Vaccine 41 (2023) 7067–7071

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Short communication

Using neural networks to calibrate agent based models enables improved regional evidence for vaccine strategy and policy

Ayush Chopra [a,e], Alexander Rodriguez [b,e], B. Aditya Prakash [f], Ramesh Raskar [a], Thomas Kingsley [c,d,e,*]

[a] Massachusetts Institute of Technology Media Lab, Cambridge, MA, United States
[b] Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, United States
[c] Mayo Clinic Department of Medicine, Rochester, MN, United States
[d] Mayo Clinic Division of Epidemiology, Rochester, MN, United States
[e] Mayo Clinic Healthcare & Epidemiology AI Lab (HEAL), United States
[f] College of Computing, Georgia Institute of Technology, Atlanta, GA 30308, United States

ARTICLE INFO

Keywords:
Deep learning
Machine learning
Agent based modeling
Epidemiology
Covid-19
Simulation modeling
Infectious diseases
Emerging disease

ABSTRACT

Distribution and administration strategy are critical to successful population immunization efforts. Agent-based modeling (ABM) can reflect the complexity of real-world populations and can experimentally evaluate vaccine strategy and policy. However, ABMs historically have been limited in their time-to-development, long runtime, and difficulty calibrating. Our team had several technical advances in the development of our GradABMs: a novel class of scalable, fast and differentiable simulations. GradABMs can simulate million-size populations in a few seconds on commodity hardware, integrate with deep neural networks and ingest heterogeneous sources. This allows for rapid and real-world sensitivity analyses. Our first epidemiological GradABM (EpiABMv1) enabled simulation interventions over real million-scale populations and was used in vaccine strategy and policy during the COVID-19 pandemic. Literature suggests decisions aided by evidence from these models saved thousands of lives. Our most recent model (EpiABMv2) extends EpiABMv1 to allow improved regional calibration using deep neural networks to incorporate local population data, and in some cases different policy recommendations versus our prior models. This is an important advance for our model to be more effective at vaccine strategy and policy decisions at the local public health level.

1. Introduction

Decision making in complex environments, such as amid a public health crisis, is challenging. During the COVID-19 pandemic, public health leaders were frequently required to make population decisions such as lockdowns, mask wearing, testing, and vaccine deployment and administration strategies. Since human clinical trials and epidemiological studies are often not feasible, decisions had to be made with a paucity of existing evidence to guide them. During the pandemic public health leaders increasingly turned to data-driven modeling and simulation which provided evidence to support decision making.

Agent-based models (ABM) are often the best choice for in-silico epidemiologic studies by bridging understanding of populations, infections and intervention. This is because ABMs allow connecting multiple, seemingly disconnected, aspects of individual demographic, socio-economic, and behavior preferences along with scientific evidence on infection dynamics and intervention to better capture complexities of the real world.

However, ABMs are conventionally slow to execute [1], difficult to scale to large populations, and tough to calibrate with real-world data [2]. This limits their application, especially during a pandemic where the duration of time-to-decision is short.

To address this, during the COVID-19 pandemic our research team set out to innovate ABM technology. The result was our ability to provide experimental evidence to aid in vaccine policy and strategy. Our goal was to design ABMs that can recreate million-size populations with detail and integrate real-world data streams to effectively analyze sensitivity of interventions. We leveraged modern advances in machine learning to first proposed EpiABMv1: a scalable and fast ABM to enable testing vaccination interventions over real-world populations and

* Corresponding author.
E-mail address: Kingsley.Thomas@mayo.edu (T. Kingsley).

https://doi.org/10.1016/j.vaccine.2023.08.060
Received 31 March 2023; Received in revised form 19 August 2023; Accepted 23 August 2023
Available online 17 October 2023

