   | Average NO2 in tract (ppbV)                      | 1 (16%; 0.964 to 1.000) | 1 (41; 0.926 to 1.000) | 1.025 (80%; 1.000 to 1.080) |
|                               | Average O3 in tract (ppbV)                       | 1 (2%; 1.000 to 1.000) | 1.016 (72%; 1.000 to 1.079) | 1 (48; 0.996 to 1.000) |
| Structural racism             | Average PM2.5 in tract (ug/m3)                   | 1 (6%; 0.980 to 1.000) | 1 (16%; 1.000 to 1.141) | 1 (85; 0.890 to 1.000) |
|                               | % population black                               | 1 (0%; 1.000 to 1.000) | 1 (0%; 1.000 to 1.000) | 1 (12%; 1.000 to 1.003) |
|                               | % population Hispanic/Latino                     | 1 (0%; 1.000 to 1.000) | 1.005 (94%; 1.000 to 1.101) | 1 (25%; 1.000 to 1.100) |
|                               | % population Asian                               | 1 (0%; 1.000 to 1.000) | 0.980 (95%; 0.975 to 1.000) | 1 (18%; 0.990 to 1.000) |
|                               | % population white                               | 1 (0%; 1.000 to 1.000) | 0.992 (98%; 0.986 to 0.999) | 1 (65; 0.999 to 1.000) |
| Model metrics                 | R²-like statistic                                 | 0.0000             | 0.2865             | 0.0547             |
|                               | Deviance ratio                                    | 0.0000             | 0.3336             | 0.0198             |

The reported regularized linear regression associations (Column B) are the rate ratios from a regularized (elastic net) Poisson regression. The dependent variable was the number of noninstitutionalized deaths of Chicago residents for a given race/ethnicity (as recorded by the medical examiner) within a census tract, with the offset equal to the log of the number of residents of that race/ethnicity living in the census tract. In parentheses is the percentage of bootstrap replications in which the variable was selected, along with bootstrap 95% confidence intervals for the estimate. Variables that were selected by more than 50% of the bootstrap replications are highlighted with a caret (^), and variables that were selected by more than 90% of the bootstrap replications are highlighted with a dagger (†). The reported R²-like statistic is: \( R^2 = \frac{\sum (y_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} \), where \( \bar{y} = \frac{\sum y_i}{n} \).
City characterized neighborhood traits associated with COVID-19 infection [38]. Consistent with our results that study found that housing value, housing density, and income were protective against infection, whereas crowded households were associated with increased risk. Other work has compiled aggregate indices of social vulnerability [9,39,40] and found that cumulative levels of social vulnerability were associated with increased mortality. Our census tract-level analysis provides additional specificity of which neighborhood characteristics may be driving that vulnerability. Our focus on deaths that occurred outside of institutionalized settings and our use of a regularized regression approach also enabled us to offer a more precise characterization of the communities with high death rates.

Although there has been speculation of the role of public transit in enabling SARS-CoV-2 infection, mainly from studies that examined county-level associations [41–43], in our analysis, public transit use was not robustly associated with COVID-19 mortality. Our analysis focused on the city of Chicago, which, like many urban areas in the spring of 2020, saw more widespread infection than nonurban counties [4]. Our tract-level single-city analysis allowed us to avoid comparing areas in which both the transportation habits and viral exposure were systematically different, and therefore avoid potential sources of confounding.

Air quality has also been proposed as a risk factor for COVID-19 mortality, with evidence coming from analyses that covered large geographic areas (U.S. counties and administrative units across Europe) [14,44]. Because our analysis examined the relationship between air quality and COVID-19 within a uniformly urban area, we avoided comparing air quality (which tends to be worse in urban areas [45]) between populations with different potential to be exposed to the virus. Our CMAQ-WRF simulated air quality exposure also assigned individual residential exposures at a fine level of spatial resolution. We observed no robust associations between NO2, O3, or PM2.5 and COVID-19 mortality. While there was moderate evidence for a higher risk of mortality associated with NO2 among Hispanic/Latino residents and with O3 among white residents, given that the findings were not robust and also differed by race, we are cautious of overinterpretation of the observed associations. As high-resolution air quality models are developed that estimate the air quality during the pandemic, future work will be able to estimate the effect of acute exposure to poor air quality.

