

Article

A Pandemic since When?

Gal Almogy

Flurensics Inc., CTO, Levy Eshkol 16, Tel Aviv, Israel; dr.galmogy@gmail.com

Abstract: late in December 2019 2019-nCoV was identified as the pathogen responsible for an outbreak of severe respiratory distress in Wuhan, China. The virus was detected in multiple countries during January, but it is believed widespread community transmission began late in February or early March. Since March the virus has caused over 100k confirmed deaths in the US, with some states more severely impacted, notably NY and NJ. Here I examine excess mortality at the national and state level from January through July 2020. I find that the increase in excess mortality began in late February, suggesting the pathogen was circulating undetected earlier than assumed. The timing and intensity of the increase in excess mortality varied across states, with two patterns emerging: an early, sharp increase reaching a peak during April-May, best exemplified by NY and NJ, and a shallower, sustained increase, reaching a peak in late July, observed mostly in the southern regions of the US.

Keywords: 2019-nCoV; COVID-19; excess mortality; all-cause deaths; case fatality ratio; CFR; early transmission; SARS

Introduction

An outbreak of pneumonia cases of unknown etymology late in 2019, in Wuhan, China, led to the identification of a novel viral pathogen belonging to the coronavirus family, 2019-nCoV. The virus was identified using metagenomic analysis of specimens from one of the first patients, and subsequently confirmed in other Wuhan patients (1-2).

The first imported covid case in the US was reported on the 21st of January, and the first local transmissions in the US were confirmed on Feb 26 in California and Feb 28th in Washington (3). The increasing incidence of covid infections in the following weeks led to the declaration of a national emergency on March 13 and the application of various prevention measures designed to curtail the spread of the virus.

One of the open questions regarding the COVID-19 pandemic is the CFR/IFR of 2019-nCoV (4-5). Initial reports from China found the CFR to be ~4% (6-7), but since then the per region CFR showed large variability, from <1% to 10% and even more in some regions, e.g. NYC, Madrid, Milano etc (8-13).

An indirect way to quantify the fatality associated with the pandemic is through examination of excess deaths that occurred in a region. Studies of excess deaths may lack the direct cause-and-effect that is assumed in standard case counts and cause-of-death reports, but considering reporting inaccuracy, and that not every person that dies from COVID-19 is identified as such, excess death data are useful in providing a more comprehensive estimate of the full burden of the 2020 pandemic (14).

Here I examine per-week, all-cause mortality data for the US, spanning Jan 2015 to August 2020, and show that:

1. Excess mortality was **observed in all age-groups** older than 25yo, and not limited to elderly population at any point of the pandemic
2. The timing of the excess mortality periods varied across states, but can be qualitatively divided into **single- and non-single-peak patterns**

3. The dramatic rise in the total number of deaths in NY and NJ that began in mid-March was preceded by a steady increase in excess deaths beginning as early as January in several states, including CA, FL, TX, WA, and WI.

These results are consistent with extensive undetected transmission across the US before/during February, reflected in the steady increase in weekly deaths in almost all states in the Feb-Mar period. Interestingly, there is a significant linear trend of increase in deaths in many states, which suggests an additional contributing factor beyond the spread of 2019-nCoV.

Methods

Per week mortality data for the US (January 2015 to Jul 25, 2020) was retrieved from the CDC website on Sep 17 (15). Data contains a breakdown of deaths according to 6 age groups: under 25yo, 25-44, 45-64, 65-75, 75-85 and people over 85yo. The under 25yo group accounted for very few deaths of the total in all states, and was ignored.

Baseline mortality for 2020 was calculated as the per week average from Jan 2015 to Dec 2019, and was not adjusted for seasonality or population growth. Hence the baseline mortality for state S, during week W is:

$$\text{Baseline}(S,W) = \text{Average number of deaths during week } W \text{ in state } S, \text{ from 2015 to 2019}$$

The difference in weekly mortality compared to baseline was calculated as the number of deaths per week minus the baseline. "Excess death" during week W in state S was defined using the condition:

$$\text{Excess deaths (S,W) IF } \text{deaths}(S,W) - \text{baseline (S,W)} - 2 \times \text{STD (S,W)} > 0$$

where $\text{STD}(S,W)$ = standard deviation of deaths in state S during week W between 2019 and 2019

The ratio of deaths per week, W, to the weekly average in state S was defined as:

$$\text{Death ratio}(S,W) = \text{Deaths}(S,W) / \text{Baseline}(S,W)$$

Death ratios greater than 1 correspond to weeks in which the number of deaths exceeded the 5-year baseline, whereas values < 1 represent weeks with less death than the average.

