Shift in racial communities impacted by COVID-19 in California

Raphael E Cuomo

University of California San Diego, La Jolla, California, USA

Correspondence to
Raphael E Cuomo, 9500 Gilman Drive #0170P La Jolla, CA 92093-0170, USA; racuomo@ucsd.edu

ABSTRACT

Introduction Since the first case of COVID-19 was recorded in California, the geospatial distribution of disease cases has fluctuated over time. Given documented racial disparities in other parts of the country, longitudinal convergence of COVID-19 rates around race groups warrants assessment.

Methods County-level cases for COVID-19 were collected from the Johns Hopkins University, and racial distributions were collected from the American Community Survey. Pearson's correlation coefficients were computed for each day since COVID-19 was first reported in California, and the longitudinal distribution of each race-specific set of correlation coefficients was assessed for stationarity, linear trend and exponential trend.

Results Earlier in the outbreak, the distribution of COVID-19 was most highly correlated with Asian American communities; after approximately 100 days, the distribution of COVID-19 most closely resembled that of African American communities. For every day in this dataset, the county-level distribution of COVID-19 was negatively correlated with the distribution of White American communities in California.

Discussion The geospatial distribution of COVID-19 in California has increasingly resembled that of African American communities within the state. Further study should be conducted to characterise potentially disproportionate impacts of the COVID-19 pandemic across race groups.

Check for updates

INTRODUCTION

Coronavirus-19 (COVID-19) presents a unique challenge for public health agencies in California. The burden of this pandemic is both pronounced, with tens of thousands of potential deaths, and rapidly fluctuating.[1] The consequence of disease burden stands squarely at the intersection of prevailing community behaviours and public health readiness.[2]

With a population of approximately 40 million people, California is home to a large number of distinct communities. Though all of these communities are the supposed equal beneficiaries of the state public health system, appreciable inequity is clearly evident. In mid-2020, one of the most salient factors dividing local communities, unfortunately, remains racial composition. Individuals in the African American race group are only half as likely to have a college degree than their White counterparts.[3] Also, African American males at the age of 65 can expect to live 10 years less than their White counterparts.[4] These statistics suggest that the effects of the US education, healthcare and public health systems have not had an equitable impact across races.

At the time of this writing, it is currently believed that COVID-19 has existed in the USA for less than 6 months.[5] Despite this short time frame, a number of racial and ethnic disparities relating to COVID-19 have already been recorded. Reports from early May documented that mortality rates among White American populations were dramatically lower than those of African American and Latino populations in various urban communities.[6] An investigation of health records from a healthcare system in Northern California revealed that Africans Americans with COVID-19 exhibited 2.7 times the odds of hospitalisation when compared with the odds of hospitalisation for White Americans. Furthermore, though only 20% of counties in the USA have a higher proportion of African Americans than the national average, these counties accounted for over half of COVID-19 cases and deaths as of mid-April.[7]

Concordantly, the spread of COVID-19 in California may be dissimilar across races. To investigate this, I assessed differences in the relationships between COVID-19 burden and the racial composition of counties in California.

METHODS

Data collection

Data on county-level confirmed cases of COVID-19 were obtained from the GitHub repository of the Johns Hopkins University Centre for Systems Science and Engineering (JHU CSSE) for each calendar day, starting with the first day that a case was recorded in California and ending 153 days into the calendar year.[8] These freely available data are updated daily with new records from state and local health departments. Data on county-level proportion race composition from 2018 were obtained from the American Community Survey, with race variables corresponding to the proportion in a county identifying as solely belonging to the queried racial identification.[9] The final dataset included values for COVID-19 cases for all 58 counties in California across 125 days, as well as county-level proportion for the American Indian, Asian American, African American, White American and Hawaiian/Pacific Islander race groups.

Data analysis

Pearson's correlation coefficients were computed between each racial distribution and COVID-19 rates on each day, thereby yielding 625 correlation coefficients from 7540 values. R (version 3.6.3) was used for statistical analysis and ArcGIS (version 10.7) was used for geospatial visualisation. Stationarity was assessed using the Augmented Dicky–Fuller test and non-stationary trends were

quantified using linear and exponential regression modelling. For exponential modelling, the Nagelkerke R^2 coefficient was computed for comparison of model fit.

RESULTS

On March 15th, only 50 days after COVID-19 was first recorded in California, the geospatial distribution of the disease (figure 1A) very closely resembled the distribution of Asian Americans (figure 1B) and was largely absent in counties highly populated by White Americans (figure 1C). Conversely, 125 days later on June 1st, the distribution of COVID-19 (figure 1D) more closely resembled the distribution of African Americans (figure 1E) than Asian Americans (figure 1B), while still bearing little resemblance to the distribution of White Americans (figure 1C).

