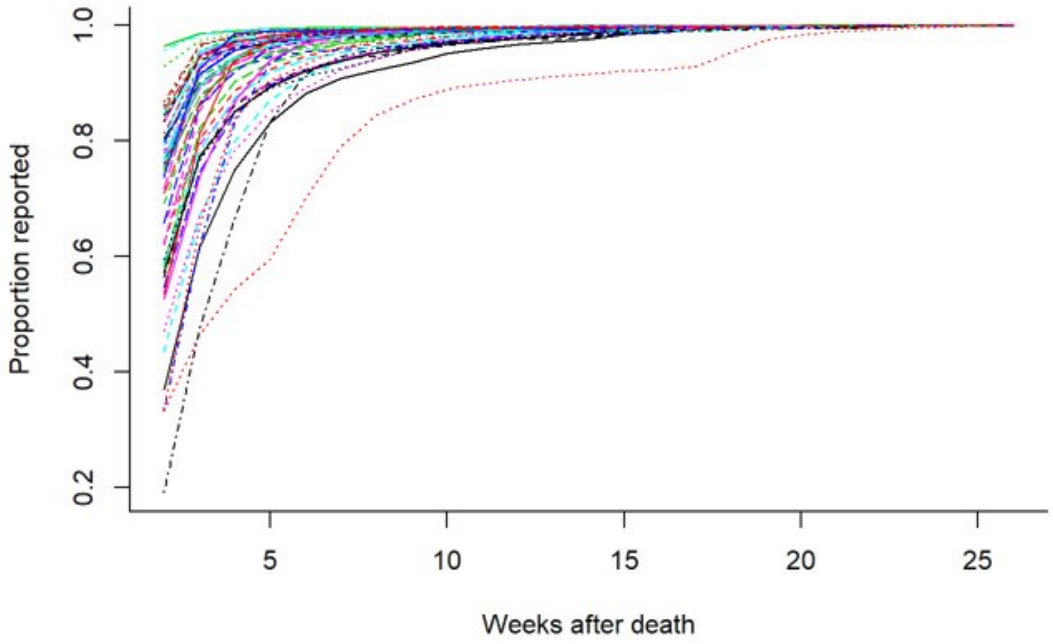
Supplementary Online Content

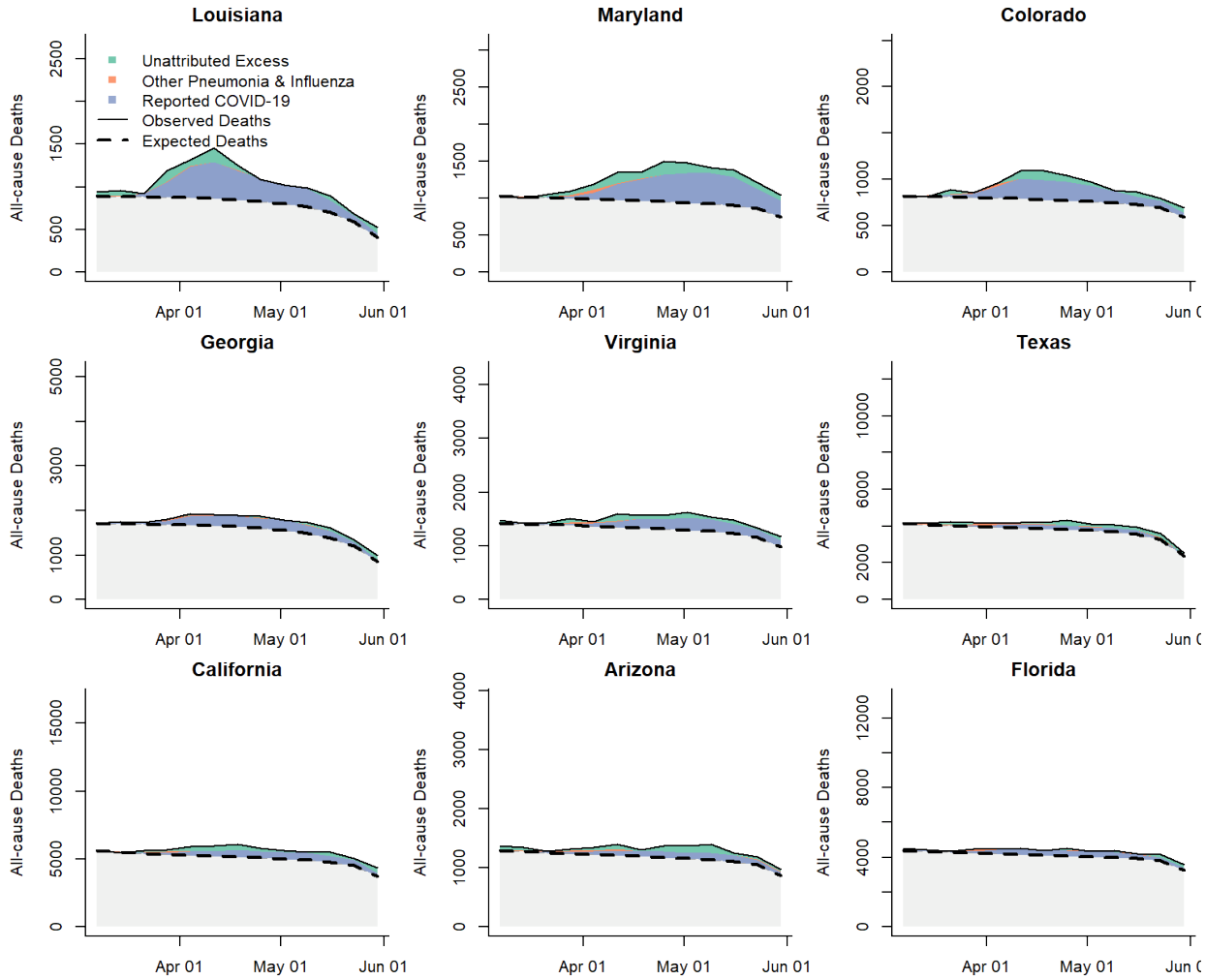
Weinberger DM, Chen J, Cohen T, et al. Estimation of excess deaths associated with the COVID-19 pandemic in the United States, March to May 2020. *JAMA Intern Med*. Published online July 1, 2020. doi10.1001/jamainternmed.2020.3391

