Themed Section: COVID-19

Where Do We Go From Here? A Framework for Using Susceptible-Infectious-Recovered Models for Policy Making in Emerging Infectious Diseases

Roy S. Zawadzki, Cynthia L. Gong, PharmD, PhD, Sang K. Cho, PharmD, PhD, Jan E. Schnitzer, MD, Nadine K. Zawadzki, Joel W. Hay, PhD, Emmanuel F. Drabo, PhD

A B S T R A C T

Objectives: Throughout the coronavirus disease 2019 pandemic, susceptible-infectious-recovered (SIR) modeling has been the preeminent modeling method to inform policy making worldwide. Nevertheless, the usefulness of such models has been subject to controversy. An evolution in the epidemiological modeling field is urgently needed, beginning with an agreed-upon set of modeling standards for policy recommendations. The objective of this article is to propose a set of modeling standards to support policy decision making.

Methods: We identify and describe 5 broad standards: transparency, heterogeneity, calibration and validation, cost-benefit analysis, and model obsolescence and recalibration. We give methodological recommendations and provide examples in the literature that employ these standards well. We also develop and demonstrate a modeling practices checklist using existing coronavirus disease 2019 literature that can be employed by readers, authors, and reviewers to evaluate and compare policy modeling literature along our formulated standards.

Results: We graded 16 articles using our checklist. On average, the articles met 6.81 of our 19 categories (36.7%). No articles contained any cost-benefit analyses and few were adequately transparent.

Conclusions: There is significant room for improvement in modeling pandemic policy. Issues often arise from a lack of transparency, poor modeling assumptions, lack of a system-wide perspective in modeling, and lack of flexibility in the academic system to rapidly iterate modeling as new information becomes available. In anticipation of future challenges, we encourage the modeling community at large to contribute toward the refinement and consensus of a shared set of standards for infectious disease policy modeling.

Keywords: policy, COVID-19, SIR modeling, health services research, epidemiology, cost benefit.

VALUE HEALTH. 2021; 24(7):917–924

Introduction

Mathematical models have been critical for developing policies to mitigate the impact of coronavirus disease 2019 (COVID-19), as in past pandemics.[1-3] Specifically, susceptible-infected, recovered (SIR) models have been widely used to develop policy recommendations for COVID-19. Nevertheless, predictions and policy recommendations from these models have not aligned well with the empirical data, leading to lasting uncertainty over the basic characteristics of the pandemic and criticism of the practice.[4-6]

On July 22, 2020, National Institute of Allergy and Infectious Disease director Anthony Fauci warned that COVID-19 is unlikely to be eradicated[7]; SIR policy modeling is therefore crucial to inform evidence-based mitigation strategies. Dynamism in pandemic modeling is more pressing than ever for fear of the public and policy makers losing confidence in science to help inform policies. An emerging pandemic that affects multiple sectors of society necessitates a systems-science approach to adequately evaluate short- and long-term impacts as well as the cost-benefit tradeoffs between different containment policies. Though not generally captured by SIR models, policy effectiveness is modulated by heterogeneities in health and economic participation of target populations, individual compliance, and private choices related to risk perception.

In our own research developing and evaluating COVID-19 models, we found it difficult both to appraise the quality of SIR models in the literature and to compare the differing and black-box results of such models. Potential frameworks for such policy modeling (eg, the ISPOR report on dynamic transmission or HPV-FRAME) are designed for and by their respective disciplines.[8,9] Nevertheless, these frameworks are limited in their ability to guide policy makers and the larger academic community at the complex, multifaceted, and unprecedented societal scale that COVID-19 has warranted. Thus, our multidisciplinary team of close

collaborators, including health economists, data scientists, epidemiologists, and clinicians, many of whom regularly evaluate health policy, convened on an ad hoc basis, to provide insights into how to close the current gap in standards and begin to create a baseline framework for widespread pandemic policy modeling.

Here, we define a set of 5 standards to increase the utility of SIR-based policy modeling (Table 1). Though not a panacea, we intend to steer the conversation toward a consensus of how we ought to improve epidemic modeling in light of lessons learned from COVID-19 to support policy making during subsequent waves and other pandemics.

For use on both preprinted and peer-reviewed articles, we develop an SIR policy modeling grading checklist that authors, readers, and reviewers can use to appraise the robustness and usability of model findings along the axes provided in Table 1. To demonstrate this checklist in action, we review peer-reviewed COVID-19 SIR policy models and share the results. The remainder of this article details our standards by using the process of developing SIR models to exemplify common problems that are present in such modeling, share specific methodology that can mitigate these issues, and offer examples in the literature that reflect our proposed practices.