Our work focused on COVID-19 mortality. While currently available data on confirmed SARS-CoV-2 infections were limited, and during the spring of 2020, the ability to confirm an infection was limited due to uneven access to testing [46,47], future work that identifies separately the influence social determinants of health on risk of infection, severe disease, and mortality will bring additional context to these findings. Similarly, future work that can identify those who died while experiencing homelessness can help to identify risk factors that may be particularly relevant for that vulnerable population. We were also limited in the neighborhood characteristics that were available at the census tract level. Other characteristics would have provided additional context, such as obesity rates and percentage of residents working in “essential” businesses. While the inclusion of these additional variables (along with possible interaction terms) might produce a model with better prediction capacity, our model does clearly describe the associations between well-measured and widely available neighborhood characteristics and COVID-19 mortality. Finally, we emphasize that the identified neighborhood characteristics may not directly influence COVID-19 mortality, and the variables selected by the regularized regression may not ultimately be causally associated with COVID-19 mortality. Instead, we expect that the identified traits broadly reflect a legacy of policy choices that have negatively impacted the health of those living in the neighborhood [22], and further work is needed to examine how the identified characteristics may cooccur and possibly modify one another.

Given the disruptions that the COVID-19–related deaths brought, this analysis identifies communities that may have unmet needs of supportive services [48,49], and may benefit from the kind of “psychiatry-palliative care liaison teams” [50] that were deployed in New York City in response to the pandemic. Because SARS-CoV-2 continues to spread, this research can be used to help identify communities that may also benefit from additional resources and education that would improve their capacity to detect and prevent infection [51,52]. The results of this research can be incorporated into communication and training to make them more relevant to the most affected populations and begin to mitigate the COVID-19–specific health disparities experienced by these communities.

Acknowledgments

The authors would like to thank Natalia Derevyanny at the Cook County Medical Examiner for answering questions, as well as Bea Malsky for her helpful discussions about the data. The authors also thank William Trick and Keeki Hinami at Cook County Health and Hospital System for their assistance in identifying congregate living facilities. We also thank the Texas A&M geocoding services for providing geocoding services for COVID-19–related research.

References

[1] WHO Director-General’s opening remarks at the media briefing on COVID-19 - 11 March 2020 [accessed 07.02.20], https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19-11-march-2020.

[2] Gupta A, Madhavan MV, Sehgal K, Nair N, Mahajan S, Sehrawat TS, et al. Extrapolation of COVID-19 in the UK using a mathematical model. BMJ Open 2020;10:e037722.

[3] STAT’s Covid-19 Drugs, Vaccines Tracker. STAT 2020 [accessed 09.28.20], https://www.statnews.com/2020/04/27/drugs-vaccines-tracker/.

[4] Home. Johns Hopkins Coronavirus Resource Center [accessed 04.21.20], https://coronavirus.jhu.edu/.

[5] Ganj S, Kim L, Whitaker M, O’Halloran A, Cummings C, Holstein R, et al. Community-level associations between COVID-19 mortality and neighborhood characteristics: a population-based study in the New York Metropolitan Area. PLoS One 2020;15.e0230647.

[6] Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. Nature 2020.

[7] Oronce CIA, Scannell CA, Kawachi I, Tsugawa Y. Association Between State-Level Income Inequality and COVID-19 Cases and Mortality in the USA. J Gen Intern Med 2020.

[8] Kim SJ, Bostwick W. Social Vulnerability and Racial Inequality in COVID-19 Deaths in Chicago. Health EducBehav 2020.

