



OPEN

A COVID-19 model incorporating variants, vaccination, waning immunity, and population behavior

Zachary LaJoie^{1,2}, Thomas Usherwood^{1,2}, Shailen Sampath^{1,2} & Vikas Srivastava^{1,2}✉

Vaccines for COVID-19 have allowed countries to combat the spread of the disease. However, new variants have resulted in significant spikes in cases and raised severe health and economic concerns. We present a COVID-19 model to predict coupled effects of vaccine multiple-dose roll-out strategies, vaccine efficacy, waning immunity, population level of caution, sense of safety, under-reporting of cases, and highly prevalent variants such as the Delta (B.1.617.2) and Omicron (B.1.1.529). The modeling framework can incorporate new variants as they emerge to give critical insights into the new cases and guide public policy decision-making concerning vaccine roll-outs and reopening strategies. The model is shown to recreate the history of COVID-19 for five countries (Germany, India, Japan, South Africa, and the United States). Parameters for crucial aspects of the pandemic, such as population behavior, new variants, vaccination, and waning immunity, can be adjusted to predict pandemic scenarios. The model was used to conduct trend analysis to simulate pandemic dynamics taking into account the societal level of caution, societal sense of safety, and the proportions of individuals vaccinated with first, second, and booster doses. We used the results of serological testing studies to estimate the actual number of cases across countries. The model allows quantification of otherwise hard to quantify aspects such as the infectious power of variants and the effectiveness of government mandates and population behavior. Some example cases are presented by investigating the competitive nature of COVID variants and the effect of different vaccine distribution strategies between immunity groups.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus that causes coronavirus disease 2019 (COVID-19), was first detected in Wuhan, China in December of 2019. Since its initial detection, the virus has resulted in a global pandemic of COVID-19 with over 500 million cases and 6 million deaths world-wide¹. In response, most countries had lock downs that negatively impacted local and global economies². The recent developments of COVID vaccines, such as BNT162b2 (Pfizer-BioNTech), mRNA-1273 (Moderna), and ChAdOx1 (AstraZeneca), have helped in curbing the spread of the disease^{3,4}. Accelerated vaccine roll-outs have been shown to significantly reduce the number of cases within countries⁵. As of March of 2022, there are currently 12 vaccines approved for use and 38 vaccines in the final stages of testing⁶. Additionally, there have been 5 billion individuals who have been vaccinated with at least one dose world-wide⁷.

However, as promising new vaccines have emerged, variants of concern (VOC) bring new uncertainties and complicate curbing of the disease. VOCs, as defined by the United States Center for Disease Control and Prevention, are more transmissible versions of SARS-CoV-2; they tend to be more severe and result in a significant reduction in vaccine effectiveness⁸. Several VOCs have emerged since December 2020. The Alpha variant (B.1.1.7) was first detected in the United Kingdom and had a 50% increased transmission, soon after the Beta variant (B.1.351) was documented in South Africa. The Gamma variant (P.1) was first detected in Brazil⁹, and the Delta variant (B.1.167.2) was first identified in India and classified on May 11th, 2021¹⁰. The Omicron variant (B.1.1.529) was first detected in South Africa and Botswana in 2021, and this VOC is even more transmissible than Delta, however, it is less severe than previous VOCs¹¹. These variants have contributed to recent large new COVID case surges across countries. When the Delta variant became the dominant VOC around the world, highly vaccinated countries saw significant increases in cases¹². Delta was more dangerous than previous variants, and vaccines are less effective in protecting against it. Getting one dose of either BNT162b2 (Pfizer-BioNTech) or ChAdOx1 (AstraZeneca) confers 25–36% immunity against the Delta compared to 45–52% immunity against the Alpha VOC¹². Receiving two doses was also shown to confer less immunity against the Delta variant¹². Lower

¹School of Engineering, Brown University, Providence, RI 02912, USA. ²Center for Biomedical Engineering, Brown University, Providence, RI 02912, USA. ✉email: vikas_srivastava@brown.edu

immunity, combined with the fact that Delta was more transmissible than previous VOCs⁸ (due to the presence of a mutation in the Delta SARS-CoV-2 spike protein's receptor-binding domain¹³) led to the immense increases in infectious cases we observed around the world in 2021. In December of 2021, the Omicron variant led to even more significant surges in cases across the globe. Omicron has been found to have 33 mutations in its spike protein compared to the wild type strain¹⁴. These large mutations that have been seen across various VOCs require that countries worldwide reconsider the best way to roll out vaccines considering vaccines become less effective and society reopens (more infectious interactions) with time.

As of Spring 2022, most current vaccines require two doses that are recommended to be given 3 to 4 weeks apart on average, require storage at chilled temperatures, and have a limited shelf life¹⁵. Currently, getting an additional one or two booster shots is necessary to retain high levels of protection¹⁶. Navigating these requirements has been difficult globally. For countries that are not producing vaccines and rely on importing doses¹⁷, the vaccine availability has been limited, and a large number of people are put at risk. As a result, many countries with lower supply levels have considered using their vaccine supply to give as many first doses as they can, which still confer reasonable immunity and delay the second dose's time compared to the recommended time frame. This vaccination strategy of prioritizing first doses has had success for other outbreaks in the past in managing delays in vaccination supplies. During a cholera outbreak in Zambia, a single-dose vaccine campaign was successfully used to help get doses to the majority of individuals even though two vaccines were recommended¹⁸. In 2016, one-fifth of the recommended dosage of the yellow fever vaccine was used to successfully respond to a yellow fever outbreak in Angola, Uganda, and the Democratic Republic of the Congo¹⁹. While most countries have prioritized using two dose deployment¹⁵ in accordance with guidelines for the COVID-19 vaccine by the World Health Organization²⁰, many countries decided to prioritize giving first doses to specific segments of the population or most of the population first when they were low on vaccine supplies. In Canada, a 4-months gap between doses was used, with the country prioritizing giving doses to the elderly and those most at risk first²¹. The United Kingdom also allowed for a 12 week delay in the second dose for both the BNT162b2 (Pfizer-BioNTech) and mRNA-1273 (Moderna) vaccines²². India also created a long gap of three to four months between doses²³. The waning of vaccine protection has further created a serious concern²⁴. Now that boosters are recommended, a dichotomy has been made between countries still trying to get the majority of their country vaccinated and countries trying to distribute additional booster shots.

Since the onset of COVID-19 many papers have investigated modeling the progression of the pandemic using modified differential equation compartment models that have compartments for susceptible, infectious, and recovered individuals, referred to as SIR models^{25–34}. For COVID-19, various factors representing more realistic scenarios can be incorporated in the SIR models to improve predictions. Mak et al.³⁵ applied a modified SEIR compartment model to take into account varying vaccine deployment strategies to predict total infections, hospitalizations, and mortality. Garcia et al.³⁶ applied a model to forecast the transmission dynamics of COVID-19 in Costa Rica considering two different vaccination rate scenarios. Gonzalez-Parra et al.³⁷ used a modified compartment model that incorporated latent, infective symptomatic, and symptomatic individuals to take into account VOC such as Alpha variant (B.1.1.7) to predict prevalence, hospitalizations, and deaths. However, despite numerous useful studies^{38–42}, there remains a gap in modeling the coupled effects of population behavioral changes, waning immunity, vaccine roll-out strategies, and new VOCs on the progression of the pandemic.

Previously Usherwood et al.²⁵ proposed a model that accounted for the population behavior and predicted infection dynamics for the original COVID-19 variant. This model was able to predict the trend of the original COVID-19 variant well but it did not consider new variants. Since then, new variants have emerged that require a model that incorporates VOCs. In this paper, we present a modeling framework that accounts for various VOCs, vaccine roll-out strategies, waning immunity, population level of caution, sense of safety, and the under-reporting of cases⁴³ to replicate and predict the number of COVID-19 cases across five countries (Germany, India, Japan, South Africa, and the United States) selected to represent a diverse geographical and societal, cultural pool. We show that the model, with its computational implementation, can reasonably accurately recreate the entire history of the COVID-19 pandemic across these countries and provide a new framework to model a variety of future trends, including the contemplative studies of COVID-19 pandemic outcomes. Such modeling studies are valuable because they enable assessments of the effects of specific aspects of the pandemic, such as behavioral changes or vaccination, in isolation from other features that would affect the infection rate.

