medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Excess Mortality in the United States During the First Peak of the COVID-19 Pandemic

R. Rivera. (1)* J. E. Rosenbaum. (2) W. Quispe. (1)

1. College of Business, University of Puerto Rico at Mayagüez, Mayagüez, Puerto Rico.
2. Department of Epidemiology and Biostatistics, School of Public Health, SUNY Downstate Health Sciences University, Brooklyn, NY.

*Corresponding author: Roberto Rivera, PhD, College of Business, University of Puerto Rico, Mayaguez, roberto.rivera30@upr.edu.

Data Availability: Data is publicly available but updated weekly. A copy of the data release will be shared through GitHub.

Acknowledgements

We thank Farida Ahmen, Michael Augenbraun, Stephan Kohlhoff, Josh Marshall, and Tracey Wilson for helpful conversations, and our spouses/partners for childcare. Work was partially supported by NSF Grant OAC-1940179

Authors have no conflict of interest to report.

Covid-19 Pandemic Excess Mortality U.S.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

1 Abstract

Deaths are frequently under-estimated during emergencies, times when accurate mortality estimates are crucial for pandemic response and public adherence to non-pharmaceutical interventions. This study estimates excess all-cause, pneumonia, and influenza mortality during the COVID-19 health emergency using the June 12, 2020 release of weekly mortality data from the United States (U.S.) Mortality Surveillance Survey (MSS) from September 27, 2015 to May 9, 2020, using semiparametric and conventional time-series models in 9 states with high reported COVID-19 deaths and apparently complete mortality data: California, Colorado, Florida, Illinois, Massachusetts, Michigan, New Jersey, New York, and Washington. The May 9 endpoint was chosen due to apparently increased reporting lags in provisional mortality counts. We estimated greater excess mortality than official COVID-19 mortality in the U.S. (excess mortality 95% confidence interval (CI) (80862, 107284) vs. 78834 COVID-19 deaths) and 6 states: California (excess mortality 95% CI (2891, 5873) vs. 2849 COVID-19 deaths); Illinois (95% CI (4412, 5871) vs. 3525 COVID-19 deaths); Massachusetts (95% CI (5061, 6317) vs. 5050 COVID-19 deaths); New Jersey (95% CI (12497, 15307) vs. 10465 COVID-19 deaths); and New York (95% CI (30469, 37722) vs. 26584 COVID-19 deaths). Conventional model results were consistent with semiparametric results but less precise.

Official COVID-19 mortality substantially understates actual mortality, suggesting greater case-fatality rates. Mortality reporting lags appeared to worsen during the pandemic, when timeliness in surveillance systems was most crucial for improving pandemic response.

Abbreviations:

NCHS	National Center for Health Statistics
US	United States
CI	confidence intervals
MSS	Mortality Surveillance System

2 Introduction

The number of COVID-19 deaths may be under-reported. The Centers for Disease Control and Prevention issues guidelines to determine cause of deaths, but underestimating the death toll of natural disasters, heatwaves, influenza, and other emergencies is common. In some cases, the underestimates can be extreme: chikungunya was associated with only 31 deaths during a 2014-15 epidemic in Puerto Rico, but time series analysis estimated excess mortality of 1310 deaths [1]. Hurricane Maria's mortality was officially only 64 deaths, but a 95% confidence interval for estimated excess mortality was between 1,069 and 1,568. [2] Deaths directly due to the COVID-19 pandemic may be underestimated due to under-diagnosis [3],insufficient postmortem SARS-CoV-2 tests, not seeking health care [4], and ascertainment bias. Deaths indirectly due to an emergency are also common, due to an overloaded health system [5] or lack of healthcare access for routine care: during the 2014 West Africa Ebola epidemic, lack of routine care for malaria, HIV/AIDS, and tuberculosis led to an estimated 10,600 additional deaths in the area.[6] Health emergencies may also lead to indirect deaths from economic, social, and emotional stress [7] and crowded emergency departments.[8]

In this study, we estimate excess all-cause, pneumonia, and influenza mortality during the COVID-19 pandemic; which estimates deaths both directly and indirectly related to COVID-19.Directly related COVID-19 deaths include deaths in patients who have undetected SARS-

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

CoV-2 due to false negative tests [9], not seeking health care [4], or being turned away from the emergency department due to emergency department crowding [8]. Testing and forensic staff shortfalls lead to lack of postmortem testing. Indirect deaths may include deaths among patients due to emergency department crowding [8]; avoidance of hospitals due to fear of the virus; or avoidance of health care due to the accompanying economic recession, such as loss of employment or income, or loss of health insurance coverage.[10]

3 Methods

3.1 Data and measures

Weekly all-cause, pneumonia, and influenza mortality for each state from September 27, 2015 (week 40) to May 9, 2020 (week 19) were obtained from the National Center for Health Statistics (NCHS) Mortality Surveillance Survey data release on June 12, 2020. The timeliness in death certificate reporting varies by region, state, and cause of death.[11] We use pneumonia and influenza mortalities because COVID-19 deaths could be misclassified as pneumonia or influenza. The NCHS divides New York State into two jurisdictions: New York City (NYC) and non-NYC New York State, and we leave them separate for plots and combine them for statistical models.

Our statistical method can estimate excess mortality in regions with considerable fatalities. In regions with small number of deaths, our method may not be able to detect the effects of the emergency because the confidence intervals are too wide. We identified 9 jurisdictions within this data that had high numbers of reported COVID-19 deaths and appeared to have complete death data, based on the NCHS measure of completeness. The states were: California, Colorado, Florida, Illinois, Massachusetts, Michigan, New Jersey, New York, and Washington.

Population estimates were from Vintage 2019 Census yearly estimates for July 1 of each year 2010–19, which were used to obtain weekly population estimates.[12]

We obtained COVID-19 mortality as of 5/9/2020 from the COVID-19 Tracking Project, the New York Times, and the Centers for Disease Control Provisional Deaths.[13, 14, 15] These three mortality counts differ, so We used the higher number in all cases to be conservative. Usually, CDC provisional counts were the highest estimate.

