Estimating Excess Deaths in the United States early in the COVID-19 Pandemic

Roberto Rivera, Ph.D. (1)*

Janet E. Rosenbaum, Ph.D (2)

Walter Quispe, Ph.D. (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.

Abstract

Importance: Deaths are frequently under-estimated during emergencies, such as hurricanes and pandemics. Accurate mortality data may improve public compliance with non-pharmaceutical interventions.

Objective: This study estimates excess all-cause, pneumonia, and influenza mortality between March 2020 and April 11, 2020 in 9 states.

Design: In this ecological study, we estimated excess mortality using two methods --- a semiparametric method and a conventional time-series method --- in weekly mortality data from 9 states from the Mortality Surveillance Survey covering September 27, 2015 to April 11, 2020.

Setting: The United States.

Participants: We estimated excess mortality in states with high reported COVID-19 deaths and apparently complete mortality data: California, Colorado, Florida, Illinois, Louisiana, Massachusetts, Michigan, New Jersey, New York, Pennsylvania, and Washington.

Exposure: The exposure is the COVID-19 health emergency.

Main outcome: Excess mortality is the difference between actual and expected all-cause, pneumonia, and influenza mortality. Deaths not explained by COVID-19 are the difference between the 95% credibility interval (CI) for excess mortality, and the number of reported COVID-19 deaths. The hypotheses were formed prior to data collection.

Results: With the semiparametric method, we estimated 257-1024 more deaths in Illinois than COVID-19 case numbers (excess mortality 95% CI (939, 1706) vs. 682 COVID-19 deaths); 218-741 more deaths in Massachusetts (95% CI (904, 1427) vs. 686 COVID-19 deaths); 313-1100 more deaths in Michigan (95% CI (1704, 2491), vs. 1391 COVID-19 deaths); 2348-3661 more

deaths in New Jersey (95% CI (4528, 5844), vs. 2183 COVID-19 deaths); and 5875-9738 more deaths in New York State (168-213% more) (95% CI (14502, 18365) vs. 8627 COVID-19 deaths).

Conclusions: Official COVID-19 mortality figures substantially understate actual mortality and also suggest substantially greater case fatality rates. Improving mortality data dissemination may improve public adherence to non-pharmaceutical interventions and reduce mortality in future pandemics.

Estimating Excess Deaths in the United States Early in the COVID-19 Pandemic

Roberto Rivera, Janet E. Rosenbaum, Walter Quispe

May 3, 2020

Introduction

The number of COVID-19 deaths is under-reported in many countries. [1, 2, 3] An early study found excess pneumonia and influenza deaths in New York, New Jersey, and New York City [4], and a modeling study suggested excess mortality due to health care worker absenteeism caused by school closures. [5] The Centers for Disease Control and Prevention issues guidelines to determine cause of deaths, but underestimating the death toll of natural disasters [6], heatwaves [7, 8], influenza during pandemic years such as 1918 and 2009 [9, 10, 11, 12] and non-pandemic years [13, 14, 15], 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.[16] Another example includes gross under-estimation of Hurricane Maria's death toll.[17, 18, 19, 20] Deaths due to the emergency may be underestimated due to under-diagnosis [21, 22], insufficient postmortem SARS-CoV-2 tests, not seeking health care [23], and incomplete disease nosology and consequent confirmation bias. Indirect deaths due to natural disasters or health emergencies are also common, due to an overloaded health system [24, 25] 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.[26] The emergency may also lead to indirect deaths from economic, social, and emotional stress [27, 28] and crowded emergency departments.[29]

In this study, we estimate excess all-cause, pneumonia, and influenza mortality during the COVID-19 pandemic; which estimates deaths both direct and indirectly related to COVID-19. We specified these hypotheses prior to analysis. Directly related COVID-19 deaths include sudden deaths in patients with undetected SARS-CoV-2, due to respiratory failure [30], organ failure [31, 32], thromboembolism [33], or other reasons. Patients may have undetected SARSCoV-2 infections due to false negative tests [34], not seeking health care [35, 23], or patient elopement due to emergency department crowding.[29] Testing and forensic staff shortfalls lead to lack of postmortem testing. Indirect deaths may include deaths due to emergency department crowding [29]; 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.[36]

Methods

Data and measures

Weekly all-cause, pneumonia, and influenza mortality for each state from September 27, 2015 (week 40) to April 11, 2020 (week 14) were obtained from the National Center for Health Statistics (NCHS) Mortality Surveillance Survey data release on May 1, 2020. The timeliness in death certificate reporting varies by region, state, and cause of death.[37] 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. [38]

We obtained COVID-19 mortality as of 4/11/2020 from the New York Times COVID-19 data repository, based on reports from state and local health agencies.[39]

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.

