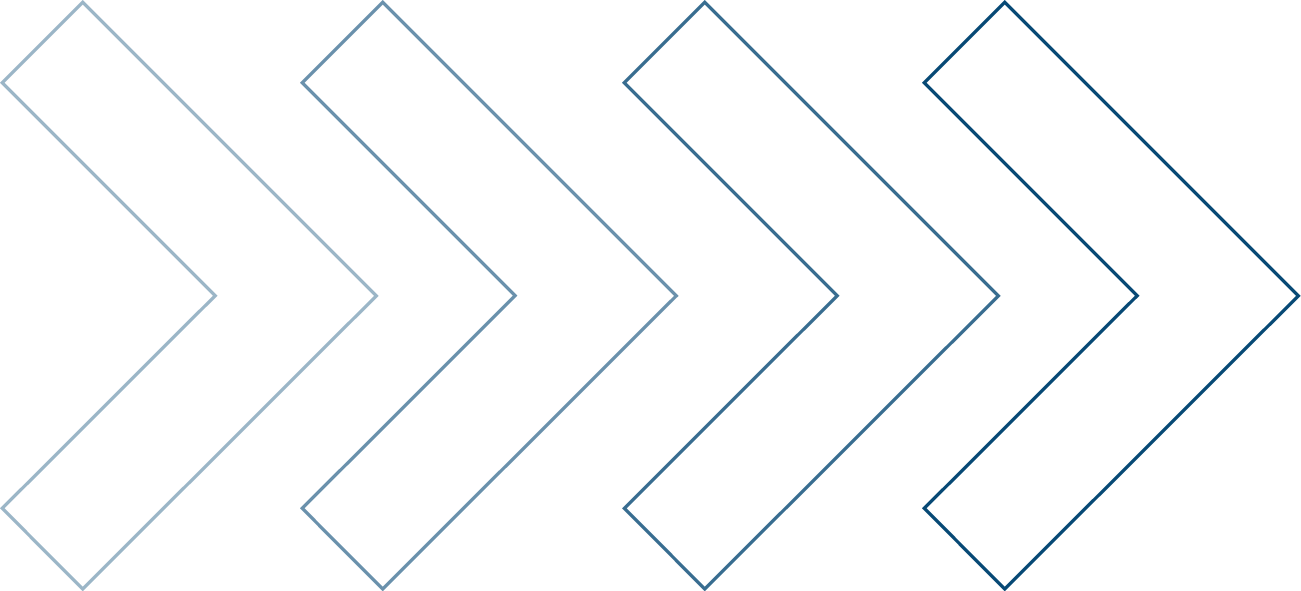
An Overview of Population Models of Epidemics



April 2020

An Overview of Population Models of Epidemics

|  |  |  |  |
| --- | --- | --- | --- |
| AUTHOR |  |  |  |
|  |  |  |

Table of Contents

[Importance of Modeling Epidemics 4](#_Toc38793139)

[The Challenges of Modeling an Outbreak Caused by a Novel Pathogen 4](#_Toc38793140)

[Key Terms that Describe Virus Tranmission and Virulence 6](#_Toc38793141)

[Overview of Modeling Approaches for Projecting Outbreaks Forward in Time 7](#_Toc38793142)

[A Closer Look at Population Models of Epidemics 8](#_Toc38793143)

[Validating and Updating Population Models 8](#_Toc38793144)

[About The Society of Actuaries 9](#_Toc38793145)

# Importance of Modeling Epidemics

The 1918 flu pandemic, oftened referred to as the “Spanish Flu”, infected an estimated 500 million persons and resulted in at least 50 million deaths, equivalent to about 3% of the world’s population at that time.[[1]](#footnote-1) While there have been major advances in medicine over the last 100 years, and vaccines have reduced the risk of outbreaks of transmissible diseases, the 1918 pandemic serves as a reminder that a novel pathogen can rapidly infect a large percentage of the human population over a short time period, with devasting consequences.

