

Kevin D. Long, MD¹;
Steven M. Albert, PhD, MS²

Objective: In the first six months of the pandemic, information on race and ethnicity was missing for half of the US COVID-19 cases. Combining case ascertainment with census-based zip code indicators may identify COVID-19 race-ethnicity disparities in the absence of individual-level data.

Design: Ecological retrospective study for the period March-July 2020.

Setting: Population-based investigation, Allegheny County, Pennsylvania.

Participants: All COVID-19 cases, adjusted for zip code area population, in the early period of the pandemic.

Main Outcome Measures: Monthly COVID-19 incidence and requests for human services by zip code level indicators of race-ethnicity and poverty.

Results: In the early period of the pandemic, COVID-19 incidence was higher in zip codes with a greater proportion of racial and ethnic minorities. Zip codes with the highest quartile of minority residents (>25.1% of population) had a COVID-19 incidence of 60.1 (95% CI: 51.7-68.5) per 10,000 in this period; zip codes with the lowest quartile of minority residents (<6.3%) had an incidence of 31.3 (95% CI: 14.4-48.2). Requests for human services during this period (volume of 211 calls and county services) confirm these disparities.

Conclusion: Use of census-defined race-ethnicity proportions by zip code offers a way to identify disparities when individual race-ethnicity data are unavailable. *Ethn Dis.* 2021;31(3):399-406; doi:10.18865/ed.31.3.399

INTRODUCTION

The United States is emerging from an unprecedented public health crisis with the novel COVID-19 pandemic. With more than 30 million cases in the United States, public health efforts need to focus on prevention and reducing disparities in risk and outcomes.¹ Research shows disparities in COVID-19 testing, infections, and deaths, as well as in social distancing and access to health care, and most recently vaccination. COVID-19 disproportionately affects individuals with lower socioeconomic status, immigrants, and racial and ethnic minority groups.²⁻⁶ Additionally, individuals with chronic diseases and people who are aged >65 years face greater risks of hos-

pitalization, ICU care, and death when infected with COVID-19.⁷

In the beginning months of the pandemic, as many as 50% of cases were missing information on race and ethnicity. The CDC COVID-19 demographic tracker first reported data on Aug 28, 2020; 51% of about 4.5 million cases lacked data on race and ethnicity.⁸ Even as late as mid-January 2021, race-ethnicity data were missing for 48% of the first 12.9 million vaccinees (in contrast to .04% missing age and 3.02% gender).^{9,10}

In order to better understand COVID-19-related disparities at the beginning of the pandemic when individual-based data on race were limited, we examined COVID-19 related outcomes and potential disparities at the zip code level, focus-

Keywords: COVID-19; Poverty; Zip Codes; Minority Groups; Ethnic Groups; Racial Disparities

¹ University of Pittsburgh, Graduate School of Public Health, Pittsburgh, PA

² University of Pittsburgh, Behavioral and Community Health Sciences, Pittsburgh, PA

Address correspondence to Steven M. Albert, PhD, MS; Department of Behavioral and Community Health Sciences, University of Pittsburgh, Public Health 6128, 130 DeSoto Street, Pittsburgh, PA 15261; smalbert@pitt.edu

ing on Allegheny County, Pennsylvania. Allegheny County was similar in underreporting of race data in the early months of the pandemic, but by the end of 2020 was able to provide race data on all but 15% of 781,708 tests reported.¹¹

Disparities in rates of infection among African American, Native American, and Latinx communities are well-documented.² Some of the reasons behind these dispari-

In this research, we examined COVID-19 disparities across zip codes in Allegheny County, Pennsylvania in the initial period of the pandemic (February-July 2020).

ties include occupational exposures, greater prevalence of chronic medical conditions, and more barriers to health care access.² Additionally, counties differ in level of social distancing as indicated by reductions in travel and rates of unemployment.³ In addition, household density, higher in low-SES neighborhoods, is associated with higher rates of COVID-19 infection.⁴

