Open Forum Infectious Diseases

REVIEW ARTICLE FIDSA

Four Challenges Associated With Current Mathematical Modeling Paradigm of Infectious Diseases and Call for a Shift

Shi Chen,[1,2,3] Patrick Robinson,[1,3] Daniel Janies,[4] and Michael Dulin[1,3]

[1]Department of Public Health Sciences, University of North Carolina at Charlotte, Charlotte, North Carolina, USA, [2]School of Data Science, University of North Carolina at Charlotte, Charlotte, North Carolina, USA, [3]Academy for Population Health Innovation, University of North Carolina at Charlotte, Charlotte, North Carolina, USA, and [4]Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, Charlotte, North Carolina, USA

Mathematical models are critical tools to characterize COVID-19 dynamics and take action accordingly. We identified 4 major challenges associated with the current modeling paradigm (SEIR) that hinder the efforts to accurately characterize the emerging COVID-19 and future epidemics. These challenges included (1) lack of consistent definition of "case"; (2) discrepancy between patient-level clinical insights and population-level modeling efforts; (3) lack of adequate inclusion of individual behavioral and social influence; and (4) allowing little flexibility of including new evidence and insights when our knowledge evolved rapidly during the pandemic. Therefore, these challenges made the current SEIR modeling paradigm less practical to handle the complex COVID-19 and future pandemics. Novel and more reliable data sources and alternative modeling paradigms are needed to address these issues.

Keywords. COVID-19; mathematical model; SEIR model; challenges.

The COVID-19 pandemic has swept the globe with unprecedented health, social, and economic consequences [1]. We have used an array of state-of-the-art science and technology breakthroughs to understand this pandemic, including next-generation sequencing to rapidly sequence the genome of SARS-CoV-2 virus, deep learning to identify COVID-19 patients from computed tomography scans, and big data to track human movements and predict hotspots of outbreaks.

Researchers, clinicians, and public health officials rely on mathematical models to characterize and predict the COVID-19 epidemic, derive critical epidemiological metrics (eg, the basic reproduction number R_0), evaluate various intervention strategies, and optimize resource needs [2]. As of July 10, 2020, >300 articles of COVID-19 modeling have been peer-reviewed and published, with many more available on preprint archives. More than 80% of the current efforts adopt the Susceptible-Exposed-Infectious-Recovered (SEIR) paradigm, which is expressed in ordinary differential equations. SEIR models are mechanistic models developed almost a century ago [3]. We have identified 4 substantial interrelated limitations of this modeling paradigm that make it inadequate to address the current COVID-19 and future pandemics. Therefore, we advocate for a modeling paradigm shift for emerging and re-emerging epidemics.

First, there is a substantial discrepancy between current molecular diagnosis of COVID-19 and definition of "case" based on host symptoms in the original SEIR model. The original definition of "case" was not imagined in light of caveats of today's molecular techniques. For COVID-19, the lag between testing and reporting, variability of reliability and access to testing across time and regions, and lack of accurate accounting of asymptomatic/presymptomatic patients all contribute to the "iceberg" phenomenon [4, 5]. These issues have a massive bearing on SEIR model formulation and consequently undermine their application to accurately characterize the COVID-19 pandemic and provide evidence-based support for decision-makers. In addition, the lack of consistency of input across regions and among various model formulations makes cross-model comparison and validation extremely difficult.

Second, SEIR models are formulated at the population level. An important discrepancy exists between patient-level clinical information and population-level modeling for public health. Especially, exposed (E) and infectious (I) states characterize the epidemic at the population level, ignoring important clinical variations among individual patients. A single and universal E state is assumed to be unable to infect others. Similarly, a single I state does not reflect varying clinical severity and prognosis of individual patients (eg, asymptomatic, presymptomatic, mild, severe, and critical stages). Current molecular diagnosis based on quantitative reverse transcription polymerase chain reaction (qRT-PCR) and antigen tests has already been able to provide much more detailed, patient-level quantitative pathogen load information. SEIR models cannot quantify the potential role of super-spreading patients who cause a disproportionately large

Received 16 July 2020; editorial decision 23 July 2020; accepted 29 July 2020.

Correspondence: Shi Chen, PhD, 9021 University City Blvd, Charlotte, NC 28223 (schen56@uncc.edu).

Open Forum Infectious Diseases®

DOI: 10.1093/ofid/ofaa333

number of new cases. Superspreading can be the consequence of individual clinical characteristics (eg, supershedding of virus, which can be identified by qRT-PCR) and/or behavioral aspects (eg, supercontacting), neither of which are properly addressed in the population-level SEIR models.

