HARVARD CENTER FOR
POPULATION AND
DEVELOPMENT STUDIES

## Working Paper Series

# Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county vs. ZIP code analyses

Jarvis T. Chen, ScD[1], Nancy Krieger, PhD[2]

April 21, 2020

HCPDS Working Paper Volume 19, Number 1

**Affiliations**

1. Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health

2. Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health

**Corresponding Author**

Jarvis T. Chen, ScD, Research Scientist, Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health
Email: jarvis@hsph.harvard.edu

**Abstract**

No national, state, or local public health monitoring data in the US currently exist regarding the unequal economic and social burden of COVID-19. To address this gap, we draw on methods of the Public Health Disparities Geocoding Project, whereby we merge county-level cumulative death counts with population counts and area-based socioeconomic measures (ABSMs: % below poverty, % crowding, and % population of color, and the Index of Concentration at the Extremes) and compute rates, rate differences, and rate ratios by category of county-level ABSMs. To illustrate the performance of the method at finer levels of geographic aggregation, we analyze data on (a) confirmed cases in Illinois ZIP codes and (b) positive test results in New York City ZIP codes with ZIP code level ABSMs. We detect stark gradients though complex gradients in COVID-19 deaths by county-level ABSMs, with dramatically increased risk of death observed among residents of the most disadvantaged counties. Monotonic socioeconomic gradients in Illinois confirmed cases and New York City positive tests by ZIP code level ABSMs were also observed. We recommend that public health departments use these straightforward cost-effective methods to report on social inequities in COVID-19 outcomes to provide an evidence base for policy and resource allocation.

Title: Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county vs. ZIP code analyses

Authors: Jarvis T. Chen, ScD
Nancy Krieger, PhD

Corresponding author: Jarvis T. Chen, ScD
Research Scientist
Department of Social and Behavioral Sciences
Harvard T.H. Chan School of Public Health
677 Huntington Avenue
Boston, MA 02115
Email: jarvis@hsph.harvard.edu

ABSTRACT

No national, state, or local public health monitoring data in the US currently exist regarding the unequal economic and social burden of COVID-19. To address this gap, we draw on methods of the Public Health Disparities Geocoding Project, whereby we merge county-level cumulative death counts with population counts and area-based socioeconomic measures (ABSMs: % below poverty, % crowding, and % population of color, and the Index of Concentration at the Extremes) and compute rates, rate differences, and rate ratios by category of county-level ABSMs. To illustrate the performance of the method at finer levels of geographic aggregation, we analyze data on (a) confirmed cases in Illinois ZIP codes and (b) positive test results in New York City ZIP codes with ZIP code level ABSMs. We detect stark gradients though complex gradients in COVID-19 deaths by county-level ABSMs, with dramatically increased risk of death observed among residents of the most disadvantaged counties. Monotonic socioeconomic gradients in Illinois confirmed cases and New York City positive tests by ZIP code level ABSMs were also observed. We recommend that public health departments use these straightforward cost-effective methods to report on social inequities in COVID-19 outcomes to provide an evidence base for policy and resource allocation.

Suggested citation: Chen JT, Krieger N. Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county vs ZIP code analyses. *Harvard Center for Population and Development Studies Working Paper Series*, Volume 19, Number 1. April 21, 2020. https://tinyurl.com/y7v72446

## INTRODUCTION

As communities in the United States (US) grapple with the COVID-19 pandemic, there is an urgent need for real-time data to better understand how particular populations are affected, including who is most at risk of infection, developing serious illness, and dying [1-2]. Informed by an awareness of the critical importance of racial/ethnic, economic, and gender inequalities in shaping individuals' exposure to and ability to protect themselves from SARS-CoV-2, as well as their ability to practice physical distancing, maintain economic wellbeing, and access appropriate healthcare when sick, there have been increasing calls for improved data to provide an evidence-base for action [1-4]. Descriptive epidemiology, which is vital to informing efforts to distribute resources, develop treatments, and coordinate public policy, is hampered by the paucity of disaggregated data by important social variables like race/ethnicity and socioeconomic position in the data reported by public health departments. For example, data from the COVID-19 tracking project [5] suggests that only ~21 states currently report COVID-19 cases or deaths disaggregated by race/ethnicity, and among those that do, substantial proportions (typically ≥50%) of cases and deaths are of unknown or missing race/ethnicity. Data tables on the US Centers for Disease Control and Prevention's own webpage reporting COVID-19 cases by race/ethnicity show upwards of 65% of reported cases with missing race/ethnicity information [6]. Furthermore, to our knowledge, no states are reporting COVID-19 cases or deaths by measures of individual socioeconomic position, though US death certificates routinely collect information on decedent's education [1-2, 7].

The Public Health Disparities Geocoding Project was established to address the absence of socioeconomic data in most routinely collected public health surveillance data [8-12]. By geocoding health records and linking them to US Census-derived data on neighborhood socioeconomic variables, we have shown that these methods can be used to compute valid estimates of socioeconomic gradients in health and, moreover, that area-based socioeconomic measures (ABSMs) can be used to characterize the influence of neighborhood socioeconomic context on health above and beyond their association with individual socioeconomic position. We have applied these techniques to a wide range of health outcomes, from birth to death and including cancer and infectious diseases, and have shown that the resulting estimates of socioeconomic gradients are valid and robust. The series of papers [8-12] stemming from this project have been cited over 3500 times and have had a demonstrable impact on US public health surveillance systems and health research more generally.

To respond to the urgent need in the United States for documentation of stark social inequities in who is affected by the COVID-19 pandemic, in this paper we quantify disparities in COVID-19 death rate in the US by county level sociodemographic attributes using currently available surveillance and US Census data. To illustrate the performance of these methods at finer levels of geographic aggregation, we additionally analyze data on (a) cumulative incidence of confirmed cases in Illinois ZIP codes and (b) cumulative incidence of positive test results in New York City ZIP codes with ZIP code level ABSMs. Our intention is to illustrate how state and local health departments can easily implement these types of analyses, using freely available US Census data, and provide tabular and graphic summaries of these social inequities to contribute to discussions on policies and interventions. In the discussion, we also discuss interpretation of these social inequities given limitations of the data and make recommendations for how public health departments can readily incorporate area-based socioeconomic measures into surveillance and monitoring.

## METHODS

### COVID-19 Data Sources

*US county death data:* We obtained publicly available data on COVID-19 deaths at the county level from the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) [13] and USA Facts [14]. Both sources report time series of cumulative confirmed cases and deaths, but notably, JHU CSSE reports a single entry for all of New York City, aggregating over the five counties corresponding to the city boroughs. Because this aggregation obscures substantial differences by boroughs (for example, death rates by borough were 128.3 per 100,000 in the Bronx, 108.1 per 100,000 in Brooklyn, 119.8 per 100,000 in Queens, 65.5 per 100,000 in Manhattan, and 87.1 per 100,000 in Staten Island), we used the USA Facts county dataset, which maintains separate reporting for New York counties. Differences were observed between JHU CSSE and USA Facts death counts on April 16, 2020 for 241 out of 2,717 matched counties, with discrepancies exceeding $\pm 10$ deaths for only 21 counties. Unmatched entries in the USA Facts datasets consisted of 421 counties with 0 deaths that did not appear in the JHU CSSE dataset, with the exception of a single death in Nantucket County, MA. Conversely, 56 unmatched entries in the JHU CSSE dataset consisted of 50 entries (298 deaths in 50 states) with "county unassigned", plus 2 entries for 152 deaths on cruise ships, and four entries for US territories (Guam, Northern Mariana Islands, Puerto Rico, and US Virgin Islands, 64 deaths). Our analytic sample consisted of 30,318 COVID-19 deaths reported in 3,144 US counties (excluding territories) as of April 16, 2020. We additionally present analyses of US COVID-19 cases as of April 16, 2020 by county characteristics in the Supplemental Appendix.

*Illinois data on confirmed cases at the zip code level*: We obtained ZIP code tabulation area (ZCTA) level data on confirmed cases in Illinois from the lookup tool developed by the Illinois Department of Public Health and the Chicago Reporter [15]. ZCTAs are US Census defined geographic units that correspond to areas roughly covered by US Postal Service (USPS) ZIP codes [16]. While there is not always a one-to-one correspondence between ZCTAs and USPS ZIP codes, the US Census ZCTAs provide a basis for linking sociodemographic and economic variables from the US Census American Community Survey to health records geocoded at the ZIP code level. As noted by the Illinois data source, infections among incarcerated populations are not fully represented in these data, including Cook County Jail (60608) and Stateville Correctional Center (60403), and possibly other ZIP codes. Illinois also reported data suppression for ZIP codes with <6 confirmed cases. Our analytic sample thus consisted of 24,675 confirmed cases reported in 372 Illinois ZCTAs as of April 16, 2020.

*New York City data on positive tests at the zip code level:* We obtained ZCTA-level data on positive tests in New York City from the New York City Department of Health and Mental Hygiene's COVID-19 GitHub repository [17]. Our analytic sample consisted of 125,422 positive tests reported in New York City 177 ZCTAs as of April 16, 2020.

**Population denominator and area attributes data**

We extracted county and ZCTA level population counts and sociodemographic attributes from the American Community Survey (ACS) 2014-2018 five-year estimates [18] using the tidycensus package in R [19]. ABSMs included: % of persons below poverty, % household crowding, and % population of color (defined as the proportion of population who are *not* White

Non-Hispanic), and a measure of racialized economic segregation, using the Index of Concentration at the Extremes [20]. This measure captures the extent to which the population in a given area is concentrated at either extreme of a social metric and ranges from -1 (everyone in the worst category) to 1 (everyone in the best category). For our analyses, we set the extremes for this ICE as: (a) high-income White population, versus (b) low-income Black population [20]. For analysis purposes, we defined categories of ABSMs using *a priori* cutpoints for % below poverty (0-4.9%, 5-9.9%, 10-14.9%, 15-19.9%, and 20-100%) and quintile cutpoints based on the distribution of county-level attributes in the US (county-level death analysis) or the distribution of ZCTA attributes within Illinois and New York City (ZCTA level analyses of confirmed cases and positive tests, respectively). Definitions, source variables from the ACS, and categorical cutpoints are presented in Table 1.

**Statistical Methods**

Drawing on the methods of the Public Health Disparities Geocoding Project [10], we merged cumulative counts of confirmed cases, positive tests, and deaths at the reported level of geography with population denominators and ABSMs. We then aggregated over areas within defined categories as described above. Since no data source currently reports disaggregated data by age and county or ZCTA, we computed crude outcome rates per 100,000 by ABSM categories rather than age-standardized rates. To quantify absolute and relative disparities, we computed rate differences and rate ratios setting the reference category to the socially most advantaged groups. We note that we use the term "death rate" in the county-level analysis to refer to cumulative deaths per 100,000 population (technically a cumulative incidence proportion); this quantity is distinct from the case fatality rate or infection fatality rate. Similarly, the rate of positive tests in the NYC ZCTA analysis is computed as the number of positive tests per 100,000 population (a cumulative incidence proportion) rather than positive tests as a proportion of all tests.

## RESULTS

### County level COVID-19 death in the US

As shown in Figures 1a-1d and Table 2, the highest COVID-19 death rates were consistently observed among those living in the most disadvantaged versus most advantaged counties in relation to: % poverty (19.3 per 100,000 vs. 9.9 per 100,000); the Index of Concentration at the Extremes for racialized economic segregation (15.0 per 100,000 vs. 13.8 per 100,000); % crowding (16.8 per 100,000 vs. 4.9 per 100,000); and % population of color (17.1 per 100,000 vs. 2.9 per 100,000). The gradient is particularly stark for % population of color, whereby populations living in counties where 61-100% of the population is of color experienced a COVID-19 death rate 6-fold greater than those living in counties where 0-17.2% of the population is of color. However, socioeconomic gradients were not always monotonic, most notably for the Index of Concentration at the Extremes, for which residents of counties in the most advantaged quintile experienced a COVID-19 death rate (13.8 per 100,000) only slightly lower than residents of counties in the lowest quintile. In contrast, residents of counties in the middle quintile of the Index of Concentration of the Extremes experienced the lowest COVID-19 death rates (3.9 per 100,000).

### ZCTA level confirmed COVID-19 cases in Illinois

As shown in Figures 2a-d and Table 3, we observed consistent and monotonic socioeconomic gradients in cumulative incidence of COVID-19 diagnoses for all ABSMs using finer resolution

ZCTA-level data in Illinois. The highest rates of COVID-19 confirmed cases were observed among the most disadvantaged compared to most advantaged categories of % poverty (367.7 per 100,000 vs. 155.3 per 100,000), Index of Concentration at the Extremes (438.3 per 100,000 vs. 155.4 per 100,000), % crowding (314.4 per 100,000 vs. 173.0 per 100,000), and % population of color (447.0 per 100,000 vs. 127.8 per 100,000). The steepest gradient was observed by quintiles of % population of color, with residents of ZCTAs in the highest quintile experiencing a rate 3.5 times that of residents in the lowest quintile.

**ZCTA level positive COVID-19 tests in New York City**

Similarly strong socioeconomic gradients were observed with finer resolution ZCTA-level data in New York City in relation to the rate of positive tests. These unequal patterns persist even in the context of New York City's substantially greater rates of infection. The population rate of positive COVID-19 tests was highest among residents in the most disadvantaged vs. most advantaged categories of the Index of Concentration at the Extremes (1603.6 per 100,000 vs. 1067.5 per 100,000), % crowding (1699.0 per 100,000 vs. 1219.4 per 100,000), and % population of color (1771.5 per 100,000 vs. 1248.6 per 100,000). Similarly, the highest rate of positive tests was observed among residents living in counties in the two most disadvantaged categories of ZCTA-level poverty (15-19.9% poverty: 1553.0 per 100,000 and 20-100% poverty: 1504.3 per 100,000, vs. 1046.7 per 100,000 in the most advantaged category, 0-4.9% poverty). These contrasts correspond to relative risks between 1.31 and 1.42.

## DISCUSSION

### The unequal burden of COVID-19

Linkage of available COVID-19 surveillance data to ABSMs at the county and ZIP code levels reveals a substantially unequal burden of COVID-19 outcomes experienced by people living in the most disadvantaged counties and ZCTAs by socioeconomic and sociodemographic characteristics. These strikingly inequitable patterns of disease burden, heretofore obscured by the lack of disaggregated reporting by race/ethnicity and socioeconomic position in publicly available US COVID-19 surveillance data, speak to the urgent need for improved testing, surveillance and monitoring, data transparency, and targeting of public health interventions for community protection and health care resources.

