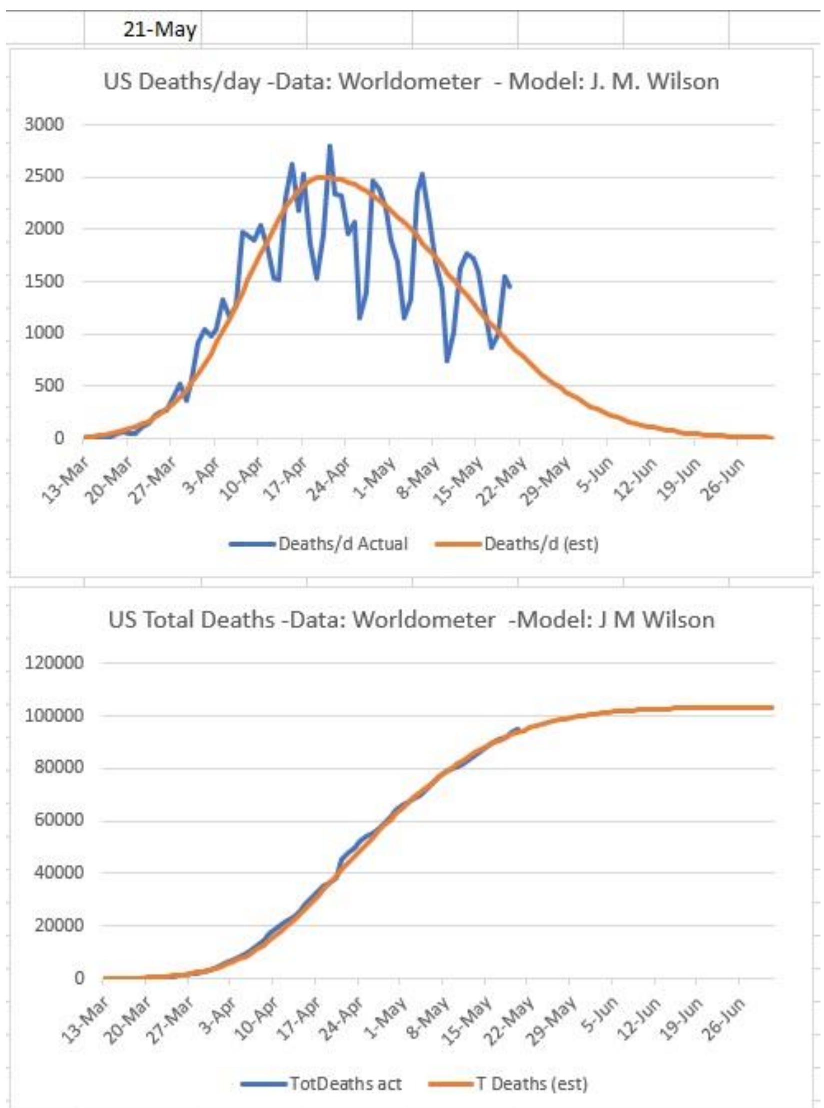
The Corona Virus Growth

Jack M. Wilson PhD

Modeling the Growth and Decline Revised 21 May 2020

I began looking at the curves of growth at the beginning of May and have followed it ever since. As I wrote in the earlier posts, the growth of the virus in the early stage approximates an exponential, but this will not continue. Eventually one expects to see some kind of rough approximation of a normal curve. In the real world all exponential growth processes eventually end in either saturation, decline, or collapse. The exponential growth phase for COVID-19 at this time has ended, The curve is looking more like the classic bell curve. We expect that any graph of total infections or total deaths will look like a saturation curve over time. It begins to rise exponentially but then levels off as deaths or infections end. If instead, we look at either deaths/day or infections/ day we expect a different curve. That curve looks like a normal curve. Again, it begins like an exponential, but then rises more slowly to a peak. After peaking the curve declines back toward zero. The Gaussian function which describes the normal distribution is symmetrical on the increase and decrease, but the rise and decline in daily infections or deaths may not be that symmetrical. This would be approximated by a skewed Gaussian. Most models use some form of the Gaussian to approximate what they expect to see. I have fit the data of the actual number of deaths per day to a skewed normal curve. The exact mathematical shape of the curve is not determined by some mathematical criteria of normal curves but is instead determined as a fit to the actual data.