In the earlier phases of the pandemic in California, Asian American communities were disproportionately burdened by COVID-19 infection (figure 2). Around mid-April, the areas most impacted by COVID-19 infection shifted to African American communities. Consistently throughout the observed period, areas which were more highly populated by White individuals had lower burden of COVID-19 than areas less populated by White individuals.

Differences were observed in the longitudinal trend of Pearson's correlation coefficients between the geospatial distribution of COVID-19 and the county-level distribution of race groups. Only the distribution of the Hawaiian and Pacific Islander race category exhibited statistically significant stationarity in its relationship to the fluctuating distribution of COVID-19 (p=0.0368). Though stationarity was not observed for the American Indian race category, neither linear nor exponential modelling resulted in statistically significant longitudinal associations. Conversely, linear and exponential modelling both exhibited statistically significant longitudinal relationships for correlation coefficients between COVID-19 distributions and the African American (linear p<0.0001; exponential p<0.0001), Asian American (linear p<0.0001; exponential p<0.0001) and White American (linear p=0.0117; exponential p=0.0345) race categories, with similar fit for both models. However, neither model involving the White American racial covariate had good fit (linear R^2=0.0503; Nagelkerke R^2=0.0351), whereas fit was appreciably improved for Asian American (linear R^2=0.2144; Nagelkerke R^2=0.3253) and African American (linear R^2=0.6701; Nagelkerke R^2=0.6686) models. Coefficients for both linear and exponential models suggested longitudinal decreases between COVID-19 and the White American and Asian American race distributions, but longitudinal increases between COVID-19 and the African American covariate.

DISCUSSION

COVID-19 in California appears to disproportionately impact minority communities. However, the typology of minority
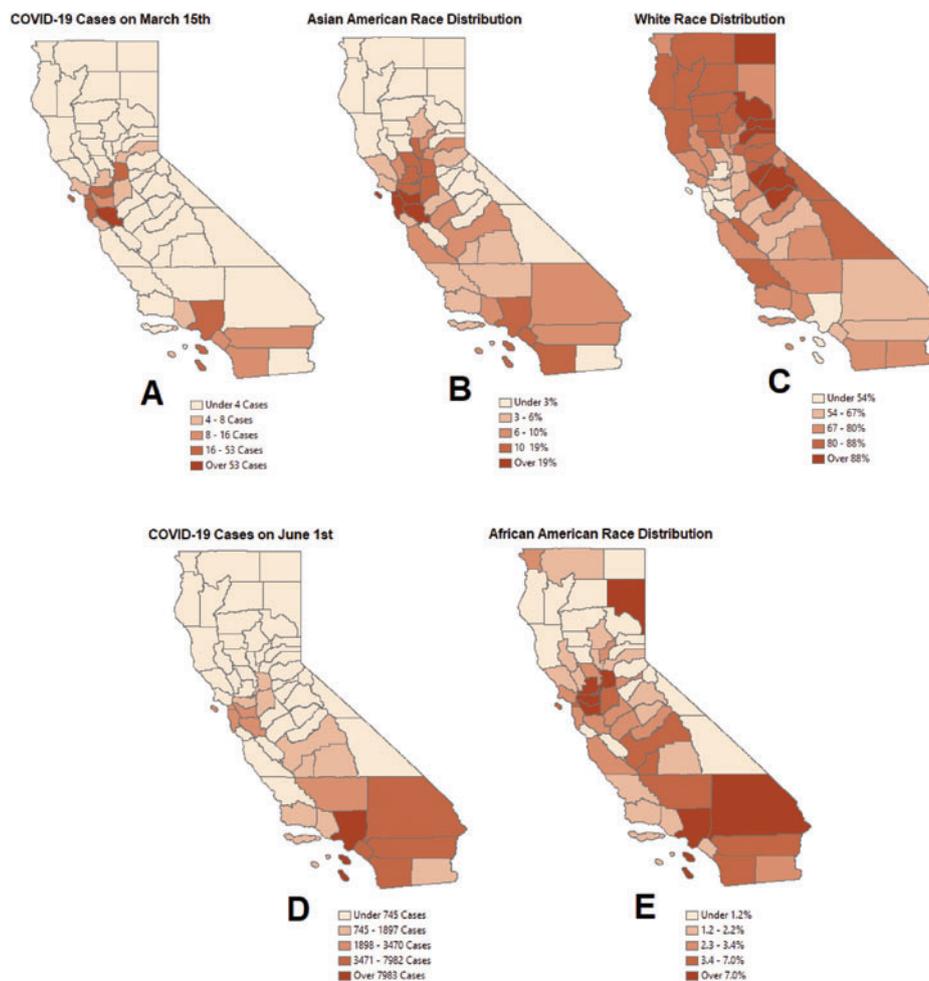