Third, there is a lack of adequate inclusion of individual behavioral and social influence in SEIR models. Infectious disease epidemics have a substantial social aspect and public health implication. It is imperative to include varying degrees of interventions such as social distancing, stay-at-home, and shelter-in-place orders at different times and across different regions. The assumption of homogenous mixing of S with I state individuals in the SEIR model is therefore invalid during COVID-19. Even with more spatially explicit metapopulation models [2], homogeneous mixing at smaller spatial scales (eg, within a state, county, or city level) is still questionable, as some people can stay home and some are essential. In addition, the regional variability of individual sentiment and behavior, for example, whether to obey or enforce these orders, is essential to determine in order to predict the trajectory of the COVID-19 pandemic, but is generally not included in the SEIR models.

Fourth, our understanding of the interrelated clinical, public health, and social system of COVID-19 has rapidly evolved. In the original paper describing the SEIR approach, the model was applied retrospectively when the epidemic had ended, revealing most clinical and epidemiological information [3]. SEIR-type mechanistic models allow little flexibility for new evidence and insights without substantially changing model structure and estimation of R_0. Compared with other alternative approaches, it is especially difficult to characterize the beginning of the ongoing COVID-19 pandemic with the SEIR model, given that many aspects of the disease remain unclear (eg, asymptomatic transmission, superspreading, and zoonotic transmission pathway) and reported data are vastly underestimated [1, 4]. Estimation of the R_0 value has been raised several iterations from 2 to 6.5 over the course of the epidemic [6]. As this is an exponential growth term, any small change of R_0 value can lead to vastly different public health consequences. This variability may lead to incorrect conclusions for decision-makers who rely on R_0 to evaluate situations and take actions such as reopening. Additionally, the conclusion that an epidemic will die out when $R_0 < 1$ holds true only for deterministic SEIR models, but is not necessarily valid for stochastic models. Given the complexity of the COVID-19 pandemic, formulating COVID-19 as a deterministic system is an oversimplification. Therefore, the lack of a comprehensive understanding of the COVID-19 pandemic and accurate data makes current modeling efforts inconsistent, contradictory, and confusing. In addition, few studies publish the accompanying codes, undergo rigorous external appraisal, and revisit the original model with updated knowledge. Unlike other aspects of COVID-19, which are less directly related to the public, numbers in epidemic trajectory often appear on headlines in major news outlets and social media and lead to unintended social consequences such as confusion, fear, and anger. Worse still, inconsistency in modeling efforts is an easy and vulnerable target for political spin [7].

We have made tremendous progress in the diagnosis, treatment, and prevention of COVID-19. However, the near 100-year-old SEIR model is a rusty weapon in our arsenal against this unprecedented pandemic. The SEIR modeling paradigm is less practical to handle the complicated clinical, public health, and social system of the COVID-19 pandemic. In addition, relying on a single R_0 value to summarize such a complicated system is fraught. In response, we suggest the following strategies. First, SEIR model assumptions should be carefully evaluated before deployment. Second, alternative modeling frameworks and novel data sources are required to accurately characterize and respond to the COVID-19 pandemic. These alternative frameworks include cross-scale models that incorporate both the pathogen and hosts, agent-based models that explicitly incorporate individual-level characteristics, and currently underexplored data-driven machine learning and deep learning models that are not prone to manmade biases in model hypotheses. Furthermore, mechanistic models (such as the SEIR model) complement data-driven models. Ensemble models across different types of models can provide a less biased characterization of the COVID-19 epidemic. New data sources such as social media, electronic health records, and molecular "-omics" can identify possible asymptomatic patients, infer human behavioral aspects, and facilitate tracking transmission chains. We urge our colleagues to consider the challenges in the current SEIR modeling paradigm, to carefully evaluate a model's effectiveness in the COVID-19 and other pandemics, and to consider the complicated clinical, behavioral, and social aspects.

Acknowledgments

Potential conflicts of interest. S.C., P.R., D.J., and M.D. have no conflicts of interest to declare. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

Patient consent. There were no human subjects involved in this study.

References

1. World Health Organization. Coronavirus disease (COVID-19) situation report 121. Published May 20, **2020**. Available at: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200520-covid-19-sitrep-121.pdf. Accessed May 20, 2020.
2. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet **2020**; 395:689–97.
3. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proc R Soc A **1927**; 115:700–21.
4. Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). Science. **2020**; 368:498–493.
5. Vogel G. Antibody surveys suggesting vast undercount of coronavirus infections may be unreliable. Science. **2020**. Available at: https://www.sciencemag.org/news/2020/04/antibody-surveys-suggesting-vast-undercount-coronavirus-infections-may-be-unreliable.
6. Liu Y, Gayle AA, Wilder-Smith A, Rocklov J. The reproduction number of COVID-19 is higher compared to SARS coronavirus. J Travel Med. **2020**; 27. doi:10.1093/jtm/taaa021.
7. Waldman S. U.S. conservatives who detest climate models add a new target: coronavirus models. Science. **2020**; https://www.sciencemag.org/news/2020/04/us-conservatives-who-detest-climate-models-add-new-target-coronavirus-models.