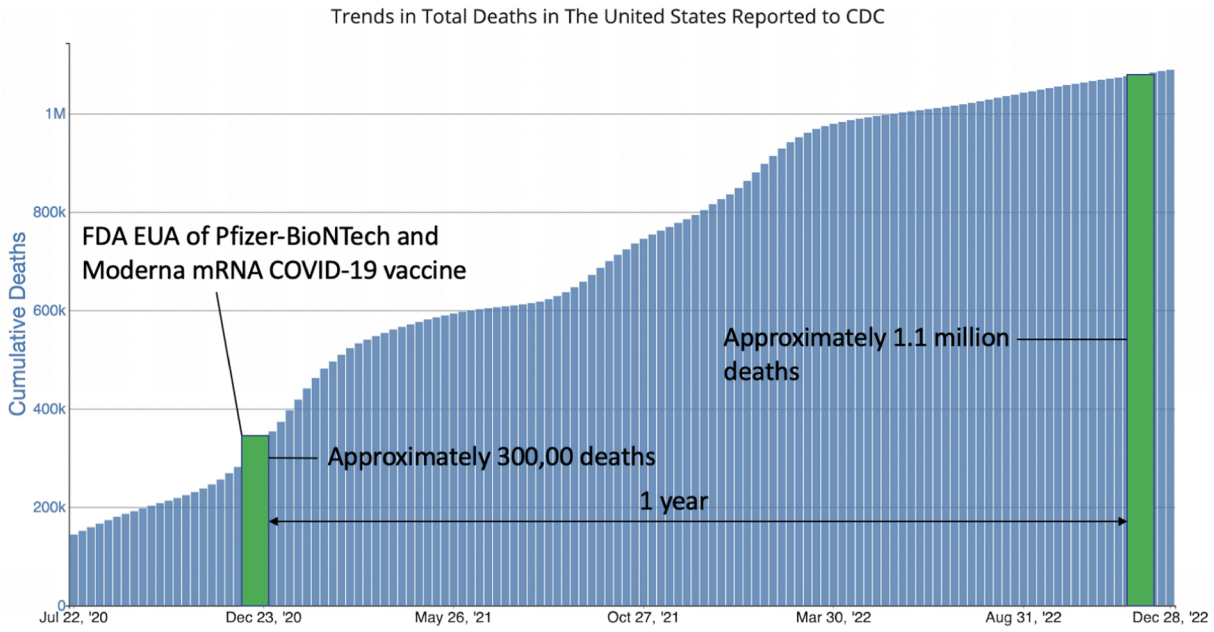

Fig. 1. Cumulative deaths due to COVID-19 infection at the time of US Food and Drug Administration's Emergency Use Authorization. Updated from CDC COVID Tracker [2].

provided evidence for COVID-19 mRNA vaccine strategy. Second, we proposed EpiABMv2 which enables improved calibration of ABMs, like EpiABMv1, to the infection dynamics in a local population using a gradient-based optimization method. In this paper, we merge our advances in EpiABMv1 and EpiABMv2 to enable design and analysis of granular vaccination policies by accounting for regional population variations.

1.1. The importance of vaccine strategy and policy in emerging disease and the need for in-silico epidemiological modeling

The development of an effective vaccine, as we observed during COVID-19, is only part of the challenge. There were more COVID-19 deaths after the FDA's emergency authorization of the Moderna and Pfizer vaccines in the same duration than before (See Fig. 1) [3]. After one year of vaccine availability, less than half of the global population was vaccinated to COVID-19 [4]. The logistics of rapidly producing and equitably delivering a vaccine on a global scale of nearly 8 billion humans is a significant challenge. This can be partly attributed to challenges with global vaccine production capacity, its transportation and storage, administration strategy, and addition of more virulent COVID variants. Additionally, in countries like the United States lack of evidence for public health or vaccine policy made it easier for speculation on social media and this has been partly attributed to the spread of misinformation leading to vaccine hesitancy [1].

Tools to evaluate vaccine strategy and policy can provide a needed source of evidence to support public health action. During the COVID-19 pandemic many public health leaders turned to in-silico modeling methods given the infeasibility of human trials [5]. However, most mechanistic and statistical simulation models cannot reflect the complexity required for this type of experiment. To do so accurately requires reflecting a real-world population of thousands to millions of individuals with different characteristics and risk factors. It requires understanding their interaction networks, the particulars of their individual and group behaviors, and various attributes from the individual, to neighborhood, to community that impact these interactions. It requires accurately simulating the dynamics of these interactions in relation to disease spread and the intervention being studied. And this list of considerations is still far from exhaustive.

1.2. Breakthrough in ABM technology and application in vaccine strategy and policy during COVID-19

ABMs are discrete simulators which comprise a collection of agents that can act and interact within a computational world. In fields such as epidemiology, these simulations are at the scale of millions of agents with large networks of interactions [2,6,10] and performing a single simulation may conventionally take several days run on a supercomputer [8]. Further, this makes calibration subpar since tuning parameters requires iteratively running the model numerous times; and sensitivity analyses unreliable due to sub-optimal calibrated parameters and high cost of re-running simulations under different scenarios [7].