This supplementary material has been provided by the authors to give readers additional information about their work.



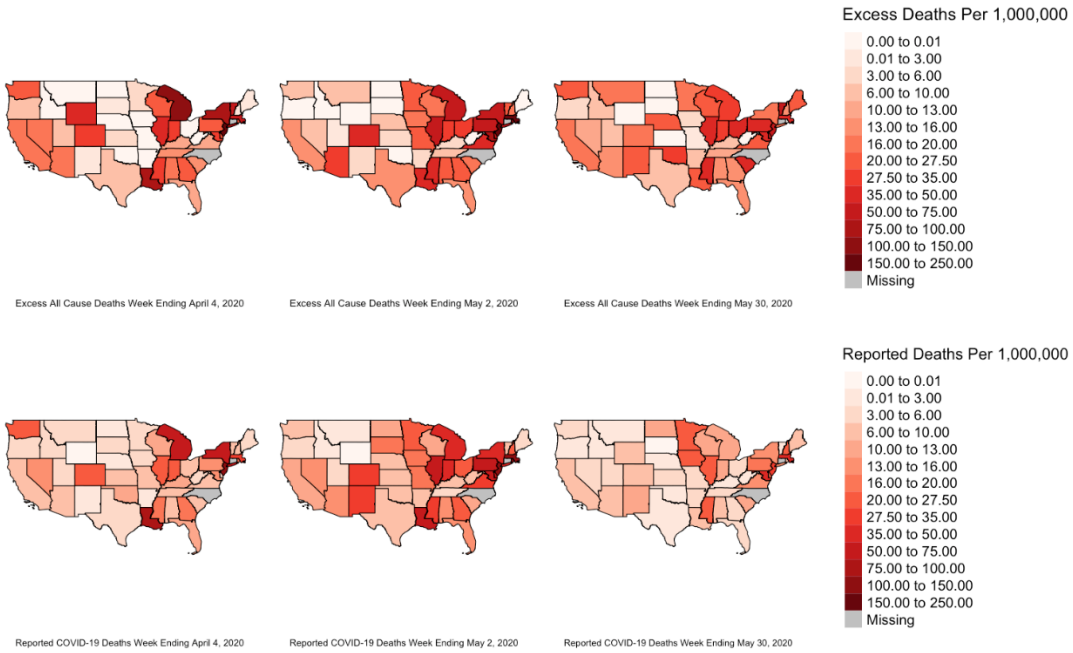
eFigure 1. Estimated reporting delays by state

Each line represents one of the states, showing the estimated proportion of deaths that have been reported relative to the time since death.