Methods

Modeling multiple-dose vaccination and emerging variants with a stacked model. Usherwood et al.²⁵ proposed a method for incorporating time-dependent population behavior changes into a SIRDV differential equation compartment model to predict original COVID-19 virus infection case trajectories for several regions of the United States. Since the publication of this work, the emergence of a large number of new virus variants, waning immunity, and a need for multiple vaccination doses has prompted the development of the new modeling framework presented in this paper. The stacked compartment framework proposed in this work is able to model the waning immunity, multiple-dose vaccinations, and new variants. The model replicates the SIRD model into the desired number of distinct immunity levels, with a fully immunized and waning level for each vaccine dosage. This results in five immunity levels/stacks: Unvaccinated (UV), First Dose (V1), First Dose Waning (V1W), Fully Vaccinated (V2), and Fully Vaccinated Waning (V2W). Each of these levels consists of a detailed compartmental model with compartments for each disease state. The model incorporates different variants by employing separate infected and recovered compartments for each variant. Waning immunity following vaccination is modeled by adding a waning immunity level for each vaccination level of the stacked compartment model. Immediately upon gaining immunity from vaccination or infection, there is flow into a compartment/layer with increased immunity; then, there is a gradual decay into the complementary waning layer with lower

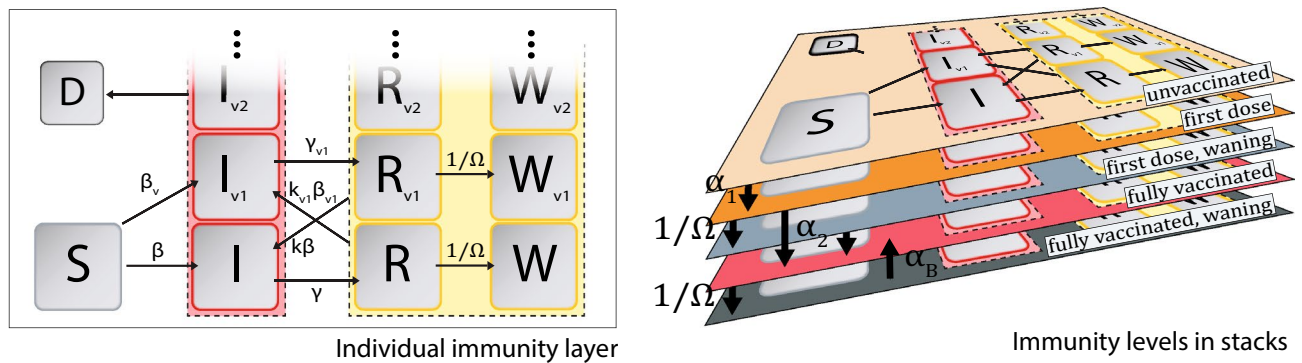


Figure 1. Schematic representation of proposed stacked compartment model. We propose a model that splits the population into a series of vaccination-related immunity levels and uses sub-compartments within each level to represent variants and infection-related disease states. The 2D figure (left) represents the compartment model within an immunity level, composed of a single susceptible (S) and deceased (D) compartment, as well as an infected (I), recovered (R), and waning recovered (W) compartment for each prevalent variant. The 3D figure (right) shows the entire model, where stacked immunity levels represent waning vaccination status. The dotted boxes contain several sub-compartments. Arrows representing flow from the dotted boxes in the 2D figure indicate proportional flow from each of the contained sub-compartments.

Parameter	Description
β_v	Transmission rate of COVID-19 variant v
α_1	Rate at which people receive first vaccine dose
α_2	Rate at which people receive second vaccine dose
α_B	Rate at which people receive booster shots
$\eta_v^{(i)}$	Efficacy of immunity level (i) to COVID-19 variant v
γ_v	Recovery rate for COVID-19 variant v —reciprocal of infectious period
k_v	Relative reinfection rate for COVID-19 variant v —ratio of variant transmission rate before and after initial infection
μ_v	Mortality rate for COVID-19 variant v
Ω	Immunity decay rate
d_I	Level of caution factor—at higher values, the population will tend to increase preventative measures more quickly in response to increases in infection
d_V	Sense of safety factor—at higher values, the population will tend to reduce preventative measures more quickly in response to increases in vaccination. This effect counteracts the effects of level of caution.
β_{v0}	Baseline transmission rate for variant v , before any behavioral response to the disease

Table 1. Parameters in the model and their physical descriptions.

immunity. A similar approach is used for infection-induced immunity, where there is flow from the recovered compartment to the waning-recovered compartment, which has a greater chance for reinfection. Each layer in the model has five main compartment categories: Susceptible (S), Infected (I), Recovered (R), Waning Recovered (W), and Diseased (D), with Infected, Recovered, and Waning Recovered composed of several compartments equal to the number of variants that are being modeled. Each compartment represents a time-dependent fractional variable of the total population in each respective country, with each sub-population of the country being identifiable with a vaccination-induced immunity level (UV, V1, V1W, V2, V2W) and an infection-induced compartment within the immunity level (S, I, R, W, D). The general architecture of our model is represented in Fig. 1 and can be described by the following equations, with the parameters described in Table 1.

$$\begin{aligned}
 \frac{dS^{(i)}}{dt} &= -S^{(i)} \sum_j^{immunity} \sum_v^{variants} \left(\beta_v I_v^{(j)} (1 - \eta_v^{(i)}) \right) \\
 \frac{dI_v^{(i)}}{dt} &= I_v^{(i)} \left(\beta_v S^{(i)} (1 - \eta_v^{(i)}) - \mu_v \gamma_v - \gamma_v \right) \\
 \frac{dR_v^{(i)}}{dt} &= \gamma_v I_v^{(i)} - \frac{R_v^{(i)}}{\Omega} \\
 \frac{dW_v^{(i)}}{dt} &= \frac{R_v^{(i)}}{\Omega} - W_v^{(i)} \sum_j^{immunity} \sum_{u \neq v}^{variants} \left(k_u \beta_u I_u^{(j)} (1 - \eta_u^{(i)}) \right) \\
 \frac{dD^{(i)}}{dt} &= \sum_v^{variants} \left(\mu_v \gamma_v I_v^{(i)} \right)
 \end{aligned} \tag{1}$$

The transmission rate of each variant β_v is modeled as proposed in the behavioral model by Usherwood et al.²⁵. This model incorporates behavioral changes to the total infections and vaccinations as a function for the transmission rate that is dependent on the sum of all infected compartments, I , and the sum of all vaccinated levels, V .

$$\begin{aligned} f_I &= e^{-d_I I} \\ f_V &= \frac{1}{f_I} + \left(1 - \frac{1}{f_I}\right) e^{-d_V V} \\ \beta_v &= \beta_{v0} f_I f_V \end{aligned} \quad (2)$$

These equations describe the change in disease transmission as a result of behavioral response to the current number of infections and vaccinations in the population. The time dependent infectious variant transmission rate $\beta_v(t)$ is given by $\beta_v = \beta_{v0} f_I f_V$, where β_{v0} is the population maximum infection transmission rate for a variant observed in the absence of any preventative societal measures. The expression for f_I models caution in a population where individuals increase their measures of reducing transmission in response to rises in the current number of infections. A population's tendency to take these measures in response to increases in infections is modeled by the population's level of caution, d_I . The mathematical formulation is such that, when there are few infections, $I \rightarrow 0$, or when the population has a low level of caution, $d_I \rightarrow 0$, f_I will approach 1 and therefore have no effect on the variant transmission rate, β_v . This represents little change toward disease safety focused behavior as a result of either a low number of infections or a low level of caution in the population. In the limiting case of very high infections and high level of caution, f_I approaches 0, thereby reducing the net transmission towards 0. This scenario occurs when a population takes specific measures to significantly reduce disease transmission, such as quarantine, mask-wearing, and improved hygiene.