Methods

A multidisciplinary team of health economists, data scientists, epidemiologists, and clinicians was convened on an ad hoc basis and in an unstructured focus group discussion to develop a 19-question modeling checklist (Table 2) that reflects the practices proposed in Table 1. To illustrate the usage of the checklist, we evaluate a selection of COVID-19 SIR modeling papers, identified through a PubMed literature search completed on July 18, 2020. Search terms included were "covid AND non-pharmaceutical interventions," and restricted to English-language journal articles that were original modeling research (no literature reviews, commentaries, etc.). Two authors graded each article according to the checklist; each reviewer's grading was then cross-checked by the other reviewer. Finally, a combined set of grades was created and the number of standards met was tallied across articles.

We caution that repeating a search with the same criteria may not produce our exact results owing to factors such as frequency of indexing of certain journals and articles published online ahead of print (ePub) given retroactive entry dates on PubMed.[11,12] We thus include more details regarding our search in the results section. We thank the reviewers of the manuscript for prompting this reproducibility exercise.

Results

Our search criteria yielded 16 articles (Fig. 1); detailed information on how these were chosen can be found in Appendix Table 1 (see Appendix Table 1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.03.005). The final list of articles and checklist grading results are shown in Appendix Table 2 (see Appendix Table 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.03.005). The average article score (%) was 6.81 of 19 (36.7%) (SD = 3.23 [17.6%]), with the best article scoring 13 of 19 (68.4%) and the worst scoring 1 of 19 (5.26%). Most authors clearly defined their research questions and study objectives (13/16 [81.3%]). Eleven articles (68.8%) clearly and thoroughly stated calibration assumptions. Only 5 articles (31.3%) adequately addressed or adjusted for biases in the data. In addition, only 4 articles (25%) included calibration parameters that accommodated heterogeneity in disease susceptibility, such as

age, while no articles explicitly accounted for economic participation and individual risk-taking. No article contained cost-benefit tradeoffs of alternative policies. Surprisingly, 5 articles (31.3%) did not publish their model code in open source. Two articles (12.5%) conducted cross-validation, and 3 articles (18.8%) conducted external validation. Overall, our findings suggest that there is significant room for improvement in the literature per our standards, though we acknowledge that this selection of articles is not representative of all epidemiological modeling policy-making literature.

Discussion

Formulation of Research Questions and Hypotheses

In rapidly evolving research environments like the current pandemic, research questions and hypotheses are often hastily and poorly defined, making model findings difficult to interpret and use. For example, if a model attempts to quantify total COVID-19 deaths as a function of some policy, what often remains ambiguous is whether the findings are meant to inform public health policy or simply convey the characteristics of the disease (eg, differential mortality rates by age). If the intent is to inform policy, focusing solely on COVID-19 deaths is misleading and biased, because it ignores additional costs and non-COVID-19 deaths that are affected by the policy of interest. Without a properly defined intent and scope, the level of standards and scrutiny that should be applied is unclear.

Virtually all studies are based on prior research but in emerging pandemic situations, authors rarely describe whether the current evidence is appropriate for the situation at hand. For example, a well-cited SIR model used influenza-based hypotheses to inform COVID-19 policy models, which proved incongruous for key COVID-19 characteristics such as virulence and mortality age distribution.[13,14] In another example, the initial COVID-19 literature from China was neither cross-validated nor critically evaluated by outside researchers before their models and findings were taken as fact by researchers and policy makers.

Before undertaking an analysis, researchers must transparently discuss whether the presently available information is reliable enough to form a tenable hypothesis. If not, they should explain how they plan to generate more robust and defensible conclusions based on their hypotheses. Such practices will anchor the hypotheses of future studies to more robust current research while also ensuring that spurious hypotheses are not tested.

Addressing and Adjusting for Data Issues

In the early stages of a pandemic, sampling biases often result in collecting data among severely symptomatic cases, thus skewing fatality and case-testing rates. For COVID-19, we now know that initial statistics inflated mortality rates because many nonfatal and asymptomatic cases were not yet measured.[15] There are also biases in test accuracy; different reverse transcription polymerase chain reaction and antibody tests have varying sensitivity and specificity rates. Furthermore, not all labs use identical tests, procedures, or testing criteria, leading to heterogeneity in the data.

Data unreliability can similarly influence COVID-19 death statistics: under a limited number of available tests, there is selection bias among whether autopsies are tested for COVID-19. Additionally, death coding standards and reporting lags vary by locale. Lack of strict international standards for coding COVID-19 cause of death combined with the role of underlying conditions can result in highly variable death attribution. In the United States, the US

Table 1. List of modeling standards.