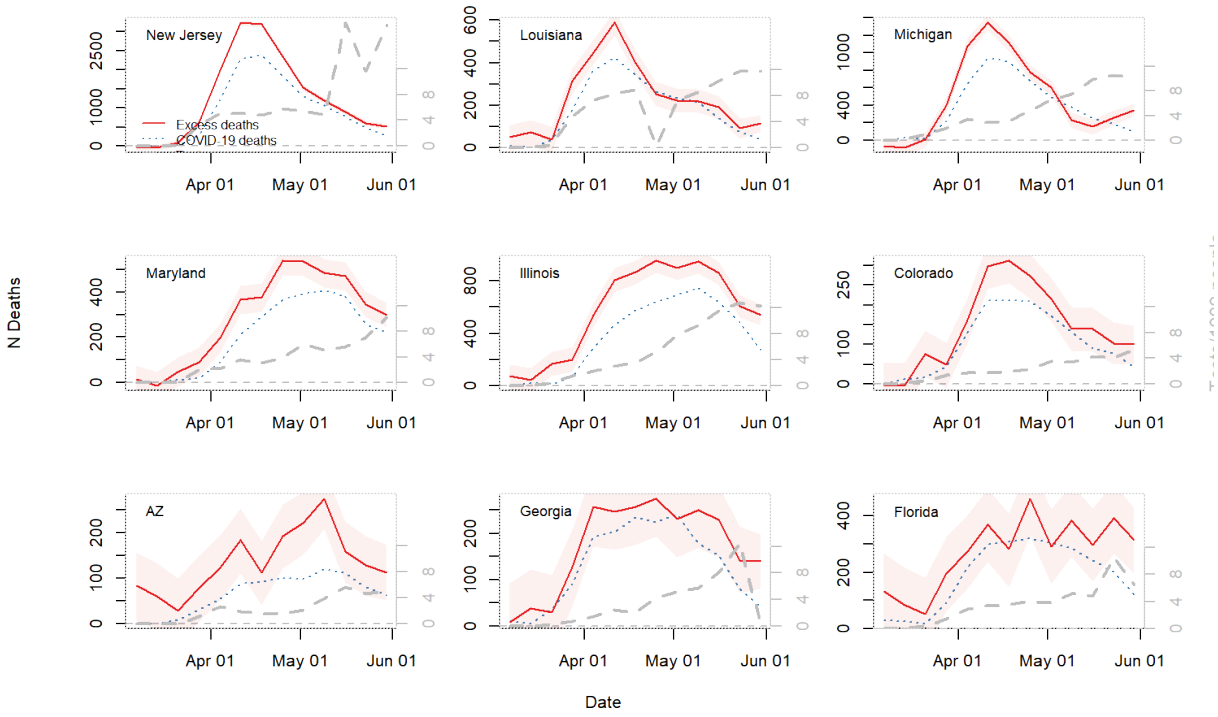


eFigure 2. Excess deaths for additional states from March 1, 2020-May 30, 2020

The observed number of deaths is indicated by the black solid line, and the expected number of deaths, adjusting for seasonality, influenza epidemics, and reporting delays, is indicated by the black dashed line. Area between these two lines represents the total number of excess deaths. The blue area represents deaths recorded as due to COVID-19, the red shaded area presents additional pneumonia and influenza excess deaths not coded as due to COVID-19, and the green shaded area represents deaths that were not attributed to COVID-19 or pneumonia or influenza. The axes are scaled relative to the first time point to make relative increases easier to compare across states.