There is a similar but competing formulation for the population's response to the current number of vaccinations, in which the population lessens their measures to reduce transmission as a larger proportion of the population becomes vaccinated. In the expression for f_V , this effect (lessening of caution due to high vaccinations) becomes negligible as the vaccinated population, V , or the population's sense of safety, d_V , approaches 0. In this case, f_V approaches 1 and the population's disease transmission is determined entirely by the population's response to infections, described above. In the limiting case of high sense of safety and high vaccination, the expression for f_V approaches $\frac{1}{f_I}$, the sense of safety from vaccination negates the sense of caution from infections and the population's disease transmission approaches its pre-pandemic level.

The net flow between immunity levels is described by first, second, and booster vaccination rates, α_1 , α_2 , and α_B , distributed equally between compartments within the immunity level. The superscript (i) represents the immunity levels so that there is a set of equations of the form in Eq. (1) for each immunity level UV, V1, V1W, V2, V2W. Vaccination-induced immunity decays naturally from the first-dose and fully-vaccinated immunity levels into their respective waning immunity levels, with an equivalent form as the decay from infected to recovered compartments. Similarly, infection-induced immunity decays from the recovered to recovered waning compartment within each immunity level. These decays between immunity levels are described by the time constant $1/\Omega$, which is described in the literature to be around 150 days, or a half-life of 100 days^{44,45}.

In the model, at the beginning of the pandemic, all individuals from a country start in the susceptible, unvaccinated compartment and then can be infected and moved to one of the infected compartments directly at a specific rate of β_v , or after being vaccinated at a reduced rate. Individuals who are vaccinated will first move into the susceptible, first dose compartment at a rate of α_1 and then into the susceptible, fully vaccinated compartment at a rate of α_2 . Unlike the conventional SIRDV models, we do not have separate compartments for susceptible vaccinated individuals and insusceptible vaccinated individuals. Instead, our stacked compartment model assumes that each vaccinated individual has some risk of getting infected due to break-through cases⁴⁶. In reality, people who are vaccinated span a spectrum of susceptibility to the virus⁴⁷. Individuals move from the susceptible compartments into the corresponding infected compartment at the unique rates of β_v times the unique $\eta_v^{(i)}$ depending on the vaccination-induced immunity level (i), and variant, v , which accounts for the protection acquired by having the vaccine against a certain variant. From each of the infected compartments, individuals will either move into the diseased compartment at a fixed rate of μ_v , or to the corresponding recovered compartment at a fixed rate of γ_v . After natural decay from the recovered to the waning recovered compartment, there will be a flow from waning recovered into each of the infected variants' compartments, except for that of the corresponding variant. This model for reinfection is based on two assumptions that are backed by empirical observations: (i) for some time after infection, there is little chance for reinfection⁴⁸; therefore, there is no flow from recovered to infected, and (ii) after natural decay of infection-induced immunity, chances of reinfection with the same variant remain low⁴⁸; thus there is no flow from waning recovered to infected of the same variant.

Since many individuals are not tested for COVID-19 and asymptomatic cases remain prevalent, there are many unreported cases, especially in countries where testing is less prevalent⁴⁹. Here we predict the actual number of cases by multiplying the reported number of cases by a factor M . The number of cases underreporting/scaling factor M is estimated through data obtained from the serological tests that were conducted in each country^{50–57} to count the number of individuals with immunoglobulin G antibodies against COVID-19. The case underreporting factor, M (actual cases/reported cases), was assumed to be inversely proportional to the number of tests reported. For this study, the underreporting factor was estimated by a population-specific constant divided by the number of tests reported on a particular day. The population-specific constants were estimated from serological testing obtained before widespread vaccination to avoid distortion in the serological testing due to vaccination-induced immunity.

Five countries around the globe representing vastly different geographical and socio-economics were selected for our model-based study. The model was implemented using MATLAB to numerically solve the complete modeling framework. Several model parameters were fitted to these 5 countries, which included Germany,

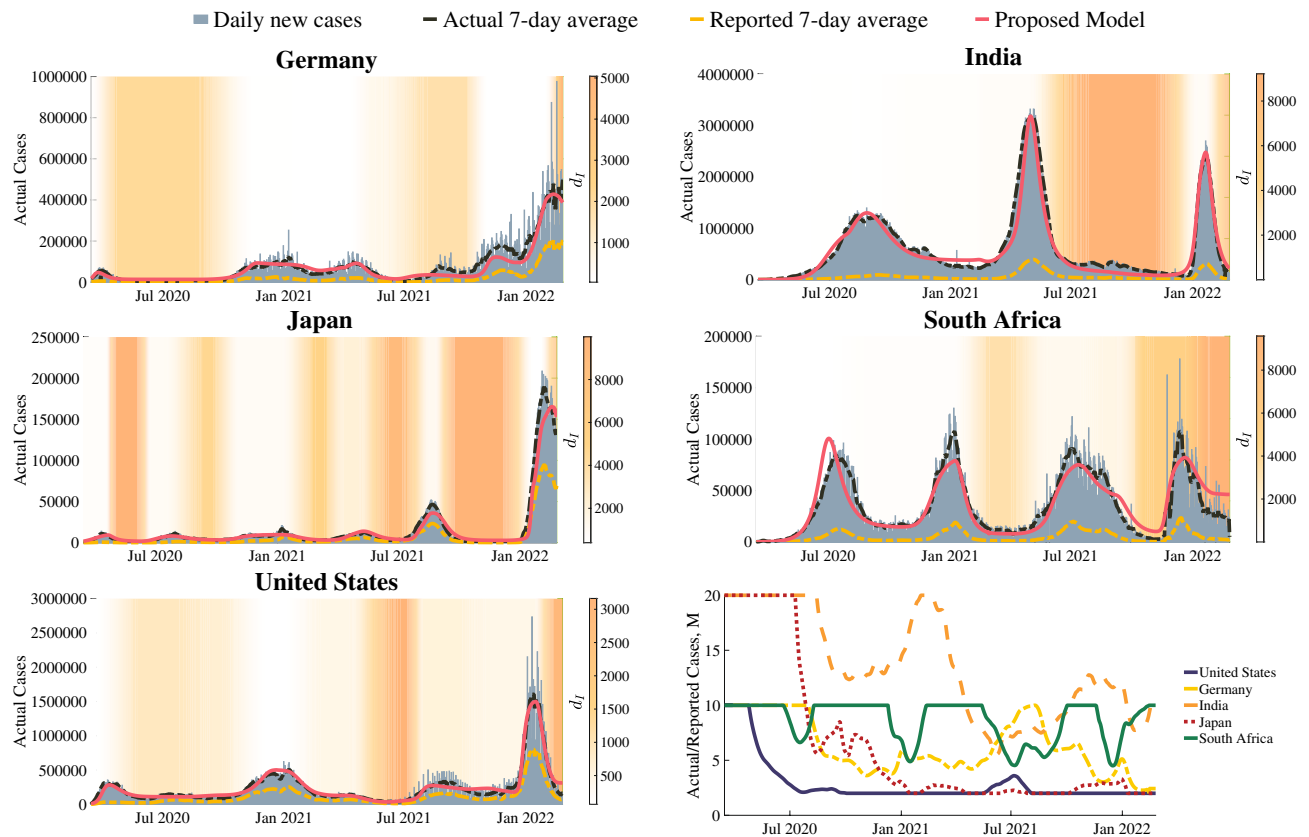


Figure 2. The results of the proposed model fit to Germany, India, Japan, South Africa, and the United States COVID-19 new cases. The model predictions (red, solid) are plotted over the reported 7-day average cases (gold, dashed), estimated actual 7-day average cases (black, dashed), and estimated actual daily cases (gray bars). The level of caution, d_I , values from the model fit are plotted as a white-to-orange vertical gradient over time in the plot background, where the color shades of vertical bars correspond to the estimated level of caution. The bottom right plot shows the case underreporting factor, M , for each of the five countries over the course of the simulation. This factor is multiplied by the reported cases to estimate actual cases.

