COVID-19 across United States congressional districts

Priyanka N deSouza, S V Subramanian

1Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA, USA
2Harvard Center for Population and Development Studies, Harvard University, Cambridge, MA, USA
3Department of Social and Behavioral Sciences, Harvard T. H. Chan School of Public Health, Boston, MA, USA

INTRODUCTION

As of July 14, 2020, confirmed coronavirus disease 2019 (COVID-19) cases and deaths in the United States (US) were close to 3.4 million, and 135,000, respectively (Fig. 1). One characteristic of the discourse related to the pandemic was the omnipresence of COVID-19 statistics at the state- and county-levels, spurring important analysis that are of public and policy significance. However, a critical geographic unit for which COVID-19 mortality and case statistics have not been presented are the US congressional districts (CDs). The CDs represent “a territorial division of a state from which a member of the US House of Representatives is elected.” The lack of reporting of COVID-19 statistics for CDs is true of other data on population health.

One of the reasons statistics on population health and well-being indicators are not routinely available for CDs is because CDs are not straight-forward aggregations of existing smaller Census-based administrative geographies, such as counties; a unit where a wide array of population health data is routinely available. Yet, having CD-level statistics can substantially enhance the ability of elected officials and the populations they represent to monitor, evaluate, and design programs specific to the needs and vision of the constituents. Other smaller administrative units, such as counties, do not have similar single elected representatives, and thus lack the same level of visibility, and therefore, accountability, that the CD has.
In this study, we provide estimates of COVID-19 cumulative and recent cases, deaths (per 1,000 people) and the case fatality rate (CFR), defined as the ratio of cumulative COVID-19 deaths to cases, for the 436 CDs in the US. The cumulative cases and deaths provide information about the overall burden of COVID-19 in a CD, while COVID-19-related statistics for last month are presented so that elected representatives and their constituents can deliberate their respective reopening strategies.

**METHODS**

All but 15 CDs are not straightforward aggregation of US counties. When only portions of counties lie in a district, it is not a simple aggregation exercise of county rates to a CD, as these portions are irregular in size and shape with different population fractions in them. We used a previously published population-weighted methodology, to derive CD-level COVID-19 statistics from county-level data. The approach we have taken falls under the category of dasymetric mapping techniques that accounts for the spatial distribution of the underlying population for performing such an interpolation.

We used county-level COVID-19 cases and deaths data from January 22, 2020 to July 13, 2020 in the US from USAFacts (https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/, last accessed July 14, 2020) that aggregates official COVID-19 statistics from the Centers for Disease Control and Prevention, state and local-level public health agencies. We focused on COVID-19 statistics of cases, deaths (per 1,000 people) and the case fatality rate (CFR) for the period January 22, 2020–July 13, 2020 (“cumulative”), and COVID-19 cases and deaths (per 1,000 people) for June 13 to July 13, 2020 (“recent”).
We retrieved the population for each of the CDs in effect for the 116th US Congress (January 2019 to January 2021) from the National Historical Geographic Information System, derived from the 2014–2018 American Community Survey (https://www.nhgis.org/, Last accessed July 14, 2020). We retrieved census block and county population data from the US Census Bureau’s 2010 decennial census (latest available at this geographic scale; https://www.nhgis.org/, Last accessed July 14, 2020). Census blocks nest inside both CDs and counties. We assigned county COVID-19 cases and deaths per capita to census blocks in an unweighted manner. We then summed the collection of block-level rates in a CD using the census block/CD population weights to obtain CD-level estimates of the different in every CD. Finally, we categorized cumulative CD-level cases and deaths per 1,000 and CFR, and the number of recent cases and deaths per 1000 residents in CD-level cases and deaths per 1,000 in each CD by decile. We provide a detailed table of the estimates for each CD in Supplementary Table 1.

RESULTS

Distribution of COVID-19 Cases
Cumulative cases per 1,000 residents across CDs varied substantially between 0.73 to 37.94 with a mean of 9.79 (Fig. 2A). States with CDs ranked in the top decile (category 10) with the highest number of cumulative cases per capita were New York (18 CDs/total CDs in the state: 27), New Jersey (7 CDs/total CDs: 12) and Arizona (5 CDs/total CDs: 9), Illinois (4 CDs/total CDs: 18), Florida (4 CDs/total CDs: 27), Massachusetts (2 CDs/total CDs: 9), Louisiana (1 CD/total CDs: 6), Maryland (1 CD/total CDs: 8) and Washington (1 CD/total CDs: 10) (Fig. 3A). States with CDs in the bottom decile were California (5 CDs/total CDs: 53) and Pennsylvania (5 CDs/total CDs: 18), among others.

