<u>Tri-County (Brewster, Jeff Davis, and Presidio Counties) and the rest of the World, CV19</u>

Please refer to accompanying two pdf files of graphs. The estimated number, and the related incidence relative to population, is based on maximums estimated from curve fits. These estimates are now adjusted to end of current wave, rather than end of year. They are subject to change of course, and should not be taken as precise projections, since there is no such thing. Another way to interpret them is as relative indices, that show a locale's progress relative to the others.[1]

US or Texas projections for confirmed cases scaled to Tri-County population:

max US daily	221,768	persons
equiv. max Tri-County	12.1	persons
max TX daily	33,027	persons
scaled to Tri-County	21.2	persons
max US cumulative	33,716,989	persons
equiv. max Tri-County	1839	persons
incidence	102:1000	
max TX cumulative	2,838,883	persons
scaled to Tri-County	1,825	persons
incidence	101:1000	
Tri-County (current)	94:1000	1/19/2021
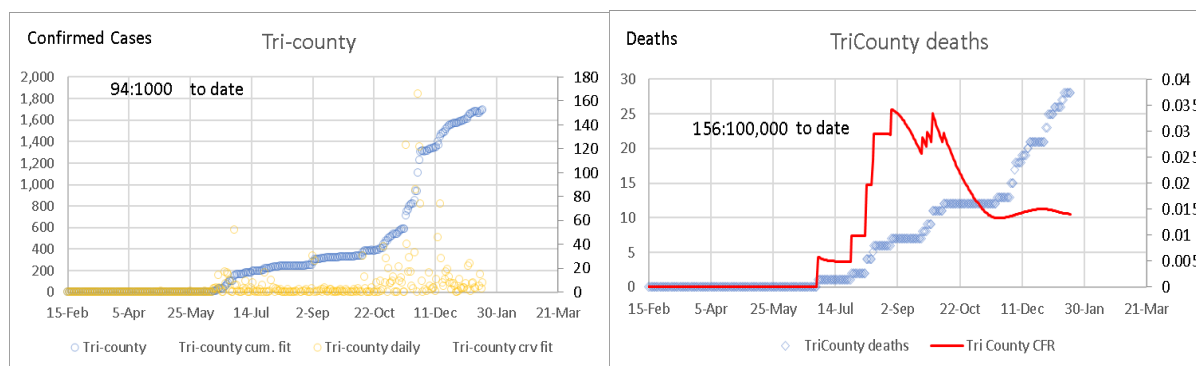


Basis of Comparison

For each curve fit, the end is extended with a gamma distribution as a rough estimate of where the last wave ends (EOW), or end of year (EOY), which ever comes first. This becomes the basis of a number per thousand or hundred thousand, in an attempt at comparison among all the locales. For example, USA:
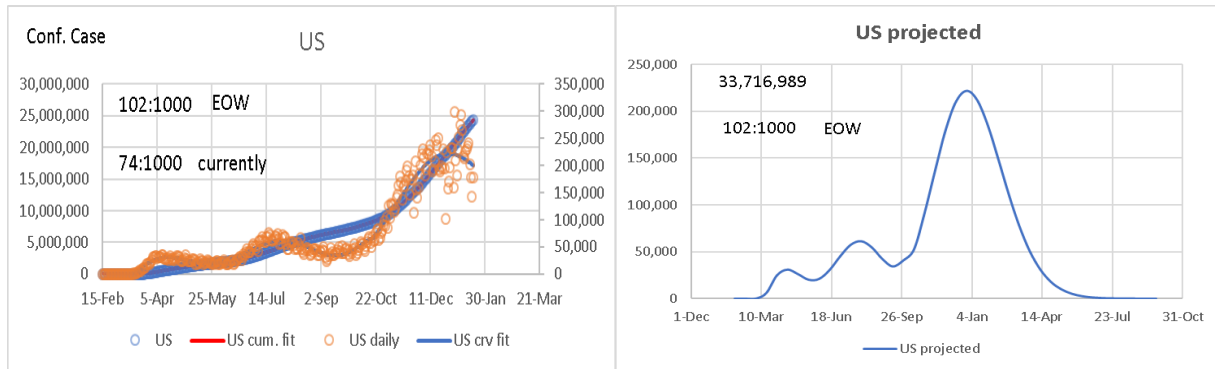