Figure 1 By county, California distributions of (A) COVID-19 cases on 15 March 2020; (B) percentage Asian Americans; (C) percentage White Americans; (D) COVID-19 cases on 1 June 2020; and (E) percentage African Americans.
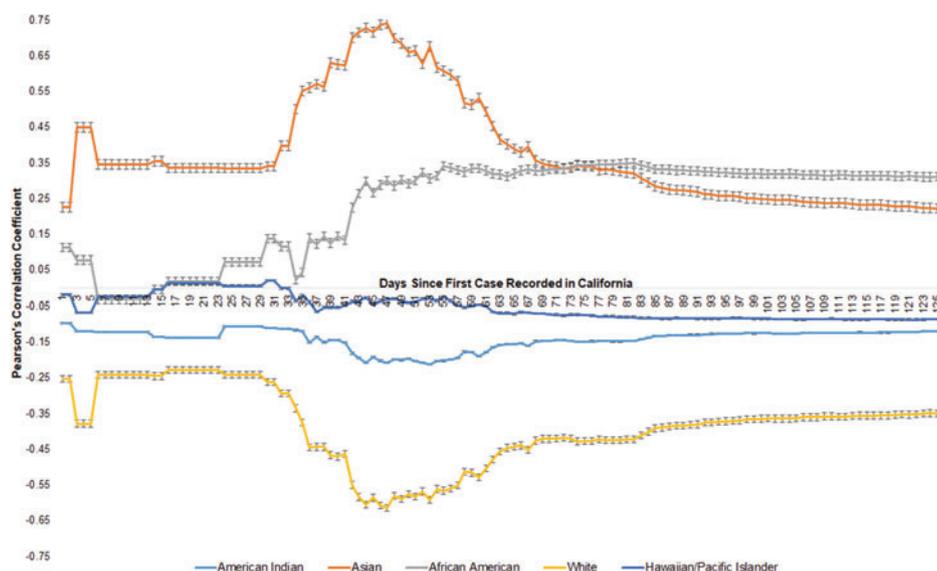
Figure 2 Longitudinal fluctuations in the correlation between race groups and COVID-19 cases, for county-level geospatial distributions, since the first case of was recorded in California. Bars denote SE.

communities most impacted did not remain static over time. During the first 100 days of the pandemic in California, the distribution of COVID-19 increasingly resembled the distribution of African Americans. The reasons for this are presently unclear and warrant further study; however, this appears more likely to be the result of socio-economic-political factors rather than SARS-CoV-2 virology. Indeed, it has been argued that racial discrepancies in *privilege* may account for differences in adherence to many of the preventative measures outlined in present recommendations for COVID-19 transmission, including social distancing strategies such as working from home and accepting work furloughs.[10]

The legacy of systemic marginalisation offers additional suggestions for the observed spatial disparity between race groups. African Americans are disproportionately likely to be employed in front-line occupations as 'essential workers', including as retail grocers, custodians or public transit employees.[11] These positions may require close contact with others, thereby facilitating the transmission of COVID-19. Furthermore, African Americans are more likely to live in areas with high housing density, which involve close physical distance between separate homes.[12] In addition, homes in high-density areas may themselves have much less living space than those in lower-density areas. Therefore, an individual with COVID-19 who lives in an area with higher housing density may be much more likely to transmit COVID-19 to his/her family members and neighbours.

Data through mid-March (ie, approximately the first 50 days of these data) were relatively unaffected by government directives. California's statewide stay-at-home order was issued on March 19th. Some county- or city-level governments issued similar directives prior to the statewide stay-at-home order. These included the counties of Alameda, Contra Costa, Marin, Monterey, San Benito, San Francisco, San Mateo, Santa Cruz, Santa Clara and Sonoma.[13] However, county- and city-level directives were issued only within the few days preceding the statewide stay-at-home order.

Geospatial studies may be useful in detecting broad disparities which would not be easily discernible in a healthcare setting. For example, this approach has been used to identify racial disparities in digestive cancer incidence in Texas,[14] Lyme disease in Maryland[15] and amyloidosis mortality in the United States.[16] A dual longitudinal and geospatial approach, as used in this study, has also been leveraged to describe worsening racial disparity in breast cancer among counties across the United States.[17]

The fluctuating nature of infectious disease epidemics, especially as pertain to those where societal reaction may influence disease transmission, warrants continued scrutiny of geospatial variations at all levels. Furthermore, given recent mass gatherings over race-related issues, serious consideration should be given to preventing an exacerbation of racial inequity in COVID-19 burden in California.