The red curve on the left is a plot of a Gaussian function with a peak of 2500 deaths per day on April 20. This curve also has a standard deviation (kind of a width) of 12 days during the rise and 21 days in the initial decline. The red curve on the right is the saturation function with the same parameters -but plotting total deaths rather than daily deaths. The blue curves are the actual results to date. We expect that COVID-19 will follow a pattern something like this. In fact, the curves above are the best fits to the existing national data as of May 21. Compared to a similar fit from April 22nd the peak is just slightly higher and wider and has shifted to the right by a few days. This suggests that the total deaths will come to just over 103,000 by June 1.

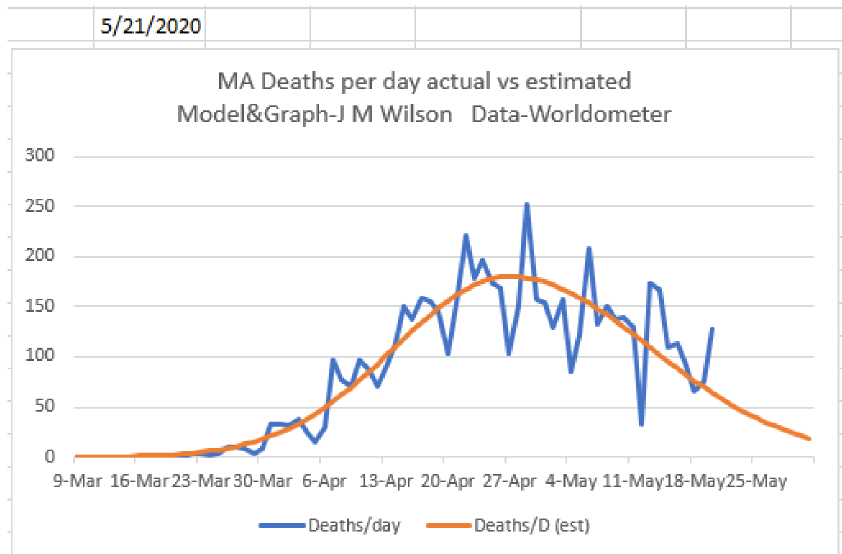
The equation for the Gaussian is written as D = Dmax*EXP(-2*((x-p)/s)^2))

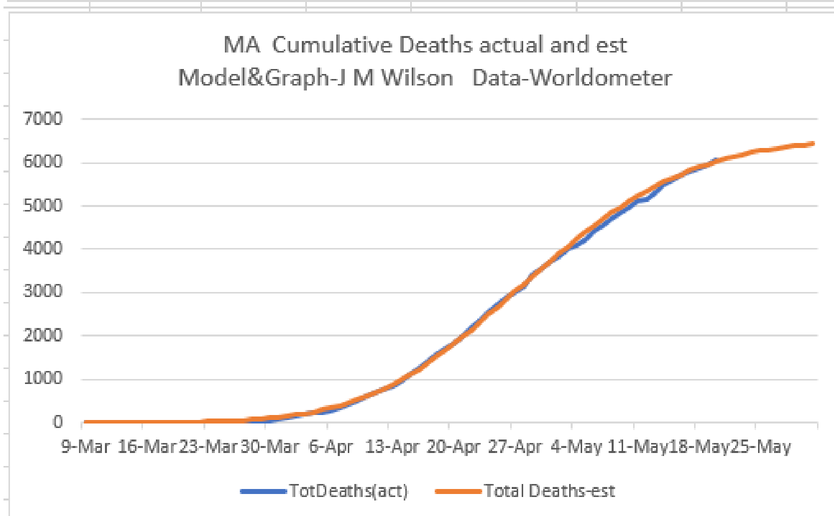
D= Daily deaths; Dmax= Max D; x=current day; p=peak day; s=standard deviation (width)

For a skewed Gaussian the width is expected to be larger during the decline than during the rise.