The ratio on the graph represents the total area under the curve in the right graph which is the total number of projected counts, divided by the locale's population. A peculiarity worth commenting on is the sequence of progressively higher waves; this pattern is very common in the rest of the graphs. Is this just a function of overall prevalence of the disease simply getting well established? What is also odd is the much higher EOW numbers of Western, so-called developed nations. There is probably something about the differing accuracies and methods of the counts reflected there, but sometimes I wonder if it's just the soft and slack living habits of these locales in general that is giving them so much a harder time with CV19. These locales have become too dependent on technological inventions like vaccines to survive, perhaps, and aren't paying enough attention to how they live.

Comments

If Texas and the USA confirmed cases are plotted as daily data concurrently, on a per capita basis, here is what you see, along with the apparent *third* wave in US confirmed cases:
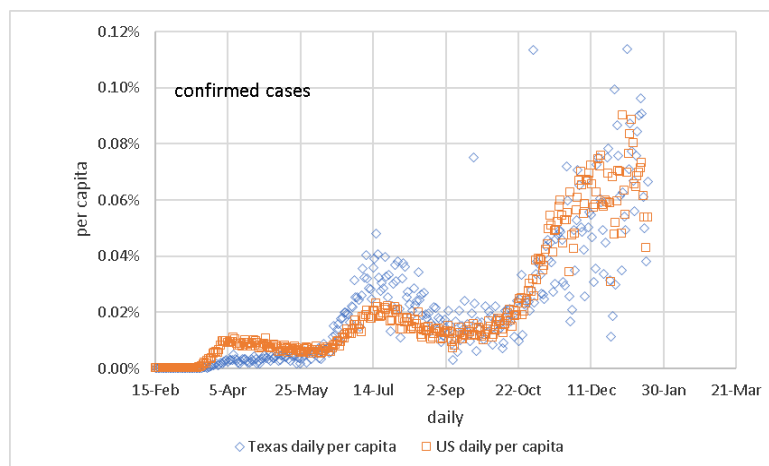
Mortality seems to have stabilized. This is an example of the confirmed case fatality ratio (CFR):

Texas CFR

$$CFR = \frac{cumul.\ fatality}{cumul.\ conf.\ case}$$
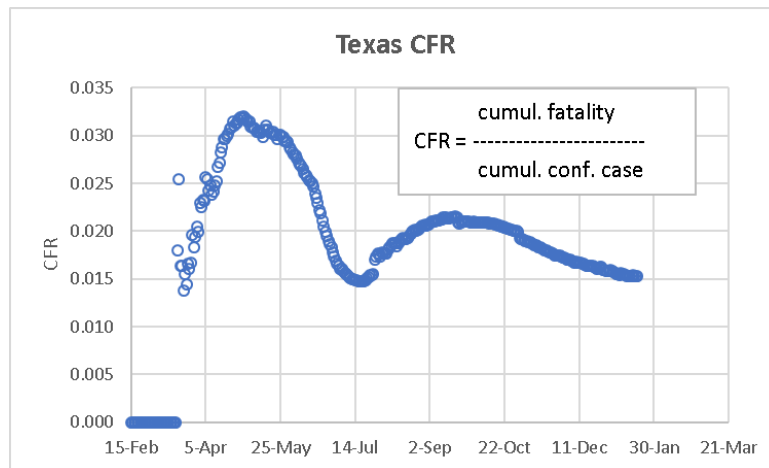
What's to be done? Your mom taught you to cover your mouth when you coughed as a act of consideration for others so you wouldn't spread germs. That is all wearing a typical cloth mask does, except more effectively than your hand (or elbow), since this virus is stronger than the usual germs.

