Spatial and Racial Group Disparities of COVID-19 Mortality: Exploring Dimensions of Structural Barriers

In the United States, the COVID-19 pandemic has disproportionately affected racial minority groups and other marginalized communities. Existing literature shows that social determinants of health (SDOH), the structural conditions that characterize where people live, work, and play, play a significant role in driving disparities in COVID-19 infection and mortality rates. However, quantifying such impact is challenging, as the Modifiable Areal Unit Problem in geography highlights issues of different findings produced by different levels of aggregation, while the ecological fallacy also exposes the challenges posed by aggregating individuals to a group. Additionally, state-level estimates can mask heterogeneous population composition and distribution, making community-level impacts difficult to evaluate. County-level data provide more fine-grained information, but race- or age- disaggregated COVID-19 data are essentially unavailable nationally.

This paper seeks to explore and uncover how the structural factors associated with racism influence racial group disparities in COVID-19 related outcomes at the county level at the national scale (n = 3142), given the data challenge. We drew on county level COVID-19 infection, mortality and vaccine data for the first two year of the pandemic in the United State reported by the Centers for Disease Control and Prevention. Sourcing data on racial/demographic using the American Community Surveys 2019 5-year estimates, we mainly focused on three groups, namely Hispanic or Latino, Black or African American alone, and white alone, not Hispanic or Latino. We used four indices from Kolak et al (2020) as measures for SDOH as they represent the most important dimensions extracted from an adapted social ecological model of health that spans over individual, interpersonal, organizational, and community spheres. Additionally, our analysis explored age distribution, rural-urban context, and multiple specific community indicators to provide more context information to help further understand and interpret the trends reflected by the SDOH indexes, as well as capturing potential uniqueness of COVID-19. For example, these community indicators include income inequality, uninsured rate, primary care physicians, preventable hospital stays, severe housing problems, and access to the Internet.

Importantly, we consider different outcomes related to COVID-19 as they reflect varied aspects of people's vulnerability and experiences during the pandemic. First, we examine the trends of both confirmed case and mortality rates and explore how various SDOH measures may be associated with different COVID outcomes differently. For example, we find the delay in mask mandate mostly relates to the infection rate but not necessarily the mortality rate. Second, we investigate and compare the patterns between the first and the second year of the pandemic, considering the fact that vaccinations have become available and played an important role in the second year.

Adopting an exploratory spatial data analysis approach, we started with co-location analysis to identify counties that are hit hardest by COVID-19, particularly areas where each racial/ethnic

group has been consistently experiencing high COVID-19 infection or mortality rates. Specifically, if a county (a) experienced COVID-19 infection or mortality rates in the top quintile for 100 days or more in the first or second year of the pandemic, and (b) is identified in the top quintile of populations of a particular racial/ethinic group, then this county is labeled as “concentrated longitudinal impact counties'' (CLI counties). As such, we identified CLI counties for each racial/ethnic group during each year of the pandemic. Stratifying COVID-19 outcomes and measures of regional structural factors by these CLI counties, we explored potential community characteristics that are statistically associated with COVID-19 disparities, based on which we hypothesized that distinct trends seem to emerge across different locales: specifically urban, rural, and suburban areas. To account for this extreme spatial heterogeneity, we modeled each place structure as a different spatial regime, with the goal of reflecting different processes through which structural factors are associated with COVID-19 outcomes. We deployed a second-order queen contiguity spatial weight for neighbor specification because first-order spatial weight still had unexplained spatial correlation in models. First, we applied models with ordinary least square estimation that do not consider spatial autocorrelation to monitor the collinearity among all predictors. Then we fit the data with a spatial autoregressive model to account for the spatial interaction effect of COVID-19 as an infectious disease using GMM estimation. Specifically, we applied the Spatial Error Model (SEM) to capture spatial spillover in our outcome of COVID-19 mortality rates caused by omitted variables.

Take the mortality rate in the first year of the pandemic as an example. Among all 3142 counties, we identified 347 CLI counties for the Black population, 198 CLI counties for the Hispanic population, and 33 CLI counties for the white, non-Hispanic population. Though all the CLI counties are disproportionately hit by COVID-19, the impact on their populations was not the same, as reflected by the demographic trends, SDOH indexes, regional structural factors measures, as well as distinct spatial trends. For example, Black CLI counties are located across the Southeastern continental US, representing a range of urban and rural counties across the Black Belt region with additional counties in metropolitan areas across the country. These counties feature a high percentage of working age people, high income inequality, high percentage of uninsured, large values of preventable hospital stays and severe housing problems. On the contrary, white non-Hispanic CLI counties are found in mostly rural areas across Appalachia and the greater Midwest and upper Great Plains region, featuring high percentage of older population, and high ratio of population to primary care physicians. The spatial regime regression results also provide statistical evidence for spatial heterogeneity across rural, suburban, and urban areas. Higher COVID-19 mortality rates are significantly associated with higher urban core opportunity and more immigrant populations with traditional family structures, multiple accessibility stressors and overcrowding housing problems in urban areas, lower socioeconomic status in suburban areas, and higher ratio of population to primary care physicians and higher proportion of younger populations and fewer persons with disabilities in rural areas. Additionally, worse access to the internet is significantly related to mortality rates in all regimes.

In the end, it is crucial to use social and structural factors to understand how various SDOH factors impact different minoritized groups differently and understanding inequalities between racial and ethnic groups is important for determining and implementing place-based interventions.