Understanding COVID-19 Mathematical Models

Countries around the world are relying on mathematical models to help forecast the future of COVID-19 epidemics. These models are influencing the decisions of government leaders with respect to social distancing orders and the allocation of health resources. To use these models in your advocacy and communication efforts, it is important to build some understanding about the purpose of the models and how the models work, as well as what they do not do.

WHAT IS A MATHEMATICAL MODEL?

A mathematical model is a real-world problem represented in mathematical form.

What are COVID-19 Mathematical Models?

These are computer simulations or tools that are used to answer questions about the possible future of the COVID-19 epidemic at global, national, state and local levels. These models do not predict the future, but they are valuable tools that can be used to prepare for the future because they can: **(1)** forecast what might happen, **(2)** project what could happen as a result of certain conditions, and **(3)** help to explain how something happened. The mathematical models being developed for COVID-19 can help to answer a range of questions.

Forecast what is to come
How many people are likely to become infected with COVID-19? How many would be symptomatic vs. asymptomatic? How many would need to be hospitalized? How many would require a ventilator? How many could potentially die? How many ventilators, nurses needed?

Project what could happen, under certain conditions
When is the peak of new infections expected? What would happen if we did nothing? What would happen with different social distancing interventions?

Explain what happened
When did the virus begin circulating in the US? What is the incubation period or transmissibility of COVID-19?

How do COVID-19 Models Work?

> "We are building simplified representations of reality. Models are not crystal balls." **Dr. Neil Ferguson**, mathematical epidemiologist at Imperial College, London

When building a COVID-19 model, modelers basically construct a **virtual mock-up** of the country (or city or region) for which the model is built, then **use mathematics** to determine how population groups move and interact with each other in space and time. Then the modelers introduce the COVID-19 virus into this virtual population and see what happens.

There are many **different types** COVID-19 models. Some try to understand how people who are susceptible to the virus become infected, then either recover or die. Others are statistical models that look at how the disease has progressed in countries that are farther along with their COVID-19 infections, and try to approximate how the disease will progress in another country. Some of the simplest models make basic assumptions (e.g., people with COVID-19 are all equally infectious or everyone on average has the same chance of getting infected), while other more complicated models subdivide people into smaller groups (e.g., by age, sex, health status, etc.).

	Statistical models	**Compartmental model**	**Simulation model**
Example	Curve-fitting/extrapolation models (e.g., Institute for Health Economics and Metrics (IHME) COVID-19 Model)	SEIR/SIR Models (e.g., Imperial College "non-pharmaceutical intervention" (NPI) model	Individual-based or agent-based models (e.g., COVID-19 Model from Northeastern University and others*)
Approach	Use experiences from epidemic in other locations to infer trends about disease spread in given location	Divide a population into groups ("susceptible", "exposed", "infected", and "removed/ recovered")	Create a simulated community in which disease is spread among individuals (or "agents")
Limitations	Projections based on past trends; correlation ≠ causality	Cannot handle too much complexity	Requires massive computing power; often include parameters for which we cannot get data (i.e., uses a lot of assumptions)

** Fogarty International Center, Fred Hutchinson Cancer Center, University of Florida*

Why do COVID-19 models show such different results?

> *"The best models [are] dynamic and they will change as more data becomes available."* **Dr. Robert Smith**, disease modeling professor at University of Ottawa, Canada

Mathematical models are **only as good as the data and assumptions used in them**, and they need to be **constantly updated** as more data become available and more research is conducted. At this time, the current COVID-19 models use many different data sources and assumptions because there is still a lot about COVID-19 that is remains unknown (e.g., the proportion of people with COVID-19 who will die, the average number of people who can be infected from one infected person, whether those who recover from COVID-19 infection are immune to re-infection in the short-term, whether people will follow public health measures such as social distancing, etc.). As such, these inputs used in COVID-19 models need to be estimated or assumed, which limits the precision of what the model is forecasting.

What Limitations of COVID-19 Models Should I be Aware of?

> *"There are things called models, and when someone creates a model, they put in various assumptions. And the model is only as good and as accurate as your assumptions."* **Dr. Anthony Fauci**, Director at National Institute of Allergy and Infectious Diseases

- The **further out** the model projects (e.g., 6 months, 12 months), the more uncertainty there is in these projections.
- There is still a lot that is **unknown** (and thus that is estimated or assumed) in current models.
- Most models were developed for **high-income settings** and also use data from high-income settings. It is unclear how these models can be used in other settings where there is a lack of data, information and evidence about COVID-19. In addition, these models might not translate well to other settings since they do not consider local norms, interaction with other diseases (such as tuberculosis and HIV) and health system constraints.
- Finally, most of the COVID-19 models are **available** on proprietary platforms with **codes and assumptions that are unclear** to people who are less empowered to challenge them.

What Questions Should I Ask about a Specific COVID-19 Model?

There are many different COVID-19 models in the media and in the literature, each with their own strengths and weaknesses. When using the findings from these models in your work, it is important to think about the question that the model seeks to answer. It is also important to consider the findings from different models, in order to understand the possible range of projections being made. Then ask yourself which model is the most appropriate for you to use in your work.

