medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Excess mortality and potential undercounting of COVID-19 deaths by demographic group in Ohio

Troy Quast, Ph.D.,[1] Ross Andel, Ph.D.[2]

[1]University of South Florida, College of Public Health,
13201 Bruce B Downs Blvd MDC 56, Tampa, FL 33612
[2]University of South Florida, College of Behavioral and Community Sciences,
4202 E Fowler Ave, Tampa, FL 33620;
Charles University and Motol University Hospital, Department of Neurology, Prague;
International Clinical Research Center, St. Anne's University Hospital, Brno

Corresponding Author: Troy Quast (University of South Florida, College of Public Health, Tampa, FL; troyquast@usf.edu; 813.974.6730)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.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/10.1101/2020.06.28.20141655; this version posted June 29, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Background: There are significant gaps in our understanding of the mortality effects of COVID-19 due to evolving diagnosis criteria, shortages of testing supplies, and challenges faced by physicians in treating patients in crisis environments. Accurate information on the number of deaths caused by COVID-19, both population wide and for demographic subgroups, is vital for policy makers and health care providers.

Methods: We performed a retrospective study of weekly data for Ohio, a large American state. To estimate expected mortality in 2020 we employed data from 2010 through 2019, adjusted for secular trends and seasonality. We estimated excess mortality as the number of observed deaths less the number of expected deaths. We conducted the analysis for the entire population and by gender, race, age, and county of residence.

Findings: We estimated 1,485 (95% CI 680-2,345) excess all-cause deaths in Ohio from March 15, 2020 through May 23, 2020. When limited to deaths due to natural causes, the estimated excess number of deaths increased to 2,504 (95% CI 1,633-3,221), reflecting the countervailing effect of a decrease in deaths due to external causes. While the largest number excess of deaths was observed in the 80+ age group, excess deaths comprised 45.3% (95% CI 21.8-60.9) of observed deaths in the groups corresponding to ages between 20 and 49 years old. Our estimate of 729 (95% CI 355-966) excess deaths for this group substantially exceeds the reported number of COVID-19 deaths of 51. We found elevated excess deaths for older individuals, blacks, and males.

Interpretation: Our methodology addressed some of the challenges of estimating the number of deaths caused by COVID-19. Our finding of high proportional levels of excess deaths among younger age groups suggests that increases in the infection rates for this cohort may have a greater mortality impact than expected.

Funding: None.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 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. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

A primary challenge in understanding COVID-19 has been the difficulty in accurately determining whether a death was caused by the novel coronavirus SARS-CoV-2. Evolving diagnosis criteria,[1] testing supply constraints and the prioritization of tests for living patients,[2] and the "fog of war" present in burdened intensive care units[3] can pose obstacles to physicians in correctly identifying COVID-19 deaths. Accurate mortality data are crucial for effective medical and policy responses to the virus.

A potentially insightful approach to estimating COVID-19 deaths is to use historical data aggregated across multiple causes of death to estimate the expected (or counterfactual) number of deaths. The number of excess deaths due to the virus can then be estimated as the observed number of deaths less the expected number. This approach abstracts from errors due to the incorrect coding of deaths if the reported cause of death falls in the same grouping as the true cause of death. This methodology has been used to estimate deaths following natural disasters[4–7] and deaths due to COVID-19.[8–15] The COVID-19 analyses found varying levels of excess mortality, but largely focused on overall mortality with the exceptions of a study of Massachusetts that found little difference in excess mortality by gender[8] and other studies that found excess deaths increased with age.[11,13,15] The authors are unaware of any studies that employ this technique to estimate excess deaths by race.

While the existing COVID-19 excess mortality retrospective studies provide important insight, they can have significant limitations. Their use of all-cause mortality likely understates the effects of COVID-19 by the inclusion of deaths by external causes, which were likely reduced during the outbreak due to reductions in accidents. Differential effects on excess deaths by race remain unknown and little is known how excess deaths compare to reported COVID-19 deaths by demographic group. Outside of studies of New York City,[9,10] there have been no analyses of sub-state areas.

The goal of this study was to estimate excess mortality in Ohio, the seventh-largest U.S. state and a state that provides comprehensive health data online. In many ways the state is representative of the country, with a similar racial and age profile and comparable median income.[16] The state has been aggressive in controlling the COVID-19 outbreak by being the first state in the country to close its schools[17] and soon thereafter closing restaurants and bars.[18] The state began allowing businesses to reopen on May 4, 2020 and for the next eight weeks experienced falling rates of new cases and COVID-19 hospitalizations.[19] We analyzed excess overall mortality as well as by demographic group. We further compared our excess mortality estimates to the number of reported COVID-19 deaths.

Methods

Study design and data sources

We performed a weekly time series analysis in which ten years of historical data were used to estimate the expected number of deaths in 2020. We defined excess mortality as the number of observed deaths less the expected number. Our mortality data were obtained from the Ohio Public Health Information Warehouse published by the Ohio Department of Health (ODH). The data for 2010-2018 were finalized, while data for 2019-2020 were provisional. The population data employed were the 2018 vintage version of bridged-race population postcensal estimates published by the National Vital Statistics System branch of the U.S. Centers for Disease Control and Prevention. Population data were not available for 2019 and 2020. To approximate those values, for each subgroup we calculated the 2013-2018 compound annual growth rate and applied the rate to the 2018 value. We obtained the number of reported COVID-19 deaths from the ODH's COVID-19 dashboard. These data are reported by county health departments to the state via the Ohio Disease Reporting System. For each death the county of residence, gender, and age are

medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 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. It is made available under a CC-BY-NC-ND 4.0 International license .

reported, as well as the dates of COVID-19 onset, hospital admission, and death. The study population was all Ohio residents from January 3, 2010 through May 23, 2020.