Standard	Rationale	Implementation	Considerations
Transparency	• Crucial because model assumptions cause variability in epidemic prediction • Clarifies whether model is designed for scientific inquiry or policy recommendation • Ensures models can be understood, reproduced, and tested • Increases credibility of results and decreases erroneous results persisting in literature	• Fully disclose all model assumptions • Include range of uncertainty in model parameters • Perform sensitivity analyses for assumptions (eg, selection bias) • Clearly define hypotheses model was designed to test • Specify intended use of model findings • Publish all code and data in open source upon publication, if possible	• Not always clear whether sensitivity analyses or uncertainty intervals are realistic and useful • Authors may not want to share proprietary methodology or intellectual property • Thorough model documentation adds time to manuscript development
Heterogeneity	• Exists in the epidemiological impacts of the disease, efficacy of policy, and costs of policy • Allows accounting for differences in disease susceptibility, economic participation (eg, essential workers), and individual risk-taking, which may also vary over time • Ignoring heterogeneity risks recommending inequitable policies	• Use heterogeneous, time-varying, and population-dependent modeling parameters • Include parameters on compliance to policy, economic participation, and need for COVID-19-specific medical resources	• Difficult to source current timely demographic and socioeconomic data; historical data may no longer be accurate • Disease heterogeneity is unknown in early stages of pandemics; assume worst case until proven otherwise • Difficult to accurately characterize individual risk-taking and compliance behaviors
Calibration and validation	• Calibration determines best fit of parameters to observed data • Fitted parameters form the basis for validation and policy analysis • Validation establishes model structural and behavioral validity[10] • Ensures consistency of model predictions with observed epidemic dynamics, assumptions, and causal linkages • Enhances model confidence and reduces revision	• Cross-validation: reproduce literature results with one's modeling process but literature parameters • External validation: apply model outside original context (eg, another country) or past epidemics	• Cross-validating results requires literature to be transparent and reproducible in the first place • Testing of external validity limited by data set availability, could be difficult to acquire for intellectual property and political reasons
Cost-benefit analysis	• Effective policy analysis evaluates full societal cost-benefits, not just pandemic-specific outcomes • Examines the influence of critical policy parameters on both epidemiological and economic outcomes	• Use epidemiological-economic models in policy analysis • Evaluate cost-benefit beyond traditional economic indicators (eg, unemployment) by including relevant downstream costs • Address potential inequity in policies (see: Heterogeneity)	• Recent data related to direct and in-direct costs largely does not exist, is inaccessible, or is proprietary • Nontraditional cost measurements (eg, mental health) difficult to measure

Table 1. Continued

Standard	Rationale	Implementation	Considerations
Model obsolescence and recalibration	• New information and policies arise constantly in emerging epidemic situations, leading to rapid irrelevance of model results over time • Consumers of literature may cite a no longer useful model • Authors have responsibility to ensure valid models past publication • The current publication environment does not accommodate rapid iteration of models • Model corrections are interpreted as errors in the peer-review process but model obsolescence in developing situations is inevitable • This publishing environment may cause corrections to be withheld, stymieing progress in developing situations	• Ex-post validation: authors routinely validate results for accuracy after publication • If results diverge from current data, disclose reasons, and, if feasible, correct the model • Journals and readers accept model corrections and updates as inevitable and useful • Journals allow authors to submit a short communication regarding modeling updates	• Models are validated only at the frequency that the authors check • Frequency and workload of validation dependent on the authors • Standard penalizes authors who lack resources for continual validation and updates (time, personnel, or funding) • Extra resources from journals are needed to publish model updates

COVID-19 indicates coronavirus disease 2019.

Centers for Disease Control and Prevention leaves much of the interpretation of a COVID-19 death to the physician, allowing probable COVID-19 deaths to be coded as confirmed COVID-19 deaths.[16]

Researchers, especially those informing policy, must transparently discuss bias caused by testing criteria and death-coding standards in their data. Studies should include detailed sample demographics, including information that can contribute to data biases such as underlying conditions, age, and socioeconomic status. If appropriate, analysis performed on observational data should be augmented by sensitivity analyses to measure and mitigate potential sources of bias such as an E-Value analysis.[17,18] These issues can be mitigated by selecting more robust outcomes than raw counts of cases and deaths (eg, measuring the first difference of cases or deaths to net out time-invariant confounders such as death-reporting lags or diagnostic testing standards).[19]

Choosing Model Parameters and Calibration

Data biases notwithstanding, SIR models are characterized by high uncertainty before the inflection point of the disease outbreak is reached, leading to volatile results.[20–22] These models rely on exponential distributions, thus parametrical errors made early in an epidemic can result in massive under- or over-predictions of future measured outcomes.