Our design, GradABM, has been shown to alleviate several of these concerns regarding scalability and data-driven calibration [1]. GradABM introduces a novel tensorized and differentiable design for ABMs, which allows simulating million-size populations in a few seconds on commodity hardware [2] and integrating with deep neural networks for faster and efficient collaboration [1]. Specifically, on the JUNE [6] epidemiology model used by the NHS and UN Global Pulse during COVID, GradABM helped reduce simulation time from 50 h to 5 min, calibration time from 10,000 CPU hours to 20 min (whilst also achieving better generalization) and sensitivity analysis time from 5,000 CPU hours to 10 s [1,10]. This computational advance directly translates to practical utility for decision making.

These novel ABM simulations by our team provide experimental evidence for optimal COVID-19 mRNA dosing and administration strategy. During the early phase of the mRNA vaccine roll, public health experts debated whether the three-week vaccine dosing strategy was the best strategy for preventing deaths. Some hypothesized delaying the second dose to administer more individuals with the first dose would save lives. Others argued strongly against this strategy, often stating lack of evidence. The clinical trials only used a three-week dosing strategy. Our GradABM work translated to an EpiABMv1 [2] model that showed delaying a second dose to 12 weeks had significant reduction in deaths if the daily vaccination rate was low [8]. These results impacted several countries' vaccine policy. For example, the UK moved to a 12-week dosing strategy and a recent publication in Lancet Public Health estimates that switching to this strategy saved 10,000 lives in the UK alone [9]. Thus, providing real-world evidence to support our model's

effectiveness.

While delaying the second COVID-19 prompted support for its potential to expedite initial protection and created tangible public health impact, it also received some skepticism given geographical disparities in disease burden, vaccine hesitancy, individual risk and resource availability. Further, quantitative variations in these dynamics can influence the qualitative design and subsequent impact of interventions. Targeted interventions are the logical progression to address specific challenges posed by demographic and geographic variations and control localized outbreaks with better management of resource constraints. Our current work facilitates this by enabling localized calibration of ABM to infection dynamics in the specific geography.

1.3. Most recent advances in GradABM model

Our most recent GradABM model, EpiABMv2 [1], allows the model to make more specific recommendations on a county level. Previously, we had shown that GradABM's EpiABMv2 outperformed other ABM models in terms of aggregate spread forecasting based on CDC guidelines, including our team's previous versions (EpiABMv1) [1,2]. In this paper we experimentally compare its performance difference compared to other models when our calibration methods are used for qualitative decisions with local geographic variables. This update allows for more regional precision in public health decision support.

2. Methods

2.1. Developing tensorization and differentiability into ABM

Our method GradABMs is a novel class of agent-based models (ABMs) that are designed to be compatible with gradient-based learning using automatic differentiation. This is the central technique which enables deep learning algorithms to optimize model parameters by learning from diverse data sources. GradABMs have two key features: tensorization and differentiability. First, GradABMs follow a tensorized design where agents are represented as vectors and interactions as (sparse) adjacency matrices. This allows scalable and fast simulations that can run rapidly, in a highly parallelized manner and using GPU hardware - simulating million-size agent populations in a few seconds [1,2]. Second, GradABM reparametrizes the gradients of discrete distributions used in the simulation with continuous approximations, allowing for end-to-end differentiability [1]. This differentiability allows GradABM to merge with deep neural networks for robust optimization and seamlessly incorporate novel data streams. Together, these allow us to build realistic simulators and efficiently calibrate them using gradient-based learning over multiple runs of the simulator.

2.2. Specification of agent states and disease spread dynamics

The state of each agent (represented as a vector of categorical variables) contains age, disease stage (S, E, I, R, D) and time since last exposure. Over timesteps, the state evolves as the agent interacts with other agents based on the clinical model in which is governed by time-dependent parameters that control the transmission of disease to create new infections (R0, initial infection rate) and the progression of disease stages of already infected agents (generation time, mortality rate). The transmission of disease is based on interactions within a day and uses a transmission function that computes the probability of infection transmission between susceptible and infected (or exposed) agents. The disease stage may update due to a new infectious interaction with an exposed or infected agent (captured by the transmission function) or the natural progression of a previously incubated infection (captured by the progression function). Specifically, the transmission function is parameterized with R_{eff}. The initial infection rate and progression function is parameterized with the variant generation time and mortality rate. The calibration of these parameters is done by matching estimates of the simulator with real-world observed cumulative death statistics.