Statistical analysis

Our baseline period was the full set of epidemiological weeks for each year from 2010 through 2019. The analytic period begins with the week ending March 21, 2020, which was the week during which the first COVID-19 death was reported in Ohio, and ends with the week ending May 23, 2020. The observed number of deaths refers to the actual number of deaths in 2020, while the expected number of deaths corresponds to the estimate based on data from the baseline period. As the variation in the number deaths is greater than the mean, an over-dispersed log-linear model was used to estimate the expected number of deaths. The exposure variable was the relevant population. Annual indicator variables were included to account for secular trends in the number of deaths, while harmonic variables consisting of four Fourier terms were used to adjust for seasonality. The model was estimated separately for all residents and by gender, age group, race, and county. The estimating equation was:

$$\log(deaths_{sit}) = \log(pop_{sit}) + \beta_0 + \beta_1 \sin\left(\frac{2\pi t}{52.1775}\right) + \beta_2 \cos\left(\frac{2\pi t}{52.1775}\right) + \beta_3 \sin\left(\frac{4\pi t}{52.1775}\right) + \beta_4 \cos\left(\frac{4\pi t}{52.1775}\right) + \gamma_i$$

where *deaths* is the number of deaths for subgroup *s* in epidemiological year *i* and week *t*. Three models were estimated for all residents: one for all-cause deaths, one for deaths due to natural causes, and one for deaths due to external causes. The models for the subgroups were based on deaths due to natural causes.

The variation in the expected number of deaths was modeled via parametric bootstrapping and follows an earlier study of COVID-19 excess mortality.[9] From each regression we obtained the asymptotic covariance matrix, which was then used along with the estimated parameter values to specify a multivariate normal distribution to approximate the sampling distribution. One hundred samples were drawn from this distribution from which the mean number of deaths was calculated for each. We then drew one hundred samples from the Poisson distribution of each of these means, resulting in 10,000 samples for each subgroup and week. The 95% confidence interval bounds were estimated as the 2.5th and 97.5th percentile of the distribution. We also calculated the number of observed deaths less the number of reported COVID-19 deaths for each subgroup.

The cause of death was based on the primary underlying cause code recorded on the death certificate. Causes of death were grouped based on the 39 selected causes as described in the Supplementary Material. By using deaths due to natural causes as our primary outcome, we abstracted from potential indirect (and likely downward) effects on the deaths due to external causes that resulted from stay-at-home behaviors including a reduction in driving and work-related accidents.

Funding source

This study did not receive any external funding.

Results

During our analytic period the state of Ohio reported 2,167 COVID-19 deaths. Over those same weeks, there 25,523 all-cause deaths and 24,359 deaths due to natural causes. For all residents, the excess numbers of all-cause deaths and deaths due to natural causes during the same period were 1,485 (95% CI 680-2,345) and 2,405 (95% CI 1,633-3,221), respectively (figure 1). The number of excess deaths due to natural causes is larger than the value across all causes because of the shortfall in the number of deaths

medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 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. It is made available under a CC-BY-NC-ND 4.0 International license .

due to external causes. The reported number of COVID-19 deaths appears to largely explain the excess deaths across all residents, as the observed number of deaths less the reported number of COVID-19 deaths falls in the 95% confidence interval for the expected number of deaths.

The expected and observed numbers of deaths shown in figures 2 through 6 are based only on deaths due to natural causes. Outside of the youngest, all of the age groups have at least one week in which the observed number of deaths exceeded the 95% confidence interval for the expected number of deaths (figure 2). The 80+ years old group had the most excess deaths at 986 (95% CI 497-1,529). While the number of deaths were substantially fewer than for the older age groups, the 20-29, 30-39, and 40-49 groups had the largest proportions of excess deaths relative to actual deaths. The corresponding values were 59.3% (95% CI 33.9-79.8), 54.7% (95% CI 29.3-68.7), and 36.3% (95% CI 14.7-49.9). The number of excess deaths across these three groups was 729 (95% CI 355-966), yet the corresponding number of reported COVID-19 deaths was only 51. To better understand the causes underlying the excess deaths for those aged 20-49, we summarized the 2010-2019 average and 2020 number of deaths for the top ten causes of death (figure 3). The excess deaths appear to be largely driven by the cause "Symptoms, signs, and abnormal findings, not elsewhere classified." Yet the ICD-10 code for deaths caused by COVID-19 (U07.1) is contained in the cause "All other diseases (Residual)", which experienced a modest increase in 2020.

The estimated number of excess deaths is higher for males (1,404, 95% CI 881-1,966) than females (965, 95% CI 456-1,555) (figure 4). Blacks appear to have suffered a disproportionate number of excess deaths (figure 5). The excess deaths as a proportion of observed deaths for blacks was 18.7% (95% CI 10.2-26.3), which is significantly higher than the corresponding proportion for whites (8.5%, 95% CI 5.1-12.1). As noted above, the number of reported COVID-19 deaths is not available by race. The estimates for the ten most populated counties in Ohio, listed in descending population order, suggest that there is not a clear association between population and the rate of excess deaths (figure 6). The most populous county (Franklin) has the most excess deaths (453, 95% CI 250-681), but the number of excess deaths in each of the next four most-populated counties does not statistically differ from zero. In terms of excess deaths as a proportion of observed deaths, the tenth most-populated county (Mahoning) has the highest point estimate at 37.0% (95% CI 14.2-51.5) while the second highest is associated with the sixth most-populated county (Lucas) at 34.4% (95% CI 15.9-47.1). The number of reported COVID-19 deaths in Franklin county was approximately 57% of the estimated excess deaths.

