Excess all-cause mortality across counties in the United States, March 2020 to December 2021

Eugenio Paglino[1], Dielle J. Lundberg[2], Ahyoung Cho[3], Joe A. Wasserman[4], Rafeya Raquib[2], Anneliese N. Luck[1], Katherine Hempstead[5], Jacob Bor[2], Irma T. Elo[1], Samuel H. Preston[1], Andrew C. Stokes[2]

[1]Department of Sociology and Population Studies Center, University of Pennsylvania, Philadelphia, PA

[2]Department of Global Health, Boston University School of Public Health, Boston, MA

[3]Department of Political Science, Boston University, Boston, MA

[4]RTI International, Research Triangle Park, NC

[5]Robert Wood Johnson Foundation, Princeton, NJ

ABSTRACT

Official Covid-19 death counts have underestimated the mortality impact of the Covid-19 pandemic in the United States. Excess mortality, which compares observed deaths to deaths expected in the absence of the pandemic, is a useful measure for assessing the total effect of the pandemic on mortality levels. In the present study, we produce county-level estimates of excess mortality for 3,127 counties between March 2020 and December 2021. We fit two hierarchical linear models to county-level death rates from January 2015 to December 2019 and predict expected deaths for each month during the pandemic. We compare observed deaths to these estimates to obtain excess deaths for each county-month. An estimated 936,911 excess deaths occurred during 2020 and 2021, of which 171,168 (18.3%) were not assigned to Covid-19 on death certificates as an underlying cause of death. Urban counties in the Far West, Great Lakes, Mideast, and New England experienced a substantial mortality disadvantage in 2020, whereas rural counties in these regions had higher mortality in 2021. In the Southeast, Southwest, Rocky Mountain, and Plains regions, there was a rural mortality disadvantage in 2020, which was exacerbated in 2021. The proportion of excess deaths assigned to Covid-19 was lower in 2020 (76.3%) than in 2021 (87.0%), suggesting that a larger fraction of excess deaths was assigned to Covid-19 later in the pandemic. However, in rural areas and in the Southeast and Southwest a large share of excess deaths was still not assigned to Covid-19 during 2021.

KEYWORDS: Covid-19, Excess mortality, Geographic inequalities, Rural mortality, Counties

CORRESPONDING AUTHOR: Andrew C. Stokes, 801 Massachusetts Avenue, Crosstown Building 362, Boston, MA, 02118, 617-358-2444, acstokes@bu.edu

SIGNIFICANCE

Deaths during the Covid-19 pandemic have been primarily monitored through death certificates containing reference to Covid-19. This approach has missed more than 170,000 deaths related to the pandemic between 2020 and 2021. While the ascertainment of Covid-19 deaths improved during 2021, the full effects of the pandemic still remained obscured in some regions. County-level estimates of excess mortality are useful for studying geographic inequities in the mortality burden associated with the pandemic and identifying specific regions where the full mortality burden was significantly underreported (i.e. Southeast). They can also be used to inform resource allocation decisions at the federal and state levels and encourage uptake of preventive measures in communities with low vaccine uptake.

Introduction

The Covid-19 pandemic has had a substantial impact on mortality in the United States, leading to declines in life expectancy not previously observed since the end of World War II (1, 2). Estimates of excess mortality, which compare observed deaths to those expected in the absence of the pandemic, suggest that the true death toll of the pandemic is much larger than indicated by the official Covid-19 death tallies (3–7). For example, one study estimated that 550,000 excess deaths occurred between March 2020 and February 2021 and that approximately one quarter of these excess deaths were assigned to causes other than Covid-19 (8).

Excess deaths may have been assigned to causes other than Covid-19 for several reasons. Lack of access to testing in the community, combined with the inconsistent use of post-mortem testing for suspected cases, likely resulted in a large share of undiagnosed Covid-19 infections and deaths, especially early in the pandemic and in rural areas (9–12). Additionally, persons with comorbid conditions, such as cardiovascular disease, diabetes, hypertension, and pulmonary disease, may have had their cause of death assigned to the comorbid condition rather than to Covid-19 (13). Unattended Covid-19 deaths occurring in the community may have been especially likely to be assigned to another cause of death (14). Individuals who developed cardiovascular disease or diabetes as a result of the post-acute sequelae of Covid-19 and subsequently died also may not have had their deaths assigned to Covid-19 (15, 16). Finally, excess deaths not assigned to Covid-19 may also reflect deaths indirectly related to the pandemic, including deaths associated with reductions in access to health care, hospital avoidance due to fear of Covid-19 infection, increases in drug overdoses, and economic hardship leading to housing and food insecurity (17–23).

There are multiple benefits to using excess mortality rather than assigned Covid-19 deaths to assess the mortality impact of the Covid-19 pandemic. First, estimates of excess mortality are more comparable spatially than Covid-19-assigned mortality, because states use different definitions to assign Covid-19 deaths and local death investigation systems may have different policies and resources that affect assignment of Covid-19 deaths.[1] Second, as many Covid-19 deaths go uncounted, excess mortality is likely to provide a more accurate measure of the pandemic impact for purposes of resource allocation. Thus, continued tracking of excess mortality across time and space is vital to clarifying the total impact of the pandemic and where its impact is greatest, and to identifying the most appropriate policy responses and interventions.

Prior studies of excess mortality have primarily focused on national and state-level estimates (5, 6), but revealing the true mortality impact of the Covid-19 pandemic at the county-level is especially valuable for several reasons. First, because counties are the administrative unit for death investigation, excess mortality estimates have the potential to help identify counties with substantial Covid-19 under-counts that would benefit from additional training in cause-of-death certification (25). Such estimates may also be valuable for informing local public health workers, community leaders, and residents of the true impact of the pandemic, thus potentially increasing vaccination and uptake of other protective measures (26). These estimates may also be used to target federal and state emergency resources, such as funeral assistance support from the Federal Emergency Management Agency (FEMA). Finally, estimating excess mortality at the county-level also enables analyses of the factors affecting mortality associated with the pandemic, including geographic dimensions like the urban/rural continuum.

[1] Recent guidance from the Council of State and Territorial Epidemiologists (CTSE) provides states with a standardized definition for Covid-19-associated deaths. Prior to the release of this guidance on December 22, 2021, there was no standard definition for reporting of Covid-19-associated deaths among Covid-19 cases and state procedures have varied widely (24).

One prior study generated predictions of excess mortality at the county level but was limited to data from 2020 and pooled small counties together to increase precision (4). Expanding these estimates to 2021 is critical because the geography of the pandemic has likely changed markedly since 2020 as a result of policy shifts, the availability of vaccines, and the emergence of new variants. In the present study, we developed two hierarchical models to estimate excess mortality for 3,127 harmonized counties[2] for the period from March 2020 to December 2021. We also evaluated the extent to which excess deaths are accounted for in official Covid-19 death tallies as an indicator of potential under-reporting of Covid-19 mortality. In addition to generating county-level estimates of excess mortality, we produced several aggregations of the county estimates to examine differences in excess mortality by U.S. Bureau of Economic Analysis (BEA) region and across metropolitan (metro) vs. non-metropolitan (nonmetro) areas.

Results

Across the United States, 459,764 excess deaths occurred between March and December 2020, and 477,147 excess deaths occurred in 2021. This equals 936,911 excess deaths during calendar years 2020 and 2021, of which 765,743 (81.7%) were assigned to Covid-19 as an underlying cause of death and 171,168 (18.3%) were not assigned to Covid-19. In 2020, excess death rates were highest in nonmetro areas (207 deaths per 100,000 residents) followed by large central metro areas (173 deaths per 100,000 residents), whereas in 2021, excess death rates were highest in nonmetro areas (227 deaths per 100,000 residents) followed by small or medium metros (163 deaths per 100,000 residents). Regionally, the excess death rate in 2020 was highest in the Mideast (206 deaths per 100,000 residents) followed by the Southwest (188 deaths per 100,000

[2] To reconcile difference in FIPS code across all of our data sources we tracked county FIPS codes back to their 1990 values reversing the changes operated by the Census Bureau over time as detailed here: https://www.census.gov/programs-surveys/geography/technical-documentation/county-changes.2020.html

residents). In 2021, however, the excess death rate was highest in the Southeast (205 deaths per 100,000 residents) and the Southwest (198 deaths per 100,000 residents). The areas with the highest excess death rates in 2020 were large central metros in the Mideast and nonmetro areas in the Southwest and Southeast. In 2021, the areas with the highest excess death rates were nonmetro areas in the Southwest, nonmetro areas in the Southeast, and small or medium metros in the Southwest and Southeast (**Table 1**). **Supplemental Table S1** provides estimates of Covid-19 and excess mortality rates for each state in 2020 and 2021. Excess death rates were highest in Mississippi (301 deaths per 100,000 residents) followed by Arizona (246 deaths per 100,000 residents) in 2020 and in West Virginia (298 deaths per 100,000 residents) followed by Mississippi (271 deaths per 100,000 residents) in 2021.

Figure 1 shows excess death rates across all counties in the United States. Between 2020 and 2021, excess deaths shifted from the Northeast and Midwest to the South and to the West and from metro to nonmetro areas. In 2020, excess mortality was higher in metro counties in the Far West, Great Lakes, Mideast, and New England (**Table 1**). In 2021, however, excess mortality was higher in nonmetro areas than metro areas in these regions. In the Southeast, Southwest, Rocky Mountain, and Plains regions, excess mortality was higher in nonmetro areas in both 2020 and 2021. Nationally, excess mortality was higher in nonmetro areas than metro areas, and the disparity between nonmetro and metro areas was greater in 2021.

Supplemental Figure S1 shows actual and expected death rates for the U.S. by month during 2020 and 2021. Three peaks in mortality are apparent: (1) early 2020, (2) end of 2020 / start of 2021, and (3) end of 2021. **Figure 2** breaks down trends in excess death rates within each BEA region by month throughout the period. Excess death rates peaked in the Mideast in early 2020, primarily in large metro areas. Excess death rates also increased markedly in New England

and the Great Lakes during this time. Around the end of 2020, a second peak resulted in high excess death rates in the Far West, Great Lakes, Southwest, and Southeast regions. A third peak was observed in most of the regions near the end of 2021. **Supplemental Figure S2** shows actual and expected death rates for the largest county in each BEA region - metro combination by month during 2020 and 2021. Some counties only experienced one peak in mortality (e.g. Kings County, New York), whereas others experienced three distinct peaks (e.g. Navajo County, Arizona).

In absolute terms, the Southeast was the region with the most excess deaths in both 2020 and 2021 followed by the Mideast and Great Lakes in 2020 and by the Southwest in 2021 (**Figure 3**). While New York was the state with the most excess deaths in 2020, Texas, California, and Florida had the most excess deaths in 2021. Excess mortality was also less concentrated in large metro areas and large fringe areas in 2021 than it was in 2020. **Supplemental Figure S3** provides a comparison of states and regions in terms of excess death rates, rather than absolute counts.

Figure 4 plots the proportion of excess deaths assigned to Covid-19 across counties in the United States. In both 2020 and 2021, counties across the country reported substantial numbers of excess deaths not assigned to Covid-19 as an underlying cause of death. In total, 76.3% of excess deaths were assigned to Covid-19 in 2020, whereas in 2021, 87.0% of excess deaths were assigned to Covid-19. This equals 109,069 excess deaths that were not assigned to Covid-19 in 2020 and 62,099 excess deaths that were not assigned to Covid-19 in 2021, for a total of 171,168 deaths. Despite the increase in assignment of Covid-19 deaths from 2020 to 2021, many regions still had areas with a low proportion of excess deaths assigned to Covid-19 during 2021, such as nonmetro areas in the Far West (69.5%) and Southeast (71.4%). The Southeast and Southwest were the regions with the lowest overall assignment of excess deaths to Covid-19 in 2021.