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 [20]. See appendix for details.

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 [17].

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 [40-42]. Week of the year is modeled non-parametrically with a penalized cyclic cubic regression spline function to capture seasonal mortality variations [41].

Preliminary analysis indicates overdispersion, so we used a quasi-Poisson model to estimate the dispersion parameter [43]. We estimated coefficients by a penalized likelihood maximization approach, where the smoothing penalty parameters are determined by restricted maximum likelihood [44, 45]. The residuals of the fitted model (1) did not present remaining temporal dependence, so we did not require a Generalized Additive Mixed Model [46, 47].

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 April 11, 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 [17], 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 (3, 7, 8, and 11) 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.[49] If the population of these community districts decreased by about 10%, the total population of New York City decreased by less than 1%[50].

The models were fit using version 1.8-28 of the *mgcv* package in R 3.6.3 between April 12, 2020 and May 3, 2020 [42, 51]. All code and data will be made publicly available.

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 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.

Unexplained excess deaths are excess deaths not attributed to COVID-19, pneumonia, or influenza. We can obtain a lower bound on unexplained excess deaths using the semiparametric model results as the difference between all-cause excess mortality and three quantities: excess pneumonia, excess influenza, and COVID-19 deaths. These categories are not mutually exclusive and they have a great deal of overlap, so this estimation gives only a lower bound for unexplained excess deaths. Assuming the all-cause excess deaths and pneumonia deaths are independent estimates, we can use the above Bayesian approximation to estimate 95% confidence intervals of unexplained excess deaths for each state.

Results

Death counts increase with each data release, especially for the penultimate week (Table 1). Provisional death counts for recent weeks can drastically underrepresent death certificate counts.

Colorado, Illinois, Massachusetts, Michigan, New Jersey, New York City, and New York State show marked increases in all-cause weekly mortality (Figure 1). All states have excess pneumonia mortality for all states, but only New York City has excess influenza deaths (Figure 2)

The variation on reporting timeliness by state makes at this moment excess deaths for the United States assessment impossible (Table 1). COVID-19 mortality is greater for older patients, which can exacerbate the dilution of excess deaths for all of the United States. Among people age 65+ aggregated across the US, we observe greater all-cause and pneumonia deaths but only for very recent weeks (Figure 3).

Table 2 presents 95% confidence intervals for all-cause, pneumonia, and influenza deaths when using the semiparametric method. We estimated 257–1024 more deaths in Illinois than COVID-19 official counts (excess mortality 95% CI (939, 1706) vs. 682 COVID-19 deaths); 218–741 more deaths in Massachusetts (95% CI (904, 1427) vs. 686 COVID-19 deaths); 313–1100 more deaths in Michigan (95% CI (1704, 2491), vs. 1391 COVID-19 deaths); 2348–3661 more deaths in New Jersey (95% CI (4528, 5844), vs. 2183 COVID-19 deaths); and 5875–9738 more deaths in New York State (95% CI (14502, 18365) vs. 8627 COVID-19 deaths). California, Colorado, and Washington present significant all-cause excess deaths but there was no evidence of them exceeding official COVID-19 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 and New Jersey (2). To evaluate whether the apparent misclassification of COVID-19 deaths as influenza in New York City is attributable to urbanicity, we evaluated influenza mortality in the District of Columbia, the other urban-only area, but influenza deaths were not higher than usual (not shown).

With the conventional model, we estimate more excess all-cause deaths in Illinois, Massachusetts, New Jersey, and New York than reported COVID-19 cases (2); 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 April 4, 2020 (not shown).

Unexplained excess deaths

We obtain a 95% confidence interval for the lower bound of unexplained excess deaths: between 185 and 1,623 unexplained excess deaths occurred in New Jersey until April 11. There was no evidence of unexplained excess deaths in other states studied, but this is likely due to substantial overlap in causes of death and our analysis covering the early stages of the pandemic in the U.S. As the emergency continued for a considerable time, we suspect excess deaths not explained by official COVID-19 and pneumonia excess deaths will become meaningful with additional data. These unexplained excess deaths may include the acute thromboses that are a forme fruste of COVID-19, among patients with no COVID-19 tests.

