Understanding the Mathematics behind the COVID-19 Challenge

Written and updated by Alan Burnham, Quest's Founder and CFO
Updated April 24, 2020

Quest's *mission is to inspire and nurture everyday exploration for lifelong engagement with science and technology. One goal is to help people understand the science behind current events. Here we explain some of the mathematics behind press coverage of the pandemic.*

You've heard about exponential growth, and probably you have seen graphs predicting the cases and deaths eventually plateauing at some value, and maybe you have heard recently about the apex. But do you understand how all these concepts relate? The object of this article is to explain the various relationships, including why we need to understand sigmoidal growth, which describes many aspects of the world around us, including algal blooms in the ocean, the growth of any plant or animal, and rush-hour traffic. This article expands upon a portion of a broader presentation created earlier on COVID-19.

In the March 27 version of this document, I plotted data through March 26 that showed an exponential increase in the number of diagnosed cases in the United States (Figure 1). An exponential plotted on a logarithmic scale forms a straight line, and the slope of that line can be used to calculate the time it takes for the cases to double, which was 2.4 days for the data after 100 cases. Continuation of that trend would have meant one million cases by April 4.

(Due to limited testing, diagnosed cases are only part of total infections. Recent antibody testing suggests that there may be far more benign cases that previously thought, but the results are not yet definitive. The rate of testing can also affect the rate of newly identified cases. In addition, data is occasionally adjusted for reporting errors, and today's update includes such an adjustment.)

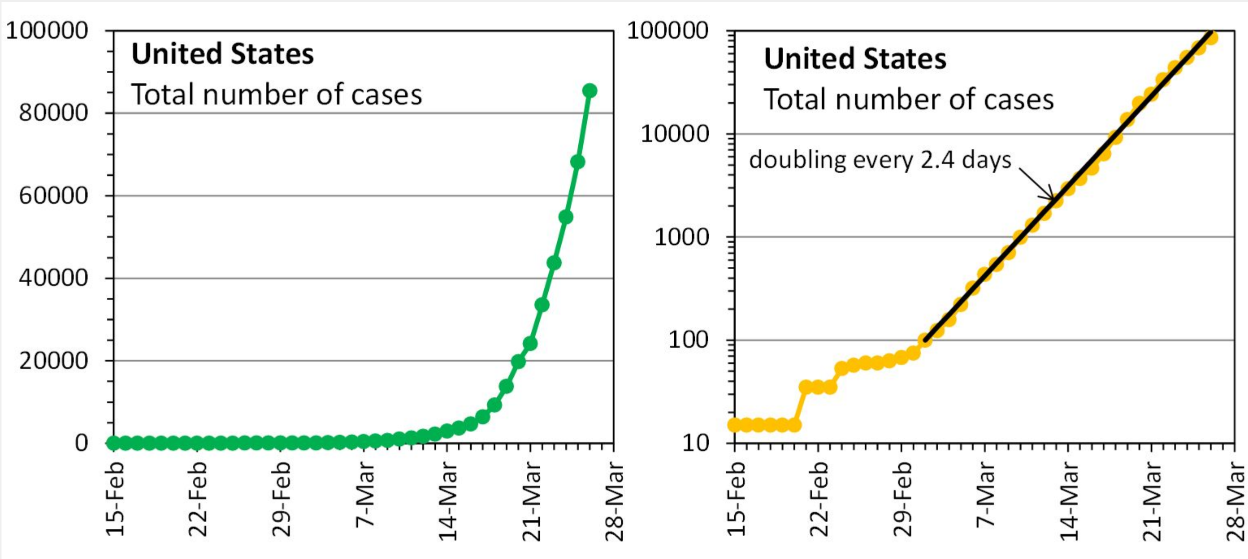


FIGURE 1

However, social distancing had started in California and other states the week before. But how effective would it be and how long would it take? We now have pretty good answers to those questions. We'll explore the math behind why social distancing worked and how well we can predict the future.

