COVID-19: Re-Visiting the Numbers in Arizona, a Week Later

More Cases. Lower Mortality Rates. In our prior analysis of Arizona's publicly available COVID-19 data we concluded (1) case counts were a poor predictor of both hospitalizations and deaths for the period April 1 through June 30 (R^2 or 0.01 and <0.01, respectively), (2) case mortality was negatively correlated with case count ($R^2 = 0.42$)—i.e., falling case mortality rates were likely to follow rising case counts, and (3) COVID-19 mortality is highly specific—people over 65 years of age are 14% of the Arizona's population, but represented more than 75% of COVID-19 Deaths. After correcting a data error and collecting updated data this analysis concludes there is a correlation between case counts and hospitalizations and mortality, and that the relationship is complicated. Hence, hospitalizations and case mortality can fall amid rising case counts.

Data were collected July 17, 2020. Mindful of delays in reporting, this analysis covers the period April 1 through July 7, 2020.

Transparency First. For new readers, our prior briefing was used in a press release that questioned the state's COVID-19 data. Our briefing did not question the state's data; to the contrary, our analyses relied on those data. We questioned the implied link between case counts, hospitalizations, and deaths in addition to the use of case counts as a basis for major public policy decisions. We believe that real-time public health data are inherently problematic as they are subject to later revision and supplementation. Records are reviewed and corrected. Corrections are not a sign of dishonesty.

The Arizona legislator that issued the press release was Mark Finchem, brother of our founder. Hearing about our analysis of Tennessee experience, he asked whether we could repeat the analyses for Arizona. We did, out of curiosity, and share it with him. We received no compensation of any kind for the work. We were not hired, commissioned, or otherwise compensated for the work.

In response to the press release, the director of the Arizona Public Health Association opined that our study was flawed because "the person that did the analysis did not recognize that there is a delay in the data…"

Operating under the common assumption that current public information is the most accurate data source, we expected to revise our analysis as more comprehensive data became available, as other institutions including the state itself should do.

Our opening graphic in that summary was copied directly from the Arizona Department of Health Services website's "Data Dashboard" which included their footnote: *"Illnesses in the last 4-7 days may not yet be reported."* We collected those data July 9. Our analyses were for the period April 1 through June 30. We were aware of the potential for delayed reporting and took at face value the ADHS's warning about the time window of delays in reporting.

Not noticed in the conversation that followed was the fact that our mortality data were just plain wrong. Wrong, like the experts' forecast of 2.2 million COVID deaths in the US and New York's need for 140,000 hospital beds. We know mistakes happen; they must be admitted honestly and corrected. We apologize for our mistake. This briefing seeks to set the record and conversation straight.

We have launched a Public Interest initiative to develop an agenda-free understanding the data are and mean for this and other important questions. This effort is resource intensive, in terms of data, tools, and people. We began a GoFundMe campaign to help us sustain this effort. To help us, visit (https://www.gofundme.com/f/just-the-numbers).

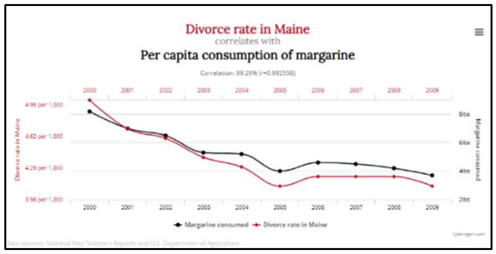
Association and Cause. Early in most statistics classes, the teacher explains, association is not necessarily causation, but causation without association doesn't happen. Sometimes finding the association is hard. The divorce rate in Maine is *associated* with per capita consumption of margarine, in the sense that they are correlated. The curves seem to mirror one another. The correlation coefficient (R^2 is nearly equal to 1.0). This association does not prove that one is a cause of the other. Establishing causation requires experiments where random application of the potential cause affects the result in question.

We use statistics, and linear regression in particular to measure the potential association between factors. A lack of association can be important.

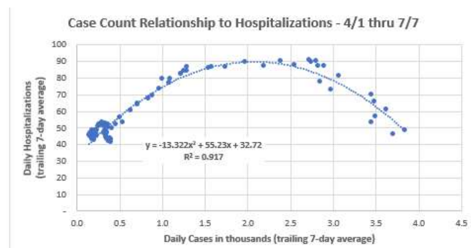
New Data & Corrections. Data were collected July 17, 2020. An error in prior treatment of case mortality was identified and corrected. The error understated the number of deaths during the study period—April 1 through June 30. As a result, it also understated the correlation between case count and case mortality.

