Integrated Modeling of Healthcare Capacity and Patient Needs to Intervene Against Human Transmitted Viral Disease

Abstract

The Covid-19 pandemic has challenged the world to reduce the presence and consequence of a highly contagious and severe disease, in the absence of vaccines and effective medical treatments. Aggressive measures to isolate populations, and restrict movement, aim to reduce the number of Covid-19 cases, and to spread such cases over a longer period of time. A benefit of this strategy is to reduce the rates at which patients present at hospitals for care, so that hospitals are better able to accommodate the flow of Covid-19 patients (i.e., so-called “flattening the curve” so that the number of cases does not exceed hospital capacity). This concept is the starting point for the proposed project, which aims to develop a comprehensive patient flow model to represent the interaction between the transmission of disease and the delivery of healthcare for afflicted patients. We recognize that in a crisis strategies exist to dynamically supplement hospital capacity (e.g., transforming spaces to add bed capacity, expanding scope of responsibility, acquiring resources, optimizing discharge, etc.), diverting patients or delaying non-urgent/emergent surgeries. Thus, we aim to reduce the consequences of viral disease through interventions that simultaneously affect the rates at which new cases occur and affect the provisioning of healthcare resources the needs to serve these cases.

Zumberge funding would support preliminary research through data integration, model development, and preparation of proposals seeking external support. Modeling will emphasize improving the sensitivity of transmission models to interventions that limit the movement of people and impose their isolation, with emphasis on skilled nursing facilities, which have experienced much higher case rates and fatalities of COVID-19 than the population at large. Skilled nursing is also an important dimension of health-care delivery for COVID-19 patients, being able to accept discharged patients from acute-care hospitals and to care for patients with co-morbidities and chronic disease. A second emphasis of modeling is representing uncertainty in model parameters, which are not known with precision at the outset of an epidemic, yet can have enormous influence on the growth in number of cases.

The Zumberge project relies on an interdisciplinary skill set in transportation, healthcare delivery, statistical modeling, and optimization. The PI has conducted research in all of these areas. By the end of the calendar year, the project team will be expanded to include healthcare economics and epidemiology for submission of a proposal for external funding from the Agency for Healthcare Research and Quality (AHRQ).

The requested funding of $41,700 will support a doctoral student for one year. During that time, he will complete his dissertation proposal, qualifying exam, and submit a proposal to AHRQ for dissertation funding. An undergraduate student worker will also be supported, and additional funding will be used for data acquisition.
**Project Narrative**

COVID-19, and prior epidemics, have shown that medical strategies are insufficient to contain outbreaks of emerging viruses. Testing might not be deployed as quickly as needed to impede the spread of disease. Vaccines are developed over a longer time-period than the spread of disease, and effective anti-viral medications or other therapies might not exist, or might not be proven effective. Therefore, non-medical public health strategies aimed at lowering the transmission rate of the disease must be considered [1], [2] in concert with medical strategies.

A viral disease may be transmitted from individual to individual when they are in close proximity, or when an uninfected person is exposed to an environment or materials that remain contaminated from an infected individual. A disease may spread locally when such proximate contact occurs. A disease may spread more widely when infected individuals travel from place to place, or when materials remain infected when moved from place to place.

A viral disease may be contained or limited in its human-to-human transmission when movement of materials or individuals is constrained. Such constraints may be highly localized, such as obligating individuals to remain in their homes; regionalized, such a restricting gathering of individuals; or nationalized, restricting movements between regions, states or countries. Shelter-in-place, quarantine, ordered closings, travel restrictions, contact tracing and the like are effectively constraints bounding freedom of movement or association. Unfortunately, such non-medical interventions may invoke significant hardship -- in freedom, mobility, economic well-being, and social connectivity -- and though they may be the only means to control the spread in the early stages of an epidemic or pandemic, they are limited in effectiveness.

In 1918, the year of the Spanish flu pandemic, 20% of the world’s population, and about half of America’s, lived in urban areas. At present, more than 80% of America’s population, and more than half of the world’s population, lives in urban areas. The total world population has more than quadrupled in this time. Because the world economy continues to shift from rural farming activities, to production and services in cities, people are more crowded and proximate than ever before, making human-to-human transmission of disease more probable. Long-distance transportation has grown tremendously as well. Air travel did not exist in 1918. Since the advent of jet travel, international tourist trips have grown by a factor of 20, far exceeding population growth, creating a much faster pathway to pandemic.

