Perspective - COVID19: A Modeling Tool for Local Epidemiological Projections
By Elvin Geng, MD, MPH
Professor, Division of Infectious Diseases, School of Medicine
Director, Center for Dissemination and Implementation

As shelter in place comes to an end in the State of Missouri and the County and City of St. Louis contemplate plans to back out of the most restrictive social distancing policies, modeling and projections can help provide a window into both where we are now in the COVID epidemic as well as what we might expect in the coming weeks and months.

Background on Epidemic Modeling

Mathematical models have been widely used to understand the potential course of many infectious epidemics. Simply put, epidemics have relatively universal features, and models can be used to create simplified and quantitative representations of such epidemic behavior. Compartmental models -- a common type -- generally conceive of individuals in a population as falling into one of several mutually exclusive and exhaustive categories in relation to the infection (e.g., susceptible, infected, recovered or removed) and then expresses the rate of movement across these compartments through mathematical equations. Other types of models simulate individuals in a population and impose epidemic rules (e.g. rate of growth) to project how the course of an infection might unfold. Mathematical models of HIV, tuberculosis, malaria and other infections have provided important insights for control of these conditions across the globe. Recently, well-known modelers such as Professor Marc Lipschitz at Harvard and Neil Ferguson at Imperial College in London have applied existing pandemic influenza models to project the course of the COVID19 epidemic. One study, from Imperial, projected 2.2 million deaths in the United States without major interventions, and is said to have influenced politicians and leaders at the highest levels of government to respond more aggressively.

The Case for Local Modeling

Well known modeling results at the global or national level, however, may not be optimal for use in Missouri or St. Louis – or any other particular regional location – because regional differences may be large and contextual information that could shape modeling is difficult to incorporate into such large scale approaches. In other words, in any given area, epidemic conditions and course differ markedly from the larger average expressed in most of these models. St. Louis and Missouri have population density, mobility patterns, employment sectors, as well as of social distancing policies (e.g., about school closure, work-from-home policies, restaurant closures) that differed in nature and timing from the national average, and these differences have implications for how and when COVID-19 spreads. In the St. Louis region, public health and health systems actors have information about such City, County and State decisions that could inform our understanding of what may come. Also, all models are improved by incorporating real data about the epidemic as it emerges, and local case counts and hospitalizations over time in our region can and should inform outputs.
Local projections may have particular relevance in the US as well because county or municipal officials must decide on policies about business closures, shelter-in-place orders and other practices in their jurisdictions. Large scale models generally do not provide this type of information.

To address this need for incorporating local information into epidemic projections, a team of epidemiologists and biostatisticians, including Maya Petersen and Josh Schwab (Biostatistics, UC Berkeley), Laura Blazer (Biostatistics, UMass), and James Peng (UCSF) as well as myself created what we call the Local Epidemiological Modeling for Management and Action model. LEMMA is a simple compartmental model of the COVID19 epidemic accompanied by an interface that allows users to enter local information about the size of the population of interest, initial assumptions about reproductive numbers, the timing and intensity of

interventions, and other factors.   Other transmission parameters can be specified as well, such as the latent period and incubation period. Version 1 ran on a web-based interface R-Shiny interface in which the user could adjust various parameters up or down and explore the implications of each adjustment on cases, hospitalizations, and outcomes. Version 2 uses an Excel spreadsheet to accept input parameters, similar to Version 1, but also accepts as an input actual hospitalization data from the area in question. Version 2 also has additional features, such as incorporating uncertainty using a pseudo-Bayesian approach, where the user can specify a range of possibilities for each input and his or her level of confidence for each of the values. The model then uses the actual hospitalization data provided by the user to constrain projections which are based on this range of inputs to those that produce estimates near to the observed data. The objective is not to create the "best" or most sophisticated model of the epidemic, but rather to create tool that can be widely used by informed public health actors at the local and regional level. The right balance between sophistication and usability remains a work in progress, but the goal is to balance relevance and rigor.