When modelers, and the media, talk about flattening the curve, they show this Gaussian curve and suggest that we want to lower the peak and increase the width. By lowering the peak, they hope to reduce the peak number of infections to avoid over-running the capacity of hospitals. The good news is that we have managed to do this for the most part. New York and Massachusetts have pressed the limits, but the rest of the country is doing better.

Let us consider the actual data from Massachusetts.





The curve of deaths per day peaks on April 27 at 180 deaths/day with a width of 13 days on the rise and 16 days on the decline. This is a model that is fit to the actual data up to May 21. Looking back a little over three weeks we can see that the situation worsened slightly. On April 22 it looked like it would peak at 165 deaths with a width of 10 days. April 22 showed a spike in deaths. This has shifted the curve a bit, but not too dramatically. It shows that the total deaths may reach about 6400 by June 1. If they open too quickly, then they will create a new growth curve and that will add to the toll.

These models are not based upon any assumptions as to the nature of the population, the degree of social distancing or any other factors like that that used in other models. It is a “follow-the-data” model. The largest assumption is that the number of cases will grow, peak, and then decline. We assume that the shape will approximate a gaussian. This is not unusual in that most of the models presented to the public when talking about the desire to “flatten the curve” are shown as gaussian. That said, any real curve is expected to deviate from this model in important ways. For example, we can see that the decay (downslope) will be slower than the rise. This creates a skewed normal curve. Since this model allows the data to tell us how it is expected to behave in the future, it can be susceptible to major changes if there are major changes in social distancing or other factors.

It is interesting to note that the shape of the curve has changed little over the last few weeks of modeling. The peak size and position have only shifted a bit.

Will it predict the future? We might hope so but need to be cautious. In fact, this may be a best-case situation.

Interestingly the model agrees closely with the modeling that the Massachusetts Government has done.

Unfortunately, we need to consider the limitations of such a model.

I am fully aware of the limitations of the models which include the following.

1. I modeled this as normally distributed -a gaussian curve. Most models do this to some extent, but it is clearly wrong. However, it DOES give the general features of the model. It also shows both the exponential growth in the early stages and the exponential decline in the later stages. It is skewed as one expects.
2. The model assumes that there is no radical change in conditions. If social distancing is widely abandoned, this will change the model drastically.
3. The data is very noisy. This is to be expected. The numbers (N) are small and represent human beings and not random events. This is why media reports of day to day changes have no meaning whatsoever. One must do a fit or a filter to see the trend over several days. In any day to day comparison, things may jump up or fall down, but that is probably only a fluctuation.
4. I use a brute force Monte Carlo method of fitting with a resolution of 1 day and about 3-5% on deaths.

All this said, the models give surprisingly realistic results. For example, I estimate 6400 total deaths in Mass by June 1. The state has a similar estimate. The University of Washington model originally gave results double that -over 8000 deaths. They improved the model on about April17, and now agrees with my model and Massachusetts model. Their modeling has some serious problems in any case. It has many arbitrary variables. Mine has no arbitrary variables, but just represents a best fit to the data. The difference means that theirs jumps all over the place with changed assumptions. Mine moves slowly and predictably with each new data point.

Despite the noisy nature of this data, the data gives a pretty good fit to a Gaussian curve. The rapid fluctuations in the number of daily deaths could be because of uneven reporting practices or it might indicate something worse. Each new day of data gives us more information as to where this curve is going. However, the noisy data also tells us that day to day fluctuations in the data tell us nothing. It is the overall pattern of the data (the curve) that tells the story. The media gets excited when the number of deaths falls (or rises) suddenly from one day to the next. That may make good or bad news stories but does not help to provide any insight into what is happening.