India, Japan, South Africa, and the United States. γ , β_0 , k , Ω , and μ are all variant-specific constants, of which γ , Ω , and μ were assumed to be consistent between variants. η_1 and η_2 were assumed to be constant for each vaccine type and variant combination. α_1 , α_2 , and α_B were determined from Our World In Data⁵⁸ and we used the time-varying reported vaccination rates to fit the model to each country's case history. The prevalence of each variant over time was estimated from the reported sequence prevalence data of each country from GISAID^{59–61}. To simplify our simulations, new variants were only considered if they reported at least 15% prevalence of all variants at their peak. All other variants were combined into the compartment referred to as the original strain. β_v estimates provide time-dependent transmission rates for each variant; β_v was determined by estimating d_I and d_V as demonstrated in (2), while the β_v of emerging variants were calculated in the same way, but were scaled using a variant-specific factor that was fit to each variant from the prevalence data. As done in previous work²⁵, we introduced several behavioral regions for fitting each country's time-series data, representing discrete times of pronounced behavioral change where d_I changes value. However, the sense of safety (d_V) was assumed to be constant over the simulated region due to the much lower rate of change in the vaccinated population compared to the infected population. With these fixed values in place and using population data⁶² to convert reported cases to a percent of the total population, a gradient-free optimization algorithm in MATLAB was used to optimize the parameters to fit our model to existing variant-specific case data⁶³. The root mean square error between our model prediction and the reported data was minimized.

Results and discussion

Model fit to country-specific daily cases and interpretation. In Fig. 2, we show the model fit to the infection histories of Germany, India, Japan, South Africa, and the United States since March of 2020. These populations were selected to represent a wide range of geographic regions, behavioral and governmental responses to the pandemic, and variant makeups. Solid color vertical stripes in the background of the plots represent behavioral regions in which the population's level of caution was relatively constant, with the gradients between them representing a smooth transition between behavioral regions. In these regions, the value of d_I changes from a previous to a new value over a set period. Within the transition period, the value of d_I varies according to cosine interpolation between its set starting and ending values, as described by Eq. (3):

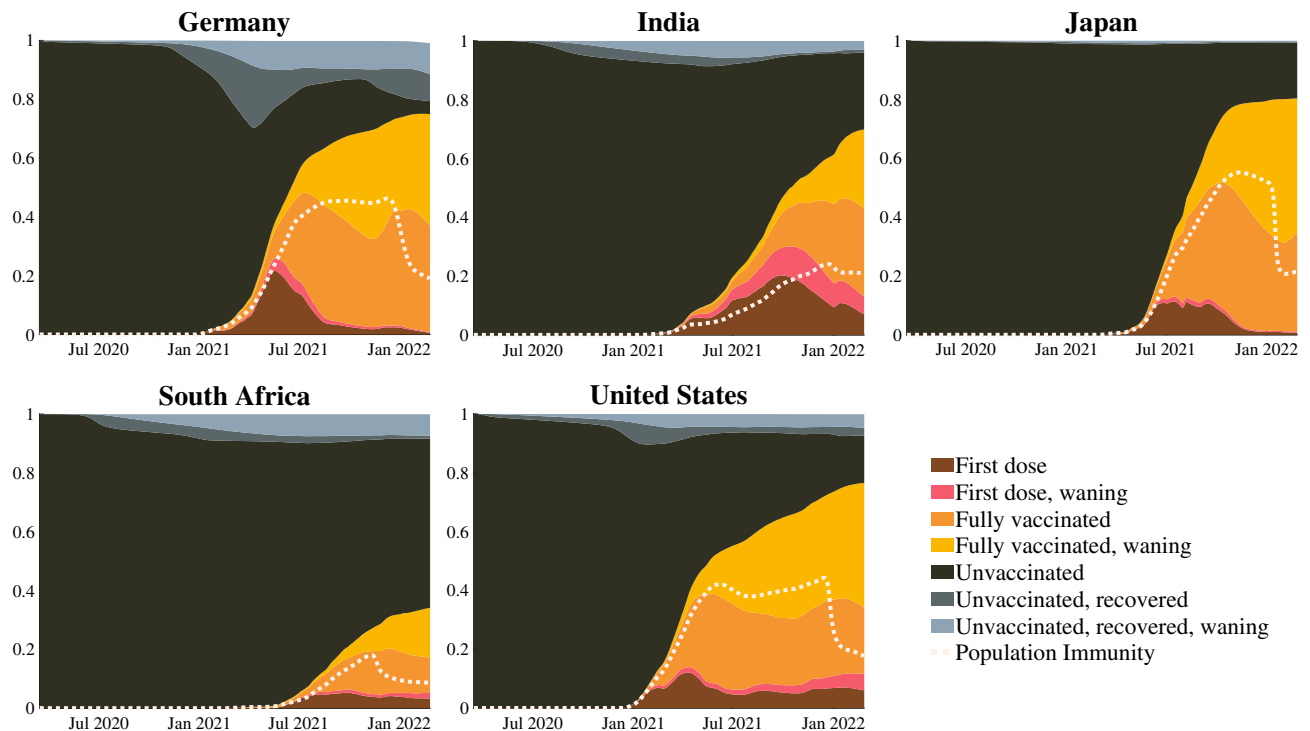


Figure 3. Proportions of various immunity levels in five countries modeled over time. Overlain in the white, dotted line measures population immunity, weighted proportionally to all population groups and prevalent variants. A value of zero represents no population immunity, and a value of one represents complete population immunity and no susceptibility to future infections.

$$d_I(t^*) = d_{I,1} \frac{1 + \cos(\varepsilon\pi)}{2} + d_{I,2} \frac{1 - \cos(\varepsilon\pi)}{2}, \quad \varepsilon = \frac{t^* - t_1}{t_2 - t_1}, \quad (3)$$

where $d_{I,1}$ and $d_{I,2}$ are the values of d_I in adjacent behavioral regions, t_1 and t_2 represent the end and beginning time of these behavioral regions, and t^* and d_{I,t^*} represent the time and value of d_I at an intermediate value between t_1 and t_2 , respectively.

In the two years that we modeled in these countries, significant behavioral shifts in their response to the pandemic occurred about every 3 months. Due to limited available testing, asymptomatic cases, and other societal reasons, the reported number of cases tend to significantly underestimate the actual number of cases in a population. To account for this, we scaled each population's reported cases by a scaling (underreporting) factor, M . This factor is inversely proportional to the number of tests reported each day and fit for each country based on serological testing. As shown in Fig. 2 for each of the modeled countries, this scaling factor was estimated to range from about 20 actual cases per reported case to about 2. It was assumed that M would never be below 2 due to an estimated 40.5% of confirmed COVID-positive individuals being asymptomatic⁶⁴.

Estimating dynamic population immunity. Valuable insights into a population's state in a pandemic can be drawn from the distribution among the compartments and immunity levels in the model. To illustrate population immunity throughout the pandemic, the proportions of the population in each vaccination level, as well as those who are unvaccinated and previously infected, are plotted in Fig. 3. As access to vaccines increases in each country, the proportion of individuals without immunity decays. This visualization of population immunity also illustrates the effects of waning immunity, both for vaccine-induced and infection-induced immunity, predicting that much of the immunized population is waning toward the lower limit of their immunity against infection. This emphasizes the importance of booster vaccinations to maintain immunity and prevent further infection.