At the beginning of the pandemic, most research examined dispari-

ties in COVID-19 risk across counties. Few looked at disparities within counties using smaller geographical areas. One exception included studies in New York City that used zip codes to determine disparities in testing. These studies indicated an increased volume of testing in zip codes with a higher proportion of White residents,⁵ and a lower volume in poorer neighborhoods and neighborhoods with more immigrants.⁶ Additionally, test positivity decreased with an increasing proportion of White individuals and also with higher socioeconomic index score.⁵ Zip code based analyses from two counties in Washington State found that vulnerability index score, chronic disease burden, greater proportion of racial and ethnic minorities, and lower socioeconomic status were associated with greater COVID-19 incidence.⁷

In this research, we examined COVID-19 disparities across zip codes in Allegheny County, Pennsylvania in the initial period of the pandemic (February-July 2020). We examined incidence as well as services provided to communities. We reasoned that in the absence of individual-based data, using zip code level detail can provide insight into potential disparities, allowing quicker public health responses to intervene on disparities before they broaden.

Methods

Data were collected from the Allegheny County Health Department (ACHD), Allegheny County Department of Human Services (DHS), the 2018 American Community

Survey 5-Year Estimates Data Profiles from the US Census Bureau, and national 211 services website. Data were retrieved for the period February 1 through July 22, 2020.

To avoid potential breaches of confidentiality in small areas, ACHD provided a count of COVID-19 cases by month for each zip code. We also accessed data from the 211 helpline, maintained by local United Way agencies across the United States; this helpline is a national surveillance and referral system for social needs, including housing, food, utility bills, transportation, child care, and health care. All and COVID-19 specific 211 calls were retrieved by zip code from the PA 211 site (<https://pa.211counts.org>). The volume of DHS social services by zip code was also available. We obtained demographic information by zip code from the US Census Bureau's website, including zip code specific profiles from the 2018 American Community Survey 5-Year Estimates.

For analyses, zip codes were grouped by demographic features. Zip codes were grouped into quartiles based on distributions of the proportion of individuals that belong to a racial or ethnic minority group (0-6.3%, 6.32-11.7%, 12.1-24.9%, 25.1-76.6%), percent of Hispanic individuals (0-.85%, .89-1.5%, 1.51-2.34%, 2.4-8.8%), percent of African Americans (0-1.7%, 1.72-4.9%, 5.2-13.7%, 14.1-68.0%), and poverty rate, as defined by the Census (0-6.9%, 7.5-11.3%, 11.3-16.8%, 16.9-53.6%). Monthly COVID incidence and 211 and DHS volume were standardized and

converted to rates per 10,000 using zip code population denominators.

ACHD provided information on 120 Allegheny County zip codes (15006-15275). Three zip codes had no inhabitants and were excluded from analysis. Of the remaining 117, the Allegheny County Health Department did not report cases for 19 zip codes. When we inquired, we were told data were unavailable. Our strategy was to drop these zip codes rather than impute a value of 0 cases. Accordingly, our final sample was 97 zip codes, and our analyses include 485 zip code months.

Analyses

Data were imported to Stata 14[®] (StataCorp, TX, USA) for analysis. Quantitative variables were described using means and standard deviations, categorical variables by frequencies and percentages. Profile plots were used to describe monthly COVID-19 incidence and COVID-specific service requests by quartiles of poverty and racial and ethnic minority status. Repeated-measures analysis of variance (ANOVA) and mixed effects models for zip code quartiles were estimated to determine the association between quartile, COVID-19 incidence, and time. We also estimated multivariable linear regression models to look at associations between COVID-19 incidence and zip code quartile adjusting for sociodemographic features. Because only group count data were assessed, the University of Pittsburgh Human Research Protections Office determined the research did not meet criteria for human subject research.