In 2020, assignment of excess deaths to Covid-19 was similar in large central metro areas (76.3%) compared to nonmetro areas (74.9%), whereas in 2021, assignment was lower in nonmetro areas (80.0%) than large central metro areas (89.1%). Despite this overall trend, several regions (Mideast and New England) had lower assignment in nonmetro areas in 2020 than in metro areas. This suggests that assignment of excess deaths to Covid-19 in nonmetro areas was low in many regions across the study period. In 2021, assignment of excess deaths to Covid-19 was particularly low in nonmetro areas in the Southeast. **Supplemental Figure S4** identifies the states with the lowest assignment of excess deaths to Covid-19. Many of these states were in the Southeast and included South Carolina, Louisiana, Mississippi, and North Carolina. In contrast, several states in New England including Massachusetts, New Jersey, Connecticut, and Rhode Island reported more Covid-19 deaths than excess deaths. Overall, few states reported more deaths assigned to Covid-19 than our estimates of excess deaths (5 states in 2020 and 14 states in 2021).

Supplemental Figure S5 plots excess death rates for each county against their Covid-19 death rates. Counties above the 45-degree reference line represent areas where the excess death rate was higher than the Covid-19 death rate, indicating there were excess deaths not assigned to Covid-19 in these counties. In 2020, the majority of counties in all regions except New England were above the 45-degree line, demonstrating that there were excess deaths not assigned to Covid-19 in these areas. Excess deaths not assigned to Covid-19 occurred in counties across the metro-nonmetro continuum. In 2021, the majority of counties in the Southeast, Southwest, Far West, and Rocky Mountains regions remained above the 45-degree line, demonstrating that they still had excess deaths not assigned to Covid-19. In particular, the counties with the highest proportion of excess deaths not assigned to Covid-19 tended to be nonmetro counties. Most counties in the Mideast, Plains, and Great Lakes, however, no longer had excess deaths not assigned to Covid-19 in 2021.

Discussion

In the present study, we produced estimates of excess deaths associated with the Covid-19 pandemic from March 2020 to December 2021 across 3,127 harmonized counties in the United States. Our study found that nearly 940,000 excess deaths occurred in the U.S. between 2020 and 2021, of which more than 170,000 were not assigned to Covid-19 on death certificates. This indicates that excess deaths were 22% higher than deaths assigned to Covid-19 during this period.

Prior studies of excess mortality have largely produced estimates for the year 2020 (3–6, 27), leaving patterns of excess mortality during 2021 under-studied. The Center for Disease Control and Prevention (CDC) however has reported a provisional estimate of approximately 944,000 excess deaths in the U.S. from March 2020 to December 2021, which is very close to our estimate (28). Woolf et al. identified 522,368 excess deaths from March 1, 2020 to January 2, 2021, which is higher than our estimate of approximately 470,000 deaths for the year 2020 (6). Islam et al. reported 458,000 excess deaths during 2020, which is close to our estimate (29). A prior estimate by Stokes et al. found 438,386 excess deaths in 2020, which is lower than our estimate due to differences in methods and time horizons for predicting expected deaths (4).

There are multiple potential advantages to using county-level data to generate estimates of excess mortality at the state and national levels. Predicting expected mortality at the state or national levels assumes that all areas within the state or nation have the same background trends of mortality. However, in actuality, different regions and metro and nonmetro counties of the U.S. have experienced varied long-term mortality trends (8, 30, 31). Our approach may produce more reliable estimates of expected mortality in the absence of the pandemic, because the projections build on historical trends at the county-level.

Another key advantage of producing excess mortality estimates at the county-level is the opportunity to examine differences across the urban-rural continuum, which is not possible with state-level data. By exploring how excess mortality varies by region and time across the urban-rural continuum, we gained several novel insights about patterns of mortality and the assignment of Covid-19 as an underlying cause on death certificates during the pandemic.

One major finding of this study is that there were similar numbers of excess deaths in 2020 and 2021, which is noteworthy as vaccinations were available for much of 2021. Despite the strong efficacy of vaccines, gaps in uptake likely contributed to high excess mortality in 2021, which may persist into the future if these vaccination gaps are not closed. This finding may also reflect federal and state governments' decision to prioritize individual-level interventions over population-based strategies designed to protect the communities at greatest risk for Covid-19 death, such as financial support for family and medical leave, improved ventilation of schools and workplaces, and vaccine delivery programs organized in coordination with community partners (32).

We found substantial variation in excess mortality by US region and across the urban-rural continuum, which differed between 2020 and 2021. In the Far West, Great Lakes, Mideast, and New England, there was a substantial urban mortality disadvantage in 2020, which was reversed in 2021 to yield a rural mortality disadvantage. In the Southeast, Southwest, Rocky Mountain, and Plains regions, there was a rural mortality disadvantage in 2020, which was exacerbated in 2021. This suggests that the pandemic has impacted rural areas heavily, especially in 2021, suggesting a need for increased preventative measures in these areas where vaccination remains low (33).

Another finding of this study is that excess death rates exceeded Covid-19 death rates across most counties in all BEA regions during 2020 except New England. This indicates that there were excess deaths not assigned to Covid-19 reported in these regions during 2020. Between 2020 and 2021, Covid-19 and excess mortality then converged in some regions of the country (e.g. Mideast and Great Lakes). In other regions (e.g. Southeast and Southwest), excess death rates continued to exceed Covid-19 death rates in most counties, indicating that excess deaths not assigned to Covid-19 persisted. This finding may have several explanations. In the Mideast, the gap between Covid-19 and excess mortality in 2020 likely reflected the fact that the pandemic affected this region early in the pandemic when access to testing was extremely limited and the clinical manifestations of Covid-19 were unclear (27). The elimination of this gap by 2021 suggests that as the pandemic progressed, the Mideast was able to more effectively capture Covid-19 deaths. It is also possible that some indirect effects of the pandemic, such as disruptions in health care access, also decreased during this time. In other regions such as the Southeast and Southwest, the gaps between Covid-19 and excess mortality persisted into 2021, which may relate to the continued lack of Covid-19 testing in many Southeastern and Southwestern counties even as the pandemic progressed (34–36).

Excess deaths exceeded deaths assigned to Covid-19 in rural areas across many regions, especially in 2021. Many counties in the Southeast and Southwest also had low levels of assignment of excess deaths to Covid-19 throughout the pandemic. Rural counties may have under-counted Covid-19 deaths throughout the pandemic as a result of the exceptionally high death rates many rural areas faced during the Winter surge of 2020-2021 and the subsequent Delta surge in Summer/Fall 2021 (37–41). Other contributing factors may have included under-resourced health care systems that were unable to care for patients with Covid-19 and/or other non-Covid-19 conditions, under-resourced death investigation systems in which medical examiners or coroners did not pursue post-mortem Covid-19 testing, and partisan beliefs regarding the Covid-19 pandemic that may have influenced cause of death assignment and the likelihood of testing (19, 25).

Although our study does not distinguish between uncounted Covid-19 deaths and deaths indirectly related to the pandemic, emerging literature suggests that a large share of the excess deaths not assigned to Covid-19 likely represent uncounted Covid-19 deaths (10, 11, 42). For example, one recent study by Lee et al. found that approximately 90% of excess mortality between March 2020 and April 2021 could be attributed to the direct effects of SARS-CoV-2 infection (42). This possibility is also supported by investigative reporting during the pandemic which has documented widespread irregularities in cause of death assignment resulting from over-burdened and under-resourced death investigation systems (43). Discrepancies between Covid-19 death rates and excess death rates are problematic because they have the potential to mislead scientists and policymakers about which areas were most heavily affected during the pandemic. Failure to accurately capture Covid-19 deaths also points to an urgent need to modernize the death investigation system in the United States, including expanding budgets for medical examiner officers and eliminating the archaic coroner system (25).

In New England, we observed a different pattern around assignment of Covid-19 deaths than in the other BEA regions. In this region, a large share of counties had higher Covid-19 than excess death rates, a pattern that became more pronounced between 2020 and 2021. Several explanations may exist for this pattern including that other causes of death (i.e. influenza) declined in these areas or that the economically privileged status of many of these counties shielded their residents from the negative indirect effect of the pandemic by allowing them to work-from-home and avoid household crowding. Finally, it is possible that deaths were over-assigned to Covid-19 in these areas due to different cause of death assignment protocols for Covid-19. For example, until March 2022, Covid-19 deaths in the state of Massachusetts included any death that occurred within 60 days of a Covid-19 diagnosis, which differed from other states and guidelines from the Council of State and Territorial Epidemiologists that recommended states use a 30 day window (44).

The study had several limitations. First, the study relied on publicly available data, which were subject to suppression of death counts fewer than 10 in a given county-month. We addressed this limitation by pooling information across different geographical levels through the use of hierarchical models and by taking advantage of the additional information provided by yearly death counts, however, our estimates remain uncertain in areas with small populations and few deaths. Second, some counties and states have experienced prolonged reporting delays of Covid-19 deaths, which could affect our estimates of the proportion of excess deaths assigned to Covid-19, particularly in more recent months. Third, we were unable to distinguish between excess deaths that represented uncounted Covid-19 deaths and excess deaths indirectly related to the pandemic. Future research should examine this distinction to clarify the extent to which excess deaths not assigned to Covid-19 reported in this study represent under-reporting of Covid-19 deaths. Fourth, our study examined all-cause mortality and did not explore differences in trends using cause-specific death rates. Assessing geographic and temporal differences in excess death rates by cause-of-death would allow for a deeper understanding of the mechanisms driving trends in excess mortality overall and is an important direction for future work. Fifth, we used underlying cause of death data to identify deaths assigned to Covid-19 and thus did not identify deaths where Covid-19 was listed as a contributing cause. Finally, due to data limitations, our model does not account for differences in age structure between counties. Since the pandemic has affected older populations more significantly, some differences in mortality observed between counties may simply reflect differences in their age distribution.

In conclusion, the present study generated novel estimates of Covid-19 and excess mortality for 3,127 harmonized county units over the period from March 2020 to December 2021. In contrast to official Covid-19 death tallies, which are subject to differential underreporting and fail to capture indirect pandemic effects, the present estimates more fully account for the true toll of the pandemic across local areas and provide a more comparable measure of the Covid-19 mortality burden. As such, these estimates may be useful for additional work to investigate the determinants of excess mortality throughout the pandemic and may also be useful for communicating Covid-19 risks with local communities where the direct tallies have hidden the full extent of the pandemic's consequences.

Materials & Methods

Yearly and monthly death counts at the county level were extracted from CDC WONDER online tool. See **Methods Supplement** for further details about data extraction procedures. We extracted death counts by all causes of death and from Covid-19. Causes of death were selected from the Multiple Cause of Death database using the provisional counts for 2020 and 2021 and the final counts for 2015-2019. A death was assigned to Covid-19 when Covid-19 was listed as the underlying cause of death using the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) code *U07.1*.[3] For all years except 2020, yearly deaths include all deaths between January to December. For 2020, yearly deaths include only deaths that occurred between March and December to better reflect the pandemic period. All death rates computed for 2020 were adjusted to account for the shorter exposure period.

To convert the number of deaths into rates, we used publicly available yearly county-level population estimates from the Census Bureau.[4] To compute monthly rates, we assumed linear growth between each two time points. For the August-December period in 2021, for which no population estimates are available, we assumed the population remained constant at its value in July.

