Editorial

by Gurumurthy Kalyanaram

COVID 19: Forecast of the diffusion

by Gurumurthy Kalyanaram

Covid-19 pandemic has devastated economies and societies. Forecasting the diffusion of the virus has been challenging. Over the last four months, several empirical models have been built. We present them here in very abbreviated manner. Then I discuss my own forecasts for the United States. Those forecasts changed substantially over time with new data and assumptions.

Empirical Models

SEIR Theory

This is the basic epidemiological theory/framework. Here, the population is described to be in one of four stochastic states: Susceptible, Exposed, Infectious and Recovered (looping back to Susceptible state). Assume S is the fraction of susceptible individuals (those able to contract the disease), E is the fraction of exposed individuals (those who have been infected but are not yet infectious), I is the fraction of infective individuals (those capable of transmitting the disease), and R is the fraction of recovered individuals (those who have become immune). We then have:

$S + E + I + R = 1$

University of Washington's Institute for Health Metrics and Evaluation (IHME) Model

This statistical model has received much attention from policy makers, scholars and media. The model was first launched on or about March 26th, 2020. Since then, there have been changes to the model and the forecasts. Most salient of these changes has been the underlying assumptions about social distancing. For more details, please see: http://www.healthdata.org/

Columbia University Model

The model produces projections for daily mortality and infections. For details about the model and estimation, please see: https://github.com/shaman-lab/COVID-19Projection

Covid Act Now

The underlying model is based on the fundamental SEIR theory. See here for more details: https://covidactnow.org/about

Covid 19 Simulator

This model combines infectious disease theory and statistical methodology. Please see here for more details: https://www.covid19sim.org/

Johns Hopkins University Model

The model offers a flexible framework to calibrate the diffusion trajectory and the impacts under various assumptions of policy and/or behavioral interventions. For more details, please see here:
https://github.com/HopkinsIDD/COVIDScenarioPipeline

Los Alamos Model

The model forecasts are probabilistic. For more details, please see: https://covid-19.bsvgateway.org/

MIT Model

This is an optimization model. For details, please see: https://www.covidanalytics.io/

Northeastern University Model

This is an individual-based stochastic model. For details, see: https://covid19.gleamproject.org/

The UT Austin Model

This model estimates the extent of social distancing using geolocation data. The details are available here: https://covid-19.tacc.utexas.edu/

UCLA Model

This is an augmented SEIR model that also takes into consideration the untested/unreported cases of the infection. For additional details, please see: https://covid19.uclaml.org/model.html

Youyang GU Model

This prediction model adds the power of data and artificial intelligence to the SEIR model. For more details, please see: https://covid19-projections.com/about/#about-the-model

My Bayesian Forecasts

Here are my forecasts as they evolved over time based on the new data/information.

Please note that these forecasts were made on May 31st, April 7th, and March 31st. They must be read in the context then. Each forecast builds on the earlier forecast in the context of new situation. So, there will be some repetition.

This exercise is to demonstrate the dynamic nature of modeling and forecasts. That's what Bayesian update is all about.

Forecast on May 31st

See here: https://gkpolicybriefs.com/briefs/7z4vc8w420sirm974tog73gla30xln
I wrote in the month of April that the mortality in US due to Covid-19 may be less than 70,000 - 80,000. I am wrong, of course. Why?

I based my forecast on model estimates. But the forecasts and models were based on a simple assumption: US as a whole and all the fifty states - at least till end of May - will mandate social distancing measures.

But that was not the case to be. President Donald Trump did not extend the national "pause" beyond April 30th, and let the States make the decisions. Social distancing measures were not mandated. That led to many States opening their economies and societies. For instance, Florida, Georgia and Texas were immediate to open their societies. Surprisingly, so did the State of California. Southern California opened the beaches end of April. They had to shut them down soon enough.

One state that has succeeded is New York. New York has been very deliberative. First, the entire State was divided into various individual regions, and each region was made independent of the other in reopening decision. Second, the State adopted a strategy of phased reopening. New York State has used four fundamental decision benchmarks to decide about reopening a region: Low Infection Rate; Adequate Health Care System Capacity; Sufficient Diagnostic Capacity; and Robust Contact Tracing Capacity. These four benchmarks have been translated into seven specific criteria.

Compare the state of New York with the state of California. California also set benchmarks. California adopted specific indicators/benchmarks to consider reopening: Ability to test, contact trace, isolate, and support the exposed; Ability to protect those at high risk for Covid-19; Surge capacity for hospital and health systems; Therapeutic development to meet the demand; and Ability of businesses, schools, and childcare facilities to support physical distancing. But the policy makers did not follow the decision making model and reopened much before the model recommended. The outcomes have been disappointing.

The real-world experience and the model outcomes tell us that social distancing measures were central to impede the spread of Covid-19 and reduce mortality. *The University of Washington IHME model* changed its forecast dramatically once it learned that Social Distancing measures would not be in effect. For instance, the mortality projections were between about 82,000 on March 26th and about 72,000 on or about April 29th. By third week of April, the assumptions changed. The US Federal Government, and the State Government began to relax social-distancing requirements. The mortality projections now increased dramatically to about 134,000 on May 4th, about 137,000 on May 10th, and about 147,000 on May 12th.

Forecast on April 7th

See here: https://medium.com/@GurumurthyKalyanaram/estimation-of-spread-of-covid-19-in-the-us-c160ba1e6576

Coronavirus (Covid-19) has proved to be very infectious, transmitted via contact and air-borne. By various estimates, there are about 30 percent of patients who are asymptomatic. And they can transmit the infection too.

Covid-19 has been spreading across the globe with speed and ferocity. Here, in the US too, the infection is inflicting a lot of hurt.