(Material posted earlier talking about exponential growth.)

Understanding Exponential Growth in the Early stages and how that would end.

I originally published material on the early exponential growth on March 11, 2020 I then revised that most recently on March 26, 2020 and following dates up until now.. This post is now superseded by the above post. Everything that I say below was written on either March 11^{th} or March 26 as noted. My plan when writing this was to show the reader how the virus would grow exponentially in the early days, but that this could not last. Eventually it must peak and decline. Because exponential growth causes things to double quickly, it is hard for many readers to see why physicians and scientists were so worried in the early days. Now, looking back a few weeks, we can see that it was even worse than we thought it might be.

Looking back on March 26 at my original post from March 11th, I recognize that what has happened in the 15 days since is far worse than I thought it would be.

When I first wrote this, I was very hesitant to post this data for two reasons. The first is that it would be alarming to some people at a time when calm rational action is what we needed. The second is that it will invariably be wrong in the details -although it is essentially correct in the pattern. In fact, our intent is to work to MAKE IT WRONG!

-

I decided to post it because it became clear that the mathematics of exponential growth is not something that is obvious to most of the general public. I puzzled at how someone could say something as foolish as "*There are only a thousand patients sick with COVID-19, so why is everybody panicking about it. Flu kills so many more*." Those not accustomed to thinking in mathematical terms (and that includes most of us), do not see the pattern. The other reason that I decided to post it is that we want to take actions to make it wrong. That is why such draconian measures are being taken in the face of what looked, at that time, to be a minor medical issue in the US.

-

The sad truth now is that I underestimated how bad it would be. I estimated a doubling time of 6.3 days and hoped that could be lengthened. It was actually three days. I estimated that there would be 4000 infections and 56 to 136 deaths by March 23. Instead there were 43,781 infections and 555 deaths. On March 26, the number of infections was 68,594 and the number of deaths was 1036. I did not think we would reach this number of deaths until April 18. Here we were on March 26 with over a thousand dead and doubling every three days.

When I published this, I had some who accused me of publishing exaggerated numbers to alarm people. Instead I did not alarm enough, and the government has not done enough.

Exponential Growth of COVID-19

Originally the exponential growth of this virus indicated that it was doubling roughly every six days (6.3). Instead it is now closer to three days. At that rate, I estimated that, without drastic action, we would have a million people ill by mid-May and we would have infected half of our population by late June. As we approach that level of infection, the growth will begin to saturate, and the doubling times will slow down. We often say that the virus is burning itself out. Eventually it will reach a static level of infection -just like the many other, already endemic, corona viruses. However, by that point millions of people (1.8 to 4.5 million) would have died. Fortunately, that will not happen. That is why we should have taken drastic actions and still should. If we do not, our hospitals and medical services will be overrun. If we can extend that doubling time through social distancing to something like 20 days, we can flatten the curve. By mid-May we would have only been looking at 8000-10,000 infections instead of about 1 million. Our hospitals can deal with that. We would also push the point where roughly half the population is infected into 2021. We do expect that we can have a vaccine ready sometime in 2021.

Corona Virus Doubling Time	Estimate (days) - 3/10	6.3

	Death rate estimates	
	Harvard	WHO
	1.40%	3.40%

	Infections		Deaths		
Date	Estimated	Actual	Actual	Deaths est	Deaths(WHO)
11-Mar-20	1,000	1,301	38		34
17-Mar-20	2,000	6,344	110	28	68
23-Mar-20	4,000	43,781	555	56	136
29-Mar-20	8,000			112	272
05-Apr-20	16,000			224	544
11-Apr-20	32,000			448	1,088
17-Apr-20	64,000			896	2,176
24-Apr-20	128,000			1,792	4,352
30-Apr-20	256,000			3,584	8,704
06-May-20	512,000			7,168	17,408
13-May-20	1,024,000			14,336	34,816
19-May-20	2,048,000			28,672	69,632
25-May-20	4,096,000			57,344	139,264
31-May-20	8,192,000			114,688	278,528
07-Jun-20	16,384,000			229,376	557,056
13-Jun-20	32,768,000			458,752	1,114,112
19-Jun-20	65,536,000			917,504	2,228,224
26-Jun-20	131,072,000			1,835,008	4,456,448

