SCIENCE & TECH SPOTLIGHT:
COVID-19 MODELING
Accessible Version

WHY THIS MATTERS

Infectious disease models can help guide major policy decisions, such as how to allocate health care resources in the COVID-19 response. However, limitations in data inputs and assumptions can lead to considerable uncertainty in model estimates. Interpreting these estimates requires understanding their purpose, limitations, and assumptions.

/// THE SCIENCE

What is it? Infectious disease models are mathematical representations that help researchers analyze the dynamics of a disease. They have played a prominent role in the COVID-19 response in the United States and abroad, for example in projecting new infections, deaths, and the potential need for health care resources.

How does it work? The equations used in a model are based on biological knowledge, data on the disease, or both. Models can be put into two broad categories: mechanistic and statistical (fig. 1).
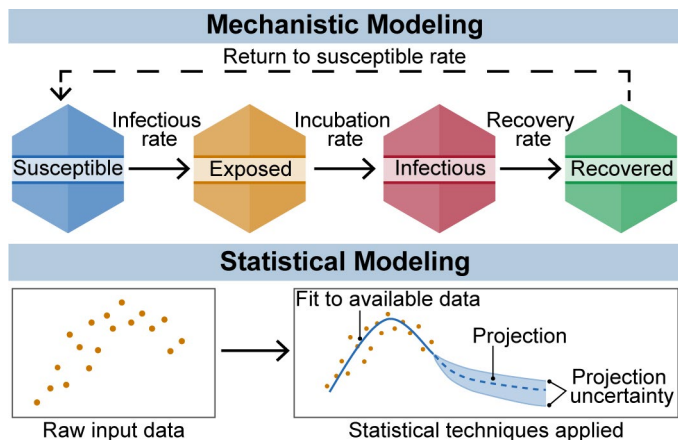
Source: GAO illustration of standard mechanistic model (Mechanistic Modeling, top); GAO illustration of standard statistical modeling (Statistical Modeling, bottom). | GAO-20-582SP

Figure 1. There are generally two broad categories of infectious disease models: mechanistic models, which use scientific understanding of disease dynamics and human behavior, and statistical models, which rely only on patterns in the data.

Mechanistic models use equations to represent the mechanics of how a disease progresses, based on scientific understanding of disease dynamics and human behavior. Researchers can estimate the effect of proposed interventions, such as social (i.e., physical) distancing, by adjusting the model and rerunning it.

One widely used mechanistic model is the susceptible-exposed-infectious-recovered (SEIR) model. It describes how a population moves through the stages of a disease. For example, movement from susceptible to exposed can be based on the rate of infections. For COVID-19, Imperial College London built an SEIR-based model that refines this rate of movement using population size, age distribution, and social contact patterns. The model projections based on this rate and other inputs can help estimate disease impacts.

Statistical models, by contrast, use only data and not disease biology or behavior. For example, they might use data on reported deaths to forecast future deaths or hospital needs.
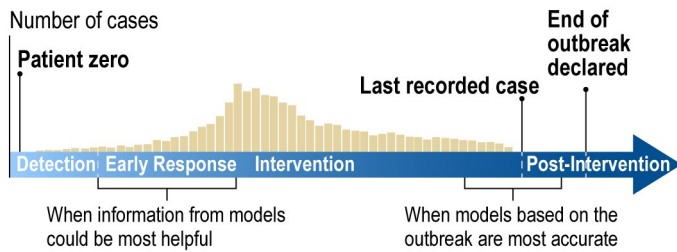
The Institute for Health Metrics and Evaluation (IHME) developed a statistical model that forecasts COVID-19 infections, deaths, hospital needs (e.g., beds or ventilators), and when states may be able to relax social distancing. To do so, researchers use data on, among other things, deaths, how the disease spread in other locations, and when states implemented social distancing.

Models may also combine statistical and mechanistic elements. For example, in May 2020, researchers added an SEIR component to the IHME model to quantify how people move through the phases of disease.

Each model has a distinct purpose that must be considered when interpreting its results. For example, statistical models may be designed to project the need for hospital beds or ventilators, based on the assumption that past trends will continue. But if human behavior changes—for instance, if social distancing is relaxed—then the forecasts are likely to be less accurate. Similarly, mechanistic models may be designed to provide insight on hypothetical scenarios, in which case their results are useful for comparing scenarios, but not as predictions.