Discussion

We analyzed administrative data and employed a rigorous methodology to provide insight into the difficult task of identifying deaths caused by COVID-19. We identified a greater number of excess deaths when we limited our estimate to deaths due to natural causes rather than deaths due to all causes, which include external causes. Our findings indicate potential disparities in the mortality effects of COVID-19, with excess mortality especially pronounced among those aged 80 and over, men, and blacks.

Our analysis of all residents demonstrates the magnitude of excess mortality due to COVID-19 and shows how the extent becomes more pronounced when the analysis is limited to deaths due to natural causes. The substantial shortfall in deaths due to external causes may be partly due to delays in determinations from the need for investigations.[20] However, the timing of the shortfall coincides with school and restaurant closures and suggests stay-at-home behavior may have obscured the true impact of COVID-19 on excess mortality. Our investigation reveals, given the potential behavioral effects on deaths due to external causes, the value of examining deaths due to natural causes rather than exclusively relying on all-cause mortality.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 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. It is made available under a CC-BY-NC-ND 4.0 International license .

Our finding of large mortality effects among older individuals is consistent with previous studies that employ alternative estimation approaches.[21-23] However, the especially large proportional amount of excess mortality in age groups with 20-49 years has not been observed elsewhere. Given younger populations have experienced higher infection rates after the initial wave of COVID-19 in many geographical areas, our finding may be especially important as officials revise guidance and policy.

The higher proportional excess mortality that we observe for blacks is in agreement with an earlier study that found counties with higher proportions of black residents had higher relative risks of COVID-19 deaths.[24] The disparity in risk of COVID-19 death echoes disparities blacks experience for many other health conditions.[24] Given the scarce reporting of COVID-19 deaths by race, our approach can be a valuable tool to obtain insight into potential differential impacts. Our findings of increased excess deaths for males is consistent with prior studies that analyzed hospital records to assess the risk factors of COVID-19 death.[25,26] The substantial variation across counties that we observed is consistent with the varying levels of infection within countries.

The number of reported COVID-19 deaths largely explains the levels of excess deaths that we estimate, but to varying degrees. The large gap between the estimated number of excess deaths and reported COVID-19 deaths for those aged 20-49 years suggests that COVID-19 deaths for this cohort may be under-reported by local health departments. Further research is warranted to determine whether this finding is specific to Ohio. If this phenomenon is more widespread, it could suggest substantial under-stating of mortality risk to younger individuals. Also worthy of further investigation is the apparent miscoding that we observed where presumably the vast majority of COVID-19 deaths in the 20-49 age group were not assigned the correct ICD-10 code (U07.1). Such miscoding can prevent analysts from correctly assessing the true mortality impact of COVID-19.

Our study has several limitations. Our analysis was limited to one, albeit relatively large, U.S. state. There has been significant variation in the effects of COVID-19 across and within countries, so our findings do not necessarily apply universally. Our data were limited to the first several months of the outbreak and thus our results may not pertain to later stages. The methodology we employ does not definitively identify excess deaths as being due to COVID-19. Stressful events such as COVID-19 can lead to negative health outcomes such as increased incidence of cardiovascular events,[27] poor medication adherence,[28] and hypertension.[29] Potential increases in mortality due to these types of effects would have been included in our estimates of excess deaths.

Our mortality data were imperfect in that the 2019 and 2020 data were provisional and not finalized. However, as noted above, undercounting of deaths in those years should largely be limited to deaths due to external causes as these deaths often require lengthy investigation before being coded. Further, we pulled the data more than three weeks after May 23, 2020, which should have provided sufficient time for nearly all deaths due to natural causes to be properly recorded. As noted in the Supplemental Material, some deaths were excluded from the analysis due to unusable data. However, the number of lost observations was minimal. The mortality data were based on state and county of residence at the time of death, rather than the location of occurrence. This choice was made to align the data with the reported number of COVID-19 deaths, but insight could be gained by exploring potential differences from basing the data on where the death occurred. The number of reported COVID-19 deaths did not include information regarding the decedent's race. As noted in the Supplementary Material, the date of death had to be approximated for six of the 2167 deaths in the reported COVID-19 deaths. Population data were only available through 2018, thus values for 2019 and 2020 had to be estimated.

The approach we employed indicated important differences in the effects of COVID-19 across demographic groups and identified potential shortcomings in published data. Our methodology can be applied in jurisdictions that provide timely access to death certificate data. While the results may not offer perfect insight into the precise causes of death, they can provide information to help craft effective policy

medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 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. It is made available under a CC-BY-NC-ND 4.0 International license .

and medical responses. Further, the methodology is robust to the imperfect coding of deaths that is common in crisis environments.

While our study is specific to Ohio, given the state's large population and similarities with the U.S. as a whole, our results may improve the understanding of COVID-19 elsewhere. Policy makers, public health officials, and health care providers may be able to apply our findings to mitigate the destructive effects of COVID-19 and better prepare for future epidemics.

Acknowledgements

We thank the Ohio Department of Health for publishing timely mortality and COVID-19 data. The Department specifically disclaims responsibility for any analyses, interpretations or conclusions.