Table 1
Forecasting results for COVID-19 over 5 runs comparing GradABM with baseline region-agnostic and region-dependent calibration techniques.

Calibration Method	Validation Error (lower is better)		
	ND	RMSE	MAE
Region-agnostic	8.75	689.92	270.13
Region-specific (surrogate-based)	2.21 ± 1.36	121.87 ± 63.97	68.20 ± 41.84
Region-specific (proposed GradABM)	0.97 ± 0.18	50.99 ± 12.12	30.02 ± 5.60

2.3. Regional variables for model calibration using deep neural networks

First, the calibration neural network (CalibNN) takes input from varied data sources and predicts simulation parameters for disease transmission and progression. The parameters are passed to the GradABM and used to run a simulation. The aggregate output generated by the simulator is compared with real-world mortality numbers for each day. We utilize the mean-squared error function to quantify the goodness of the solution. After that, we learn the parameters of the neural network using the backpropagation algorithm. We note that the CalibNN based approach presented in this paper is fundamentally distinct from emulation or surrogate models. Emulation models take the same input as the ABM's input and predict the output of the ABM, without simulating any agent behavior. In contrast, the output of CalibNN serves as the input to GradABM, which then simulates the behavior of the agents. Thus, CalibNN extends the simulation pipeline, by enabling us to learn the correct inputs for GradABM.

2.4. Validation to real-world scenarios by transmission forecasting

Since there is no ground truth on simulation parameters, we evaluate calibration based on the quality of infection forecasts produced by the simulator. Following CDC forecasting guidelines [1], we make weekly forecasts of cases and deaths for 1–4 weeks ahead in the future. In our evaluation, we work with the following counties in Massachusetts: 25001, 25003, 25005, 25009, 25011, 25013, 25015, 25021, 25023, 25027. The specific evaluation period is determined with epidemic weeks which is the standard in CDC's epidemic prediction initiatives. For COVID-19 these are 202014, 202016, 202018, 202020, 202022, 202024, 202026, 202028, 202030. To evaluate performance, we use several standard metrics for evaluating epidemic predictions. Specifically, normal deviation (ND), root mean squared error (RMSE) and mean absolute error (MAE). A better calibration technique will produce parameters with lower measurement errors in the simulated forecasts.

2.5. Comparison with baseline techniques

We compare our proposed pipeline with several methods for region-agnostic and region-specific calibration. Results are summarized in Table 1. Conventionally, *region-agnostic calibration* (row 1) techniques have been used which estimate the infection dynamics parameters using in-situ control trials and reuse the same parameters across different geographical and demographic regions. The parameters are a good representation of the population but may not generalize to other geographies. More recently, advances in data availability and modeling techniques have enabled *region-specific calibration* which utilize aggregate infection data (cases, deaths, hospitalizations) to generate localized estimates for the dynamics parameters. Prior region-specific methods estimate (row-2) parameters by building simplified surrogate models which do not capture the heterogeneity and complexity of real-world interactions (usually assume perfect mixing of people) but are easy to model. Our proposed method (row-3) alleviates computational

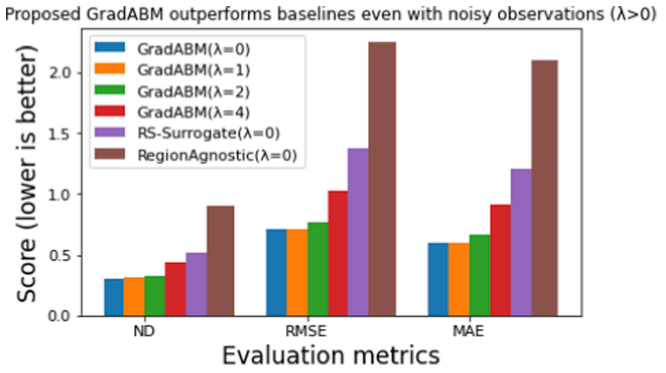

Fig. 2. Our proposed calibration method is more robust to observation error. GradABM achieves lower forecasting error than all baselines even when it is trained with noisy data ($\lambda > 0$) while the baselines receive original data. This is achieved due to the joint scalable (EpiABMv1) and differentiable (EpiABMv2) design which allow modeling real-world populations and incorporating localized data-sources for calibration.

challenges by allowing region-specific calibration using an agent-based model which can capture heterogeneous interactions in the population through calibration of parameters using both local data and neural networks. This improved calibration is shown by better validation performance in the Table 1 below.