Offsetting the danger of human proximity and travel, we also live in a world where technology offers the possibility to continue significant economic and social activity in the absence of physical proximity. The consequences of a shelter-at-home order would have been vastly different in 1918 than it is at present, and possibly vastly different from what it might mean in the future [3].

In the 21st century, prior to the outbreak of COVID-19, the world experienced five large epidemics: 1) In 2002, Severe Acute Respiratory Syndrome (SARS) emerged in Guangdong, China and killed 774 people. 2) In 2003, Avian flu broke out in Southeast Asia, causing more than 400 deaths. 3) In 2009, 18,500 people died of “Swine flu” or H1N1, which originated in Mexico and the U.S.. 4) In 2002, Middle East Respiratory Syndrome (MERS) emerged in Saudi Arabia, resulting in nearly 1,000 deaths. 5) In 2013, the Ebola Virus Disease emerged in West Africa and lasted more than two years, killing more than 11,300 people. All of these pale in comparison to the Spanish Flu and COVID-19. Non-medical public health intervention has played a key role on a population level in fighting epidemics through various interventions: surveillance of cases and health indicators, travel restrictions, mandatory quarantine, promotion of hand hygiene [4]. In particular, research on SARS has shown that reducing the onset-to-hospitalization interval and effective quarantine are key factors in controlling transmission [5]–[8].

[9] argued that measures aimed at limiting human interaction during a pandemic have the benefits of reducing the peak number of active infections and spreading infections over a longer period of time, helping prevent the health care system from being overwhelmed (i.e., so called “flattening the curve”). These are key (but simplified) concepts of a broader strategy of matching health care capacity to time.
varying, and stochastic, patient needs. We created Figure 1 to more completely depict patient flow during an epidemic, from time of transmission to patient recovery (or death), along with options for intervention and measures of success. Capacity, as with need (resulting from number and severity of cases) is time varying, as resources, such as ICU beds, medicine, PPE, medical staff and ventilators, are acquired or reallocated [10]. Likewise, patient requirements are met by reducing the need for hospitalization or reducing length-of-stay. With the goal of preventing the health system from being overwhelmed as the afflicted are served, a complete strategy should optimize across all options. This means matching capacity/resources to need, recognizing the rate of transmission will vary from day to day, the consequence of transmission will vary from person to person, the requirements for health services will vary during an individual’s course of disease, and the availability of health capacity can be controlled and optimized through multiple interventions. It also means recognizing that a failure to optimize may result in surpluses in some locations while severe shortages exist elsewhere, and recognizing that a failure to control spread of disease may be the cause of shortages. Thus:

The goal of our research is to optimally, dynamically and globally balance healthcare capacity with patient needs during the course of a pandemic through coordinated intervention, balancing regional supply with regional need.

To achieve this goal, we will predict success measures associated with Figure 1 as consequences of interventions through integrated modeling. The Zumberge grant would initiate this effort. As a starting point, the epidemic models play a pivotal role in predicting the transmission of disease, and hence predicting the need for medical resources. Epidemic models can be categorized into three approaches: 1) Insights into observed epidemiological trends, to identify the underlying mechanisms that influence the transmission; 2) Prediction of future epidemiologic dynamics; 3) Assessment of alternative intervention policy[10]–[12]. Researchers have developed a variety of mathematical frameworks, ranging from relatively simple curve fitting to sophisticated dynamic models, incorporating various geographic and population heterogeneities.