Limitations

Guidance for COVID-19 testing may have shifted as testing capacity increased later in the study period. Variation in testing capacity may have influenced the case identification rate throughout the study period. Also, the existence of relatively high cases in a county with high proportions of a given race group does not necessitate that the individuals in that race group were infected with COVID-19. Moreover, the comparison of COVID-19 cases and racial density within counties could not be observed longitudinally as these data were not made publically available. Lastly, as the data analysed are derived from units within a contiguous spatial plane, it should be noted that the presence of spatial autocorrelation may introduce a bias to correlation coefficients computed in this study. Nevertheless, the results of this study suggest that further research should be conducted to better characterise the potentially fluctuating burdens of COVID-19 across race categories in California.

What is already known on this subject

► Geospatial approaches have been leveraged to identify racial disparities in numerous diseases, both communicable and non-communicable. Racial disparities in incidence, hospitalisation, and mortality from COVID-19 have been documented. However, geospatial approaches have not been used to discern longitudinal convergence among communities within California.

Short report

What this study adds

► This study identifies the longitudinal trend describing the disproportionate impacts of COVID-19 on specific race groups in California.

Map disclaimer The depiction of boundaries on the map(s) in this article does not imply the expression of any opinion whatsoever on the part of BMJ (or any member of its group) concerning the legal status of any country, territory, jurisdiction or area or of its authorities. The map(s) are provided without any warranty of any kind, either express or implied.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; internally peer reviewed.

ORCID iD
Raphael E Cuomo http://orcid.org/0000-0002-8179-0619

REFERENCES

1 Gates B. Responding to COVID-19: a once-in-a-century pandemic? *N Engl J Med* 2020;382:1677–9.
2 Pan A, Liu L, Wang C, *et al.* Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. *JAMA* 2020;323:1915.
3 Deming D, Dynarski S. College aid. In: Levine Phillip B, Zimmerman David J (eds.). *Targeting investments in children: fighting poverty when resources are limited*. University of Chicago Press, 2010: 283–302.
4 Guralnik JM, Land KC, Blazer D, *et al.* Educational status and active life expectancy among older blacks and whites. *N Engl J Med* 1993;329:110–16.
5 Ueda M, Martins R, Hendrie PC, *et al.* Managing cancer care during the COVID-19 pandemic: agility and collaboration toward a common goal. *J National Compr Cancer Netw* 2020;1:1–4.
6 Webb Hooper M, Nápoles AM, Pérez-Stable EJ. COVID-19 and racial/ethnic disparities. *JAMA* 2020;323:2466–7.
7 Millett GA, Jones AT, Benkeser D, *et al.* Assessing differential impacts of COVID-19 on black communities. *Ann Epidemiol* 2020;47:37–44.
8 CSSE JHU. Novel coronavirus (COVID-19) cases. *GitHub Repository*. Available https://github.com/CSSEGISandData/COVID-19 (accessed 5 Apr 2020)
9 U.S. Census Bureau. *2014-2018 American community survey 5-year public use microdata samples*. 2019. Available https://data.census.gov/cedsci (accessed 5 Apr 2020)
10 Yancy CW. COVID-19 and African Americans. *JAMA* 2020;323:1891–2.
11 Dorn AV, Cooney RE, Sabin ML. COVID-19 exacerbating inequalities in the US. *Lancet (London, England)* 2020;395:1243–4.
12 DiMaggio C, Klein M, Berry C, *et al.* Blacks/African American communities are at highest risk of COVID-19: spatial modeling of New York city zip code-level testing results. *Ann Epidemiol* 2020;51:7–13.
13 Shalby C, Lin II RG, Vives R, *et al.* As officials grapple with coronavirus pandemic, more than 8 million Californians are living under shelter-in-place orders. *Los Angeles Times*, 2020.
14 Sparks C. An examination of disparities in cancer incidence in Texas using Bayesian random coefficient models. *PeerJ* 2015;3:e1283.
15 Fix AD, Peña CA, Strickland GT. Racial differences in reported; Lyme disease incidence. *Am J Epidemiol* 2000;152:756–9.
16 Alexander KM, Orav J, Singh A, *et al.* Geographic disparities in reported US amyloidosis mortality from 1979 to 2015: potential underdetection of cardiac amyloidosis. *JAMA Cardiol* 2018;3:865–70.
17 Coughlin SS. Epidemiology of breast cancer in women. In: Ahmad A (ed.). *Breast cancer metastasis and drug resistance*. Springer, 2019: 9–29.