Table 1. Demographics and COVID-19 cases for Allegheny County, Pennsylvania

Demographics ^a	
Population	1,225,561
Female, %	51.7
Age (median years)	40.8 ± 0.1
Poverty, %	12.1%
Race	
White, %	80.1
Black or African American, %	12.9
American Indian and Alaska Native, %	.1
Asian alone, %	3.7
Native Hawaiian and Other Pacific Islander, %	0.0
Other, %	.5
Two or more races, %	2.7
Hispanic or Latino, %	2.0
Language other than English Spoken at home, %	7.3
Uninsured, %	4.3
Education (HS graduate or higher), %	94.3
COVID-19 Specific, March-July 22, 2020 ^b	
Individuals tested	144,081
Total cases	9,932
Deaths	306

a. Population information from <https://data.census.gov/cedsci/profile?g=0500000US42003>.
b. COVID estimates from Allegheny County COVID-19 Individual Test Results, 2020. Western PA Regional Data Center.¹¹

RESULTS

Allegheny County has a population of more than 1.2 million individuals, with a greater proportion of females (51.7%) than the national average (50.8%).^{12,13} The population has a mean poverty rate of 12.1%, average age of 40.8 years, and is mostly non-Hispanic White (80.1%). Through the middle of July, there were more than 144,000 tests, nearly 10,000 cases of COVID-19, and more than 300 deaths. Table 1 provides an overview of the demographics and

COVID-19 burden of the county.

Total COVID incidence, volume of 211 calls (total and COVID-specific), and county social service use over the initial pandemic period are shown in Table 2, broken out by quartile of minority population across the 97 zip codes. COVID incidence ranged from 31.3 per 10,000 (95%CI: 14.4-48.2) in zip codes with the lowest proportion of individuals belonging to racial or ethnic minority groups (0-6.3%), to 60.1 (95%CI: 51.7-68.5) in zip codes with the highest proportion (25.1-76.6%). Similar differences

COVID-19 Race Disparities by Zip Code - Long and Albert

Table 2. Total DHS services, 211 calls, and COVID-19 cases by zip code areas: quartiles defined by proportion of minority residents

	Proportion of Minority Residents in Zip Codes ^a			
	Quartile 1, 0-6.3%	Quartile 2, 6.32-11.7%	Quartile 3, 12.1-24.9%	Quartile 4, 25.1-76.6%
	n (95%CI)	n (95%CI)	n (95%CI)	n (95%CI)
DHS services ^b	23,766 (16,070-31,463)	18,402 (14,358-22,447)	26,042 (18,845-33,240)	48,098 (39,982-56,214)
211 calls ^c	143.6 (37.6-249.7)	103.8 (76.6-131.1)	224.2 (98.5-349.9)	391.1 (324.6-457.5)
COVID-specific 211 calls ^c	30.1 (12.6-47.6)	29.4 (21.2-37.6)	76.5 (4.0-148.9)	81.7 (69.8-93.5)
COVID-19 cases ^c	31.3 (14.4-48.2)	33.0 (25.8-40.1)	52.5 (39.5-65.4)	60.1 (51.7-68.5)

a. Zip code quartiles based on proportion of individuals belonging to racial or ethnic minority groups in those zip codes.

b. February 1, 2020- June 30, 2020.

c. March 1, 2020- July 22, 2020.

P<.001 by one way ANOVA except for COVID-specific 211 calls (P=.13).

were evident for the volume of 211 calls and COVID-19-specific 211 calls, as well as county-provided social services. Likewise, in comparing

the zip code quartiles with the highest and lowest poverty rates, those with a higher poverty rate also had a higher incidence of COVID-19

cases, 60.3 per 10,000 vs 40.2, as well as a greater volume of 211 calls and county social DHS services (results available upon request).