But that is not what happened. Instead the rate of infections and death has gone up. Many argue in print that the increase in testing is responsible for the rapid growth in the count of infections. That is certainly partially true. However, that is not the whole story. The deaths are also doubling at this increased rate. The number of deaths is an actual number. That is NOT because of increased testing. That is a real number. We cannot relax and trick ourselves into thinking it is a testing artifact. Yes, the number of infections will go up as the number of tests increases. For that reason, I would suggest that you watch the growth in the deaths instead. That is a real number that cannot be manipulated by changes in measurement.

-

Update to March 26

I have been checking these numbers regularly since March 11. For example, on March 17 after one doubling period there were 5243 infections instead of the 2000 that I had predicted. There were also 94 actual deaths instead of the predicted 25-68. This means the doubling time has worsened to about 2 days. On March 20th, the US has reported 14,366 infections and 217 deaths. My forecasts, with the 6.3-day doubling time, did not predict this many deaths or infections until early April. Clearly things are moving much faster than originally predicted. We had hoped that we could slow things down, but we have not. The doubling time is now less than three days instead of the original 6. Some have suggested that this is because of the increased testing leading to an increased identification of cases. This is certainly true for the number of infections, but testing does not affect the number of deaths. (The one caveat: if people had previously died but not been counted as COVID-19 this would have a small effect on the death rate doubling.) Thus, the doubling time of approximately three days (actually less than three days), based upon the fatality rates, is fairly accurate.

-

Myth: Youth are Immune

There is a myth that young people do not get this virus or that they do not die of this virus. Both are **false**. The truth is that their rates are probably lower, although research is beginning to question that assertion. They certainly do acquire the virus at a high rate, and some do die, but not at as high a fatality rate. In fact, the age group of 20-44 has the largest number of infections.
(https://www.bloomberg.com/news/articles/2020-03-19/coronavirus-in-young-people-is-it-dangerous-data-show-it-can-be) or
(https://www.statnews.com/2020/03/18/coronavirus-new-age-analysis-of-risk-confirms-young-adults-not-invincible/)

Here is the latest CDC data.

TABLE. Hospitalization, intensive care unit (ICU) admission, and case-fatality percentages for reported COVID-19 cases, by age group — United States, February 12–March 16, 2020

Age group (yrs) (no. of cases)	%*		
	Hospitalization	ICU admission	Case-fatality
0–19 (123)	1.6–2.5	0	0
20–44 (705)	14.3–20.8	2.0–4.2	0.1–0.2
45–54 (429)	21.2–28.3	5.4–10.4	0.5–0.8
55–64 (429)	20.5–30.1	4.7–11.2	1.4–2.6
65–74 (409)	28.6–43.5	8.1–18.8	2.7–4.9
75–84 (210)	30.5–58.7	10.5–31.0	4.3–10.5
≥85 (144)	31.3–70.3	6.3–29.0	10.4–27.3
Total (2,449)	**20.7–31.4**	**4.9–11.5**	**1.8–3.4**

* Lower bound of range = number of persons hospitalized, admitted to ICU, or who died among total in age group; upper bound of range = number of persons hospitalized, admitted to ICU, or who died among total in age group with known hospitalization status, ICU admission status, or death.

Credit: U.S. Centers for Disease Control of Prevention

-

Fatality Rate

The data on the expected fatality rate has also changed a lot. This is more strongly affected by the increased testing. As more infections are discovered, the percent that result in death is decreased. The Harvard generated estimate, that I used above, attempted to correct for the undertesting and now looks much closer to the actual death rate than the WHO estimate. In fact, it now looks like the mortality rate is close to 1%. However, that fatality rate varies significantly depending upon age as you see in the CDC data above and in the Imperial College London data below.