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. We find substantial excess all-cause mortality that exceeds the number of documented COVID-19 deaths in New York, New Jersey, Michigan, Massachusetts, and Illinois. The CDC has established guidelines for certifying COVID-19 deaths [52] and whether to collect postmortem specimens for SARS-CoV-2 testing [53]. However, 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.

We quantify higher pneumonia mortality than expected in all 9 states and higher influenza deaths primarily in New York. However, emergencies such as pandemics also result in considerable indirect fatalities, and our results indicate this had already started occurring in New Jersey.

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 [39]. COVID-19 test access, as measured by decreased percent of COVID-19 tests that are positive, has increased with time in the US [54], but substantial heterogeneity across states [55] 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 may become apparent in subsequent data releases.

Some excess mortality may be due to patients not seeking care or being 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 influenza-like illness (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 [35], and during a flu pandemic care-seeking increases by only about 10 percentage points [23]. Illness severity and/or dyspnea/breathlessness predicts health care seeking for ILI [56, 57], but patients infected with SARS CoV-2 may not have these symptoms. Most severe or fatal COVID-19 cases do not present with dyspnea [58, 59], and lung damage can be substantial even without dyspnea [60].

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 [30, 32, 33], 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]. A wide variety of clinical conditions with a false negative test or no test may in fact be attributable to COVID-19: cardiac injury [61], kidney or liver injury [62], or acute thromboses caused by hypercoagulability [33].

Excess mortality during heatwaves and influenza pandemics often shows mortality displacement (or harvesting effect), where subsequent mortality subsequently declines [63, 64]. 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.

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 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 due to 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. However, social distancing and stay-at-home orders are not expected to reduce overall mortality due to less driving and lower economic activity [65, 66], so the vast majority of excess mortality is not attributable to social distancing or the recession itself.

Implications for policy and practice

Excess mortality may be reduced by future pandemic planning that prioritizes research during the pandemic, including allocation of tests for research, random sample testing of people without symptoms, and postmortem tests [67]. 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 [68], 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. Bereaved families often agree to research participation to prevent future morbidity and mortality [69, 70], so bioethicists may conclude that releasing more detailed mortality data would be considered acceptable in pandemic emergencies in order to save lives [71]. Research using detailed data will increase type I error at the cost of reducing type II errors (i.e., excess mortality), but reducing type II error may save lives by identifying risks that physicians and public health should address.

Conclusions

Excess all-cause mortality exceeding the number of reported COVID-19 deaths is evident in many high-COVID-19 states; in New York State, deaths associated with COVID-19 appear to be twice as high as reported. 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 nonpharmaceutical interventions to reduce transmission.

References

[1] Modi C, Boehm V, Ferraro S, Stein G, Seljak U. Total COVID-19 Mortality in Italy: Excess Mortality and Age Dependence through Time-Series Analysis. *medRxiv*, 2020;page 2020.04.15.20067074.

[2] Docherty K, Butt J, de Boer R, Dewan P, Koeber L, Maggioni A, McMurray J, Solomon S, Jhund PS. Deaths from Covid-19: Who are the forgotten victims? *medRxiv*. 2020; 2020.04.21.20073114.

[3] Vandoros S. Has mortality due to other causes increased during the Covid-19 pandemic? Early evidence from England and Wales. *medRxiv*. 2020; 2020.04.14.20065706.

[4] Weinberger D, Cohen T, Crawford F, Mostashari F, Olson D, Reich NG, Pitzer VE, Russi M, Simonsen L, Watkins A, Viboud C. Estimating the early death toll of COVID-19 in the United States. medRxiv. 2020.

[5] Bayham J, Fenichel EP. Impact of school closures for COVID-19 on the US health-care workforce and net mortality: A modelling study. *The Lancet. Public Health* 2020.

[6] Hammer CC. Understanding excess mortality from not-so-natural disasters. *The Lancet. Planetary Health*. 2018;2(11):e471–e472.

[7] Guo Y, Gasparrini A, Armstrong BG, et al. Heat Wave and Mortality: A Multicountry, Multicommunity Study. *Environmental Health Perspectives*. 2017;125(8):087006.