We grouped counties into 4 metropolitan-nonmetropolitan categories (large central metro, large fringe metro, medium or small metro, nonmetro) based on the 2013 NCHS Rural-Urban Classification Scheme for Counties (45) and into 8 Bureau of Economic Analysis (BEA) regions (Far West, Great Lakes, Mideast, New England, Plains, Rocky Mountains, Southeast, Southwest) (30). Next, we stratified each region by each metropolitan- nonmetropolitan category, leading to 32 geographic units. We also grouped small counties into county-sets according to the United States Census Bureau's County Sets classification. County-sets were used, in addition to counties, within our model to improve the precision of estimates for small counties in the study. See **Methods Supplement** for further details about the geographic classifications used in this study.

[3]Covid-19 is listed as the underlying cause of death in approximately 92% of cases in which it is mentioned somewhere on the death certificate (28).
[4]Population data were obtained from the following sources for 2010-2020 and 2020-2021:
https://www2.census.gov/programs-surveys/popest/datasets/2010-2020/counties/totals/co-est2020.csv
https://www2.census.gov/programs-surveys/popest/datasets/2020-2021/counties/totals/co-est2021-alldata.csv

To model the monthly county-level number of deaths for 100,000 residents (DR), we estimated two different hierarchical linear models. The first model predicts monthly DR for a county directly while the second one predicts yearly DR for a county and then distributes the corresponding number of deaths over the year by month according to weights computed over the All-Causes of Death data at the national level. Both models were estimated using the lme4 package for the R language (46) and fitted using monthly mortality data for the period January 2015 - December 2019. The need for these two different models arises as a result of the suppression procedure applied in the CDC WONDER tool to all data points (county-months in our case) with fewer than 10 deaths. As a result of suppression, small counties that rarely exceed 10 deaths in a given month have very few data points, and these data points are not representative of normal mortality conditions (that is why they were not suppressed). In other words, data points are not missing at random and non-missing data points for small counties reflect higher-than-normal mortality. The yearly-level model, by making use of the additional information on the yearly number of deaths, leads to more accurate predictions in counties with a high proportion of missing data points. Combining these two models led to a better overall performance compared to using either one in isolation.

The monthly-level model expresses the monthly number of deaths for every 100,000 residents (DR) as a function of time (in years), month (with dummy variables), and an intercept allowed to vary across counties, county-sets, and states. To make the model more flexible, the time slope is also allowed to vary across county sets. Formally[5]:

[5] Here we use the notation employed in Raudenbush and Bryk (2002) (47).

$$\begin{aligned}
\text{Level 1: } & DR_{c,cs,s,m,y} = \pi_{0c,cs,s} + \pi_{1cs} Year_y \\
& \qquad + \sum_{m=1}^{12} \theta_m Month_m + \varepsilon_{c,cs,s,m,y} \\
\text{Level 2: } & \pi_{0c,cs,s} \sim \mathcal{N}\left(\gamma_{0cs,s}, \sigma^2_{0C}\right) \\
\text{Level 3: } & \begin{bmatrix} \gamma_{0cs,s} \\ \pi_{1cs} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \gamma_{0s} \\ \gamma_{1CS} \end{bmatrix}, \begin{bmatrix} \sigma^2_{0CS} & \sigma_{01CS} \\ \sigma_{10CS} & \sigma^2_{1CS} \end{bmatrix}\right) \\
\text{Level 4: } & \gamma_{0s} \sim \mathcal{N}\left(\gamma_0, \sigma^2_{0S}\right)
\end{aligned}$$

Where:

$$\varepsilon_{c,cs,s,m,y} \sim \mathcal{N}\left(0, \sigma^2_{\varepsilon}\right)$$

In the equations above, the subscripts *c*, *cs*, *s*, *m*, and *y* indicate county, county-set, state, month, and year respectively. The capital letters C, CS, and S indicate that a term is specific to the county, county-set, and state level equations respectively. Using county-sets as an intermediate level between counties and states helped us overcome estimation difficulties with counties with few data points due to suppression. Even when the county level intercept cannot be estimated with precision, the estimate will be pulled toward the mean of counties in the same state and county-set, which we were able to estimate more precisely. The yearly-level model follows a structure similar to the monthly model but, due to the smaller sample size, it is less complex. Yearly death rates are modeled as a function of time (in years) and an intercept, both allowed to vary across counties.

$$\text{Level 1: } DR_{c,y} = \pi_{0c} + \pi_{1c} Year_y + \varepsilon_{c,y}$$
$$\text{Level 2: } \begin{bmatrix} \pi_{0c} \\ \pi_{1c} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \gamma_0 \\ \gamma_1 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right)$$

Where:

$$\varepsilon_{c,y} \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$$

In both models, the intercepts and the slopes are allowed to be correlated. To obtain the number of deaths, we multiplied the estimated death rate by the corresponding population. To obtain the monthly number of deaths from the annual number of deaths for the yearly model, we first computed the average proportion of deaths occurring in each month over the 2015-2019 period and then distributed annual deaths accordingly.

To decide whether to use the yearly model or the monthly model, we computed the average percentage difference between the predicted yearly deaths for the period 2015-2019 and the actual yearly deaths. We then used the estimate from the yearly model for all counties in which the difference was larger than 10%. Applying this decision rule, we used the yearly model for 790 counties and the monthly model for the remaining 2322 counties. Further details about the model section are provided in the **Methods Supplement.**

We obtained confidence intervals for the death rates by sampling from the distribution of the models' fixed effects 1000 times and using these samples to compute 1000 different predicted rates. We then computed the 2.5 and 97.5 percentiles of the resulting distribution. The intervals thus obtained do not reflect all of the models' uncertainty but only the portion due to fixed effects. However, they are consistent with the historical variability in the death rates and can be used to get a sense of the estimates' variability.

This study used de-identified publicly available data and was exempted from review by the Boston University Medical Center Institutional Review Board. Programming code was developed using R, version 4.1.0 (R Project for Statistical Computing) and Python, version 3.7.13 (Python Software Foundation).

Acknowledgements

The authors would like to thank Robert N. Anderson and Farida B. Ahmad from the National Center for Health Statistics for technical assistance with the provisional mortality files used in the present study. We also thank Elif Coskun, Mikas Hansen, Sylvia Lutze, Pia MacDonald, and Steele Myrick for additional technical and administrative support. Additionally, we are grateful to Nahid R. Bhadelia and Erica Augustine from the Center for Emerging Infectious Diseases Policy and Research at Boston University for their guidance and support with translation and dissemination of study findings; Gus Wezerek from the New York Times for input on the analytic strategy and data visualization; and members of the Documenting Covid-19 Project team at Muckrock for their insightful investigative reporting on uncounted Covid-19 deaths. The authors gratefully acknowledge financial support from the Robert Wood Johnson Foundation (77521), the National Institute on Aging (R01-AG060115-04 and R01-AG060115-04S1) and the W.K. Kellogg Foundation. The views expressed in this paper are those of the authors and not necessarily the views of the study sponsors.

Conflicts of Interest

The authors report that they have no conflicts of interests to disclose.

Data Availability

Data used in the present manuscript are publicly available from the Centers for Disease Control and Prevention and Census Bureau. Additional details about the data and programming code for replicating the analyses of the present study can be accessed at the linked GitHub repository: https://github.com/Vital-Stats-Integrity-Proj/monthly_county_level_excess_mortality

References

1. T. Andrasfay, N. Goldman, Reductions in 2020 US life expectancy due to COVID-19 and the disproportionate impact on the Black and Latino populations. *Proc. Natl. Acad. Sci. U. S. A.* **118** (2021).

2. S. H. Woolf, R. K. Masters, L. Y. Aron, Effect of the covid-19 pandemic in 2020 on life expectancy across populations in the USA and other high income countries: simulations of provisional mortality data. *BMJ* **373**, n1343 (2021).

3. A. C. Stokes, *et al.*, COVID-19 and excess mortality in the United States: A county-level analysis. *PLoS Med.* **18**, e1003571 (2021).

4. C. A. Ackley, *et al.*, County-level estimates of excess mortality associated with COVID-19 in the United States. *SSM - Population Health* **17**, 101021 (2022).

5. L. M. Rossen, A. M. Branum, F. B. Ahmad, P. D. Sutton, R. N. Anderson, Notes from the Field: Update on Excess Deaths Associated with the COVID-19 Pandemic - United States, January 26, 2020-February 27, 2021. *MMWR Morb. Mortal. Wkly. Rep.* **70**, 570–571 (2021).

6. S. H. Woolf, D. A. Chapman, R. T. Sabo, E. B. Zimmerman, Excess Deaths From COVID-19 and Other Causes in the US, March 1, 2020, to January 2, 2021. *JAMA* (2021) https:/doi.org/10.1001/jama.2021.5199 (April 2, 2021).

7. C. J. Ruhm, Excess Deaths in the United States During the First Year of COVID-19. *NBER Working Paper* (2021) https:/doi.org/10.3386/w29503.

8. S. H. Woolf, *et al.*, Changes in midlife death rates across racial and ethnic groups in the United States: systematic analysis of vital statistics. *BMJ* **362**, k3096 (2018).

9. J. R. Gill, M. E. DeJoseph, The Importance of Proper Death Certification During the COVID-19 Pandemic. *JAMA* **324**, 27 (2020).

10. A. D. Iuliano, *et al.*, Estimating under-recognized COVID-19 deaths, United States, march 2020-may 2021 using an excess mortality modelling approach. *Lancet Reg Health Am* **1**, 100019 (2021).

11. Y. Zhang, H. H. Chang, A. D. Iuliano, C. Reed, Application of Bayesian spatial-temporal models for estimating unrecognized COVID-19 deaths in the United States. *Spat Stat*, 100584 (2022).

12. J. T. Mueller, *et al.*, Impacts of the COVID-19 pandemic on rural America. *Proc. Natl. Acad. Sci. U. S. A.* **118** (2020).

13. S. Elezkurtaj, *et al.*, Causes of death and comorbidities in hospitalized patients with COVID-19. *Sci. Rep.* **11**, 4263 (2021).

14. E. B. Pathak, R. B. Garcia, J. M. Menard, J. L. Salemi, Out-of-Hospital COVID-19 Deaths: Consequences for Quality of Medical Care and Accuracy of Cause of Death Coding. *Am. J. Public Health* **111**, S101–S106 (2021).

15. Y. Xie, E. Xu, B. Bowe, Z. Al-Aly, Long-term cardiovascular outcomes of COVID-19. *Nat. Med.* **28**, 583–590 (2022).

16. Y. Xie, Z. Al-Aly, Risks and burdens of incident diabetes in long COVID: a cohort study. *Lancet Diabetes Endocrinol* (2022) https:/doi.org/10.1016/S2213-8587(22)00044-4.

17. J. Zhang, Hospital Avoidance and Unintended Deaths during the COVID-19 Pandemic. *American Journal of Health Economics* **7**, 405–426 (2021).

18. K. E. Anderson, E. E. McGinty, R. Presskreischer, C. L. Barry, Reports of Forgone Medical Care Among US Adults During the Initial Phase of the COVID-19 Pandemic. *JAMA Network Open* **4**, e2034882 (2021).

19. A. C. Stokes, *et al.*, Association of Health Care Factors With Excess Deaths Not Assigned to COVID-19 in the US. *JAMA Network Open* **4**, e2125287 (2021).

20. N. D. Volkow, Collision of the COVID-19 and Addiction Epidemics. *Ann. Intern. Med.* **173**, 61–62 (2020).

21. M. J. Dubey, *et al.*, COVID-19 and addiction. *Diabetes Metab. Syndr.* **14**, 817–823 (2020).

22. J. Raifman, J. Bor, A. Venkataramani, Association Between Receipt of Unemployment Insurance and Food Insecurity Among People Who Lost Employment During the COVID-19 Pandemic in the United States. *JAMA Netw Open* **4**, e2035884 (2021).