To assess whether the COVID-19 emergency was associated with reduced emergency department (ED) utilization, we used data from the National Syndromic Surveillance Program (NSSP). We identified 3 of the United States top 5 causes of death that present with acute symptoms that require immediate treatment, for which the choice not to seek health care may result in death: heart disease, chronic lower respiratory diseases, and cerebrovascular disease. The only such data available were ED visits in New York City for asthma symptoms. We obtained daily counts of asthma ED visits, age group (5–17, 18–64, 65+), borough, and date from the New York City Department of Health and Mental Hygiene's EpiQuery website from January 1, 2016 to May 9, 2020. We were unable to obtain data from the other 58 NSSP jurisdictions, or for acute coronary syndromes that could suggest stroke or heart attack.

This study is an analysis of publicly available data in broad categories such that individuals cannot be identified, so it is not human subjects research and is exempt from requiring human subjects board review.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

3.2 Statistical Analysis

We construct two models to capture the temporal behavior of death certificate data to estimate excess mortality during the COVID-19 pandemic: a semiparametric model and a conventional model estimating the difference in current death totals starting from the beginning of a health emergency and the projected deaths under normal conditions [16].

3.3 Semiparametric model

We use a semiparametric model to capture the temporal behavior of all-cause mortality while measuring how the beginning of an emergency alters the 'normal' mortality pattern. This model was successfully deployed to estimate excess deaths due to Hurricane Maria in Puerto Rico [2].

We use a general additive model with the following covariates: the natural logarithm of population as an offset, a smooth function of week of the year, year, and an indicator function that presents the possibility that the mean death rate has increased after the start of the emergency in a location. Week of the year is modeled non-parametrically with a penalized cyclic cubic regression spline function to capture seasonal mortality variations [15]. Preliminary analysis indicated overdispersion, so we used a quasi-Poisson model to estimate the dispersion parameter [17]. We estimated coefficients by a penalized likelihood maximization approach, where the smoothing penalty parameters are determined by restricted maximum likelihood. The residuals of the fitted model did not present remaining temporal dependence.

To estimate cumulative excess deaths, we sum the coefficients of the indicator for the start of the emergency. Approximate simulations from the Bayesian posterior density are performed to obtain 95% credibility intervals, which we will refer to as confidence intervals throughout.

For each state, we determine excess deaths due to the COVID-19 emergency from a starting point through May 9, the date of the most recent complete mortality data. The starting point was the date after the most recent inflection point in all-cause mortality, suggesting the onset of the COVID-19 emergency, chosen to balance concerns of deaths prior to the official first cases and the sensitivity of the model to detect small excess mortality in the limited available data, exacerbated by the provisional counts being lower than actual deaths. This semiparametric model can incorporate population displacement [2], but we are not aware of significant displacement during the pandemic. Data from New York City Department of Sanitation indicates that 4 wealthy Manhattan community districts — community districts 3, 7, 8, and 11, with a combined population of under 800,000 people – had 3–5% lower tonnage for the month of March 2020 than March 2019, whereas New York City as a whole had 4.1% more waste collected in March 2020 than March 2019.[18] If the population of these community districts decreased by about 10%, the total population of New York City decreased by less than 1%.

The models were fit using version 1.8-28 of the *mgcv* package in R 3.6.3 [19, 20]. All code and data will be made publicly available.

3.4 A more conventional excess deaths analysis

We also estimate excess mortality using a conventional excess mortality method: we fit a quasi-Poisson semiparametric model as above until February 1, 2020. The deaths from February 8, 2020 forward should follow a Poisson distribution with some expected rate; the maximum likelihood estimator of such rate is the mean weekly deaths during this period. Weekly deaths

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

will be approximately normal with a variance that accounts for overdispersion according to the scale parameter in the fitted model. From this distribution, we simulate 10,000 weekly deaths and subtract the results from the posterior distributions of the fitted results. The 95% confidence interval is the range between the 2.5% and 97.5% percentiles of all excess deaths.

3.5 Assessing reduction in emergency department utilization

To evaluate whether there was a reduction in emergency department visits during the COVID-19 emergency, we fit a quasi-Poisson regression for daily number of asthma-related emergency department visits in New York City syndromic surveillance data, controlling for borough, age group, and date, with an indicator for date after 3/21/20, which estimated a dispersion parameter of 1.95.

4 Results

Provisional death counts increase with each data release, especially for recent weeks. The variation on reporting timeliness by state hinders excess death assessment for the United States. Figure 1 illustrates how the pandemic appeared to increase mortality reporting lags and the latest release of weekly mortality counts 'blankets' each previous release.

California, Illinois, Massachusetts, New Jersey, New York City, and New York State show greater all-cause weekly mortality exceeding official COVID-19 death toll (Figure 2). All states have excess pneumonia mortality, but only New York City has excess influenza deaths (Figure 3)

In the Unite States, we observe greater all-cause and pneumonia deaths (Figure 4). Using the semiparametric model, between March 21, 2020 and May 9, 2020 we are 95% confident that all-cause excess deaths in the United States were between 80862 and 107284, compared with 78834 reported COVID deaths. This is at least 2028 more deaths than official COVID-19 deaths for the same time period. Pneumonia excess deaths were between 36750 and 43777. Using the conventional method we found a 95% confidence interval of 80666 and 116859 for all-cause excess deaths, and 37782 and 43934 for pneumonia deaths.

Table 1 presents for each state 95% confidence intervals for all-cause, pneumonia, and influenza excess deaths. We estimated greater excess mortality than COVID-19 deaths in 5 of 9 states: California, Illinois, Massachusetts, New Jersey, and New York. Colorado, Michigan, Washington and Florida present significant all-cause excess mortality but there was no evidence of them exceeding official COVID-19 death counts. However, relative to expected numbers of pneumonia and influenza deaths, we observe excess pneumonia deaths in all states studied, and excess influenza deaths in New York (Table 1). We evaluated influenza mortality in the District of Columbia, the other urban-only area, but influenza deaths were not higher than usual (not shown), in contrast to New York City.

With the conventional model, we estimate more excess all-cause deaths in Illinois, Massachusetts, New Jersey, and New York than reported COVID-19 cases (Table 1); and once more excess pneumonia deaths are observed for all states. These results are also consistent with the excess mortality results for the earlier data release covering through May 9, 2020 (not shown).