Recent COVID-19 cases per 1,000 residents ranged between 0.25 to 15.71 with a mean of 3.65 (Fig. 2B). States with CD ranked in the top decile with the greatest increase in cases per capita in this month were Florida (19 CDs/total CDs in the state: 27), Arizona (9 CDs/total CDs: 9), Texas (5 CDs/total CDs: 36), South Carolina (3 CDs/total CDs: 7), Alabama (2 CDs/total CDs: 7), Tennessee (2 CDs/total CDs: 9), Louisiana (1 CD/total CDs: 6), Mississippi (1 CD/total CDs: 4), Washington (1 CD/total CDs: 10) (Fig. 3B). States with CDs in the bottom decile were Connecticut (5 CDs/total CDs: 5), New York (7 CDs/total CDs: 27) and New Jersey (5 CDs/total CDs: 12), among others.

Distribution of COVID-19 Deaths
Cumulative deaths per 1,000 residents across CDs, varied between 0.014 and 3.23 with a mean of 0.41 (Fig. 2C). Of the 43 CDs ranked in the top decile, states with the greatest number of CDs in this category were New York (17 CDs/total CDs in the state: 27), New Jersey (10 CDs/total CDs: 12), Massachusetts (7 CDs/total CDs: 9), Connecticut (4 CDs/total CDs: 5), Michigan (3 CDs/total CDs: 14), Louisiana (1 CD/total CDs: 6), and Pennsylvania (1 CD/total CDs: 18) (Fig. 3C). CDs in the bottom decile were in California (7 CDs/total CDs: 53) and Texas (4 CDs/total CDs: 36). States with CDs in the lowest decile were California (8 CDs/total CDs: 53) among others.

Recent COVID-19 deaths per 1,000 residents ranged between 0.001 and 0.46 with a mean of 0.074 (Fig. 2D). States with CDs ranked in the top decile with the greatest increase in deaths per capita occurred in New Jersey (12 CDs/total CDs in the state: 12), New York (5 CDs/total CDs: 27), Massachusetts (5 CDs/total CDs in the state: 9), Arizona (5 CDs/total CDs: 9), Illinois (4 CDs/total CDs: 18), among others.
CDs: 9), Illinois (3 CDs/total CDs: 19), among others (Fig. 3D). States with CDs with the lowest increase in deaths per capita that ranked in the bottom decile were California (7 CDs/total CDs in the state: 53), Oregon (4 CDs/total CDs: 5), West Virginia (3 CDs/total CDs: 3), Washington (3 CDs/total CDs: 10), among others.

Distribution of COVID-19 Case Fatality Rate
CFR, for the cumulative period across CDs varied between 0.38% and 11.92% with a mean of 3.37% (Fig. 4). States with the highest CFR, ranking in the top decile were New York (11 CDs/total CDs in the state: 27), New Jersey (10 CDs/total CDs: 12), Michigan (7 CDs/total CDs: 14), Pennsylvania (5 CDs/total CDs: 18), Connecticut (4 CDs/total CDs: 5), Massachusetts (4 CDs/total CDs: 9), Ohio (1 CD/total CDs: 16) and Rhode Island (1 CD/total CDs: 2) (Fig. 4). States with CDs in the bottom decile were in Texas (13 CDs/total CDs: 36), Florida (7 CDs/total CDs: 27), among others.

Relationship between recent and past periods for COVID-19 statistics
The correlation between recent cases per capita after states reopened and the cases per capita before reopening (January 22–June 13, 2020) was negative (~0.24) across CDs. This
implies that on average, CDs with formerly low cases per capita, saw a relative increase from mid-June onwards and vice versa.

Fig. 3. Decile CDs for (A) cumulative COVID-19 cases per 1,000 residents as of July 13, (B) Recent COVID-19 cases per 1,000 residents between June 13 and July 13 and (C) COVID-19 deaths per 1,000 residents, (D) Recent COVID-19 deaths per 1,000 residents between June 13 and July 13.

The correlation between recent deaths per capita after states reopened, and the deaths per capita before reopening was higher (0.58), indicating that CDs with high deaths per capita before the states reopened, also had high deaths per capita from mid-June onwards. This could be largely because patients with COVID-19 are sick for a few days/weeks before succumbing to the virus. (Fig. 5).