Based on the mathematical framework, epidemiological parameters - such as the reproductive number (R₀), latent period, transmission rate, and fatality ratio - are fundamental to insight into the epidemic and forecasts of the trend of the transmission under various scenarios. For example, the basic reproduction number, R₀, predicts the average number of the secondary cases caused by a typical infected individual in a wholly susceptible population [13]. It indicates the potential for infectious disease transmission. If R₀>1, the number of infected individuals tends to self-sustain, and the epidemic will grow. In contrast, if R₀<1, the transmission could die out due to the lack of infection replacement. For COVID-19, [14] used an iterated filter-ensemble adjustment Kalman filter (IF-EAKF) framework to predict R₀ to be 2.38 (95%CI: 2.03 to 2.77) in China. However, estimation of the reproduction number depends on the model, choice of input parameter values, and input data [10], [15]. [16] estimated the R₀ to be 2.2 (95%CI: 1.4 to 3.9) from the data on the first 425 confirmed cases in Wuhan, using a statistical exponential growth model under the assumption that COVID-19 shared the same serial interval as SARS.
To predict the future after intervention, simple curve fitting alone is inadequate. One must predict how the rate of transmission is changed as the result of the intervention, with sensitivity to the particulars of the intervention [17]. Also, to assess the effect of restrictions on travel or social distancing, one may need to incorporate spatial information in the model [18]. Further, transmission is an uncertain stochastic process. For simplicity, a deterministic model, such as the typical Susceptible-Infective-Removal (SIR) model, can provide an average estimate under assumptions of inputs. However, the predictions can be highly sensitive to input parameters. For example, at the early stage of an epidemic, even a small number of infections can be decisive to future development [14], [19], [20]. Because early stage infections are under-reported due to the lack of knowledge and lack of monitoring, model predictions can be very uncertain. How to deal with the uncertainty in the transmission process and how to enhance the robustness of the model is an underexplored domain.

For a newly emerging virus, like SARS-CoV-2, 2019 Novel Coronavirus (causing the COVID-19 disease), vaccines or proven drug therapies will not exist. Quarantine, combined with other interventions, contained the transmission of SARS [21]–[24]. However, for COVID-19, quarantine and travel restrictions were too late and insufficient to prevent a pandemic. Until a vaccine is developed, we must seek to minimize the harm of the pandemic through integrated public health intervention and medical assistance for the most severely afflicted [18], [24], [25]. Therefore, adequate medical resources and specifically hospital capacity are needed to save lives. Unfortunately, demand for medical services may exceed hospital capacity in many major cities in the US over the next four months due to COVID-19 [26], [27]. In situations like this, models for dynamically predicting the trend of the transmission and the need for medical resources under different scenarios of intervention are important. Furthermore, healthcare organizations must optimize the delivery of medical resources, considering location, time, and cost.

**Research Description**

The one-year Zumberge project will: (1) assemble data sets to refine and parameterize the patient flow model of Figure 1; (2) develop a dynamic probabilistic model to assess the effects of uncertainty in parameters during the course of an epidemic; (3) develop a proposal to AHRQ to further the research to optimize interventions, in light of uncertainties and the effectiveness of COVID-19 interventions.

**Task 1, Assemble Data Sets:** We have identified these public data sets to support the project: (1) COVID-19 confirmed cases and fatalities, by region and date, internationally from WHO website and the COVID-19 tracking project lead by the Atlantic [28]. Specifically, our data structure includes the COVID-19 cases in long-term care facilities reported by Kaiser Family Foundation (KFF) to address impact of COVID-19 on nursing homes and other long-term care facilities [29]; (2) hospital capacity and utilization, by region, within the United states from the American Hospital Association and the DHC Primary Research of the Definitive Healthcare [30], [31]; (3) transportation data from the Airline on-time statistics reported by Bureau of Transportation Statistics (BTS) showing the daily airline data between the cities in the United States and from the Baidu Intelligent Map indicating the daily migration index between the cities in China collected by the Baidu Location Based Service [32], [33]; (4) timeline data of the intervention policy complied from the national and international government website, including international travel restriction, the beginning and ending date of the stay-at-home policy for each region, and other policy that influence the transmission potentially (e.g. extension of the Tokyo Olympics, immigration ban) [34]. We will find and integrate additional data sets, for instance mobile phone data demonstrating how interventions have affected travel and data on population testing positive for antibodies to estimate the proportion of infected individuals with document cases. We will also produce a public website that links to data sources.