The results of the semiparametric and conventional models are generally compatible. But since our conventional model also reflects random fluctuations in mortality, prediction intervals are obtained, which naturally present larger variation than confidence intervals. Treating the

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

deaths as fixed [16] would lead to intervals that are misleadingly short on width. The larger uncertainty in the conventional model makes it harder to interpret excess deaths.

5 Discussion

Mortality underestimation may reduce the public's willingness to adhere to costly and stressful non-pharmaceutical interventions, such as governors' orders to stay at home, wear masks, and engage in social distancing. The conventional method generally resulted in intervals of larger width. Since this hinders interpretation of excess deaths, we focus on interpreting results from the semiparametric model. We find substantial excess all-cause mortality that exceeds the number of documented COVID-19 deaths in California, Illinois, Massachusetts, New Jersey, and New York. The CDC has established guidelines for certifying COVID-19 deaths [21] and whether to collect postmortem specimens for SARS-CoV-2 testing [22]. Underestimation of mortality may be due to subjective interpretations of guidelines, infrastructure limitations caused by the emergency, and undetected SARS-CoV-2 due to lack of outpatient and postmortem testing that could allow medical examiners to discriminate between deaths from COVID-19 and deaths among people infected with SARS-CoV-2.

The CDC also estimates point estimates of excess mortality, but our methods quantify uncertainty. The semiparametric method yields more precise confidence intervals than a conventional approach, which must estimate prediction intervals. The semi parametric model is less affected by under reporting than the conventional approach. Under reporting of death counts occur when health departments do not submit all death certificates during the pre-pandemic period, as sometimes occurs even during non-emergency periods.

We quantify higher pneumonia mortality than expected in all 9 states and higher influenza deaths primarily in New York. For New Jersey, even when adjusting for COVID-19 official deaths, pneumonia and influenza mortality, excess deaths are still significant. This raises the prospect of COVID-19 deaths being misclassified or indirect deaths being substantial in this state. With only limited cause of death data, our analysis is unable to distinguish between COVID-19 deaths being misclassified and indirect deaths.

Many of the excess influenza deaths in New York appear to be misclassified COVID-19 deaths. The number of positive influenza tests in New York City decreased steadily throughout March 2020, with no steep drop around the date of the COVID-19 emergency declaration [23], consistent with the end of seasonal influenza rather than people with influenza not seeking care, suggesting that many excess influenza deaths were misclassified COVID-19 deaths, rather than a resurgence of undiagnosed influenza. The apparent misclassification of COVID-19 deaths as influenza in New York City does not appear attributable to urbanicity, as we observed no excess influenza mortality in the District of Columbia, the other urban-only area.

Our findings suggest large excess deaths due to the COVID-19 emergency, even with incomplete death counts during the mid-April mortality peaks due to the reporting lag [14]. COVID-19 test access, as measured by decreased percent of COVID-19 tests that are positive, has increased with time in the US [24], but substantial heterogeneity across states [13] means that excess mortality may not decrease substantially unless test availability increases in high-prevalence states. Although we find significant influenza excess deaths primarily for New York state, influenza excess deaths in other states may become apparent in subsequent data releases.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Some excess mortality may include indirect deaths from chronic lower respiratory diseases, such as deaths due to not seeking care for asthma syndrome due to fear of becoming infected with SARS-CoV-2 during the COVID-19 emergency. Evidence suggests that patients with heart attacks and stroke delayed seekubg care due to the COVID-19 emergency [25], which would explain additional excess mortality. However, emergency department visit data for acute coronary syndromes are not released in the NSSP, as for asthma syndrome ED visits. When these data are released, we would expect more deaths at home coded as chronic lower respiratory diseases, cardiovascular disease, and cerebrovascular diseases; normally, these are three of the top five causes of death in the US, so reduced care seeking could contribute substantially to excess mortality [26].

Patients who suspect COVID-19 may also be turned away from emergency departments: these patients would not have SARS CoV-2 test results, so they would not be coded as COVID-19 deaths. Many patients with ILI never seek health care, including patients likely to have severe effects: about 45% of patients with heart disease and 52% with COPD delay at least 3 days [27], and during a flu pandemic care-seeking increases by only about 10 percentage points [4]. Dyspnea/breathlessness predicts health care seeking for ILIs [28]. However, most severe or fatal COVID-19 cases do not present with dyspnea [29], and lung damage can be substantial even without dyspnea [30].

In addition, excess mortality may be due to sudden health declines, which have been observed in COVID-19 inpatient populations: patients decline in minutes from ambulatory/conversant to unresponsive and requiring resuscitation [31],and such sudden health declines could also occur in outpatient populations.

Some but not all of the excess COVID-19 mortality is captured by pneumonia excess mortality, but some sudden deaths seem to occur among patients without apparent pneumonia [32]. These deaths may be due to cardiac injury [33], kidney or liver injury [34], or hypercoagulability [35].

Excess mortality during heatwaves and influenza pandemics often shows mortality displacement (or harvesting effect), where subsequent mortality declines [36]. Longer observation periods will be able to detect mortality displacement, but inconsistent non-pharmaceutical intervention policies across the United States may cause elevated mortality due to COVID-19 to persist over longer durations than during typical influenza pandemics.

COVID-19 appears to be misclassified as influenza primarily in New York City. The misclassification may be due to heterogeneity in cause of death determination and/or COVID-19 presentations between New York City and the states examined. Alternatively, the misclassification may occur in many jurisdictions, but it is more detectable in NYC due to the large number of deaths.

5.1 Strengths and Limitations

Lags in death reporting in these provisional data suggest that this analysis understates excess all-cause, pneumonia, and influenza mortality, and these lags are differential by state. The data at the time of retrieval likely excludes deaths not yet reported and/or included in provisional data releases. Completeness varies by state because states submit death certificates at different rates using various methods. The measure of completeness included in these data — total weekly deaths divided by expected deaths based on average weekly deaths over the

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

entire year during the previous 3 years — cannot accurately assess completeness during any period of elevated mortality, including seasonal influenza increases.