New virus strains arise on a regular basis through mutation. In addition, a virus that previously existed only in animal hosts may jump to the human population. The jump could be facilitated by a mutation that increases the virus’ capacity to enter human cells, or via the merger or “recombination” of genetic material from two separate viruses that, by chance, simultaneously infect a single animal cell, leading to the creation of an entirely new virus.[[2]](#footnote-2)

A new virus that finds its way into the human population holds a temporary competitive advantage: its novel structure makes it less likely to encounter a strong immune defense. Few people may possess immunity to the virus, or perhaps none at all. Meeting little or no resistence, a new virus has the potential to spread rapidly, causing an epidemic or pandemic.[[3]](#footnote-3)

In the event of an outbreak, simulation models can provide policymakers, governments and citizens with a means to assess strategic options. One course of action is to do nothing, letting the outbreak run its natural course. If simulations indicate that this could lead to an unacceptably high loss of life, additional simulations can be run to evaluate the impact of actions intended to decelerate or halt the outbreak. Models can provide a rough sense of the effect of options such as school closures, closures of bars and restaurants, travel restrictions, and stay-at-home or shelter-in-place orders. For each option, a forecast of the daily demand for hospital services can be compared against hospital capacity, the goal being to avoid a scenario in which hospitals are overwhelmed with an abrupt surge in the number of new patients.

When forecasting an outbreak caused by a novel pathogen, model builders must contend with many unknowns. The speed of an outbreak may leave the scientific community scrambling to collect and analyze data needed to understand the pathogen’s risk characteristics. Data limitations compel modelers to use assumptions that may have a wide range of uncertainty, and, as a consequence, outbreak forecasts also have a wide range of uncertainty. A particular forecast could, for example, indicate that anywhere from 0.3% to 1.0% of a region’s population could die in the absence of interventions to slow the spread of an infection. A wide range of predicted outcomes may frustrate policymakers and citizens who seek a more precise quantification of risk. However, a range of possible outcomes is preferable to the alternative of “flying blind”, without any guidance and insight from data-driven forecasts. Without forecasts, the human race could be caught completely offguard, deprived of the opportunity to reshape its fate via timely interventions.

# The Challenges of Modeling an Outbreak Caused by a Novel Pathogen

For any outbreak caused by a novel pathogen, the starting point for researchers is usually a small dataset – perhaps gathered from a single city or region – that is likely be inadequate to confidentally provide answers to the following key questions:

* How easily or rapidly is the virus transmitted, and under what conditions?
* What is the average length of the infectious period, and how does this period vary from one individual to another?
* How long can the virus survive outside of a human host (for example, on a kitchen counter)?
* What is the risk of death for those who are infected?
* What factors influence the risk of death, such as age, gender and comorbities?
* Are there individuals who have been infected but who exhibit no symptoms, or whose symtoms are so mild that they don’t need to visit a doctor?
* Is immunity conferred to those who survive an infection?
* Are there country, region and city-specific factors that influence the rate of transmission and/or the fatality rate?
* Does weather influence the transmission rate? For example, is transmission correlated with temperature or humidity levels?

The passage of time gradually increases the body of evidence available to researchers, and, eventually, reliable answers can usually be provided to these questions. But when faced with a rapidly progressing outbreak, time is a luxury that decision-makers do not have. In the early stages of an outbreak, a virus may potentially spread at a geometric rate (Table 1), leaving little time for model builders to gather data and refine their forecasts.

Table 1

number of days required for a virus to infect 50% of a population,

Assuming A geometric rate of growth

|  |  |
| --- | --- |
| Days Required for Infection Rate to Double | Number of Days for Infection to Spread  to 50% of a Region’s Population,  Given an Initial Infection Rate of 1 out of 1 Million |
|
| 2 | 37.9 |
| 3 | 56.8 |
| 4 | 75.7 |
| 5 | 94.7 |
| 6 | 113.6 |
| 7 | 132.5 |

Consider an outbreak in which the number of infected persons doubles every 4 days. Assuming an initial infection rate of just one out of a million persons, in merely 76 days the infection will have spread to 50% of the population. For point of reference, some studies indicate that COVID-19 had a doubling time of 2 to 5 days in the early stages of the epidemic.[[4]](#footnote-4) While geometric growth rates must necessarily decline as the number of infected persons increases and the number of uninfected persons decreases, the results above are useful for roughly illustrating the time constraints under which model builders and decision-makers must operate.

The adage “a stitch in time saves nine” is applicable to a virus that can potentially spread at a geogmetric rate of growth.

When dealing with a virus that could potentially spread at a geometric rate, delay can be quite costly.

Examples of geometric growth 1 day – 4 days – 8 days – time to infect 10% of the population

Meeting little or no resistence, a new virus has the potential to spread rapidly, causing an epidemic or pandemic.

With respect to SARS-CoV-2, the coronavirus that causes the illness known as COVID-19, questions abound regarding its transmission rate, the risk of death for those who are infected, the influence of comorbidities, the length of the infectious period, the number of people who have been infected but who have experienced no symptoms or mild symptoms that didn’t require a trip to a doctor, how long the virus can survive on a surface, whether immunity is acquired by those who have been infected, and the influence of country, region and city-specific factors on both the rate of transmission and the case fatality rate.