The question is: what is a reasonable estimate of mortality? Estimates have ranged from 2 million to 100,000 by end of July. One of the widely cited models is the University of Washington Model.

At the end of March, the model projected about 93,000 deaths by about August 4, 2020. See here: https://covid19.healthdata.org/

I wrote then: "While this is plausible, even 93,000 appears to be on higher end. Here is why."

My arithmetic was straightforward. The aggregate numbers of infections and mortality for the US (on March 30th) were thus. Total number of infections was 139,773. Total number of deaths was 2,429. So, the percentage of mortality was about 1.7. For 100,000 deaths, assuming mortality rate of 2 percent, the number of total infections must be about 5 million. Based on the evidence thus far, those numbers appeared to be very high. On March 20th/31st, there were about 150,000 infections or so in about two weeks. The data from all of US, and from the State of New York, which was most afflicted, appeared to already show a leveling off. See: https://twitter.com/Kalyanaram_G/status/1244810297669554183?s=20

The current data increasingly suggests that the initial estimates certainly appear to be on the higher end. I have been arguing this for sometime.

On April 6th, the University of Washington lowered the mortality numbers to about 82,000 (from about 93,000). Even that is likely to be revised because the said model projected 1,967 deaths for April 6th, and the actual numbers are 1,255.

The raw data for the USA and New York for the last one week is as follows.
Mon 4/6-1182 (USA); 599 (NY); 583 (Rest of USA)
Sun 4/5-1184 (USA); 594 (NY); 590 (Rest of USA)
Sat 4/4-1352 (USA); 630 (NY); 722 (Rest of USA)
Fri 4/3-1178 (USA); 562 (NY); 616 (Rest of USA)
Thu 4/2-1084 (USA); 432 (NY); 652 (Rest of USA)
Wed 4/1-954 (USA); 391 (NY); 563 (Rest of USA)
Tue 3/31-807 (USA); 332 (NY); 475 (Rest of USA)

New York is certainly showing a leveling off, so does the rest of US.

In light of this, it is reasonable to hope that there may be fewer than 82,000 deaths (by August 4th). Here is the simple arithmetic.

The current total number of infections is 363,220, and the number of deaths is 10,847. So, the rate of mortality is about 3 percent. Using this peak rate and projecting it in the future - even though we can observe the leveling off - the total number of infections must be 2.1 million to cause additional 70,000 deaths. That is over 500,000 infections a month, when we have reached the peak and will begin to taper off.

All of this assumes social distancing till end of May. The US government has mandated social distancing only till of April. So, there are many questions:

What would happen if the US government does not extend social distancing beyond April 30th?

What would happen after end of May when social distancing ends (as assumed in the model)?

We do not yet know the answers. There is much uncertainty, and it is imperative on all of us to be thoughtful and careful.

An observation. Leveling off occurs because the epidemiological phenomenon is a S-shaped curve. The phenomenon increases at increasing rate and then levels off (asymptotically). But this leveling is not obvious in the charts of the number of infections and deaths presented by many publications. The reason is this: almost all of them are using logarithmic metric and when logarithmic metric is employed, the inflection points are lost.

Forecast on March 31st

See here: https://gkpolicybriefs.com/briefs/covid-19-estimates-of-mortality-in-the-us

Covid-19 has been spreading across the globe with speed and ferocity. Here, in the US too, the infection is inflicting a lot of hurt. The question is: What is a reasonable estimate of mortality? Estimates have ranged from 2 million to 100,000 by end of July. One of the widely cited models is the University of Washington Model. The model projects about 93,000 deaths by about August 4, 2020. See: https://covid19.healthdata.org/

While this is plausible, even 93,000 appears to be on the higher end. Here is why. The aggregate numbers of infections and mortality for the US (on March 30th) are thus. Total number of infections are 139,773. Total number of deaths is 2,429. So, the percentage of mortality is about 1.7. For 100,000 deaths, assuming mortality rate of 2 percent, the total number of infections must be about 5 million.

Are we going to get 5 million deaths in the months of April-May-June-July? That is about 1.25 million infections a month, about 30,000 infections a week on a continuous basis for the next four months. That's the question.

Based on the evidence thus far, those numbers appear high. So far, we have had about 150,000 infections or so in about two weeks. Of course, a dramatic increase in the pace of infection is plausible but appears to be relatively lower probability event. The data from all of US, and from the State of New York, which is most afflicted, appear to already show a leveling off. See: https://twitter.com/Kalyanaram_G/status/1244810297669554183?s=20

***Dr. Gurumurthy Kalyanaram (GK)**, a resident-scholar alumnus of the Wilson Center and a doctoral alumnus of MIT, is a global academic, a management and policy consultant, and an Education Counselor to MIT. He has contributed and continues to contribute as Adviser, Counselor, Dean, Director, and Professor in academe. Dr. Kalyanaram serves on the advisory boards of several scholarly journals, think-tanks and businesses.*

Dr. Kalyanaram is a well-cited scholar and an institution builder. He has had affiliations with many prestigious universities such as MIT, The University of Texas, Frankfurt School of Finance and Management, International University of Japan, Jiang Xi School of Economics and Finance, KIMEP (Kazakhstan), London School of Economics, MIT Asia School of Business, Saint Petersburg University (Russia), and Vienna University of Economics and Business globally, and Amrita University, NMIMS University and Tata Institute of Social Sciences in India.

For his services, MIT has recognized Dr. Kalyanaram with Harold Lobdell Award, and Honor Roll of Service.
http://www.gkeducationalservices.com/
https://gkpolicybriefs.com/

Dr. Kalyanaram serves as the Editor-in-Chief of NMIMS Management Review, NMIMS Economics and Public Policy, and NMIMS Engineering and Technology Review.