Our hypothesized mechanisms for excess mortality are based on existing research, but the available mortality data are insufficient to test these hypotheses. For example, deaths at home are not published weekly or systematically. The available U.S. data includes data aggregated over all states stratified by age (0– 17, 18–64, 65+) and region; however, these data do not allow identification of all-cause mortality trends by region or age because some states are incomplete, even by the available completeness measure, which understates completeness during periods of elevated mortality.

Our analysis examines excess deaths associated with the pandemic emergency. It is likely that the longer the period since the beginning of the emergency, the more indirect causes of death are included in our excess deaths estimate. The vast majority of excess mortality is not attributable to social distancing or the pandemic-induced recession because recessions are generally associated with lower mortality [37].

We can evaluate whether emergency department visits for asthma in New York City decreased during the COVID-19 emergency, but we can't evaluate whether the decrease in asthma ED visits increased mortality because chronic lower respiratory mortality data from this period are not currently available. We also cannot evaluate for other syndromes or for other jurisdictions because of limitations in the National Syndromic Surveillance Program data and its public availability.

5.2 Implications for policy and practice

Research can identify populations at risk for mortality that may not be included in the official COVID-19 counts, and the overlap with groups affected by other health disparities, so resources can be allocated to the most affected communities. Excess mortality may be reduced by prioritizing research including random sample testing of people without symptoms and postmortem tests [38]. To avoid ascertainment bias, random sample postmortem testing can identify atypical disease presentations that may be under-recognized in clinical settings. In the face of testing limitations, postmortem tests may be viewed as expendable, but postmortem testing can improve patient care by identifying gaps in diagnosis and treatment.

To identify under-use of health care for cardiovascular and cerebrovascular conditions, National Surveillance System Program data should include these indications and current data should be available for all NSSP locations; New York City's system suggests that this change is feasible. Future federal pandemic planning should include upgrades to state vital statistics infrastructure, so that all states can report deaths in a timely fashion. As suggested after earlier pandemics [39], pandemic planning should identify how to release more detailed data so that research can discriminate between mechanisms of excess mortality, especially for high-vulnerability jurisdictions. Research using detailed data will find some spurious findings, including cases who died with the virus but not of the virus, but more detailed data will save lives by identifying vulnerable communities to allocate resources to, as well as atypical symptoms of the disease.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

5.3 Conclusions

Excess all-cause mortality exceeding the number of reported COVID-19 deaths is evident in many high-COVID-19 states. Greater test availability, including postmortem tests, can yield more accurate mortality counts and case-fatality ratios, and increase the public's willingness to adhere to non-pharmaceutical interventions to reduce transmission.

6 References

[1] Andr´e Ricardo Ribas Freitas, Maria Rita Donalisio, and Pedro Mar´ıa Alarc´on-Elbal. Excess mortality and causes associated with chikungunya, puerto rico, 2014–2015. *Emerging infectious diseases*, 24(12):2352, 2018.

[2] Roberto Rivera and Wolfgang Rolke. Modeling excess deaths after a natural disaster with application to hurricane maria. *Statistics in medicine*, 38(23):4545–4554, 2019.

[3] Steven G. Krantz and Arni S. R. Srinivasa Rao. Level of underreporting including underdiagnosis before the first peak of COVID-19 in various countries: Preliminary retrospective results based on wavelets and deterministic modeling. *Infection Control and Hospital Epidemiology*, pages 1–3, April 2020.

[4] Wang Ma, Xiang Huo, and Minghao Zhou. The healthcare seeking rate of individuals with influenza like illness: A metaanalysis. *Infectious Diseases*, 50(10):728–735, October 2018. eprint: https://doi.org/10.1080/23744235.2018.1472805.

[5] Lone Simonsen, Peter Spreeuwenberg, Roger Lustig, Robert J Taylor, Douglas M Fleming, Madelon Kroneman, Maria D Van Kerkhove, Anthony W Mounts, W John Paget, et al. Global mortality estimates for the 2009 influenza pandemic from the glamor project: a modeling study. *PLoS medicine*, 10(11), 2013.

[6] Alyssa S Parpia, Martial L Ndeffo-Mbah, Natasha S Wenzel, and Alison P Galvani. Effects of response to 2014–2015 ebola outbreak on deaths from malaria, hiv/aids, and tuberculosis, west africa. *Emerging infectious diseases*, 22(3):433, 2016.

[7] Sunil M. Shah, Iain M. Carey, Tess Harris, Stephen Dewilde, Christina R. Victor, and Derek G. Cook. The effect of unexpected bereavement on mortality in older couples. *American Journal of Public Health*, 103(6):1140– 1145, June 2013.

[8] Nathan R. Hoot and Dominik Aronsky. Systematic Review of Emergency Department Crowding: Causes, Effects, and Solutions. *Annals of Emergency Medicine*, 52(2):126–136.e1, August 2008.

[9] Safiya Richardson, Jamie S. Hirsch, Mangala Narasimhan, James M. Crawford, Thomas McGinn, Karina W. Davidson, Douglas P. Barnaby, Lance B. Becker, John D. Chelico, Stuart L. Cohen, Jennifer Cookingham, Kevin Coppa, Michael A. Diefenbach, Andrew J. Dominello, Joan Duer-Hefele, Louise Falzon, Jordan Gitlin, Negin Hajizadeh, Tiffany G. Harvin, David A. Hirschwerk, Eun Ji Kim, Zachary M. Kozel, Lyndonna M. Marrast, Jazmin N. Mogavero, Gabrielle A. Osorio, Michael Qiu, and Theodoros P. Zanos. Presenting

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *JAMA*, April 2020.

[10] Steffie Woolhandler and David U. Himmelstein. Intersecting U.S. Epidemics: COVID-19 and Lack of Health Insurance. *Annals of Internal Medicine*, April 2020.

[11] MR Spencer and F Ahmad. Timeliness of death certificate data for mortality surveillance and provisional estimates. vital statistics rapid release special report no. 001. us department of health and human services, centers for disease control and prevention. *National Center for Health Statistics, National Vital Statistics System. https://www. cdc. gov/nchs/data/vsrr/report001. pdf*, 2016.

[12] US Census Bureau. National Population by Characteristics: 2010-2019. https://www.census.gov/data/tables/time-series/demo/popest/2010snational-detail.html.