---

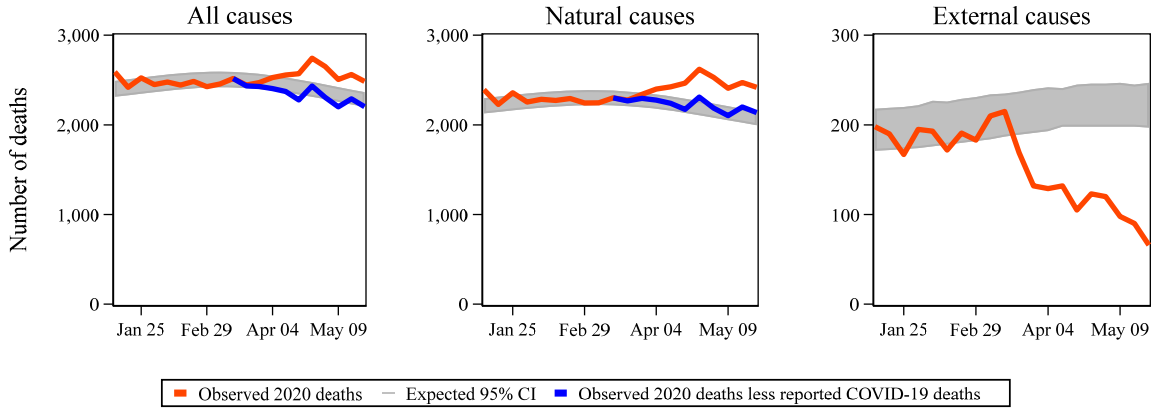
[1] Council of State & Territorial Epidemiologists. Position Statement: Standardized surveillance case definition and national notification for 2019 novel coronavirus disease (COVID-19). April 5, 2020. https://cdn.ymaws.com/www.cste.org/resource/resmgr/2020ps/Interim-20-ID-01_COVID-19.pdf (accessed May 23, 2020).
[2] Resnick B, Scott D. America's shamefully slow coronavirus testing threatens all of us. *Vox*. March 12, 2020. https://www.vox.com/science-and-health/2020/3/12/21175034/coronavirus-covid-19-testing-usa (accessed June 22, 2020).
[3] Goh KJ, Wong J, Tien J-CC, et al. Preparing your intensive care unit for the COVID-19 pandemic: practical considerations and strategies. *Critical Care* 2020; **24**:215.
[4] Stephens KU, Grew D, Chin K, et al. Excess Mortality in the Aftermath of Hurricane Katrina: A Preliminary Report. *Disaster Med Public Health Prep* 2007; **1**:15–20.
[5] Santos-Burgoa C, Sandberg J, et al. Differential and persistent risk of excess mortality from Hurricane Maria in Puerto Rico: a time-series analysis. *Lancet Planet Health* 2018; **2**:e478–e488.
[6] Santos-Lozada AR, Howard JT. Use of Death Counts From Vital Statistics to Calculate Excess Deaths in Puerto Rico Following Hurricane Maria. *JAMA* 2018; **320**:1491–1493.
[7] Whitman S, Good G, Donoghue ER, Benbow N, Shou W, Mou S. Mortality in Chicago attributed to the July 1995 heat wave. *Am J Public Health* 1997; **87**:1515–1518.
[8] Krieger N, Chen JT, Waterman PD. Excess mortality in men and women in Massachusetts during the COVID-19 pandemic. *Lancet* 2020; **395**:1829.
[9] Weinberger D, Cohen T, Crawford F, et al. Estimating the early death toll of COVID-19 in the United States. *medRXiv* 2020; published online April 29. DOI:10.1101/2020.04.15.20066431 (preprint).
[10] US Centers for Disease Control and Prevention. Excess Deaths Associated with COVID-19. June 17, 2020. https://www.cdc.gov/nchs/nvss/vsrr/covid19/excess_deaths.htm (accessed June 21, 2020).
[11] Freitas ARR, Medeiros NM, Frutuoso L, et al. Use of excess mortality associated with the COVID-19 epidemic as an epidemiological surveillance strategy - preliminary results of the evaluation of six Brazilian capitals. *medRxiv* 2020; published online May 12. DOI: 10.1101/2020.05.08.20093617 (preprint).
[12] Vieira A, Peixoto VR, Aguiar P, Abrantes A. Rapid estimation of excess mortality in times of COVID-19 in Portugal - Beyond reported deaths. *medRxiv* 2020; published online May 19. DOI: 10.1101/2020.05.14.20100909 (preprint).
[13] Kontopantelis E, Mamas MA, Deanfield J, et al. Excess mortality in England and Wales during the first wave of the COVID-19 pandemic. *medRxiv* 2020; published online June 18. DOI: 10.1101/2020.05.26.20113357 (preprint).
[14] McLeay S. Measuring Excess Mortality: A Second Look at Germany. Social Science Research Network 2020; published online May 27. DOI: 10.2139/ssrn.3626618 (preprint).
[15] Modi C, Boehm V, Ferraro S, et al. How deadly is COVID-19? A rigorous analysis of excess mortality and age-dependent fatality rates in Italy. *medRxiv* 2020; published online May 14. DOI: 10.1101/2020.04.15.20067074 (preprint).
[16] US Census Bureau. QuickFacts: Ohio; United States. https://www.census.gov/quickfacts/fact/table/OH,US/PST045219 (accessed June 9, 2020)
[17] Meckler L. Seven states, D.C. order all schools closed in effort to prevent spread of covid-19. *Washington Post*, March 13, 2020. https://www.washingtonpost.com/local/education/ohio-maryland-order-all-schools-closed-in-

medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 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. It is made available under a CC-BY-NC-ND 4.0 International license .