Another good thing to do is avoid enclosed areas with low or no ventilation where there are lots of people. This would include most restaurants and bars, and probably schools, too. Or step up the air changes in such places. For example, Alpine has done some nice work on granting special use rules so restaurants can serve food outside on tables on the sidewalk. And, for better or for worse, Germany is now taking action on fresh air in buildings.[2]

A serious claim by the CDC is that 42% of the US adult population is obese, a very big comorbity with this virus. Sounds like some low hanging fruit to me; stop pushing sugary foods and drinks on the young![3]

---

[2] https://www.bbc.co.uk/news/world-europe-54599593
[3] https://www.cdc.gov/obesity/data/adult.html

COVID-19 deaths

The USA CV-19 death statistics are based on 7th order curve fit, combined with a last gamma distribution curve fit of US deaths that generates a cumulative maximum for the two combined (minus their overlap). See "*Experimental Page*" for a look at ratios of the curve fits, rather than the data itself:
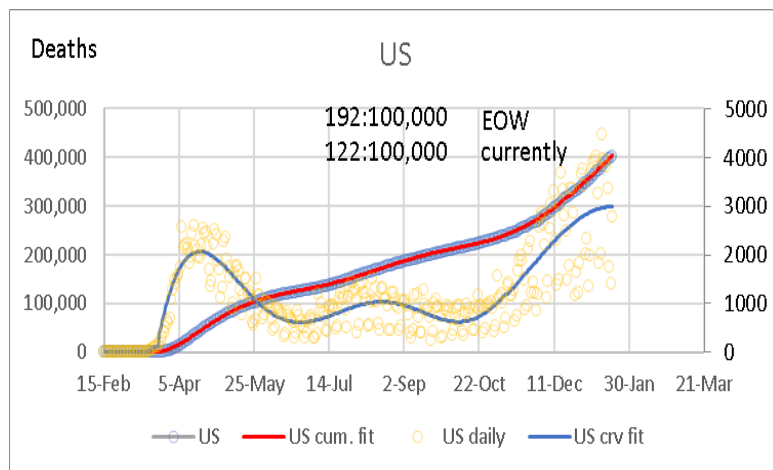
A few months ago the Wall Street Journal published data on annual causes of death which is plotted with the USA CV-19 death statistics. The possible death value is plotted as deaths per 100,000 with WSJ data of annual averages of leading causes. CDC data also shown, for 2020, as an alternative estimate. See "*Some Graphs*" for data and plots.

It's right up there with the major killers, although it's known many who died from CV-19 already had these potentially fatal issues. It's not clear where dying from old age fits here, either. (It's strange how nobody dies of old age anymore—it's always some other cause.) It's interesting to read off the mortality table the IRS uses that the expected age of death is 77.3 yrs (median is 80.3). The probability of dying within age bin 77 to 78 is 0.026, and the 77 to 78 age bin proportion of population is 2.7%, so fraction dying is 15:100,000. Similarly, the probability of dying in each year age, from 77 on up, is 62%, with population proportion 29.9%, so fraction dying in that group is 257:100,000. Total deaths are around 860:100,000 every year, so old people are a significant chunk of that, not surprisingly. See table towards end of this document.*

Recently, a lot of discussion has centered around excess deaths. The CDC has taken a stab at this, while recognizing the difficulties of getting accurate death certificate information (e.g. did this person die *from* CV-19, or die while *having* CV-19, etc.). They took weighted data from the previous three years and averaged it. I found the difference between the three year avg observed and for 2020, up to the most current week, and annualized it. It's interesting to speculate that if 5% of the population is to have confirmed cases this year, but there is a five to seven multiple for actual cases, the US would have more than 35% infected in the year, and get closer to herd immunity. There is tentative evidence that immunity may last longer than was thought before. (See graph on *"Experimental Page."*)

Estimating Actual Incidence

John Ioannidis, a highly respected epidemiologist now at Stanford, recently took a stab at estimating CV19 infection fatality rates; for example, infection fatality rate for deaths > 50:100,000 (e.g. USA) at 0.57%.[4] He is famous for his critique on research statistics released in 2005 which is available for free download on www.plosmedicine.org. I have taken a cross-plot of data he included in the paper cited below, to generate an estimate of the actual incidence of COVID19 in the Tri-County area.