Authors must fully discuss their assumptions when calibrating model parameters, source calibration parameters from real-world evidence and adjusted for bias, and include appropriate uncertainty intervals around each parameter. Researchers should include the full list of calibration parameters for grid search and the results of their cross-validation procedure to choose the final parameters, such as a calibration checklist.[23,24]

In selecting model parameters, researchers should not assume population homogeneity regarding susceptibility, spreading of the disease, viral shedding, participation in the economy (eg, essential versus nonessential workers), and risk-taking. In addition, there are marked differences in race and income level between those working essential jobs and those working from home.[25] Without considering these factors, models can overestimate key metrics such as medical resources and create inequitable policy.[26,27] Importantly, each of these factors contributing to heterogeneity is time varying and dependent on previously enacted policy. As a result, key parameters, like $ß_t$, the transmission rate, should be made time-varying and dependent on individual behaviors, as well as the effectiveness of a given policy.

Models should not simply assume that policy implementation leads to full individual compliance—such assumptions can confound the estimated effectiveness of policy mandates. For example, an analysis by FiveThirtyEight found that many stayed home before official shelter-in-place mandates were passed.[28] A survey found that only 44% of Hawaiians were practicing social distancing.[29]

Models should account for observed compliance rates or, at minimum, conduct sensitivity analyses with varying levels of individual compliance. Compliance can be measured using survey data or anonymized cell phone mobility data.[30,31]

Understanding Findings Through Cost-Benefit

It is misleading for models to use total number of predicted COVID-19 deaths as the sole metric of an intervention's success; many have voiced concerns regarding this one-dimensional evaluation of widespread lockdown mandates.[32-34] Although total predicted death count can quantify survival benefits, few

Table 2. Checklist of good SIR policy modeling practices.

Standard	Yes	No
Clearly defined research questions and study objectives		
Clearly defined study population		
Clearly stated who should use findings (eg, policymakers)		
Usefulness of study hypotheses contextualized against current literature		
Adjustments for potential data biases and/or discussion of this in limitations		
Clearly and thoroughly stated calibration process assumptions		
Detailed calibration grid search process description (ie, a calibration checklist)		
Calibration parameters explicitly include range of uncertainty		
Calibration parameters allow for time variation		
Calibration parameters accommodate heterogeneity in disease susceptibility (eg, by age or pre-existing conditions)		
Calibration parameters accommodate heterogeneity in economic participation and individual risk-taking		
Calibration assumptions tested via sensitivity analysis		
The policy or treatment variable analyzed at the individual level (not the macro-policy level)		
Includes a cost-benefit analysis for policies		
Cost-benefit criteria include metrics beyond traditional economic indicators (see text for more details)		
Calibration process validated with the parameters from other papers (cross-validity)		
Model code is available open source		
Modeling process applied to situations outside the immediate modeling context (external validity)		
After publication, authors provide updates on present model validity		

SIR indicates susceptible-infectious-recovered.

models evaluate other benefits and costs. If existing costs are ignored, policy makers will unknowingly select suboptimal, nearsighted interventions that superficially appear highly beneficial by predicting the fewest COVID-19 deaths but are so costly they ultimately result in a net loss to society. These include decreases in quality of life; delayed screenings, treatments, and vaccines; financial hardships; food insecurity; hospital insolvency; and stressed societal inequities.[35–40] Special attention should be given to modeling the impacts of interventional policies on children and young adults, because their development heavily influences and predicts their future health, education, and employment outcomes.[41,42]

Economic epidemiology is a growing field that combines SIR models with economic outcomes while considering heterogenous populations and age-differential risk-taking. Acemoglu et al focus on finding an optimal balance of policy between economic loss and COVID-19 deaths by constructing a Pareto efficient frontier curve.[43] This methodology improves standard SIR models by comprehensively framing the effects of a policy. Many similar analyses exist both in the context of COVID-19 and past epidemics,[44–46] yet these models are not often used by policy makers. Econonomic-epidemiological models should be the gold standard for measuring the impacts of emerging infectious disease policy. Kim and Neumann describe a diverse set of axes to be considered by such cost-benefit analyses.[32]

Validation of Modeling Results

Model results used for policy making must be thoroughly validated. Validation should encompass 3 levels: (1) internal model validation, (2) third-party cross-validation, and (3) external validation. This section will cover cross-validation and external validation. For internal validity, please refer to sections "Addressing and Adjusting for Data Issues" and "Choosing Model Parameters and Calibration."