[13] Alexis Madrigal, Jeffrey Hammerbacher, Erin Kissane, and COVID Tracker team. COVID Tracking Project. https://covidtracking.com/data.

[14] New York Times. Coronavirus in the U.S.: Latest Map and Case Count. https://www.nytimes.com/interactive/2020/us/coronavirus-uscases.html, April 2020.

[15] CDC. Daily Updates of Totals by Week and State: Provisional Death Counts for Coronavirus Disease 2019 (COVID-19). https://www.cdc.gov/nchs/nvss/vsrr/COVID19/index.htm, 2020.

[16] Carlos Santos-Burgoa, John Sandberg, Erick Su´arez, Ann Goldman-Hawes, Scott Zeger, Alejandra Garcia-Meza, Cynthia M P´erez, Noel EstradaMerly, Uriyoan Col´on-Ramos, Cruz Mar´ıa Nazario, et al. Differential and persistent risk of excess mortality from hurricane maria in puerto rico: a time-series analysis. *The Lancet Planetary Health*, 2(11):e478–e488, 2018.

[17] Jay M Ver Hoef and Peter L Boveng. Quasi-poisson vs. negative binomial regression: how should we model overdispersed count data? *Ecology*, 88(11):2766–2772, 2007.

[18] Rosa Gabriel Sandoval Goldensohn, Ann Choi. Garbage Pickups Tell a Tale of Two Cities, With Part of Manhattan Shrinking. http://thecity.nyc/2020/04/garbage-pickups-tell-tale-of-two-cities-asmanhattan-shrinks.html, April 2020.

[19] S. N. Wood. *Generalized additive models: an introduction with R*. Chapman and Hall/CRC, 2017.

[20] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.

[21] National Vital Statistics System. Guidance for Certifying Deaths Due to Coronavirus Disease 2019 (COVID–19). *National Vital Statistics System: Vital Statistics Reporting Guidance*, (Report 3), April 2020.

[22] CDC. Coronavirus disease 2019 (COVID-19): Collection and submission of postmortem specimens from deceased persons with known or suspected COVID-19, march 2020

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

(interim guidance). https://www.cdc.gov/coronavirus/2019-ncov/hcp/guidance-postmortemspecimens.html, 2020.

[23] New York City Department of Health and Mental Hygiene. Influenza Surveillance Report Week ending April 18, 2020 (Week 16).

[24] CDC. COVIDView: A weekly surveillance summary of U.S. COVID-19 activity. https://www.cdc.gov/coronavirus/2019-ncov/coviddata/covidview/index.html, May 2020.

[25] M Boukhris, A Hillani, F Moroni, and et al. Cardiovascular implications of the covid-19 pandemic: A global perspective. *Can J Cardiol.*, S0828282X(20):30464–5, 2020.

[26] CDC. Leading Causes of Death: Data are for the U.S. https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm, May 2020.

[27] Matthew Biggerstaff, Michael A. Jhung, Carrie Reed, Alicia M. Fry, Lina Balluz, and Lyn Finelli. Influenza-like illness, the time to seek healthcare, and influenza antiviral receipt during the 2010-2011 influenza season-United States. *The Journal of Infectious Diseases*, 210(4):535–544, August 2014.

[28] Mark B. Parshall, Richard M. Schwartzstein, Lewis Adams, Robert B. Banzett, Harold L. Manning, Jean Bourbeau, Peter M. Calverley, Audrey G. Gift, Andrew Harver, Suzanne C. Lareau, Donald A. Mahler, Paula M. Meek, and Denis E. O'Donnell. An Official American Thoracic Society Statement: Update on the Mechanisms, Assessment, and Management of Dyspnea. *American Journal of Respiratory and Critical Care Medicine*, 185(4):435–452, February 2012.

[29] Ruoqing Li, Jigang Tian, Fang Yang, Lei Lv, Jie Yu, Guangyan Sun, Yu Ma, Xiaojuan Yang, and Jianqiang Ding. Clinical characteristics of 225 patients with COVID-19 in a tertiary Hospital near Wuhan, China. *Journal of Clinical Virology: The Official Publication of the Pan American Society for Clinical Virology*, 127:104363, April 2020.

[30] Long-Quan Li, Tian Huang, Yong-Qing Wang, Zheng-Ping Wang, Yuan Liang, Tao-Bi Huang, Hui-Yun Zhang, Weiming Sun, and Yuping Wang. COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis. *Journal of Medical Virology*, March 2020.

[31] Francois-Xavier Lescure, Lila Bouadma, Duc Nguyen, Marion Parisey, Paul-Henri Wicky, Sylvie Behillil, Alexandre Gaymard, Maude Bouscambert-Duchamp, Flora Donati, Quentin [Le Hingrat, Vincent Enouf, Nadhira Houhou-Fidouh, Martine Valette, Alexandra Mailles, Jean-Christophe Lucet, France Mentre, Xavier Duval, Diane Descamps, Denis Malvy, Jean-Fran¸cois Timsit, Bruno Lina, Sylvie van-der-Werf, and Yazdan Yazdanpanah. Clinical and virological data of the first cases of COVID-19 in Europe: A case series. *The Lancet Infectious Diseases*, 2020.

[32] Ling Lin, Lianfeng Lu, Wei Cao, and Taisheng Li. Hypothesis for potential pathogenesis of SARS-CoV-2 infection-a review of immune changes in patients with viral pneumonia. *Emerging Microbes & Infections*, 9(1):727– 732, December 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

[33] Tao Guo, Yongzhen Fan, Ming Chen, Xiaoyan Wu, Lin Zhang, Tao He, Hairong Wang, Jing Wan, Xinghuan Wang, and Zhibing Lu. Cardiovascular Implications of Fatal Outcomes of Patients With Coronavirus Disease 2019 (COVID-19). *JAMA cardiology*, March 2020.

[34] Yichun Cheng, Ran Luo, Kun Wang, Meng Zhang, Zhixiang Wang, Lei Dong, Junhua Li, Ying Yao, Shuwang Ge, and Gang Xu. Kidney disease is associated with in-hospital death of patients with COVID-19. *Kidney International*, 97(5):829–838, May 2020.