effort-to-prevent-spread-of-covid-19/2020/03/12/e4078b3a-6499-11ea-845d-e35b0234b136_story.html (accessed June 21, 2020).
[18] O'Brien RD, Bauerlein V. How Coronavirus Remade American Life in One Weekend. *Wall Street Journal*, March 15, 2020. https://www.wsj.com/articles/coronavirus-remakes-american-life-in-a-weekend-11584293065 (accessed June 20, 2020).
[19] Borchardt J. Ohio started reopening six weeks ago, hasn't seen a coronavirus case surge. *USA TODAY*, June 18, 2020. https://www.usatoday.com/story/news/health/2020/06/18/ohio-reopening-one-month-later-no-surge-in-coronavirus-cases/3208493001/ (accessed June 20, 2020).
[20] Spencer MR, Ahmad F. Timeliness of Death Certificate Data for Mortality Surveillance and Provisional Estimates. *Vital Statistics Rapid Release*. 2016; Report No. 001:8.
[21] Mills JP, Kaye KS, Mody L. COVID-19 in older adults: clinical, psychosocial, and public health considerations. *JCI Insight*. 2020; **5**:e139292.
[22] Applegate WB, Ouslander JG. COVID-19 Presents High Risk to Older Persons. *J Am Geriatr Soc*. 2020; **68**:681.
[23] Le Couteur DG, Anderson RM, Newman AB. COVID-19 Through the Lens of Gerontology. *J Gerontol A Biol Sci Med Sci*. 2020.
[24] Millett GA, Jones AT, Benkeser D, et al. Assessing Differential Impacts of COVID-19 on Black Communities. *Ann Epidemiol*. in press.
[25] Zhang J, Wang X, Jia X, et al. Risk factors for disease severity, unimprovement, and mortality in COVID-19 patients in Wuhan, China. *Clin Microbiol Infect*. 2020; **26**:767–772.
[26] Palaiodimos L, Kokkinidis DG, Li W, et al. Severe obesity, increasing age and male sex are independently associated with worse in-hospital outcomes, and higher in-hospital mortality, in a cohort of patients with COVID-19 in the Bronx, New York. *Metabolism*. 2020; **108**:154262.
[27] Steptoe A, Kivimäki M. Stress and cardiovascular disease. *Nat Rev Cardiol*. 2012; **9**:360–370.
[28] Kretchy IA, Owusu-Daaku FT, Danquah SA. Mental health in hypertension: assessing symptoms of anxiety, depression and stress on anti-hypertensive medication adherence. *Int J Ment Health Syst*. 2014; **8**:25.
[29] Sparrenberger F, Cichelero FT, Ascoli AM, et al. Does psychosocial stress cause hypertension? A systematic review of observational studies. *J Hum Hypertens*. 2009; **23**:12–19.

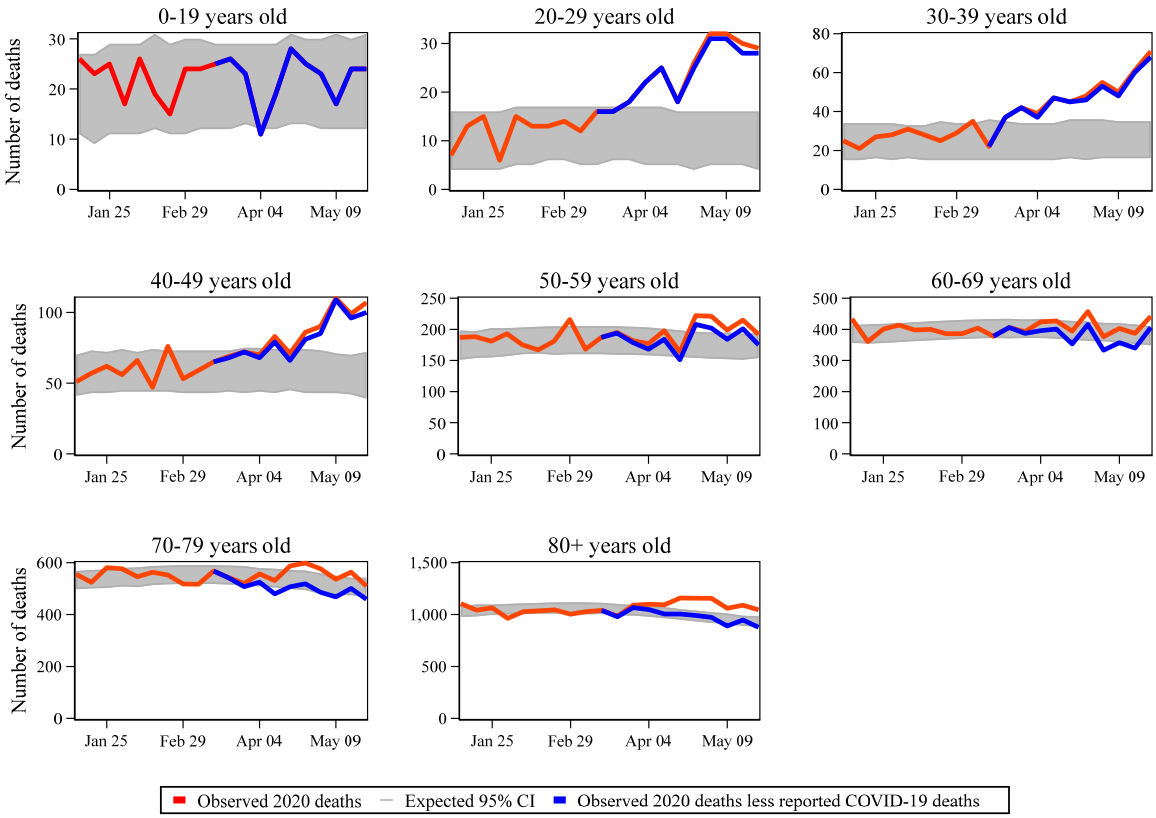
medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 1. Excess mortality for all-cause deaths, deaths due to natural causes, and deaths due to external causes.