Third-party (cross) validation, in which authors apply parameters from similar models and verify that results remain consistent despite an alternative modeling approach, should be performed to improve the clarity and reproducibility of models and mitigate costly mistakes from errors. For example, Drabo et al validate their model with parameters derived from a similar SIR cost-benefit analysis in the literature.[46] If results do not match, the reasons should be further investigated and described in the article. To facilitate cross-validation, published models should be transparent and/or open sourced. Peng et al provide a useful checklist for reproducible epidemiological models.[47] Code can be shared using computational notebooks such as Jupyter notebooks and services like GitHub.

Complex models may overfit the idiosyncrasies of the input data and thus may have limited generalizability or external validity. Researchers should apply the entire modeling process to other locations and points to verify whether their methodology maintains accuracy. In addition, researchers can apply their methods to past pandemics using only the data that were available at a similar point of that epidemic. For example, if a scientist wanted to construct an SIR model for COVID-19 when the pandemic was at an early stage, he or she should validate the model using only the data that were available when another pandemic (eg, severe acute respiratory syndrome) was at a comparable stage.

Real-Time Validity and Modeling Improvement in an Emerging Pandemic

In emerging pandemics such as COVID-19, published models can quickly become outdated and unrepresentative of the real world owing to new discoveries, new data, or new implemented

Figure 1. Literature search tree.
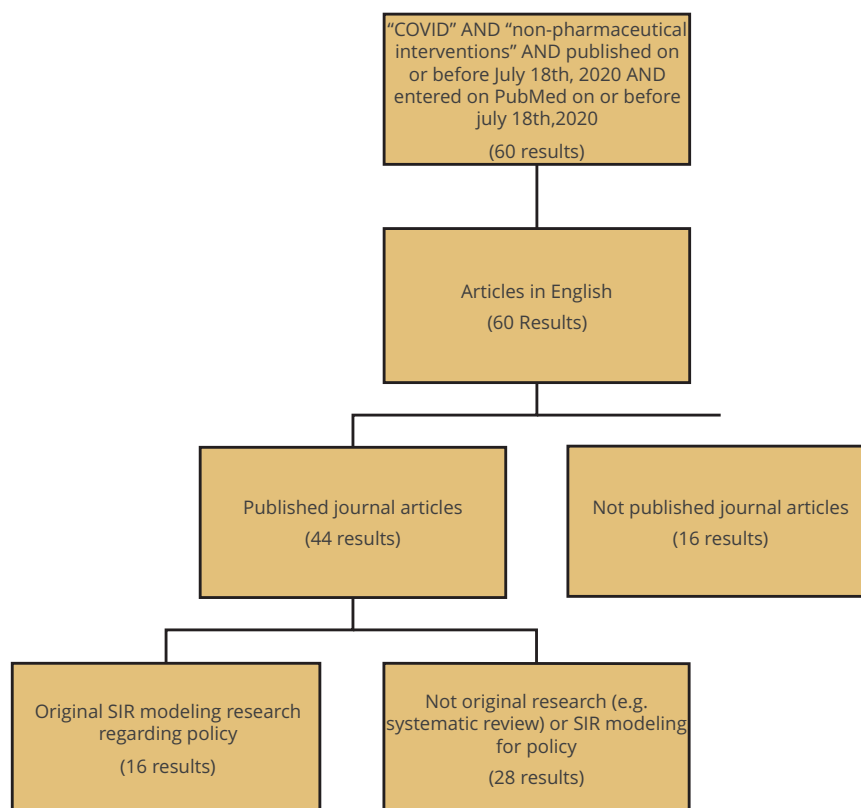
COVID-19 indicates coronavirus disease 2019; SIR, susceptible-infectious-recovered.

policy. Such events can be impactful; researchers and policy makers examining the literature may fail to question the relevancy of a model in light of new data or policies, making it more likely to misinterpret and misuse the model's findings.

Researchers have an onus to routinely confirm whether their model results remain accurate as the time elapsed since publication increases (ie, ex-post validation). The frequency at which researchers ex-post validate is subject to the current conditions. If there is a divergence in fit, the authors should disclose their suspicions as to why this occurred (eg, death coding standards changed). Better yet, they should develop a flexible model infrastructure to accommodate issues if/as they arise and/or keep models up to date (eg, an updating dashboard).

Currently, incentives for the aforementioned practices are low. Corrections are often regarded as failures in the scientific method and peer-review process as well as reputation-tarnishing. Nevertheless, in the context of a rapidly evolving research environment like a pandemic, reasonable assumptions and conditions change constantly, potentially rendering an SIR model inaccurate. Rather than criticizing those who regularly and reliably update their work, which discourages authors from innovating their models, the scientific community should accept the changing of models as a scientific inevitability.