Results

Examining crude US mortality data reveals a highly unusual mortality pattern in the first half of 2020 when compared to previous years (figure 1a). Overall mortality in the first weeks of 2020 was average but by week 8 reached higher levels than in any of the previous 5 years, and kept increasing during March, reaching a very pronounced peak in mid-April, and remaining at elevated level through July 25. The increase in mortality was not limited to older demographics, and is clearly evident in the 25-44yp age group when compared to previous years, showing overall similar trend to mortality in the older populations (figure 1b).

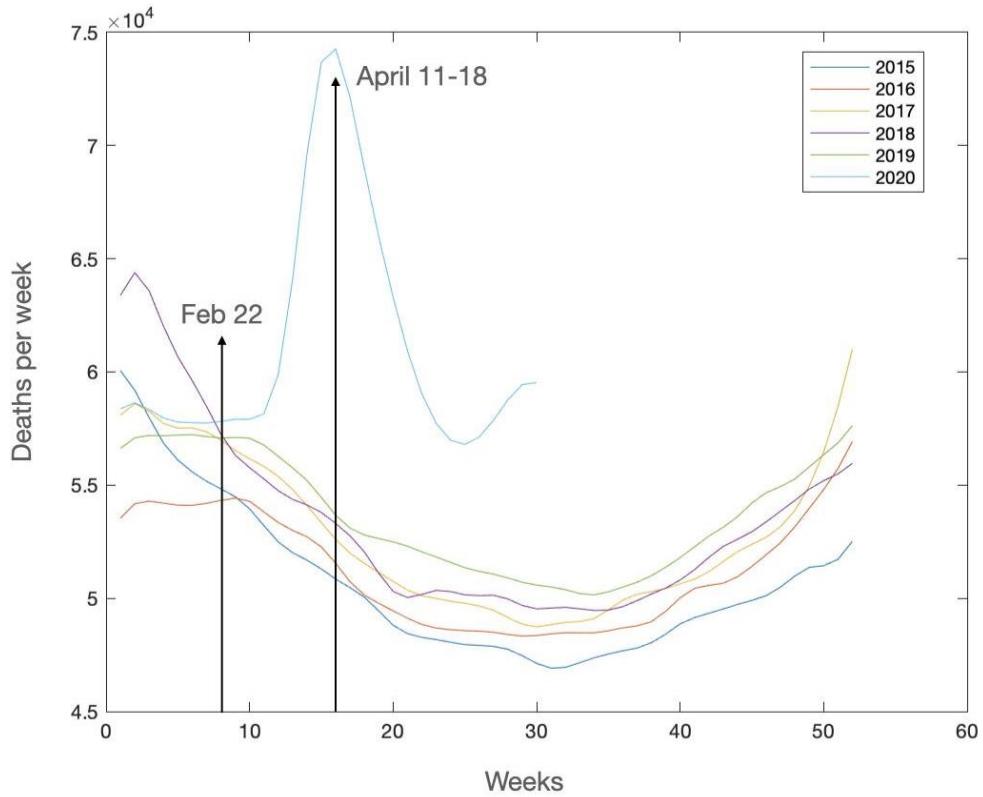


Figure 1a. Legend: crude death/week for the entire US, ages 25+, 2015 (dark blue) to first 30 weeks of 2020 (light blue). Number of deaths exceeded previous years from week 8 (February 22), remaining elevated through July 25.