23. S. Sandoval-Olascoaga, A. S. Venkataramani, M. C. Arcaya, Eviction Moratoria Expiration and COVID-19 Infection Risk Across Strata of Health and Socioeconomic Status in the United States. *JAMA Netw Open* **4**, e2129041 (2021).

24. Council of State and Territorial Epidemiologists, "Interim Guidance for Public Health Surveillance Programs for Classification of Covid-19-associated Deaths among Covid-19 Cases" (2021).

25. A. C. Stokes, D. J. Lundberg, J. Bor, K. Bibbins-Domingo, Excess Deaths During the COVID-19 Pandemic: Implications for US Death Investigation Systems. *Am. J. Public Health* **111**, S53–S54 (2021).

26. A. Bilinski, E. Emanuel, J. A. Salomon, A. Venkataramani, Better Late Than Never: Trends in COVID-19 Infection Rates, Risk Perceptions, and Behavioral Responses in the USA. *J. Gen. Intern. Med.* **36**, 1825–1828 (2021).

27. D. M. Weinberger, *et al.*, Estimation of Excess Deaths Associated With the COVID-19 Pandemic in the United States, March to May 2020. *JAMA Intern. Med.* **180**, 1336–1344 (2020).

28. , Excess Deaths Associated with COVID-19: Provisional Death Counts for Coronavirus Disease (COVID-19). *Center for Disease Control and Prevention* (2022) (April 18, 2022).

29. N. Islam, *et al.*, Excess deaths associated with covid-19 pandemic in 2020: age and sex disaggregated time series analysis in 29 high income countries. *BMJ* **373**, n1137 (2021).

30. I. T. Elo, A. S. Hendi, J. Y. Ho, Y. C. Vierboom, S. H. Preston, Trends in Non-Hispanic White Mortality in the United States by Metropolitan-Nonmetropolitan Status and Region, 1990-2016. *Popul. Dev. Rev.* **45**, 549–583 (2019).

31. National Academies of Sciences, Engineering, and Medicine, Division of Behavioral and Social Sciences and Education, Committee on National Statistics, Committee on Population, Committee on Rising Midlife Mortality Rates and Socioeconomic Disparities, *High and Rising Mortality Rates Among Working-Age Adults* (National Academies Press, 2021).

32. E. Wrigley-Field, K. M. Berry, A. C. Stokes, J. P. Leider, "Pandemic of the unvaccinated"? At midlife, white people are less vaccinated but still at less risk of Covid-19 mortality in Minnesota. *medRxiv* (2022) https:/doi.org/10.1101/2022.03.02.22271808.

33. R. Saelee, *et al.*, Disparities in COVID-19 Vaccination Coverage Between Urban and Rural Counties - United States, December 14, 2020-January 31, 2022. *MMWR Morb. Mortal. Wkly. Rep.* **71**, 335–340 (2022).

34. K. Brandt, *et al.*, SARS-CoV-2 testing in North Carolina: Racial, ethnic, and geographic disparities. *Health Place* **69**, 102576 (2021).

35. U.S. Department of Health & Human Services, COVID-19 Diagnostic Laboratory Testing (PCR Testing) Time Series (2020) (April 20, 2022).

36. H. Cohen-Cline, *et al.*, Major disparities in COVID-19 test positivity for patients with non-English preferred language even after accounting for race and social factors in the United States in 2020. *BMC Public Health* **21**, 2121 (2021).

37. T. L. Saitone, K. Aleks Schaefer, D. P. Scheitrum, COVID-19 morbidity and mortality in U.S. meatpacking counties. *Food Policy* **101**, 102072 (2021).

38. G. A. Millett, *et al.*, Assessing differential impacts of COVID-19 on black communities. *Ann. Epidemiol.* **47**, 37–44 (2020).

39. D. F. Cuadros, A. J. Branscum, Z. Mukandavire, F. D. Miller, N. MacKinnon, Dynamics of the COVID-19 epidemic in urban and rural areas in the United States. *Ann. Epidemiol.* **59**, 16–20 (2021).

40. S. A. Karim, H.-F. Chen, Deaths From COVID-19 in Rural, Micropolitan, and Metropolitan Areas: A County-Level Comparison. *J. Rural Health* **37**, 124–132 (2021).

41. H. N. Grome, *et al.*, Disparities in COVID-19 Mortality Rates: Implications for Rural

Health Policy and Preparedness. *J. Public Health Manag. Pract.* (2022) https:/doi.org/10.1097/PHH.0000000000001507.

42. W.-E. Lee, *et al.*, Direct and indirect mortality impacts of the COVID-19 pandemic in the US, March 2020-April 2021. *medRxiv* (2022) https:/doi.org/10.1101/2022.02.10.22270721.

43. Documenting COVID-19 project and USA TODAY Network, Uncounted: Inaccurate death certificates across the country hide the true toll of COVID-19. *USA Today* (2021) (April 19, 2022).

44. Massachusetts Department of Public Health, Department of Public Health updates COVID-19 death definition. *Massachusetts Department of Public Health* (May 3, 2022).

45. U.S. Department of Agriculture, Economic Research Service, 2013 Rural-Urban Continuum Codes (May 3, 2022).

46. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).

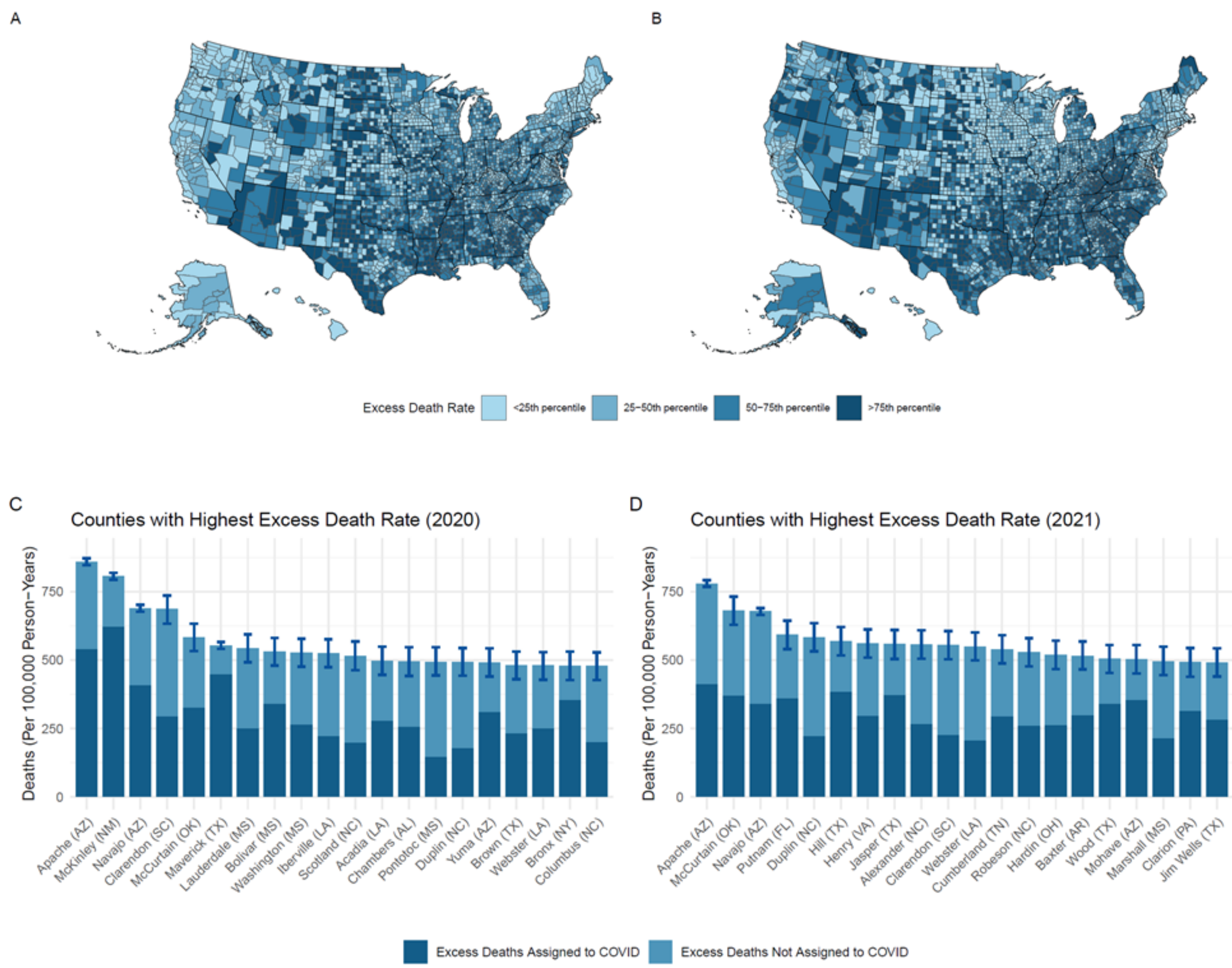
47. S. W. Raudenbush, A. S. Bryk, *Hierarchical Linear Models: Applications and Data Analysis Methods* (SAGE, 2002).

Table 1. Excess mortality, Covid-19 mortality, and the ratio of Covid-19 to excess mortality by metropolitan-nonmetropolitan status and BEA region