One solution benefitting both authors and journals is permitting authors to submit short communications with model updates. If readers can be confident a given model is up-to-date and valid at any time of access, they are more likely to continue to reference it many months after its publication. Authors will receive additional citations on their short update if it includes information that can assist other modelers with similar issues. For example, the National Bureau of Economics Research has a "Working Paper Series"

where authors keep a public "change log" as they update their articles based on feedback.

Conclusion

When used carefully and with full acknowledgement of all limitations, SIR modeling can be an invaluable tool to understand novel epidemic situations and suggest rational policy. The numerous COVID-19 modeling issues in the past year and our informal literature review suggest that there is significant room for improvement. This article attempts to outline key issues in epidemiological modeling for policy making and suggest potential solutions in the literature.

We acknowledge that it is far easier to point out solutions than it is to implement them in tumultuous times. Processes and sets of standards like those presented here must be established before these events to facilitate optimal policy. It may not be immediately possible to check every box in our list of standards, but the implications and associated costs that rely on such modeling cannot be overstated. The standards and practices shared in this article represent a starting point for developing an agreed-upon set of standards similar to other fields such as the International Society for Pharmacoeconomics and Outcomes Research's Consolidated Health Economic Evaluation Reporting Standards and we call for more formal work to address these issues.[48]

In addition to wider community input, our suggested framework could benefit largely from a more quantitative and precise version of the evaluation tool seen in Table 2. Rather than a simple binary yes or no, the checklist could include a Likert scale response for each question. Such changes would expand the

present applicability of our standards as a simple first-pass check for modelers, readers, and reviewers to a tool that can be used for more rigorous systematic reviews and meta-analyses.

The evolution of emerging epidemic modeling must be embraced by all: authors, readers, and publishers. Authors should make transparent efforts to prevent the issues raised that affect SIR modeling, readers should be vocal when models are being misused or when they can be improved, and publishers should uphold an environment where rapid iteration and revision is encouraged. The next large disease outbreak is only a matter of time. Fundamentally, modelers should be asking themselves what they can do today to be ready for the epidemiological challenges of tomorrow.

Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2021.03.005.

Article and Author Information

Author Affiliations: Department of Statistics, Donald Bren School of Information and Computer Sciences, University of California, Irvine, CA, USA (Zawadzki); Fetal and Neonatal Institute, Division of Neonatology, Department of Pediatrics, Children's Hospital Los Angeles, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA (Gong); College of Pharmacy, University of Houston, Houston, TX, USA (Cho); Proteogenomics Research Institute for Systems Medicine (PRISM), San Diego, CA, USA (Schnitzer); Schaeffer Center for Health Policy & Economics, School of Pharmacy, University of Southern California, Los Angeles, CA, USA (Zawadzki, Hay); Department of Health Policy and Management, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA (Drabo).

Correspondence: Cynthia L. Gong, PharmD, PhD, Department of Pediatrics, University of Southern California, Division of Neonatology, Children's Hospital Los Angeles, 4650 Sunset Blvd, MS #31, Los Angeles, CA, USA 90027. Email: gongc@usc.edu

