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.
Convergence of COVID-19 and chronic air pollution risks: Racial/ethnic and socioeconomic inequities in the U.S

Jayajit Chakraborty

Department of Sociology and Anthropology, University of Texas at El Paso; 500 West University Avenue, El Paso, TX 79968, USA

ARTICLE INFO

Keywords:
COVID-19
Air pollution
Environmental justice
Race/ethnicity
Socioeconomic status

ABSTRACT

Recent research suggests greater COVID-19 prevalence in areas burdened with higher exposure to chronic air pollution, but previous studies have not examined if socially disadvantaged populations are more likely to reside in communities located at the convergence of both COVID-19 and air pollution health risks. This article presents a national scale U.S. study that investigates whether racial/ethnic minorities, socioeconomically deprived residents, and other vulnerable groups are significantly overrepresented in counties where significantly higher COVID-19 incidence spatially coincides with higher respiratory health risks from outdoor exposure to hazardous air pollutants (HAPs). COVID-19 data from the Johns Hopkins Center for Systems Science and Engineering database are linked to respiratory risk estimates from the U.S. Environmental Protection Agency’s National Air Toxics Assessment and variables from the 2018 American Community Survey. Bivariate local measures of spatial association are implemented to identify county clusters representing relationships between COVID-19 incidence rate and respiratory risk from HAP exposure. Socio-demographic characteristics of these clusters are compared using bivariate statistical tests and multivariable generalized estimating equations. Counties where greater COVID-19 incidence coincides significantly with higher HAP respiratory risk contain disproportionately higher percentages of non-Hispanic Black, socioeconomically deprived, and uninsured residents than all other U.S. counties, after controlling for spatial clustering, population density, older age, and other contextual factors. These significant socio-demographic inequities represent an important starting point for more detailed investigations of places facing the double burden of elevated COVID-19 prevalence and air pollution exposure, and also emphasize the urgent need to develop mitigation strategies for addressing both COVID-19 and chronic air pollution in socially vulnerable communities.

1. Introduction

Since the novel coronavirus disease (COVID-19) is primarily a respiratory disease characterized by respiratory illness, recent studies suggest that exposure to outdoor air pollution can increase the risks of contracting COVID-19, severity of COVID-19 infection symptoms, and worsen the prognosis of COVID-19 patients (Hendryx and Luo 2020a; Wu et al., 2020). Chronic air pollution has been reported to be a contributing factor to higher COVID-19 cases, hospitalizations, and/or deaths in China (Zhu et al., 2020), Italy (Contricini et al., 2020), Spain (Ogen 2020), and The Netherlands (Cole et al., 2020). Increased rates of COVID-19 have also been found in counties with elevated air pollution levels in the U.S. For example, Wu et al. (2020) have shown that a small increase in long-term exposure to fine particulate matter leads to a large increase in the COVID-19 death rate, while Hendryx and Luo (2020a) found COVID-19 prevalence and fatality rates to be significantly associated with greater exposure to air pollution emissions from multiple sources.

Before the COVID-19 outbreak, numerous studies had used a distributive environmental justice (EJ) research framework (Holifield et al., 2018) to analyze spatial and social disparities in the distribution of air pollution in the U.S. A majority of these empirical studies found racial/ethnic minorities, individuals of lower socioeconomic status (SES), and other socially disadvantaged groups in the U.S. to be disproportionately exposed to outdoor air pollution (Apelberg et al., 2005; Bell and Ebisu 2012; Chakraborty 2009; Clark et al. 2014, 2017; Collins et al., 2017; Wilson et al., 2015), and that inequitable air pollution exposure contributes to health disparities (Brender et al., 2011; Grineski et al., 2013; Hackbart et al., 2011). Recent reports on COVID-19 morbidity and mortality also indicate a higher burden among minorities and lower SES residents (Devakumar et al., 2020; Eligon et al., 2020; Ray 2020; Yancy 2020). Greater COVID-19 incidence or
fatality rates have also been documented in U.S. counties or ZIP codes containing higher percentages of non-White residents, people in poverty, and other socially vulnerable groups (Finch and Finch 2020; Hendryx and Luo 2020b; Karaye and Horney 2020; Mahajan and Larkins-Pettigrew 2020; Maroko et al., 2020).

While prior studies on the EJ implications of air pollution and recent research on COVID-19 outcomes have both found socially disadvantaged communities to be adversely impacted, racial/ethnic and socioeconomic disparities among people residing at the convergence of both COVID-19 and chronic air pollution health risks have not been systematically examined. This article addresses this gap by: (1) identifying counties facing significantly higher exposure to both COVID-19 and respiratory risks from hazardous air pollutants (HAPs) in the continental U.S.; and (2) investigating if socially disadvantaged population groups are more likely to reside in such counties. The central research question is framed as follows: are racial/ethnic minorities, socioeconomically deprived residents, and other vulnerable groups significantly over-represented in counties where significantly higher COVID-19 prevalence coincides with higher respiratory risks from outdoor HAP exposure? PUBLICLY AVAILABLE SECONDARY DATA AND LOCAL MEASURES OF SPATIAL ASSOCIATION are used to identify geographic clusters of counties that represent statistical relationships between COVID-19 incidence rate and HAP respiratory risks. Socio-demographic characteristics of counties representing the spatial coincidence of significantly higher COVID-19 incidence and higher respiratory risk from HAP exposure are then statistically compared to those in: (1) all remaining counties in the continental U.S.; and (2) counties where significantly lower COVID-19 prevalence coincides with lower HAP respiratory risk, using both bivariate statistical tests and multivariable generalized estimating equations that account for spatial clustering in the data.

