VIEWPOINT

An Impact-Oriented Approach to Epidemiological Modeling

Nirav R. Shah, MD, MPH[1] ![ORCID]*, Debbie Lai[2], and C. Jason Wang, MD, PhD[3]*

[1]Clinical Excellence Research Center, Stanford University, Stanford, CA, USA; [2]Covid Act Now, Walnut, CA, USA; [3]Center for Policy, Outcomes and Prevention, Stanford University, Stanford, CA, USA.

T he COVID-19 pandemic has propelled epidemiological modeling into the public and political consciousness, beyond the strict purview of scientific and public health experts. Models have emerged as crucial tools for decision-makers, with calls for government-mandated non-pharmaceutical interventions (NPIs) such as stay-at-home orders to be based on data-driven thresholds such as case numbers and transmission rates.[1] And it goes both ways: data drives use of NPIs, which then affect models in an iterative process. Meanwhile, the outputs of COVID-19 models have become a subject of public fixation and mainstay of media headlines.

There is a growing body of evidence supporting the efficacy of NPIs such as shelter-in-place and mask-wearing, which are affected by the extent of the public's buy-in and compliance. Studies have shown that NPIs averted a 67× increase in cases in China by February 29, 2020, and even lax compliance can reduce transmission by as much as 25%.[2] Other studies suggest that suppression will minimally require social distancing by the entire population.[3] Under such circumstances, public awareness and consensus become paramount, particularly in the USA, where societal and cultural norms may limit imposed lockdowns akin to those that occurred in Wuhan and other parts of China.

Thus, there emerges an unprecedented need to build a shared understanding of the disease, not just among experts and policymakers but also for the public. Those who develop epidemiological models are no longer only creating specialty tools, but consumer products as well, and thus face a new, non-traditional, set of considerations. We propose that this requires an impact-oriented approach, i.e., what is the cumulative *impact* of their models upon the public? We call this impact-oriented modeling.

Traditionally, epidemiological models have been valued for their ability to inform decision-makers who possess prior knowledge of disease management.[4] In the wake of the H1N1 pandemic in 2009, the World Health Organization (WHO) convened a mathematical modeling network of public health experts and academics.[5] The Centers for Disease Control and Prevention (CDC) recently added policy development as a sixth item in its list of the major tasks of epidemiology in public health, but there remains no mention of the impact on the general public.[6]

Impact-oriented modeling values more than accuracy, which remains non-negotiable. Beyond simply the outputs of such a model, consideration must be given to the presentation of these outputs, including design, visualization, and supporting content, all of which affect the utility, user experience, downstream policy, and, ultimately, *impact*.

To this end, we outline a set of 8 key considerations for impact modeling. Though these eight considerations will not be easily met in totality, we recommend incorporating as many possible into modeling for the COVID-19 pandemic (Table 1).

1. Agility: Is the data and model providing timely information? The fast-changing nature of COVID-19 highlights the need for models to reflect the most recent information, which may differ drastically from recent, prior information. With COVID-19, journalists have become an active source not only of news but also of data. The *New York Times*' repository of COVID-19 cases (available at https://github.com/nytimes/covid-19-data), collected by reporters who monitor news conferences, analyze data releases, and seek clarification from public officials, is updated daily and is among the best sources of this fundamental metric.

2. Responsiveness: Do the data and model respond to new evidence? Not only do models allow the public and decision-makers to react to data, but the models themselves should also react to data in an iterative fashion. A feedback loop of action-information-reaction should drive models to continuously evolve, along with COVID-19 and our knowledge of it. For instance, on May 4, 2020, over five weeks after it first launched on March 26, the Institute for Health Metrics and Evaluation (IHME) pivoted from a poorly-performing curve-fitting model drawing on prior death reporting, to a traditional SEIR model (available at https://covid19.

Table 1 Considerations for Impact-Oriented COVID-19 Modeling

	Considerations	Description
1	Agility	Is the data and model providing timely information?
2	Responsiveness	Does the data and model respond to new evidence?
3	Transparency	Are the data and model's mechanisms and data sources publicly available for fact-checking and validation?
4	Usability	Can the data and model be used easily, effectively, and efficiently?
5	Accessibility	Can the data and model be understood and used by a broad audience, irrespective of scientific, technical, and other capabilities?
6	Universality	Does the data and model draw on inputs that are defined and measured consistently?
7	Adaptability	Can the model be easily modified and adapted?
8	Actionability	Are there clear calls-to-action that reflect current government policies?

healthdata.org/), which led to a substantial increase in forecasted COVID-19 deaths and more accurate outputs.

3. Transparency: Are the data and model's mechanisms and data sources publicly available for fact-checking and validation? This issue has already been raised in the field of machine learning, where the plethora of options likewise renders the task of selection difficult. In the absence of perfect knowledge and the presence of myriad approaches, open-source models and databases enable users of these models to make more informed choices between models and data sources. They also enable the validation of models and data sources, which is critical not only for verifying the accuracy, but for enabling iteration and improvement. For instance, the Covid Act Now (CAN) model is fully open-source, along with its data inputs (available at https://covidactnow.org). The mechanisms of its models, its assumptions, and references, are made publicly available.[6] This enables the public and experts to escalate questions and concerns that have enabled the model to be refined, such as by ingesting more accurate data.