[9] Yancy CW. COVID-19 and African Americans. JAMA 2020.

[10] Pettirini CM, Jones SA, Yang J, Rajagopalan H, O’Donnell LR, Chernyak Y, et al. Factors associated with hospitalization and critical illness among 4,103 patients with COVID-19 disease in New York City. MedRxiv 2020.

[11] Kullar R, Marcelin JR, Swartz TH, Piggott DA, Macias Gil R, Mathew TA, et al. Racial Disparity of Coronavirus Disease 2019 (COVID-19) in African American Communities. J Infect Dis 2020.

[12] Pinter-Wollman N, Jelić E, Wells NM. The impact of the built environment on health behaviours and disease transmission in social systems. Philosoph Trans R Soc B: Biol Sci 2018;373.

[13] Oyen Y. Assessing nitrogen dioxide (NO2) levels as a contributing factor to coronavirus (COVID-19) fatality. Sci Total Environ 2020;726:138605.

[14] Wu X, Nethery RC, Sabath BM, Braun D, Dominici F. Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study. MedRxiv 2020-2020.

[15] Rocklov J, Sjödin H. High population densities catalyse the spread of COVID-19. J Trav Med 2020;27.

[16] van Dam A, Cooney RE, Sabin ML. COVID-19 exacerbating inequalities in the US. Lancet 2020;395:1243–4.

[17] Bureau UC. Population Estimates and Projections (2010–2019). The United States Census Bureau [accessed 04.21.20], https://www.census.gov/data/developers/data-sets/popest-poprj.html.
Appendix

Figure s1. Daily deaths from novel coronavirus in Chicago through July 22, 2020.

Figure s2. Correlation between tract-level characteristics in Chicago.
Figure s3. Trends in racial and ethnic composition of COVID-19 deaths in Chicago.

Figure s4. Trends in age of death from COVID-19 deaths in Chicago.
Figure S5. Racial and ethnic composition of Chicago.
Figure S6. Variability in neighborhood characteristics by race. Variability by race of the of two neighborhood characteristics robustly associated with COVID-19 mortality among white residents: (A, B) percent of adults without a high school diploma, (C, D) Percent of population that is Hispanic/Latino, (E, F) Percent of population that is Asian, and (G, H) percent of the population that is white. To further investigate whether neighborhoods with high death rates had an outsized influence on the results, the plots were separated based on whether they occurred in neighborhoods with low to moderate COVID-19 death rates (i.e., less than 1.5 deaths per thousand) (right side) or in neighborhoods with high COVID-19 death rates (i.e., more than 1.5 deaths per thousand, left side).
Table S1
Regularized regression associations between neighborhood characteristics and COVID-19 death rates in the noninstitutionalized population of Chicago June 3 (1544 deaths) and June 26 (2118 deaths)