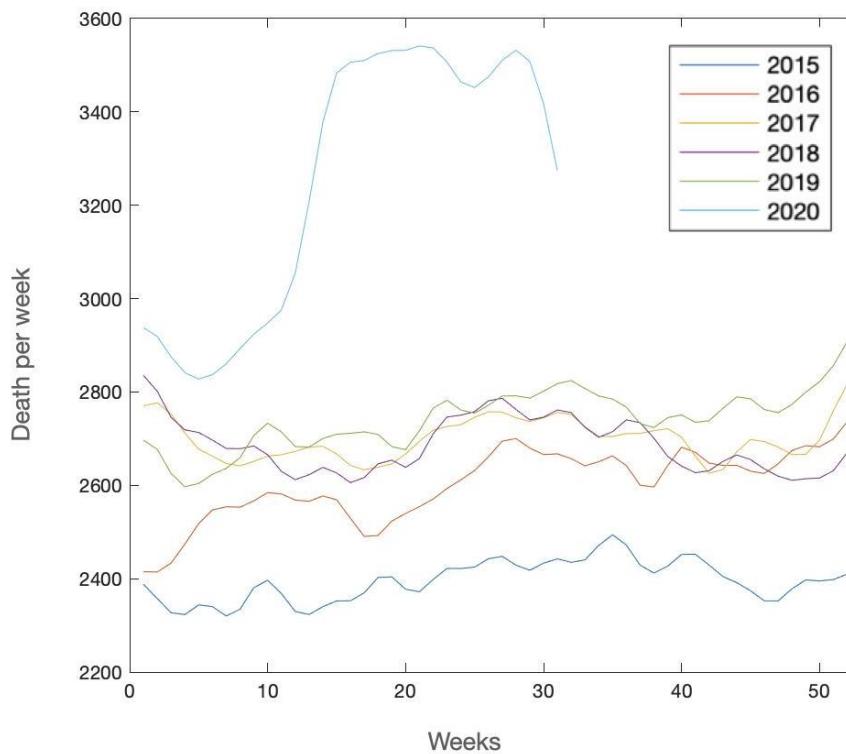
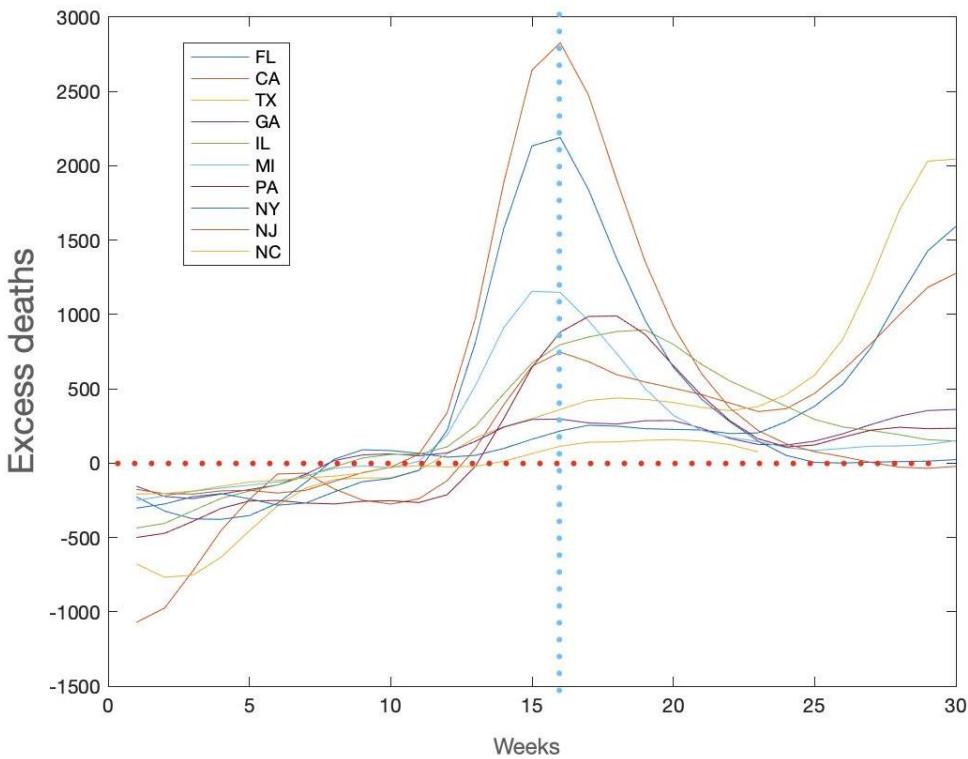


Figure 1. b. Legend: crude death/week for the entire US, ages 25 to 44, 2015 (dark blue) to first 30 weeks of 2020 (light blue).

Excess mortality at the state level

The first states to show excess mortality were FL and GA beginning week 8, followed by IL (week 9), MI and NJ (week 11), NY (week 12), TX and CA (week 13), and finally PA and NC on week 14 (Figure 2a).

Figure 2a. Legend: Excess per-week deaths in most populated states, week1 to week 30, 2020. Horizontal dashed red line indicates 5-year baseline.