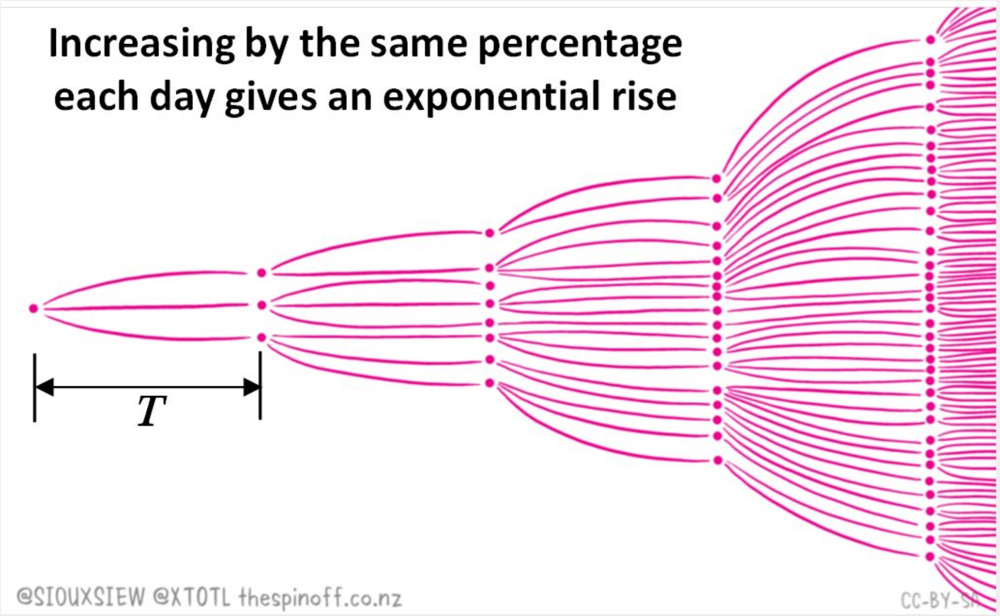
This article emphasizes the use of models to make sense of the world around us. Models can be conceptual, physical, or mathematical. Mathematical models are often incorporated into computer programs. They can be phenomenological or mechanistic models. They are also called macroscale and microscale models. Phenomenological models are empirical but try to honor the fundamental characteristics of the system, while mechanistic models try to describe how the system works in detail. They both have strengths and weaknesses, and when used well, are synergistic. The sigmoidal curves introduced in this article are phenomenological models, and epidemiologists often use microscale models, such as Kissler and coworkers in a recent issue of Science.

Why did Cases Grow Exponentially?

We all know that the virus spreads by human contact, so the more people a sick person comes in close contact with, the more people they will infect. Early studies showed three things:

(1) the average sick person infects 2.4 other people
(2) the average time for symptoms to appear is 5 days
(3) the average daily increase in cases is 33%, and sometimes closer to 40%

This leads to multiplication over time as shown schematically in Figure 2. This multiplication is actually a geometric series, which leads to exponential growth. This growth could be for rabbits breeding or weeds taking over a lawn as well as viral infections.



A good video about how social distancing reduces exponential growth
FIGURE 2

The equation for a geometric series is simply the number either created or existing at each time interval to the power of the number of time intervals:

Number of new cases = $N^{t/T}$

where t is time and T is the average time it takes to infect other people and N is the number of people infected. This equation creates more cases at a rate that depends on the values of T and N. We can create a table and graph of how many new infections occur over time as the value of N changes. The following table uses 3 rather than 5 days for T, because transmittance can occur before symptoms and T =3 causes the typical 33% daily increase.

It turns out that N=2.4 infections and T=3 days is only one of many pairs of N and T that gives that daily increase. For example, a sick person infecting 4.3 other people in 5 days gives the same result. The general relationship between N and T is $\ln(N)/T$ = 0.292, where ln is the natural logarithm.)

Days	Cycles	Transmitted infections/person (N)			
	t/ T	**2.4**	**1.6**	**1.2**	**0.9**
0	0	1.0	1.0	1.0	1.0
3	1	2.4	1.6	1.2	0.9
6	2	5.8	2.6	1.4	0.8
9	3	13.8	4.1	1.7	0.7
12	4	33.2	6.6	2.1	0.7
15	5	79.6	10.5	2.5	0.6
18	6	191.1	16.8	3.0	0.5
21	7	458.6	26.8	3.6	0.5
etc.					

The 459 new infections after 21 days in the base case don't sound like many, but it is on a trajectory to attain 40 million cases by 60 days. Such is the nature of exponential growth. This can be seen in the following logarithmic graph (Figure 3), which also shows that merely cutting the number of transmitted infections in half cuts the number of new infections to less than 200 after 60 days. Perhaps it is obvious, but if an infected person infects less than one person on average, 0.9 in the example given, the infection gradually dies out. The slope of these curves gives the exponential growth coefficient, $1/\tau$, which is also $\ln(N)/T$. This also means that one can slow the growth rate by either decreasing the number (N) each person infects or increasing the time (T) before they come in contact and get infected.