2. Materials and methods

2.1. COVID-19 and air pollution respiratory risks

Data on COVID-19 cases were obtained from the Johns Hopkins University Center for Systems Science and Engineering database (JHU 2020) on November 13, 2020, for all counties in the continental U.S. This database provides the most comprehensive and reliable county-level information on COVID-19 incidence from the Centers for Disease Control and Prevention and state health departments. The total number of COVID-19 cases in the 3108 counties of the continental U.S. (lower 48 states and Washington DC) was 10,355,623 on the date this information was downloaded. The number of confirmed cases per 100,000 people in each county (from January 22 to November 13, 2020), was used to represent the COVID-19 incidence rate for this study.

The analysis of outdoor air pollution risks focuses on HAPs, which are also referred to as air toxics or toxic air pollutants. These include 187 specific airborne substances identified by Section 112 of the Clean Air Act (1990) that are known to cause cancer and respiratory health problems, and other serious health impacts (USEPA 2020a). The USEPA’s latest National Air Toxics Assessment (NATA), released in August 2020, was used as the data source for county level estimates of cumulative respiratory risk from HAPs (USEPA 2020b). The NATA is considered to be the best available secondary data source for spatially explicit characterization of HAP exposure in the U.S. and has been used to measure cancer and respiratory health risks associated with the inhalation of HAPs in previous EJ and environmental health research (Apelberg et al., 2005; Chakraborty 2009; Collins et al., 2017; Grineski et al., 2013; McCarthy et al., 2009; Morello-Frosch and Jossdale 2006).

Information on HAP emissions from a wide range of sources are collected from the 2014 National Emissions Inventory and multiple steps are utilized in the latest NATA to generate modeled estimates of respiratory risk, as detailed elsewhere (USEPA 2020c). For each HAP known to cause adverse respiratory health outcomes, the NATA calculates a hazard quotient, which is estimated as the ratio of ambient air concentration to the chemical’s inhalation reference concentration, defined as the amount of toxicity below which long-term exposure to the general population is not expected to result in adverse respiratory effects. Respiratory risk estimates in the NATA are scaled in terms of a respiratory hazard index, which represents the sum of hazard quotients for all HAPs that cause adverse respiratory effects. This index summarizes the cumulative impacts of all relevant HAPs for which respiratory effects were the key negative health impact. The NATA respiratory hazard index includes risks associated with inhalation exposure to HAPs released by major stationary sources (e.g., industrial factories), smaller stationary sources (e.g., small manufacturers), on-road mobile sources (e.g., cars and trucks), non-road mobile sources (e.g., trains, construction vehicles), non-point sources (e.g., agricultural livestock, residential wood combustion), fires, biogenic sources, secondary sources, and background concentrations (i.e., contributions from distant or natural sources).

For this study, values of the total respiratory hazard index were obtained from the NATA Assessment Results website (USEPA 2020d) for 3107 counties in the continental U.S. Since respiratory hazard index values were unavailable for Oglala Lakota County, North Dakota, this county was not included in the statistical analysis. Descriptive statistics for both COVID-19 incidence rates and the NATA respiratory hazard index are provided in Table 1.

2.2. County socio-demographic characteristics

Data on selected population characteristics of U.S. counties were obtained from the 2018 American Community Survey (ACS) five-year estimates. This study encompassed variables that are commonly used in distributive EJ studies conducted in the U.S., as well as variables that have been utilized in recent research on the social distribution of COVID-19 outcomes. The names and summary statistics for all these variables are provided in Table 1. To examine race/ethnicity, the analysis included variables for the percentage of individuals in each county

| Health risk variables: | Min | Max | Mean | SD |
|-----------------------|-----|-----|------|----|
| COVID-19 incidence rate (confirmed cases per 100,000 people) | 0.0 | 17961.4 | 3428.9 | 1997,685 |
| Cumulative respiratory risk from HAP exposure (NATA respiratory hazard index) | 0.1 | 1.16 | 0.36 | 0.14 |
| Socio-demographic variables: | | | | |
| % White (non-Hispanic) | 0.7 | 100.0 | 76.8 | 19.9 |
| % Hispanic | 0.0 | 99.1 | 80.4 | 11.9 |
| % Black (non-Hispanic) | 0.0 | 87.4 | 90.0 | 14.5 |
| % American Indian (Non-Hispanic) | 0.0 | 89.6 | 1.6 | 6.5 |
| % Asian (non-Hispanic) | 0.0 | 35.7 | 1.3 | 2.3 |
| % Other (non-Hispanic non-White) | 0.0 | 5.5 | 0.2 | 0.3 |
| Socioeconomic deprivation (component loadings) | -2.1 | 6.3 | 0.0 | 0.0 |
| % Persons (age 25+) with no high school diploma (0.81) | 1.2 | 66.3 | 13.4 | 6.3 |
| % Persons age 5+ speak English “less than well” (0.34) | 0.0 | 30.4 | 1.7 | 2.8 |
| % Persons below poverty (0.87) | 2.3 | 55.1 | 15.6 | 6.5 |
| % Civilians (age 16+) unemployed (0.76) | 0.0 | 26.4 | 5.7 | 2.8 |
| % Households with no vehicle available (0.66) | 0.0 | 77.0 | 6.2 | 3.6 |
| % Persons aged 65 and older | 3.8 | 55.6 | 18.4 | 4.5 |
| % Civilian non-institutionalized with a disability | 3.8 | 33.7 | 16.0 | 4.4 |
| % Civilian non-institutionalized with no health insurance | 1.7 | 42.4 | 10.0 | 5.0 |
| Population density (persons per square mile) | 0.2 | 72053.0 | 272.9 | 1813.1 |

n = 3107 counties.
identified as Hispanic/Latino (of any race) and each of the following non-Hispanic groups: White, Black, American Indian, Asian, and multi/race. The term ‘non-Hispanic’ is not used to refer to these non-Hispanic racial/ethnic groups in subsequent sections of this article. The percentage of non-Hispanic White residents was used in the bivariate statistical comparisons but excluded from multivariable models so that results for the minority racial/ethnic groups can be interpreted relative to non-Hispanic Whites.