		2020					2021				
		Number of Deaths			Rates per 100,000 PY		Number of Deaths			Rates per 100,000 PY	
BEA Region	Metro Status	Excess	COVID	Ratio	Excess	COVID	Excess	COVID	Ratio	Excess	COVID
Total	Total	459,764	350,695	0.763	167	127	477,147	415,048	0.870	144	125
Total	Large Central	146,578	111,101	0.758	173	131	123,302	109,824	0.891	122	108
Total	Large Fringe	98,568	82,432	0.836	140	117	87,332	87,734	1.005	103	103
Total	Medium/Small	135,180	97,689	0.723	164	119	161,891	133,750	0.826	163	134
Total	Non Metro	79,438	59,473	0.749	207	155	104,622	83,740	0.800	227	182
Far West	Region Total	52,939	39,938	0.754	112	85	72,562	60,076	0.828	128	127
Far West	Large Central	32,389	24,486	0.756	128	97	37,827	32,883	0.869	126	109
Far West	Large Fringe	6,852	5,022	0.733	95	69	9,258	7,456	0.805	106	86
Far West	Medium/Small	11,853	8,935	0.754	97	73	19,858	15,831	0.797	136	108
Far West	Non Metro	1,845	1,495	0.810	71	58	5,619	3,906	0.695	179	124
Great Lakes	Region Total	73,063	54,660	0.748	186	139	61,766	56,302	0.912	131	143
Great Lakes	Large Central	22,117	16,654	0.753	204	154	13,113	12,153	0.927	101	94
Great Lakes	Large Fringe	17,662	13,111	0.742	165	123	15,597	13,644	0.875	121	106
Great Lakes	Medium/Small	19,546	13,919	0.712	182	130	17,686	16,211	0.917	138	126
Great Lakes	Non Metro	13,738	10,976	0.799	192	154	15,370	14,294	0.930	179	167
Mideast	Region Total	85,924	76,661	0.892	206	184	40,782	55,015	1.349	82	132
Mideast	Large Central	39,938	33,275	0.833	288	240	11,885	16,404	1.380	72	100
Mideast	Large Fringe	30,257	29,937	0.989	174	172	10,283	19,796	1.925	49	94
Mideast	Medium/Small	11,875	10,504	0.885	149	132	12,222	13,237	1.083	127	138
Mideast	Non Metro	3,854	2,945	0.764	157	120	6,392	5,578	0.873	217	189
New England	Region Total	15,752	18,016	1.144	125	144	5,839	11,269	1.930	39	90
New England	Large Central	3,739	4,012	1.073	191	205	709	1,750	2.468	30	75
New England	Large Fringe	5,295	6,731	1.271	116	148	1,208	3,852	3.189	22	71
New England	Medium/Small	5,846	6,579	1.125	128	144	2,187	4,254	1.945	40	77
New England	Non Metro	872	694	0.796	59	47	1,735	1,413	0.814	98	80
Plains	Region Total	29,544	24,705	0.836	164	137	21,150	21,814	1.031	98	121
Plains	Large Central	3,561	2,764	0.776	150	117	2,433	2,167	0.891	86	77
Plains	Large Fringe	5,968	4,578	0.767	142	109	4,712	4,645	0.986	93	92
Plains	Medium/Small	9,518	7,464	0.784	157	123	6,005	6,305	1.050	82	86
Plains	Non Metro	10,497	9,899	0.943	196	184	8,000	8,697	1.087	124	135
Rocky Mountain	Region Total	11,815	8,649	0.732	113	83	14,193	12,516	0.882	112	118
Rocky Mountain	Large Central	1,717	1,223	0.712	108	77	907	1,031	1.137	48	54
Rocky Mountain	Large Fringe	2,889	2,114	0.732	116	85	2,749	2,078	0.756	91	69
Rocky Mountain	Medium/Small	4,174	3,068	0.735	102	75	5,776	5,282	0.914	115	105
Rocky Mountain	Non Metro	3,035	2,244	0.739	132	98	4,761	4,125	0.866	171	148
Southeast	Region Total	124,271	81,065	0.652	175	114	175,879	130,234	0.740	205	182
Southeast	Large Central	19,341	12,753	0.659	148	98	24,800	18,120	0.731	159	116
Southeast	Large Fringe	23,284	16,312	0.701	129	91	33,283	26,869	0.807	153	123
Southeast	Medium/Small	48,893	30,717	0.628	177	111	72,189	52,697	0.730	216	158
Southeast	Non Metro	32,753	21,283	0.650	262	170	45,607	32,548	0.714	304	217
Southwest	Region Total	66,456	47,001	0.707	188	133	84,976	67,822	0.798	198	190
Southwest	Large Central	23,776	15,934	0.670	150	100	31,628	25,316	0.800	165	132
Southwest	Large Fringe	6,361	4,627	0.727	110	80	10,242	9,394	0.917	143	131
Southwest	Medium/Small	23,475	16,503	0.703	254	178	25,968	19,933	0.768	233	179
Southwest	Non Metro	12,844	9,937	0.774	288	223	17,138	13,179	0.769	320	246

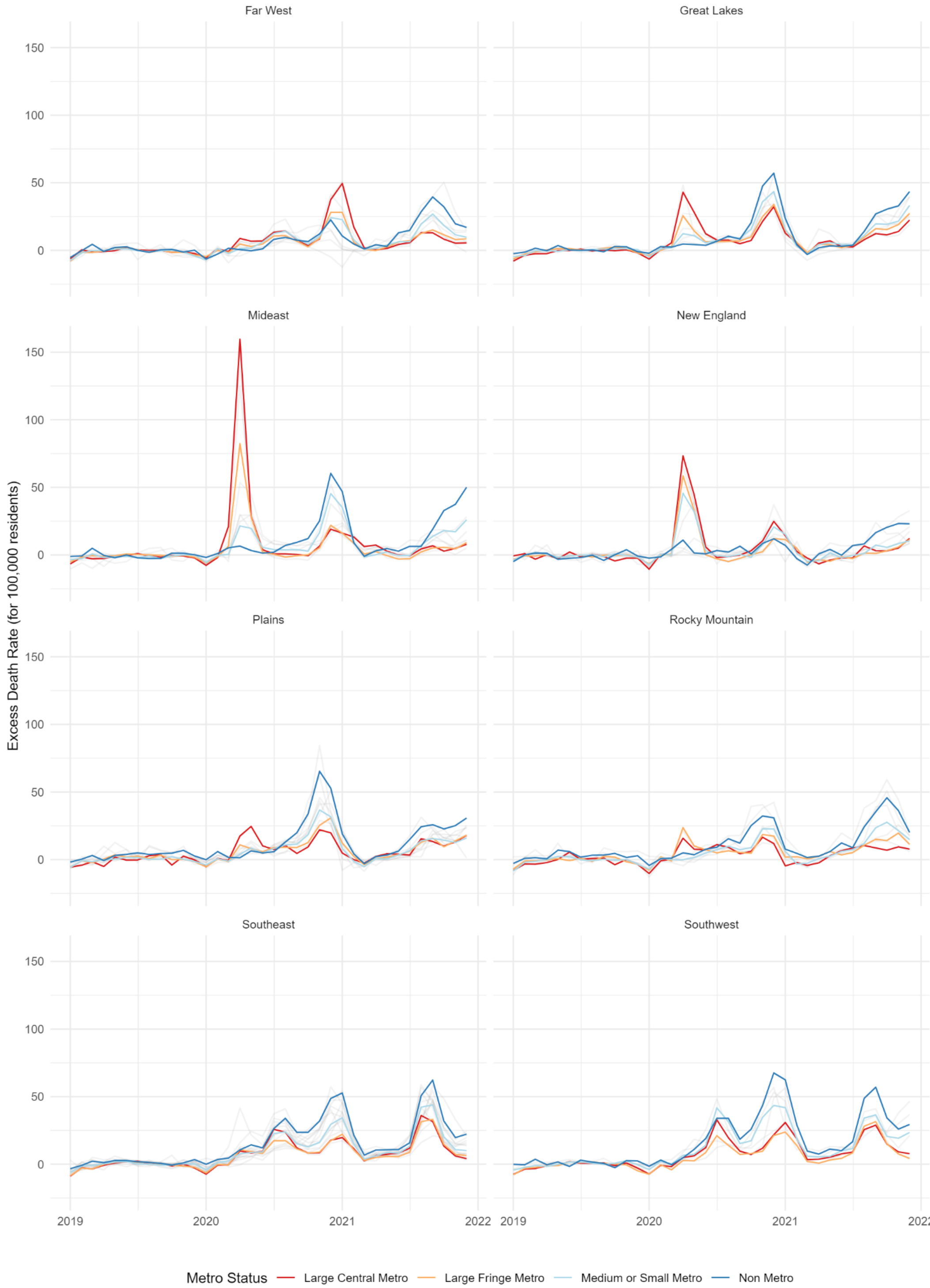
Notes: Death rates were calculated by aggregating deaths and population over counties within each BEA region and metropolitan-non-metropolitan area. Estimates for 2020 correspond to the period March - December 2020. Estimates for 2021 refer to the period January to December, 2021.

Figure 1. County excess mortality rates per 100,000, 2020-2021



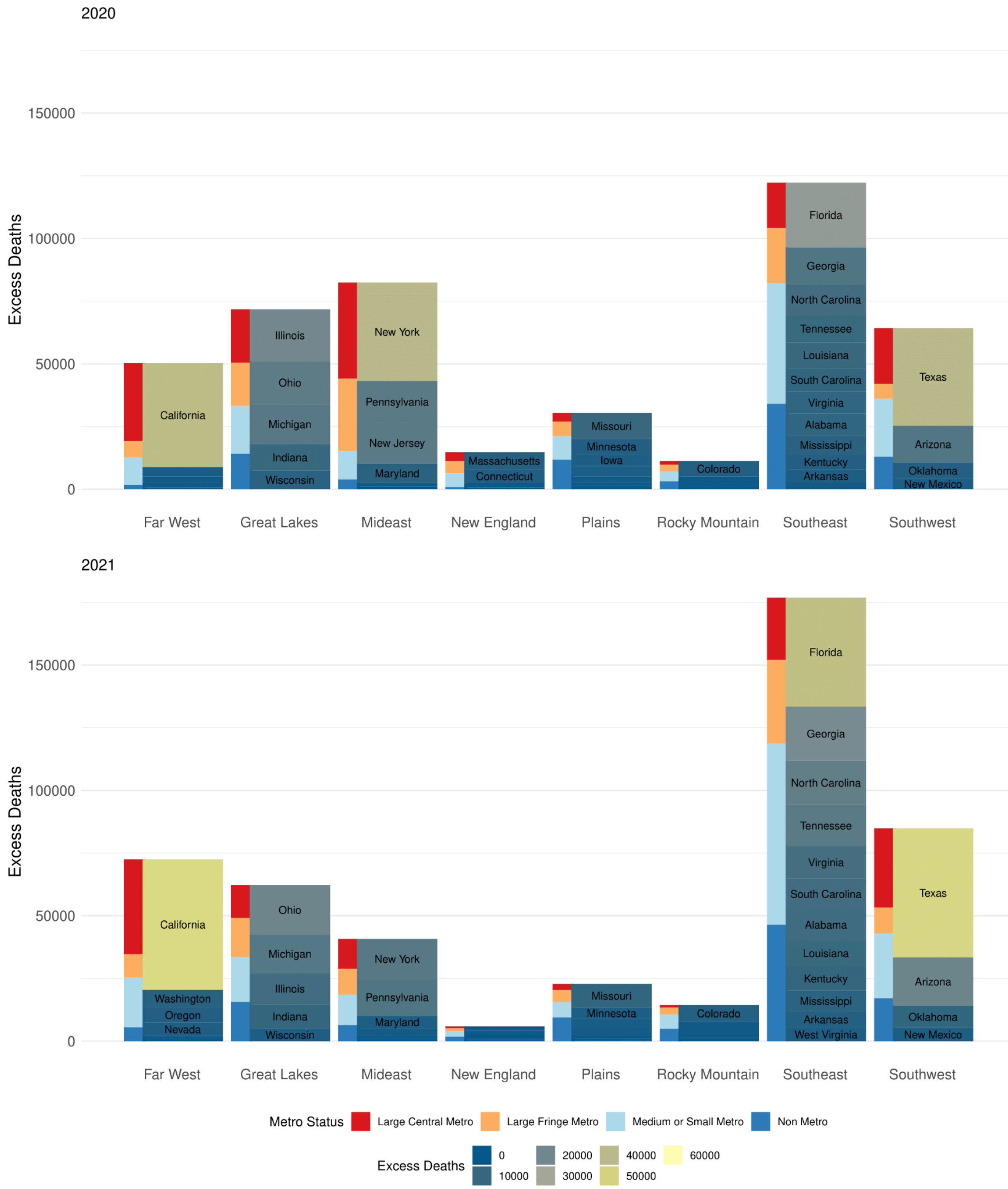
Notes: Panels A and B show the geographic distribution of excess death rates in 2020 (A) and 2021 (B) as estimated by comparing the expected number of deaths from our model to the actual number of deaths. Panels C and D report excess and COVID deaths rates for the counties with the highest excess deaths rates 2020 and 2021 respectively. Counties with less than 30,000 residents and less than 60 COVID deaths across the two years were excluded from the rankings in the barplots.