Using the LEMMA Model in the real world

Over the last month, we've been using the LEMMA model – along with my colleagues here at WashU Karen Joynt Maddox, Keith Woeltje, Christina Hohner, Abby Barker and others-- with groups from the BJC Health System and the State of Missouri to better understand the epidemic trajectory. These groups use the modeling work as one component of the broader conversation about epidemiological awareness – alongside the availability and accuracy of data on testing availability, test positivity, hospitalizations and other issues. Also, other modeling efforts are brought in and the group will discuss their results. For example, the Institute for Health Evaluation and Metrics in Seattle made projections for Missouri, and other regional teams have made their own projections as well. One feature of LEMMA model which has been particularly useful is its transparency: the ability to customize the inputs makes it possible for a group of stakeholders and experts to systematically discuss assumptions that go into such a model and therefore to achieve some consensus about these inputs. Agreement on the inputs enhances the credibility of outputs. Over time, the model projections have tracked relatively closely to actual emerging data, which creates more credibility as well.

Figure 1: Estimates of the reproductive number in the St. Louis region (Dr. Aaloke Mody) using EpiEstim package in R software and using hospitalizations to parametrize. Sensitivities analyses were conducted by removing cases from nursing homes.
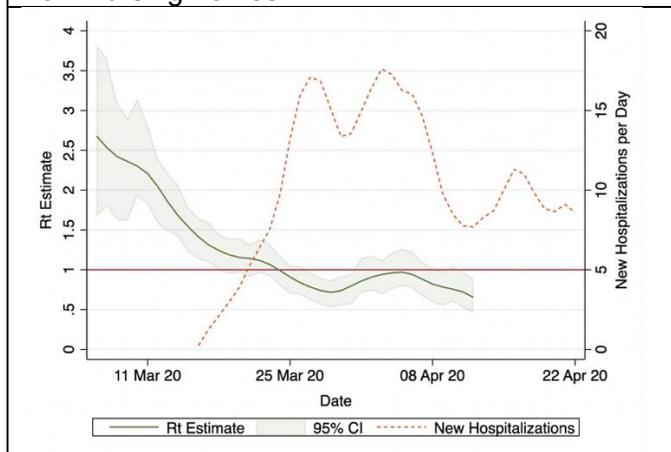
Although a tremendous amount about the future remains unknown, our discussion over the last month has brought several features of the epidemic in the St. Louis region into view, and positions us to examine the implications of choices that we now confront.

First, these analyses suggest that the St. Louis region has undoubtedly flattened the curve.   Given the vagaries of testing availability, we have used hospitalizations as a proxy for epidemic size (assuming those newly hospitalized today represent a fixed fraction of those newly infected about 7 to 10 days ago).   In late March, we saw a rapid rise in hospitalizations, which represented the growth of the epidemic during the first half of March, before major social distancing policies were enacted in the middle of the month.   The pace of hospitalizations at the time suggested a reproductive number in the region around 3 (suggesting that for each initial infection, three additional people were infected).   Without further interventions, by May 1, this reproductive number would have yielded a total of hundreds of thousands of active cases, with tens of thousands in the hospital. Such a scenario would have been catastrophic for the people of St. Louis and would have completely overwhelmed current hospital capacity.