| Domain                  | Neighborhood characteristic                                      | June 3                               | June 26                              |
|-------------------------|-----------------------------------------------------------------|--------------------------------------|--------------------------------------|
| Infection               | % population 30+ living w/grandchildren <18                      | 1 (8%; 0.992 to 1.000)               | 1 (10%; 0.991 to 1.000)             |
|                         | % occupied units w/out complete kitchen                         | 1 (20%; 0.999 to 1.022)              | 1 (17%; 0.993 to 1.014)             |
|                         | % occupied units > 1 resident per room                          | 1 (21%; 1.000 to 1.014)              | 1 (34%; 1.000 to 1.020)             |
| Transportation          | % occupied units w/no car available                              | 1 (2%; 1.000 to 1.000)               | 1 (25%; 1.000 to 1.000)             |
|                         | % workers > 16 commuting public transit                         | 1 (26%; 1.000 to 1.006)              | 1 (29%; 1.000 to 1.006)             |
|                         | % workers > 16 commuting carpool                                 | 1 (7%; 1.000 to 1.003)               | 1 (10%; 1.000 to 1.005)             |
| Dense housing           | % units in buildings w/20+ units                                 | 1 (12%; 0.999 to 1.000)              | 1 (34%; 0.997 to 1.000)             |
|                         | % population without health insurance                           | 1 (40%; 0.999 to 1.000)              |                                      |
|                         | % households w/SNAP in last year                                | 1 (64%; 1.000 to 1.006)              | 1 (56%; 1.000 to 1.006)             |
|                         | % adults > 25 w/ high school diploma                            | 1 (68%; 1.000 to 1.010)              | 1 (79%; 1.000 to 1.010)             |
|                         | % adults > 25 w/ bachelor’s degree or higher                    | 1 (48%; 0.997 to 1.000)              | 0.999 (64%; 0.996 to 1.000)         |
| Transport              | % workers > 16 who work from home                               | 1 (8%; 1.000 to 1.009)               | 1 (7%; 1.000 to 1.005)              |
| Barriers to social distancing | % households w/broadband internet                             | 0.988 (100%; 0.982 to 0.994)         | 0.991 (100%; 0.985 to 0.998)        |
|                          | % population without health insurance                           | 1.015 (98%; 1.003 to 1.029)          | 1.014 (99%; 1.004 to 1.027)         |
| Mortality              | % population w/a primary care provider                          | 1 (16%; 0.996 to 1.000)              | 1 (16%; 0.995 to 1.000)             |
| Health care access      | Rate of heart disease deaths                                    | 1 (1%; 1.000 to 1.000)               | 1 (0%; 1.000 to 1.000)              |
| Comorbid conditions     | Rate of diabetes-related deaths                                 | 1 (26%; 1.000 to 1.003)              | 1 (17%; 1.000 to 1.001)             |
| Age and gender          | Rate of nephrotic disease deaths                                | 1.003 (81%; 1.000 to 1.010)          | 1.001 (62%; 1.000 to 1.008)         |
| % population male       | 1 (12%; 1.000 to 1.005)                                        | 1 (24%; 1.000 to 1.017)              |                                      |
| % population between 65 and 74 | 1.006 (67%; 1.000 to 1.030)                                    | 1.008 (72%; 1.000 to 1.040)          |                                      |
| % population 75+        | 1.003 (62%; 1.000 to 1.025)                                     | 1.004 (64%; 1.000 to 1.032)          |                                      |
| % population below poverty | 1 (2%; 1.000 to 1.002)                                        | 1 (18%; 1.000 to 1.003)              |                                      |
| Indicators of poverty   | % population > 16 unemployed                                    | 1 (18%; 1.000 to 1.004)              | 1 (6%; 1.000 to 1.001)              |
| Air quality             | % occupied units w/rent > 35% income                            | 1 (12%; 1.000 to 1.002)              | 1 (19%; 1.000 to 1.003)             |
|                          | Rated C/D by the HOLC                                           | 1 (38%; 1.000 to 1.135)              | 1 (34%; 1.000 to 1.134)             |
| Structural             | Average NO2 in tract (ppbV)                                    | 1 (36%; 0.967 to 1.000)              | 1 (28%; 0.975 to 1.000)             |
| Structural racism       | Average summertime O3 in tract (ppbV)                           | 1 (4%; 1.000 to 1.000)               | 1 (8%; 1.000 to 1.009)              |
|                          | Average PM2.5 in tract (ug/m3)                                  | 1 (28%; 0.913 to 1.000)              | 1 (35%; 0.908 to 1.000)             |
| Model metrics          | % population black                                              | 1 (32%; 1.000 to 1.002)              | 1 (12%; 1.000 to 1.001)             |
|                          | % population Asian                                             | 0.997 (88%; 0.999 to 1.000)          | 0.995 (96%; 0.988 to 1.000)         |
|                          | % population white                                             | 0.994 (100%; 0.990 to 0.997)         | 0.993 (100%; 0.990 to 0.996)        |
|                          | R²-like statistic                                               | 0.2600                               | 0.2525                               |