4. Usability: Can the data and model be used easily, effectively, and efficiently? Intuitively, we know that when users are not able to easily access and use a product, they are less likely to continue using it. Developers of consumer products are thus familiar with the need to consider user expectations, desires, and requirements.[7] COVID-19 models may benefit from doing the same. For example, user research, a common component in the development of consumer products, may become increasingly important in order to better understand the barriers that prospective users of epidemiological models face.

5. Accessibility: Can the data and model be understood and used by a broad audience, irrespective of scientific, technical, and other capabilities? The majority of the USA has not received training in epidemiology or data science. Elderly populations that are more vulnerable to

infection typically have less experience using technology. As progress containing the virus depends on the cumulative behavior of millions of individuals, a broad understanding of a model results in success or failure, and hence, models must use language and visuals that forgo specialized jargon and excessive complexity.

6. Universality: Do data and the model draw on inputs that are defined and measured consistently across geographies? Given the unprecedented nature of COVID-19, countries, states, counties and cities depend upon learning from each other, and what happens across artificial political boundaries matters across a region. Standardization and consistency of data across regions can enable this. For example, the COVID Tracking Project is an open-source initiative of *The Atlantic* and provides one of the most complete data sets available about COVID-19 in the USA (available at https://covidtracking.com/).

7. Adaptability: Can the model be modified and adapted? In particular, efforts to provide useful COVID-19 data for the USA have run into the following quandary: even as the implementation of tactical strategies exists primarily at the local level, it is also at the local level that the big data required to feed epidemiological models becomes most difficult to obtain. It may be that the models most easily customized by cities and counties will be those that have the greatest impact. The COVID-19 Hospital Impact Model for Epidemics (CHIME) model allows for custom inputs, such as estimates of the regional population, hospital market share, and currently hospitalized COVID-19 patients, in order to assist local officials with hospital capacity planning (available at https://penn-chime.phl.io/).

8. Actionability: Does the model reflect current government policies? Given the role of epidemiological models in shaping public discourse and behavior, there is a responsibility to also inform actionability. Models that fail to do so may contribute to anxiety, confusion, or even actions that violate federal, state, or local regulations. On the other hand, models that clearly communicate the actionable implications of their outputs can contribute to a positive rather than a negative impact. Both the New York Times and Georgetown University's Center for Global Health, Science, and Security (available at https://covidamp.org/) have begun to collect data on COVID-19 policies by state and effective dates, including shelter-in-place and reopening orders.

To our knowledge, no data source or model currently fulfills all the considerations that we have set forth. These eight considerations may enable COVID-19 data and models to become better harbingers of actionable, behavior-changing, and even life-saving information; to bridge the gap between scientific public health expertise and mainstream, layperson

knowledge; and to generate more positive impact than noise. As the British statistician George Box said, "All models are wrong. But some are useful."

Corresponding Author: *Nirav R. Shah, MD, MPH; Clinical Excellence Research Center, Stanford University, Stanford, CA, USA (e-mail: nirav.shah@stanford.edu).*

Compliance with Ethical Standards:

REFERENCES

1. **Holmdahl I, Buckee C.** Wrong but useful — what Covid-19 epidemiologic models can and cannot tell us. N Engl J Med. 2020;15. https://doi.org/10.1056/NEJMp2016822.
2. **Lai S, Ruktanonchai NW, Zhou L, et al.** Effect of non-pharmaceutical interventions to contain COVID-19 in China. Nature. 2020. https://doi.org/10.1038/s41586-020-2293-x.
3. **Ferguson NM, Laydon D, Nedjati-Gilani G, et al.** Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial Coll Lond. 2020. https://doi.org/10.25561/77482. Available at https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-03-16-COVID19-Report-9.pdf. Accessed May 30, 2020.
4. **Rosenfeld R, Grefenstette J, Burke D.** A Proposal for Standardized Evaluation of Epidemiological Models. Developing the Theory and Practice of Epidemiological Forecasting (DELPHI), Carnegie Mellon University; 2018. https://delphi.midas.cs.cmu.edu/files/StandardizedEvaluation_Revised_12-11-09.pdf. Accessed July 18, 2020.
5. World Health Organization. Mathematical modelling of the pandemic H1N1 2009. Wkly Epidemiol Rec. 2009; 84:341-8.
6. Principles of Epidemiology in Public Health Practice, Third Edition: An Introduction to Applied Epidemiology and Biostatistics. The Centers for Disease Control and Prevention. Published on May 18, 2012. https://www.cdc.gov/csels/dsepd/ss1978/lesson1/section4.html. Accessed July 18, 2020.
7. **Mittler JN, Volmar KM, Shaw BW, Christianson JB, Scanlon DP.** Using websites to engage consumers in managing their health and healthcare. Am J Manag Care. 2012;18(6 Suppl).