, decisions can not be

Along with these variables, outbreak modelers must estimate the impact of social distancing measures intended to alter the trajectory of the outbreak, as well as assess the capacity of the health care system(s) to absorb the projected case load.

ven if a simulation model correctly predicts the short-range trajectory of an outbreak, this doesn’t necessarily imply that the model’s key assumptions are correct, and, as a consequence, the model’s medium and long-range projections could prove to be off-the-mark. In a paper released on March 26 by the University of Oxford’s Department of Zoology, using data from the United Kingdom and Italy, researchers demonstrated that the time series of COVID-19 deaths from late January through mid-March can be replicated (via simulation) using either an assumed infection fatality rate (IFR) of 1%, or a much lower assumption of 0.1%.[[5]](#footnote-5) To replicate the UK and Italian time series of deaths using either an IFR of 0.1% or 1%, the researchers slightly varied the estimated starting date for the outbreak: the lower the assumed IFR, the earlier the estimated starting date for the outbreak, and the larger the estimated number of individuals who have, to date, been exposed to the virus and thus acquired immunity.

Obviously, it matters a great deal whether the IFR is 0.1% or 1% (or some other value). While either assumption is plausible in the view of the authors, and either assumption can be reconciled with January through March COVID-19 data, the long-range implications of these two assumptions are vastly different. An IFR of 0.1% is equal to that of the seasonal flu, and therefore represents a routine risk-level that society has successfully managed year after year. In contrast, an IFR of 1% is an order of magnitude greater, the equivalent of 10 seasonal flu seasons compressed into single year. Given that seasonal flu typically kills anywhere from 20 to 60 thousand Americans a year[[6]](#footnote-6), and 300 to 650 thousand persons worldwide[[7]](#footnote-7), a virus with an IFR of 1%, all else equal[[8]](#footnote-8), could potentially kill hundreds of thousands of Americans and several million persons worldwide.

To reduce the level of uncertainty associated with forecasts

the Oxford paper argues that serological surveys are needed to determine the percentage of the general population that has already been exposed to the virus. The greater this percentage, the lower must be the IFR. However, most countries do not yet have the capacity to randomly test the general population; instead, testing has been focused almost exclusively on individuals who exhibit severe symptoms. Until testing can be expanded to cover the general population, the Oxford paper implies that IFR estimates will be fraught with uncertainty, and, consequently, the level of risk posed by COVID-19 will be difficult to ascertain.

# Key Terms that Describe Virus Tranmission and Virulence

Basic reproduction number

Effective reproduction number

Infectious period

Case fatality rate

Infection fatality rate

Susceptible population

# Overview of Modeling Approaches for Projecting Outbreaks Forward in Time

Because policymakers, governments and citizens require some sense of the danger posed by a new pathogen, outbreak simulation models are necessary despite the fact that they can be difficult to construct and calibrate due to data limitations, and despite the fact that their forecasts may be subject to significant uncertainty. Of the COVID-19 models that researchers have shared with the public and described in online reports, there are two main types: 1) statistical models[[9]](#footnote-9) and (2) mechanistic models. The Institute for Health Metrics and Evaluation (IHME) model[[10]](#footnote-10), which has been frequently cited by the media as well as by the White House[[11]](#footnote-11), is an example of a statistical model[[12]](#footnote-12), while the Imperial College of London’s model[[13]](#footnote-13),[[14]](#footnote-14) is an example of a susceptible-infected-recovered (SIR) model, which is a type of mechanistic model.

A statistical model uses correlations or patterns in data to forecast the propagation of a virus. A common approach is to focus on the time series of virus-related deaths, separately by city or geographic region, fitting this data to a curve the describes the anticipated rise, peak and fall of the number of daily deaths. The curve is extracted from cities or regions that have already passed through the outbreak, such as Wuhan, China. The assumption is that, in each different region, the outbreak will follow a similar “shape”, curve or pattern across time. A model may tweak or adjust the assumed outbreak shape to account for region or city-specific factors, such as delays associated with implementing social distancing measures.