[35] Evangelos Terpos, Ioannis Ntanasis-Stathopoulos, Ismail Elalamy, Efstathios Kastritis, Theodoros N. Sergentanis, Marianna Politou, Theodora Psaltopoulou, Grigoris Gerotziafas, and Meletios A. Dimopoulos. Hematological findings and complications of COVID-19. *American Journal of Hematology*, April 2020.

[36] Theodore Lytras, Katerina Pantavou, Elisavet Mouratidou, and Sotirios Tsiodras. Mortality attributable to seasonal influenza in Greece, 2013 to 2017: Variation by type/subtype and age, and a possible harvesting effect. *Euro Surveillance: Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin*, 24(14), April 2019.

[37] Claire Margerison-Zilko, Sidra Goldman-Mellor, April Falconi, and Janelle Downing. Health Impacts of the Great Recession: A Critical Review. *Current epidemiology reports*, 3(1):81–91, March 2016.

[38] Marc Lipsitch, David L. Swerdlow, and Lyn Finelli. Defining the Epidemiology of Covid-19 — Studies Needed. *New England Journal of Medicine*, 382(13):1194–1196, March 2020. eprint: https://doi.org/10.1056/NEJMp2002125.

[39] Diana M. Prieto, Tapas K. Das, Alex A. Savachkin, Andres Uribe, Ricardo Izurieta, and Sharad Malavade. A systematic review to identify areas of enhancements of pandemic simulation models for operational use at provincial and local levels. *BMC public health*, 12:251, March 2012.

7 Acknowledgements

We thank Farida Ahmen, Michael Augenbraun, Stephan Kohlhoff, Josh Marshall, and Tracey Wilson for helpful conversations, and our spouses/partners for childcare.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

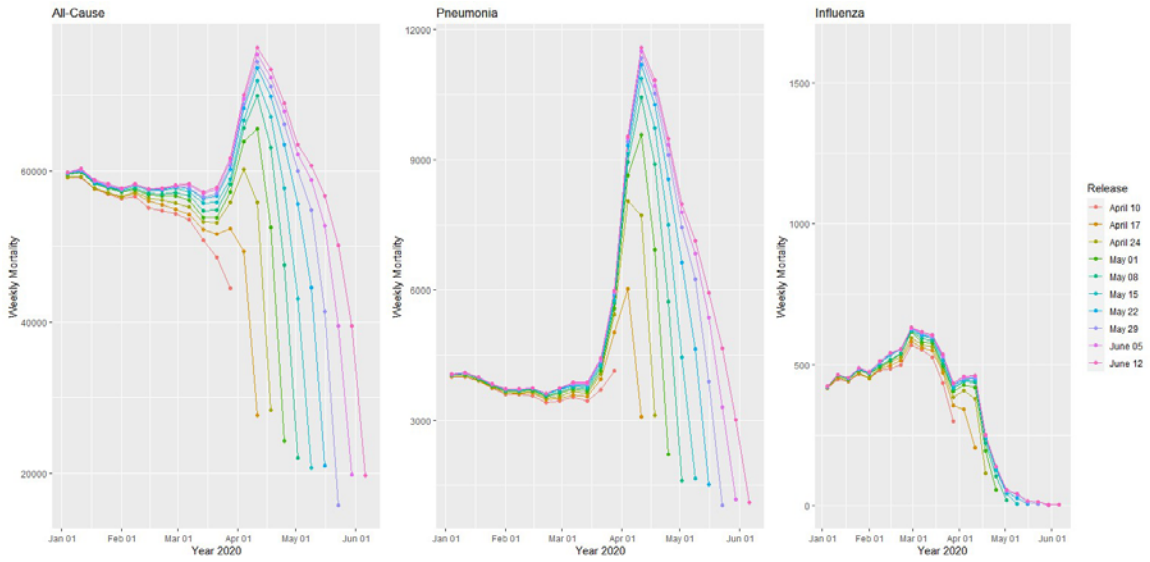

Figure 1: Provisional weekly mortality all-cause, pneumonia and influenza counts for the United States from weekly data releases April 10–June 12, 2020.

		COVID-19	All-cause Excess Deaths		Pneumonia Excess Deaths		Influenza Excess Deaths	
Jurisdiction	Start Date	Deaths	Semiparametric	Conventional	Semiparametric	Conventional	Semiparametric	Conventional
California	3/21/2020	2849	(2891, 5873)	(2458, 8086)	(1606, 2228)	(1661, 2695)	(-75, 52)	(89, 356)
Colorado	3/21/2020	1130	(1106, 1657)	(868, 1787)	(586, 768)	(512, 766)	(-26, 15)	(-62, 34)
Florida	3/21/2020	1840	(1150, 2731)	(210, 3745)	(1075, 1411)	(1344, 1925)	(-40, 19)	(-206, 5)
Illinois	3/21/2020	3525	(4412, 5871)	(4717, 6994)	(1888, 2324)	(1983, 2459)	(-20, 43)	(0, 127)
Massachusetts	3/28/2020	5050	(5061, 6317)	(5173, 6801)	(1819, 2196)	(1896, 2319)	(-28, 13)	(-113, 32)
Michigan	3/21/2020	5036	(4728, 6320)	(4494, 6457)	(2176, 2694)	(2214, 2672)	(-51, 28)	(-47, 129)
New Jersey	3/21/2020	10465	(12497, 15307)	(13026, 14749)	(5255, 6226)	(5572, 6052)	(0, 43)	(-50, 57)
New York	3/21/2020	26584	(30469, 37722)	(33363, 36599)	(11239, 13463)	(12033, 12778)	(674, 889)	(631, 912)
Washington	2/29/2020	925	(486, 1555)	(124, 1575)	(353, 614)	(333, 647)	(-3, 72)	(-26, 76)
United States	3/21/2020	78834	(80862, 107284)	(80666, 116859)	(36750, 43777)	(37782, 43934)	(-27, 1331)	(-126, 2784)

Table 1: Excess all-cause, pneumonia, and influenza mortality 95% confidence intervals from 2 models, from states with the largest reported COVID-19 mortality as of 5/9/2020, and with official COVID-19 toll. All data used included mortality until 5/9/2020 included in the 6/12/2020 data release.