The information on the immunity levels and distribution among compartments is used to estimate the average population immunity over time. This is a sum of the immunity of each compartment towards infection, weighted by the distribution of the population amongst these compartments and the prevalence of each of the disease variants. For both vaccination-induced and infection-induced immunity, a significant waning of immunity causes the population's aggregate immunity to decay towards the lower waning levels, requiring regular vaccination to maintain population immunity. Before the advent and distribution of vaccines, each population's average immunity was around zero, indicating complete susceptibility. Through the widespread distribution of the vaccines, average immunity rose fairly consistently in each modeled country until late 2021, when the Omicron variant emerged with reduced protection from the overall vaccines and an increased risk of reinfection. This correlates directly to severe peaks in cases globally. By combining the effects of vaccine-induced immunity, infection-induced

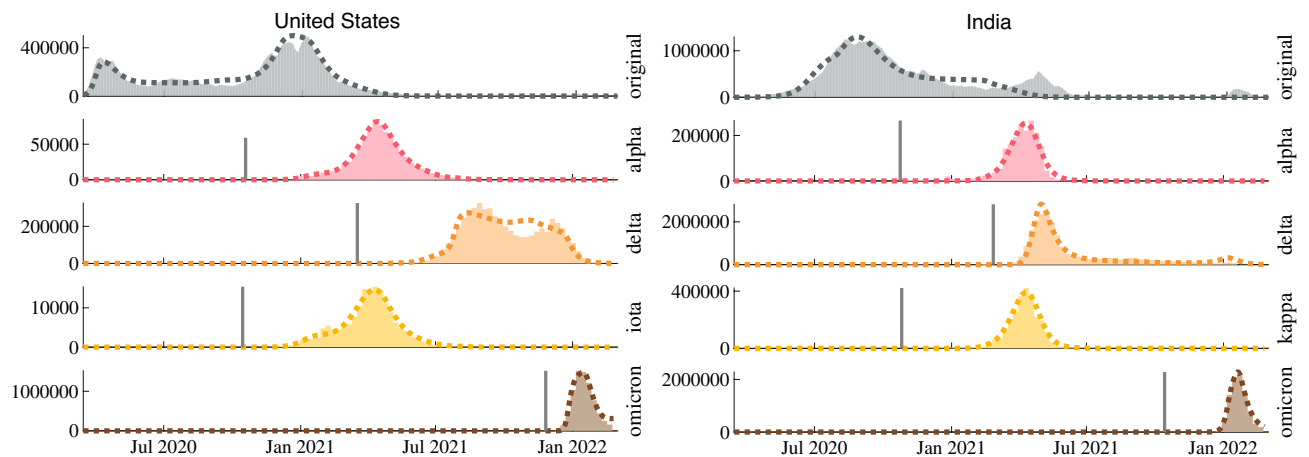


Figure 4. Results of the model fit to each of the dominant variants in the United States and India (dotted), plotted against the estimated actual daily cases (bar). The transmission start date of each variant is indicated by a vertical black bar.

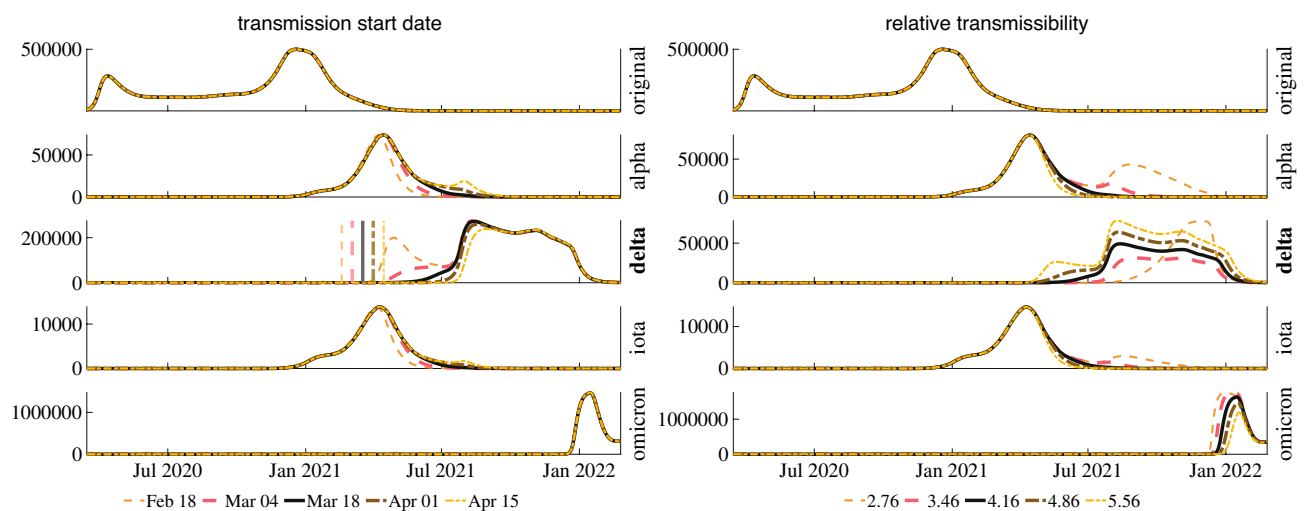


Figure 5. Effect of transmission start date and relative transmissibility of the Delta variant on COVID-19 cases among all dominant variants in the United States.

immunity, waning immunity, reinfection risks, and variant prevalence, the aggregate immunity estimate presented in this paper provides a valuable metric for assessing a population's susceptibility to further infection.

Variant dynamics and sensitivity. The infection histories of each of the dominant variants in the United States and India are shown in Fig. 4, along with the model's ability to reproduce these curves. We selected the United States and India as examples from developed and developing countries. By fitting the model parameters to the observed variant curves, we can deduce helpful insight into the relevant disease characteristics of each variant. For example, in the United States, our model estimates that the Alpha, Delta, and Omicron variants are about 1.8, 4.2, and 21.8 times more transmissible than the original strain of COVID-19. In India, a similar analysis estimates the Alpha, Delta, and Omicron variants to be around 2.0, 5.5, and 11.8 times more transmissible than the original strain. It is important to note that transmissibility is interrelated with vaccine efficacy and infectious period in our model equations. The net population transmissibility estimates are based on reported values for vaccine efficacy against each variant and an infectious period of ten days assumed on an average for all the variants. In addition to estimating transmissibility, the proposed model can estimate the date of the initial spread of each variant in a population. The model's estimate is more specific than estimates such as the first reported case of a variant because it is likely that either many instances of a particular variant were not identified, resulting in an estimate later than the actual first case, or the first identified instance was well contained and did not lead to further spread, resulting in an estimate before the real beginning of the variant spread. Figure 5 explores a sensitivity analysis of the transmission start date and relative transmissibility of the Delta variant, revealing that the model is quite sensitive to small changes in these parameters. This sensitivity further supports our estimates of these variant parameters.