To assess county SES, five specific ACS variables were used: the percentages of the population aged 25 or more years with no high school education, population aged 5 or more years with limited English language proficiency (i.e., do not speak English well), individuals with an annual income below the family poverty level (i.e., persons below poverty), civilians aged 16 and over who are unemployed, and households with no vehicles available (i.e., zero-car households). Following recent studies utilizing these five variables to measure SES (Chakraborty et al., 2019a, 2019b), they were combined into a single factor for multivariable analysis because these variables are significantly correlated with each other. Specifically, a single and robust index of socioeconomic deprivation was computed with principal components analysis (PCA) using IBM SPSS Statistics software and these five SES variables (Cronbach’s $\alpha = 0.74$). This single component extracted through PCA explained 61% of the variable’s total variance. The results of the PCA, or component loadings for each of the five variables comprising this factor (socioeconomic deprivation index), are provided in Table 1.

Several additional variables that are relevant to analyzing both COVID-19 prevalence and air pollution exposure were included. The first of these variables was the percentage of people aged 65 years or more, since older populations have been linked to higher likelihood of COVID-19 infections and greater susceptibility to health problems caused by air pollution (Bell and Ebisu 2012; Sandolu 2020). The second variable was the percentage of civilian non-institutionalized persons reporting a disability in the ACS. Recent studies have demonstrated that people with disabilities are disproportionately exposed to environmental health hazards in the U.S. and also adversely impacted by COVID-19 (Chakraborty, 2020; Turk and McDermott 2020). The third variable used was the percentage of civilian non-institutionalized persons without health insurance, a well-documented measure of socioeconomic vulnerability that has also been linked to disproportionate COVID-19 and respiratory health outcomes due to HAP exposure (Grenneski et al., 2013; King 2020). Finally, the multivariable analysis also includes and adjusts for population density, calculated by dividing the total county population by the area of the county in square miles, because densely populated areas are more likely to have higher COVID-19 cases (Maroko et al., 2020) as well as contain emission-generating activities that increase health risks from HAPs (Chakraborty 2009; Collins et al., 2017).

### 2.3. Geospatial and statistical analysis

Local Indicators of Spatial Association (LISA) are implemented to identify geographic clusters of counties based on the statistical relationship between COVID-19 incidence rate and the respiratory hazard index for HAP exposure. Unlike global measures of spatial association that evaluate relationships for the entire study area, LISA statistics provide a spatial correlation measure for each unit of analysis. Specifically, the local bivariate Moran’s I statistic (Anselin 1995) was estimated to determine how the association between COVID-19 incidence rate and the HAP respiratory hazard index varies across counties in the continental U.S. To identify spatial clusters of significant associations and counties where these associations are statistically significant, 999 random permutations and a significance level of 0.05 ($p < 0.05$) were used.

The local Moran’s I statistic is used to test the extent of significant spatial clustering of similar or dissimilar values around each analytic unit. Four statistically significant categories of spatial association can be derived from this statistic, with high-high and low-low categories for spatial clustering of similar values (positive relationships), and high-low and low-high categories for spatial clustering of dissimilar values (negative relationships). The fifth category comprises analytic units where spatial correlations are not statistically significant ($p > 0.05$). In this case, the high-high category implies that high COVID-19 incidence rate in a county is significantly associated with high HAP respiratory risk in neighboring counties ($p < 0.05$), while low-low category indicates that a county with low COVID-19 incidence is significantly associated with low respiratory risk in neighboring counties ($p < 0.05$). Spatial contiguity was measured based on the frequently used first-order queen criterion, which defines neighbors as adjacent spatial units (counties) that share a common border or vertex with the spatial unit (county) of interest. All analyses were conducted using the open-source spatial analysis software program GeoDa (Anselin et al., 2006).

To determine whether socially disadvantaged groups are overrepresented in counties where higher COVID-19 exposure spatially coincides with higher HAP respiratory risk (high-high type), both bivariate and multivariable analyses are used. The first phase focuses on estimating and comparing mean values of socio-demographic variables in the high-high cluster of counties to: (a) all other counties of the continental U.S.; and (b) counties in the low-low cluster; based on the independent two-sample t-test for comparing group means.

The second phase utilizes generalized estimating equations (GEEs), a multivariable modeling technique that extends the generalized linear model to accommodate clustered data (Garson 2012; Liang and Zeger 1986; Nelder and Wedderburn 1972). GEEs are appropriate for this study since they are useful for analyzing clustered dichotomous variables and relax several assumptions of traditional regression models (Chakraborty et al., 2019a, 2019b; Maldonado et al., 2016). An important assumption for estimating a GEE is that observations (i.e., counties) within a cluster are correlated, while observations from different clusters are independent. The state in which a county is located was used to define clusters to account for potential correlation in counties within the same state, because of similarities in COVID-19 response and testing policies, as well socio-cultural and healthcare system characteristics (Wu et al., 2020). This cluster definition resulted in 49 clusters, with number of counties per cluster ranging from 1 (Washington DC) to 254 (Texas). GEEs also require the specification of an intra-cluster dependency correlation matrix (Garson 2012). After experimentation with several different specifications available in IBM SPSS Statistics software, the ‘independent’ correlation matrix was selected since this specification yielded the best statistical fit.