Figure 2. Monthly Excess Deaths Rates by BEA Region and Metro Status, 2019-2021



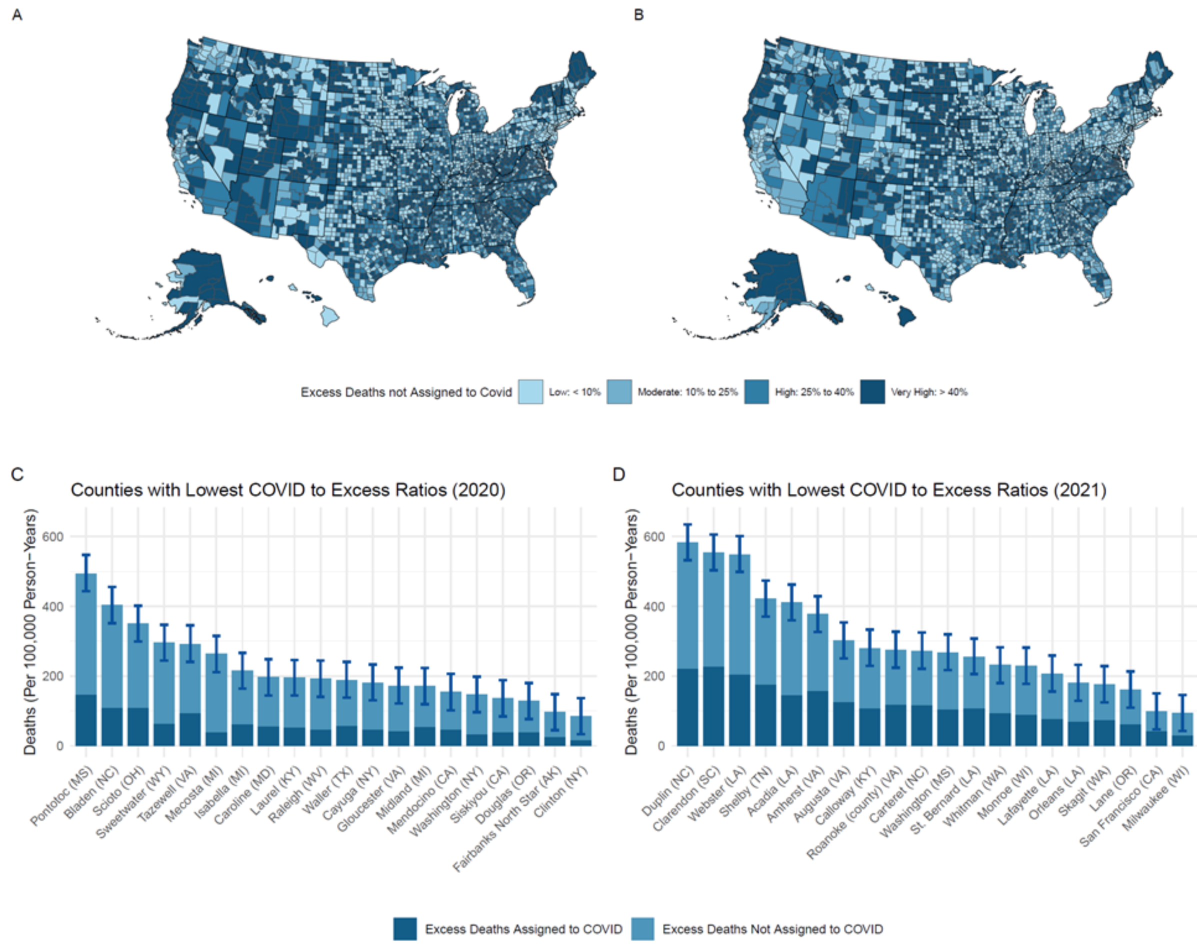
Notes: This graph shows aggregated trends in excess mortality at the monthly level between 2019-2021 stratified by BEA region and metro status.

Figure 3. Excess Deaths by BEA Region and Metro Status, and by State, 2020-2021



Notes: The top panel presents absolute excess deaths by BEA region disaggregated by metro status (left) and state (right) for 2020; the bottom panel shows the same for 2021.

Figure 4. Percentage of Excess Deaths not Assigned to Covid-19, 2020-2021



Notes: Panels A (2020) and B (2021) show the geographic distribution of proportion of excess deaths not assigned to COVID in 2020 and 2021. Panels C (2020) and D (2021) report excess and COVID deaths rates for the counties with the lowest COVID to excess ratios in 2020 and 2021. Counties with less than 30,000 residents and less than 60 COVID deaths across the two years were excluded from the rankings in the barplots.

Supplementary Information

Paglino E, Lundberg DJ, Cho A, Wasserman JA, Raquib R, Luck AN, Hempstead K, Bor J, Elo IT, Preston SH, Stokes AC. Excess all-cause mortality across counties in the United States, March 2020 to December 2021.

Table S1. Excess mortality, Covid-19 mortality, and the ratio of Covid-19 to excess mortality by US state.

Table S2. Performance of different models on training and test data.

Figure S1. Actual and expected mortality trends by month at the national level, 2015-2021

Figure S2. Time-series plots for the largest counties in each BEA region and metro-status combination

Figure S3. Excess Death Rates by BEA Region and Metro Status, and by State, 2020-2021

Figure S4. State-level rankings of COVID-19 to excess mortality ratios for 2020 and 2021

Figure S5. A Comparison of Excess and Covid-19 Deaths across BEA Regions and Metro Status.

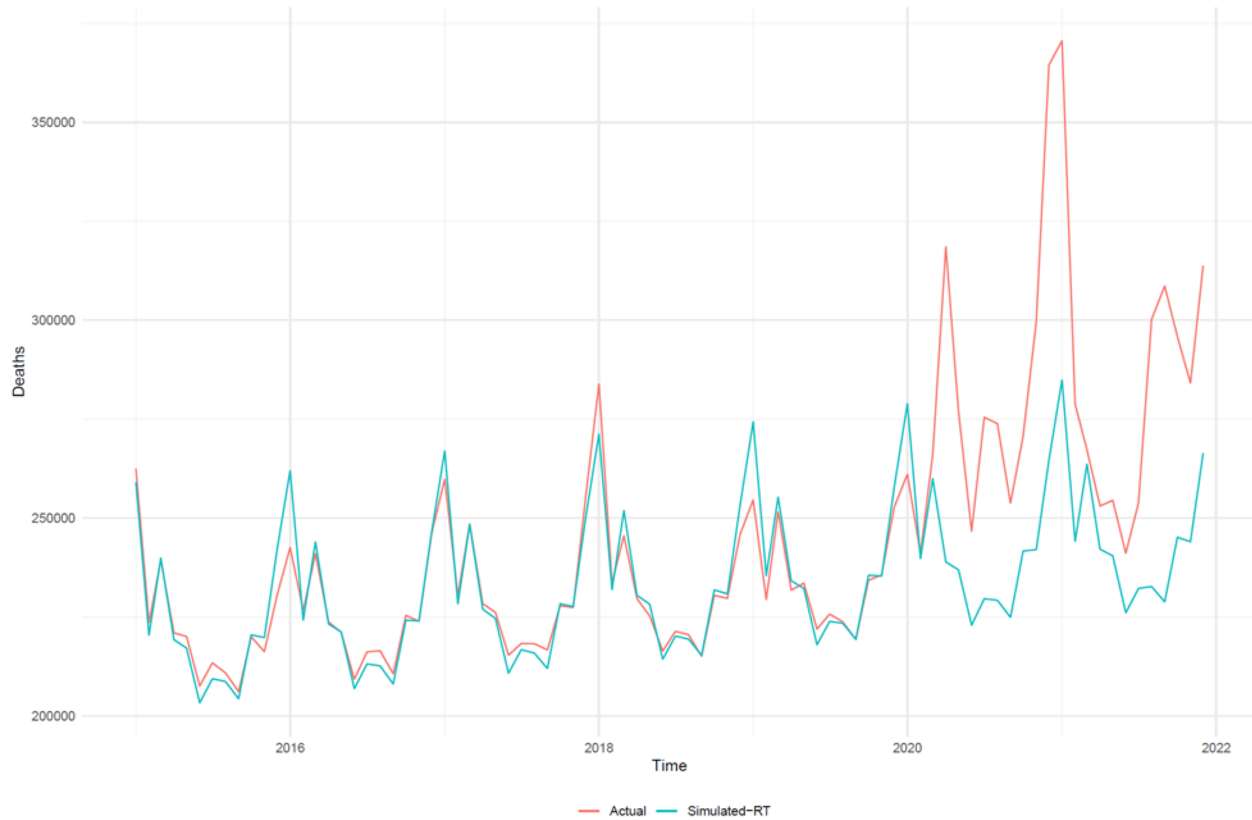
Methods Supplement. Details on data extraction, geographic classifications, model selection and estimation

Table S1. Excess mortality, Covid-19 mortality, and the ratio of Covid-19 to excess mortality by US state

	2020					2021				
	Number of Deaths			Rates per 100,000 PY		Number of Deaths			Rates per 100,000 PY	
State	Excess	COVID	Ratio	Excess	COVID	Excess	COVID	Ratio	Excess	COVID
Alabama	9,002	6,543	0.727	216	157	11,959	9,474	0.792	237	188
Alaska	478	275	0.575	78	45	1,368	729	0.533	187	100
Arizona	14,755	8,447	0.572	246	141	19,246	12,690	0.659	265	175
Arkansas	4,938	3,520	0.713	197	140	6,455	4,726	0.732	213	156
California	42,806	31,339	0.732	130	95	51,982	44,435	0.855	132	113
Colorado	6,432	4,346	0.676	133	90	6,708	5,277	0.787	115	91
Connecticut	5,826	5,782	0.992	194	193	1,304	2,719	2.085	36	75
Delaware	1,430	1,008	0.705	173	122	1,341	1,111	0.828	134	111
District of Columbia	1,121	831	0.741	195	144	810	463	0.572	120	69
Florida	27,150	19,230	0.708	151	107	43,434	34,509	0.795	200	159
Georgia	14,921	9,431	0.632	167	106	21,443	15,774	0.736	199	146
Hawaii	199	342	1.719	17	28	795	682	0.858	55	47
Idaho	1,327	1,350	1.017	86	88	2,418	2,387	0.987	128	126
Illinois	20,979	15,717	0.749	197	148	12,299	11,267	0.916	97	89
Indiana	10,496	8,520	0.812	186	151	9,516	8,545	0.898	140	126
Iowa	4,492	4,315	0.961	169	163	2,281	3,059	1.341	71	96
Kansas	3,964	3,287	0.829	162	135	3,579	3,567	0.997	122	122
Kentucky	5,925	4,112	0.694	158	110	9,488	7,434	0.784	210	165
Louisiana	10,394	6,520	0.627	268	168	10,573	6,322	0.598	228	137
Maine	454	419	0.923	40	37	1,737	1,330	0.766	127	97
Maryland	8,275	5,995	0.724	161	117	5,073	5,251	1.035	82	85
Massachusetts	7,222	9,329	1.292	124	160	1,002	4,870	4.860	14	70
Michigan	16,246	11,398	0.702	194	136	15,365	13,588	0.884	153	135
Minnesota	5,861	5,211	0.889	123	110	4,152	4,429	1.067	73	78
Mississippi	7,431	4,460	0.600	301	181	8,007	5,076	0.634	271	172
Missouri	10,296	7,129	0.692	201	139	9,269	7,717	0.833	150	125
Montana	1,315	1,125	0.856	145	124	1,703	1,594	0.936	155	145
Nebraska	2,148	2,079	0.968	132	127	1,051	1,653	1.573	54	84
Nevada	3,937	3,242	0.823	152	125	5,338	5,136	0.962	170	164
New Hampshire	747	766	1.025	65	67	823	1,111	1.350	59	80
New Jersey	17,156	16,497	0.962	223	215	2,928	8,402	2.870	32	91
New Mexico	4,192	2,851	0.680	238	162	5,277	3,464	0.656	249	164
New York	40,721	35,723	0.877	244	214	16,006	21,650	1.353	81	109
North Carolina	12,842	7,874	0.613	147	90	17,638	13,551	0.768	167	129
North Dakota	1,218	1,203	0.988	188	186	159	629	3.956	21	81
Ohio	17,793	13,605	0.765	181	139	19,658	18,244	0.928	167	155
Oklahoma	6,300	4,832	0.767	191	146	8,783	7,230	0.823	221	182
Oregon	2,315	1,444	0.624	66	41	6,088	3,664	0.602	143	86
Pennsylvania	17,221	16,607	0.964	159	154	14,624	18,138	1.240	113	140
Rhode Island	1,283	1,585	1.235	141	174	481	978	2.033	44	89
South Carolina	9,367	5,277	0.563	219	123	12,448	9,062	0.728	240	175
South Dakota	1,565	1,481	0.946	211	200	659	760	1.153	74	85
Tennessee	10,929	6,831	0.625	190	119	16,309	11,874	0.728	234	170
Texas	41,209	30,871	0.749	169	127	51,670	44,438	0.860	175	151
Utah	1,879	1,367	0.728	69	50	1,988	2,241	1.127	60	67
Vermont	220	135	0.614	41	25	492	261	0.530	76	40
Virginia	8,805	5,789	0.657	123	81	12,807	8,874	0.693	148	103
Washington	3,204	3,296	1.029	50	51	6,991	5,430	0.777	90	70
West Virginia	2,567	1,478	0.576	172	99	5,318	3,558	0.669	298	199
Wisconsin	7,549	5,420	0.718	154	111	4,928	4,658	0.945	84	79
Wyoming	862	461	0.535	179	96	1,376	1,017	0.739	238	176