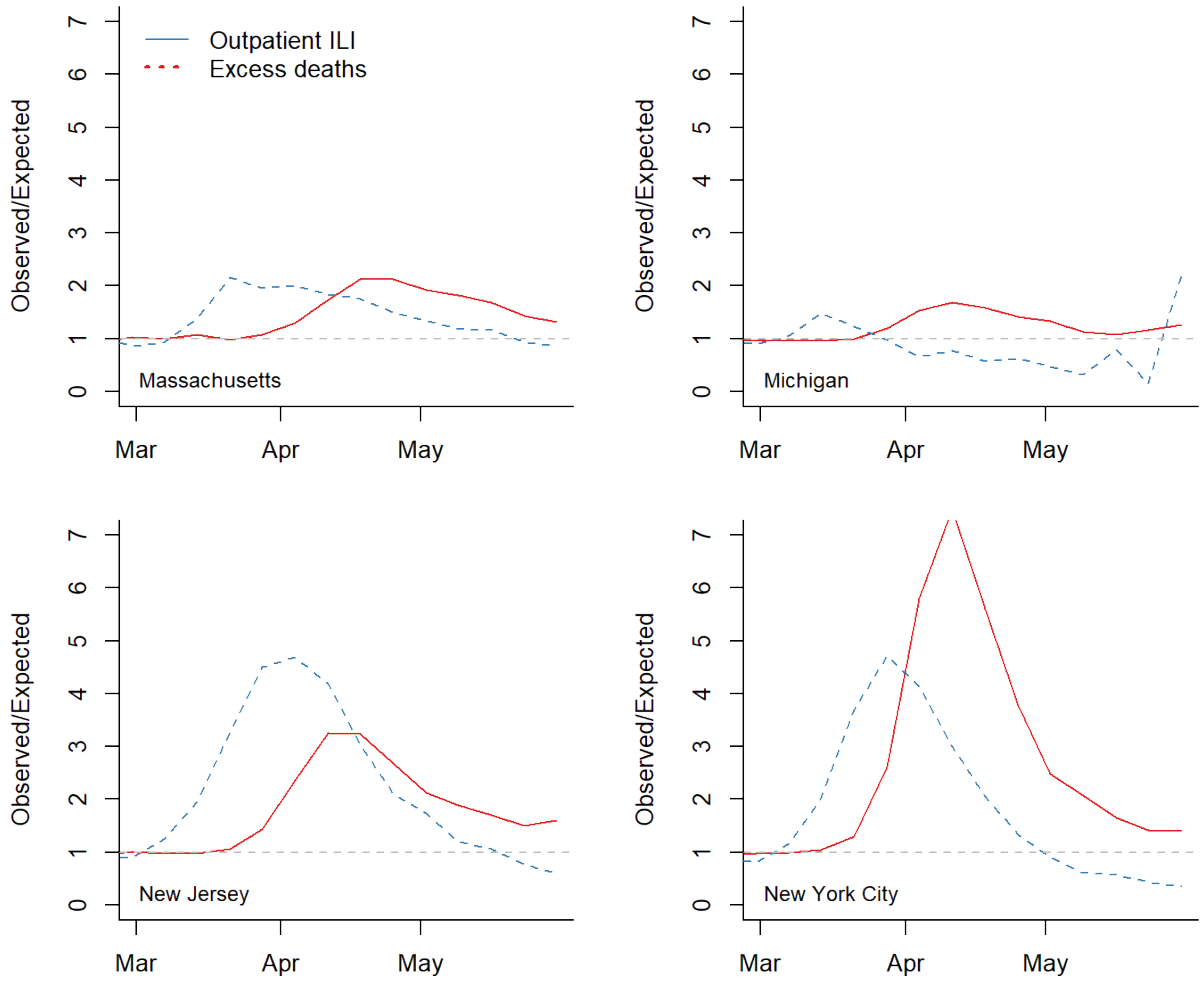


eFigure 3. Map of Excess all-cause death deaths by state, and COVID-19 deaths reported by NCHS

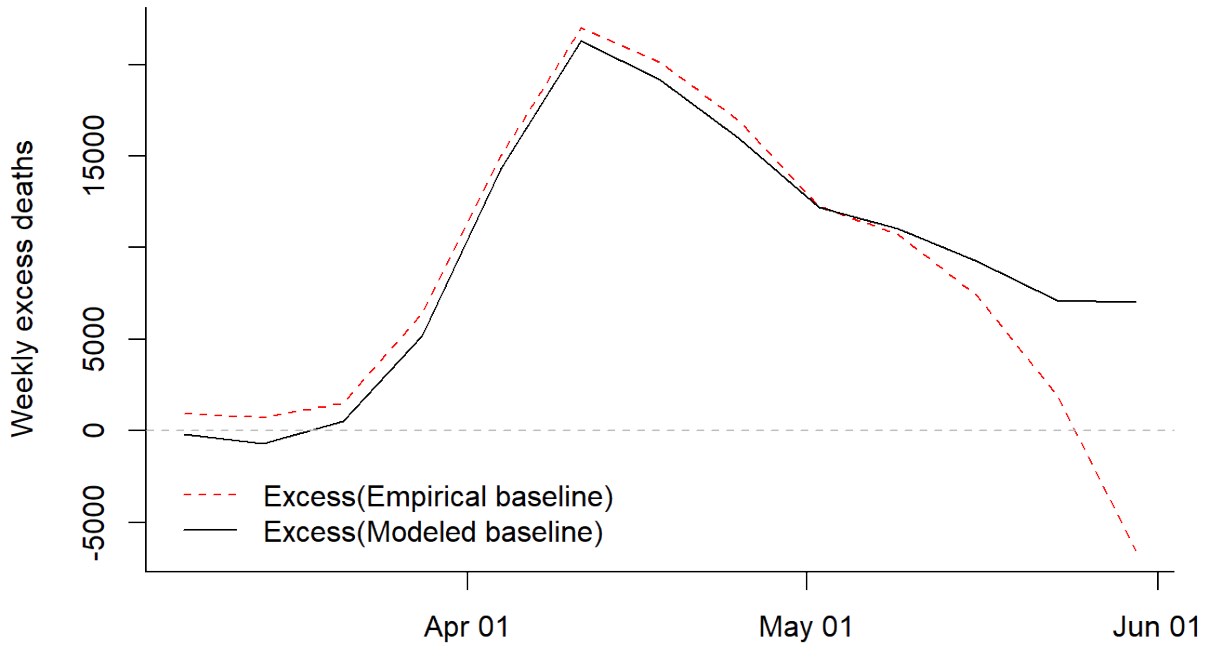


eFigure 4. Trends in excess mortality due to all causes (red solid line +/- 95% prediction intervals) or reported deaths due to COVID-19 (blue dotted line) for March 1, 2020 to May 30, 2020