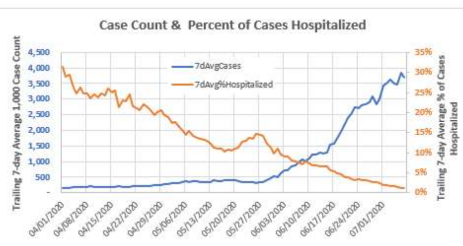
The present analyses consider the period April 1 thru July 7—a cut-off 10 days prior to the data collection date.

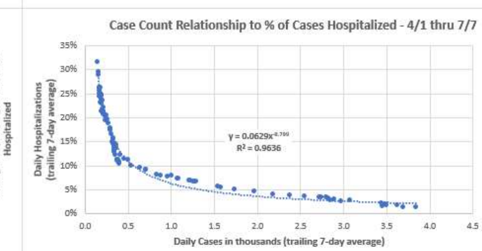
Correcting for the prior mortality under-count, and re-visiting the questions, *Is Case Count important? Does it predict important events—e.g., Hospitalizations? Deaths?* the answer is "maybe."



Cases & Hospitalization. The 7-day average of confirmed cases rose, beginning late May and continued to increase through June. The correlation coefficient (R^2) was 0.92 for April 1 thru July 7, but the relationship was not linear. It was a more complex (polynomial) relationship.







Counterintuitively, the maximum of hospitalizations—about 90 per day—occurred between 2,000 and 2,500 daily cases (7-day trailing average basis). *Hospitalizations did not consistently increase with rising case counts.* In fact, hospitalization began to slow amid rising case counts in mid-June. This could be because more recent cases included a relatively large number of less severe cases. Average daily case count also appears to have a complex relationship to the percent of cases requiring hospitalization.

Notwithstanding the high, complex correlation between case count and hospitalizations, case count is not a strong predictor of hospitalizations. Dramatic increases in case count will not necessarily translate to dramatically more hospitalizations.

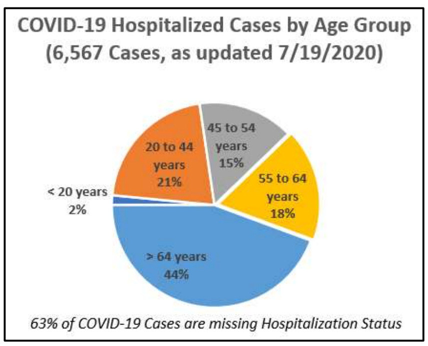
Capacity Shortages. Much has been written about hospitalizations in the context of available beds, IC Units, and ventilators. Unused capacity varies daily, and it is fair to say the state—on a combined basis—was near 90% of available capacity on some days during the last month. This can happen for several non-mutually exclusive reasons:

- Wide-spread disease, requiring considerable in-hospital care, and/or
- Insufficient hospital capacity.

It is worth noting that Arizona is among the bottom 20% of US states, in terms of hospital beds per population. According to a 2018 Kaiser Family Foundation study (see Kaiser). Arizona has 1.9 beds per 1,000 population compared to a

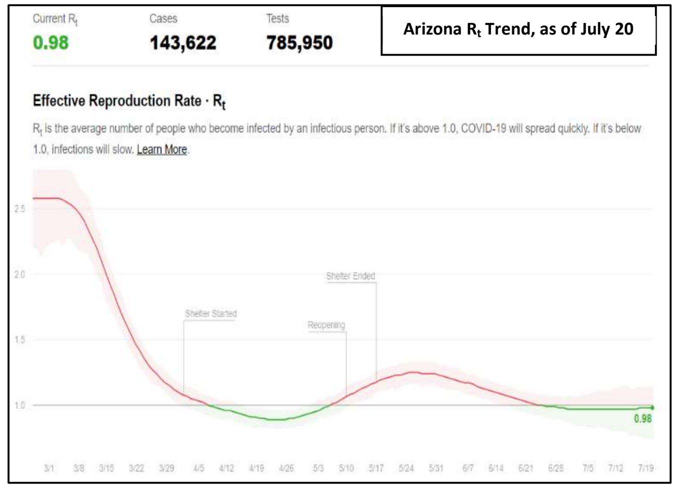
US national average of 2.4 beds. In March 2020, AZCentral.com reported about 16,000 licensed in-patient hospital beds in Arizona. Increasing to the US national average would provide an additional 3,650 beds, a 23% increase in capacity.

Statewide daily hospitalizations peaked June 15, and have generally followed a downward trend. As of July 19, 44% of hospital cases were patients age 65 or older (14% of Arizona's population).



The AZDHS COVID-like-illness surveillance dashboard, which is a non-comprehensive summary of emergency department (ED) and in-patient activity, reports cases peaked, as a percent of ED visits at 15.5% on June 28 and as a percent of in-patient "visits" at 20.9% on the same day. The dashboard also reports the values have fallen to 8.7% of emergency department visits and 12.8% or in-patient visits, thru of July 12, as of the July 19 data. AZDHS reported that 5% of the state's COVID cases were hospitalized, as of July 19.