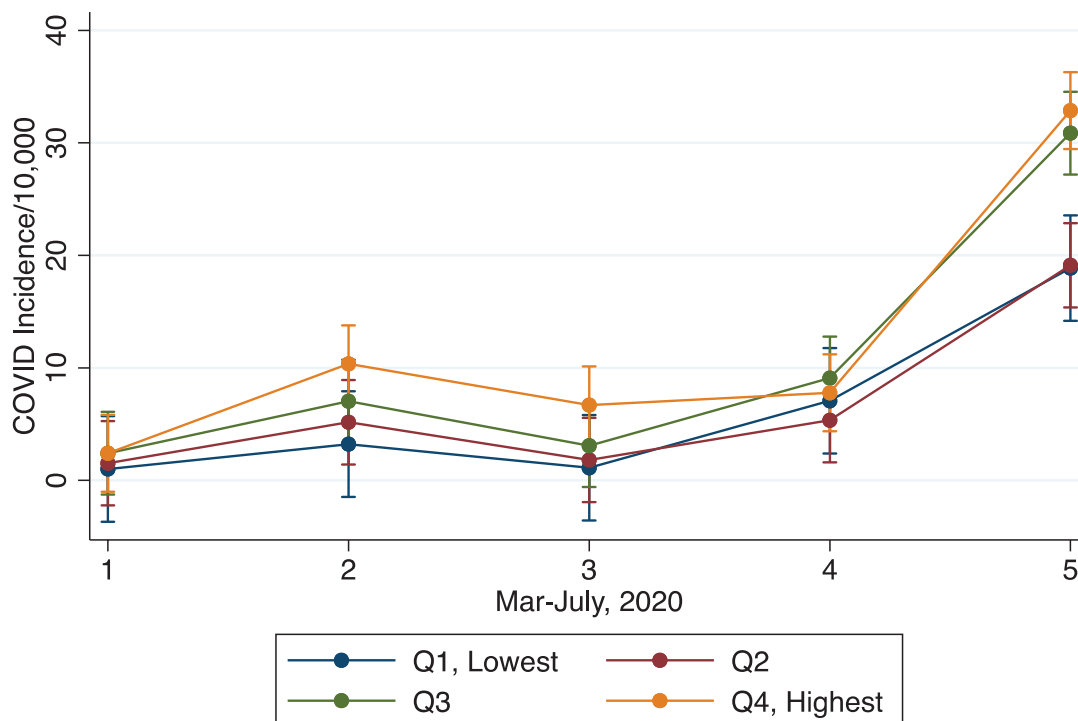


Figure 1. COVID incidence/10,000 by zip code minority quartiles

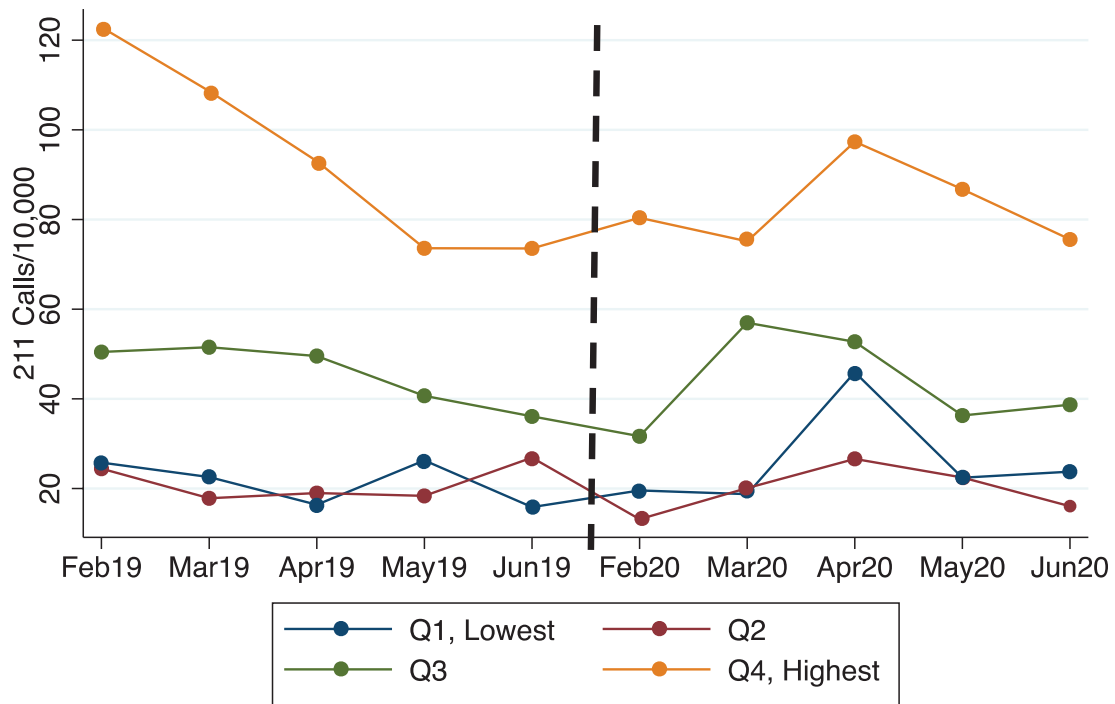


Figure 2. 211 calls, 2019 and 2020, by zip code minority quartiles

Figure 1 shows COVID-19 cases in each month (March-July 2020) per 10,000 individuals in the zip code quartiles. The plot shows mean incidence and 95%CI for each of the minority population quartiles per month. Zip codes with a greater proportion of minorities had higher COVID-19 incidence over the 5 months ($F=6.7$, $P<.0001$), the time trend for cases increased over the five months ($F=108.6$, $P<.0001$), and a test of the interaction showed a greater increase over time in quartiles with a higher proportion of minorities ($F=2.59$, $P<.003$). Results from a mixed model were similar.

To assess the independent effect of the proportion of individuals belonging to racial or ethnic

minority groups on COVID incidence, we estimated a multivariable linear regression model. The model showed that a higher percentage of racial and ethnic minorities in zip codes was associated with greater COVID-19 incidence per 10,000 ($\beta=.281$, $P=.036$) adjusting for age, sex, and poverty rate.

Overall, 211 requests for utilities and food needs increased as COVID-19 became more prominent in March and April, with a decrease in May and June. COVID-19-specific housing and shelter needs were not prominent until July. Rates peaked in April, with COVID-19-specific calls increasing again in July. 211 requests for housing services, food, employment

and income, and disaster relief were higher in zip codes with the highest proportion of individuals belonging to racial or ethnic minority groups.

To assess trends in service use, we examined the volume of 211 calls by quartile of individuals belonging to racial or ethnic minority groups over comparable months in 2019 and 2020. Results are shown in Figure 2. Zip codes with the highest proportion of minorities utilized the 211 helpline to a greater extent in 2019 as well as 2020. Across 2019-2020, we see that quartile 1, with the lowest proportion of minority population, was mostly unaffected by the pandemic; the volume of 211 calls across the two periods did not change. Quartile 4, with the