The first states to reach a peak in weekly number of deaths were NJ, NY and MI (week 15), followed by PA and IL (week 16). Interestingly, following the April peak excess, mortality went down in all these states, but except for NYC, NY, and NJ remained elevated, and resumed the upward trend in CA, TX, FL, and to a lesser extent in GA and PA from June.

When expanding the analysis to smaller states (Figure 2b), it is clear that most states follow one of two patterns: a single peak around week 16 (dashed blue line), followed by steady decline in the number of deaths, as in NY and NJ (termed here “the NY-like” pattern), is almost perfectly replicated, with slight delays, in neighboring states like MA, CT, and MD, and a second pattern, more similar to FL and TX, entering the excess mortality period early (week 9), but with more pronounced increase in mortality beginning around June, as in AZ, AR, AL, and SC.

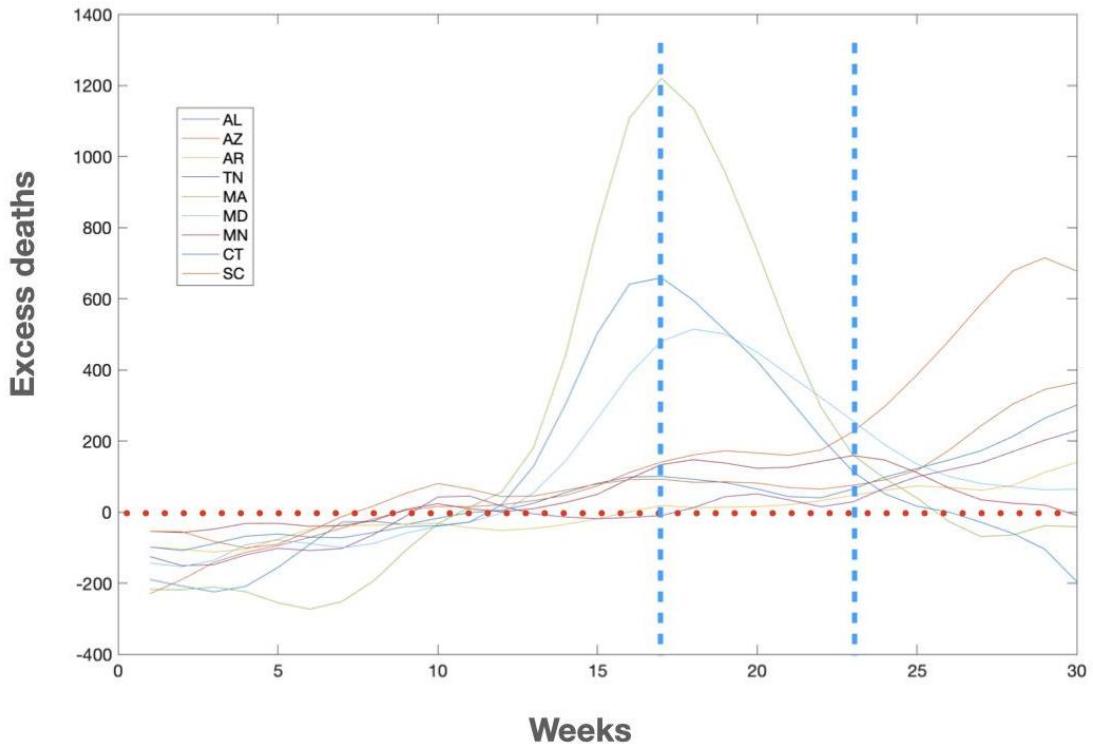


Figure 2b. Legend: Excess per-week deaths, week1 to week 30, 2020. Horizontal dashed red line indicates 5-year baseline.

Per week death ratios at the national and regional levels

The ratio of per-week deaths in 2020 to the 2015-2019 baseline (see methods) offers a convenient way to examine the normalized per week changes in mortality across multiple states.

From mid-January to March 2020 there was a moderate, yet steady increase in the death ratio in the US as a whole, followed by a sharper increase, especially in the NY-like states from week 12 (see figure 3a, red and yellow curves). Death ratio in the US when ignoring the NY-like states began from a higher starting point, and shows a more moderate increase without the massive peak observed in the NY area, but increased continuously through July (figure 3a, blue curve).