[8] Gasparrini A, Guo Y, Hashizume M, et al. Mortality risk attributable to high and low ambient temperature: A multicountry observational study. *Lancet. 2015;*386(9991):369–375.

[9] Cilek L, Chowell G, Fariñas DR. Age-Specific Excess Mortality Patterns During the 1918-1920 Influenza Pandemic in Madrid, Spain. *American Journal of Epidemiology*. 2018; 187(12):2511–2523.

[10] Bengtsson T, Dribe M, Eriksson B. Social Class and Excess Mortality in Sweden During the 1918 Influenza Pandemic. *American Journal of Epidemiology*. 2018; 187(12):2568–2576.

[11] Yang L, Chan KP, Cowling BJ, et al. Excess mortality associated with the 2009 pandemic of influenza A(H1N1) in Hong Kong. *Epidemiology and Infection*. 2012; 140(9):1542–1550.

[12] Yang L, Wong CM, Chiu SS, Cowling BJ, Peiris JS. Estimation of excess mortality and hospitalisation associated with the 2009 pandemic influenza. *Hong Kong Medical Journal = Xianggang Yi Xue Za Zhi*. 2018; 24 Suppl 6(5):19–22.

[13] Park M, Wu P, Goldstein E, Kim WJ, Cowling BJ. Influenza-Associated Excess Mortality in South Korea. *American Journal of Preventive Medicine*. 2016;50(4):e111–e119.

[14] Nielsen J, Vestergaard LS, Richter L, et al. European all-cause excess and influenza-attributable mortality in the 2017/18 season: Should the burden of influenza B be reconsidered? *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*. 2019; 25(10):1266–1276.

[15] Rosano A, Bella A, Gesualdo F, et al. Investigating the impact of influenza on excess mortality in all ages in Italy during recent seasons (2013/14-2016/17 seasons). *International journal of infectious diseases. 2019*; 88:127–134.

[16] Freitas ARR, Donalisio MR, Alarcón-Elbal PM. Excess mortality and causes associated with chikungunya, puerto rico, 2014–2015. *Emerging infectious diseases*. 2018; 24(12):2352.

[17] Rivera R, Rolke W. Modeling excess deaths after a natural disaster with application to hurricane Maria. *Statistics in medicine*. 2019;38(23):4545–4554.

[18] Rivera R, Rolke W. Estimating the death toll of hurricane Maria. *Significance*. 2018; 15(1):8–9.

[19] Kishore N, Marqúes D, Mahmud A, et al. Mortality in Puerto Rico after hurricane maria. *New England journal of medicine*. 2018; 379(2):162–170.

[20] Santos-Burgoa C, Sandberg J, Súarez E, et al. Differential and persistent risk of excess mortality from hurricane Maria in Puerto Rico: a time-series analysis. *The Lancet Planetary Health*. 2018; 2(11):e478–e488.

[21] Krantz SG, Rao ASRS. 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*. 2020; 1–3.

[22] Lau H, Khosrawipour V, Kocbach P, et al. Internationally lost COVID-19 cases. *Journal of Microbiology, Immunology, and Infection = Wei Mian Yu Gan Ran Za Zhi*. 2020 (March).

[23] Ma W, Huo X, Zhou M. The healthcare seeking rate of individuals with influenza like illness: A meta-analysis. *Infectious Diseases*. 2018; 50(10):728–735. DOI:10.1080/23744235.2018.1472805.

[24] Simonsen L, Spreeuwenberg P, Lustig R, et al. Global mortality estimates for the 2009 influenza pandemic from the glamor project: a modeling study. *PLoS medicine*. 2013; 10(11).

[25] Jamison DT, Gelband H, Horton S, et al. *Disease Control Priorities, (Volume 9): Improving Health and Reducing Poverty*. The World Bank, 2017.

[26] Parpia AS, Ndeffo-Mbah ML, Wenzel NS, Galvani AP. Effects of response to 2014–2015 Ebola outbreak on deaths from malaria, HIV/AIDS, and tuberculosis, West Africa. *Emerging infectious diseases*. 2016; 22(3):433.

[27] Vandoros S, Avendano M, Kawachi I. The association between economic uncertainty and suicide in the short-run. *Social Science & Medicine.* 2019; 220:403–410.