The idea is to use the following regression on his data, since from the known all-mortality values one could generate an estimated IFR. For example, all-mortality for Tri-County as of 12 Sep 2020 is 0.000385 deaths per capita (7 out of 18,000 people, which has at least tripled since that date). That predicts an IFR of 0.4% ± 0.5% (based on two standard deviation of scatter in this plot):
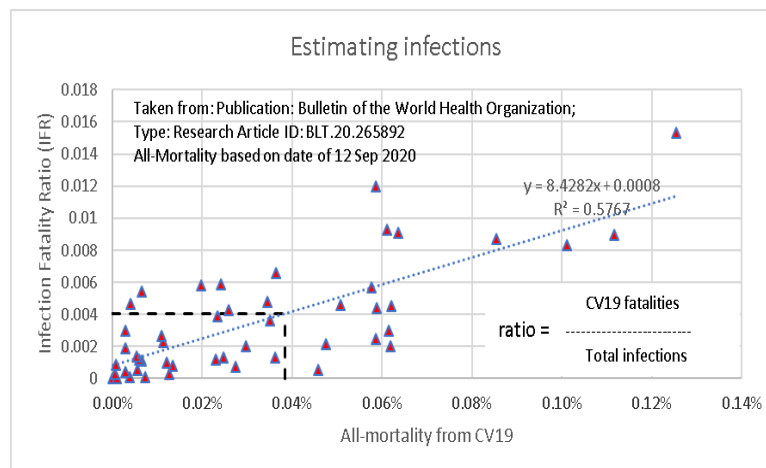
The IFR is defined as:

$$IFR \equiv \frac{Fatality}{Infections}$$

Likewise, the CFR is defined as

$$CFR \equiv \frac{Fatality}{Cases}$$

where *Cases* is the number of confirmed cases, known, always something less than the actual number of *Infections*, unknown. The ratio of *CFR*/*IFR* is the same as the ratio of *Infections* to *Cases*, *Fatality* cancels:

$$\frac{CFR}{IFR} = \frac{\dfrac{Fatality}{Cases}}{\dfrac{Fatality}{Infections}} = \frac{Infections}{Cases}$$

If you take the known average *daily* confirmed case incidence and apply this actual per confirmed infection factor, it results in an estimate for the actual number of infections. In this case, it's estimated at around three to five times the case incidence for Tri-County.

---

[4] https://www.who.int/bulletin/online_first/BLT.20.265892.pdf

Dave Leet, P.E.   20-Jan-21

Another band of researchers in Colima, Mexico, took IFR estimates and applied them to the concept of using the ratio of CV19 death incidence to IFR to estimate percentage of population already infected, which in turn is applied to an estimate of how close to herd immunity.[5] Let $x(t)$ = cumulative number of CV19 deaths, N the population of a locale, and n_{inf} number of actual infections (the number that is hard to count):

$$\frac{x(t)/N}{IFR} = \frac{x(t)/N}{x(t)/n_{inf}}$$

This leads to fraction infected a:

$$a = \frac{n_{inf}}{N} = \frac{x(t)/N}{IFR}$$

The Colima researchers took R_o to be about 5.7 for New York and New Jersey, as an example, which corresponds to γ^{-1} = 4.7 days and the slope of the exponential part of the early part of their CV19 growth curves, so is plausible. I ran the number of cumulative deaths to 17 January '21, and got $x(t)/N$ = 0.21% and 0.23%, respectively. The researchers quoted an IFR of 0.25% (Ioannidis, again):

	New York	New Jersey
Estimated infected	$a = 84.3\%$	$a = 92\%$
Estimated not infected	$1 - a = 15.7\%$	$1 - a = 8.0\%$
Herd Imm. $(1 - 1/R_o)$	82.2%	82.4%

So their conclusion is we are most likely at herd immunity already, for which the vaccines, lockdowns, and masks will get credit! Interesting idea, but keep in mind the enormous sensitivity to the estimate of IFR for such conclusions.