highest proportion of minorities, by contrast, had a higher volume of calls across both periods, with a substantial uptick in April 2020, the height of the first wave of the pandemic. Its use of 211 calls was twice as high as the next quartile.

DISCUSSION

This retrospective analysis highlights disparities in COVID-19 risk and the need for COVID-related

Our study shows that rates of COVID-19 were higher in zip codes with a greater prevalence of individuals belonging to racial and ethnic minorities.

services at the beginning of the pandemic in Allegheny County. Given the substantial proportion of missing individual-level racial and ethnic data in COVID testing, we show an alternative approach using zip codes to analyze potential disparities. Our study shows that rates of COVID-19 were higher in zip codes with a greater prevalence of individuals belonging to

racial and ethnic minorities. Missing race-ethnicity data is a problem for COVID testing; for example, CDC's national COVID Data Tracker included race-ethnicity for only 53% of cases by mid-September (<https://covid.cdc.gov/covid-data-tracker/#demographics>) when we completed this analysis. Lack of complete COVID-19 data by race and ethnicity may minimize COVID-19 disparities.¹⁴ Use of Census-defined race-ethnicity proportions by zip code offers a way to estimate disparity effects when individual race-ethnicity data are unavailable.

Additionally, zip codes with higher proportions of racial and ethnic minorities and greater poverty sought social services at higher rates, evident in 211 calls. These needs were primarily for health care, employment and income, food, and later housing and shelter. We showed as well that county social services were also more concentrated in zip codes with higher rates of poverty and a greater proportion of racial and ethnic minorities. Additional research is required to see if receiving these resources reduces COVID-19 incidence and social impact.