In contrast to statistical models, mechanistic models focus on the dynamic processes through which a virus propagates through a population. Estimates of the transmissibility and lethality of the virus are used to simulate the progression of an outbreak across time. A SIR model, for example, projects shifts in the population from “susceptible” (i.e. not yet infected) to “infected”, and from “infected” to either “recovered” or deceased. Some SIR models are quite simple, assuming that all persons have an equal chance of becoming sick, that infected persons are equally likely to transmit the virus, and that infected persons share the same probability of death. More complicated SIR models subdivide the population into groups, each group having distinct characteristics with respect to risk of infection, risk of transmission, and risk of death. Some SIR models go a step further, using an agent-based method to simulate unique individuals (as opposed to groups of individuals), each interacting with other unique simulated individuals.

To project hospital visits, models use data to assess the probability that an infected person will become seriously ill. Some models, such as the IHME’s model, assume that hospital services are a function of the projected number of deaths. Many models not only forecast the demand for hospital services, but also compare that demand to available hospital beds in each geographic region, thereby developing a sense of where hospital strain is most likely to occur.

Relative to statistical models, SIR models are generally more “data hungry”, but they provide greater modeling flexibility. For example, a SIR model’s key parameters can be sensitivity-tested to produce a range of simulated outcomes, such as low, most-likely and high estimates for the number of hospitalizations and deaths. In addition, the impact of social distancing measures can be estimated by adjusting the basic reproduction number (“R0”) which describes the rate of virus transmission.

Both statistical models and mechanistic models can quickly become “stale” during the early stages of an outbreak. With little data to draw upon, initial modeling efforts necessitate the use of assumptions that have a wide range of uncertainty. As an outbreak progresses, the pool of available data expands, providing researchers with valuable information that can be used to revise their models.

Inevitably, model revisions result in shifts in outbreak forecasts. Large shifts could potentially undermine the public’s faith in a model. However, revisions of forecasts do not, in general, arise from a lack of modeling expertise, but rather from data limitations that are part and parcel of dealing with a new pathogen. Revisions to forecasts are a sign that modelers are paying attention to the continuous influx of new data produced by researchers around the world, and diligently adjusting their models to reflect the most current available information about the pathogen.

In the coming weeks, this report will be periodically updated, providing a summary of the ongoing efforts of modelers to forecast the course of the outbreak.

# A Closer Look at Population Models of Epidemics

Types of population models: SIR, SEIR, SEIRS

Number of “compartments” and subgroups

Homogenous vs non-homogenous mixing models

Difficulty of determining models parameters during the early stage of an epidemic: basic reproduction number, IFR

Effective reproduction number can vary from country to country and region to region

Effective reproduction number is affected by interventions such as social distancing

# Validating and Updating Population Models

ronavirus, there are four periods associated with the disease:

Infected, but not contagious and not symptomatic.

Infected and contagious, but not symptomatic.

Infected, contagious, and symptomatic.

Recovering (assuming survival), where you may still have symptoms but are no longer contagious.

Without a widespread test available to us all, pa

# About The Society of Actuaries

With roots dating back to 1889, the [Society of Actuaries](https://www.soa.org/) (SOA) is the world’s largest actuarial professional organizations with more than 31,000 members. Through research and education, the SOA’s mission is to advance actuarial knowledge and to enhance the ability of actuaries to provide expert advice and relevant solutions for financial, business and societal challenges. The SOA’s vision is for actuaries to be the leading professionals in the measurement and management of risk.

The SOA supports actuaries and advances knowledge through research and education. As part of its work, the SOA seeks to inform public policy development and public understanding through research. The SOA aspires to be a trusted source of objective, data-driven research and analysis with an actuarial perspective for its members, industry, policymakers and the public. This distinct perspective comes from the SOA as an association of actuaries, who have a rigorous formal education and direct experience as practitioners as they perform applied research. The SOA also welcomes the opportunity to partner with other organizations in our work where appropriate.

The SOA has a history of working with public policymakers and regulators in developing historical experience studies and projection techniques as well as individual reports on health care, retirement and other topics. The SOA’s research is intended to aid the work of policymakers and regulators and follow certain core principles:

**Objectivity:** The SOA’s research informs and provides analysis that can be relied upon by other individuals or organizations involved in public policy discussions. The SOA does not take advocacy positions or lobby specific policy proposals.

**Quality:** The SOA aspires to the highest ethical and quality standards in all of its research and analysis. Our research process is overseen by experienced actuaries and nonactuaries from a range of industry sectors and organizations. A rigorous peer-review process ensures the quality and integrity of our work.

**Relevance:** The SOA provides timely research on public policy issues. Our research advances actuarial knowledge while providing critical insights on key policy issues, and thereby provides value to stakeholders and decision makers.