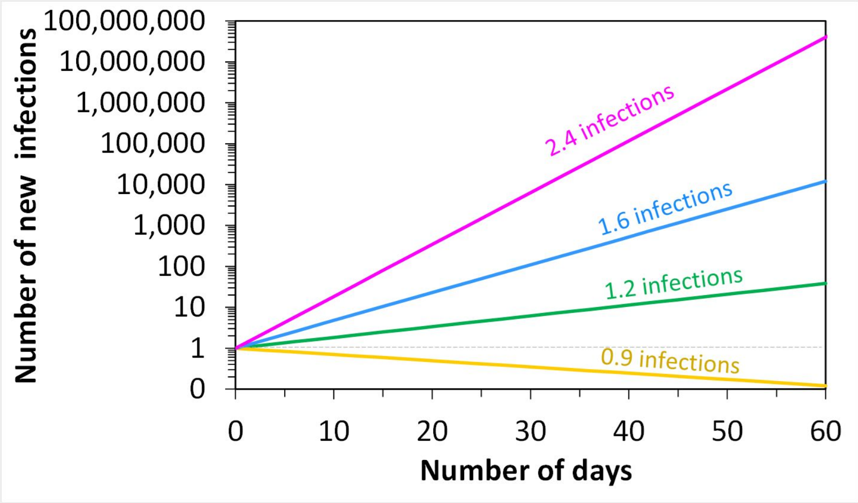


FIGURE 3

Exponential growth can't go on forever. Eventually some constraint comes into play, like running out of resources or changing the underlying driving forces for growth. As mentioned earlier, this could be either by decreasing N or increasing T for coronavirus transmission or by infecting all those not having natural immunity. A sigmoidal function initially increases exponentially but then flattens to model this behavior.

On March 16, the six Bay Area Counties implemented a stay-at-home order followed by similar state orders by California and other states. By March 28, the logarithmic graph of infections in the United States started to show favorable curvature, but low testing and reporting frequency made that trend speculative. By April 1 (30 days since 100 cases), the curvature was unmistakable, as shown in Figure 4. Also shown are the exponential growth curve and three sigmoidal curves, which capture the concept of stagnating growth, that give three different extrapolations. All three sigmoidal curves work well during the exponential growth phase, but the simple sigmoids predicting 400,000 and 1,000,000 ultimate cases diverge from the data shortly thereafter. The reason is that the infection started at different times in different regions, so a single sigmoid doesn't have quite the right shape. This can be modeled by using 10 parallel sigmoids spaced over a month. It approximately tracks case growth up to the current time.