**Task 2, Probabilistic Modeling:** To infer the transmission dynamics over the course of an epidemic, we will develop a metapopulation model, dividing the population into the susceptible, exposed, documented infected, undocumented infected and recovered groups in each region, which are especially important in predicting transmission dynamics at the onset of an epidemic. To enhance the predictability and accuracy of the model, uncertainties in model parameters caused by the stochastic process in transmission and
reporting, we will model each parameter through random sampling from probabilistic distributions. Then an iterative algorithm will be applied to optimize solutions for the parameters. To assess the validity of the model, both the simulated outbreak and a case study of Wuhan, China will be evaluated. After validation, we will use the framework and solver algorithm to model transmission in the United States, and then assess the effect of interventions such as quarantine, travel restriction, and barrier precautions. Next, we will model disease transmission, serving as an input to predicting requirements for medical resources (e.g., ICU beds, ventilators, and professional medical staff) by location, associated with interventions.

The Zumberge research will focus on robustness of the model to underreported cases, seeking to understand the effects of timing and choice of interventions in light of uncertainty. We will simulate the transmission dynamic of the epidemics with different characteristics (e.g., reproduction number, incubation period, and latent period). Then, with the simulated data, we will analyze the minimal sufficient amount of time-series data for a trustable transmission model and find the most efficient intervention to contain the expansion, which could support emergency guidance for future epidemics.

As a starting point, we will enhance the SEIR transmission model through inclusion of regional migration parameters, thus creating a spatial model for regional interaction. We will also model the proportion of infected individuals who have documented cases of the disease, emphasizing the difference in transmission rate between the documented infected patients and the undocumented infected patients[14], [24], [35]:

\[
\begin{align*}
\frac{dS_i}{dt} & = -\frac{\beta_i S_i I_i^r}{N_i} + \frac{\mu_i \beta_i S_i I_i^u}{N_i} + \theta \sum_j \frac{M_{ij} S_j}{N_j - I_j^r} - \theta \sum_j \frac{M_{ij} S_j}{N_i - I_j^r} \\
\frac{dE_i}{dt} & = \frac{\beta_i S_i I_i^r}{N_i} + \frac{\mu_i \beta_i S_i I_i^u}{N_i} - \sigma E_i + \theta \sum_j \frac{M_{ij} E_j}{N_j - I_j^r} - \theta \sum_j \frac{M_{ij} E_j}{N_i - I_j^r} \\
\frac{dI_i^r}{dt} & = \alpha_i \sigma E_i - \gamma_r I_i^r \\
\frac{dI_i^u}{dt} & = (1 - \alpha_i) \sigma E_i - \gamma_u I_i^u + \theta \sum_j \frac{M_{ij} I_j^u}{N_j - I_j^r} - \theta \sum_j \frac{M_{ij} I_j^u}{N_i - I_j^r} \\
N_i & = N_i + \theta \sum_j M_{ij} - \theta \sum_j M_{ij}
\end{align*}
\]

The parameters are:

- \(S_i, E_i, I_i^r, I_i^u, N_i\) are susceptible, exposed, documented infected, undocumented infected and total population in region \(i\).
- \(\beta_i\) is the transmission rate from an infected individual to a susceptible individual in region \(i\), \(\mu_i\) is the adjustment factor of the transmission rate of an undocumented infected individual relative to a documented individual in region \(i\).
- \(\sigma\) is the rate at which an exposed person (carrying the virus) transitions to the infectious state.
- \(\alpha_i\) is the proportion of infected persons who are documented as infected in region \(i\).
- \(\gamma_r, \gamma_u\) are the rates of infected individuals becoming non-infectious, of documented patients and undocumented patients respectively.
- \(M_{ij}\) is the daily number of people traveling from region \(i\) to region \(j\), \(\theta\) is introduced to covered the rated of the underreported transportation data.

We recognize that the various parameters may differ from one region to another, reflecting interventions, demographics, crowding and other characteristics of the region. A particular area of interest will be the presence of disease, and policies, surrounding skilled nursing facilities, which are locations where disease is prevalent and fatality rates are high. They are also locations that provide care to Covid-19 patients, thus making it particularly important to model the relationship between transmission of disease, the ability to quarantine the afflicted, and the ability to accept discharged patients from hospitals, thus adding to the capacity to serve more patients in acute-care settings.
We also note that model accuracy is limited by the precision of parameter estimates made from available data, which in turn depends on the accuracy and completeness of reporting, as well as understanding how interventions affect parameters. We will focus on investigating these relationships through probabilistic modeling, through a learning framework in which model parameters are continuously updated to reflect the current state of knowledge and the current state of interventions.