Measuring the Epidemic. R_t is a factor used in epidemiology to measure (and compare) the spread of a disease. It represents the average number of people an infected person has infected. R_t of more than 1.0 means the disease will spread—i.e., case count will increase. R_t less than 1.0 means spread will slow. Current R_t values for each state can be seen at R_t COVID-19 (https://rt.live).



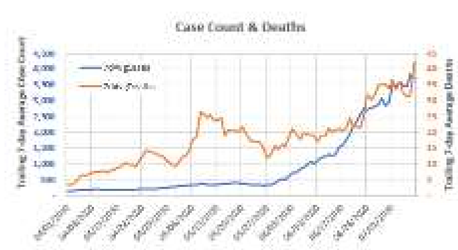
The site's graphics include interesting features: significant dates—start/stop for shelter orders, re-opening dates, uncertainty bands—are marked as shaded areas around the curves. In the case of Arizona, R_t was declining *before* sheltering started and was rising *before* re-opening; R_t has been 1.00 or less since June 22. It is currently stable at about 0.98

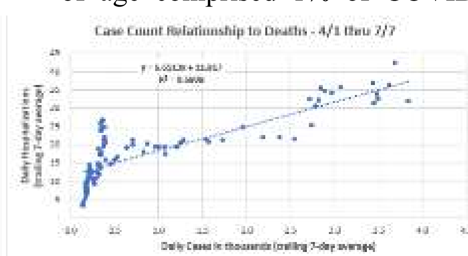
Cases & Mortality. Contrary to our prior analyses, *the correlation between case count (on a trailing 7-day average basis) and the trailing 7-day average of deaths is strong ($R^2 = 0.69$)*, with about 6.7 deaths per 1,000 new cases to be expected.

Interestingly, the correlation between case count and case mortality rate is also strong ($R^2 = 0.58$), but *negative* (i.e., as case count rises, a smaller percent of patients is likely to die). Most recently, the mortality rate has stabilized at about 1% of cases, down from a 7-day trailing peak of 7% in early May.

As before, this makes sense. Sadly, the weakest in a population are likely to succumb first during a disease outbreak; those remaining are stronger and more resistant.

Vulnerable Groups. It remains the case that COVID-19 mortality is more prevalent in older populations. People under 25 years of age comprised 1% of COVID





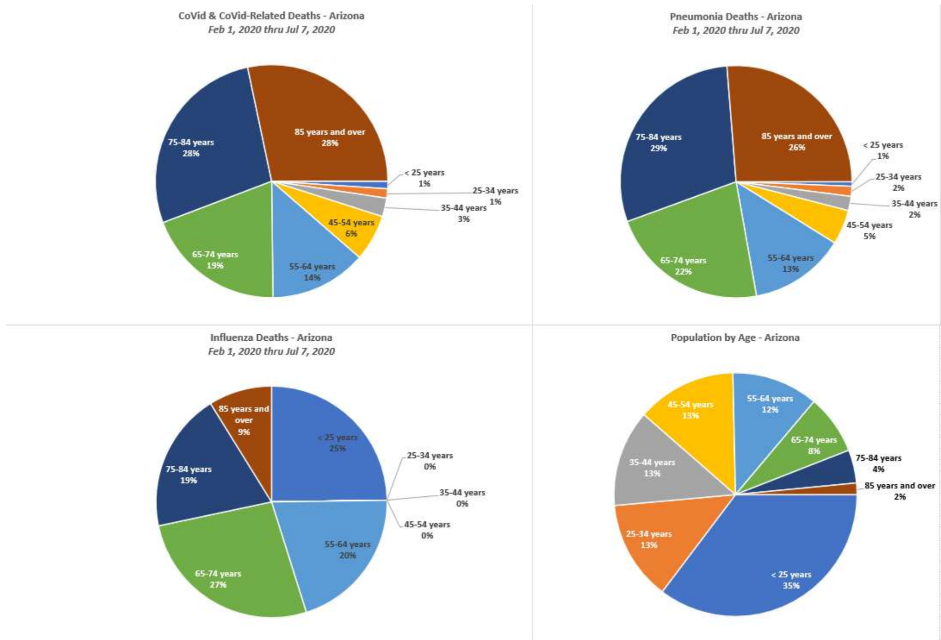
and pneumonia/COVID deaths since February 1, despite the fact the group comprises about 35% of the state's population. People 65-years and older comprised 75% of COVID and pneumonia/COVID deaths, while being 14% of the state's population. Age is an important determinant.