How mature is it? The field of infectious disease modeling is well established, but the maturity of a specific model depends upon the quality, completeness, and accuracy of the data. Early in an outbreak, data and knowledge of the disease may be scarce (as is the case with COVID-19), and model predictions are unlikely to be accurate. However, as data and knowledge of the disease become more available, the model matures as its predictions can become more accurate and precise (fig. 2). Models can still be useful early in an outbreak because they can provide quantitative estimates to help decision makers respond to a range of possible scenarios. This has been the case for the COVID-19 pandemic.

Models can also be useful even when they do not provide an accurate estimate of the final results. For example, the Imperial College model explored several scenarios, including one without intervention, which projected over 2 million U.S. deaths by October 2020. This estimate is no longer useful as a prediction because that scenario did not occur, but it did prove helpful in showing the potential benefit of early intervention.

Number of cases

Patient zero

End of outbreak declared

Last recorded case

Detection | Early Response | Intervention | Post-Intervention

When information from models could be most helpful

When models based on the outbreak are most accurate

Source: GAO analysis of documents. | GAO-20-582SP

Figure 2. During the outbreak of a new disease, models can be most helpful early in the response, but are most limited by a lack of data. Later in the outbreak, more data becomes available, but there is less time to implement an optimal response for ending the outbreak.

/// OPPORTUNITIES

When used and interpreted correctly, models can provide:

- **Support for early response.** Models can inform decision-making early in an outbreak, when data and knowledge of the disease are limited. For example, they can help prepare for worst-case hospital bed needs and show which steps might best slow the outbreak.

- **Scenario comparison.** Models can provide a range of scenarios based on the potential behavior of the disease, helping policymakers prepare for multiple outcomes or identify the most likely.

- **Evaluation of options.** Models can estimate the possible impact of different policy decisions, such as relaxing social distancing or accelerating the development of a vaccine.

/// CHALLENGES

Models are only as good as the data and assumptions that go into them. The following issues may make results less precise and harder to explain:

- **Data.** Models are inherently limited by the underlying data, which can be scarce and inconsistent, especially early in a novel disease outbreak, lessening the precision and accuracy of model estimates.

- **Uncertainty.** Especially in the early stages of a novel infectious disease outbreak, there are multiple sources of uncertainty, such as overreliance on a few data points and lack of information on behavioral changes that might occur among the population.

- **Communication.** Models can be complex and thus difficult to explain to lay audiences. As a result, their methods and outputs may be oversimplified and thus misunderstood. For example, the public may assume model outputs are intended as predictions, when in some cases they illustrate what-if scenarios that may not occur.

/// POLICY CONTEXT AND QUESTIONS

The British statistician George Box once said, "All models are wrong, but some are useful." In other words, all outbreak models simplify reality but can still support decisions and improve our understanding of a disease. This context raises many policy questions, such as:

- How can modelers and policymakers collaborate to identify the most appropriate and reliable models to consider for decision-making?

- What steps could improve data quality, data sharing, and transparency for models used to inform decisions?

- How can model results be better communicated to the public?

- How can agencies further coordinate with each other and with academic and other modelers?

/// SELECTED GAO WORK

In addition to two forthcoming reports on modeling and ongoing work related to COVID-19, GAO has previously issued the following products:

- Science & Tech Spotlight: Coronaviruses, GAO-20-472SP.

- Science & Tech Spotlight: Social Distancing During Pandemics, GAO-20-545SP.

- Science & Tech Spotlight: COVID-19 Vaccine Development, GAO-20-583SP.

- Science & Tech Spotlight: COVID-19 Testing, GAO-20-584SP.

/// SELECTED REFERENCES

Dembek Z.F., T. Chekol, and A. Wu. 2018. "Best practice assessment of disease modelling for infectious disease outbreaks." *Epidemiology and Infection*, vol. 146, no. 10 (2018): pp. 1207–1215.

Department of Health and Human Services. Centers for Disease Control and Prevention. "COVID-19 Forecasts." In *Coronavirus Disease 2019 (COVID-19)*. Atlanta, Ga.: Center for Disease Control and Prevention, 2020. Accessed May 22, 2020. https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html.