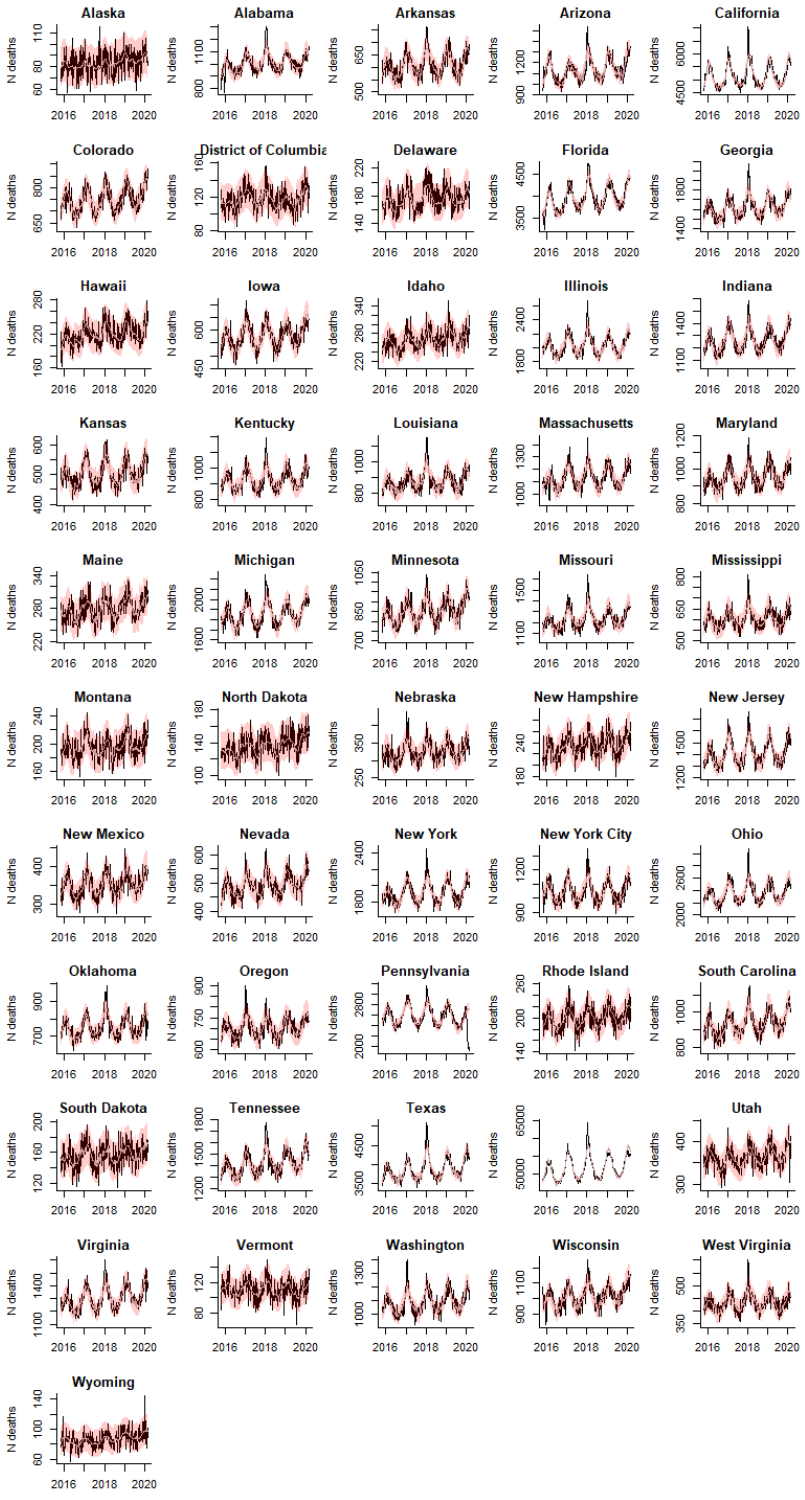
The thick dashed gray line shows the volume of tests performed/1000 people in that week.



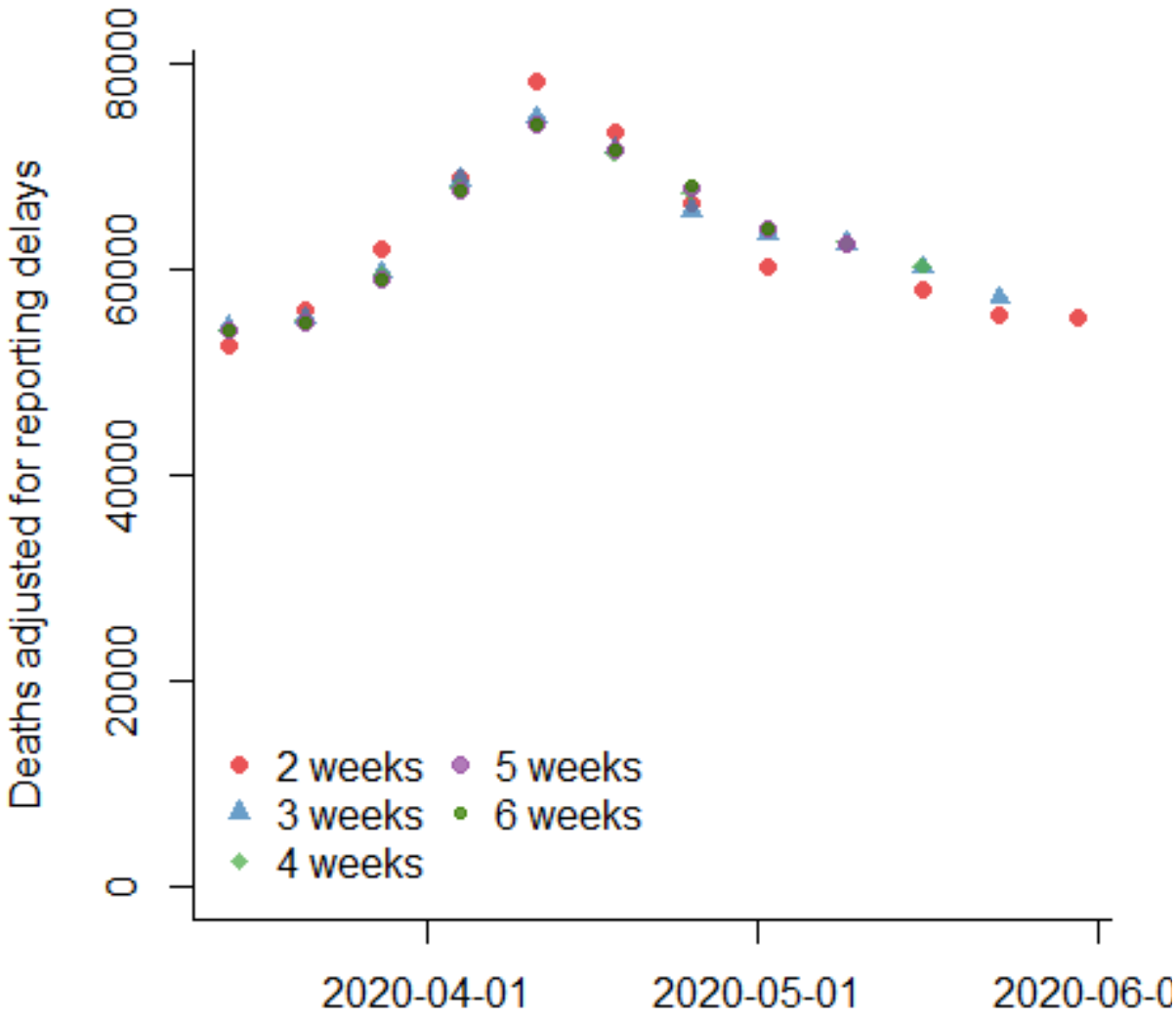
eFigure 5. Relative increase (observed/expected) for influenza-like illness (blue dashed line) and all-cause deaths (red solid line) in select states with large epidemics



