CRAIN'S DETROIT BUSINESS

April 07, 2020 05:54 PM | UPDATED 5 HOURS AGO

Q&A: Why pandemic modeling is an imperfect, but important, tool

DUSTIN WALSH

Wayne State University

Hengguang Li

The nearly minute-by-minute onslaught of COVID-19-related news coming out of Detroit, Southeast Michigan hospitals and the unemployment office is enough to leave most of us crestfallen.

But Michiganders largely bunkering in place, peering out to a world that's more dangerous than it was only weeks ago, still wonder when this deadly pandemic will run is course. When will we be able to roam free once more? And, ultimately, how many of us will have succumbed to the disease?

Gov. Gretchen Whitmer is using a model developed by the University of Michigan that shows COVID-19 deaths and cases per day peaking in early May — another month of rising deaths and bad news. The Institute for Health Metrics and Evaluation model, however, says Michigan will level off Thursday at 190 deaths that day — the state recorded 118 deaths Tuesday — peaking at a total of about 2,963 total COVID-19 deaths.

Those are very different projections with very different outcomes for Michiganders.

So who's right? Everyone. No one. The trouble is that tracking the apex of COVID-19 is an imperfect science built on an unknown trajectory.

Crain's interviewed Hengguang Li, chair of Wayne State University's mathematics department to discuss the flaws of mathematical modeling and why it's not much a predictor of the carnage COVID-19 will ultimately leave behind.

Crain's Detroit Business: Let's start with the question everyone wants to know. Why are the models so different?

Li: The short answer is there isn't enough data. Mathematical models are made by observing the world, getting some data and making sense of it. Then they try to make predictions.

The way I would explain it to students is imagine you're driving a car. You can observe your

speed and within an hour you will know how far you've traveled. It's a rigid, simple mathematical model. For epidemics, or pandemics, mathematical modeling is not that complicated, but the variables are. Here, we don't know the velocity of how fast the disease is spreading because of a lack of available tests and time.

Models can be unstable. To be accurate, they need very accurate data in the very beginning, which we didn't have here. The only thing you can give is probability.

The data is always small in the beginning. Why is it so important to be accurate then if the model is always changing?

Mathematical models are told in terms of equations and while the models are simple, the equations are not. Those equations may be unstable and that's the nature of the problem. We keep hearing about how the number of cases in the country doubled every three days or something very similar. That's exponential, like compounded interest. But the rate of infection is very unstable. A very small change will change the prediction a lot. If the infection doubles every three days, then the math says in 30 days, that infections will be 500 times bigger after that one month. (For example, if there are 10 known COVID-19 positive cases at the start of the month and those cases double every three days, after a month there will be 5,120 cases. If there are 15 cases at the start of the month, by the end there would be 15,360 cases.) For this type of rapid change, there needs to really accurate data to make a real prediction.

What factors make for more accurate projections?

There are two big factors people usually need — the infection rate and fatality rate. This is a new disease and we have no idea how fast it can spread. We don't know its velocity. So they are using predictive models to set the parameters, which is why the infection projections vary. Or the velocity.

If we don't know how many people were infected, because some are asymptomatic or haven't been tested or to a hospital, we can't just take the number who died divided by the number of infected. And for the fatality rate, there are many inputs people don't know. The socioeconomic factors, for example. Those data points require more time to figure out. So many factors weigh in to see which one is most important to determine the fatality rate. The situation is rapidly changing. Everybody is trying to get as much data as possible and the models will be refined and more accurate over time.

But we have data from China, Italy and Spain, etc.?

People can look into what's happening in European countries, Italy and Spain, and try to get some taste of it. But that doesn't mean (the virus) will behave the same in America. Viruses mutate all the time. Some people suspect the virus in Italy is more fatal than in other areas of the world, but we don't know that to be true yet. Also, we have to take into account politics and other methods of controlling the virus in different countries. This isn't all mathematical. There are many ciphers and aspects to think about. That's why you see updates every day, so gradually we should expect the numbers to become more and more accurate.

But if we can't model accurately now, why do it?

Take weather predictions. We have lots and lots of data and many engineering working on predicting it. Yet it's not always right. Human beings have limits and even computers have limits predicting nature. We try to make use of all the data available, but at this stage of COVID-19, this is the best that people can do. And that's important for planning, just like knowing the weather.

Inline Play