Looking across the US, people living in the most impoverished, crowded, and racially and economically polarized counties are experiencing substantially elevated rates of COVID-19 infection and death. We chose to focus our main analysis on COVID-19 death at the county level because this is the geographic level at which comprehensive data on COVID-19 for all parts of the US are being reported. We focus on death in particular because, unlike confirmed case counts, these numbers are less likely to be affected by well-documented inconsistencies in testing eligibility, procedures, and availability [21-22]. (We do, however, include a county-level analysis of COVID-19 cases in Supplemental Appendix 1). Reported deaths due to COVID-19 nonetheless may not capture the potentially large burden of mortality due to unexplained deaths among individuals who were not tested for SARS-CoV-2, who might have died at home or in nursing facilities, or who might have died of a pre-existing condition whose disease course was exacerbated by coronavirus infection [23-25]. If individuals living in disadvantaged counties were less likely to have been tested for SARS-CoV-2, to have accessed healthcare given infection, or generally less likely to have had their death recorded as COVID-19 related, we would expect that our analyses underestimated the magnitude of inequities across categories of ABSMs.

In spite of these data limitations, we saw strong associations of COVID-19 death rates with all four county-level ABSMs. These inequities are fundamentally related to the material circumstances in which people live and work. For example, individuals living in low income areas may be more likely to be classified as "essential workers" who are less able to practice physical distancing and may not have access to personal protective equipment (PPE) [1-3, 26-27]. "Essential workers" also include many healthcare professionals including nurses, home health aides, and nursing home employees whose risk of occupational exposure to SARS-CoV-2 is high and who live in working class communities [28-30]. Moreover, we noted a strong association with county % crowding, defined as the proportion of households in an area with more than one person per room (excluding bathrooms and kitchens) [31]; by this definition, a one-bedroom apartment with 1 bedroom, 1 dining room, and 1 living room would be categorized as crowded only if 4 or more persons were in the household.

Socioeconomic gradients in COVID-19 death rates by county poverty and the Index of Concentration at the Extremes exhibited more complex patterns. This likely reflects the contribution of particularly large counties with high levels of transmission. Depending on the stratum of county-level ABSM in which it falls, a county with a large number of deaths will tend to dominate the computed rate for that stratum. Table 5 shows the top 25 counties by cumulative count of deaths, along with population and ABSM estimates. These counties include all five boroughs of New York City as well as surrounding areas with high death counts in New York state, New Jersey, and Connecticut. The list also includes other large US urban areas with substantial transmission. Together, these 25 counties account for over 53% of reported COVID-19 deaths in the US. Examination of this list suggests that the higher death rates observed in the 5-9.9% category of county poverty and the most advantaged quintile of the Index of Concentration at the Extremes reflects the contribution of counties like Nassau, Suffolk, Westchester, and New York (Manhattan) Counties, NY to these strata. It is also important to note that county-level analyses gloss over important socioeconomic heterogeneity within counties, which may further contribute to the more complex socioeconomic gradients seen here. Also potentially relevant are changing class dynamics of COVID-19 infections, whereby early cases may have arisen from travelers who could afford international travel, followed by increased risk among essential workers and working class communities with crowded housing.

**ZIP code level analyses**

To illustrate the utility of using finer levels of geography, we additionally presented analyses of confirmed COVID-19 cases in Illinois and positive tests in New York City in the ZCTA level, the only two COVID-19 outcomes for which ZCTA-level data were available in these localities. ZCTA-level analyses revealed more consistently monotonic gradients for all ABSMs, though the magnitude of disparities comparing the top to the bottom socioeconomic categories was smaller on the relative disparity scale. Together, these results suggest that analyzing inequities in COVID-19 outcomes at finer levels of geographic aggregation is feasible and can provide important information about the unequal spread and impact of COVID-19 within counties and cities. As with the county-level death analysis, the results suggest that areas with higher rates of poverty, crowded housing, and populations of color are being disproportionately affected. Moreover, given unequal patterns of testing, if residents of these neighborhoods are not able to access testing, these results may be understanding the true magnitude of inequities in COVID-19 infection.

**Recommendations for public health departments**

The results we have presented reaffirm the urgency of documenting how historically disadvantaged communities are being unequally affected by the devastation of the COVID-19 pandemic. In the absence of national leadership and in the wake of chronic underfunding of public health infrastructure, state and local health departments have been left to fend for themselves in fulfilling the vital functions of public health surveillance in providing an evidence base for action and ensuring accountability [1-2]. The methods of the Public Health Disparities Geocoding Project [8-12] provide a well-validated, robust, and cost-effective methodology by which public health departments can enhance their reporting of disparities in COVID-19 outcomes.

Based on the analyses we have presented here, we recommend that state and local public health departments adopt reporting of COVID-19 outcomes minimally by ZCTA-level characteristics, which we consider preferable to county-level reporting. In our earlier work, we originally recommended routine reporting by socioeconomic characteristics of census tracts [10,16]. While we stand by that recommendation, we recognize that it may be more feasible for surveillance systems to implement ZCTA-level analyses in the short term, since ZIP code is easy to ask of individuals as they are being tested, is already recorded on death certificates, and does not require additional steps for geocoding, compared to census tracts [1]. We emphasize that reporting of disparities by ZCTA characteristics need not entail risk of individual data disclosure due to small numbers in small areas: because our methodology involves aggregating over

ZCTAs with similar socioeconomic characteristics, summary statistics are reported for aggregations of ZCTAs and typically have large enough numbers not to require data suppression [24]. Because of this, we additionally recommend that, whenever possible, public health departments report summary statistics by race/ethnicity, gender, and age within strata of ZCTA-level ABSMs in order to paint a fuller picture of the extent of inequities in COVID-19 outcomes. To assist public health departments who wish to implement these types of analyses, we direct interested readers to the Public Health Disparities Geocoding Project website at http://www.hsph.harvard.edu/thegeocodingproject/.

**Statistical considerations**

Aggregation over areas is analogous to how state and local health departments typically report disease rates by sex and race/ethnicity and avoids problems with statistical instability in the estimation of small area rates at the county and ZCTA levels by essentially assuming that populations within strata of ABSMs have a common disease experience. While marginalizing over disease counts and population at risk may obscure meaningful area differences important to questions of disease etiology or, in the case of COVID-19, infectious disease transmission dynamics, we maintain that cumulative incidence proportions computed for strata of ABSMs still provide an important description of what populations are impacted by COVID-19 and where disease burdens are most substantial.

The analyses we have presented here can be easily implemented by state and local health departments using existing surveillance data and an Excel spreadsheet or similar software. We argue that these simple descriptive analyses of inequities are vital to identifying the communities who are experiencing the most serious impacts of the pandemic and to holding government leaders and policy makers accountable for directing resources to those in need. Throughout, we have presented confidence limits based on traditional formulas for the variance of an incidence rate [25], which assumes that the count of events is Poisson distributed and arises from a homogenous pool of person-time. Given county variation in SARS-CoV-2 transmission dynamics (including when infected cases were seeded in these communities and how the pace of transmission has been affected by containment and mitigation strategies) as well as variation in the susceptibility of populations in these counties above and beyond what is explained by the area-based socioeconomic measures considered here, the assumption of homogeneity is likely unrealistic. More sophisticated statistical models can be employed to model area-level variation in rates, including overdispersed Poisson, negative binomial, mixed models, and zero-inflated models [26-28]. In our experience, however, estimates of socioeconomic inequities can be sensitive to the modelling approach taken, and the interpretation of summary measures of health disparities at the population level may be complicated by model assumptions. Even when there are variations in area-level rates within strata of ABSMs, estimates from the aggregated method still have relevant interpretation as the "average" health experience of persons living in areas with particular socioeconomic characteristics. While our future work will address small-area estimation and appropriate models for handling spatial heterogeneity in COVID-19 outcomes, we should not lose sight of the immediate need for timely data on economic and social inequities to inform policy and interventions.

## REFERENCES


1. Krieger N, Gonsalves G, Bassett MT, Hanage W, Krumholz HM. "The Fierce Urgency Of Now: Closing Glaring Gaps In US Surveillance Data On COVID-19" *Health Affairs Blog*, April 14, 2020. DOI: 10.1377/hblog20200414.238084 Accessed April 21, 2020.

2. Krieger N. COVID-19, Data, and Health Justice. Commonwealth Fund, To The Point, April 16, 2020. https://www.commonwealthfund.org/blog/2020/covid-19-data-and-health-justice Accessed April 21, 2020

3. Kendi IX, What the racial data show, The Atlantic, April 6, 2020. https://www.theatlantic.com/ideas/archive/2020/04/coronavirus-exposing-our-racial-divides/609526 Accessed April 21, 2020.

4. Kendi IX, Why don't we know who the coronavirus victims are? The Atlantic, April 1, 2020. https://www.theatlantic.com/ideas/archive/2020/04/stop-looking-away-race-covid-19-victims/609250/ Accessed April 21, 2020

5. The COVID Tracking Project. Available at: https://covidtracking.com/about-data. Accessed April 21, 2020.

6. US Centers for Disease Control and Prevention. Cases of Coronavirus (COVID-19) in the US. Available at: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html. Accessed April 20, 2020.

7. Rehkopf DH, Haughton L, Chen JT, Waterman PD, Subramanian SV, Krieger N. Monitoring socioeconomic disparities in death: comparing individual-level education and area-based socioeconomic measures. Am J Public Health 2006; 96:2135-2138; doi:10.2105/AJPH.2005.075408. (erratum: AJPH 2007; 97:1543).

8. Krieger N, Chen JT, Waterman PD, Rehkopf DH, Subramanian SV. Painting a truer picture of US socioeconomic and racial/ethnic health inequalities: The Public Health Disparities Geocoding Project. Am J Public Health 2005; 95: 312-323.

9. Krieger N, Chen JT, Waterman PD, Rehkopf DH, Subramanian SV. Race/ethnicity, gender, and monitoring socioeconomic gradients in health: a comparison of area-based socioeconomic measures—the *Public Health Disparities Geocoding Project*. Am J Public Health. 2003; 93: 1655-1671.

10. Krieger N, Chen JT, Waterman PD, Soobader M-J, Subramanian SV, Carson R. Geocoding and Monitoring US Socioeconomic Inequalities in Mortality and Cancer Incidence: Does Choice of Area-Based Measure and Geographic Level Matter?—*The Public Health Disparities Geocoding Project.* Am J Epidemiol 2002; 156(5):471-82.

11. Krieger N, Chen JT, Waterman PD, Soobader M-J, Subramanian SV, Carson R. Monitoring socioeconomic inequalities in sexually transmitted diseases, tuberculosis, and

violence: geocoding and choice of area-based socioeconomic measures -- *The Public Health Disparities Geocoding Project*. Public Health Report. 2003; 118:240-260.

12. Krieger N, Chen JT, Waterman PD, Soobader M-J, Subramanian SV, Carson R. Choosing area-based socioeconomic measures to monitor social inequalities in low birthweight and childhood lead poisoning –*The Public Health Disparities Geocoding Project* (US). J Epidemiol Community Health 2003; 57:186-199.

13. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis*; published online Feb 19. https://doi.org/10.1016/S1473-3099(20)30120-1. Data available on Github: https://github.com/CSSEGISandData/COVID-19 Accessed April 16, 2020.

14. COVID-19 deaths dataset. USA Facts. Available at: https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/ Accessed April 16, 2020.

15. Eads D, Friedrich P. MAP: Find the number of COVID-19 cases in Illinois by ZIP code. Available at: https://www.wbez.org/stories/map-by-zip-code-of-coronavirus-covid-19-cases-illinois/90ca85cd-bdf4-423a-a7bc-924fcee9d0f3 Accessed April 16, 2020.

16. Krieger N, Waterman P, Chen JT, Soobader M-J, Subramanian SV, Carson R. ZIP Code caveat: bias to spatiotemporal mismatches between ZIP Codes and US census-defined areas -- *The Public Health Disparities Geocoding Project*. *Am J Public Health* 2002; 92:1100-1102.

17. New York City Department of Health and Mental Hygiene COVID-19 GitHub repository. Available at: https://raw.githubusercontent.com/nychealth/coronavirus-data/master/tests-by-zcta.csv. Accessed April 17, 2020.

18. US Census Bureau. 2014-2018 American Community Survey 5-year estimates. Accessed April 17, 2020.

19. Tidy census package: https://github.com/walkerke/tidycensus

20. Krieger N, Waterman PD, Spasojevic J, Li W, Maduro G, Van Wye G. Public Health monitoring of privilege and deprivation using the Index of Concentration at the Extremes (ICE). *Am J Public Health* 2016; 106:256-263. doi: 10.2105/AJPH.2015/302955

21. Kost GJ. Geospatial Hotspots Need Point-of-Care Strategies to Stop Highly Infectious Outbreaks: Ebola and Coronavirus. *Arch Pathol Lab Me*d 2020 Apr 16. doi: 10.5858/arpa.2020-0172-RA [epub ahead of print]

22. CDC COVID-19 Response Team. Geographic differences in COVID-19 cases, deaths, and incidence – United States, February 12-April 7, 2020. MMWR 2020 Apr 17; 69(15):465-471. doi: 10.15885/mmwr.mm6915e4

23. Wu J, McCann Al. 28,000 missing deaths: tracking the true toll of the coronavirus crisis. *New York Times*, April 21, 2020. https://www.nytimes.com/interactive/2020/04/21/world/coronavirus-missing-deaths.html

24. Watkins A, Rashbaum WK. How many people have actually died from coronavirus in New York? *New York Times*, April 10, 2020. https://www.nytimes.com/2020/04/10/nyregion/new-york-coronavirus-death-count.html

25. CDC COVID-19 Response Team. Geographic differences in COVID-19 cases, deaths, and incidence – United States, February 12-April 7, 2020. MMWR 2020 Apr 17; 69(15):465-471. doi: 10.15885/mmwr.mm6915e4

26. García M. Chelsea, city of the working class Latino immigrant, emerges as COVID-19 hotspot. *Boston Globe*, April 7, 2020. https://www.bostonglobe.com/2020/04/07/opinion/chelsea-city-working-latino-immigrant-emerges-covid-19-hotspot/?p1=Article_Inline_Related_Link Accessed April 21, 2020.

27. Hawkins D. The coronavirus is falling heavily on black Americans. Why? *The Guardian*, April 16, 2020. https://www.theguardian.com/commentisfree/2020/apr/16/black-workers-coronavirus-covid-19 Accessed April 21, 2020.

28. Bailey M, Gee A, Jewett C, Renwick, Varney S. COVID-19: Lost on the Frontline. Kaiser Health News, April 15, 2020. https://khn.org/news/lost-on-the-frontline-health-care-worker-death-toll-covid19-coronavirus/ Accessed April 21, 2020.

29. Campbell R, Gebeloff R. How millions of women became the most essential workers in America. *New York Times*, April 18, 2020. https://www.nytimes.com/2020/04/18/us/coronavirus-women-essential-workers.html Accessed April 21, 2020.