Two multivariable GEEs were utilized to investigate the simultaneous effects of the independent variables on the odds of a county being classified as the high-high type, or where significantly higher COVID-19 exposure coincides with higher cumulative respiratory risk from HAP exposure based on the LISA results. Both these models were specified using the binomial distribution with the logit function. The first GEE model encompassed all counties in continental U.S. The dichotomous dependent variable was coded as ‘1’ if the county was classified as high-high and ‘0’ for all other counties (non-high-high). The second GEE model included counties that were categorized as either high-high or low-low. The dependent variable for this model was coded as ‘1’ if the county was classified as high-high and ‘0’ if low-low. All independent variables were standardized before inclusion in these models. Potential multicollinearity among the variables was also checked based on variance inflation factor, tolerance, and condition index criteria. These diagnostics indicated that inferences from the GEEs are not affected by multicollinearity. The statistical significance of independent variables is estimated using two-tailed $p$-values from the Wald chi-square test.

### 3. Results

The LISA results based in the bivariate local Moran’s I statistic are
shown in Fig. 1 and summarized in Table 2. In Fig. 1, each county in the continental U.S. is classified into five clusters based on the nature and significance of the spatial correlations between COVID-19 incidence rate and the HAP respiratory hazard index. The high-high category, which represents significantly high COVID-19 incidence spatially associated high HAP respiratory risks, contains 422 or almost 14% of counties in the continental U.S. These counties are located mainly in the U.S. S. Southeast, as well as in California. The top five counties in the high-high category, based the bivariate local Moran’s I values, comprise Chattahoochee and Stewart in Georgia, Lincoln in Arkansas, Lafayette in Florida, and East Felicia in Louisiana. In these five counties, COVID-19 incidence rates are at least three times higher and cumulative respiratory risks from HAPs are almost twice as high compared to their respective national averages. The low-low category also comprises about 14% of all U.S. counties that are located primarily in the Mountain region of the U.S. West, northern areas of the Midwest (e.g., Michigan and Minnesota), and several states in the northeast. Less than 25% of all counties in the continental U.S. are collectively classified as either low-high or high-low. Counties with significantly low COVID-19 incidence surrounded by high HAP respiratory risk counties (low-high) can be found mainly in California, Arkansas, and Georgia, while those in the high-low category are scattered throughout the Mountain and Midwest regions.

Table 2 provides the mean values of the socio-demographic variables in each of the five spatial clusters depicted in Fig. 1. The high-high category is characterized by substantially higher percentages of Black residents, socioeconomically deprived persons, and other socially vulnerable groups, compared to the other four clusters and the continental U.S. as a whole. The first set of t-tests compares the means in high-high category to those in the rest of the continental U.S., referred to as ‘non high-high’ in the tables. With regards to race/ethnicity, the high-high category contains significantly higher percentages of Black residents, but significantly lower percentages of White, Hispanic, Asian, and American Indian residents, when compared to the non high-high category of counties. With the exception of English language proficiency, all socioeconomic deprivation variables show significantly higher means in the high-high category than the rest of the U.S. The mean percentage of people aged 65 or years is significantly lower, while the percentages of people with a disability and without health insurance are significantly higher, among counties in the high-high category than those in the non high-high category.

The second set of t-tests focuses on comparing the means in high-high category to only those in the low-low category. For race/ethnicity, the high-high category indicates significantly higher percentages of Black and Asian residents, but significantly lower percentages of White, Hispanic, and American Indian residents, when compared to counties in the low-low category. This disparity is particularly staggering for the mean Black percentage in the high-high category, which is about 18 times higher than that in the low-low category. All five socioeconomic deprivation variables, including the combined index, show significantly higher means in the high-high category. The largest disparities are observed for the percentages of adults without high school education and people in poverty; both these means are almost twice as high in the high-high category compared to the low-low category. The mean percentages of people with a disability and without health insurance are again significantly higher for counties in the high-high category, while the mean percentage of those aged 65 or more years is significantly lower.

Results from the multivariable GEEs are shown in Table 3, which includes coefficients and their significance (Wald’s chi-square test) for the independent variables when entered simultaneously in each model. In the first model, the odds of high-high classification are examined with respect to all other categories in the continental U.S. (non high-high category). This GEE reveals that counties with higher percentages of Black, socioeconomically deprived, and uninsured residents are significantly more likely, while those with higher percentages of American Indians and people above 65 years of age are less likely, to be located in the high-high category. More specifically, a one standard deviation increase in the Black percentage, socioeconomic deprivation index, and uninsured percentage is respectively associated with approximately 155%, 59%, and 83% increases in the odds of being classified as a high-high (vs. non high-high) county, after controlling for spatial clustering and other explanatory variables.