Beyond capturing population heterogeneity, real-world utility also requires robustness to measurement error in region data used for calibration. Centers for Disease Control and Prevention (CDC) estimates on deaths, cases, hospitalizations are likely to be noisy due to data reporting and collection issues [11]. To investigate the robustness of our proposed calibration procedure, we run experiments using ground truth data distorted by gaussian noise. More specifically, we add gaussian noise to each ground truth target with mean $\mu = 0$ and varying scales of standard deviation s. To set the standard deviation of noise for each county, we first compute the standard deviation of the ground truth data and multiply it by a λ factor. Even for a large degree of noise ($\lambda = 4$), we observe that our method outperforms both region-specific with surrogate and region-agnostic calibration techniques. Results are summarized in Fig. 2 below.

2.6. *Experimentally evaluating sensitivity to clinical & geographical variations with GradABM versus prior ABM models*

The experimental pipeline is summarized in Fig. 3. In this experiment we incorporate regional variables (census information, mobility patterns, facebook symptom surveys, CDC case statistics) into an ABM we term EpiABMv1 and our latest GradABM model, EpiABMv2. The EpiABMv2 uses the CalibNN on these regional variables while EpiABMv1 does not. We randomly selected Franklin County in Massachusetts to parameterize regional variables into both models. We then compare vaccine strategy recommendations between the two models across variable 1st dose vaccine efficacies.

The delay vs not delay strategy can be evaluated from the model output by computing the ratio of cumulative deaths of P2 by P1, which we denote as relative mortality. Basically, if the relative mortality is less than 1, then policy P2 is better (can delay the second COVID-19 dose); while relative mortality greater than 1 implies that policy P1 is better (do not delay the second COVID-19 dose). The goal of our experiment is to evaluate whether calibrating on regional variables impacts strategy at a local level. Although delaying the second dose of mRNA vaccine is not a current consideration, we use this example in our experiment because our prior work impacting this decision and having been validated in real-world outcomes.

3. Results

Fig. 2 reports the sensitivity of the relative mortality for Franklin County in MA to the protection offered by the first vaccine dose. We do this using our prior and most recent ABM (different calibration ABMs): EpiABMv1 and EpiABMv2. For vaccine efficacy of 80 %, both models recommend delaying the 2nd dose of the COVID-19 vaccine, which is consistent with real policy recommendations deployed in the UK and as validated by prior clinical work. However, at vaccine protection of 60 %, EpiABMv1 and EpiABMv2 provide qualitatively different recommendations for Franklin County. EpiABMv2, the model, which was calibrated using CalibNN and granular county data to achieve a superior calibration fit, recommends not delaying the 2nd dose while EpiABMv1 recommends delaying the second dose. This difference reflects calibrating regional variables using our CalibNN method that can change outcomes.