- What was the **purpose** of the model? What **questions** is it trying to answer?
- In what **context** was the model applied?
- What **data** were used in the model?
- What **assumptions** did the model make?
- What are the **limitations** of the model?
- How well does this model project **outcomes so far**?

SUMMARY

- ✓ Models can be important tools for understanding COVID-19 and policy responses
- ✓ To appreciate the value of COVID-19 models, it is important to understand the overall purpose of the model and the assumptions on which it was built
- ✓ Models are not meant to accurately predict the future; they are meant to be used for advocacy and policy purposes
- ✓ Models are inherently uncertain; the produce ranges or estimates and outputs
- ✓ The true value of a model might only be understood when looking back at it months or years from now

WANT MORE INFORMATION ON MATHEMATICAL MODELS?

Check out the **free, self-paced course** *Build your skills to understand and use mathematical models* at https://engage.avac.org/!

This work was supported by the Coalition to Accelerate and Support Prevention Research (CASPR), (AID-OAA-A-16-00031), made possible by the generous support of the American people through the US President's Emergency Plan for AIDS Relief (PEPFAR) and the US Agency for International Development (USAID). The contents do not necessarily reflect the views of PEPFAR, USAID or the United States Government.

	Imperial College (Dr. Neil Ferguson)	IHME (Dr. Chris Murray, University of Washington)	Imperial College (Dr. Azra Ghani)	COVID-19 Scenarios (University of Basel)
What is the purpose of model?	Understand the impact of policies and actions on the course of the disease	Project the time of peak hospital strain in an area Project the number of deaths	Project the health impact of the COVID-19 pandemic globally	This model serves as a planning tool for COVID-19 outbreaks in communities across the world
How does the model work?	Uses data about COVID-19 transmissibility[1] to project the range of deaths if the UK and USA did nothing, versus based on impact of various social distancing policies	Uses data from how COVID-19 outbreaks have progressed around the world to project what the epidemic curve will look as new outbreaks form in new areas, based on the social distancing actions being taken	Uses estimates of mortality and healthcare demand to compare projected mortality with and without interventions or social distancing	Breaks the population into age groups and different categories (e.g., susceptible, infected, dead, recovered, etc.)
In what context was the model applied?	UK, USA	US	Global	Global, includes low- and middle-income countries
What type of data are used in the model?	Populations in households, schools, workplace and wider community	Reported worldwide COVID-19 deaths	Population size and age distributions by country, contact patterns, hospitalization rates, infection fatality ratio	Population size, age-structure, reported COVID-19 cases
What assumptions did the model use?	Asymptomatic transmission occurs Transmissibility of 2.4 ranging from 2.0 to 2.6 No difference in general health and pre-existing conditions prevalence between Chinese and other populations	Asymptomatic transmission occurs Transmissibility of 3.0 ranging from 2.4 to 3.3 Similar conditions to those in other parts of the world Assumes social distancing behavior will continue[2]	Transmissibility of 3.0 ranging from 2.4 to 3.3 Potential benefit from a range of mitigation strategies Assume no difference in general health and pre-existing conditions prevalence between Chinese and other populations	Transmissibility could have seasonal variation No difference in general health and pre-existing conditions prevalence between Chinese and other populations

[1] i.e., the average number of new cases expected to be spawned by each case of an illness

[2] Note: the current IHME model goes until the beginning of August and does not project what would happen if social distancing is not fully implemented, or if there is a resurgence after August.

	Imperial College (Dr. Neil Ferguson)	**IHME (Dr. Chris Murray, University of Washington)**	**Imperial College (Dr. Azra Ghani)**	**COVID-19 Scenarios (University of Basel)**
What were the main findings from the model?	Population-wide social distancing, along with home isolation of cases and school/ university closure has the potential to have the largest impact in suppressing transmission	As of May 4, 2020, the model estimates 134,475 deaths (estimated range of 95,092 to 242,890) deaths is projected in the US by August 4, 2020	"In the absence of interventions, COVID-19 would have resulted in 7 billion infections and 40 million deaths globally this year."	Shows that outbreak depends on interventions and parameters
What else is of interest?	This was the model that spurred the UK government to take more action to address COVID-19	Earlier versions of this model primarily used data from Wuhan, China including a high level of social distancing. Current versions of the model rely on the current social distancing actions in the US	Considers the difference in health system capacity in low- and high-income country settings	Uses current and country-specific parameters Introduces seasonality and impact of different climates
What are the limitations of the model?	Built off a model designed for the flu Does not consider additional deaths if health system becomes overwhelmed	Assumes the effects of social distancing policies are the same everywhere Reported global death counts might be unreliable and different by country Underestimates the number of daily deaths Greater uncertainty in individual states other than NY	Projections in low-income settings might in reality be substantially higher than what is predicted due to health system capacity	Model outputs are based on the assumptions and parameter choices defined by the user Does not consider the difference in social mixing patterns across countries