[28] Shah SM, Carey IM, Harris T, et al. The effect of unexpected bereavement on mortality in older couples. *American Journal of Public Health*. 2013; 103(6):1140– 1145.

[29] Hoot NR, Aronsky D. Systematic Review of Emergency Department Crowding: Causes, Effects, and Solutions. *Annals of Emergency Medicine*. 2008; 52(2):126–136.e1.

[30] Lescure FX, Bouadma L, Nguyen D, et al. Clinical and virological data of the first cases of COVID-19 in Europe: A case series. *The Lancet Infectious Diseases*, 2020.

[31] Libby P. The Heart in COVID19: Primary Target or Secondary Bystander? *JACC. Basic to translational science*, April 2020.

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

[33] Terpos E, Ntanasis-Stathopoulos I, Elalamy I, et al. Hematological findings and complications of COVID-19. *American Journal of Hematology*, April 2020.

[34] Richardson S, Hirsch JS, Narasimhan M, et al. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *JAMA*, April 2020.

[35] Biggerstaff M, Jhung MA, Reed C, Fry AM, Balluz L, Finelli L. 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*. 2014;210(4):535–544.

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

[37] Spencer MR, Ahmad F. 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.

[38] US Census Bureau. National Population by Characteristics: 2010-2019. https://www.census.gov/data/tables/time-series/demo/popest/2010snational-detail.html. Accessed April 30, 2020.

[39] 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. Accessed April 30, 2020.

[40] Dobson AJ, Barnett AG. *An introduction to generalized linear models*. CRC press, 2018.

[41] Ruppert D, Wand MP, Carroll RJ. *Semiparametric Regression*. Cambridge: Cambridge Univ. Press., 2003.

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

[43] Ver Hoef JM, Boveng PL. Quasi-poisson vs. negative binomial regression: how should we model overdispersed count data? *Ecology*. 2007;88(11):2766–2772.

[44] Wood SN. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2011;73(1):3– 36.

[45] Reiss PT, Ogden RT. Smoothing parameter selection for a class of semiparametric linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2009;71(2):505–523.

[46] Chien LC, Bangdiwala SI. The implementation of Bayesian structural additive regression models in multi-city time series air pollution and human health studies. *Stochastic environmental research and risk assessment*. 2012; 26(8):1041–1051.

[47] Zwack LM, Paciorek CJ, Spengler JD, Levy JI. Modeling spatial patterns of traffic-related air pollutants in complex urban terrain. *Environmental Health Perspectives*. 2011; 119(6):852.

[48] Casella G, Berger RL. *Statistical inference*, volume 2. Duxbury Pacific Grove, CA, 2002.

[49] Goldensohn RGS, Choi A. 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. Accessed

[50] Population Division, New York City Department of City Planning. New York City Population By Community Districts — NYC Open Data. https://data.cityofnewyork.us/City-Government/New-York-CityPopulation-By-Community-Districts/xi7c-iiu2, July 2020. Accessed April 29, 2020.

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

[52] 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. https://www.cdc.gov/nchs/data/nvss/vsrg/vsrg03-508.pdf Accessed May 3, 2020.

[53] Coronavirus Disease 2019 (COVID-19): Collection and submission of postmortem specimens from deceased persons with known or suspected COVID. https://www.cdc.gov/coronavirus/2019-ncov/hcp/guidance-postmortem-specimens.html Accessed May 3, 2020.

[54] Centers for Disease Control and Prevention. COVIDView: A weekly surveilance summary of U.S. COVID-19 activity. https://www.cdc.gov/coronavirus/2019-ncov/coviddata/covidview/index.html, May 2020. Accessed May 1, 2020.

[55] Madrigal A, Hammerbacher J, Kissane E, COVID Tracker team. COVID Tracking Project. https://covidtracking.com/data. Accessed May 3, 2020.

[56] Peppa M, Edmunds WJ, Funk S. Disease severity determines health-seeking behaviour amongst individuals with influenza-like illness in an internet-based cohort. *BMC infectious diseases*. 2017; 17(1):238.

[57] Parshall MP, Schwartzstein RM, Adams L, et al. An Official American Thoracic Society Statement: Update on the Mechanisms, Assessment, and Management of Dyspnea. *American Journal of Respiratory and Critical Care Medicine.* 2012; 185(4):435–452.

[58] Li R, Tian J, Yang F, et al. 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*. 2020; 127:1043-63.