(https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf)

Table 1: Current estimates of the severity of cases. The IFR estimates from Verity et al.[12] have been adjusted to account for a non-uniform attack rate giving an overall IFR of 0.9% (95% credible interval 0.4%-1.4%). Hospitalisation estimates from Verity et al.[12] were also adjusted in this way and scaled to match expected rates in the oldest age-group (80+ years) in a GB/US context. These estimates will be updated as more data accrue.

Age-group (years)	% symptomatic cases requiring hospitalisation	% hospitalised cases requiring critical care	Infection Fatality Ratio
0 to 9	0.1%	5.0%	0.002%
10 to 19	0.3%	5.0%	0.006%
20 to 29	1.2%	5.0%	0.03%
30 to 39	3.2%	5.0%	0.08%
40 to 49	4.9%	6.3%	0.15%
50 to 59	10.2%	12.2%	0.60%
60 to 69	16.6%	27.4%	2.2%
70 to 79	24.3%	43.2%	5.1%
80+	27.3%	70.9%	9.3%

https://www.imperial.ac.uk/media/imperial-college/
medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf

We should not fret too much that this data will change and adjust as more and better data come in and as the pandemic evolves. That is just the way it is. Real science is always statistical. It does not give exact answers, it gives probabilities. The REAL number is the number of deaths, and that is increasing exponentially at a rate greater than anyone estimated in the early days.

-

There are also many who write opinion articles suggesting that the death rates will be much lower than the original estimates by WHO. I agree with that. However, some of them give rates that are completely unsupported by the data. As testing increases, we will find that there are many more people infected than we knew about. Since death rates are calculated by dividing the number of deaths (a real number) by the number of infections (estimated from testing) then an increased infection rate will lead to a decreased death RATE.

[Death rate] = [#Deaths[/ [estimated infections]

-

It does NOT lead to a decrease in deaths!

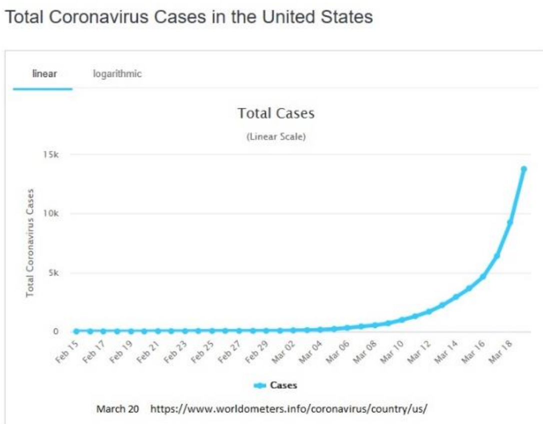
-

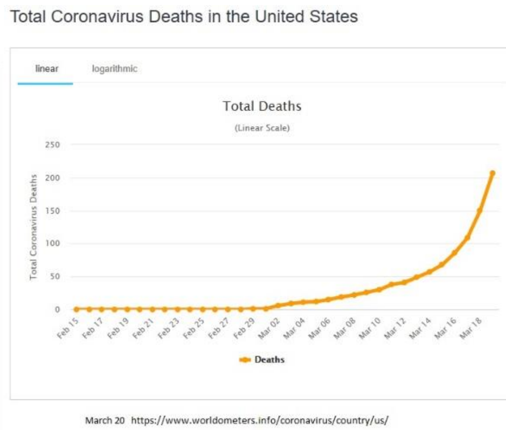
Need for Action

The actions may seem overly draconian. The paradox is that, if these actions are successful, many will see the good result and conclude “See there was nothing to worry about!” However, if we do not take the drastic actions, then many will fall ill and too many will die. That is what leadership is about. Leaders need to do the right thing -even though many will never give them credit for what they had to do.