One major finding from our study is that the rates of COVID-19 cases were rising at quicker rates for areas with more racial and ethnic minorities. These easily accessible data can be used to predict areas most in need of services to target outreach to reduce disparities. This finding is consistent with studies showing dynamic changes in the impact of COVID-19 by so-

cial vulnerability across counties.¹⁵

Minority zip code quartiles in Allegheny County mostly involved Black residents. For example, in the quartile with the highest proportion of minorities, 26 of these 30 zip codes also had the highest proportion of Black residents, ranging from 20%-68% for each zip code population, consistent with the substantial segregation of the county. The low proportion of Hispanic residents (no more than 8.8% of residents in any zip code) prevents separate analyses by race and ethnicity. However, other analyses have shown higher rates of COVID among Hispanics.⁴ Like Blacks, Hispanic individuals are also more likely to have essential jobs with fewer opportunities to work from home.¹⁶ Minorities face lower socioeconomic status, more overcrowding, more racial segregation, more food insecurity, and less access to health insurance, all of which are associated with less opportunity for social distancing.¹⁷ Potential policy changes to reduce the impact of the pandemic on minorities in the United States would include expanding access to Medicaid and Children's Health Insurance Program (CHIP), increasing access to employer-sponsored insurance, and reducing language barriers.

Given limitations of data collection on race and ethnicity, aggregate zip code level data can help correct potential underestimates of disparities. Studies have shown that minorities and individuals living in poverty have higher rates of hospitalization and other worse COVID-19 outcomes, showing that findings using aggregate data are supported by studies based on individual-level data.¹⁸⁻²⁰

While the ecological fallacy is always a concern when using aggregate indicators, this convergence of findings provides reassurance that area level rates can fill in gaps when individual level data are limited. A limitation of the study is our use of census data, which may underrepresent non-citizens or undocumented citizens, as well as restriction to a single geographic area and limited time interval.

CONCLUSIONS

This study offers essential information on disparities that were widening for COVID-19 cases and pandemic-related social needs at the beginning of the pandemic. We show that the incidence of COVID-19 increased earlier and faster for areas with a higher proportion of individuals belonging to ethnic and racial minorities. Health departments can rely on area indicators, such as zip code of residence, to identify these inequities even with incomplete ascertainment of race among cases. Areas with a higher proportion of racial and ethnic minorities requested more services even when controlling for age, sex, and poverty status. Providing services and public health strategies to mitigate disparities will be critical as the United States and the rest of the world attempt to control the pandemic. Aggregate level data should be considered in the absence of quality individual level data.

ACKNOWLEDGMENTS

The authors thank the Allegheny County Health Department for data access.

CONFLICT OF INTEREST

No conflicts of interest to report.

AUTHOR CONTRIBUTIONS

Research concept and design: Long, Albert; Acquisition of data: Long, Albert; Data analysis and interpretation: Long, Albert; Manuscript draft: Long, Albert; Statistical expertise: Long, Albert; Acquisition of funding: Albert; Administrative: Long, Albert; Supervision: Albert