The second GEE directly compares the odds of the high-high to the low-low classification. The odds of belonging to the high-high category in this GEE are also significantly greater in counties with higher percentages of Black, socioeconomically deprived, and uninsured residents. A one standard deviation increase in the Black percentage, socioeconomic deprivation index, and uninsured percentage is associated with approximately 6174%, 202%, and 299% increases, respectively, in the odds of being classified as a high-high (vs. low-low) county, after controlling for spatial clustering and other explanatory variables. The only variables that yielded a significantly negative coefficient in this GEE model include the percentage of people who are American Indians and aged 65 or more years.

Fig. 1. Significant spatial clusters based on bivariate local Moran’s I for correlation between COVID-19 incidence rate and cumulative respiratory risk from hazardous air pollutants (HAPs).
identifying and comparing socio-demographic characteristics of this high-high category of counties, when compared to the rest of the Black population, socioeconomically deprived residents, people with disabilities, about one out of seven counties in the continental U.S. are facing this exposure. LISA measures and related significance tests indicated that convergence of COVID-19 and air pollution risks burdens in the U.S. by counties are located primarily in southeastern states of the U.S. and double burden of COVID-19 prevalence and HAP respiratory risks. These

4. Discussion

*GEEs are based on binomial logit specification and an independent correlation matrix.

| Variables | Cont. USA Mean | Groups means | Non High-High Mean | High-High vs Non High-High | High-High vs Low-Low |
|-----------|----------------|--------------|-------------------|---------------------------|---------------------|
| % White (non-Hispanic) | 76.8 | 60.0 | 85.5 | 70.3 | 83.3 | 78.2 | 79.5 | -19.5 | -0.001 | -25.5 | -0.001 |
| % Hispanic | 9.3 | 6.6 | 8.8 | 10.1 | 8.5 | 10.3 | 9.7 | -3.1 | -0.001 | -2.1 | -0.01 |
| % Black (non-Hispanic) | 9.0 | 29.8 | 1.6 | 13.7 | 1.1 | 7.0 | 5.7 | 24.1 | -0.001 | 28.2 | -0.001 |
| % American Indian (non-Hispanic) | 1.6 | 0.8 | 1.5 | 1.1 | 4.5 | 0.9 | 1.6 | -0.8 | -0.001 | -0.7 | -0.01 |
| % Asian (non-Hispanic) | 1.3 | 1.0 | 0.8 | 2.4 | 0.9 | 1.4 | 1.3 | -0.3 | -0.01 | 0.3 | -0.01 |
| % Other (non-Hispanic non-White) | 0.2 | 0.2 | 0.1 | 0.2 | 0.1 | 0.2 | 0.2 | 0.0 | 0.164 | 0.1 | -0.01 |

Socioeconomic deprivation index
| % Persons speak English “less than well” | 0.0 | 0.9 | -0.5 | 0.4 | -0.6 | -0.1 | -0.1 | 1.0 | -0.001 | 1.4 | -0.001 |
| % Persons below poverty | 15.6 | 21.8 | 13.4 | 17.0 | 12.7 | 15.2 | 14.6 | 7.2 | -0.001 | 8.4 | -0.001 |
| % Civilians (age 16+) unemployed | 5.7 | 7.7 | 4.9 | 7.2 | 3.8 | 5.8 | 5.4 | 2.2 | -0.001 | 2.8 | -0.001 |
| % Households with no vehicle available | 6.2 | 7.9 | 5.7 | 6.6 | 4.8 | 6.2 | 5.9 | 2.0 | -0.001 | 2.2 | -0.001 |
| % Persons aged 65 and older | 18.4 | 16.7 | 20.2 | 18.4 | 18.9 | 18.2 | 18.7 | -2.0 | -0.001 | 3.5 | -0.001 |
| % Civilian non-inst. with a disability | 16.0 | 17.4 | 15.5 | 17.3 | 13.1 | 16.3 | 15.7 | 1.6 | -0.001 | 1.9 | -0.001 |
| % Civilian non-inst. with no insurance | 10.0 | 12.7 | 8.2 | 10.6 | 8.9 | 10.0 | 9.6 | 3.2 | -0.001 | 4.5 | -0.001 |

Population density
| Number of counties (n) | 272.9 | 311.3 | 60.1 | 735.5 | 58.3 | 310.1 | 267.0 | 44.2 | 0.641 | 251.2 | -0.05 |
| Percent of counties (%) | 100.0 | 13.6 | 14.3 | 8.5 | 14.9 | 48.8 | 86.4 |

*Based on two-sample t-test of means.

Table 3
Multivariable generalized estimating equations (GEE) for predicting the odds of county location in high-high cluster.

| Variables | High-High Vs. Non High-High Counties<sup>1</sup> | High-High Vs Low-Low Counties<sup>1</sup> |
|-----------|-----------------------------------------------|------------------------------------------|
| % Hispanic | -0.52 | 0.13 | 0.09 | 0.59 | 0.095 | -0.54 | -0.12 | 0.20 | 0.58 | 0.155 |
| % Black (non-Hispanic) | 0.94 | 0.52 | 1.36 | 2.55 | <0.001 | 4.14 | 2.32 | 5.96 | 62.74 | <0.001 |
| % American Indian (Non-Hispanic) | -0.37 | -0.07 | -0.08 | 0.69 | <0.05 | -0.09 | -0.15 | -0.05 | 0.55 | -0.05 |
| % Asian (non-Hispanic) | -0.07 | -0.30 | 0.16 | 0.93 | 0.540 | 0.50 | 0.19 | 1.19 | 1.65 | 0.156 |
| % Other (non-Hispanic non-White) | 0.05 | -0.06 | 0.17 | 1.05 | 0.373 | -0.16 | -0.55 | 0.23 | 0.85 | 0.415 |
| Socioeconomic deprivation index | 0.46 | 0.20 | 0.90 | 1.59 | <0.05 | 1.10 | 0.11 | 2.10 | 3.02 | -0.05 |
| % Persons aged 65 and older | -0.72 | -1.07 | -0.37 | 0.49 | <0.001 | -0.95 | -1.88 | -0.03 | 0.39 | -0.05 |
| % Civilian non-institutionalized with a disability | 0.32 | -0.13 | 0.78 | 1.28 | 0.164 | 0.80 | -0.12 | 1.72 | 2.23 | 0.087 |
| % Civilian non-institutionalized with no insurance | 0.60 | 0.18 | 1.03 | 1.83 | <0.01 | 1.38 | 0.43 | 2.34 | 3.99 | <0.01 |