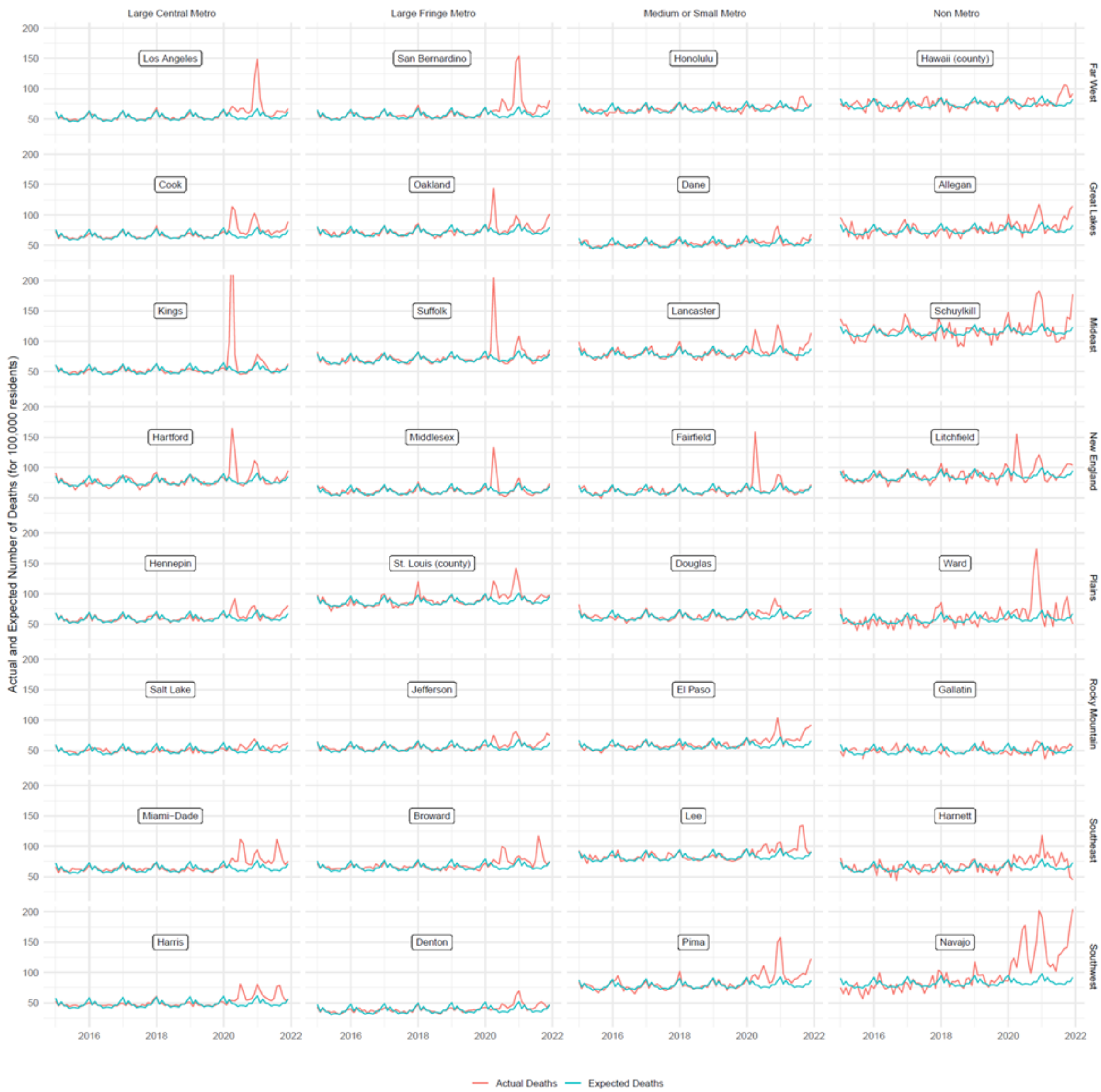
Notes: Death rates were calculated by aggregating deaths and population over counties within each state. Estimates for 2020 correspond to the period March - December 2020. Estimates for 2021 refer to the period January to December, 2021.

Figure S1. Actual and expected mortality trends by month at the national level, 2015-2021



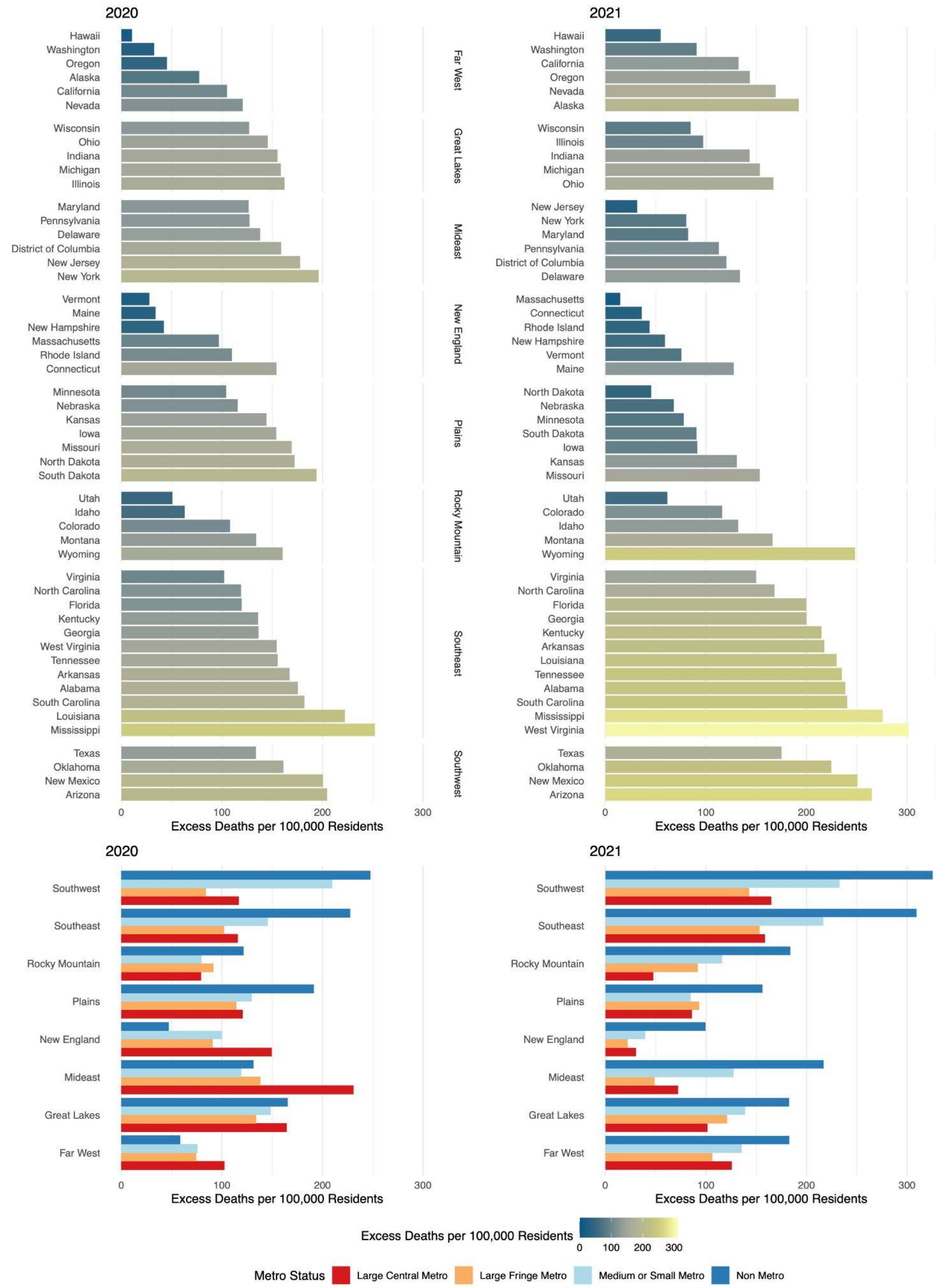
Notes: Monthly trends in actual and expected deaths at the national level are obtained by aggregating the estimates from the county level.

Figure S2. Time-series plots for the largest counties in each BEA region and metro-status combination



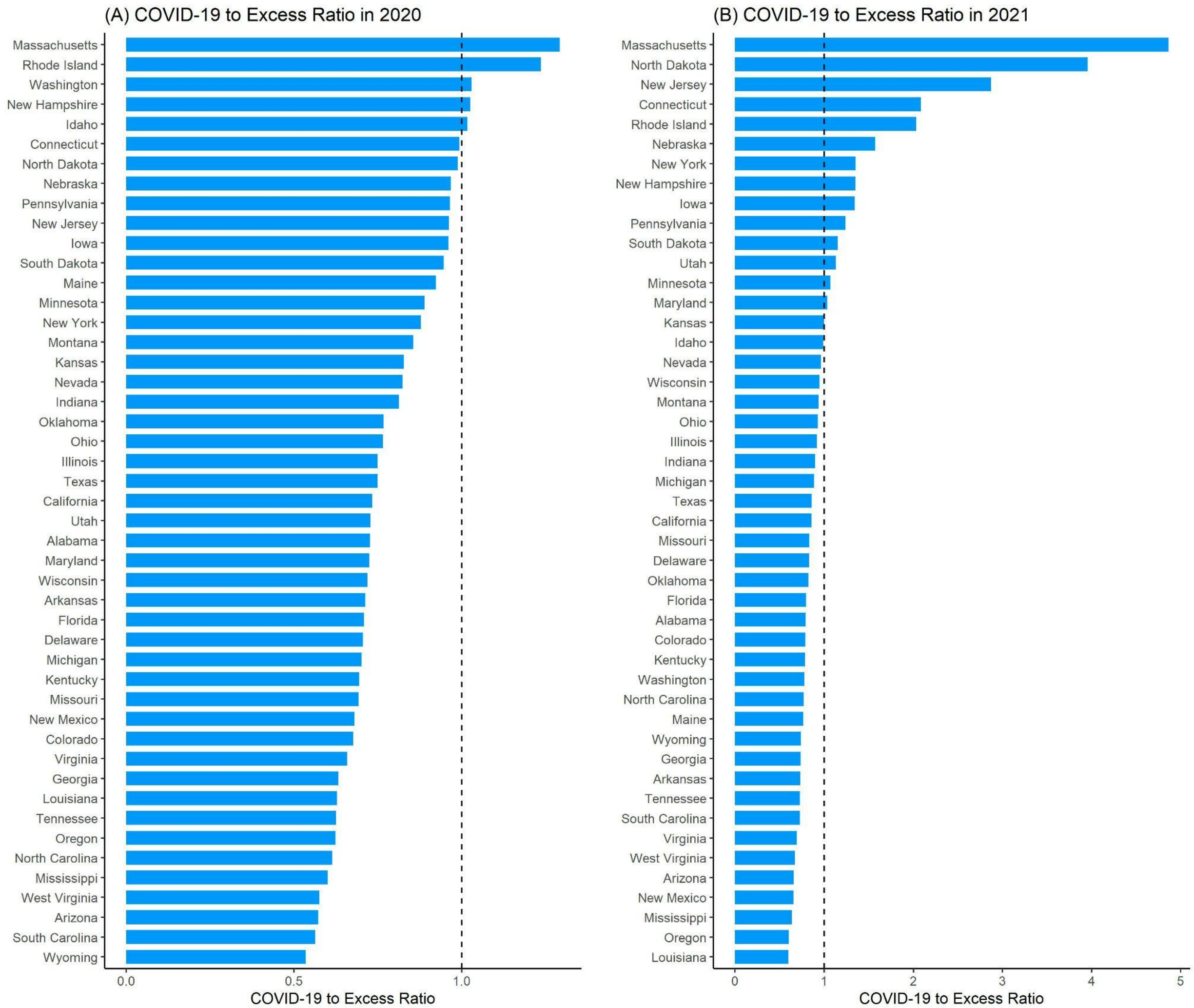
Notes: The plot shows the largest counties within each BEA region and metro status category. Time series of actual and expected all-cause deaths per 100,000 are plotted over 2015-2021.