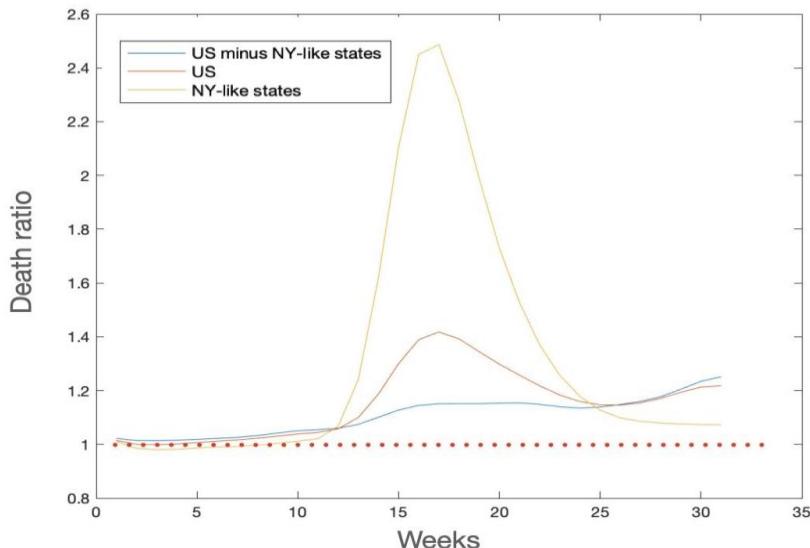


Figure 3a. Legend: death ratio in the US, the NY-like states, and the US minus the NY-like states (NY, NJ, MI, PA, MA). Horizontal dashed red line indicates 5-year baseline.

When the states are divided by geographical location into NE, NEC, and southern US, it is clear that the main increase in mortality, centered around week 16 in the NE area, occurs a few weeks later in the NEC region, and approx 2 months later (mid-June) in the most southern states TX and FL, as well as in CA and GA (figure 3b-d).

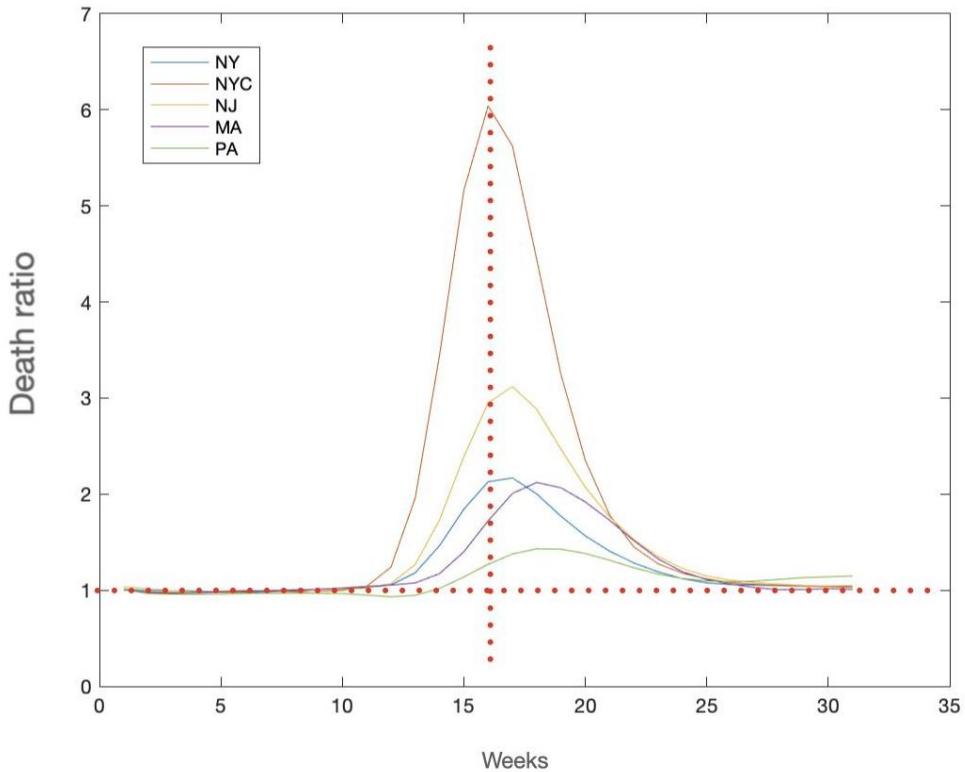


Figure 3b. Legend: death ratio is NE states. Horizontal dashed red line indicates 5-year baseline.