30. O'Donnell J. How do health care staffers learn if co-workers have coronavirus? Word of mouth. *USA Today*, April 17, 2020. https://www.usatoday.com/story/news/health/2020/04/17/coronavirus-health-care-workers-warning-covid-19-colleagues/2982026001/ Accessed April 21, 2020.

31. US Department of Housing and Urban Development Office of Policy Development and Research (2007). *Measuring Overcrowding in Housing*. Available at: https://www.census.gov/content/dam/Census/programs-surveys/ahs/publications/Measuring_Overcrowding_in_Hsg.pdf. Accessed April 21, 2020

32. Chen, JT. Merging survey data with aggregate data from other sources: opportunities and challenges. In: Timothy P. Johnson (ed.) *Handbook of health survey methods*. New York, NY: John Wiley & Sons, 2015.

33. Rothman KJ, Greenland S. (1998) *Modern Epidemiology*, Chapter 15. Philadelphia, PA: Lippincott-Raven

34. Chen, JT. Multilevel and hierarchical models for disease mapping. In: Boscoe, FP (ed). *Geographic health data: fundamental techniques for analysis*. Oxfordshire, UK: CABI, 2013: 183-208.

35. Wu X, Nethery RC, Sabath BM, Braun D, Dominici F. Exposure to air pollution and COVID-19 mortality in the United States. medRxiv 2020.04.05.20054502; doi: https://doi.org/10.1101/2020.04.05.20054502

Table 1: Population counts and area based socioeconomic measures, source variables, and cutpoints computed from the 2014-2018 American Community Survey 5-year estimates

| **Variable** | **Formula: Source Variables** | **US County Cutpoints** | **ZCTA cutpoints (Illinois)** | **ZCTA cutpoints (NYC)** |
|---|---|---|---|---|
| *Population Counts* | | | | |
| Total population | B01003_001E | | | |
| White Non-Hispanic Population | B01001H_001E | | | |
| *Area-based socioeconomic measures* | | | | |
| % of persons below poverty | B17001_002E/B17001_001E | 0-4.9%<br>5-9.9%<br>10-14.9%<br>15-19.9%<br>20-100% | 0-4.9%<br>5-9.9%<br>10-14.9%<br>15-19.9%<br>20-100% | 0-4.9%<br>5-9.9%<br>10-14.9%<br>15-19.9%<br>20-100% |
| Index of Concentration at the Extremes (high income white households versus low income black households) | ((B19001A_014E + B19001A_015E + B19001A_016E + B19001A_017E) - (B19001B_002E + B19001B_003E + B19001B_004E + B19001B_005E))/B19001_001E, | Q1: (-0.522,0.114]<br>Q2: (0.114,0.159]<br>Q3: (0.159,0.205]<br>Q4: (0.205,0.283]<br>Q5: (0.283,0.536] | Q1: (-0.612,0.0175]<br>Q2: (0.0175,0.171]<br>Q3: (0.171,0.289]<br>Q4: (0.289,0.403]<br>Q(5: 0.403,0.721] | Q1: (-0.385,-0.102]<br>Q2: (-0.102,0.0212]<br>Q3: (0.0212,0.141]<br>Q4: (0.141,0.29]<br>Q5: (0.29,0.7] |
| % crowding (>1 person per room) | (B25014_005E + B25014_006E + B25014_007E + B25014_011E + B25014_012E + B25014_013E) / B25014_001E | Q1: (0,0.0147]<br>Q2: (0.0147,0.0212]<br>Q3: (0.0212,0.0306]<br>Q4: (0.0306,0.0491]<br>Q5: (0.0491,0.493] | Q1: (0,0.00975]<br>Q2:(0.00975,0.0177]<br>Q3:(0.0177,0.0274]<br>Q4: (0.0274,0.0472]<br>Q5: (0.0472,0.143] | Q1:(0.00942,0.0478]<br>Q2: (0.0478,0.0698]<br>Q3: (0.0698,0.0978]<br>Q4: (0.0978,0.138]<br>Q5: (0.138,0.297] |
| % population of color (not White Non-Hispanic) | B01003_001E - B01001H_001E)/ B01003_001E | Q1: (0,0.172]<br>Q2: (0.172,0.302]<br>Q3: (0.302,0.443]<br>Q4: (0.443,0.61]<br>Q5: (0.61,1] | Q1: (0.0318,0.197]<br>Q2:c(0.197,0.315]<br>Q3: (0.315,0.46]<br>Q4: (0.46,0.744]<br>Q5: (0.744,0.99] | Q1: (0.0839,0.402]<br>Q2: (0.402,0.584]<br>Q3: (0.584,0.826]<br>Q4: (0.826,0.957]<br>Q5: (0.957,0.992] |

Table 2: US COVID-19 death rate per 100,000 by county characteristics as of 4/16/2020

| | Number of counties | Number of deaths | Population | Death rate per 100,000 | (95% CI) | | Rate difference per 100,000 | (95% CI) | | Rate ratio | (95% CI) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| % poverty (categories) | | | | | | | | | | | | |
| 0-4.9% | 41 | 443 | 4,495,932 | 9.9 | (8.9 | ,10.8) | 0.0 | (reference) | | 1.00 | (reference) | |
| 5-9.9% | 558 | 7,877 | 71,157,744 | 11.1 | (10.8 | ,11.3) | 1.2 | (0.3 | ,2.2) | 1.12 | (1.02 | ,1.24) |
| 10-14.9% | 1,023 | 8,031 | 108,820,591 | 7.4 | (7.2 | ,7.5) | -2.5 | (-3.4 | ,-1.5) | 0.75 | (0.68 | ,0.82) |
| 15-19.9% | 860 | 6,654 | 101,961,251 | 6.5 | (6.4 | ,6.7) | -3.3 | (-4.3 | ,-2.4) | 0.66 | (0.60 | ,0.73) |
| 20-100% | 659 | 7,034 | 36,428,205 | 19.3 | (18.9 | ,19.8) | 9.5 | (8.4 | ,10.5) | 1.96 | (1.78 | ,2.16) |
| missing | | 279 | | | | | | | | | | |
| Index of Concentration at the Extremes (high income white households versus low income black households) | | | | | | | | | | | | |
| (-0.522,0.114] | 974 | 9,314 | 61,949,063 | 15.0 | (14.7 | ,15.3) | 1.3 | (0.8 | ,1.7) | 1.09 | (1.06 | ,1.12) |
| (0.114,0.159] | 701 | 4,941 | 64,942,197 | 7.6 | (7.4 | ,7.8) | -6.2 | (-6.5 | ,-5.8) | 0.55 | (0.53 | ,0.57) |
| (0.159,0.205] | 696 | 2,564 | 65,113,354 | 3.9 | (3.8 | ,4.1) | -9.8 | (-10.2 | ,-9.5) | 0.29 | (0.27 | ,0.30) |
| (0.205,0.283] | 515 | 4,082 | 64,525,801 | 6.3 | (6.1 | ,6.5) | -7.4 | (-7.8 | ,-7.1) | 0.46 | (0.44 | ,0.48) |
| (0.283,0.536] | 255 | 9,138 | 66,333,308 | 13.8 | (13.5 | ,14.1) | 0.0 | (reference) | | 1.00 | (reference) | |
| missing | | 279 | | | | | | | | | | |
| % crowding (quintiles) | | | | | | | | | | | | |
| (0,0.0147] | 1,047 | 3,189 | 65,273,354 | 4.9 | (4.7 | ,5.1) | 0.0 | (reference) | | 1.00 | (reference) | |
| (0.0147,0.0212] | 709 | 3,973 | 64,425,866 | 6.2 | (6.0 | ,6.4) | 1.3 | (1.0 | ,1.5) | 1.26 | (1.20 | ,1.32) |
| (0.0212,0.0306] | 656 | 6,739 | 63,510,499 | 10.6 | (10.4 | ,10.9) | 5.7 | (5.4 | ,6.0) | 2.17 | (2.08 | ,2.27) |
| (0.0306,0.0491] | 443 | 5,423 | 65,654,959 | 8.3 | (8.0 | ,8.5) | 3.4 | (3.1 | ,3.7) | 1.69 | (1.62 | ,1.77) |
| (0.0491,0.493] | 244 | 10,715 | 63,913,934 | 16.8 | (16.4 | ,17.1) | 11.9 | (11.5 | ,12.2) | 3.43 | (3.30 | ,3.57) |
| missing | | 279 | | | | | | | | | | |
| % percent population of color | | | | | | | | | | | | |
| (0,0.172] | 1,635 | 1,862 | 65,219,459 | 2.9 | (2.7 | ,3.0) | 0.0 | (reference) | | 1.00 | (reference) | |
| (0.172,0.302] | 549 | 3,981 | 65,166,967 | 6.1 | (5.9 | ,6.3) | 3.3 | (3.0 | ,3.5) | 2.14 | (2.03 | ,2.26) |
| (0.302,0.443] | 468 | 7,034 | 69,376,152 | 10.1 | (9.9 | ,10.4) | 7.3 | (7.0 | ,7.6) | 3.55 | (3.37 | ,3.74) |
| (0.443,0.61] | 280 | 6,534 | 60,922,155 | 10.7 | (10.5 | ,11.0) | 7.9 | (7.6 | ,8.2) | 3.76 | (3.57 | ,3.96) |
| (0.61,1] | 209 | 10,628 | 62,217,817 | 17.1 | (16.8 | ,17.4) | 14.2 | (13.9 | ,14.6) | 5.98 | (5.70 | ,6.29) |
| missing | | 279 | | | | | | | | | | |

Table 3: Illinois rate of confirmed COVID-19 cases per 100,000 population by ZCTA characteristics as of 4/16/2020

| | Number of ZCTAs | Number of confirmed cases | Population | Confirmed case rate per 100,000 | (95% CI) | | Rate difference per 100,000 | (95% CI) | | Rate ratio | (95% CI) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| % poverty (categories) | | | | | | | | | | | | |
| 0-4.9% | 65 | 2,378 | 1,531,569 | 155.3 | (149.0 | ,161.5) | 0.0 | (reference) | | 1.00 | (reference) | |
| 5-9.9% | 138 | 6,442 | 3,357,448 | 191.9 | (187.2 | ,196.6) | 36.6 | (28.8 | ,44.4) | 1.24 | (1.18 | ,1.30) |
| 10-14.9% | 65 | 4,682 | 2,052,094 | 228.2 | (221.6 | ,234.7) | 72.9 | (63.9 | ,81.9) | 1.47 | (1.40 | ,1.54) |
| 15-19.9% | 39 | 3,085 | 1,225,648 | 251.7 | (242.8 | ,260.6) | 96.4 | (85.6 | ,107.3) | 1.62 | (1.54 | ,1.71) |
| 20-100% | 63 | 8,041 | 2,186,595 | 367.7 | (359.7 | ,375.8) | 212.5 | (202.3 | ,222.7) | 2.37 | (2.26 | ,2.48) |
| missing | | 47 | | | | | | | | | | |
| Index of Concentration at the Extremes (high income white households versus low income black households) | | | | | | | | | | | | |
| (-0.612,0.0175] | 63 | 9,077 | 2,070,809 | 438.3 | (429.3 | ,447.3) | 283.0 | (272.5 | ,293.5) | 2.82 | (2.71 | ,2.94) |
| (0.0175,0.171] | 72 | 4,258 | 2,087,542 | 204.0 | (197.8 | ,210.1) | 48.6 | (40.5 | ,56.8) | 1.31 | (1.25 | ,1.37) |
| (0.171,0.289] | 75 | 4,582 | 2,070,229 | 221.3 | (214.9 | ,227.7) | 66.0 | (57.6 | ,74.3) | 1.42 | (1.36 | ,1.49) |
| (0.289,0.403] | 77 | 3,502 | 2,058,711 | 170.1 | (164.5 | ,175.7) | 14.7 | (7.0 | ,22.5) | 1.09 | (1.04 | ,1.15) |
| (0.403,0.721] | 82 | 3,196 | 2,057,150 | 155.4 | (150.0 | ,160.7) | 0.0 | (reference) | | 1.00 | (reference) | |
| missing | | 60 | | | | | | | | | | |
| % crowding (quintiles) | | | | | | | | | | | | |
| (0,0.00975] | 87 | 3,370 | 1,948,122 | 173.0 | (167.1 | ,178.8) | 0.0 | (reference) | | 1.00 | (reference) | |
| (0.00975,0.0177] | 82 | 3,131 | 2,060,973 | 151.9 | (146.6 | ,157.2) | -21.1 | (-29.0 | ,-13.2) | 0.88 | (0.84 | ,0.92) |
| (0.0177,0.0274] | 64 | 5,009 | 2,052,139 | 244.1 | (237.3 | ,250.8) | 71.1 | (62.2 | ,80.0) | 1.41 | (1.35 | ,1.47) |
| (0.0274,0.0472] | 68 | 6,386 | 2,101,938 | 303.8 | (296.4 | ,311.3) | 130.8 | (121.4 | ,140.3) | 1.76 | (1.68 | ,1.83) |
| (0.0472,0.143] | 54 | 6,450 | 2,051,676 | 314.4 | (306.7 | ,322.0) | 141.4 | (131.7 | ,151.0) | 1.82 | (1.74 | ,1.89) |
| missing | | 329 | | | | | | | | | | |
| % percent population of color | | | | | | | | | | | | |
| (0.0318,0.197] | 99 | 2,651 | 2,073,667 | 127.8 | (123.0 | ,132.7) | 0.0 | (reference) | | 1.00 | (reference) | |
| (0.197,0.315] | 78 | 2,992 | 2,023,605 | 147.9 | (142.6 | ,153.2) | 20.0 | (12.8 | ,27.2) | 1.16 | (1.10 | ,1.22) |
| (0.315,0.46] | 77 | 4,071 | 2,159,499 | 188.5 | (182.7 | ,194.3) | 60.7 | (53.1 | ,68.2) | 1.47 | (1.40 | ,1.55) |
| (0.46,0.744] | 60 | 5,731 | 2,038,179 | 281.2 | (273.9 | ,288.5) | 153.3 | (144.6 | ,162.1) | 2.20 | (2.10 | ,2.30) |
| (0.744,0.99] | 55 | 9,172 | 2,051,861 | 447.0 | (437.9 | ,456.2) | 319.2 | (308.8 | ,329.5) | 3.50 | (3.35 | ,3.65) |
| missing | | 58 | | | | | | | | | | |

Chen JT, Krieger N. Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county vs ZIP code analyses. *Harvard Center for Population and Development Studies Working Paper Series*, Volume 19, Number 1. April 21, 2020. https://tinyurl.com/y7v72446