eFigure 6. Comparison of excess deaths per week estimated from the main regression model (black solid line) or using an empirical baseline (mean deaths per week from previous years)



eFigure 7. Observed deaths per week, compared to the model fitted deaths (red line) +/- 95% prediction intervals, through end of February 2020



eFigure 8. Evaluation of how reporting delay adjustments influence the magnitude of deaths at the national scale

Data reported weekly in the CDC fluview reports from mid-March onwards were taken, and the reported number of deaths in each week was divided by the estimated proportion complete based on whether it was 2,3,4, 5, or 6 weeks after the death. Each dot represents a corrected estimate taken from a different vintage of CDC fluview report. This demonstrates that nationally, the reporting delay correction modestly under-corrected for the reporting delay by 5-8% 2 weeks after the death (red dot) for dates of death in May 2020, and then the estimates stabilized 3+ weeks after the death.

State	Excess deaths, adjusted for influenza	Excess deaths, unadjusted for influe
AK	-100 (-200,0)	-70 (-160,0)
AL	900 (600,1200)	710 (420,1000)
AR	100 (-200,300)	80 (-170,300)
AZ	1800 (1400,2100)	1250 (900,1570)
CA	6800 (6100,7500)	6190 (5500,6850)
CO	1900 (1600,2100)	1730 (1450,1980)
DC	500 (400,600)	490 (400,590)
DE	400 (300,600)	440 (320,550)
FL	3500 (2900,4100)	2770 (2140,3370)
GA	2200 (1800,2600)	1960 (1580,2310)
HI	-100 (-200,100)	-80 (-230,50)
IA	100 (-100,400)	150 (-90,380)
ID	0 (-100,200)	10 (-160,160)
IL	7500 (7100,7900)	7370 (6910,7760)
IN	2100 (1800,2500)	2120 (1780,2440)
KS	0 (-200,200)	10 (-200,220)
KY	100 (-200,300)	-40 (-330,220)
LA	3000 (2700,3300)	2680 (2410,2940)
MA	7400 (7100,7700)	7160 (6830,7490)
MD	3700 (3500,4000)	3680 (3380,3960)
ME	0 (-200,100)	-70 (-230,90)
MI	6100 (5700,6600)	5950 (5540,6330)
MN	900 (600,1200)	790 (500,1050)
MO	700 (400,1000)	660 (320,960)
MS	1100 (900,1400)	1120 (880,1340)
MT	0 (-100,100)	10 (-120,130)
ND	-500 (-700,-400)	-560 (-680,-450)
NE	100 (0,300)	130 (-40,290)
NH	300 (100,400)	270 (100,400)
NJ	16200 (15800,16500)	15950 (15600,16300)
NM	100 (-100,300)	50 (-140,220)
NV	200 (0,400)	200 (-20,410)
NY	12300 (11900,12700)	11990 (11580,12400)
NYC	25100 (24800,25400)	25000 (24700,25280)
OH	1600 (1200,2000)	1520 (1060,1940)
OK	400 (100,600)	370 (110,600)
OR	300 (0,500)	240 (-10,480)
PA	5400 (4900,5900)	5470 (4980,5910)
RI	500 (400,700)	550 (410,660)
SC	1400 (1100,1700)	1330 (1020,1600)
SD	-100 (-200,0)	-80 (-200,30)
TN	800 (400,1100)	440 (90,790)
TX	3600 (3000,4200)	2170 (1590,2750)
UT	100 (-100,300)	100 (-90,280)
VA	2300 (1900,2600)	2070 (1740,2410)
VT	200 (100,300)	190 (80,280)
WA	1300 (1000,1700)	1160 (860,1470)
WI	600 (300,900)	490 (170,790)
WV	-600 (-800,-500)	-640 (-830,-470)
WY	-300 (-400,-300)	-330 (-410,-250)

eTable. Comparison of baselines that are or are not adjusted for influenza

Observed and excess deaths due to pneumonia & influenza, and COVID-19, from March 1, 2020 through May 30, 2020. The excess estimates represent a median and 95% prediction intervals.