REFERENCES

1. Gupta S, Anderson RM, May RM. Mathematical models and the design of public health policy: HIV and antiviral therapy. *SIAM Rev.* 1993;35(1):1–6.
2. Fisman D, Khoo E, Tuite A. Early epidemic dynamics of the West African 2014 Ebola outbreak: estimates derived with a simple two-parameter model. *PLoS Curr.* 2014;6.
3. Ferguson NM, Cummings DA, Fraser C, Cajka JC, Cooley PC, Burke DS. Strategies for mitigating an influenza pandemic. *Nature.* 2006;442(7101):448–452.
4. Ioannidis JP, Cripps S, Tanner MA. Forecasting for COVID-19 has failed. *Int J Forecast.* 2020. https://doi.org/10.1016/j.ijforecast.2020.08.004.
5. Holmdahl I, Buckee C. Wrong but useful—what covid-19 epidemiologic models can and cannot tell us. *N Engl J Med.* 2020 May 15;383:303–305.
6. Jin J, Agarwala N, Kundu P, Wang Y, Zhao R, Chatterjee N. "Transparency, Reproducibility, and Validation of COVID-19 Projection Models." Johns Hopkins Bloomberg School of Public Health Expert Insights. 2020. www.jhsph.edu/covid-19/articles/transparency-reproducibility-and-validation-of-covid-19-projection-models.html. Accessed June 26, 2020.
7. Lovelace Berkeley. "Dr. Anthony Fauci Warns the Coronavirus Won't Ever Be Eradicated." CNBC. 2020 June 22. www.cnbc.com/2020/07/22/dr-anthony-fauci-warns-the-coronavirus-wont-ever-be-totally-eradicated.html. Accessed July 23, 2020.
8. Pitman R, Fisman D, Zaric GS, et al. Dynamic transmission modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–5. *Value Health.* 2012;15(6):828–834.
9. Canfell K, Kim JJ, Kulasingam S, et al. HPV-FRAME: A consensus statement and quality framework for modelled evaluations of HPV-related cancer control. *Papillomavirus Res.* 2019;8:100184.
10. Barlas Y. Formal aspects of model validity and validation in system dynamics. *Syst Dyn Rev.* 1996;12(3):183–210.
11. Nahin AM, Tybaert SJ. Why Citations to Older Articles May Display Before More Recent Ones in PubMed®. *NLM Tech Bull.* 2002;325:e3.
12. Help - PubMed. (2020, November 2). https://pubmed.ncbi.nlm.nih.gov/help/#edat. Accessed January 6, 2020.
13. Adam D. Special report: The simulations driving the world's response to COVID-19. *Nature.* 2020;580(7803):316.
14. "Similarities and Differences between Flu and COVID-19." Centers for Disease Control and Prevention, 2020 July 10. www.cdc.gov/flu/symptoms/flu-vs-covid19.htm.
15. Gandhi M, Yokoe DS, Havlir DV. Asymptomatic transmission, the Achilles' heel of current strategies to control COVID-19. *N Engl J Med.* 2020;382:2158–2160.
16. Schwartz S. "New ICD Code Introduced for COVID-19 Deaths." National Center for Health Statistics, CDC. www.cdc.gov/nchs/data/nvss/coronavirus/Alert-2-New-ICD-code-introduced-for-COVID-19-deaths.pdf?fbclid=IwAR0-aMmi5bjnfXwvUZDa8iv6sTPEyXu. Accessed June 4, 2020.
17. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. *Ann Intern Med.* 2017;167(4):268–274.
18. Drabo Emmanuel F, Kang So-Yeon, Gong Cynthia L. Guarding against seven common threats to the credible estimation of COVID-19 policy effects. *Am J Public Health.* 2020;110:1724–1725.
19. Wooldridge JM. *Econometric analysis of cross section and panel data.* MIT press; 2010:279–291.
20. Yang W, Zhang D, Peng L, Zhuge C, Hong L. Rational evaluation of various epidemic models based on the COVID-19 data of China. arXiv preprint arXiv: 2003.05666. 2020 Mar 12.
21. Magal P, Webb G. The parameter identification problem for SIR epidemic models: identifying unreported cases. *J Math Biol.* 2018;77(6-7):1629–1648.
22. Chowell G. Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts. *Infect Dis Model.* 2017;2(3):379–398.
23. Hazelbag CM, Dushoff J, Dominic EM, Mthombothi ZE, Delva W. Calibration of individual-based models to epidemiological data: A systematic review. *PLoS Comput Biol.* 2020;16(5):e1007893.
24. Stout NK, Knudsen AB, Kong CY, McMahon PM, Gazelle GS. Calibration methods used in cancer simulation models and suggested reporting guidelines. *Pharmacoeconomics.* 2009;27(7):533–545.
25. Kearney M, Pardue L. *Exposure on the Job: Who are the Essential Workers Who Likely Cannot Work from Home?* Brookings Institution Report; 2020. Available at: https://www.brookings.edu/research/exposure-on-the-job/. Accessed June 4, 2020.
26. Kashyap S, Gombar S, Yadlowsky S, et al. Measure what matters: Counts of hospitalized patients are a better metric for health system capacity planning for a reopening. *J Am Med Inform Assoc.* 2020;27(7):1026–1131.
27. Glover A, Heathcote J, Krueger D, Ríos-Rull JV. *Health versus wealth: On the distributional effects of controlling a pandemic.* National Bureau Econ Res; 2020 Apr 23. Available at: https://www.nber.org/papers/w27046. Accessed June 4, 2020.
28. Malone C, Bourassa K. Americans didn't wait for their governors to tell them to stay home because of COVID-19. FiveThirtyEight. 2020 May 8. fivethirtyeight.com/features/americans-didnt-wait-for-their-governors-to-tell-them-to-stay-home-because-of-covid-19/. Accessed June 4, 2020.
29. "Survey Finds Only 44% of People Are Social Distancing: Hawaii Tribune." Hawaii Tribune. 2020 May 30. www.hawaiitribune-herald.com/2020/05/30/hawaii-news/survey-finds-only-44-of-people-are-social-distancing/. Accessed June 4, 2020.
30. "Covid-19 Social Distancing Scoreboard." Unacast. www.unacast.com/covid19/social-distancing-scoreboard. Accessed July 23, 2020.
31. COVID-19 Mobility Data Network. www.covid19mobility.org/. Accessed July 23, 2020.