Variability across congressional districts by states
There was substantial variation in cases, deaths (per 1,000 people) and CFR across CDs within states (Fig. 6). States with a high mean of COVID-19 cases, deaths per 1,000 residents and CFR also display great variability in these outcomes. For example, New York which has the highest median number of cases (min, 2.005; max, 37.937; median, 22.887) and deaths (min, 0.112; max, 3.228; median, 1.603) per 1,000 residents also displays the greatest variability in both outcomes across CDs (Fig. 6A and B). Michigan, which has the highest median CFR value (min, 2.973%; max, 11.923%; median, 7.584%), also displays the greatest variability in CFR across CDs.
COVID-19 across United States congressional districts

Fig. 6. Boxplot of coronavirus disease 2019 (A) cumulative cases per 1,000 residents, (B) cumulative deaths per 1,000 residents, (C) case fatality rate as of July 13, 2020. Boxes represent the 25th to 75th percentile, central dark line the median, bars outside the box represent 1.5 × interquartile range, and the circles are the outliers. States with the highest variance in congressional district cases per 1,000 residents are shown from left to right.

Fig. 5. Correlations between the different coronavirus disease 2019 outcomes for different time periods. CFR = case fatality rate.

https://e-jghs.org

https://doi.org/10.35500/jghs.2020.2.e22
Fig. 6. (Continued) Boxplot of coronavirus disease 2019 (A) cumulative cases per 1,000 residents, (B) cumulative deaths per 1,000 residents, (C) case fatality rate as of July 13, 2020. Boxes represent the 25th to 75th percentile, central dark line the median, bars outside the box represent 1.5 × interquartile range, and the circles are the outliers. States with the highest variance in congressional district cases per 1,000 residents are shown from left to right.
COMMENT

Although state and local governments are at the front line of the COVID-19 crisis, the US Congress plays an important role in designing federal legislation to help the country counter COVID-19. The CDs form a critical link between the local, state, and federal instruments of government, and offer real opportunities for convergence between the administrative and political aspects of policy governance.

More than 800 legislative bills designed to combat COVID-19 are pending in Congress. These bills range from designing the means to improve emergency unemployment relief for governmental entities, establishing loan programs for affected businesses, instituting emergency housing protections and so on. Such federal legislation remains key to the safe reopening of the US economy. Data at the CD-level can equip congressional representatives to further customize legislation and vote for bills most suited for their constituent’s needs.

Access to COVID-19 statistics can also help representatives decide which bills to push for, how to effectively communicate with their constituents, make decisions about the required increasing capacity of healthcare centers in their districts and the kind of socially distancing measures the need to enforce. Data is integral to monitoring and evaluating the plans to reopen their respective districts safely. Such data should also facilitate learning from other CDs, regardless of their “success” or “failures”. Equally importantly, access to data at the CD-level can inform and empower the constituents to engage with their representatives and foster accountability and improve governance mechanisms.

The substantial variation in CD-level COVID-19 statistics within states suggests that access to only state-level information is insufficient to capture the nature of the disease burden in a CD.

The COVID-19 data presented here can be layered on with other data related to clinical and non-clinical risk factors of COVID-19, as well as health care capacities or socioeconomic indicators of well-being for better targeting of strategies to contain and mitigate the impact of COVID-19. As COVID-19 outcomes are dynamic and change every day, a dynamic national-level monitoring system needs to be set up to monitor changes in CDs for representatives to better evaluate their plans.

Indeed, in all likelihood, congressional representatives or their constituents are perhaps aware of the burden in a general sense. However, data can sharpen and bring precision to the intensity of the COVID-19 related burden, and in turn, triggering targeted action. For example, Senator Andrew Jones, when presented with CD-level opioid prescription rates in his district said, “I always knew it was a serious problem here. But to learn that we were ground zero for prescribing rates was eye-opening.”

In conclusion, our aim is to fill an important data void by providing an evidence-base for US congressional representatives and their constituents, which we hope can be leveraged to decide on the containment and mitigation strategies for the COVID-19 infection, as well as to ensure a safe opening up of our economy and society.
ACKNOWLEDGMENT

We are grateful to Jack Cordes for his advice and guidance with the methodology we use in this paper.

SUPPLEMENTARY MATERIAL

Supplementary Table 1
Estimates of the CD-level COVID-19 outcomes: cumulative cases per 1,000 residents as of July 13, cumulative deaths per 1,000 residents as of July 13, CFR as of July 13, 2020, cases per 1000 between June 13 and July 13, cases per 1000 between June 13 and July 13, the ranking of reach CD for each outcome and the decile in which each CD falls in for each outcome

Click here to view

REFERENCES