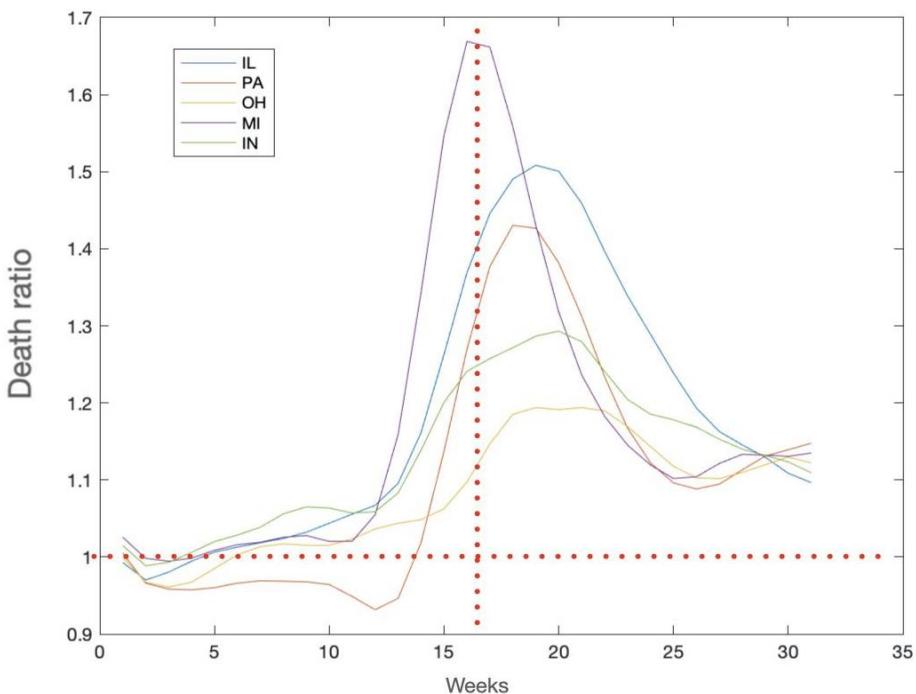


Figure 3c. Legend: death ratio is NEC states. Horizontal dashed red line indicates 5-year baseline.