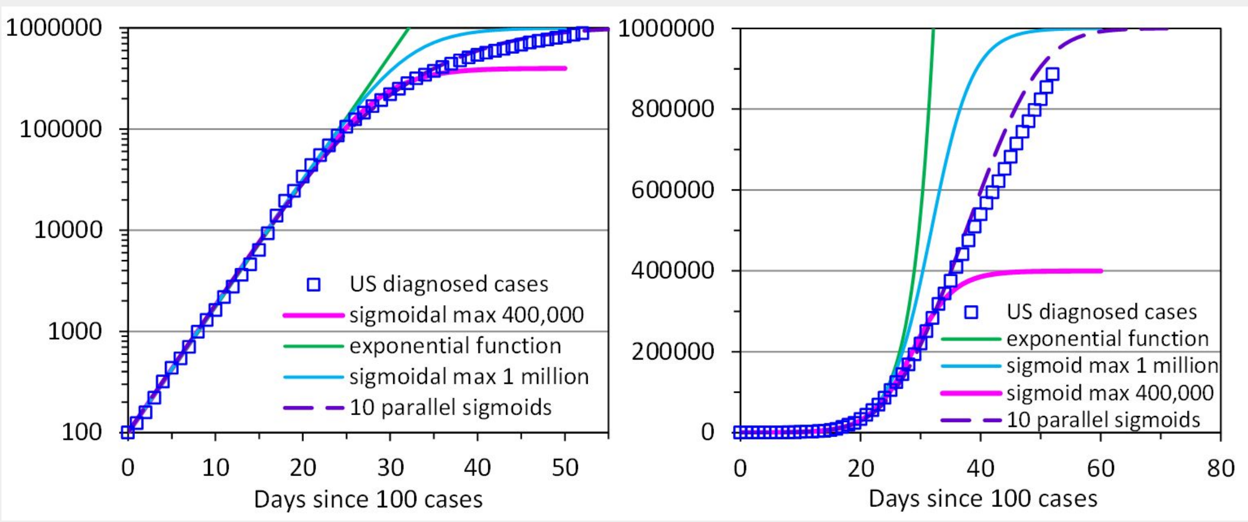


FIGURE 4

The 10-parallel-sigmoid model was finalized on April 3 before data a few days later showed definitively that the US was approaching the maximum of the US daily new cases. The prediction of the parallel sigmoid model is compared to the data in Figure 5. The apex of the daily rate curve was predicted within a day by the model. The bar graph shows the daily data along with a 3-day moving average and an exponential curve. The exponential curve, which was derived by matching data before March 22 as shown in the left figure, keeps accelerating and quickly diverges from the data.

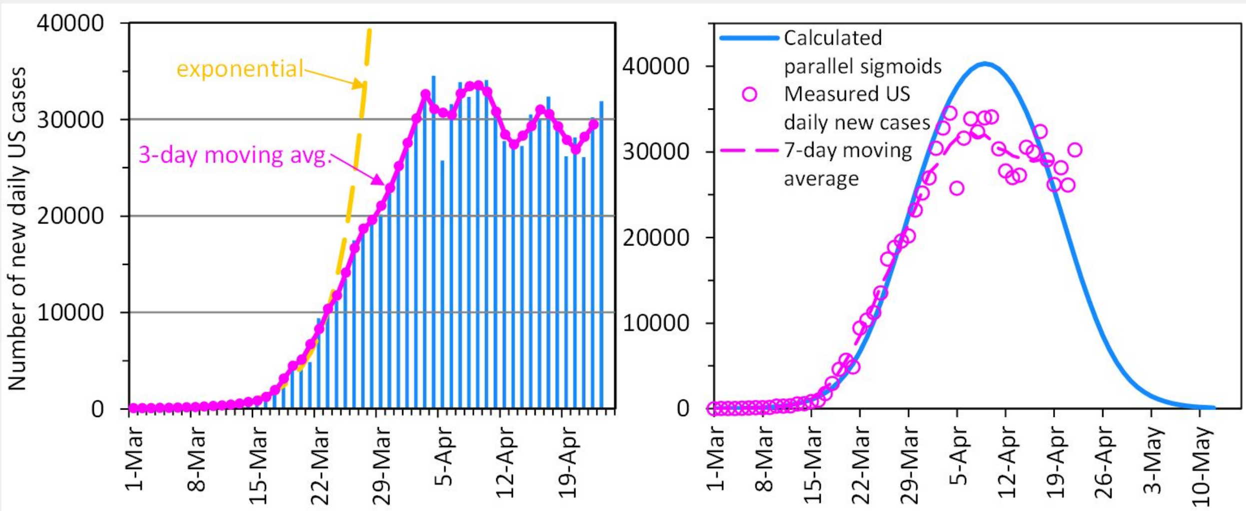


FIGURE 5

Increasingly, newspapers are explaining the status of case growth by the doubling time, that is, the time over which the cases have doubled. We next explore the mathematical relationships among exponential and sigmoidal functions and the geometric series described earlier. Figure 6 shows the shapes and equations for the rate, cumulative amount, and ratio of the rate and cumulative amount for exponentials and sigmoids. ($\boldsymbol{a}_o$ is the initial amount of infections and $\boldsymbol{a}_\infty$ is the final amount of infections, but you can ignore the equations and just look at the shapes if you wish.) The ratio of the rate and the cumulative amount is called a rate constant, and it is simply related to the percent increase per day and the doubling time.

For exponential growth (top two curves), both the rate and the cumulative amount increase exponentially, and the ratio of the rate and amount has a constant value, $1/\tau$, where τ is the time it takes for the value to grow by a factor of 2.72. The doubling time is 0.692 times τ.

For sigmoidal growth (bottom three curves), the cumulative amount is a stretched "s" shape and the rate is a bell-shaped curve. The ratio of the two is no longer a constant value, so we can describe the ratio as an effective rate constant, $1/\tau_{eff}$, that is the mirror image of the cumulative amount. It is approximately constant for the first 10% of the growth (just like exponential growth), but then it gradually declines and eventually becomes zero as the cumulative amount becomes constant.