Table 4: New York City rate of positive COVID-19 tests per 100,000 population by ZCTA characteristics as of 4/16/2020

| | Number of ZCTAs | Number of positive tests | Population | Rate per 100,000 | (95% CI) | | Rate difference per 100,000 | (95% CI) | | Rate ratio | (95% CI) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| % poverty (categories) | | | | | | | | | | | | |
| 0-4.9% | 9 | 1,362 | 130,121 | 1046.7 | (991.1 | ,1102.3) | 0.0 | (reference) | | 1.00 | (reference) | |
| 5-9.9% | 41 | 20,609 | 1,506,286 | 1368.2 | (1349.5 | ,1386.9) | 321.5 | (262.8 | ,380.1) | 1.31 | (1.24 | ,1.38) |
| 10-14.9% | 48 | 30,294 | 2,100,915 | 1441.9 | (1425.7 | ,1458.2) | 395.2 | (337.3 | ,453.1) | 1.38 | (1.30 | ,1.45) |
| 15-19.9% | 27 | 22,359 | 1,439,746 | 1553.0 | (1532.6 | ,1573.3) | 506.3 | (447.1 | ,565.5) | 1.48 | (1.40 | ,1.57) |
| 20+% | 52 | 48,982 | 3,256,108 | 1504.3 | (1491.0 | ,1517.6) | 457.6 | (400.4 | ,514.8) | 1.44 | (1.36 | ,1.52) |
| missing | | 1,816 | | | | | | | | | | |
| Index of Concentration at the Extremes (high income white households versus low income black households) | | | | | | | | | | | | |
| (-0.385,-0.102] | 28 | 25,855 | 1,612,266 | 1603.6 | (1584.1 | ,1623.2) | 536.2 | (511.1 | ,561.2) | 1.50 | (1.47 | ,1.53) |
| (-0.102,0.0212] | 30 | 28,209 | 1,749,736 | 1612.2 | (1593.4 | ,1631.0) | 544.7 | (520.3 | ,569.2) | 1.51 | (1.48 | ,1.54) |
| (0.0212,0.141] | 29 | 26,844 | 1,623,732 | 1653.2 | (1633.5 | ,1673.0) | 585.8 | (560.6 | ,611.0) | 1.55 | (1.52 | ,1.58) |
| (0.141,0.29] | 39 | 23,751 | 1,692,826 | 1403.0 | (1385.2 | ,1420.9) | 335.6 | (311.9 | ,359.3) | 1.31 | (1.29 | ,1.34) |
| (0.29,0.7] | 50 | 17,913 | 1,678,089 | 1067.5 | (1051.8 | ,1083.1) | 0.0 | (reference) | | 1.00 | (reference) | |
| missing | | 2,850 | | | | | | | | | | |
| % crowding (quintiles) | | | | | | | | | | | | |
| (0.00942,0.0478] | 47 | 20,428 | 1,675,260 | 1219.4 | (1202.7 | ,1236.1) | 0.0 | (reference) | | 1.00 | (reference) | |
| (0.0478,0.0698] | 37 | 23,808 | 1,688,963 | 1409.6 | (1391.7 | ,1427.5) | 190.2 | (165.7 | ,214.7) | 1.16 | (1.13 | ,1.18) |
| (0.0698,0.0978] | 38 | 24,507 | 1,679,177 | 1459.5 | (1441.2 | ,1477.7) | 240.1 | (215.3 | ,264.8) | 1.20 | (1.17 | ,1.22) |
| (0.0978,0.138] | 31 | 25,783 | 1,682,708 | 1532.2 | (1513.5 | ,1550.9) | 312.8 | (287.8 | ,337.9) | 1.26 | (1.23 | ,1.28) |
| (0.138,0.297] | 23 | 28,434 | 1,673,537 | 1699.0 | (1679.3 | ,1718.8) | 479.6 | (453.8 | ,505.5) | 1.39 | (1.37 | ,1.42) |
| missing | | 2,462 | | | | | | | | | | |
| % population of color (quintiles) | | | | | | | | | | | | |
| (0.0839,0.402] | 43 | 21,166 | 1,695,113 | 1248.6 | (1231.8 | ,1265.5) | 0.0 | (reference) | | 1.00 | (reference) | |
| (0.402,0.584] | 38 | 20,554 | 1,678,144 | 1224.8 | (1208.1 | ,1241.6) | -23.8 | (-47.6 | ,-0.1) | 0.98 | (0.96 | ,1.00) |
| (0.584,0.826] | 38 | 25,541 | 1,708,248 | 1495.2 | (1476.8 | ,1513.5) | 246.5 | (221.6 | ,271.4) | 1.20 | (1.18 | ,1.22) |
| (0.826,0.957] | 29 | 27,231 | 1,708,722 | 1593.6 | (1574.7 | ,1612.6) | 345.0 | (319.7 | ,370.3) | 1.28 | (1.25 | ,1.30) |
| (0.957,0.992] | 28 | 29,042 | 1,639,409 | 1771.5 | (1751.1 | ,1791.9) | 522.8 | (496.4 | ,549.3) | 1.42 | (1.39 | ,1.44) |
| missing | | 1,888 | | | | | | | | | | |

Table 5: Deaths, population, crude death rate, and county-level area-based measures for counties with the largest cumulative death counts as of 4/16/2020

| FIPS code | County Name | State | Deaths | Population | Crude death rate per 100,000 | % below poverty | Index of Concentration at the Extremes (white/black race + income) | % crowding (>1 person per room) | % population of color |
|---|---|---|---|---|---|---|---|---|---|
| 36081 | Queens County | NY | 37,918 | 2,298,513 | 1649.7 | 0.130 | 0.117 | 0.095 | 0.747 |
| 36047 | Kings County | NY | 33,521 | 2,600,747 | 1288.9 | 0.211 | 0.070 | 0.103 | 0.638 |
| 36059 | Nassau County | NY | 27,772 | 1,356,564 | 2047.2 | 0.057 | 0.412 | 0.026 | 0.392 |
| 36005 | Bronx County | NY | 25,932 | 1,437,872 | 1803.5 | 0.291 | -0.065 | 0.123 | 0.907 |
| 36103 | Suffolk County | NY | 24,182 | 1,487,901 | 1625.2 | 0.071 | 0.416 | 0.026 | 0.319 |
| 36119 | Westchester County | NY | 21,828 | 968,815 | 2253.1 | 0.092 | 0.336 | 0.041 | 0.460 |
| 17031 | Cook County | IL | 18,087 | 5,223,719 | 346.2 | 0.151 | 0.138 | 0.034 | 0.575 |
| 36061 | New York County | NY | 17,091 | 1,632,480 | 1046.9 | 0.166 | 0.289 | 0.058 | 0.531 |
| 26163 | Wayne County | MI | 13,002 | 1,761,382 | 738.2 | 0.231 | -0.022 | 0.022 | 0.504 |
| 34003 | Bergen County | NJ | 11,409 | 929,999 | 1226.8 | 0.070 | 0.356 | 0.024 | 0.427 |
| 6037 | Los Angeles County | CA | 10,854 | 10,098,052 | 107.5 | 0.160 | 0.168 | 0.114 | 0.737 |
| 34017 | Hudson County | NJ | 9,165 | 668,631 | 1370.7 | 0.163 | 0.175 | 0.075 | 0.711 |
| 34013 | Essex County | NJ | 9,084 | 793,555 | 1144.7 | 0.164 | 0.072 | 0.042 | 0.692 |
| 36087 | Rockland County | NY | 8,752 | 323,686 | 2703.9 | 0.143 | 0.337 | 0.066 | 0.367 |
| 36085 | Richmond County | NY | 8,684 | 474,101 | 1831.7 | 0.128 | 0.293 | 0.043 | 0.383 |
| 12086 | Miami-Dade County | FL | 8,326 | 2,715,516 | 306.6 | 0.180 | 0.127 | 0.063 | 0.866 |
| 34039 | Union County | NJ | 7,904 | 553,066 | 1429.1 | 0.098 | 0.227 | 0.045 | 0.597 |
| 42101 | Philadelphia County | PA | 7,684 | 1,575,522 | 487.7 | 0.249 | -0.040 | 0.026 | 0.654 |
| 34031 | Passaic County | NJ | 7,317 | 504,041 | 1451.7 | 0.167 | 0.220 | 0.071 | 0.582 |
| 25017 | Middlesex County | MA | 7,206 | 1,595,192 | 451.7 | 0.079 | 0.400 | 0.019 | 0.275 |
| 34023 | Middlesex County | NJ | 6,994 | 826,698 | 846.0 | 0.085 | 0.238 | 0.042 | 0.562 |
| 25025 | Suffolk County | MA | 6,820 | 791,766 | 861.4 | 0.193 | 0.192 | 0.036 | 0.550 |
| 9001 | Fairfield County | CT | 6,816 | 944,348 | 721.8 | 0.088 | 0.379 | 0.027 | 0.376 |
| 36071 | Orange County | NY | 5,888 | 378,227 | 1556.7 | 0.118 | 0.289 | 0.037 | 0.351 |
| 22071 | Orleans Parish | LA | 5,847 | 389,648 | 1500.6 | 0.246 | -0.134 | 0.015 | 0.694 |

Chen JT, Krieger N. Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county vs ZIP code analyses. *Harvard Center for Population and Development Studies Working Paper Series*, Volume 19, Number 1. April 21, 2020. https://tinyurl.com/y7v72446

## Supplemental Appendix Table A.1: US COVID-19 cases per 100,000 by county characteristics as of 4/16/2020

| | Number of counties | Number of deaths | Population | Death rate per 100,000 | (95% CI) | Rate difference per 100,000 | (95% CI) | Rate ratio | (95% CI) |
|---|---|---|---|---|---|---|---|---|---|
| % poverty (categories) | | | | | | | | | |
| 0-4.9% | 41 | 9,236 | 4,495,932 | 205.4 | (201.2 ,209.6) | 0.0 | (reference) | 1.00 | (reference) |
| 5-9.9% | 558 | 200,112 | 71,157,744 | 281.2 | (280.0 ,282.5) | 75.8 | (71.4 ,80.2) | 1.37 | (1.34 ,1.40) |
| 10-14.9% | 1023 | 177,196 | 108,820,591 | 162.8 | (162.1 ,163.6) | -42.6 | (-46.9 ,-38.3) | 0.79 | (0.78 ,0.81) |
| 15-19.9% | 860 | 161,502 | 101,961,251 | 158.4 | (157.6 ,159.2) | -47.0 | (-51.3 ,-42.8) | 0.77 | (0.76 ,0.79) |
| 20-100% | 659 | 112,604 | 36,428,205 | 309.1 | (307.3 ,310.9) | 103.7 | (99.1 ,108.2) | 1.50 | (1.47 ,1.54) |
| missing | | 31 | | | | | | | |
| Index of Concentration at the Extremes (high income white households versus low income black households) | | | | | | | | | |
| (-0.522,0.114] | 974 | 160,588 | 61,949,063 | 259.2 | (258.0 ,260.5) | -82.8 | (-84.7 ,-80.9) | 0.76 | (0.75 ,0.76) |
| (0.114,0.159] | 701 | 103,896 | 64,942,197 | 160.0 | (159.0 ,161.0) | -182.1 | (-183.8 ,-180.4) | 0.47 | (0.46 ,0.47) |
| (0.159,0.205] | 696 | 70,626 | 65,113,354 | 108.5 | (107.7 ,109.3) | -233.6 | (-235.2 ,-232.0) | 0.32 | (0.31 ,0.32) |
| (0.205,0.283] | 515 | 98,635 | 64,525,801 | 152.9 | (151.9 ,153.8) | -189.2 | (-190.9 ,-187.5) | 0.45 | (0.44 ,0.45) |
| (0.283,0.536] | 255 | 226,905 | 66,333,308 | 342.1 | (340.7 ,343.5) | 0.0 | (reference) | 1.00 | (reference) |
| missing | | 31 | | | | | | | |
| % crowding (quintiles) | | | | | | | | | |
| (0,0.0147] | 1047 | 75,149 | 65,273,354 | 115.1 | (114.3 ,116.0) | 0.0 | (reference) | 1.00 | (reference) |
| (0.0147,0.0212] | 709 | 95,224 | 64,425,866 | 147.8 | (146.9 ,148.7) | 32.7 | (31.4 ,33.9) | 1.28 | (1.27 ,1.30) |
| (0.0212,0.0306] | 656 | 160,008 | 63,510,499 | 251.9 | (250.7 ,253.2) | 136.8 | (135.3 ,138.3) | 2.19 | (2.17 ,2.21) |
| (0.0306,0.0491] | 443 | 142,573 | 65,654,959 | 217.2 | (216.0 ,218.3) | 102.0 | (100.6 ,103.4) | 1.89 | (1.87 ,1.90) |
| (0.0491,0.493] | 244 | 187,660 | 63,913,934 | 293.6 | (292.3 ,294.9) | 178.5 | (176.9 ,180.0) | 2.55 | (2.53 ,2.57) |
| missing | | 67 | | | | | | | |
| % percent population of color | | | | | | | | | |
| (0,0.172] | 1635 | 44,958 | 65,219,459 | 68.9 | (68.3 ,69.6) | 0.0 | (reference) | 1.00 | (reference) |
| (0.172,0.302] | 549 | 95,876 | 65,166,967 | 147.1 | (146.2 ,148.1) | 78.2 | (77.1 ,79.3) | 2.13 | (2.11 ,2.16) |
| (0.302,0.443] | 468 | 177,223 | 69,376,152 | 255.5 | (254.3 ,256.6) | 186.5 | (185.2 ,187.9) | 3.71 | (3.67 ,3.74) |
| (0.443,0.61] | 280 | 155,758 | 60,922,155 | 255.7 | (254.4 ,256.9) | 186.7 | (185.3 ,188.2) | 3.71 | (3.67 ,3.75) |
| (0.61,1] | 209 | 186,845 | 62,217,817 | 300.3 | (298.9 ,301.7) | 231.4 | (229.9 ,232.9) | 4.36 | (4.31 ,4.40) |
| missing | | 21 | | | | | | | |

Chen JT, Krieger N. Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county vs ZIP code analyses. *Harvard Center for Population and Development Studies Working Paper Series*, Volume 19, Number 1. April 21, 2020. https://tinyurl.com/y7v72446

**LIST OF FIGURES:**

Figure 1a: US COVID−19 deaths per 100,000 population
by county % below poverty (categories) (as of 4.16.2020)