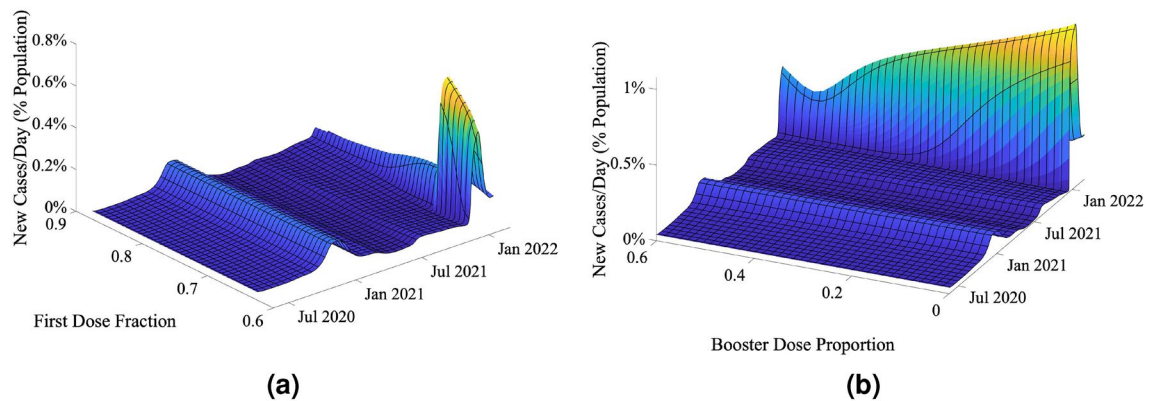


Figure 6. Evaluation of vaccine allocation strategies, assuming that only a set number of the first, second, and booster vaccine doses combined can be administered ($\alpha_{1,prop} + \alpha_{2,prop} + \alpha_{B,prop} = 1$). Sensitivity of pandemic trajectories with respect to (a) a fixed number of doses for the booster vaccines (60% of the total doses being administered), and the relative proportions of first and second vaccine doses being varied, and (b) a variable number of booster doses, where changes in the amount of booster doses were accommodated by changing the proportions of the first and second doses equally.

By conducting simulations using the model with a range of parameters for a single variant and population, we can estimate the sensitivity of our simulation results to perturbations in the variant parameters. Figure 5 shows a sensitivity analysis of the transmission start date and relative transmissibility ($\beta_v/\beta_{original}$). Transmission start date, as described above, is the date used in the model fit to predict the first case that led to widespread transmission, rather than the actual first case or first reported case of a variant, as described in the previous paragraph. The relative transmissibility of a variant is the ratio of the estimated transmissibility of the variant to the estimated transmissibility of the original strain of the virus. The delta variant was chosen to perform the sensitivity analysis because it reached nearly complete dominance over other variants at its peak, and its spread occurred before and after other preeminent variants. With this analysis of all key variants, a clear competitive interaction between the prevalent variants is observed. As the transmissibility and infectious period of a particular variant was increased, or it was introduced at an earlier date, its infections increased, and the infections of all other variants that occur following its initial spread decreased. Notably, a similar sensitivity analysis on the infectious period of a particular variant shows the same trends as relative transmissibility, supporting a metric such as R_0 that combines transmissibility and infectious period to describe the nature of a variant.

Vaccine allocation study. Given the model's ability to predict pandemic trends concerning different variants, we investigated its ability to draw conclusions for differing vaccine distribution rates using the United States as an example case. Figure 6 shows several contemplative/retrospective studies performed using the model. To perform these simulations, at each time point, the total number of vaccine doses available was determined from past vaccine data, and these vaccines were redistributed among the first, second, and booster doses, given a certain proportion of each. While the first and second doses of the vaccine in question were imposed at the first recorded date of vaccination in the US (December 14, 2020), the booster dose was started much later (August 17, 2021) when the booster doses became available. Thus, before the advent of the booster dose, vaccines were just distributed between the first and second vaccine doses. In addition, a parameter was added to account for the fact that a certain proportion of the population will refuse to take any vaccine dose. A threshold value was placed on the layer within the stacked model representing unvaccinated individuals to set α_1 to 0 if the proportion of the population in this group dropped below the threshold. The same value was used to threshold the layers of the stacked model containing individuals with the first dose of the vaccine (i.e., first dose and first dose, waning). If the total individuals in these compartments dropped below this threshold, α_2 was set to 0. For this part of the trend study, the population that does not take vaccination was assumed to be 10% of the total population.

In Fig. 6a, dose allocation between the first and second doses of the vaccine is examined. The pandemic trajectories over time are plotted for different values of the fraction of the first dose, defined in Eq. (4) as

$$x_{first\ dose} = \frac{\alpha_{1,prop}}{\alpha_{1,prop} + \alpha_{2,prop}}, \quad (4)$$

where $\alpha_{1,prop}$ and $\alpha_{2,prop}$ are the proportions of the total number of vaccine doses at each time point allocated to the first and second vaccine doses, respectively. In Fig. 6a, the pandemic trajectories are relatively stable for large values of $x_{first\ dose}$ close to 1. However, the peak in infectious cases around January 2022 grows rapidly for $x_{first\ dose} \approx 0.67$. This indicates that when the first dose of the vaccine is prioritized significantly over the second dose (by a ratio of approximately 2:1), the pandemic trajectories tend to be relatively stable and do not significantly vary when the dose prioritization is changed a small amount. However, when the first dose becomes less prioritized, there is a critical threshold below which infectious cases can proliferate significantly.

Figure 6b shows the results of the booster vaccine dose being varied in relation to the other doses. The proportions of the first, second, and booster doses, $\alpha_{1,prop}$, $\alpha_{2,prop}$, and $\alpha_{B,prop}$ respectively, were initially set to

0.25, 0.15, and 0.60. The sum of the proportions of the three doses was always set to be 1. The proportion of the booster dose was varied, and each change in proportion was reflected by changing the proportions of the first and second dose proportions in equal amounts. From Fig. 6b, it is clear that as booster doses increase, the peak in infectious cases around January 2022 begins to decrease until a booster dose proportion of approximately 0.5, where the number of infectious cases begins to increase. This could happen due to insufficient numbers of individuals being given the base levels of immunity from the first and second doses, with too much prioritization given to individuals who had already been vaccinated twice, leading to a proliferation in infectious cases.

Closing remarks

The proposed modeling method can recreate the history of the COVID-19 pandemic up to the spring of 2022. It can estimate future case trends accounting for variants, waning immunity, population behavior, and vaccine roll-out strategies. This infectious disease dynamic modeling framework can be used for future population disease studies to study the effects of several critical aspects of the pandemic, such as behavioral mandates, the advent and interaction of variants, and the dispersion of multi-dose vaccines with waning immunity. In this paper, we propose methods of quantifying various population-level metrics that are useful in the analysis of a pandemic, such as the relative strength of variants, the aggregate immunity of a population, and the effectiveness of various vaccination rollout strategies. We have included a few examples that provide insight into the competitive nature of disease variants and the outcomes of a range of vaccine allocation strategies. As with any mathematical model, several limitations exist. While the level of caution and sense of safety are accounted for, the model does not fully account for complex, locally heterogeneous social structures. We assume homogeneous mixing of populations and a fixed population level that does not reckon large-scale emigration and immigration of individuals. As travel resumes and if conflicts or economic opportunity-drive mass migration, ignoring the movement of populations between countries may limit the compartmental model's ability to predict future cases accurately. Heterogeneity in the susceptibility of individuals within a population can have a significant effect on the infection trajectory, so further studies can investigate how heterogeneity based models⁶⁵ interact with the model presented here. Another potential limitation of the proposed model is how we account for the underreporting of cases. Since we do not have serological data that can provide insight into the under-reporting of cases after the introduction of the vaccine, the ability to estimate the actual number of infected cases reduces for these circumstances. The model's predictive capabilities will be limited by the uncertainties associated with evolving COVID-19. Some of the infection, variants and population specific estimated COVID parameters may get adjusted over time as new information becomes available. The model and its framework have the capability to incorporate new and more accurate data as they become available. With correct model parameter inputs, it is one of the first models that allow studying coupled effects of vaccine roll-out strategies, vaccine efficacies, VOCs, waning immunity, population level of caution, sense of safety, and the under-reporting of cases to predict potential infectious disease trends.