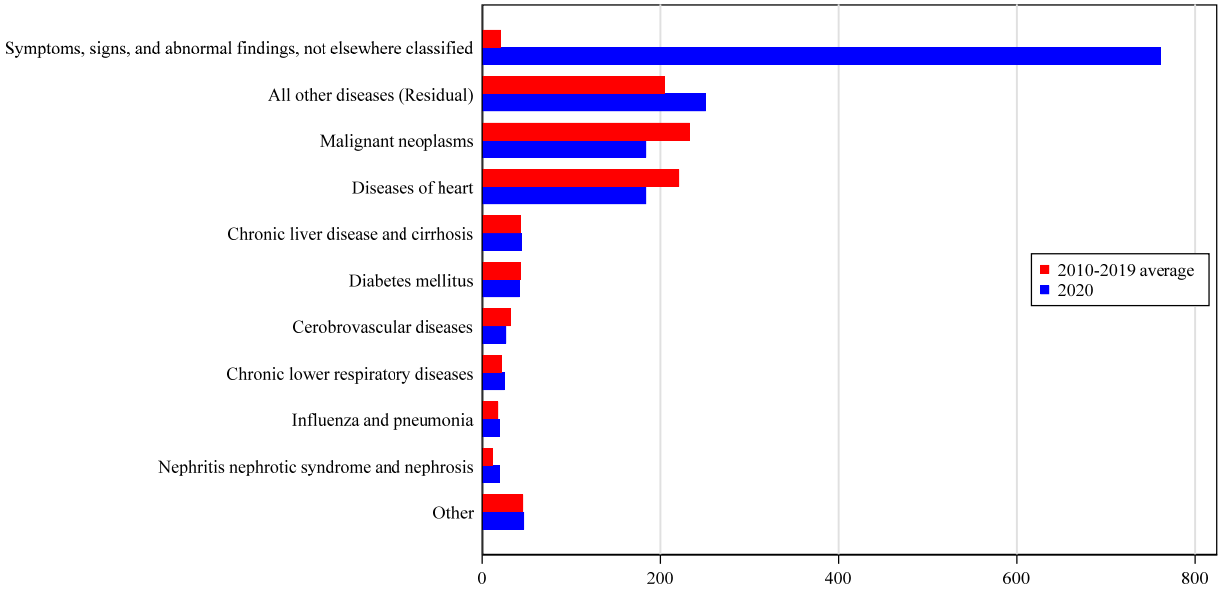
medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 2. Excess mortality for deaths due to natural causes by age group.



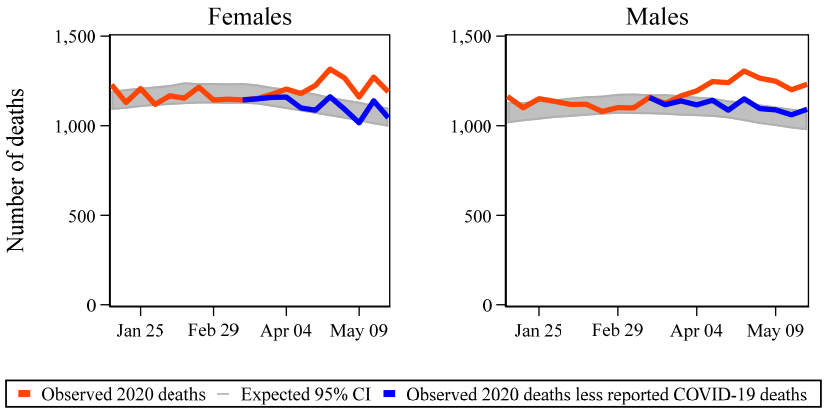
medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 3. Number of deaths by cause of death group, 2010-2019 average and 2020.



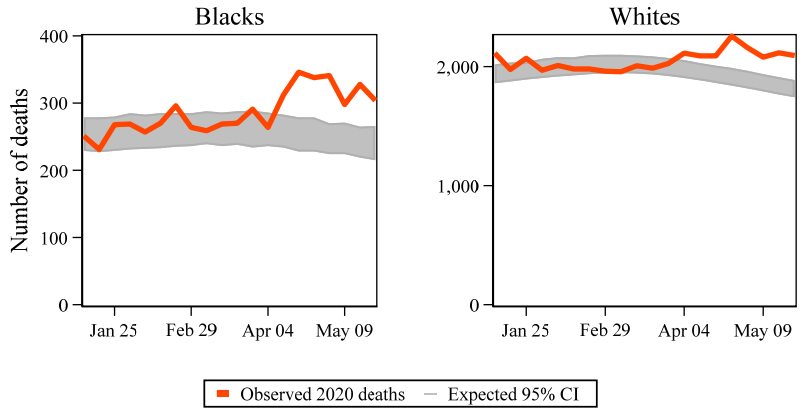
medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 4. Excess mortality for deaths due to natural causes by gender.