This is very speculative, but an example of what things have to be done to find what the actual (as opposed to confirmed) infection incidence is, approximately. Cumulative is summed off the daily data:
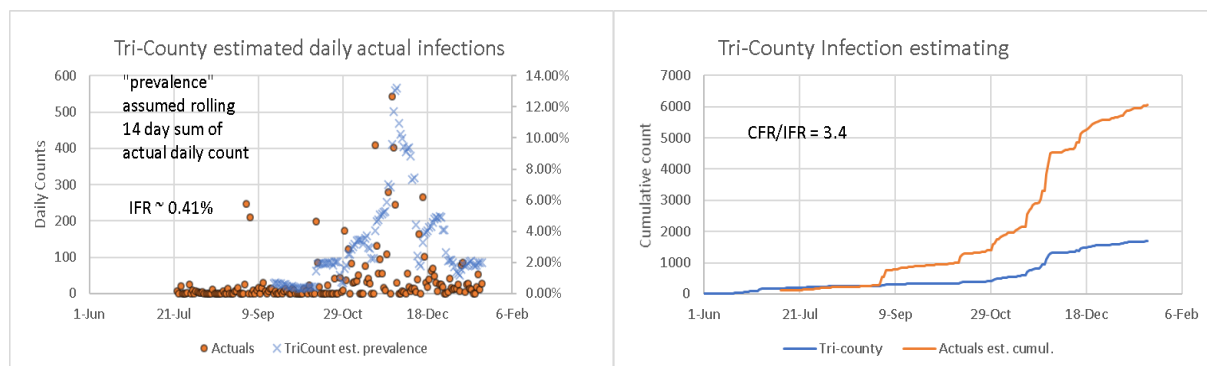
This is just an estimate, obviously, and the original source lists a whole bunch of caveats that go with it. Certainly our small population size carries more uncertainty with it, too. But as you can see, having some idea of what the actual infection incidence is, when it can't be directly measured, has a whole lot of uses.

---

[5] https://www.medrxiv.org/content/10.1101/2020.12.19.20248571v1

Dave Leet, P.E.   20-Jan-21

Masks

There's a lot of hullaballoo about masks, but it's useful to remember that the "95" rated masks, like the N95 or KN95, are rated for the most difficult particle size to trap. Above this size, it's relatively easy to make a material that will act like a seive and trap particles and still be breathable. Below this size, it's interesting that the effectiveness of the mask relies on the random kinetic motion of molecules to shove the particles *into* the material. That's how small viruses are! The viruses are on the order of 100 nanometers; the critical size most difficult to trap because it falls between the two filtration modes is about 300 nanometers. So the viruses are best trapped by kinetic motion of molecules. The N95 is rated to trap 95% of particles of 300 nanometer size, which is where the "95" designation comes from. Since this size is the hardest to trap, this implies efficacy is better than 95% for particles bigger or smaller than this critical size.[6]

However, how much virus travels in an aerosol (that is, suspended in the air by itself)? It turns out a large portion of them ride on relatively large water droplets that people cough, sneeze, or otherwise expel. The fraction traveling the one way or the other way is not well understood. An N95 is "tight" enough that it offers resistance to air flow, so if it is not fitted properly, the air you breathe will bypass the filter and it will do no good. It also will get saturated sooner or later with various particles, which increases flow resistance and increases the tendency for bypass, too. So, it must be fitted properly and changed regularly.

The cloth and surgical masks everyone wears don't offer much protection to the wearer; it's safest to consider the protection none and act accordingly. However, they do significantly impede spraying of water droplets and individual viruses in aerosols, which is their real value in protecting others *from* you, and I think that is not to be discounted, in spite of the inconvenience.