eAppendix. Supplemental Methods

Datasets:

Details on the choice of mortality indicators

We used multiple cause of death data to extract deaths with PIC causes (pneumonia, influenza , coronavirus) listed anywhere on the death certificate. Excess mortality from pneumonia and influenza has been used in the US to monitor the severity of influenza since the 1918 pandemic. Here we concentrate on PIC mortality rather than pneumonia alone, or P&I alone, to be more comprehensive. Deaths coded as 'influenza' or 'coronavirus' do not necessarily require laboratory confirmation of infection, and there is overlap of symptoms between influenza and COVID-19. The PIC grouping includes individuals with a cause of death listed as COVID-19 (either with or without a P&I code) as well as people who did not have COVID-19 listed but did have a cause of death of pneumonia or influenza. PIC codes could be present alone or in combination.

Reported COVID-19 deaths

The number of COVID-19 deaths reported to NCHS were used for most states. NCHS suppresses data when there are 1-9 counts for a particular week/state—this was an issue for a few of the smaller states. In these instances, the data from the Covid tracking project for that week and state were substituted in. In more recent weeks, the data from the Covid tracking project were higher than the official tallies in many states, likely due to shorter reporting delays. Therefore, this substitution would have the effect of shrinking the gap between excess deaths and reported COVID-19 deaths. Because this is only an issue in states with small counts, it does not meaningfully change the overall estimates.

Baseline data on pneumonia and influenza deaths

For the pre-pandemic period, there were two datasets that were combined to get the number of pneumonia and influenza deaths. NCHS provides data on all-cause deaths and pneumonia and influenza deaths by week and state for previous years based on *underlying cause*. CDC fluview produces estimates of the number and proportion of deaths due to pneumonia and influenza by state and week that is based on *multiple-cause* of death coding (pneumonia/influenza code anywhere on the death certificate). But the CDC fluview data are based on state of residence, rather than state of occurrence. To get an estimate of pneumonia and influenza deaths by state of occurrence and by multiple cause of death that would be comparable with the pneumonia/influenza/coronavirus definition produced for 2020, we multiplied the proportion of deaths due to P&I from CDC FluView with the number of all-cause deaths reported by NCHS.

Data on morbidity and circulation of other respiratory pathogens:

Weekly state-level ILI data were obtained from the CDC's ILINet system[10], which aggregates data from a network of outpatient providers. To adjust for activity of non-SARS-CoV-2 respiratory pathogens, we used state-level data on laboratory-confirmed influenza activity from the CDC's National Respiratory and Enteric Virus Surveillance System (NREVSS)[11]. This dataset captures the number of tests performed for influenza and the number that were positive by week and state. The ILI data provide the percent of visits to participating outpatient providers that were for ILI. ILINet and NREVSS data are available with a 1-week lag.

The ILI, NREVSS, and P&I mortality datasets were accessed through the CDC's FluView portal using the cdcfluview package in R. Data from NCHS, ILINet and NREVSS were obtained for the weeks ending January 5, 2013 through May 30, 2020.

Comparison with changes in influenza-like illness

To get a measurement of COVID-19 epidemic intensity in the outpatient setting, the same model was fit to ILI data. Time series for excess ILI were compared with times series for excess all-cause deaths.

Adjustments to the influenza confirmed cases time series

In our main analyses, it was desirable to adjust for influenza activity when estimating the seasonal baseline. There are two reasons for this. First, failure to adjust for influenza epidemics that were still ongoing in the Spring would lead to over-attributing excess deaths to influenza. Second, influenza epidemics in previous years could bias the seasonal baseline upwards. However, starting in March 2020 the number of influenza cases and hospitalizations in the US dropped to historically low levels. Because of this drop, including influenza activity in the model could bias the baseline downward and lead to an over-estimate of excess deaths. To avoid this issue, we quantified increases in influenza activity above a seasonal baseline. Briefly, the percent of specimens positive for influenza were log transformed, we used a "Serfling" regression approach to set a seasonal baseline. Data for flu season (December-February) were set to missing, and we fit a linear regression to the remaining data, with harmonics with periods of 52 and 26 weeks. We then subtracting the fitted harmonic baseline from the observed data, and any negative values were set to 0. This provides a time series of increases of influenza above the seasonal baseline while ignoring recent decline in influenza below the baseline. This time series was used as a covariate in the main analysis.