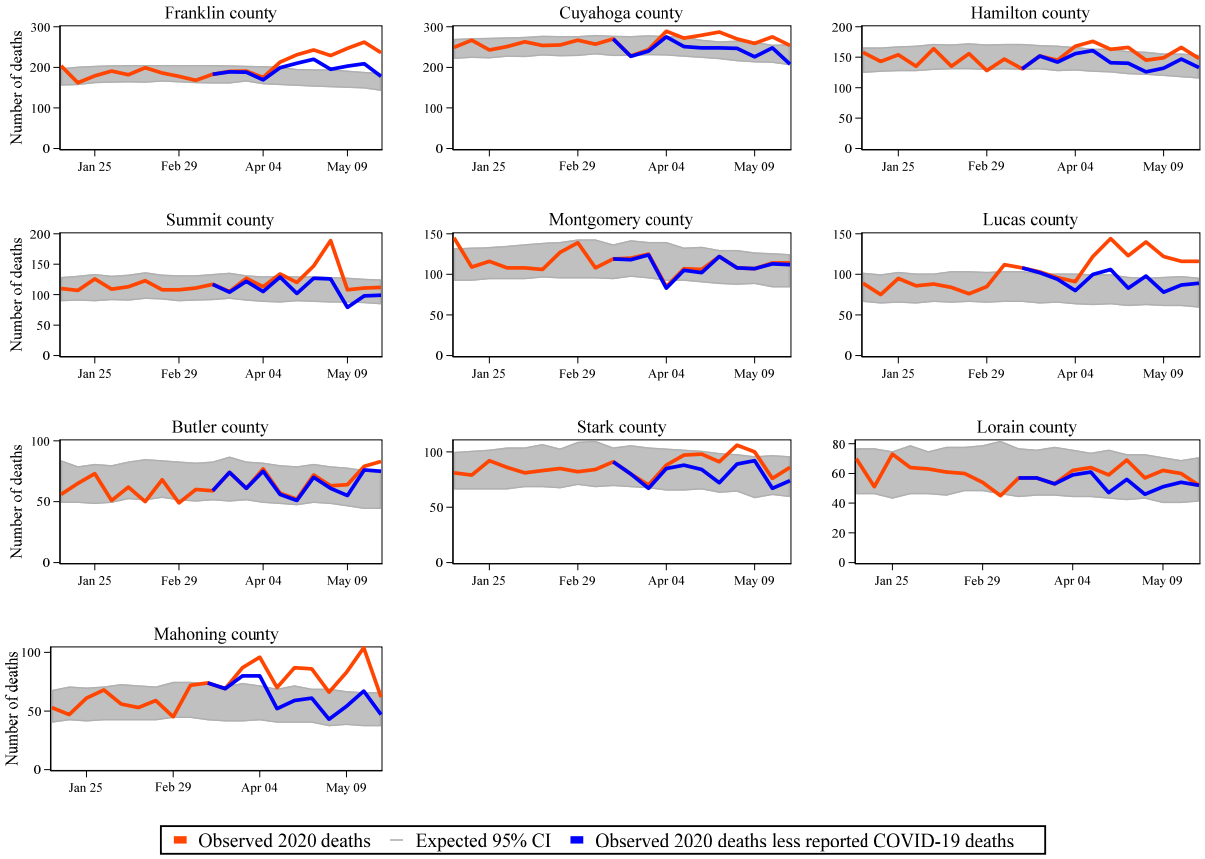
medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 5. Excess mortality for deaths due to natural causes by race.



medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 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. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 6. Excess mortality for deaths due to natural causes by county of residence.



medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 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. It is made available under a CC-BY-NC-ND 4.0 International license .

Supplementary Material

Data sources

The Ohio Public Health Information warehouse is located at http://publicapps.odh.ohio.gov/EDW/DataCatalog. The data for the analyses of the aggregated number of all causes of death, deaths due to natural causes, and deaths due to external causes were obtained on June 13, 2020. The data for the specific causes of death for 20-49 years old was obtained on June 17, 2020.

The Ohio COVID-19 dashboard is located at https://coronavirus.ohio.gov/wps/portal/gov/covid-19/dashboards/overview. These data were obtained on June 15, 2020 and were limited to deaths that occurred through May 23, 2020. There were valid dates of death during the analytic period for 2161 deaths. Twelve observations were listed as "Unknown" dates of death. Two of these observations had a valid hospital admission date. For these observations, the date of death was approximated as seven days after the admission date. Of these two observations, one had an approximated date that fell within the analytic period and was thus included in our data. The ten remaining observations did not have a valid hospital admission date but did have a valid onset date value. For these observations, the date of death was approximated as fourteen days after the onset date. Of these ten observations, five had an approximated date that fell within the analytic period and was thus included in the sample. The final dataset contained 2,167 deaths.

The data and code employed in the analysis are available at https://github.com/troyquast/covid19_ohio.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 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. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 1. Cause-of-death types and ICD-10 codes.

Cause type	Cause of death	ICD-10 codes
Natural	Tuberculosis	A16-A19
Natural	Syphilis	A50-A53
Natural	Human immunodeficiency virus (HIV) disease	B20-B24
Natural	Malignant neoplasm of stomach	C16
Natural	Malignant neoplasms of colon rectum and anus	C18-C21
Natural	Malignant neoplasm of pancreas	C25
Natural	Malignant neoplasms of trachea bronchus and lung	C33-C34
Natural	Malignant neoplasm of breast	C50
Natural	Malignant neoplasms of cervix uteri corpus uteri and ovary	C53-C56
Natural	Malignant neoplasm of prostate	C61
Natural	Malignant neoplasms of urinary tract	C64-C68
Natural	Non-Hodgkin's lymphoma	C82-C85
Natural	Leukemia	C91-C95
Natural	Other malignant neoplasms	C00-C15,C17,C22-C24,C26-C32,C37-C49,C51-C52,C57-C60,C62-C63,C69-C81,C88,C90,C96-C97
Natural	Diabetes mellitus	E10-E14
Natural	Alzheimer's disease	G30
Natural	Hypertensive heart disease with or without renal disease	I11,I13
Natural	Ischemic heart diseases	I20-I25
Natural	Other diseases of heart	I00-I09,I26-I51
Natural	Essential hypertension and hypertensive renal disease	I10,I12,I15
Natural	Cerebrovascular diseases	I60-I69
Natural	Atherosclerosis	I70
Natural	Other diseases of circulatory system	I71-I78
Natural	Influenza and pneumonia	J9-J18
Natural	Chronic lower respiratory diseases	J40-J47
Natural	Peptic ulcer	K25-K28
Natural	Chronic liver disease and cirrhosis	K70,K73-K74
Natural	Nephritis nephrotic syndrome and nephrosis	N00-N07,N17-N19,N25-N27
Natural	Pregnancy childbirth and the puerperium	O00-O99
Natural	Certain conditions originating in the perinatal period	P00-P96
Natural	Congenital malformations deformations and chromosomal abnormalities	Q00-Q99
Natural	Sudden infant death syndrome	R95
Natural	Symptoms signs and abnormal clinical and laboratory findings not elsewhere classified excluding Sudden infant death syndrome)	R00-R94,R96-R99
Natural	All other diseases (Residual)	A00-A09,A20-A49,A54-B19,B25-B99,D00-E07,E15-G25,G31-H93,I80-J06,J20-J39,J60-K22,K29-K66,K71-K72,K75-M99,N10-N15,N20-N23,N28-N98,U04.9,U07.1
External	Motor vehicle accidents	V02-V04,V09.0,V09.2,V12-V14,V19.0-V19.2,V19.4-V19.6,V20-V79,V80.3-V80.5,V81.0-V81.1,V82.0-V82.1,V83-V86,V87.0-V87.8,V88.0-V88.8,V89.0,V89.2
External	All other and unspecified accidents and adverse effects	V01,V05-V06,V09.1,V09.3-V09.9,V10-V11,V15-V18,V19.3,V19.8-V19.9,V80.0-V80.2,V80.6-V80.9,V81.2-V81.9,V82.2-V82.9,V87.9,V88.9,V89.1,V89.3,V89.9,V90-X59,Y40-Y86,Y88
External	Intentional self-harm (suicide)	*U03,X60-X84,Y87.0
External	Assault (homicide)	*U01-*U02,X85-Y09,Y87.1
External	All other external causes	Y10-Y36,Y87.2,Y89