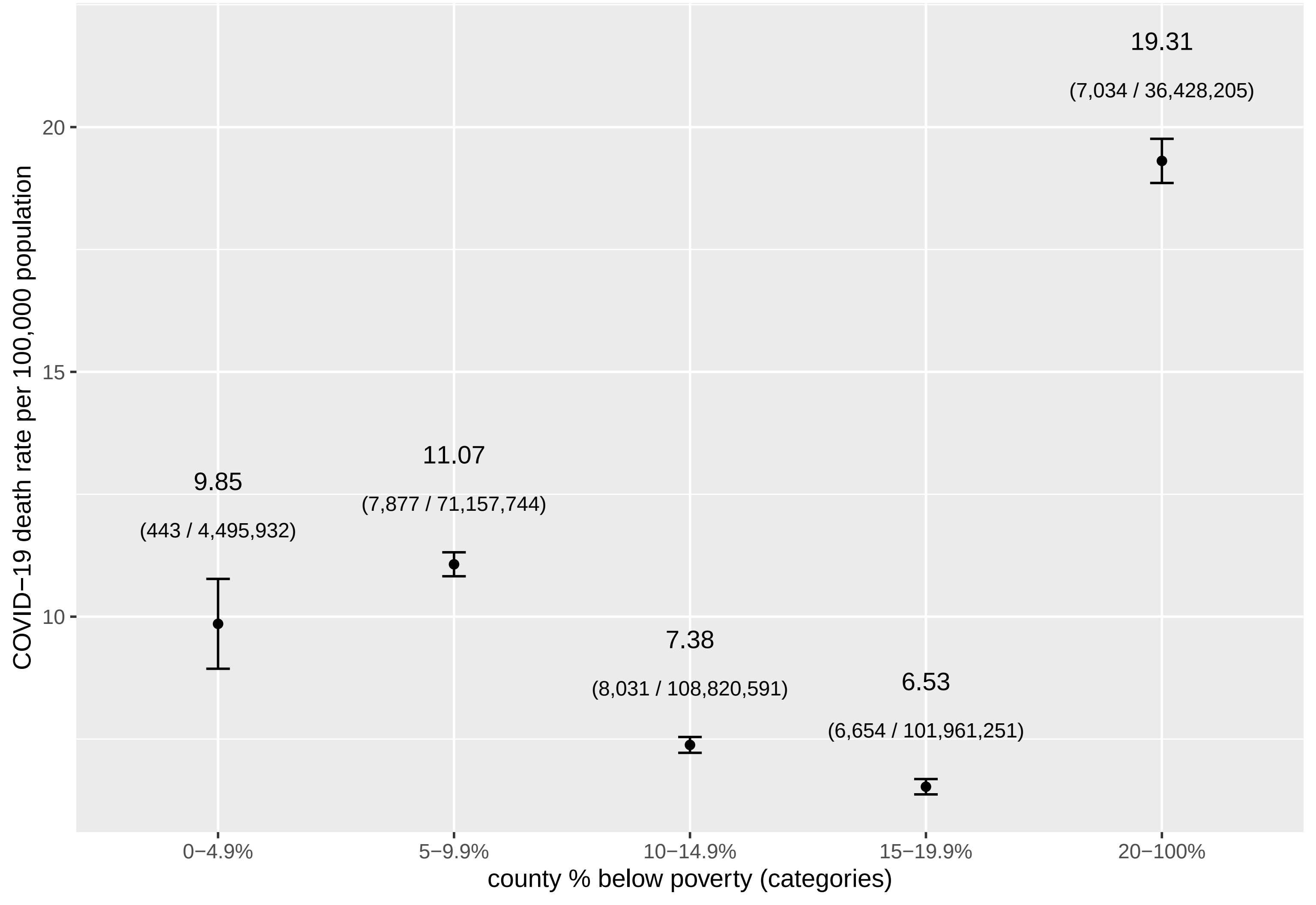


Figure 1b: US COVID−19 deaths per 100,000 population
by county Index of Concentration at the Extremes (white/black race + income)
(as of 4.16.2020)

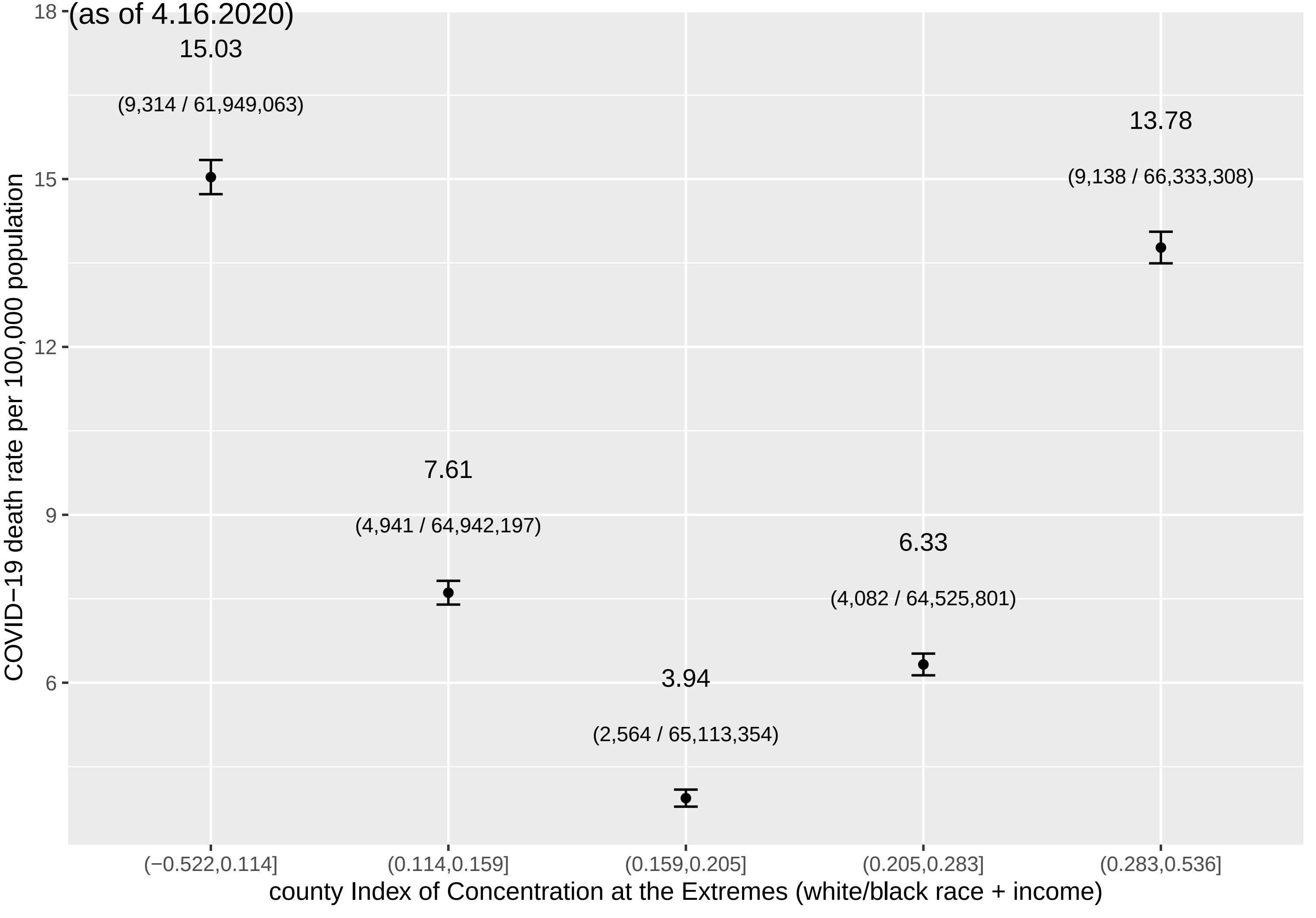


Figure 1c: US COVID−19 deaths per 100,000 population
by county % crowding (>1 person per room) (as of 4.16.2020)

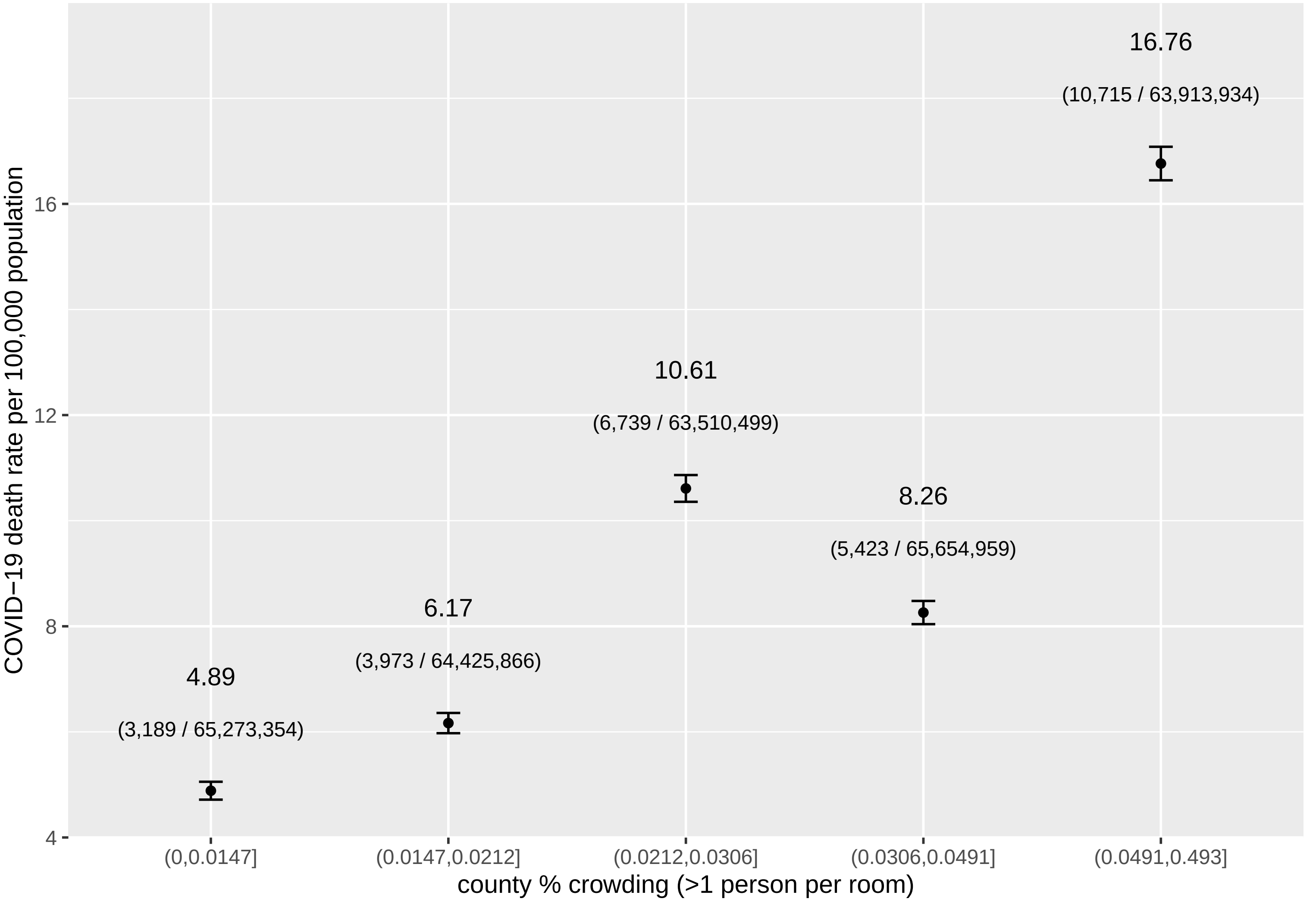


Figure 1d: US COVID−19 deaths per 100,000 population by county % population of color (as of 4.16.2020)