The semiparametric method uses a starting date, but the conventional method does not.

The table presents prediction intervals using the conventional method, although we refer to confidence intervals throughout for consistency.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

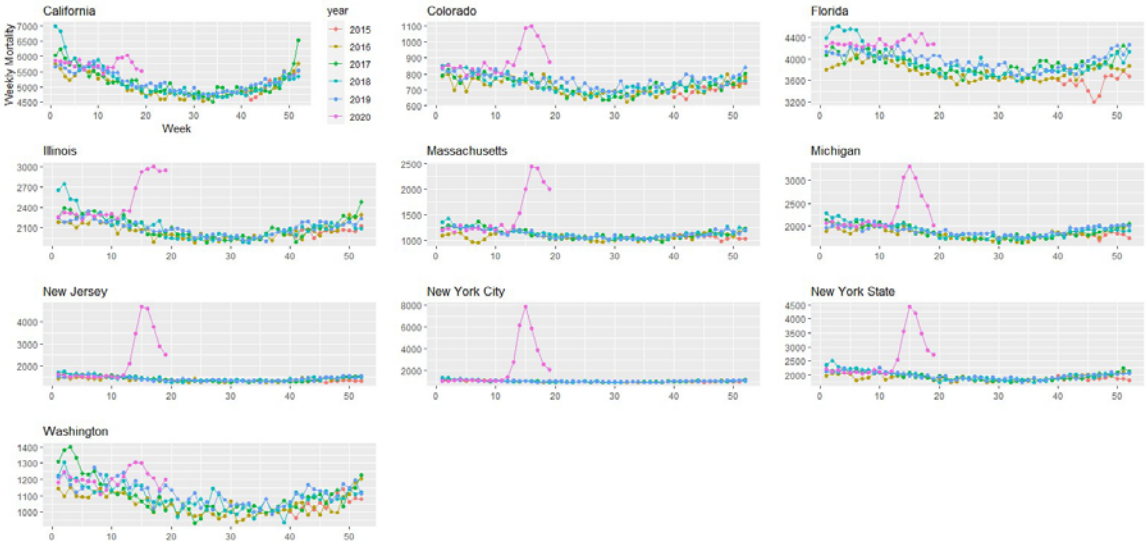

Figure 2: Weekly all-cause mortality grouped by year and state starting on week 40 of 2015 until 5/9/2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

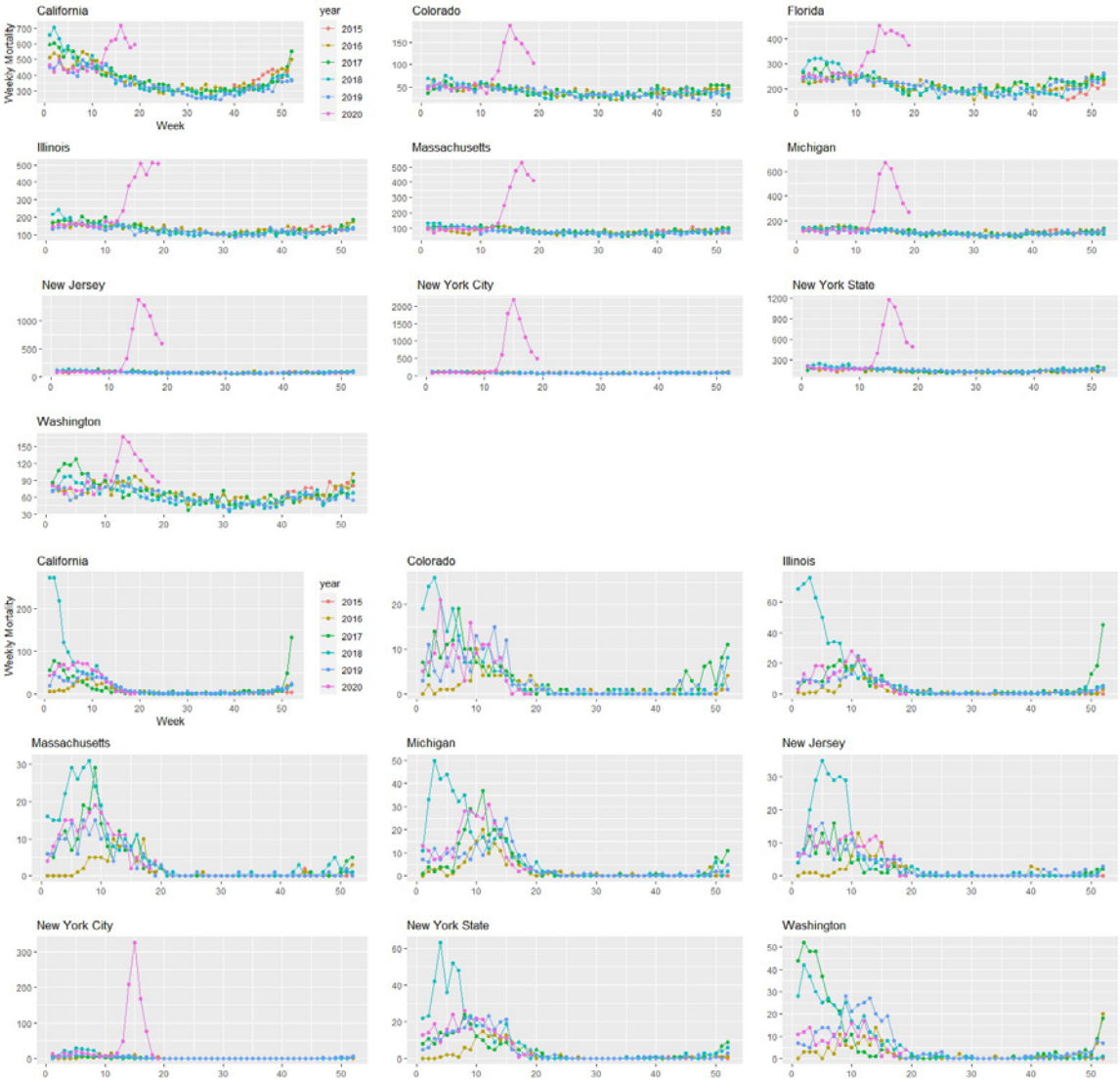

Figure 3: Weekly (1) pneumonia and (2) influenza mortality grouped by year and state starting on week 40 of 2015 until 5/9/2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