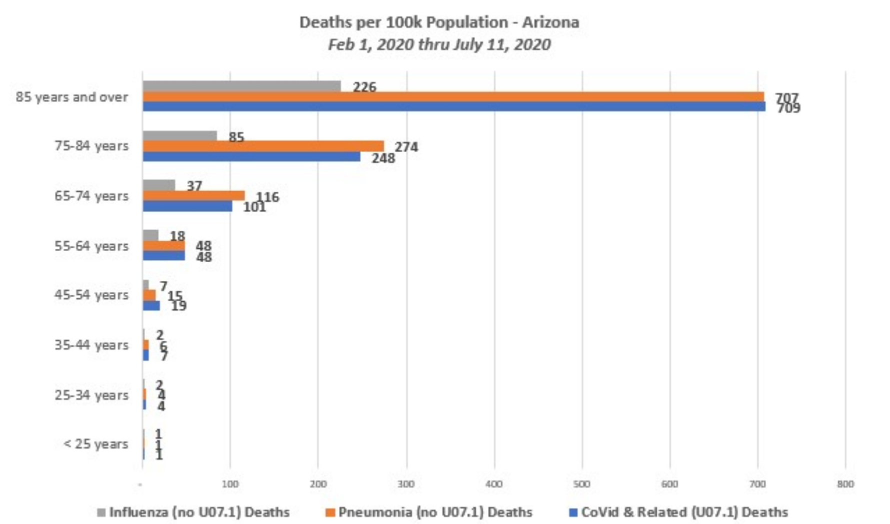
These results mimic the mortality distribution of other major respiratory diseases—influenza and pneumonia.

Better Reporting. We have discovered some authorities that are developing more helpful reporting. In our view, the best reporting should include both context and the ability to download data, so those that are most affected by policy—citizens—can develop an informed understanding. Lubbock, Texas implemented a dashboard that provides important context to case counts and death (Lubbock).

Numbers that Matter. The mortality rate (deaths per 100,000 population) of COVID-19 is like that of pneumonia if one includes COVID/pneumonia deaths in the COVID total. There has been considerable debate about whether COVID/pneumonia deaths might have been prevented, absent the presence of COVID-19 virus.

We conclude that hospitalizations, in the context of available capacity, and case mortality rates are important numbers. Simple case count, lacking age demographics, imply a random, probabilistic risk of illness and death. From a public policy perspective, it is important to know if new cases are high- or low-





risk groups, whether the cases are symptomatic or asymptomatic.

Also, case counts can increase due to more extensive testing AND a spreading disease. It is not clear that a newly discovered case offers decisive implications for an important public health (nor policy) issue.

Hospitalization and case mortality rates are both decreasing. This seems to be very good news.

Hospital bed and ICU capacity shortages can result from an epidemic growing beyond the preparedness plans and actions of policy makers. Under-investment or contagion would seem to be an important distinction.

More Data. More Context. W. Edward Deming is frequently crediting with saying, "In God we trust. All others must bring data." To that we would add, data that serves as the basis for public policy should be freely shared—in a useful format—and clearly explained.

Public health recommendations and actions should consider both health benefits and costs (e.g., reducing COVID deaths versus increasing isolation-related depression), balancing the interests of the broader public.

Public health agencies have used the language of "epidemic" for some time—epidemics of obesity, hypertension, domestic violence, substance abuse and addiction, suicide, depression, etc. Unfortunately, statistics about these "other epidemics" are published well after the fact. As an example, the latest suicide statistics available from the CDC, as of April 2020, are through the year 2018. Those data showed Arizona had the 16th largest suicide rate in the US (19.2 deaths per 100,000 population).

Delaying "elective procedures," as many states have done has left an as-yet unknown number of cancer, stroke, and heart disease cases undiagnosed and/or untreated. The future hospitalizations for these cases may increase their number and acuity due to COVID-related delays.

The impact of COVID-19 on concurrent "epidemics" would seem to be important. Oddly, it does not seem to be part of any public policy conversations. By the time data is published, policy decisions on public health matters to which the data refers will already be in place and their impacts unchangeable. Getting preliminary data sooner rather than later would better inform the public and build trust in policy decisions through strengthened data-driven context on which policy makers can justify their conclusions.

Note: This work was completed without commercial sponsorship of any kind from any source.

We established a GoFundMe site (https://gofundme.com/f/just-the-numbers) to help underwrite our effort to develop independent, politics-free analyses.

Anchor & Helm Decision Advisors helps its clients understand their data and develop actionable insights. We help clients develop business plans and budgets, better analyze data and communicate results, and implement reporting tools.

INFO@ANCHOR-HELM.COM