Instead, on May 1, we had approximately 600 cases in the hospital in the region. While 600 and 10,000 seem inconceivably far apart, recall that epidemics grow on an exponential scale. For 600 cases to grow to 10,000 is less than four doubling cycles (which with COVID is scarcely a three-week difference). While the tragedy of the
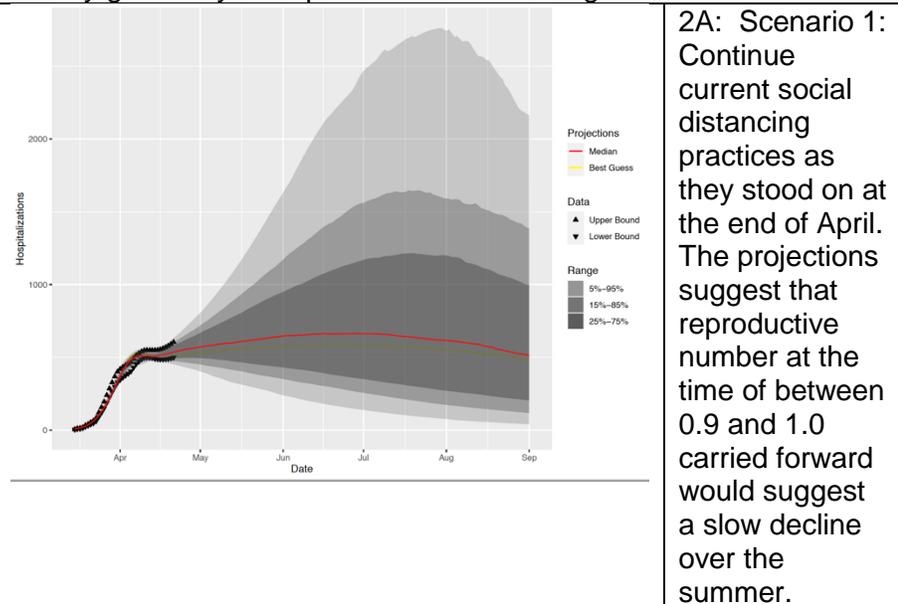
last weeks should not be understated (including unacceptable obvious racial disparities), plausible counterfactual scenarios would have been far worse.
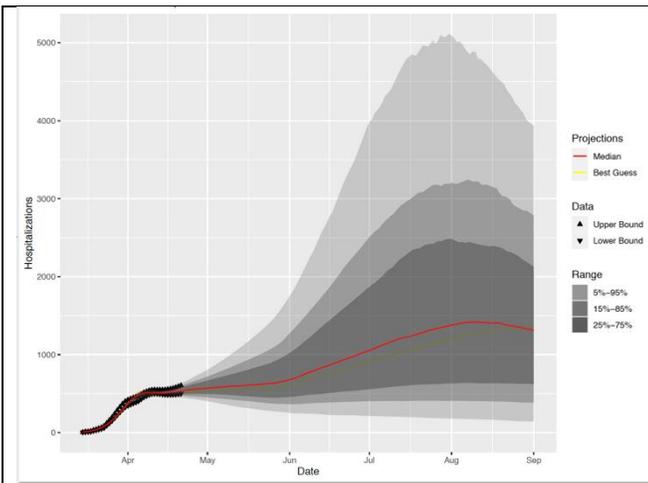
How did we avert such a crisis? The answer is through social distancing. With closure of schools on 3/17, closure of restaurants and bars on 3/19 and shelter in place on 3/22, the LEMMA models suggest that the reproductive number dropped markedly from 3, but has hovered somewhere just less than 1 (Figure 1). If we believe that the shelter in place on enacted on March 22 had its maximal effects on new infections 7-10 days later (given 7 days for the new reproductive number to be reached, and another 10 days for new infections to reach hospitalization), the maximal effect on new hospitalizations should be seen in the second half of April. The continued trickle of new admissions suggests that even with our maximal interventions, we have settled out at a reproductive number not much lower than 1.

While we can be applauded as a community for flattening the curve, this reproductive number just less than one suggests that we remain in a precarious position in the St. Louis region. As long as each infected individual in a largely susceptible population gives rise to more than one additional infected person, the epidemic continues to expand – and even if that number is only slightly greater than 1. Disease expansion can be explosive because it happens on an exponential scale.

What LEMMA tells us about next steps

Figure 2: Projections from LEMMA modeling effort under three scenarios for the St. Louis area. The red line is the median trajectory of all combinations of supplied inputs and constrained by observed hospital case series (triangles) from the region. Hospitalization numbers are entered as a range given uncertainties on any given day from persons under investigation.
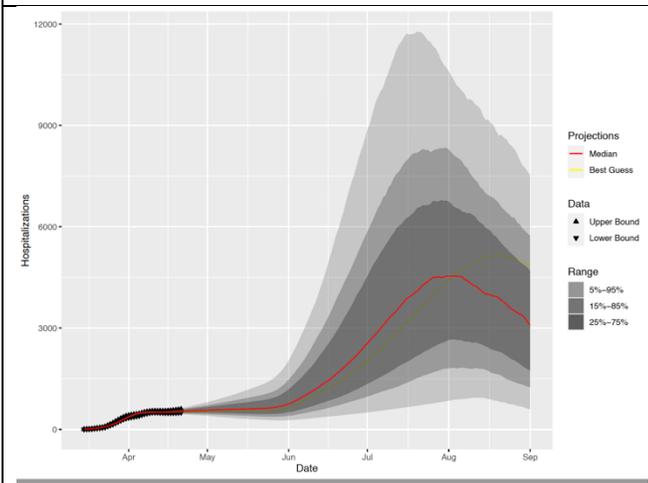

2A: Scenario 1: Continue current social distancing practices as they stood on at the end of April. The projections suggest that reproductive number at the time of between 0.9 and 1.0 carried forward would suggest a slow decline over the summer.