Data availability

The datasets analyzed during the current study are currently available in the following repositories. COVID case data were obtained from the Center for Systems Science and Engineering (CSSE) COVID-19 Data Repository at Johns Hopkins University⁶², <https://github.com/CSSEGISandData/COVID-19>. Country-specific prevalence of each variant was obtained from the GISAID EpiCovTM database^{59–61}, <https://www.gisaid.org>. Daily vaccinations from each country were provided by Our World In Data⁶⁶, <https://ourworldindata.org/coronavirus>.

Code availability

The MATLAB source code for the model is available on GitHub, <https://github.com/zlajoie/covidmodeling.git>.

Received: 26 June 2022; Accepted: 22 November 2022

Published online: 27 November 2022

References

1. Johns Hopkins University Coronavirus Resource Center. Accessed 30 Apr 2022. <https://coronavirus.jhu.edu/> (Johns Hopkins Coronavirus Resource Center, 2022).
2. Ke, X. & Hsiao, C. Economic impact of the most drastic lockdown during COVID-19 pandemic—The experience of Hubei, China. *J. Appl. Economet.* **37**, 187–209. <https://doi.org/10.1002/jae.2871> (2022).
3. Chen, X., Huang, H., Ju, J., Sun, R. & Zhang, J. Impact of vaccination on the COVID-19 pandemic in US states. *Sci. Rep.* **12**, 1554 (2022).
4. Telenti, A. *et al.* After the pandemic: Perspectives on the future trajectory of COVID-19. *Nature* **596**, 495–504 (2021).
5. Sah, P. *et al.* Accelerated vaccine rollout is imperative to mitigate highly transmissible COVID-19 variants. *eClinicalMedicine* **35**, 100865 (2021).
6. Zimmer, C., Corum, J., Wee, S.-L. & Kristoffersen, M. Coronavirus vaccine tracker. *The New York Times*. <https://www.nytimes.com/interactive/2020/science/coronavirus-vaccine-tracker.html> (2020).
7. Ritchie, H. *et al.* Coronavirus pandemic (covid-19). In *Our World in Data*. Accessed 30 Apr 2022. <https://ourworldindata.org/covid-vaccinations> (2020).
8. CDC. *Coronavirus Disease 2019 (COVID-19)*. Accessed 30 Apr 2022. <https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html> (Centers for Disease Control and Prevention, 2022).
9. Nasreen, S. *et al.* Effectiveness of COVID-19 vaccines against variants of concern in Ontario, Canada. *MedRxiv*. <https://doi.org/10.1101/2021.06.28.21259420v2> (2021).
10. Lustig, Y. *et al.* Eurosurveillance | neutralising capacity against delta (b.1.617.2) and other variants of concern following comirnaty (BNT162b2, BioNTech/pfizer) vaccination in health care workers, Israel. *Eurosurveillance* **26**, 2100557. <https://doi.org/10.2807/1560-7917.ES.2021.26.26.2100557> (2021).

11. Katella, K. *Omicron, Delta, Alpha, and More: What to Know About the Coronavirus Variants*. Accessed 30 Apr 2022. <https://www.yalemedicine.org/news/covid-19-variants-of-concern-omicron> (Yale Medicine, 2022).
12. Bernal, J., Andrews, N., Gower, C. & Gallagher, E. Effectiveness of covid-19 vaccines against the b.1.617.2 (delta) variant. *N. Engl. J. Med.* **385**, 585–594. <https://doi.org/10.1056/NEJMoa2108891> (2021).
13. Cherian, S. *et al.* Convergent evolution of SARS-CoV-2 spike mutations, I452R, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India. *BioRxiv*. <https://doi.org/10.1101/2021.04.22.440932v2> (2021).
14. Takashita, E. *et al.* Efficacy of antibodies and antiviral drugs against covid-19 omicron variant. *N. Engl. J. Med.* **386**, 995–998. <https://doi.org/10.1056/NEJMc2119407> (2022).
15. Matrajt, L. *et al.* Optimizing vaccine allocation for COVID-19 vaccines shows the potential role of single-dose vaccination. *Nat. Commun.* **12**, 3449 (2021).
16. Bar-On, Y. M. *et al.* Protection of BNT162b2 vaccine booster against covid-19 in Israel. *N. Engl. J. Med.* **385**, 1393–1400. <https://doi.org/10.1056/NEJMoa2114255> (2021).
17. Kumar, V. M., Pandi-Perumal, S. R., Trakht, I. & Thyagarajan, S. P. Strategy for COVID-19 vaccination in India: The country with the second highest population and number of cases. *NPJ Vaccines* **6**, 1–7 (2021).
18. Ferreras, E. *et al.* Single-dose cholera vaccine in response to an outbreak in Zambia. *N. Engl. J. Med.* **378**, 577–579. <https://doi.org/10.1056/NEJMc1711583> (2018).
19. Casey, R. *et al.* Immunogenicity of fractional-dose vaccine during a yellow fever outbreak—Final report. *N. Engl. J. Med.* **381**, 444–454. <https://doi.org/10.1056/NEJMoa1710430> (2019).
20. The Pfizer BioNTech (BNT162b2) COVID-19 Vaccine: What You Need to Know. Accessed 30 Apr 2022. <https://www.who.int/news-room/feature-stories/detail/who-can-take-the-pfizer-biontech-covid-19-vaccine-what-you-need-to-know>. (World Health Organization, 2022).
21. Balch, B. *Canada Took a Risk Delaying Second COVID-19 Vaccine Doses. Now, Its Vaccination Campaign is One of the Best in the World*. <https://www.aamc.org/news-insights/canada-took-risk-delaying-second-covid-19-vaccine-doses-now-its-vaccination-campaign-one-best-world> (Association of American Medical Colleges, 2021).
22. Zimmer, C. To speed vaccination, some call for delaying second shots. *The New York Times*. Accessed 30 Apr 2022. <https://www.nytimes.com/2021/04/09/health/covid-vaccine-second-dose-delay.html> (2021).
23. Ujjainia, R. *et al.* Effect monitoring and insights from vaccination program of healthcare workforce from a tertiary level hospital in India against SARS-CoV-2. *MedRxiv*. <https://doi.org/10.1101/2021.02.28.21252621v2> (2021).
24. Chemaitelly, H., Tang, P., Hasan, M. R. & Almkudat, E. A. Waning of bnt162b2 vaccine protection against sars-cov-2 infection in Qatar. *N. Engl. J. Med.* **385**, e83. <https://doi.org/10.1056/NEJMoa2114114> (2021).
25. Usherwood, T., LaJoie, Z. & Srivastava, V. A model and predictions for COVID-19 considering population behavior and vaccination. *Sci. Rep.* **11**, 12051 (2021).
26. Steinegger, B., Arola-Fernández, L., Granell, C., Gómez-Gardeñes, J. & Arenas, A. Behavioural response to heterogeneous severity of covid-19 explains temporal variation of cases among different age groups. *Philos. Trans. R. Soc. A* **380**, 20210119. <https://doi.org/10.1098/rsta.2021.0119> (2022).
27. Roda, W. C., Varughese, M. B., Han, D. & Li, M. Y. Why is it difficult to accurately predict the COVID-19 epidemic? *Infect. Dis. Model.* **5**, 271–281. <https://doi.org/10.1016/j.idm.2020.03.001> (2020).
28. Liu, M., Thomadsen, R. & Yao, S. Forecasting the spread of COVID-19 under different reopening strategies. *Sci. Rep.* **10**, 20367 (2020).
29. Ram, V. & Schaposnik, L. P. A modified age-structured sir model for covid-19 type viruses. *Sci. Rep.* **11**, 15194. <https://doi.org/10.1038/s41598-021-94609-3> (2021).
30. Bubar, K. M. *et al.* Model-informed covid-19 vaccine prioritization strategies by age and serostatus. *Science* **371**, 916–921. <https://doi.org/10.1126/science.abe6959> (2021).
31. Kim, D., Keskinocak, P., Pekgün, P. & Yildirim, İ. The balancing role of distribution speed against varying efficacy levels of covid-19 vaccines under variants. *Sci. Rep.* **12**, 7493 (2022).
32. Postnikov, E. B. Estimation of COVID-19 dynamics on a back-of-envelope: Does the simplest SIR model provide quantitative parameters and predictions? *Chaos Solitons Fractals* **135**, 109841 (2020).
33. Brauer, F. Mathematical epidemiology: Past, present, and future. *Infect. Dis. Model.* **2**, 113–127 (2017).
34. Kermack, W. O., McKendrick, A. G. & Walker, G. T. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A* **115**, 700–721. <https://doi.org/10.1098/rspa.1927.0118> (1927).
35. Mak, H.-Y., Dai, T. & Tang, C. S. Managing two-dose COVID-19 vaccine rollouts with limited supply: Operations strategies for distributing time-sensitive resources. In *SSRN Scholarly Paper 3790836, Social Science Research Network*. Accessed 30 Apr 2022. <https://papers.ssrn.com/abstract=3790836> (2022).
36. García, Y. E. *et al.* Projecting the impact of covid-19 variants and vaccination strategies in disease transmission using a multilayer network model in Costa Rica. *Sci. Rep.* **12**, 2279 (2022).
37. Gonzalez-Parra, G., Martínez-Rodríguez, D. & Villanueva-Micó, R.-J. Impact of a new SARS-CoV-2 variant on the population: A mathematical modeling approach. *MedRxiv*. <https://doi.org/10.1101/2021.02.24.21252406v1> (2021).
38. Lipsitch, M. & Dean, N. E. Understanding covid-19 vaccine efficacy. *Science* **370**, 763–765. <https://doi.org/10.1126/science.abe5938> (2020).
39. Read, J. M., Bridgen, J. R. E., Cummings, D. A. T., Ho, A. & Jewell, C. P. Novel coronavirus 2019-ncov (covid-19): Early estimation of epidemiological parameters and epidemic size estimates. *Philos. Trans. R. Soc. B Biol. Sci.* **376**, 20200265. <https://doi.org/10.1098/rstb.2020.0265> (2021).
40. Faucher, B. *et al.* Agent-based modelling of reactive vaccination of workplaces and schools against covid-19. *Nat. Commun.* **13**, 1414. <https://doi.org/10.1038/s41467-022-29015-y> (2022).
41. Liu, Q.-H. *et al.* Model-based evaluation of alternative reactive class closure strategies against covid-19. *Nat. Commun.* **13**, 322. <https://doi.org/10.1038/s41467-021-27939-5> (2022).
42. Gleeson, J. P. *et al.* Calibrating covid-19 susceptible-exposed-infected-removed models with time-varying effective contact rates. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **380**, 20210120. <https://doi.org/10.1098/rsta.2021.0120> (2022).
43. Lau, H. *et al.* Evaluating the massive underreporting and undertesting of COVID-19 cases in multiple global epicenters. *Pulmonology* **27**, 110–115 (2021).
44. Goldberg, Y. *et al.* Waning immunity after the BNT162b2 vaccine in Israel. *N. Engl. J. Med.* **385**, e85. <https://doi.org/10.1056/NEJMoa2114228> (2021).
45. Levin, E., Lustig, Y., Cohen, C. & Fluss, R. Waning immune humoral response to BNT162b2 covid-19 vaccine over 6 months. *N. Engl. J. Med.* **385**, e84. <https://doi.org/10.1056/nejmoea2114583> (2021).
46. Birhane, M. *et al.* COVID-19 vaccine breakthrough infections reported to CDC—United States, January 1–April 30, 2021. *Morb. Mortal. Wkly. Rep.* **70**, 792–793 (2021).
47. Jeyanathan, M. *et al.* Immunological considerations for COVID-19 vaccine strategies. *Nat. Rev. Immunol.* **20**, 615–632 (2020).
48. Altarawneh, H. N. *et al.* Protection against the omicron variant from previous SARS-CoV-2 infection. *N. Engl. J. Med.* **386**, 1288–1290. <https://doi.org/10.1056/NEJMc2200133> (2022).
49. Biswas, R. K., Afiaz, A. & Huq, S. Underreporting COVID-19: The curious case of the Indian subcontinent. *Epidemiol. Infect.* **148**, e207 (2020).