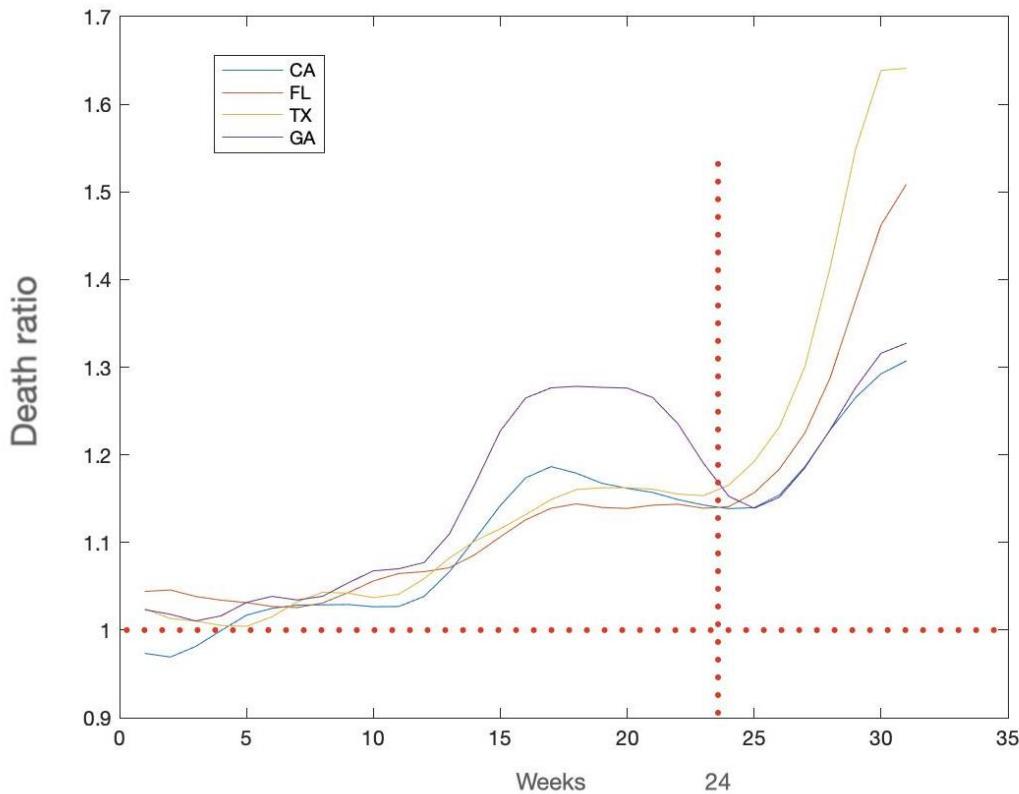


Figure 3. d. Legend: death ratio is southern states and CA. Horizontal dashed red line indicates 5-year baseline.

There are exceptions to these regional patterns, e.g. LA in the south follows a very similar pattern to the NE states, while the death ratio in OR peaks later (week 17), and continues increasing through July like in the southern states. Interestingly, the pattern in CA and GA appears to be a combination of the patterns observed in the NE/NEC states and the southern states, rising slightly more sharply in April, but sharing the same increase trend as in the southern states from mid-June.

Discussion

Longitudinal mortality data provides important context when studying the impact of the covid pandemic in the US. Here I show that the trend of increased mortality affected all age groups older than 25yo, and began before March 2020 in most large states, leading to two general patterns of excess mortality in the following months.

Mortality dynamics in MI, NY, and NJ (figure 3b) provide the clearest example of the single-peak pattern, with weekly death numbers rising dramatically from mid-March, reaching a peak within 3-4 weeks, and declining to roughly pre-pandemic levels by June. The majority of large states in the NE and NEC regions of the US mimic this single-peak pattern, but differ from MI, NY, and NJ in that the peak is somewhat delayed, less pronounced, and in that mortality rates in these states remained elevated through July (figure 3c).

On the other end of the spectrum are FL and TX (figure 3d), where the increase in deaths, compared to a 2015-2019 baseline, increased more or less linearly from late February to early June, and at a much higher rate from week 24 through July. Most southern states mimic the late increase in deaths, some skipping the first peak in April-May, e.g. AL and AK (not shown), and other states showing a mix of the two patterns, like GA and CA (figure 3d), with a first peak between weeks 15 and 19, followed by a sharp increase in deaths from June through July. Given recent mortality data may still be incomplete, July 25 was used as the cutoff in the above analysis. When extending the analysis to Aug 15 it seems that excess mortality in the southern states had reached its peak at the

end of July and is trending to baseline levels, and no increase in excess mortality is apparent in other states (figure 4).