**Quantification:** The SOA leverages the diverse skill sets of actuaries to provide research and findings that are driven by the best available data and methods. Actuaries use detailed modeling to analyze financial risk and provide distinct insight and quantification. Further, actuarial standards require transparency and the disclosure of the assumptions and analytic approach underlying the work.

Society of Actuaries

475 N. Martingale Road, Suite 600

Schaumburg, Illinois 60173

[www.SOA.org](http://www.SOA.org)

1. https://www.npr.org/2020/04/02/826358104/the-1918-flu-pandemic-was-brutal-killing-as-many-as-100-million-people-worldwide [↑](#footnote-ref-1)
2. An overview of virus mutation is available here: <https://www.ncbi.nlm.nih.gov/books/NBK8439/>. An overview of how viruses can jump from one species to another is available here: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2546865/> [↑](#footnote-ref-2)
3. <https://www.dictionary.com/e/epidemic-vs-pandemic/> [↑](#footnote-ref-3)
4. <https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30567-5/fulltext> and <https://wwwnc.cdc.gov/eid/article/26/8/20-0219_article> [↑](#footnote-ref-4)
5. Lourenco, Jose, et al. (2020, March 26). “Fundamental Principles of Epidemic Spread Highlight the Immediate Need for Large-Scale Serological Surveys to Assess the Stage of the SARS-CoV-2 Epidemic.” MedRxiv. [*https://www.medrxiv.org/content/10.1101/2020.03.24.20042291v1*](https://www.medrxiv.org/content/10.1101/2020.03.24.20042291v1). [↑](#footnote-ref-5)
6. Centers for Disease Control and Prevention. “Past Seasons Estimated Influenza Disease Burden.” [*https://www.cdc.gov/flu/about/burden/past-seasons.html*](https://www.cdc.gov/flu/about/burden/past-seasons.html)*.* [↑](#footnote-ref-6)
7. Centers for Disease Control and Prevention. “Seasonal Flu Death Estimates Increase Worldwide.” [*https://www.cdc.gov/media/releases/2017/p1213-flu-death-estimate.html*](https://www.cdc.gov/media/releases/2017/p1213-flu-death-estimate.html). [↑](#footnote-ref-7)
8. The IFR is just one of the various key metrics that is used to quantify the threat posed by a virus. Another key metric is the “basic reproduction number” or “R0”. The greater the value of R0, the greater the rate of virus transmission from person to person. In various studies, the R0 of seasonal flu has been estimated in the neighborhood of 1.3. In contrast, the R0 for COVID-19 has been estimated in the range of 2 to 3, indicating that it is more easily transmitted than the seasonal flu. [↑](#footnote-ref-8)
9. An earlier version of this report used the term “outbreak curve model”. However, feedback received from several epidemiologists suggests that “statistical model” is the preferred term. [↑](#footnote-ref-9)
10. Murray, Christopher JL. (2020, March 30). “Forecasting COVID-19 Impact on Hospital Bed-Days, ICU-Days, Ventilator-Days and Deaths by U.S. State in the Next Four Months.” MedRxiv. [*https://www.medrxiv.org/content/10.1101/2020.03.27.20043752v1*](https://www.medrxiv.org/content/10.1101/2020.03.27.20043752v1)*.* [↑](#footnote-ref-10)
11. Aizenman, Nurith. (2020, April 1). “Five Key Facts Not Explained in White House COVID-19 Projections.” NPR. [*https://www.npr.org/sections/health-shots/2020/04/01/824744490/5-key-facts-the-white-house-isnt-saying-about-their-covid-19-projections*](https://www.npr.org/sections/health-shots/2020/04/01/824744490/5-key-facts-the-white-house-isnt-saying-about-their-covid-19-projections)*.* [↑](#footnote-ref-11)
12. While the IMHE model uses a statistical approach to project the number of deaths, the component of the model that projects hospital service utilization is best described as mechanistic. Thus, the IMHE model has both statistical and mechanistic components. [↑](#footnote-ref-12)
13. Flaxman, Seth, et al. (2020, March 30). “Report 13—Estimating the Number of Infections and the Impact of Non-Pharmaceutical Interventions on COVID-19 in 11 European Countries.” Imperial College London. [*https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-13-europe-npi-impact/*](https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-13-europe-npi-impact/). [↑](#footnote-ref-13)
14. Adam, David. (2020, April 2). “Special Report: The Simulations Driving the World’s Response to COVID-19.” Nature. [*https://www.nature.com/articles/d41586-020-01003-6*](https://www.nature.com/articles/d41586-020-01003-6)*.* [↑](#footnote-ref-14)