ICD-10: Tenth Revision of the International Classification of Diseases

*: Code was introduced by the United States, but is not officially part of the World Health Organization's International Classification of Diseases

medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 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. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2. Number of deaths by sample after restricting to usable observations.

Sample	Restriction	Number of deaths
State all causes	Initial	1,221,053
	Drop week equal to "99"	1,221,041
Natural causes by county	Initial	1,125,202
	Drop week equal to "99"	1,125,197
	Drop county equal to "NonOH" and "Unknown"	1,124,983
Natural causes by gender	Initial	1,125,214
	Drop week equal to "99"	1,125,209
	Drop gender equal to "Unknown"	1,123,823
Natural causes by race	Initial	1,125,214
	Drop week equal to "99"	1,125,209
	Limit to race equal to "Black" or "White"	1,114,833
Natural causes by age	Initial	1,125,212
	Drop week equal to "99"	1,125,207
	Age group equal to "Unknown"	1,123,857

Number of deaths reflect data obtained for all epidemiological weeks in 2010 through 2019 and for weeks one through 21 in 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20141655; this version posted June 29, 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. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 3. Observed and excess deaths due to natural causes and reported COVID-19 deaths in Ohio from March 15, 2020 through May 23, 2020.

Subgroup	Observed deaths	Excess deaths (95% CI)		Excess deaths as a percentage of observed deaths (95% CI)		Reported COVID-19 deaths
.						
All	24,359	2,405	(1,633 to 3,221)	9.9%	(6.7 to 13.2)	2,167
By gender						
Female	12,137	965	(456 to 1,555)	8.0%	(3.8 to 12.8)	1,031
Male	12,221	1,404	(881 to 1966)	11.5%	(7.2 to 16.1)	1,136
By age						
0-19	220	31	(-79 to 97)	14.1%	(-35.9 to 44.1)	0
20-29	248	147	(84 to 198)	59.3%	(33.9 to 79.8)	6
30-39	495	271	(145 to 340)	54.7%	(29.3 to 68.7)	12
40-49	857	311	(126 to 428)	36.3%	(14.7 to 49.9)	33
50-59	1,965	184	(-19 to 412)	9.4%	(-1.0 to 21.0)	119
60-69	4,109	197	(-139 to 473)	4.8%	(-3.4 to 11.5)	314
70-79	5,523	228	(-144 to 588)	4.1%	(-2.6 to 10.6)	529
80+	10,941	986	(497 to 1529)	9.0%	(4.5 to 14.0)	1,154
By race						
Blacks	3092	578	(316 to 812)	18.7%	(10.2 to 26.3)	NR
Whites	21,014	1,783	(1,069 to 2,528)	8.5%	(5.1 to 12.1)	NR
By county						
Franklin	2,217	453	(250 to 681)	20.4%	(11.3 to 30.7)	251
Cuyahoga	2,657	224	(-26 to 548)	8.4%	(-1.0 to 17.2)	239
Hamilton	1,578	140	(-41 to 348)	8.9%	(-2.6 to 22.1)	148
Summit	1,266	223	(-31 to 387)	17.6%	(-2.4 to 30.6)	175
Montgomery	1,108	44	(-219 to 211)	4.0%	(-19.8 to 19.0)	14
Lucas	1,173	404	(186 to 552)	34.4%	(15.9 to 47.1)	256
Butler	682	79	(-117 to 206)	11.6%	(-17.2 to 30.2)	28
Stark	892	110	(-110 to 256)	12.3%	(-12.3 to 28.7)	94
Lorain	595	44	(-142 to 168)	7.4%	(-23.9 to 28.2)	59
Mahoning	810	300	(115 to 417)	37.7%	(14.2 to 51.5)	198

Observed deaths taken from the Ohio Department of Health's Ohio Public Health Information Warehouse. Excess deaths and excess deaths as percentage of observed deaths based on author's calculations. Reported COVID-19 deaths based on data from the Ohio Department of Health's COVID-19 dashboard.
NR: Not reported.