[59] Recasens BB, Martinez-Llorens JM, Rodriguez-Sevilla JJ, Rubio MA. Lack of dyspnea in Covid-19 patients; another neurological conundrum? *European Journal of Neurology*. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ene.14265.

[60] Li LQ, Huang T, Wang YQ, et al. COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis. *Journal of Medical Virology*, March 2020.

[61] Guo T, Fan Y, Chen M, et al. Cardiovascular Implications of Fatal Outcomes of Patients With Coronavirus Disease 2019 (COVID-19). *JAMA cardiology*, March 2020.

[62] Cheng Y, Luo R, Wang K, et al. Kidney disease is associated with in-hospital death of patients with COVID-19. *Kidney International*. 2020; 97(5):829–838.

[63] Schwartz J. Is there harvesting in the association of airborne particles with daily deaths and hospital admissions? *Epidemiology. 2001;*12(1):55–61.

[64] Lytras T, Pantavou K, Mouratidou E, Tsiodras S. 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*. 2019;24(14).

[65] Ruhm CJ. Health Effects of Economic Crises. *Health Economics*. 2016; 25 Suppl 2:6–24.

[66] Margerison-Zilko C, Goldman-Mellor S, Falconi A, Downing J. Health Impacts of the Great Recession: A Critical Review. *Current epidemiology reports*. 2016; 3(1):81–91.

[67] Lipsitch M, Swerdlow DL, Finelli L. Defining the Epidemiology of Covid-19 — Studies Needed. *New England Journal of Medicine*. 2020;382(13):1194–1196. eprint: https://doi.org/10.1056/NEJMp2002125.

[68] Prieto DM, Das TK, Savachkin AA, Uribe A, Izurieta R, Malavade S. A systematic review to identify areas of enhancements of pandemic simulation models for operational use at provincial and local levels. *BMC public health*. 2012; 12:251.

[69] Germain A, Mayland CR, Jack BA. The potential therapeutic value for bereaved relatives participating in research: An exploratory study. *Palliative and Supportive Care*. 2016; 14(5):479–487.

[70] Kentish-Barnes N, McAdam JL, Kouki S, et al. Research Participation for Bereaved Family Members: Experience and Insights From a Qualitative Study. *Critical Care Medicine*, 2015;43(9):1839–1845.

[71] Yee A, Zubovic E, Yu J, et al. Ethical considerations in the use of Pernkopf's Atlas of Anatomy: A surgical case study. *Surgery*. 2019; 165(5):860–867.

Acknowledgements

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

Author contributions

RR conceived the study and devised the model; RR and JR obtained data; RR and WQ implemented the model; RR, JR, and WQ interpreted the results; JR conducted literature review and wrote the manuscript; all authors approved the final version.

	New York City											
Week	**April 10 Release**			**April 17 Release**			**April 24 Release**			**May 1 Release**		
	All	Pneu.	Infl.	All	Pneu.	Infl.	All	Pneu.	Infl.	All	Pneu.	Infl.
12	1,317	156	13	1,353	159	14	1,367	159	14	1,370	160	14
13	2,231	476	40	2,474	528	42	2,602	570	43	2,670	589	45
14	-	-	-	4,408	1,231	158	5,336	1,544	189	5,747	1,652	200
15	-	-	-	3,426	698	158	5,889	1,473	284	6,858	1,838	302
16	-	-	-	-	-	-	2,777	569	82	4,515	1,140	141
17	-	-	-	-	-	-	-	-	-	1,778	386	41
	United States											
	All	Pneu.	Infl.	All	Pneu.	Inf.	All	Pneu.	Inf.	Al	Pneu.	Inf.
12	48,488	3,697	435	51,602	3,947	474	53,130	4,071	485	53,811	4129	495
13	44,402	4,135	300	52,285	5,016	355	55,778	5,433	386	57,147	5580	406
14	-	-	-	49,292	6,017	342	60,195	8,045	402	63,809	8629	428
15	-	-	-	27,688	3,081	206	55,773	7,723	380	65,524	9563	419
16	-	-	-	-	-	-	28,374	3,121	115	52,521	6917	197
17	-	-	-	-	-	-	-	-	-	24,318	2,214	58

Table 1: New York City and United States provisional mortality counts for recent weeks according to the data released on April 10, April 17 April 24, and May 1, 2020.
All = All-cause mortality; Pneu. = pneumonia mortality; infl. = influenza mortality