50. Yamamoto, S. *et al.* Seroprevalence of SARS-CoV-2 antibodies in a national hospital and affiliated facility after the second epidemic wave of Japan. *J. Infect.* **83**, 237–279 (2021).
51. Nawa, N. *et al.* Seroprevalence of SARS-CoV-2 IgG antibodies in Utsunomiya city, greater Tokyo, after first pandemic in 2020 (u-CORONA): A household- and population-based study. *MedRxiv*. <https://doi.org/10.1101/2020.07.20.20155945v1> (2020).
52. Silveira, M. F. *et al.* Population-based surveys of antibodies against SARS-CoV-2 in Southern Brazil. *Nat. Med.* **26**, 1196–1199 (2020).
53. Song, S.-K. *et al.* IgG seroprevalence of COVID-19 among individuals without a history of the coronavirus disease infection in Daegu, Korea. *J. Korean Med. Sci.* **35**, e269 (2020).
54. Nah, E.-H., Cho, S., Park, H., Hwang, I. & Cho, H.-I. Nationwide seroprevalence of antibodies to SARS-CoV-2 in asymptomatic population in South Korea: A cross-sectional study. *BMJ Open* **11**, e049837 (2021).
55. Stout, R. & Rigatti, S. Seroprevalence of SARS-CoV-2 antibodies in the US adult asymptomatic population as of September 30, 2020 | infectious diseases | JAMA network open | JAMA network. *JAMA Netw. Open* **4**, e211552 (2021).
56. Murhekar, M. V. *et al.* Prevalence of SARS-CoV-2 infection in India: Findings from the national serosurvey, May–June 2020. *Indian J. Med. Res.* **152**, 48–60 (2020).
57. Harries, M. *et al.* SARS-CoV-2 seroprevalence in Germany—A population based sequential study in five regions. *MedRxiv*. <https://doi.org/10.1101/2021.05.04.21256597v1> (2021).
58. Mathieu, E. *et al.* A global database of COVID-19 vaccinations. *Nat. Hum. Behav.* **5**, 947–953 (2021).
59. Khare, S. *et al.* GISAID's role in pandemic response. *China CDC Wkly.* **3**, 1049–1051 (2021).
60. Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob. Challenges (Hoboken, NJ)* **1**, 33–46 (2017).
61. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data—From vision to reality. *Eurosurveillance* **22**, 30494 (2017).
62. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).
63. *Bound Constrained Optimization Using Fminsearch; Fminsearchbnd, Fminsearchcon*. <https://www.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd-fminsearchcon> (MathWorks, 2022).
64. Ma, Q., Liu, J. & Liu, Q. Global percentage of asymptomatic SARS-CoV-2 infections among the tested population and individuals with confirmed COVID-19 diagnosis: A systematic review and meta-analysis | global health | JAMA network open | JAMA network. *JAMA Netw. Open* **4**, e2137257 (2021).
65. Rose, C. *et al.* Heterogeneity in susceptibility dictates the order of epidemic models. *J. Theor. Biol.* **528**, 110839 (2021).
66. Ritchie, H. *et al.* Coronavirus pandemic (COVID-19). In *Our World in Data*. <https://ourworldindata.org/coronavirus> (2020).

Acknowledgements

The authors would like to thank Brown University's School of Engineering and the Center for Biomedical Engineering for their support.

Author contributions

All authors contributed to the modeling and analysis development and wrote the main manuscript. V.S. conceptualized and supervised the research. Z.L. coded the model and performed variant and sensitivity analysis. T.U. and S.S. performed vaccine allocation analysis. All authors reviewed the manuscript.

Funding

This work was supported by Brown University School of Engineering support to PI V. Srivastava. No external funding was used.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to V.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022