32. Daeho KD, Neumann PJ. Analyzing the Cost Effectiveness of Policy Responses for COVID-19: The Importance of Capturing Social Consequences. *Med Decis Making*. 2020;40(3):251–253.

33. "The Unequal Cost of Social Distancing." Johns Hopkins Coronavirus Resource Center. 2020 Mar 30. coronavirus.jhu.edu/from-our-experts/the-unequal-cost-of-social-distancing. Accessed July 23, 2020.

34. Barbera RJ, Dowdy DW, Papageorge NW. Economists and Epidemiologists, Not at Odds, but in Agreement: We Need a Broad Based COVID-19 Testing Survey. Johns Hopkins University, Coronavirus Resource Center. 2020. https://coronavirus.jhu.edu/from-our-experts/economists-and-epidemiologists-not-at-odds-but-in-agreement-we-need-a-broad-based-covid-19-testing-survey. Accessed June 4, 2020.

35. Barzilay R, Moore TM, Greenberg DM, et al. Resilience, COVID-19-related stress, anxiety and depression during the pandemic in a large population enriched for healthcare providers. *Translational Psychiatry*. 2020;10(1):1–8.

36. Blanco GD, Calabrese E, Biancone L, Monteleone G, Paoluzi OA. The impact of COVID-19 pandemic in the colorectal cancer prevention. *Int J Colorectal Dis*. 2020;35:1951–1954. https://doi.org/10.1007/s00384-020-03635-6.

37. Santoli JM. Effects of the COVID-19 pandemic on routine pediatric vaccine ordering and administration—United States, 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(19):591–593.

38. Kinsey EW, Kinsey D, Rundle AG. COVID-19 and Food Insecurity: an Uneven Patchwork of Responses. *J Urban Health*. 2020;97(3):332–335.

39. Khullar D, Bond AM, Schpero WL. COVID-19 and the Financial Health of US Hospitals. *JAMA*. 2020;323(21):2127–2128.

40. Hardy B, Logan TD. *Racial economic inequality amid the COVID-19 crisis*. The Brookings Institute; 2020 August 13. https://www.brookings.edu/research/racial-economic-inequality-amid-the-covid-19-crisis/. Accessed August 18, 2020.

41. Steuerle Eugene, Jackson Leigh Miles, National Academies of Sciences, Engineering, and Medicine. *Advancing the power of economic evidence to inform investments in children, youth, and families*. Washington DC: National Academies Press; 2016 Aug 25.

42. Bonnie RJ, Stroud CE, Breiner HE, Committee on Improving the Health, Safety, and Well-Being of Young Adults. *Investing in the health and well-being of young adults*. Washington DC: National Academies Press; 2015.

43. Acemoglu D, Chernozhukov V, Werning I, Whinston MD. *A multi-risk SIR model with optimally targeted lockdown*. National Bureau Econ Res; 2020. https://www.nber.org/papers/w27102. Accessed May 21, 2020.

44. Alvarez FE, Argente D, Lippi F. A simple planning problem for covid-19 lockdown. *Natl Bureau Econ Res*; 2020. https://www.nber.org/papers/w26981. Accessed May 21, 2020.

45. Fernández-Villaverde J, Jones CI. Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities. *Natl Bureau Econ Res*; 2020. https://www.nber.org/papers/w27128. Accessed May 21, 2020.

46. Drabo EF, Hay JW, Vardavas R, Wagner ZR, Sood N. A cost-effectiveness analysis of preexposure prophylaxis for the prevention of HIV among Los Angeles County men who have sex with men. *Clin Infect Dis*. 2016;63(11):1495–1504.

47. Peng RD, Dominici F, Zeger SL. Reproducible epidemiologic research. *Am J Epidemiol*. 2006;163(9):783–789.

48. Husereau D, Drummond M, Petrou S, et al. Consolidated health economic evaluation reporting standards (CHEERS)—explanation and elaboration: a report of the ISPOR Health Economic Evaluations Publication Guidelines Good Reporting Practices Task Force. *Value Health*. 2013;16(2):231–250.