---

*This is an approximation of deaths by age out of 100k, normalized to an average 860:100,000 each year:

Age range	Probability of dying in age range (IRS tables)	Fraction of US population	Fraction dying
0 - 5	0.01474	6.62%	11:100,000
5 - 10	0.00079	6.38%	2:100,000
10 - 15	0.00099	6.37%	2:100,000
15 - 20	0.00334	6.48%	7:100,000
20 - 25	0.00462	6.86%	10:100,000
25 - 30	0.00452	7.14%	10:100,000
30 - 35	0.00551	6.77%	11:100,000
35 - 40	0.0078	6.46%	15:100,000
40 - 45	0.01151	6.11%	22:100,000
45 - 50	0.01677	6.34%	33:100,000
50 - 55	0.0238	6.48%	48:100,000
55 - 60	0.03616	6.65%	74:100,000
60 - 65	0.05371	6.07%	100:100,000
65 - 70	0.0743	5.05%	115:100,000
70 - 75	0.10233	3.87%	121:100,000
75 - 80	0.13742	2.67%	112:100,000
80 - 85	0.16348	1.77%	89:100,000
85 - 90	0.15999	1.11%	55:100,000
90 - 95	0.11601	0.56%	21:100,000
95 - 100	0.05394	0.18%	3:100,000
100+	0.00827	0.03%	0:100,000
Total	1	100.00%	860:100,000

---

[6] Millimeter is a thousandth of a meter (mm), micrometer (or micron, or μm) is a millionth of a meter, and a nanometer is a billionth of a meter (nm). So, a nanometer is one thousandth of a micron. A human hair is measured in the micron range, for example, say from 20 μm to 200 μm; 300 nm is 0.3 μm.

Explanation of statistic used for comparative purposes:

Often the generation of confirmed cases or deaths tend to be in a gamma distribution, which is a slightly skewed bell-shaped curve. A least squares procedure is used to get the closest fit to the data. Generally, the data is provided in a *cumulative* format, so the number increases each day, until the wave is over, at which point it is at its maximum and no longer changes. This can be converted to a daily format just by finding the difference between each pair of successive days. In mathematical terms, the daily format is the time derivative of the cumulative format. In any case, the area under the daily curve is identical to the last, largest and unchanging value in the cumulative format. This is the value that is used here as a comparative statistic, on a per capita basis, for a completed curve. Each locale's "wave" has a different beginning, develops at a different rate, with a different population, so comparison of absolute values on a particular date among any given datasets doesn't tell you anything, but nonetheless that is how most of this data is usually presented to the public.

Cases are described here as incidents per thousandths; deaths are described as incidents per hundred thousandths. In this way, one can get a better feel for the relative performance of locales, as long as the current population figure is available. It's also a way to gauge one's own personal risk. Keep in mind this statistic can and will change over time, not as a direct result of increasing cases or deaths, but because of the possible change in shape of the distribution. It should be thought a qualitative measure, since it's impossible to put error bounds on it, at least in its earlier stages before the inflection point.

R_o

Basic Reproduction Rate (said as R-sub-0, or just R-0). R_o is a measure of transmissibility: $R_o < 1$, disease disappears; $R_o = 1$, it's endemic; $R_o > 1$, epidemic. R_o is mentioned a lot during this epidemic, along with flattening of curves, with not a lot of understanding or relevance. It's not even a rate, but usually defined as a ratio of two different rates. It is often referred to as "infection rate", to be confused with infectious rate and infected rate; I don't think that is a good usage. It is found in the cumulative data at the very beginning of an epidemic, when the curve is exponential. More discussion in the "*Details on R_o you may or may not be interested in*" file found elsewhere on this site. "*Experimental page*" file shows some demonstrations of susceptible-infected-recovered and susceptible-exposed-infected-recovered models, and a what-if showing what a gradual moderation of the effective reproduction number (R_e) might look like.