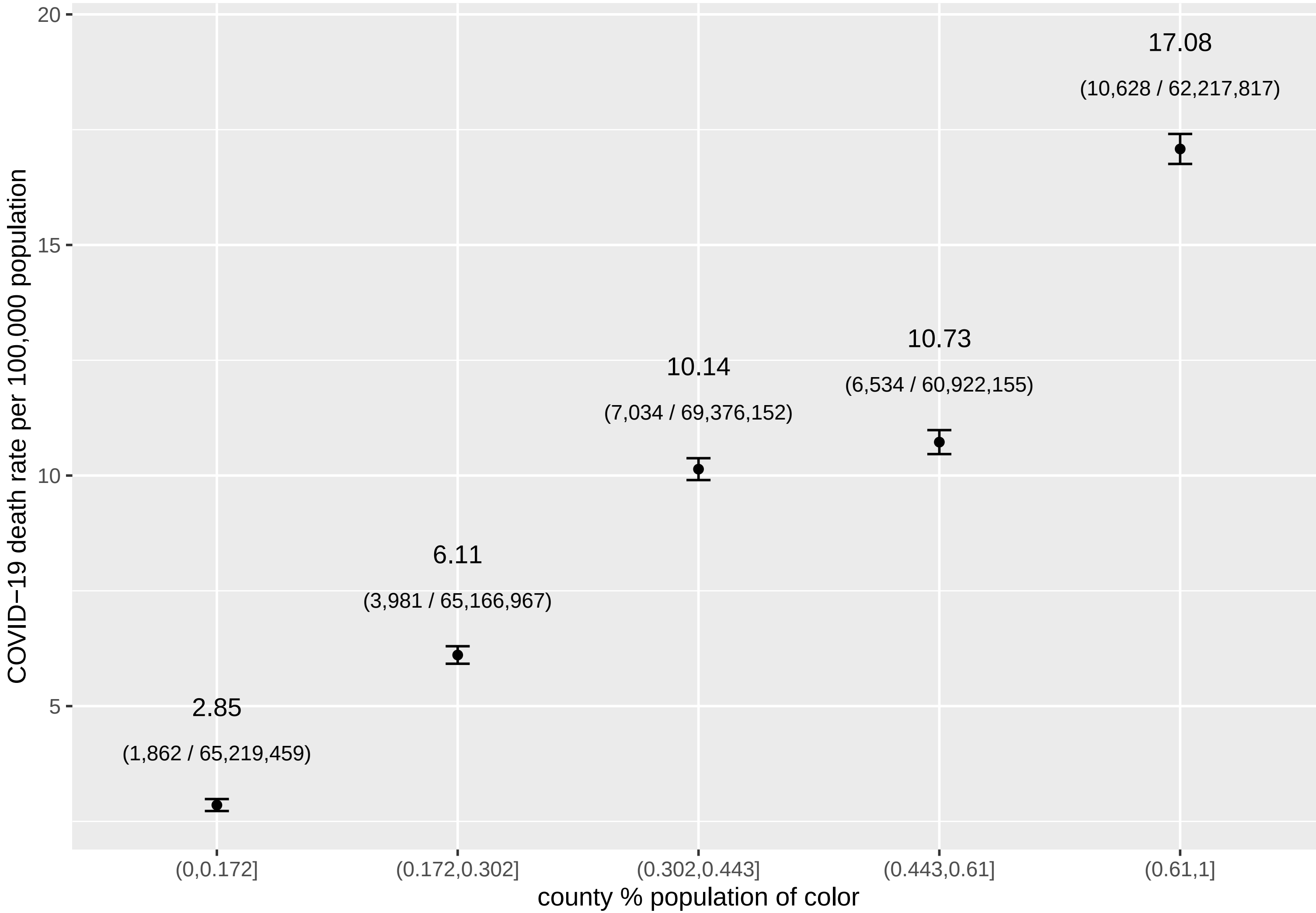
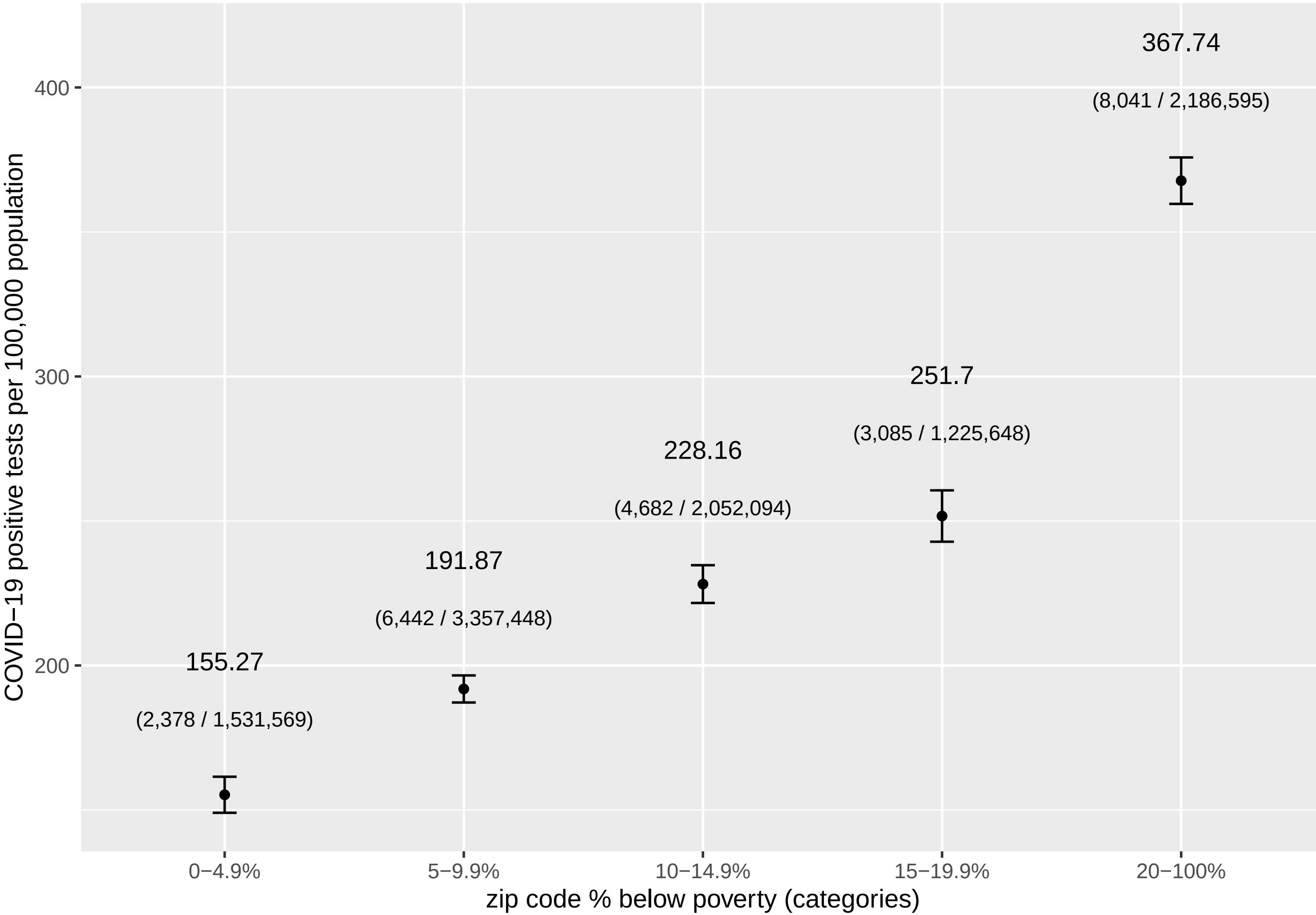


Figure 2a: Illinois COVID−19 confirmed cases per 100,000 population by ZIP code % below poverty (categories) (as of 4.16.2020)



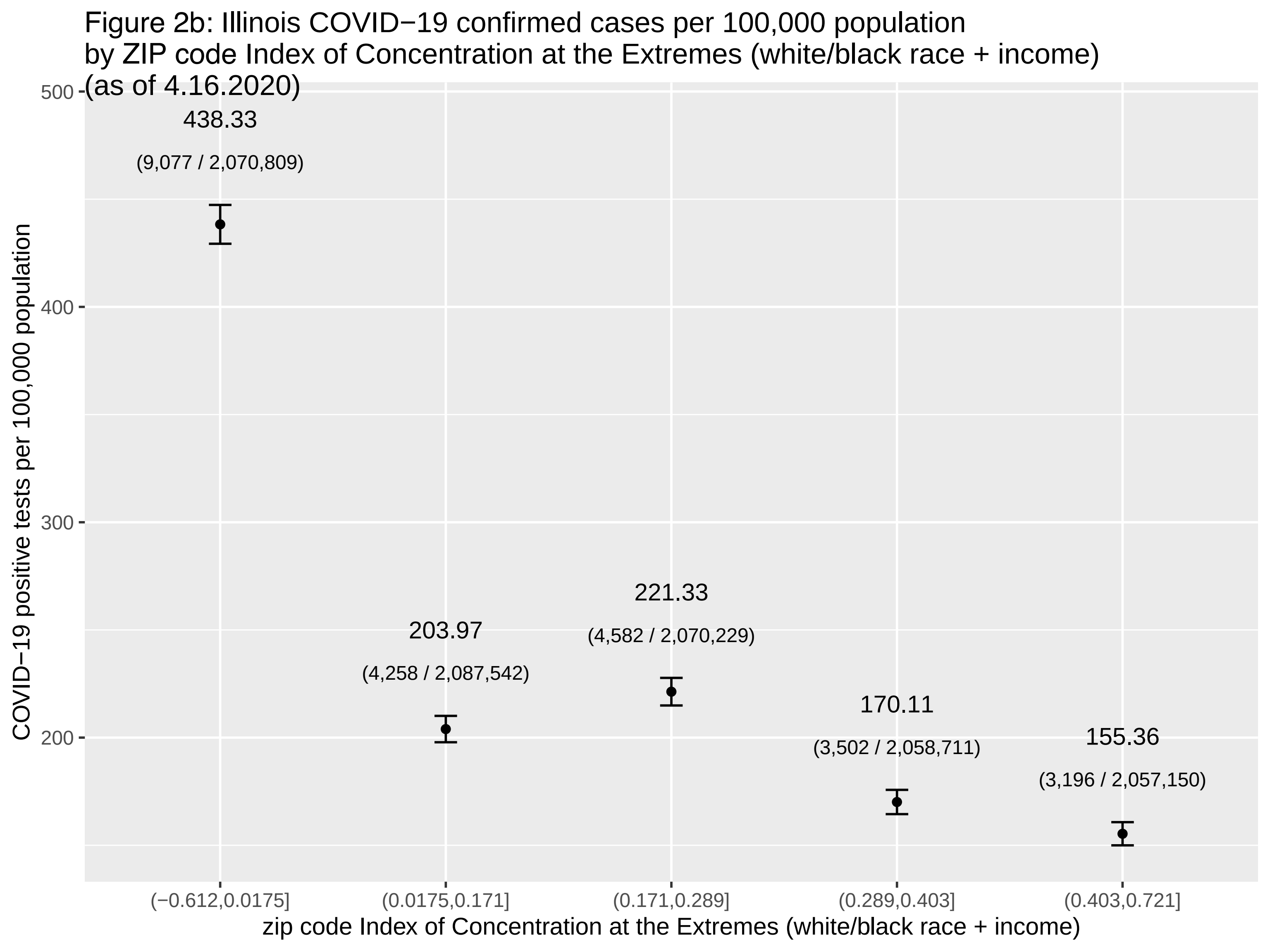


Figure 2b: Illinois COVID−19 confirmed cases per 100,000 population
by ZIP code Index of Concentration at the Extremes (white/black race + income)
(as of 4.16.2020)
COVID−19 positive tests per 100,000 population
500
400
300
200
438.33
(9,077 / 2,070,809)
203.97
(4,258 / 2,087,542)
221.33
(4,582 / 2,070,229)
170.11
(3,502 / 2,058,711)
155.36
(3,196 / 2,057,150)
(−0.612,0.0175]
(0.0175,0.171]
(0.171,0.289]
(0.289,0.403]
(0.403,0.721]
zip code Index of Concentration at the Extremes (white/black race + income)

Figure 2c: Illinois COVID−19 confirmed cases per 100,000 population
by ZIP code % crowding (>1 person per room) (as of 4.16.2020)

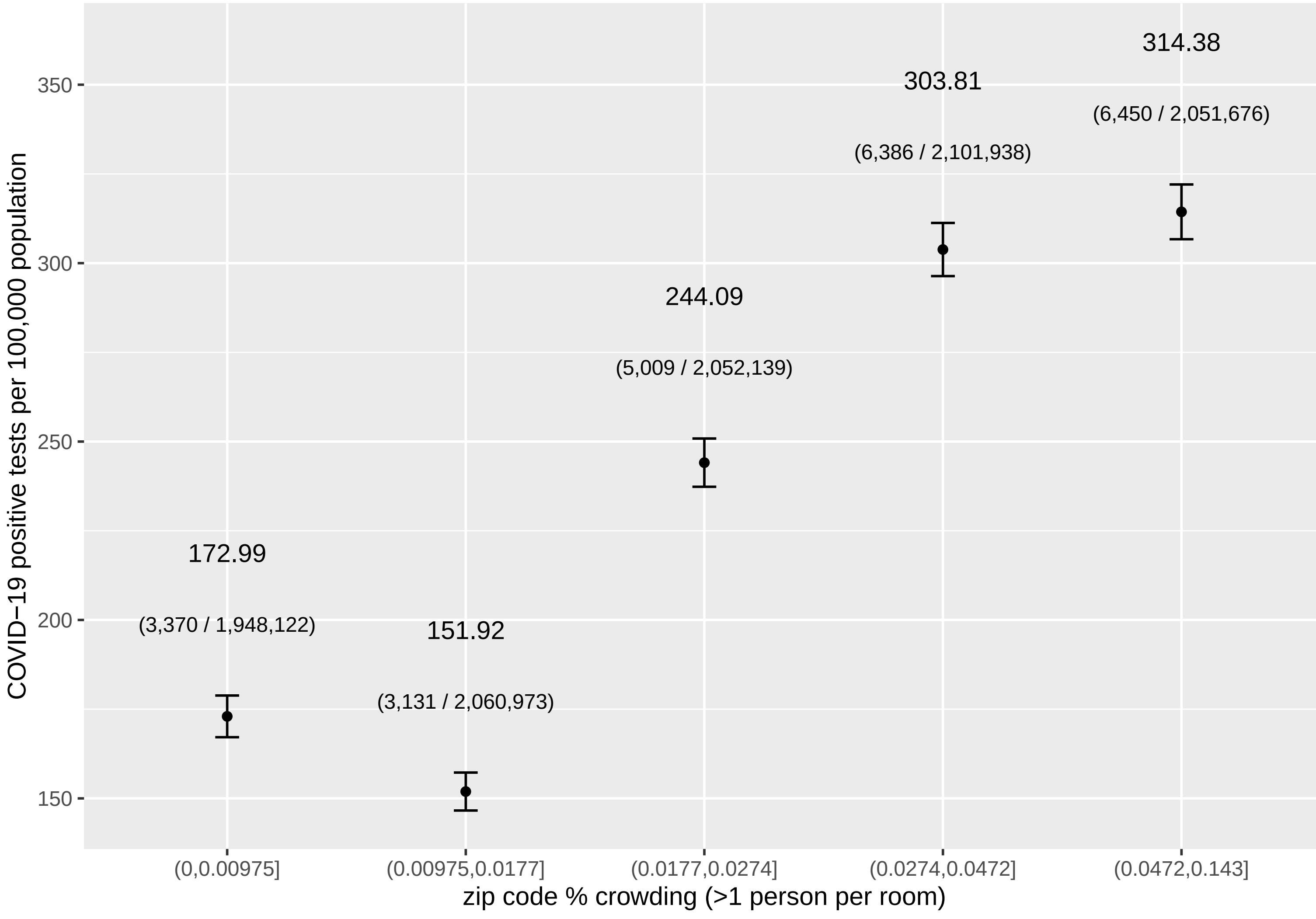


Figure 2d: Illinois COVID−19 confirmed cases per 100,000 population by ZIP code % population of color (as of 4.16.2020)

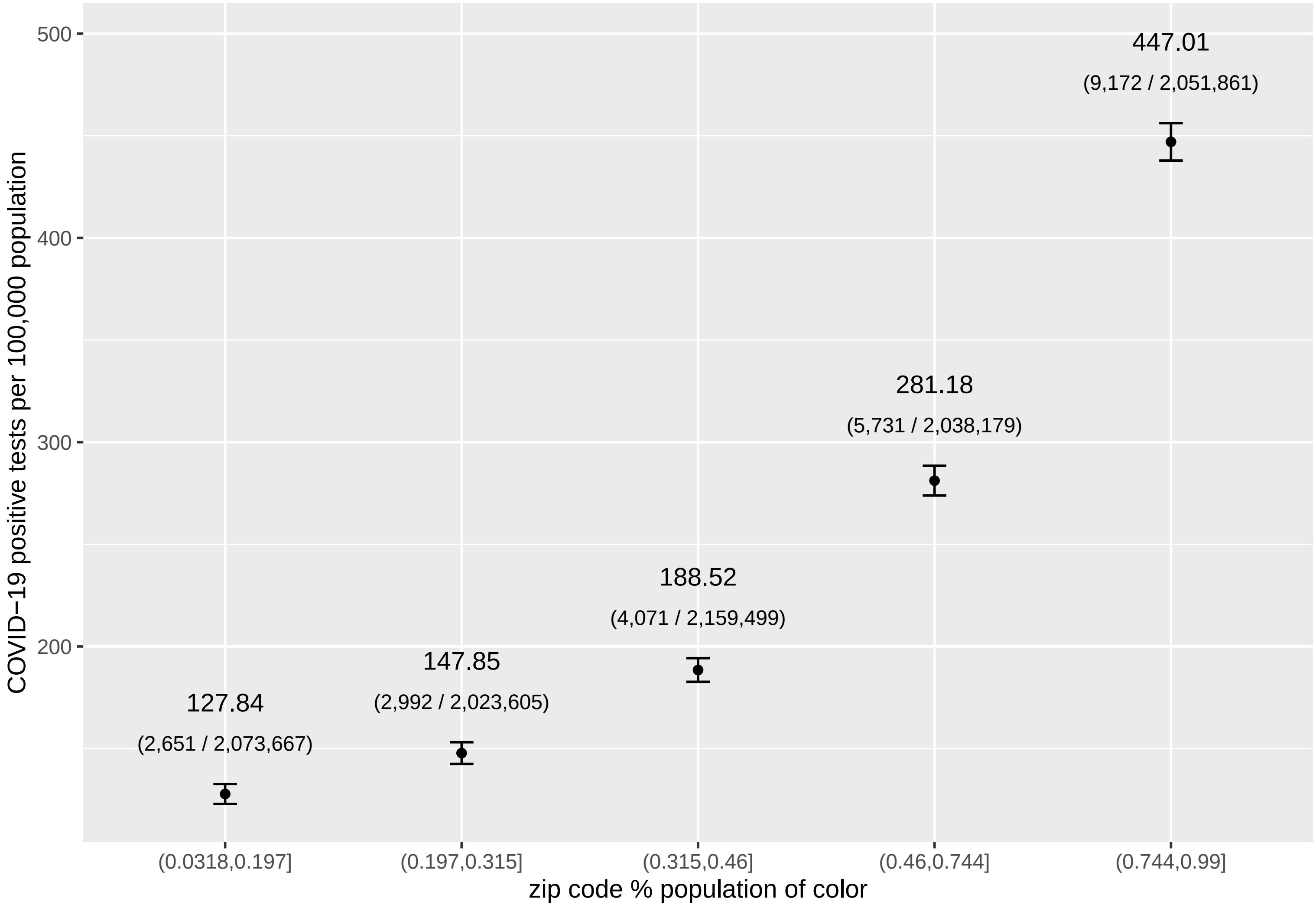


Figure 3a: NYC COVID−19 positive tests per 100,000 population
by ZIP code % below poverty (categories) (as of 4.16.2020)