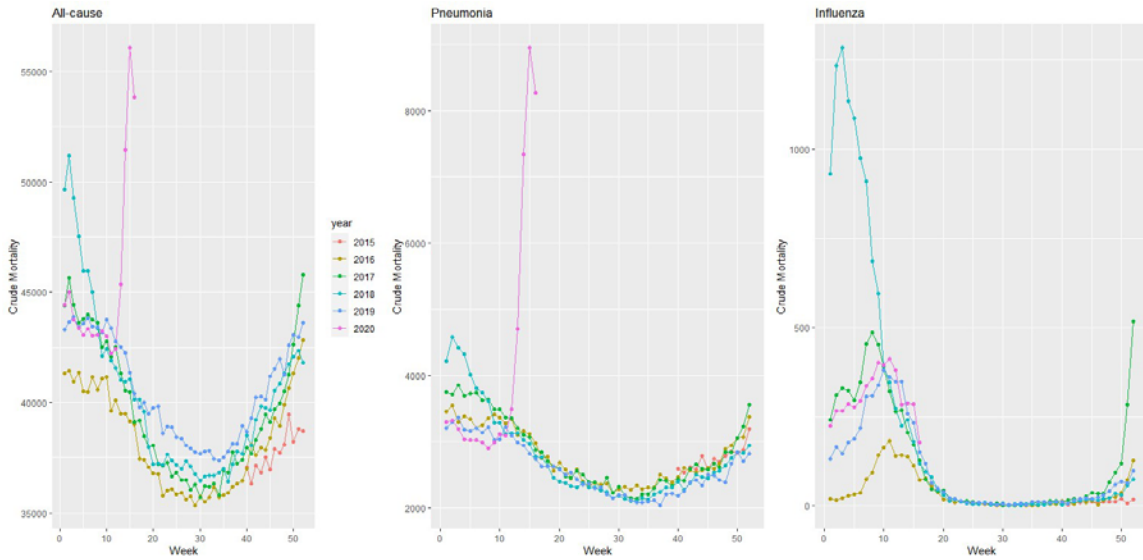

Figure 4: Weekly mortality in the United States by year starting on week 40 of 2015 until 5/9/2020.

9 Appendix: Detailed methods

9.1 Semiparametric excess mortality model

Covariates include week of the year, year, an indicator function that presents the possibility that the mean death rate has increased after the start of the emergency in a location, and the natural logarithm of population as an offset. Week of the year is modeled non-parametrically to capture seasonal mortality variations. Let D_t = number of certified deaths at time index t, N_t = population size. We use p_t as an indicator of time period t falls in. The indicator variable permit us to estimate excess deaths. Specifically, $p_t = 0$ represents the pre-emergency period; $p_t = 1$, the period after the emergency. Moreover, let $week_t$ = week of year, and $year_t$ = a categorical year effect.

Assuming D_t follows a Poisson distribution, we fit a semiparametric model;

$$\log(\mu_t) = \log(N_t) + \beta_o + \beta_1 p_t + f(week_t) + year_t \qquad (1)$$

where $\mu_t = E(D_t|t,p_t,N_t,week_t,year_t)$. The natural logarithm of N_t is an offset variable; while f is a smooth function of *week*, which accounts for within year variation. f is fit using a penalized cyclic cubic regression spline [19].

Preliminary analysis indicates overdispersion, so we used a quasi-Poisson model to estimate the dispersion parameter [17]. We estimated coefficients by a penalized likelihood maximization approach, where the smoothing penalty parameters are determined by restricted maximum likelihood. The residuals of the fitted model (1) did not present remaining temporal dependence.

The fit of model (1) can be used to estimate excess deaths through the difference between the estimated model with $p_t = 1$, versus the estimated model with $p_t = 0$. Let $\widehat{\mu_t} = E(D_t|t,p_t =$

medRxiv preprint doi: https://doi.org/10.1101/2020.05.04.20090324; this version posted June 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

1,N_t,$week_t$,$year_t$), $\widehat{\psi_t}$ = $E(D_t|t,p_t = 0,N_t,week_t,year_t)$, and $\widehat{b_o}$, $\widehat{b_1}$ estimate β_o,β_1 respectively. Then [2],

$$\hat{\mu}_t - \hat{\psi}_t = e^{\hat{b}_o + \widehat{f}(week_t) + year_t}\left(e^{\log(N_t) + \hat{b}_1} - 1\right) \quad (2)$$

When $p_t = 0$, then $\widehat{\mu_t} - \widehat{\psi_t} = 0$. Equation (2) is the maximum likelihood estimator for expected excess deaths at t.

To estimate cumulative excess deaths we use (2),

$$\sum_{t=q}^{r}(\widehat{\mu_t} - \widehat{\psi_t}) \quad (3)$$

for any time period starting at index q and ending at r. Approximate simulations from the Bayesian posterior density are performed to obtain 95% credibility intervals, which we will refer to as confidence intervals throughout.

9.2 Conventional mortality method

Conventional excess mortality methods build a temporal mortality model until time $m < T$, then use this model to predict deaths from time $m + 1$ to time T, and excess deaths will be the difference between the deaths between times $m + 1$ and T, and the predicted deaths from the model. Uncertainty is usually quantified through the uncertainty of the parameters on the regression model [16], which ignores natural random fluctuations in mortality from time $m + 1$ to time T. We tackle this problem combining the Bayesian posterior density described in section 3.3 with uncertainty on weekly mortality.

First, we fit a quasi-Poisson semiparametric model similar to the one presented in the previous section, but here we only use data until February 1, 2020. The deaths from February 8, 2020 forward should follow a Poisson distribution with some expected rate; the maximum likelihood estimator of such rate is the mean weekly deaths during this period. This weekly deaths will be approximately normal with a variance that accounts for overdispersion according to the scale parameter in the fitted model. From this distribution, we simulate 10,000 weekly deaths and subtract the results from the posterior distributions of the fitted results. The 95% confidence interval is the range between the 2.5% and 97.5% percentiles of all excess deaths.