**Task 3, Proposals:** The one-year project will be the first phase of a multi-year effort (seeking to secure external AHRQ funding), which will provide a retrospective analysis of the interventions associated with COVID-19 pandemic, as well as develop techniques to model uncertainty and optimize the timing and selection of interventions. External support will be used toward combinatorial optimization of interventions, including both medical and non-medical resources, containing the spread of disease and serving patient needs to reduce the fatality rate. In addition, the Ph.D. student supported under this project, Mingdong Lyu, will prepare his dissertation proposal and complete his qualifying exam by the end of spring semester, 2021. As Mingdong develops his proposal, the project team will expand to include his entire dissertation committee, who will be invited to participate in the external proposal. The committee will draw from faculty in epidemiology and health economics.

**Project Design**

Mingdong Lyu, doctoral student in the Epstein Department of Industrial and Systems engineering (50% support for 12 months), will have primary responsibility for Tasks 1 and 2, under the direction of PI Randolph Hall. Work will be initiated in summer of 2020, with support from Dr. Hall’s unrestricted funds. Dr. Hall will be responsible for preparing the external proposal in Task 3, and will supervise Mingdong Lyu’s dissertation proposal. An undergraduate student worker will support Task 1.

**Sustainment**

We will seek sustainment funding through application to the Agency for Healthcare Research and Quality, PA-18-795 and PA-18-765, with emphasis on “Harnessing Data and Technology To Improve Health Care Quality and Patient Outcomes” (2/5/2021 and 2/1/2021 due dates). The Zumberge grant will provide a modeling framework, data sets and data analysis as the basis for future in-depth analysis of the COVID-19 pandemic and future optimization of interventions. An application will also be submitted to the Russell Sage Foundations solicitation on Covid-19, 5/21/2020 deadline, emphasizing aspects of inequality in patient outcomes associated with Covid-19.

**Expected Results and Impact**

(1) Assembled data sets representing the dynamic nature of patient flow (including transmission and healthcare resources) represented in Figure 1, and public website with links to data sources.

(2) Modeling framework for assessment of parameter uncertainty coupled to guide interventions during the onset of an epidemic, resulting in peer reviewed publications.

(3) Submission of PhD thesis proposal by Mingdong Lyu, and passage of qualifying exam in the Epstein Department of Industrial and Systems Engineering

(4) Submission of proposals to AHRQ and the Russell Sage Foundation for sustainment.

**Timeline**

Work on data sets will begin in May of 2020, and continue throughout the project. The initial probabilistic modeling framework will be completed and documented in a research paper by December of 2020. The AHRQ proposals will be completed and submitted by 2/2021. The dissertation proposal will be completed by 4/2021. A second research paper, and final report, will be completed by 8/2021.
**Budget Justification**

A total budget of $41,700 will support the effort of a doctoral graduate research assistant, a student worker, and data licensing expenses.

**Randolph Hall**, professor in the Epstein Department of Industrial and Systems Engineering, will serve as Principal Investigator. Dr. Hall’s interdisciplinary experience in transportation, health care and risk assessment includes editing the *Handbook of Transportation Science*, the *Handbook of Healthcare Scheduling* and *Patient Flow: Reducing Delay in Healthcare Delivery*. He also served as founding principal investigator/director for two federally funded centers: METRANS (national center for metropolitan transportation research) and CREATE (Center for Risk and Economic Analysis of Terrorism Events). This background provides a unique interdisciplinary perspective on linking interventions aimed at restricting movement and association of people (e.g., “social distancing”) with delivery of healthcare. Dr. Hall will oversee all research, and write follow-on proposals, but will receive no salary support.

**Mingdong Lyu**: doctoral student in the Epstein Department, will receive 12 months of 50% graduate research assistant support, for fall 2020, and spring/summer in 2021. Mingdong has been an extremely productive graduate student, focus on statistical modeling of dynamic systems. He recently changed research directions to apply his modeling skills to pandemics. Mingdong will be responsible for assembling data sets in support of the patient flow model, and developing the modeling framework to represent parameter uncertainty. Mingdong will also develop his dissertation proposal.

**Undergraduate Student Worker**: an undergraduate student worker will be engaged to assist in the creation of data sets, including the development of online resources for public access to models developed in the project.
References


9