Population density
| Intercept | -0.09 | 0.49 | 0.32 | 0.92 | 0.674 | -0.31 | -0.77 | 0.15 | 0.73 | 0.181 |
| QIC 1708.3 | 328.9 |
| N (counties) | 3107 | 865 |

-1.608

4. Discussion

This study sought to provide new and important insights on the convergence of COVID-19 and air pollution risks burdens in the U.S. by identifying and comparing socio-demographic characteristics of counties where higher COVID-19 incidence is spatially and statistically associated with higher respiratory health risks from outdoor HAP exposure. LISA measures and related significance tests indicated that about one out of seven counties in the continental U.S. are facing this double burden of COVID-19 prevalence and HAP respiratory risks. These counties are located primarily in southeastern states of the U.S. and California. Bivariate statistical analysis revealed that non-Hispanic Blacks, socioeconomically deprived residents, people with disabilities, and those without health insurance are significantly overrepresented in this high-high category of counties, when compared to the rest of the continental U.S. and the low-low category, in particular. The largest relative disparities were observed for the percentages of non-Hispanic Black population, adults without high school education, and people in poverty. White and older (age 65 or more) residents, in contrast, are significantly overrepresented in counties with both low COVID-19 prevalence and low HAP respiratory risks. Multivariable GEE analysis confirmed that the odds of locating in the high-high category remain significantly greater in counties with higher percentages of Black, socioeconomically deprived, and uninsured residents, even after controlling for the effects of spatial clustering, population density, older age, and other relevant factors.

For other racial/ethnic minority groups, the GEEs indicated a significantly lower Hispanic percentage in counties associated with high-high category and the Asian percentage showed a non-significant relationship. However, both Asians and Hispanics are relatively

Table 2
Socio-demographic characteristics of counties within spatial clusters representing bivariate associations between COVID-19 incidence rate and cumulative respiratory risk from hazardous air pollutants.

J. Chakraborty

Environmental Research 193 (2021) 110586
overrepresented in the low-high category of counties (e.g., California) where respiratory risks from air pollution are substantially high, regardless of their COVID-19 prevalence. The American Indian percentage also revealed a significantly negative association with the odds of locating in the high-high category, in part, because of the over-representation of this group in the high-low category of counties. Although large proportions of American Indians reside in counties characterized by smaller population density and lower air pollution exposure that are located in southwestern or midwestern states, many of them are facing significantly high risks from exposure to COVID-19 in these counties.

This study contributes to the existing literature on distributive EJ analysis of air pollution and emerging literature on the socially inequitable impacts of the COVID-19 pandemic by documenting the significant overrepresentation of non-Hispanic Blacks, socioeconomically deprived residents, and those without health insurance in U.S. counties located at the intersection of COVID-19 and respiratory risks from HAPs. The significant social disparities revealed in this research highlight the urgent need to develop mitigation and intervention strategies that focus on addressing both COVID-19 and chronic air pollution risks for minority and socioeconomically vulnerable communities. The geospatial methodology used in this study can also be applied to identify COVID-19-impacted areas and populations burdened with other types of non-respiratory health risks (e.g., cancer or neurological), as well as other kinds of pollution (e.g., exposure to criteria air pollutants or proximity to hazardous waste facilities) that are likely to amplify the negative environmental and health impacts.

While this study represents an important starting point for documenting the double burden of exposure to both COVID-19 and air pollution risks for socially disadvantaged populations, it is important to consider certain limitations. First, the statistical analyses are restricted to county level associations since COVID-19 data for smaller geographic units (e.g., census tracts) and individual cases are currently unavailable for the entire U.S. These county level results cannot be used for inferences regarding COVID-19 outcomes or respiratory risk from HAP exposure for individuals residing in these counties. Future research should focus on specific U.S. states, counties, and urban areas, as well as utilize more detailed data at the neighborhood and household levels to clarify the relationships between COVID-19 susceptibility, chronic air pollution exposure, and socio-demographic characteristics of impacted residents. Second, data on respiratory risks from long-term exposure to HAPs were based on the latest NATA (2014), but these estimates may not represent the most recent conditions or acute exposures to HAPs. Recent reports suggest that total HAP emissions could be declining nationwide, although many urban areas are experiencing local or neighborhood level increases (Trinity Consultants, 2019). The county level distribution of HAP respiratory risk can be expected to have remained relatively unchanged, an assumption supported by previous EJ studies utilizing the NATA and comparison of NATA risk estimates from 2014 to 2011. Third, the analysis of COVID-19 prevalence did not consider the potential impacts of the complex lockdown rules that were in effect in some states relative to others, and in specific counties within certain states. Finally, additional factors that explain COVID-19 burdens and their county level patterns such as testing capacity, access, and availability also need to be included in future research.