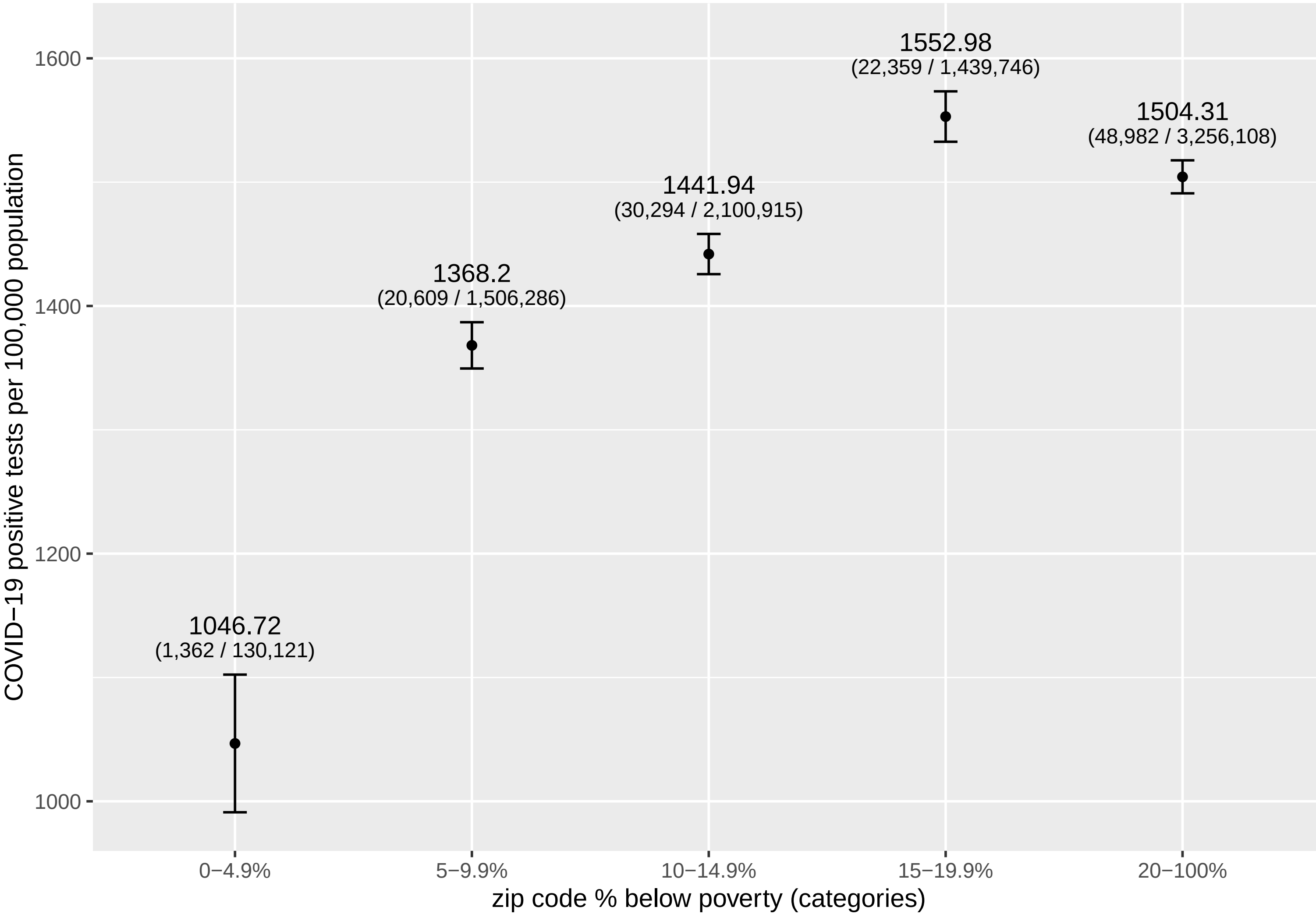


Figure 3b: NYC COVID−19 positive tests per 100,000 population
by ZIP code Index of Concentration at the Extremes (white/black race + income)
(as of 4.16.2020)

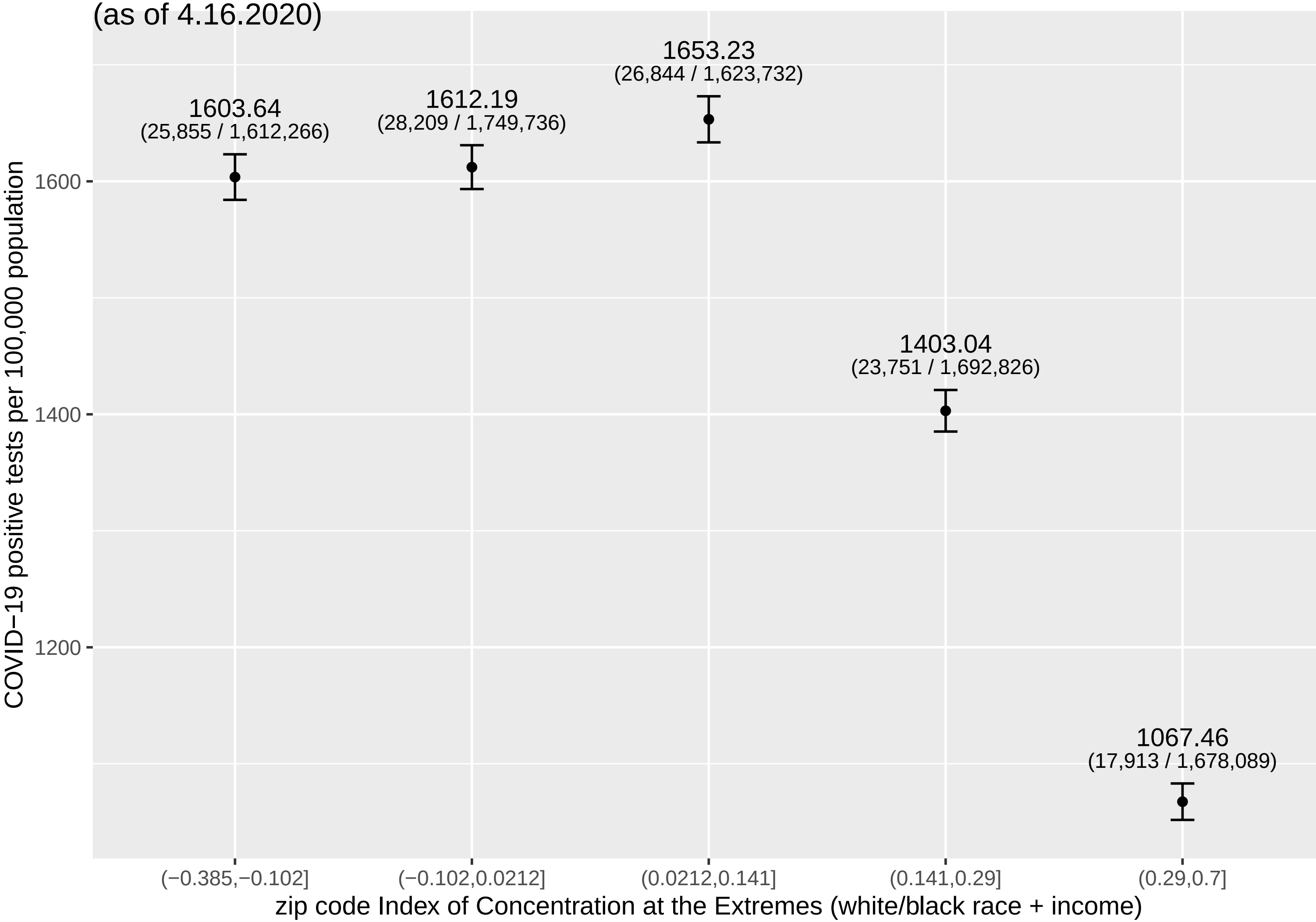


Figure 3c: NYC COVID−19 positive tests per 100,000 population
by ZIP code % crowding (>1 person per room) (as of 4.16.2020)

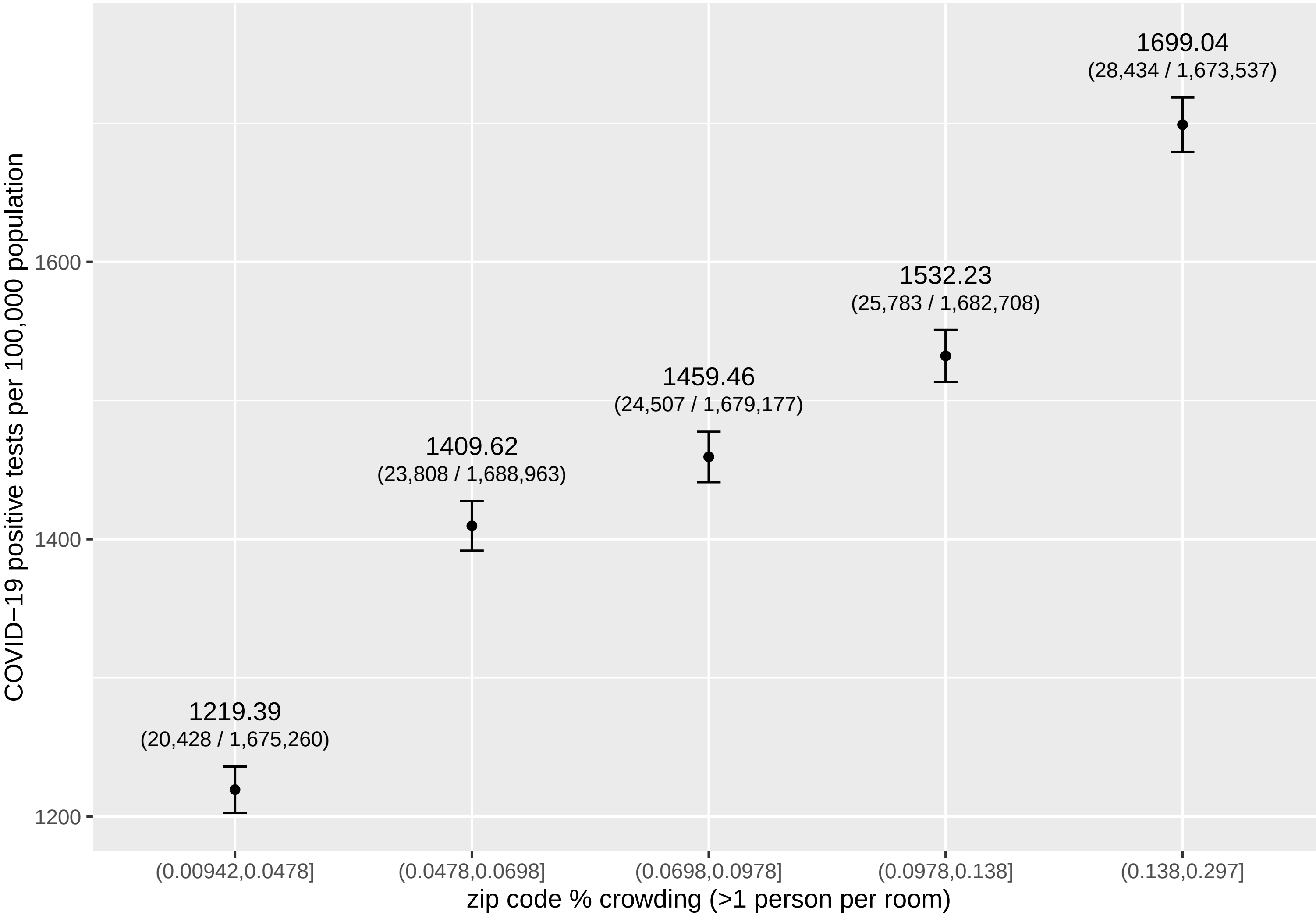


Figure 3d: NYC COVID−19 positive tests per 100,000 population
by ZIP code % population of color (as of 4.16.2020)