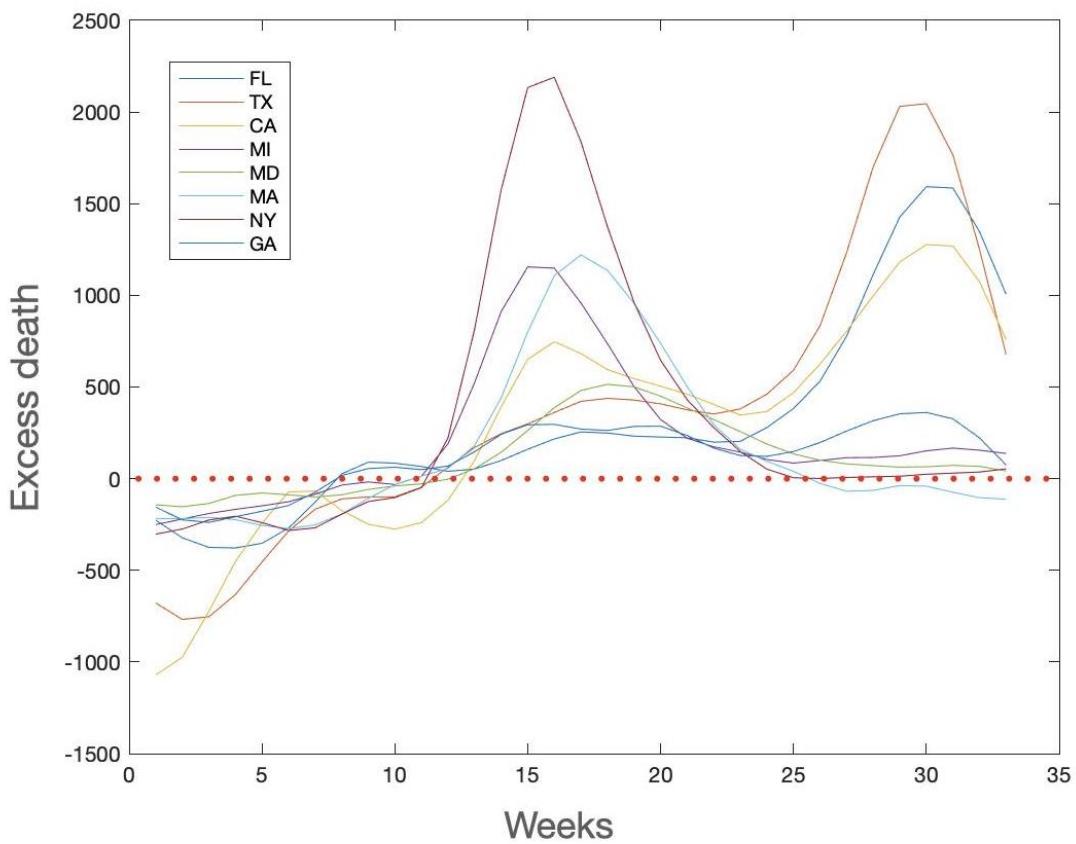


Figure 4. Legend: excess deaths in 8 states, week 1 to week 33 of 2020. Dashed red line indicates 5-year baseline.

In conclusion, the increase in excess mortality in some states, e.g FL and TX began around February, several weeks earlier than expected considering very little 2019-nCoV transmission was documented during that period. In the following months states displayed one of two patterns: an early, pronounced peak around April-May followed by a return to near baseline levels by June, as in NY and NJ, and a slower, but more extended increase in excess mortality observed mainly in southern states, which currently seems to trend towards baseline mortality.

It would be of interest to see how these patterns correlate with COVID-19 incidence in those states.

References

- 1 Wu, Fan, et al. "A new coronavirus associated with human respiratory disease in China." *Nature* 579.7798 (2020): 265-269.
- 2 Zhou, Peng, et al. "A pneumonia outbreak associated with a new coronavirus of probable bat origin." *nature* 579.7798 (2020): 270-273.
- 3 Jorden, Michelle A., et al. "Evidence for limited early spread of COVID-19 within the United States, January–February 2020." (2020).
- 4 Rockx, Barry, et al. "Comparative pathogenesis of COVID-19, MERS, and SARS in a nonhuman primate model." *Science* 368.6494 (2020): 1012-1015.
- 5 Estimating case fatality rates of COVID-19 - Piotr Spycharski, Agata Błażyńska-Spycharski, Jarek Kobiela, *The Lancet Infectious Diseases* - 2020

6 Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395(10223):497-506. doi:10.1016/S0140-6736(20)30183-5

7 Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020;395(10223):507-513. doi:10.1016/S0140-6736(20)30211-7

8 Sood, Neeraj, et al. "Seroprevalence of SARS-CoV-2-Specific Antibodies Among Adults in Los Angeles County, California, on April 10-11, 2020." *Jama* (2020).

9 Ioannidis, John. "The infection fatality rate of COVID-19 inferred from seroprevalence data." *medRxiv* (2020).

10 Erikstrup, Christian, et al. "Estimation of SARS-CoV-2 infection fatality rate by real-time antibody screening of blood donors." *medRxiv* (2020).

11 De Natale, Giuseppe, et al. "The COVID-19 infection in Italy: a statistical study of an abnormally severe disease." *Journal of Clinical Medicine* 9.5 (2020): 1564.

12 Covid, C. D. C., and Response Team. "Severe outcomes among patients with coronavirus disease 2019 (COVID-19)—United States, February 12–March 16, 2020." *MMWR Morb Mortal Wkly Rep* 69.12 (2020): 343-346.

13 Rossi, Paolo Giorgi, et al. "Survival of hospitalized COVID-19 patients in Northern Italy: a population-based cohort study by the ITA-COVID19 Network." *medRxiv* (2020).

14 Weinberger, Daniel M., et al. "Estimation of excess deaths associated with the COVID-19 pandemic in the United States, March to May 2020." *JAMA Internal Medicine* (2020).

15 https://www.cdc.gov/nchs/nvss/vsrr/covid19/excess_deaths.htm (retrieved Sep 17).



© 2020 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).