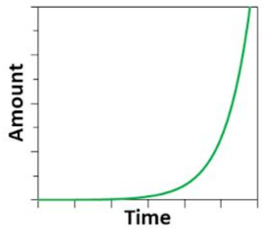
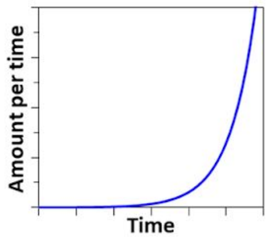
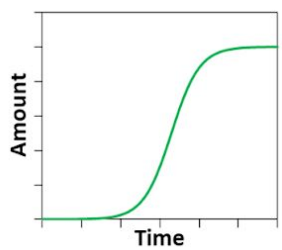
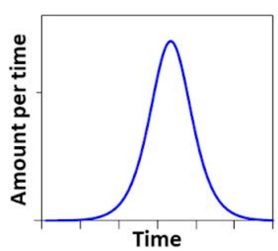
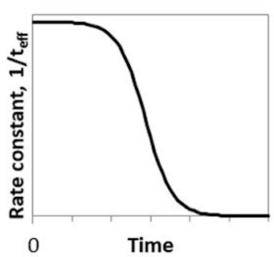
Shapes	Equations
	<u>Exponential growth</u> $amount = a_o e^{t/\tau}$
	$rate = (a_o/\tau)e^{t/\tau}$
constant value	$\frac{rate}{amount} = 1/\tau$
	<u>Sigmoidal growth</u> $amount = a_\infty/(1 + \frac{e^{-t/\tau}}{a_o/a_\infty})$ for small a_o/a_∞
	$rate = (a_\infty^2/a_o\tau)e^{-t/\tau}/(1 + \frac{e^{-t/\tau}}{a_o/a_\infty})^2$
	$\frac{rate}{amount} = (a_\infty/a_o\tau)e^{-t/\tau}/(1 + \frac{e^{-t/\tau}}{a_o/a_\infty})$

FIGURE 6

In the first few weeks of the news, the coverage was about exponential growth. A common source of confusion was that when people talked about flattening the curve, it was not always clear which curve was being described and whether it was on a linear or logarithmic plot. The simple relationship is that when the rate curve reaches its maximum, it becomes relatively flat for several days, but the cumulative amount is increasing at a rapid rate. If social distancing is maintained, then roughly half the cases will have been detected when the daily rate reaches its maximum. Only when the pandemic is nearly over does the cumulative amount curve become flat. In addition, there can be a low residual infection rate that causes the number of cases to grow slowly with time, which occurred in South Korea.

Figure 7 shows the improvement in measures of US case growth, with the points showing the US data and the lines showing the 10-parallel-sigmoid model. The blue points are $1/\tau$ and the pink points are the doubling time.