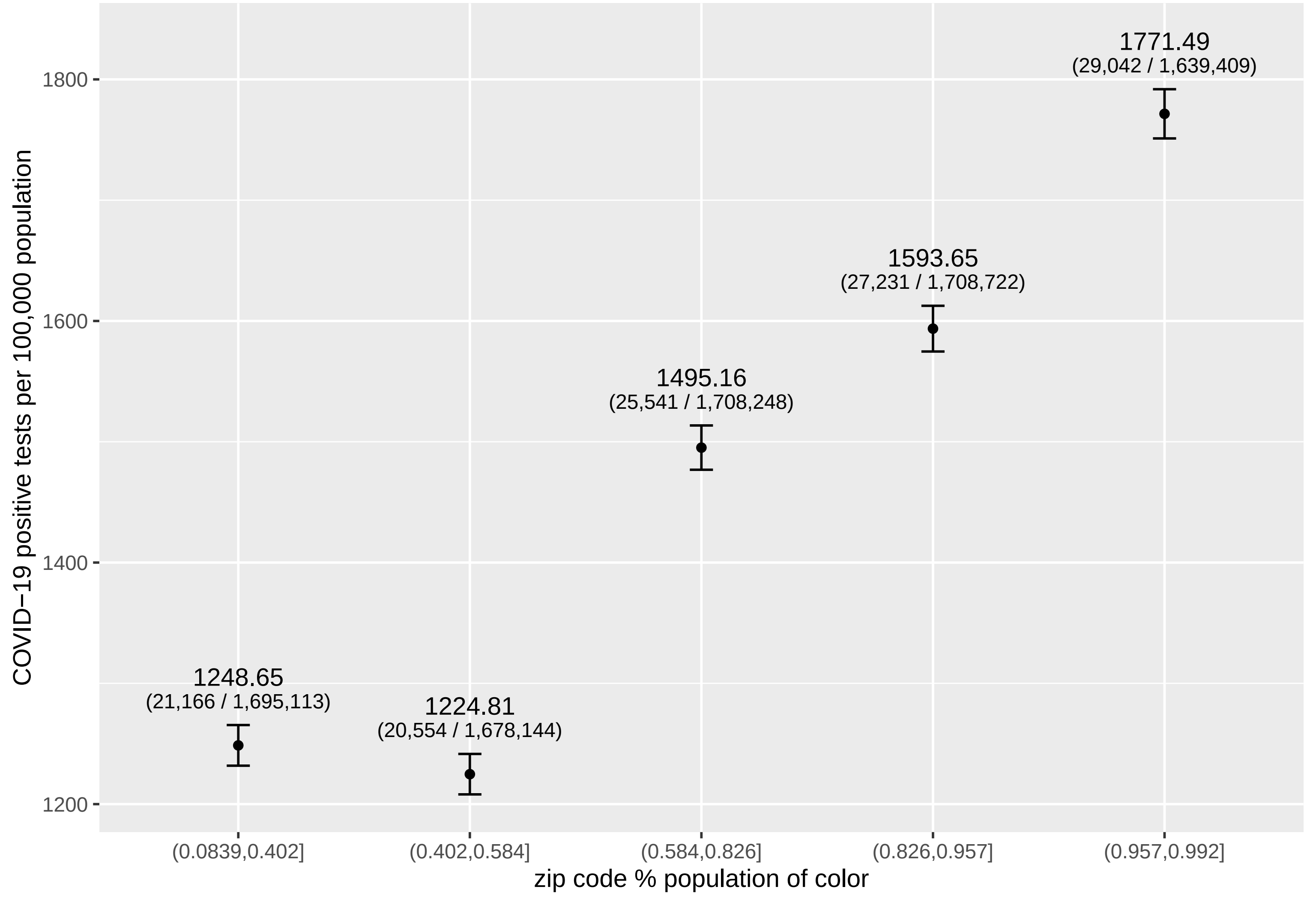


Figure A.1a: US COVID−19 cases per 100,000 population
by county % below poverty (categories) (as of 4.16.2020)

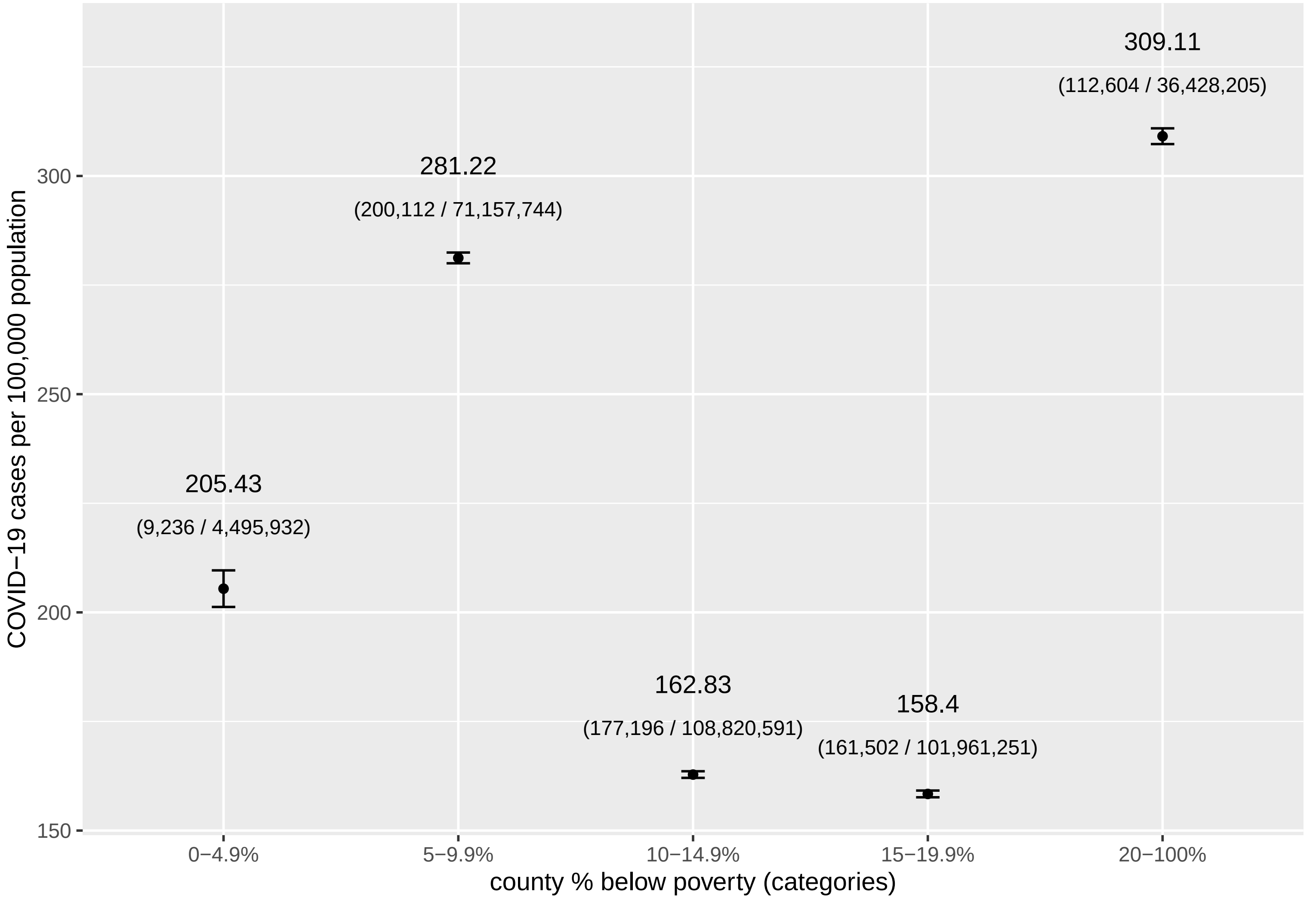


Figure A.1b: US COVID−19 cases per 100,000 population
by county Index of Concentration at the Extremes (white/black race + income)
(as of 4.16.2020)

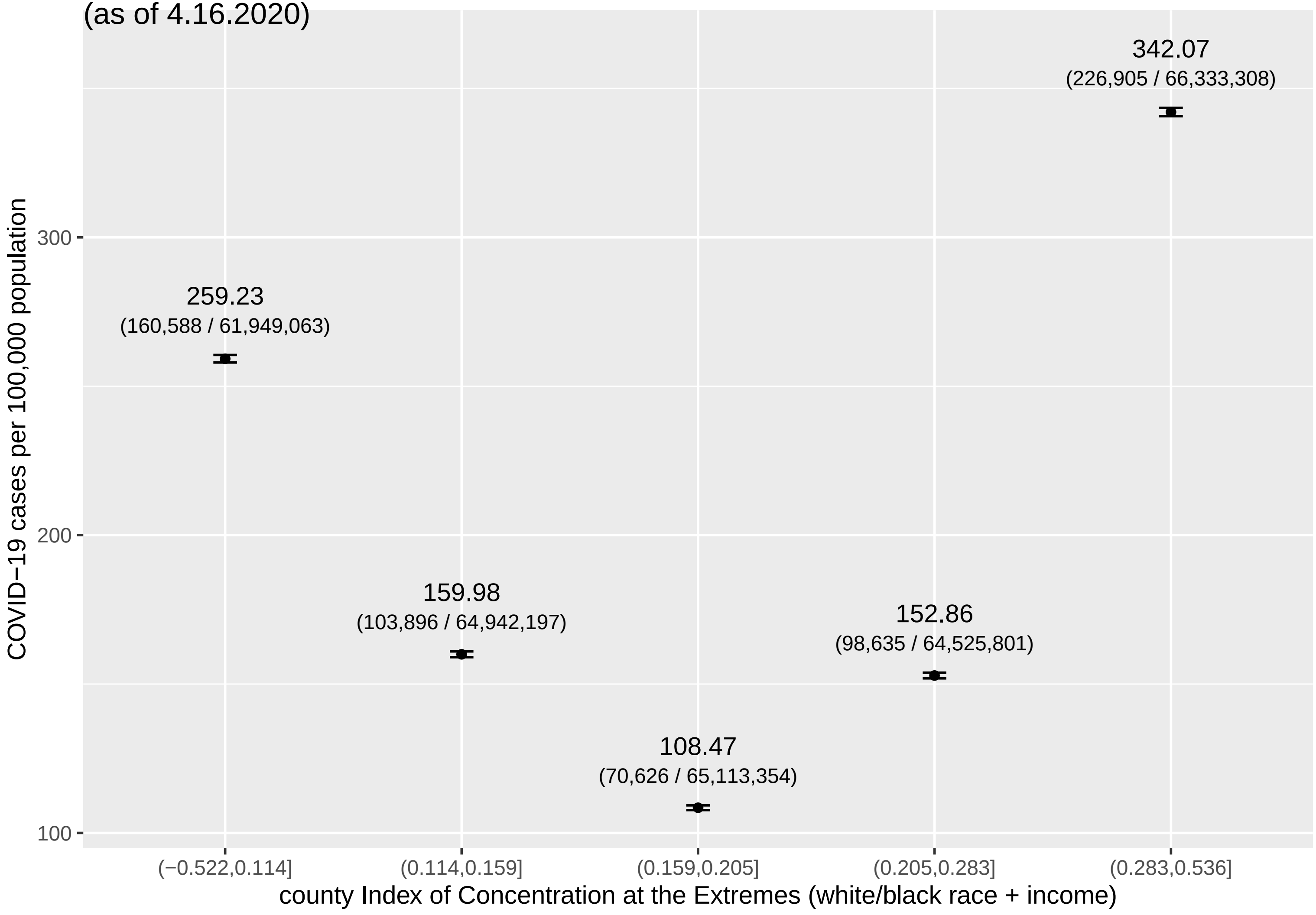


Figure A.1c: US COVID−19 cases per 100,000 population by county % crowding (>1 person per room) (as of 4.16.2020)

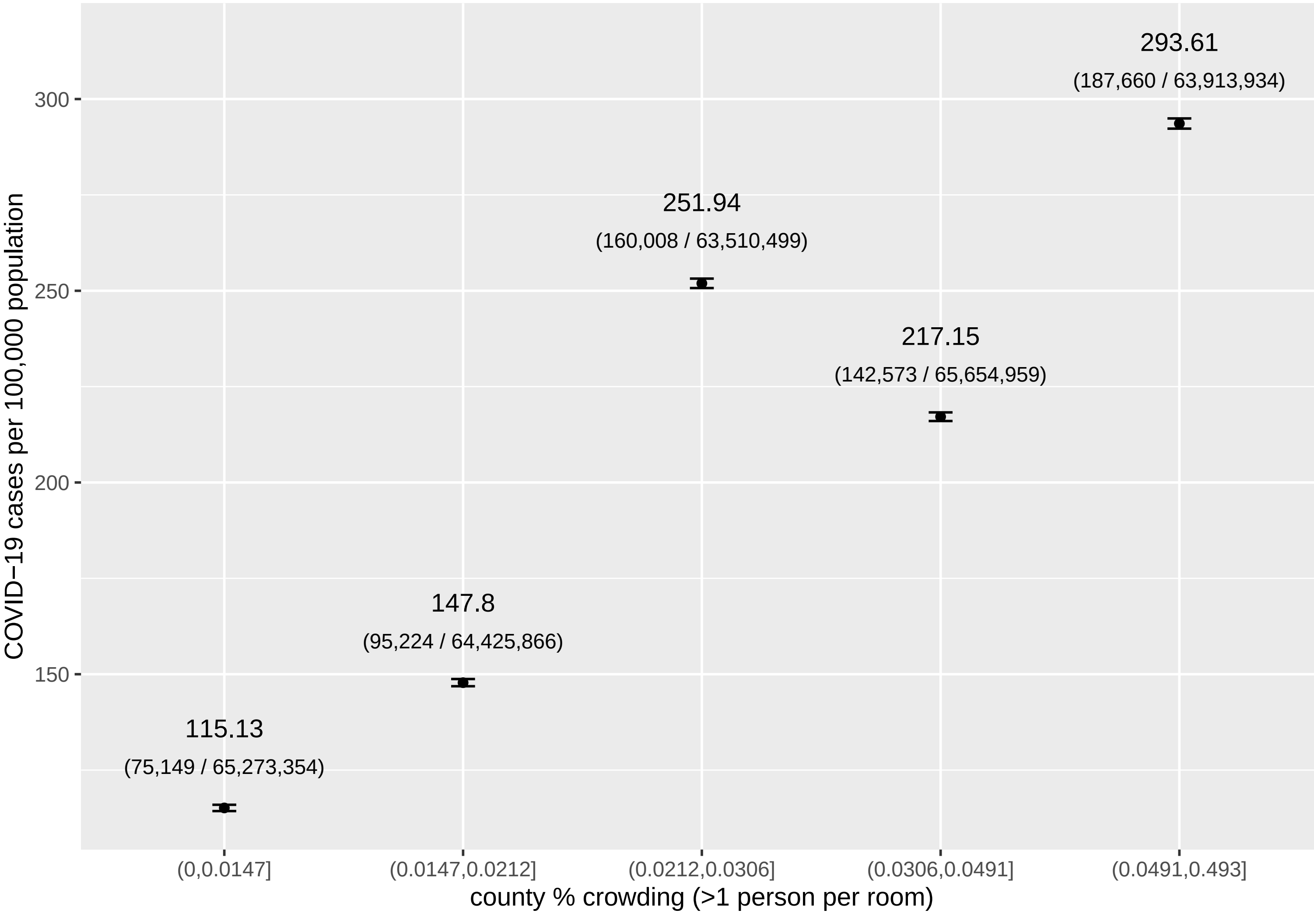


Figure A.1d: US COVID−19 cases per 100,000 population
by county % population of color (as of 4.16.2020)