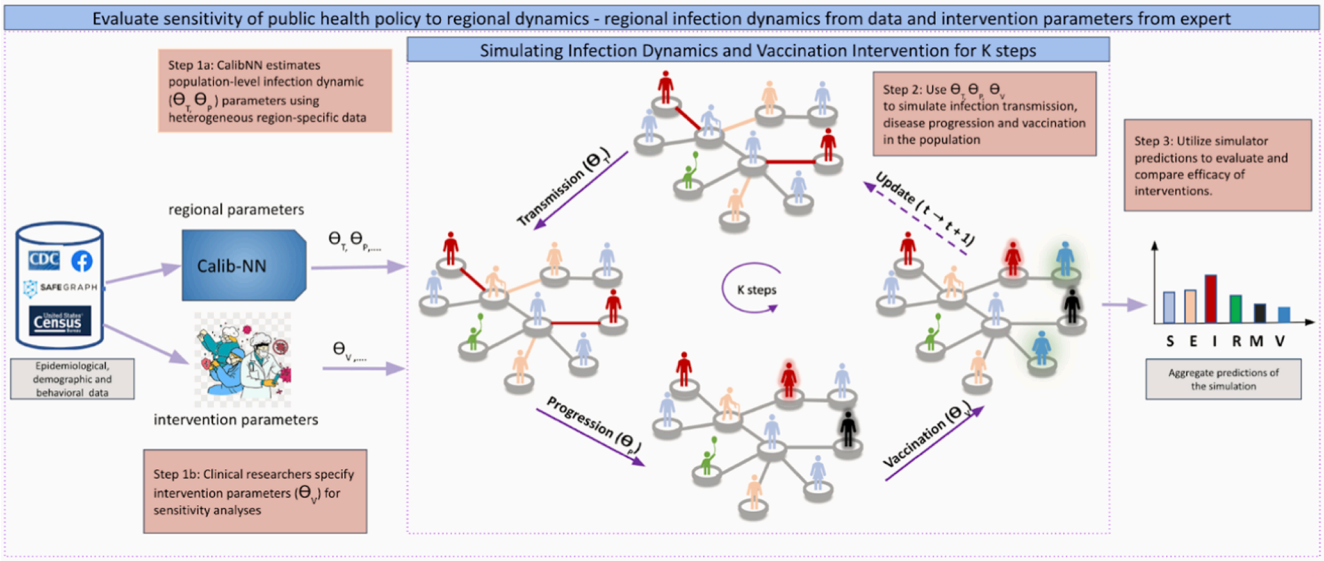

Fig. 3. Overview of simulation model parameterization and calibration pipeline with neural networks.

4. Discussion and future Work

GradABM allows learning regional calibration parameters for different counties. Because of inter-county variation, the same policy may not be optimal in both locations. With this advance in our ABM, we can provide more effective evidence for vaccine strategy and policy decisions at the local public health level.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] Chopra A, Rodríguez A, Subramanian J, Krishnamurthy B, Prakash BA, Raskar R. Differentiable agent-based epidemiology. Adaptive Agents Multi-agent Systems (AAMAS) 2023.

[2] Chopra A, Raskar R, Subramanian J, et al. DeepABM: scalable and efficient agent-based simulations via geometric learning frameworks - a case study for Covid-19 spread and interventions. In: *2021 Winter simulation conference (WSC)*; 2021. p. 1. doi: 10.1109/WSC52266.2021.9715507.

[3] CDC. COVID data tracker. Centers for Disease Control and Prevention. Published March 28, 2020. Accessed March 30, 2023. https://covid.cdc.gov/covid-data-tracker.

[4] Watson OJ, Barnsley G, Toor J, Hogan AB, Winskill P, Ghani AC. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. Lancet Infect Dis 2022;22(9):1293–302. https://doi.org/10.1016/S1473-3099(22)00320-6.

[5] Becker AD, Grantz KH, Hegde ST, Bérubé S, Cummings DAT, Wesolowski A. Development and dissemination of infectious disease dynamic transmission models during the COVID-19 pandemic: what can we learn from other pathogens and how can we move forward? Lancet Digit Health 2021;3(1):e41–50. https://doi.org/10.1016/S2589-7500(20)30268-5.

[6] Aylett-Bullock J, Cuesta-Lazaro C, Quera-Bofarull A, Icaza-Lizaola M, Sedgewick A, Truong H, et al. June: open-source individual-based epidemiology simulation. Royal Soc Open Sci 2021;8, 7(July2021):21050. https://doi.org/10.1098/rsos.210506. Publisher: Royal Society.

[7] Squazzoni F, Polhill JG, Edmonds B, et al. Computational models that matter during a global pandemic outbreak: a call to action. J Artif Soc Soc Simul 2020;23(2):10.

[8] Romero-Brufau S, Chopra A, Ryu AJ, et al. Public health impact of delaying second dose of BNT162b2 or mRNA-1273 covid-19 vaccine: simulation agent based modeling study. BMJ 2021;373:n1087. https://doi.org/10.1136/bmj.n1087.

[9] Imai N, Rawson T, Knock ES, et al. Quantifying the effect of delaying the second COVID-19 vaccine dose in England: a mathematical modelling study. Lancet Public Health 2023;8(3):e174–83. https://doi.org/10.1016/S2468-2667(22)00337-1.

[10] Quera-Bofarull A, Chopra A, Aylett-Bullock J, Cuesta-Lazaro C, Calinescu A, Raskar R, et al. Don't simulate twice: one-shot sensitivity analyses via automatic differentiation. Adaptive Agents Multi-agent Systems (AAMAS) 2023.

[11] Pian W, Chi J, Ma F. The causes, impacts and countermeasures of COVID-19 "Infodemic": a systematic review using narrative synthesis. Inf Process Manag 2021;58(6):102713.