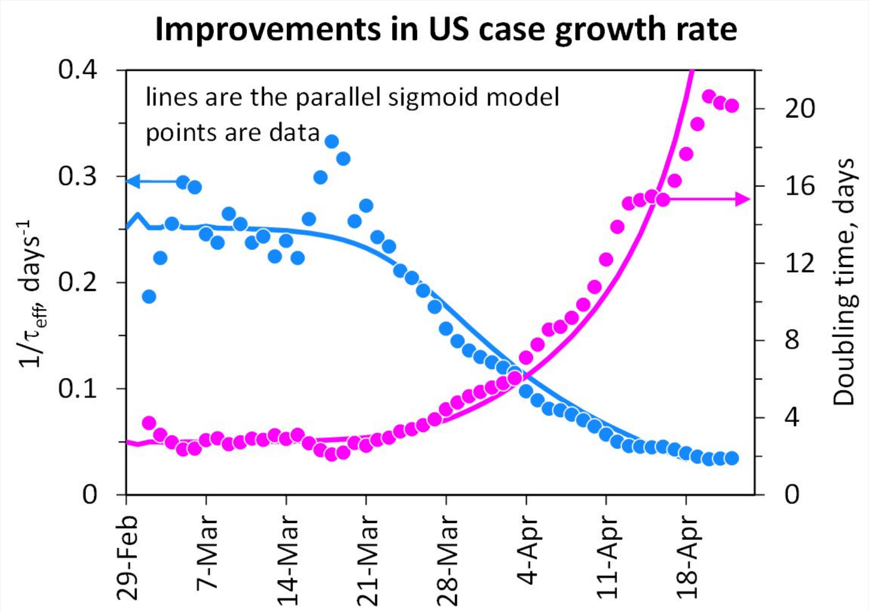


FIGURE 7

The decrease in growth rate and increase in doubling time were accomplished by some combination of decreasing the number of people each infected person infects and increasing the time that transmission occurs. Mathematics shows how the social distancing policies should and do work. The question now becomes how quickly new cases will drop to a small enough rate that the country can get back to work. It took only 2 weeks for South Korea to drop to 15% of its peak value, but it will take Germany about 4 weeks. Italy is only down to 50% after nearly 5 weeks. How fast the drop occurs depends on the size and heterogeneity of the country and the diligence in maintaining social distancing. And until herd immunity is attained, some level of social distancing will be required to prevent a major resurgence.

It is not yet clear how fast the new case rate will drop in the United States, but recent data is not highly encouraging. Figure 8 compares the US trends with Germany and Italy. The curves have been aligned so that the rise portions overlay. It is too early to be sure, but it appears that the US is dropping more slowly than Italy.

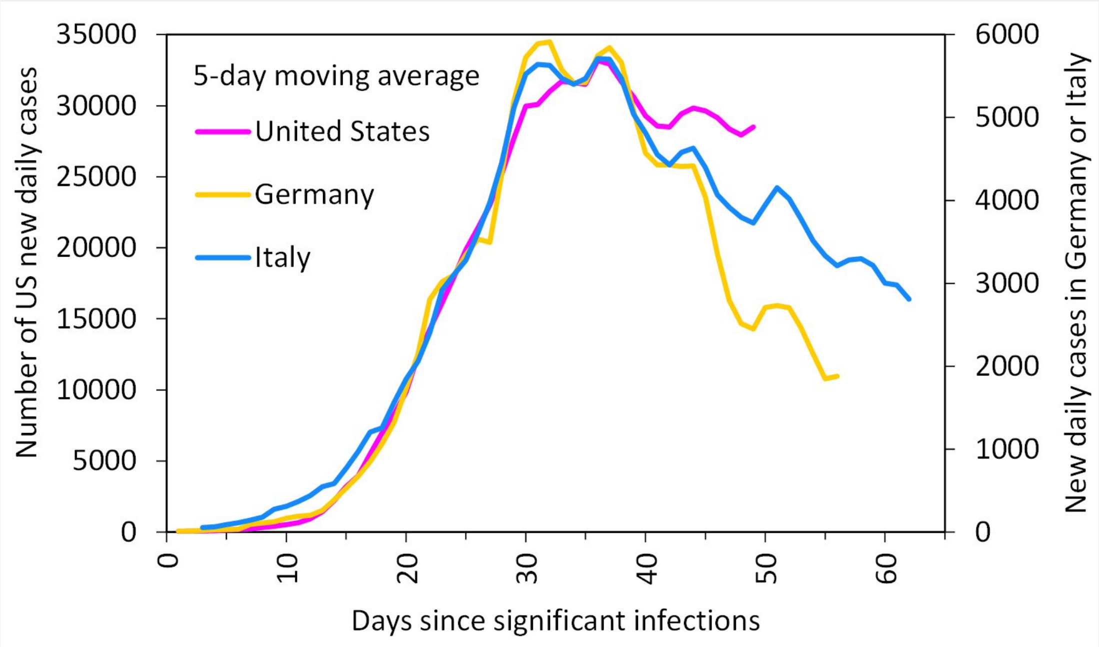


FIGURE 8

References

Wikipedia math articles: Exponential Growth, Sigmoid Function, Natural Logarithm

Types of models: Common Core High School Standard, Phenomenological Models, Microscale and Microscale Models, Epidemic Calculator

US COVID-19 data and projections: Worldometer, University of Washington models, Imperial College Study, Harvard Study (Science Magazine), Tomas Pueyo (Hammer and Dance)

Quest Science Center is a nonprofit organization based in downtown Livermore that serves the Tri-Valley region and beyond. Quest develops outreach activities and community learning experiences that connect everyone to the value of science and technology. With its dedicated volunteers, Quest champions everyday exploration and lifelong science engagement to help build a skilled workforce and inspire future innovators.