Statistical model

We fit Poisson regression models to the weekly state-level death counts from January 5, 2015 to January 25, 2020. The baseline was then projected forward until May 30, 2020; excess mortality

was estimated as observed minus baseline deaths. Models were fit separately for each state. We adjusted for seasonality, year-to-year baseline variation, and influenza activity in the previous week (details above). The 1-week lag between the influenza data and the mortality data accounts for the delay between the time when the influenza test is performed and death. Influenza data for Florida were not reported, so we used influenza data for the other states in region 4 (southeastern US) instead.

We regress all-cause deaths and PIC deaths in epidemiological year *i* (July-June) and week *t,* using the equation below (shown for all-cause deaths). The PIC model was the same, except that all-cause deaths were used as an offset-term, and there was no adjustment for Proportion Complete.

Let $\text{Total\_Deaths}_{i,t}$ be the number of deaths and let $\text{Flu\_Epidemic}_{i,t-1}$ be the time series of influenza epidemic activity (see above), and $\text{Proportion\_Complete}_{i,t}$ the estimated proportion of deaths that have been reported for that state and week (see next section) We modeled

$$\text{Total\_Deaths}_{i,t} \sim \text{Poisson}(\boldsymbol{\lambda}_{i,t,}, \phi)$$

where

$$\log(\boldsymbol{\lambda}_{i,t}) = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1*\sin(\Theta_t) + \boldsymbol{\beta}_2*\cos(\Theta_t) + \boldsymbol{\beta}_3*\sin(\Theta_t/2) + \boldsymbol{\beta}_4*\cos(\Theta_t/2) + \\ + \boldsymbol{\beta}_6*\log(\text{Flu\_Epidemic}_{i,t-1}) + \gamma_i + \alpha_i*\log(\text{Flu\_Epidemic}_{i,t-1}) + \log(\text{Proportion\_Complete}_{i,t})$$

and

$$\Theta_t = 2*\boldsymbol{\pi}*t/52.1775$$

To compute prediction intervals, we used the following procedure. Once the regression coefficients were estimated, we extracted the estimated asymptotic covariance matrix for the parameters and constructed a multivariate normal distribution approximating the sampling distribution, centered at the estimated parameter values. We drew 100 samples from this parameter distribution, and for each sample computed the resulting mean value $\boldsymbol{\lambda}_{i,t}$, and then drew 100 samples from the Poisson distribution with this mean. This resulted in 10,000 samples from an empirical predictive distribution of $\text{Total\_Deaths}_{i,t}$. Empirical 95% prediction intervals were computed by taking the 2.5th and 97.5th percentiles of this resulting distribution. In sensitivity analyses, we evaluated a model in which Flu_Epidemic was excluded altogether. Using a Poisson model might under-estimate the variance for some of the larger states, resulting in prediction intervals that are too narrow. The dispersion parameter for a negative binomial model was not identifiable for many of the smaller states; therefore, we used the Poisson model for all states for consistency.

Reporting delays analysis

We used an empirical and model-based approach to estimate the reporting delays.

Empirical approach: First, to determine whether reporting delays changed between 2019 and 2020 outside of the period of intense pandemic activity, we compared the provisional death reports filed in week 11 in 2019 and 2020 and found that they did not diverge (**Figure S1A**). This demonstrates that the reporting delays did not change in 2020. We then compared the provisional data reported for weeks 10-19 in week 20 of 2020 in the CDC FluView report with the data reported for the same period in the week 20 of 2019. The area between these two curves is an empirical estimate of excess deaths but is not adjusted for year-year variations in deaths. In

January and February 2020, there were an average of 829 more deaths in 2020 compared to 2019, representing increases not related to COVID-19. We summed the area between the 2020 and 2019 curves for weeks 10-19 and subtracted 829*(10 weeks). This provides an empirical estimate of excess deaths, adjusted for reporting delays and year-over-year increases.

Model-based approach: We used the state-level FluView reports for weeks 11-20 in 2020 to construct reporting triangles for each state. Archived versions of the FluView data were not publicly available for earlier weeks. We used the nowcasting with Bayesian smoothing method described by McGough et al (*PLOS Comp Biol 2020)*, using a negative binomial model. The only modification was that the reporting triangle was not complete because we did not fully observe the data. The earliest week of data we considered was 2 weeks after death. Any deaths reported in the first 2 weeks were combined into a single category. For dates of death where the first N weeks were not observed the weekly reporting probabilities for weeks 2:N were combined to estimate the reporting fraction of the observed count. The modified JAGS code is here: https://github.com/weinbergerlab/excess_pi_covid/blob/master/functions/jags_negbin4.R. The proportion reported by each week after death was estimated from the model. The proportion reported each week was combined to get a cumulative estimate of the proportion of deaths that were reported by a particular week. The median of 10,000 draws from the posterior distribution for each delay week was estimated. This estimate of proportion complete was used as a denominator in the main analysis. For earlier weeks with complete data, this value is 1 or close to 1. For more recent weeks, this value is smaller (indicating incomplete data), and the baseline of expected reported deaths is adjusted down accordingly. The smoothed nowcast estimates of deaths from the NobBS model were not directly used in our analyses. Sensitivity analyses using these nowcasted estimates in place of the observed counts found national estimates of excess deaths that were approximately 4% higher than the estimates presented here.