REFERENCES

- Centers for Disease Control and Prevention; U.S. Department of Health & Human Services. *Coronavirus Disease 2019 (COVID-19): United States COVID-19 Cases and Deaths by State*. Last accessed February 12, 2021 from https://covid.cdc.gov/covid-data-tracker/#cases_totalcases
- Tai DBG, Shah A, Doubeni CA, Sia IG, Wieland ML. The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States. *Clin Infect Dis*. 2021;72(3):703-706. <https://doi.org/10.1093/cid/ciaa815> PMID:32562416
- Chang S, Pierson E, Koh PW, et al. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*. 2021;589(7840):82-87 <https://doi.org/10.1038/s41586-020-2923-3>
- Rodriguez-Diaz CE, Guilamo-Ramos V, Mena L, et al. Risk for COVID-19 infection and death among Latinos in the United States: examining heterogeneity in transmission dynamics. *Ann Epidemiol*. 2020;52:46-53.e2. <https://doi.org/10.1016/j.annepidem.2020.07.007> PMID:32711053
- Lieberman-Cribbin W, Tuminello S, Flores RM, Taioli E. Disparities in COVID-19 testing and positivity in New York City. *Am J Prev Med*. 2020;59(3):326-332. <https://doi.org/10.1016/j.amepre.2020.06.005> PMID:32703702
- Borjas GJ. *Demographic Determinants of Testing Incidence and COVID-19 Infections in New York City Neighborhoods*. April 2020. National Bureau of Economic Research Working Paper 26952; <https://www.nber.org/papers/w26952>; <https://doi.org/10.3386/w26952>
- Amram O, Amiri S, Lutz RB, Rajan B, Monsivais P. Development of a vulnerability index for diagnosis with the novel coronavirus, COVID-19, in Washington State, USA. *Health Place*. 2020; Epub 2020 Jun 26. <https://doi.org/10.1016/j.healthplace.2020.102377>.
- Krieger N, Testa C, Hanage WP, Chen JT. US racial and ethnic data for COVID-19 cases: still missing in action. *Lancet*. 2020;396(10261):e81. [https://doi.org/10.1016/S0140-6736\(20\)32220-0](https://doi.org/10.1016/S0140-6736(20)32220-0) PMID:33169681
- Krieger N, Waterman PD, Chen JT, Testa C, Hanage WP. Missing again: US racial and ethnic data for COVID-19 vaccination. *Lancet*. 2021;397(10281):1259-1260. [https://doi.org/10.1016/S0140-6736\(21\)00465-7](https://doi.org/10.1016/S0140-6736(21)00465-7) PMID:33743218
- Painter EM, Ussery EN, Patel A, et al. Demographic characteristics of persons vaccinated during the first month of the COVID-19 vaccination program—United States, December 14, 2020–January 14, 2021. *Morb Mortal Wkly Rep*. 2021;70(5):174-177. <https://doi.org/10.15585/mmwr.mm7005e1> PMID:33539333
- Allegheny County COVID-19 Individual Test Results, 2020. Western PA Regional Data Center. Last accessed June 1, 2021 from <https://data.wprdc.org/dataset/allegheny-county-covid-19-tests-cases-and-deaths>
- United States Census Bureau. ACS Demographic and Housing Estimates. Last accessed September 30, 2020 from <https://data.census.gov/cedsci/table?q=us&tid=ACSDP1Y2019.DP05&hidePreview=false>
- United States Census Bureau. Allegheny County, Pennsylvania. Last accessed September 20, 2020 from <https://data.census.gov/cedsci/profile?g=0500000US42003>
- Gross CP, Essien UR, Pasha S, Gross JR, Wang SY, Nunez-Smith M. Racial and ethnic disparities in population-level COVID-19 mortality. *J Gen Intern Med*. 2020;35(10):3097-3099. <https://doi.org/10.1007/s11606-020-06081-w> PMID:32754782
- Neelon B, Mutiso F, Mueller NT, Pearce JL, Benjamin-Neelon SE. Spatial and temporal trends in social vulnerability and COVID-19 incidence and death rates in the United States. *PLoS ONE*. 2021; 16(3): e0248702. <https://doi.org/10.1371/journal.pone.0248702>
- Selden TM, Berdahl TA. COVID-19 and racial/ethnic disparities in health risk, employment, and household composition: study examines potential explanations for racial-ethnic disparities in COVID-19 hospitalizations and mortality. *Health Aff (Millwood)*. 2020;39(9):1624-1632. <https://doi.org/10.1377/hlthaff.2020.00897>
- Kavanagh NM, Goel RR, Venkataramani AS. Association of county-level socioeconomic and political characteristics with engagement in social distancing for COVID-19. *Am J Prev Med*. 2021;000(000):1-7.
- Azar KMJ, Shen Z, Romanelli RJ, et al. Disparities in outcomes among COVID-19 patients in a large health care

COVID-19 Race Disparities by Zip Code - Long and Albert

- system in California. *Health Aff (Millwood)*. 2020;39(7):1253-1262. <https://doi.org/10.1377/hlthaff.2020.00598>
PMID:32437224
19. Muñoz-Price LS, Nattinger AB, Rivera F, et al. Racial disparities in incidence and outcomes among patients with COVID-19. *JAMA Netw Open*. 2020;3(9):e2021892. <https://doi.org/10.1001/jamanetworkopen.2020.21892> PMID:32975575
 20. Kim HN, Lan KF, Nkyekyer E, et al. Assessment of disparities in COVID-19 testing and infection across language groups in Seattle, Washington. *JAMA Netw Open*. 2020;3(9):e2021213. <https://doi.org/10.1001/jamanetworkopen.2020.21213>
PMID:32970156