2B: Scenario 2: On May 15, relax social distancing by ~ 10%-20% for an effective reproductive number of between ~ 1.1. This scenario would give rise to a slow but increasingly burden of hospitalizations that could exceed 1000 by late summer.

Scenario 3: On May 15, relax social distancing by approximately 30% so that effective reproductive number is between 1.2 and 1.3. Even a 30% rise in reproductive number would quickly lead to unacceptable levels of hospitalizations.

Given that we are currently in a precarious point in the epidemic, lifting or relaxing social distancing incurs risks that we must be alert to even if countervailing economic considerations demand that we do begin to consider how to re-open the city and county. LEMMA projections suggest that the most likely scenario under current conditions – continuation of current social distancing polities – would be a slight drop in cases over the summer from a peak around July (Figure 2, Panel 1).

However, recall that we are close to a reproductive number of 1. Therefore, we have to anticipate that opening up will most likely lead to a reproductive number greater than 1. The reproductive number is a function of (1) infections per contact, (2) the average number of contacts and (3) the duration of infectivity as well as the other factors (e.g., the fraction of a population susceptible). These relationships suggest that a 10% increase in contacts due to relaxing of social distancing – an amount that seems entirely plausible – will lead to a ~ 10% rise in the reproductive number without concomitant changes in the other factors. If on May 15, the reproductive number rises to 1.1, the hospital census could exceed 1000 by late summer, and possibly higher (Figure 2, Panel 2). Higher rises in contact and reproductive numbers of 25%, 50% and 90% from our current reproductive number also seem entirely plausible. Any of these scenarios (Figure 2, Panel 3) would be catastrophic.

These projections bring some of the implications of opening up into focus. First, by looking at the other factors driving the reproductive number, we must offset the rise in contacts with reducing the successful infections per contact and the duration of infectivity. For the former, masking, fastidious hand hygiene, and other recommended activities would help. For the latter, contact tracing and isolation of infected persons and quarantine of contact can decrease the effective duration of infectivity. We need to invest robustly in these activities to keep the lid on.

Second, we need to anticipate that these measures will not keep the reproductive number under 1 and anticipate a second wave. Even if COVID19 is kept under control, given the prevalence of other respiratory viruses in the Fall that could be mistaken for COVID19, we may well be intensifying social distancing practices again by then if not before. Planning ahead can diminish the costs of doing so when we do have to intensify.

Third, planning for both current and future control measures mean that we need to equip citizens, especially the most vulnerable, to respond to the prospects of a positive test. One of the hallmarks of this epidemic in America is well known burden of disease, hospitalizations and deaths among African Americans. The notion that this epidemic is not over must motivate urgent efforts to ameliorate differential effects on social distancing

recommendations on different communities. Mitigation going forward means recognizing that the most vulnerable, who are often persons of color, who work in particular employment sectors, are least likely to be able to isolate given congregate living conditions and the least likely to be able to take time away from work. Social protections, including case management, assessment for housing and food insecurity, and cash transfer must be considered to protect the most vulnerable as well as all of us collectively.

Finally, we must invest in a multi-pronged monitoring architecture to detect changes in the epidemic course. This means rapidly scaling up testing, monitoring of cases and new hospitalizations by ZIP, enhancement of syndromic surveillance, and exploration of mobility data (Figure 3). No single source of information will be perfect, and we need to put up as many antennae as we can to detect where and when COVID is beginning to spread.

The COVID19 pandemic has introduced vast uncertainty in the lives of nearly every person on earth. The model itself as well as the data that we have are both imperfect, which could affect the projections. For example, we have some remaining uncertainty about who has been admitted from nursing homes, our estimates in scenario one could be too pessimistic. We have known unknowns: does the presence of antibodies actually confer reliable clinical immunity (many assume it will, but we await more data); will there be seasonal effects (if there are, scenarios above will change dramatically). There are also unknown unknowns – things that few anticipate to be questions at the present, but which will be obvious in retrospect. While not a crystal ball, use of local inputs to carry out modeling of the epidemic can help to make our known unknowns slightly more concrete and tangible. At the moment, these projections suggest that we have succeeded in flattening the curve, but our gains are precarious, and opening up needs to be accompanied by continued vigilance wherever possible. In the words of Maya Angelou, while we hope for the best, we must prepare for the worst, and not be surprised by anything in between.