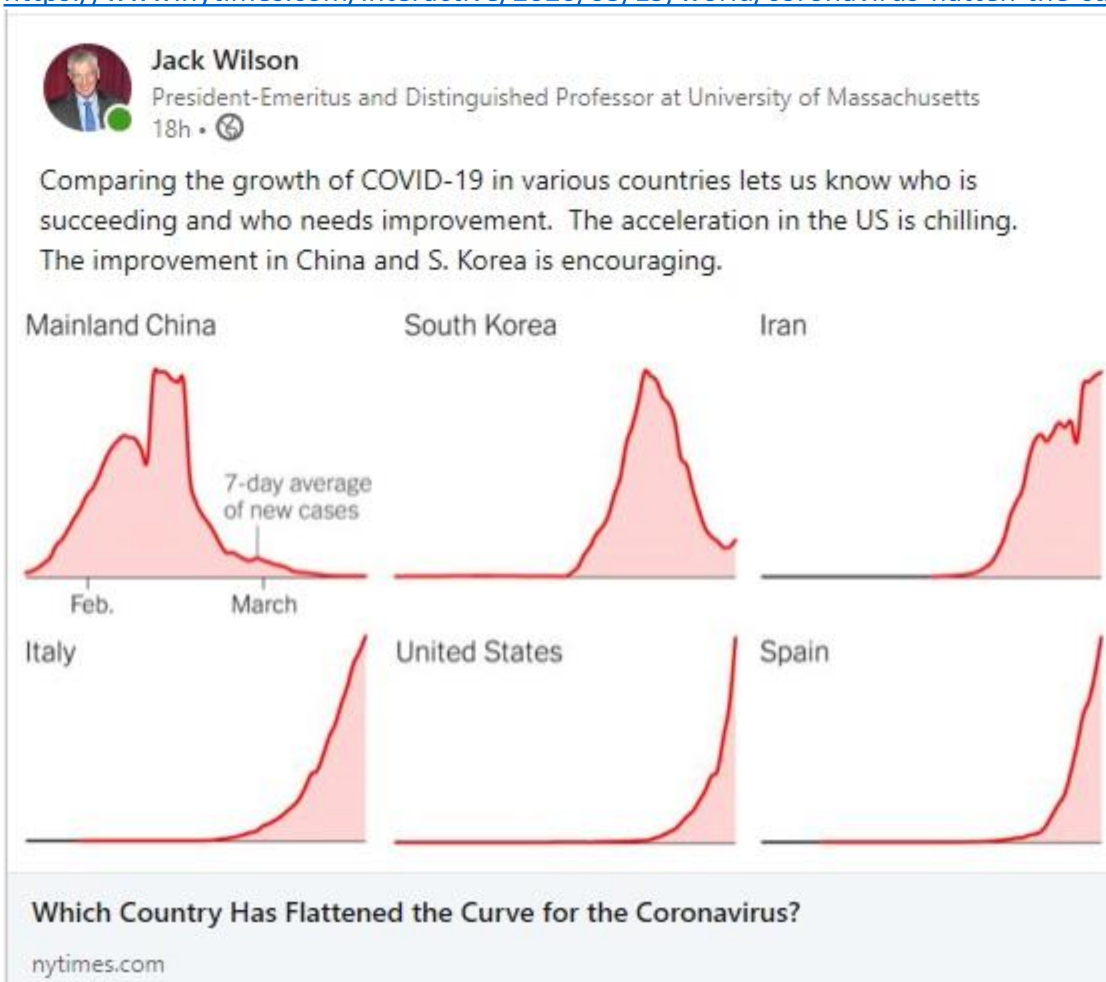
Graphical Appendix

Showing the actual exponential growth on graphs.
https://www.worldometers.info/coronavirus/country/us/





Some countries, particularly China and South Korea have managed to flatten the curve. We are not managing to do that yet.
https://www.nytimes.com/interactive/2020/03/19/world/coronavirus-flatten-the-curve-countries.html



"Honest Government Ad" Amusing but accurate take on the situation. Warning! Potentially offensive language. Do not watch if you are easily offended. https://www.youtube.com/watch?v=Hks6Nq7g6P4