5. Conclusions

Both air pollution and COVID-19 are well-documented to cause or exacerbate respiratory health problems. Recent studies argue that long-term exposure to air pollution adversely affects respiratory systems and increases mortality risk, in addition to influencing COVID-19 susceptibility and severity. This study addressed the growing need to identify places and people at the convergence of both COVID-19 and respiratory health risks from air pollution, as well as determine if socially disadvantaged groups are more likely to reside in these communities. Results revealed that non-Hispanic Black, socioeconomically deprived, and uninsured residents are significantly overrepresented in counties where significantly higher COVID-19 incidence coincide with high respiratory risk from HAP exposure, even after controlling for spatial clustering, population density, and other relevant socio-demographic factors. The mean non-Hispanic Black percentage, in particular, is about five times higher in counties at the intersection of high COVID-19 prevalence and HAP respiratory risks, compared to other counties in the continental U.S.

These findings regarding the location and social characteristics of counties suffering from the double burdens of COVID-19 and outdoor HAP risks can be used to inform more equitable public policies that focus on addressing racial/ethnic and socioeconomic disparities, as well as allocate resources to mitigate the adverse effects of the pandemic in areas facing elevated respiratory health risks from long-term air pollution exposure. The results of this study also emphasize the need to continue enforcing existing air pollution regulations, some of which were recently relaxed by the USEPA in response to the COVID-19 pandemic (Wu et al., 2020). The failure to mitigate toxic air pollution could not only increase the number of COVID-19 hospitalizations and fatalities, but also contribute to other respiratory illnesses and deaths that draw resources away from COVID-19 patients and further amplify the racial/ethnic and social disparities reported in this study.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Anselin, L., 1995. Local indicators of spatial association— LISA. Geogr. Anal. 27 (2), 93-115.
Anselin, L., Syabri, I., Kho, Y., 2006. GeoDa: an introduction to spatial data analysis. Geogr. Anal. 38 (1), 5–22.
Apelberg, B.J., Buckley, T.J., White, R.H., 2005. Socioeconomic and racial disparities in cancer risk from air toxics in Maryland. Environ. Health Perspect. 113 (6), 693–699.
Bell, M.I., Ebisu, K., 2012. Environmental inequality in exposures to airborne particulate matter components in the United States. Environ. Health Perspect. 120, 1699–1705.
Brender, J.D., Maantay, J.A., Chakraborty, J., 2011. Residential proximity to environmental hazards and adverse health outcomes. Am. J. Publ. Health 101 (S1), S37–S52.
Chakraborty, J., 2009. Automobiles, air toxics, and adverse health risks: environmental inequities in Tampa Bay, Florida. Ann. Assoc. Am. Geogr. 99 (4), 674–697.
Chakraborty, J., 2020. Unequal exposure to environmental pollution: an intersectional analysis of people with disabilities in Harris County, Texas. Prof. Geogr. 72 (4), 521–534.
Chakraborty, J., Collins, T.W., Grinecki, S.E., 2019a. Exploring the environmental justice implications of hurricane harvey flooding in greater houston, Texas. Am. J. Publ. Health 109 (2), 244–250.
Chakraborty, J., Grinecki, S.E., Collins, T.W., 2019b. Hurricane harvey and people with disabilities: disproportionate exposure to flooding in houston, Texas. Soc. Sci. Med. 226, 176–181.
Clark, L.P., Millet, D.B., Marshall, J.D., 2014. National patterns in environmental injustice and inequality: outdoor NO2 air pollution in the United States. PLoS One 9 (4), e94482.
Clark, L.P., Millet, D.B., Marshall, J.D., 2017. Changes in transportation-related air pollution exposures by race-ethnicity and socioeconomic status: outdoor nitrogen dioxide in the United States in 2000 and 2010. Environ. Health Perspect. https://doi.org/10.1289/EHP1509.
Cole, M.A., Ozgen, C., Strohl, E., 2020. Air Pollution Exposure and COVID-19. In: Discussion Paper Series-IZA DP No. 13367. IZA.
Cole, M.A., Ozgen, C., Strohl, E., 2020. Air Pollution Exposure and COVID-19. In: Discussion Paper Series-IZA DP No. 13367. IZA – Institute of Labor Economics, Bonn, Germany.
Collins, T.W., Grinecki, S.E., Morales, D., 2017. Environmental injustice and sexual minority health disparities: a national study of inequitable health risks from air pollution among same-sex partners. Soc. Sci. Med. 191, 38–47.
Conticini, E., Frediani, B., Caro, D., 2020. Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy? Environ. Pollut. 114465 https://doi.org/10.1016/j.envpol.2020.114465.
Devakumar, D., Shannon, G., Bhupal, S.S., Abubakar, I., 2020. Racism and discrimination in COVID-19 responses. Lancet 395 (10231), 1194.

Eligon, J., Burch, A.D.S., Searcey, D., Oppel, R.A., 2020. Black Americans face alarming rates of coronavirus infection in some states. N. Y. Times. April 14. www.nytimes.com/2020/04/07/us/coronavirusrace.html.

Finch, W.H., Finch, H.M., 2020. Poverty and Covid-19: rates of incidence and deaths in the United States during the first 10 weeks of the pandemic. Front Sociol. 5, 47.

Garson, G., 2012. Generalized Linear Models and Generalized Estimating Equations. Statistical Associates Publishing, Asheboro, NC.