Figure S3. Excess Death Rates by BEA Region and Metro Status, and by State, 2020-2021



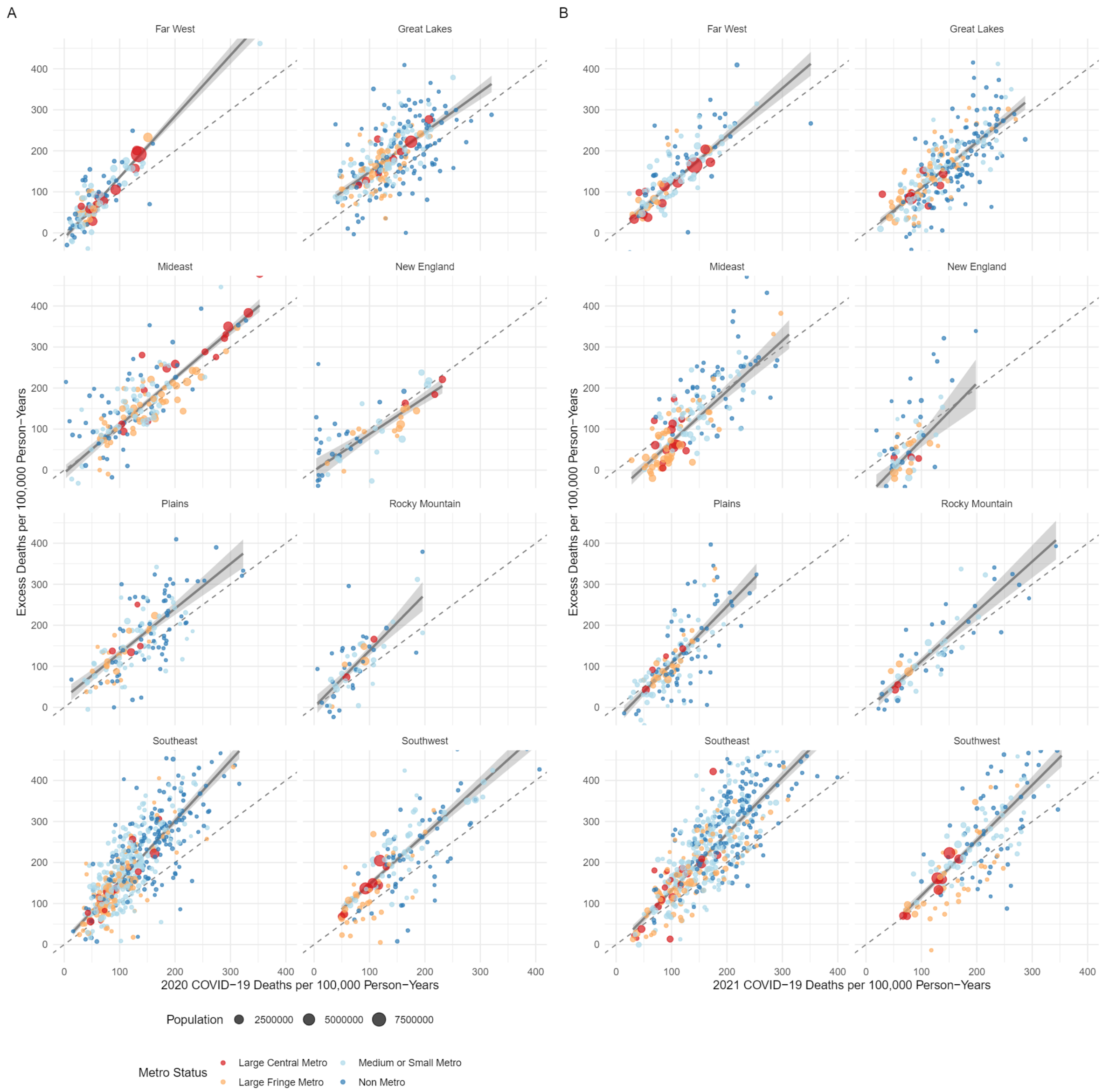
Notes: The top panel shows excess death rates by state organized by BEA region for 2020 (left) and 2021 (right). The bottom panel shows excess death rates by metro status organized by BEA region for 2020 (left) and 2021 (right).

Figure S4. State-level rankings of COVID-19 to excess mortality ratios for 2020 and 2021.



Notes: These graphs present state-level rankings of Covid-19 to excess mortality ratios for 2020 and 2021. The dotted vertical line represents a Covid-19 to excess ratio of 1.0, indicating that Covid-19 deaths fully account for estimated excess deaths. States with very few Covid-19 deaths over 2020-2021 were excluded (Alaska, Hawaii, Vermont) from the graph.

Figure S5. A Comparison of Excess and Covid-19 Deaths across BEA Regions and Metro Status.



Notes: Panels A and B show the relationship between COVID and excess deaths in 2020 and 2021 respectively. Counties with less than 30,000 residents were excluded. The regression lines reflect the coefficients of a set of regressions of COVID deaths on excess deaths within each BEA region using population as weights.

Methods Supplement

Data Extraction from CDC WONDER

The CDC WONDER online database query system found at https://wonder.cdc.gov/ was used to extract all mortality data used in this project. To obtain death counts for all-causes mortality, we used the Multiple Cause of Death (Final) database from 1999-2020. We obtained two sets of extracts, one for data at the county-year level and one for data at the county-month level.

For county-year extracts, the data request was submitted for the time period of interest using the request form with the following settings changed from the default:

- Tab 1. Organize table layout: Group results by County and by Year
- Tab 4. Select time period of death: specific period
- Tab 6. Select underlying cause of death: *All* (All Causes of Death)
- Tab 8. Other options: checking Export Results and Show totals. The request generates a text file.

For county and month for the years 1999 through 2019, the data was extracted using six month time periods due to limitations of the CDC Wonder servers. The data request was submitted for each time period of interest using the request form with the following settings changed from the default:

- Tab 1. Organize table layout: Group results by County and by Month
- Tab 4. Select time period of death: specific period (6 months at a time)
- Tab 6. Select underlying cause of death: *All* (All Causes of Death)
- Tab 8. Other options: checking Export Results and Show totals. The request generates a text file.

To extract data for the time periods of March 2020 to December 2021, we used the Multiple Cause of Death (Provisional) database from 2018 – Last Month database. The data requests were submitted for each time period of interest using the request form with the following settings changed from the default:

- Tab 1. Organize table layout: Group results by County and by Year
- Tab 4. Select time period of death: March 2020 to December 2021
- Tab 6. Select underlying cause of death: *All* (All Causes of Death)
- Tab 8. Other options: checking Export Results and Show totals. The request generates a text file.

To extract counts of COVID deaths for 2020 and 2021 at the county-year level we used the Multiple Cause of Death (Provisional) database from 2018 – Last Month database. The data requests were submitted for each time period of interest using the request form with the following settings changed from the default:

- Tab 1. Organize table layout: Group results by County and by Year
- Tab 4. Select time period of death: March 2020 to December 2021
- Tab 6. Select underlying cause of death: U07.1 (COVID-19)
- Tab 8. Other options: checking Export Results and Show totals. The request generates a text file.

Geographic Classifications

<u>USDA/ERS/NCHS Metropolitan-Nonmetropolitan Categories:</u>

- **Large central metros**: counties in metropolitan statistical areas with a population of more than 1 million.

- **Large fringe metros**: counties that surrounded the large central metros
- **Small or medium metros**: counties in metropolitan statistical areas with a population between 50,000 and 999,999.
- **Nonmetropolitan areas**: all other counties.

BEA Regions:

- **New England**: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont
- **Mideast**: Delaware, District of Columbia, Maryland, New Jersey, New York, and Pennsylvania
- **Great Lakes**: Illinois, Indiana, Michigan, Ohio, and Wisconsin
- **Plains**: Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota
- **Southeast**: Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia
- **Southwest**: Arizona, New Mexico, Oklahoma, and Texas
- **Rocky Mountain**: Colorado, Idaho, Montana, Utah, and Wyoming
- **Far West**: Alaska, California, Hawaii, Nevada, Oregon, and Washington

County-Sets:

- Counties with populations greater than or equal to 50,000 were not grouped into county-sets
- Counties with populations less than 50,000 were grouped with contiguous counties to form county-sets with populations greater than 50,000
- County-sets did not cross state borders

Model Selection and Estimation

Before settling on the models used in the paper (Random Time), we tested a simple model (Base), as well as a more complex one (Random Time Plus). Model Base shares the fixed effect of the Random Time model described in the methods section but has a simpler random effects structure including only random intercepts for counties. Model Random Time Plus borrows the random effects structure of the Random Time model but allows the month dummy variables (used to capture seasonality) to vary across census regions. We hypothesized that, because seasonal patterns in mortality are not uniform across the United States, such a model would have achieved a better fit. To evaluate the model performance, we trained them on data for the period 2014-2018 and then tested them on data for 2019. Table S2 reports the Root Mean Squared Error for all models. While the Random Time Plus model outperformed the two simpler alternatives in the training data, the Random Time model achieved the best performance in the test data. This suggests that the Random Time Plus Model was overfitting the training data to a larger extent compared to the simpler models. Looking at the AIC and BIC confirms our insights. According to the BIC, we should select the Random Time model. According to the AIC the Random Time Plus model should instead be preferred. Overall, because the Random Time model outperformed the Base Model in all data sets, has the lowest BIC, and is the best performing model on the test data, we chose to use it as the final model in the paper.

Table S2. Model Comparison

Performance	Base	Random Time	Random Time Plus
Training Data (2014-2018)	17.037	16.989	16.972
Test Data (2019)	17.981	17.955	17.961
Overall	17.231	17.187	17.175
BIC	1002787	1002431	1002579
AIC	1002642	1002248	1002048