		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	632	(185, 1998)	(-575, 3714)	(486, 865)	(453, 1231)	(-53, 59)	(100, 356)
Colorado	3/21/2020	274	(267, 579)	(11, 697)	(225, 340)	(147, 336)	(-11, 24)	(-51, 37)
Florida	3/21/2020	445	(-142, 945)	(-1014, 1683)	(461, 680)	(507, 950)	(-30, 27)	(-170, 19)
Illinois	3/21/2020	682	(939, 1706)	(896, 2633)	(460, 668)	(441, 791)	(-11, 43)	(1, 124)
Massachusetts	3/28/2020	686	(904, 1427)	(693, 1900)	(382, 545)	(366, 657)	(-15, 13)	(-94, 34)
Michigan	3/21/2020	1391	(1704, 2491)	(1105, 2566)	(763, 1020)	(689, 1013)	(-26, 40)	(-27, 130)
New Jersey	3/21/2020	2183	(4528, 5844)	(4318, 5622)	(1839, 2350)	(1931, 2256)	(3, 38)	(-45, 53)
New York	3/21/2020	8627	(14602, 18365)	(15329, 17772)	(5124, 6350)	(5404, 5947)	(449, 636)	(374, 620)
Washington	2/29/2020	498	(71, 819)	(-290, 833)	(200, 391)	(167, 411)	(-4, 65)	(-26, 69)

Table 2: Excess all-cause, pneumonia, and influenza mortality 95% confidence intervals from 2 models, from states with the largest reported COVID-19 mortality as of 4/11/2020, and with official COVID-19 toll. All data used included mortality until 4/11/2020 included in the 5/1/2020 data release.
The semiparametric method uses a starting date, but the conventional method does not.

Figure 1: Weekly all-cause mortality grouped by year and state starting on week 40 of 2015 until 4/11/2020.

Figure 2: Weekly (1) pneumonia and (2) influenza mortality grouped by year and state starting on week 40 of 2015 until 4/11/2020.

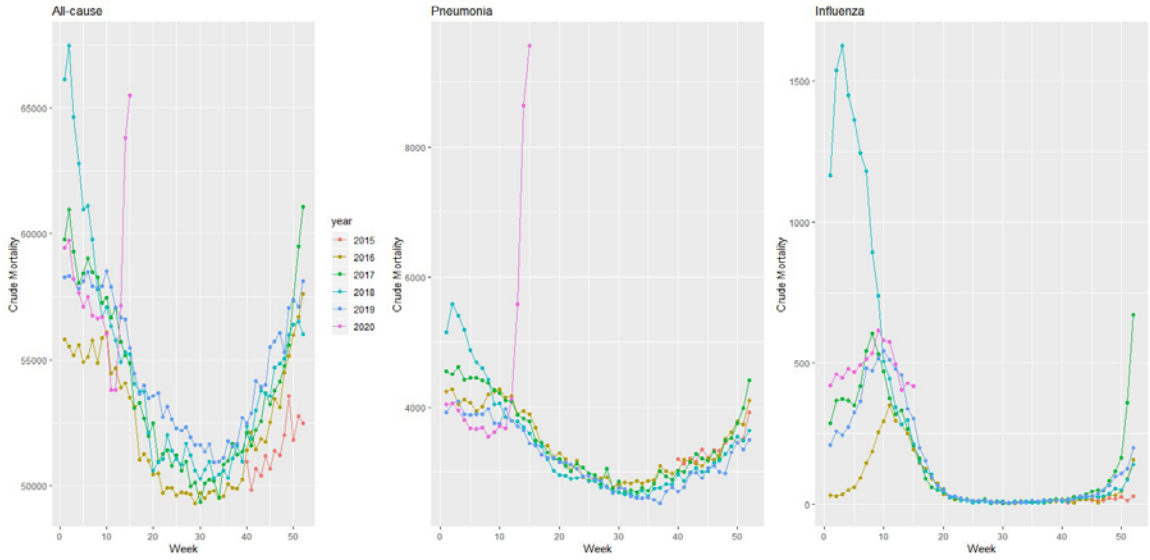

Figure 3: Weekly mortality in the United States by year for the age 65+ group starting on week 40 of 2015 until 4/11/2020.