Grineski, S.E., Collins, T.W., Chakraborty, J., McDonald, Y., 2013. Environmental health injustices: exposure to air toxics and children’s respiratory hospital admissions in El Paso, Texas. Prof. Geogr. 65 (1), 31–46.

Hackbarth, A., Romley, J., Goldman, D., 2011. Racial and ethnic disparities in hospital care resulting from air pollution in excess of federal standards. Soc. Sci. Med. 73 (8), 1163-1168.

Hendryx, M., Luo, J., 2020b. COVID-19 prevalence and mortality rates in association with Black race and segregation in the United States April 1 to April 15, 2020. The Lancet Public Health. https://doi.org/10.21293/srn.3582857.

Hendryx, M., Luo, J., 2020a. COVID-19 prevalence and fatality rates in association with air pollution emission concentrations and emission sources. Environ. Pollut. 265, 115126. https://doi.org/10.21293/srn.3582857.

Hendryx, M., Luo, J., 2020b. COVID-19 prevalence and mortality rates in association with Black race and segregation in the United States April 1 to April 15, 2020. The Lancet Public Health. https://doi.org/10.21293/srn.3582857.

Holifield, R., Chakraborty, J., Walker, G. (Eds.), The Routledge Handbook of Environmental Justice. Routledge, New York, pp. 1–8.

JHU (Johns Hopkins University), 2020. Center for Systems Science and Engineering. Johns Hopkins University. https://www.arcgis.com/home/user.html?user=CSSEGISData.

Karaye, I.M., Horney, J., 2020. The impact of social vulnerability on COVID-19 in the U. S.: an analysis of spatially varying relationships. Am. J. Prev. Med. https://www.ajpmonline.org/article/S0749-3797(20)30259-2/fulltext.

King, J.S., 2020. Covid-19 and the need for health care reform. N. Engl. J. Med. 382, e104. https://doi.org/10.1056/NEJMp2000821.

Li, M., Zhang, D., 2020. A spatial analysis of New York City and Chicago hot spots. J. Urban Health. https://doi.org/10.3390/ijerph13080775.

Maldonado, A., Collins, T.W., Grineski, S.E., Chakraborty, J., 2016. Exposure to flood hazards in Miami and Houston: are Hispanic immigrants at greater risk than other social groups? Int. J. Environ. Res. Publ. Health 13 (8), 775. https://doi.org/10.3390/ijerph13080775.

Maroko, A.R., Nash, D., Pavilonis, B.T., 2020. COVID-19 and inequity: a comparative spatial analysis of New York City and Chicago hot spots. J. Urban Health. https://doi.org/10.1007/s11524-020-00468-0.

McCarthy, M.C., O’Brien, T.E., Charriere, J.E., Hafner, H.R., 2009. Characterization of the chronic risk and hazard of hazardous air pollutants in the United States using ambient monitoring data. Environ. Health Perspect. 117, 790–796.

Morello-Frosch, R., Jesdale, B.M., 2006. Separate and unequal: residential segregation and estimated cancer risks associated with ambient air toxics in US metropolitan areas. Environ. Health Perspect. 114 (3), 386e393.

Nelder, J., Wedderburn, R., 1972. Generalized linear models. J Royal Stat Soc, Series A 135, 370–384.

Oge, Y., 2020. Assessing nitrogen dioxide (NO2) levels as a contributing factor to coronavirus (COVID-19) fatality. Sci. Total Environ. 726, 138605.

Ray, R., 2020. Why are Blacks dying at higher rates from COVID-19?Brookings. http://www.brookings.edu/blog/bfix/2020/04/09/why-are-blacks-dying-at-higher-rates-from-covid-19/.

Sandola, A., 2020. The Impact of the COVID-19 Pandemic on Older Adults. Medical News Today. https://www.medicalnewstoday.com/articles/the-impact-of-the-covid-19-pandemic-on-older-adults.

Trinity Consultants, 2019. Latest NATA released: its role and implications. https://www.trinityconsultants.com/news/federal/latest-nata-released—its-role-and-implications.

Turk, M.A., McDermott, S., 2020. The Covid-19 pandemic and people with disability. Disability Health J. 13 https://doi.org/10.1016/j.dhjo.2020.100944.

USEPA (Environmental Protection Agency), 2020a. Hazardous air pollutants. https://www.epa.gov/hap.

USEPA, 2020b. National air toxics assessment. https://www.epa.gov/national-air-toxics-assessment.

USEPA, 2020c. 2014 NATA: technical support document. https://www.epa.gov/national-air-toxics-assessment/2014-nata-technical-support-document.

USEPA, 2020d. 2014 NATA: assessment results. https://www.epa.gov/national-air-toxics-assessment/2014-nata-assessment-results.

Wilson, S., Burwell-Nane, K., Jiang, C., Zhang, H., Samantapudi, A., Murray, R., Dalemarre, L., Rice, L., Williams, E., 2015. Assessment of sociodemographic and geographic disparities in cancer risk from air toxics in South Carolina. Environ. Res. 140, 562–568.

Wu, X., Nethery, R.C., Sabath, B., Braun, D., Dominici, F., 2020. Exposure to Air Pollution and COVID-19 Mortality in the United States. medRxiv. https://doi.org/10.1101/2020.04.05.20054502.

Yance, C.W., 2020. COVID-19 and african Americans. J. Am. Med. Assoc. 323 (19), 1891–1892.

Zhu, Y., Xie, J., Huang, F., Cao, L., 2020. Association between short-term exposure to air pollution and COVID-19 infection: evidence from China. Sci. Total Environ. 727, 138704.