medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Protect or prevent? A practicable framework for the dilemmas of COVID-19 vaccine prioritization

Raghu Arghal[1]*, Harvey Rubin[2], Shirin Saeedi Bidokhti[1], Saswati Sarkar[1],

1 Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA, United States
2 Division of Infectious Diseases, Department of Medicine, University of Pennsylvania School of Medicine, Philadelphia, PA, United States
* rarghal@seas.upenn.edu

1 Abstract

Determining COVID-19 vaccination strategies presents many challenges in light of limited vaccination capacity and the heterogeneity of affected communities. Who should be prioritized for early vaccination when different groups manifest different levels of risks and contact rates? Answering such questions often becomes computationally intractable given that network size can exceed millions. We obtain a framework to compute the optimal vaccination strategy within seconds to minutes from among all strategies, including highly dynamic ones that adjust vaccine allocation as often as required, and even with modest computation resources. We then determine the optimal strategy for a large range of parameter values representative of various US states, countries, and case studies including retirement homes and prisons. The optimal is almost always one of a few candidate strategies, and, even when not, the suboptimality of the best among these candidates is minimal. Further, we find that many commonly deployed vaccination strategies, such as vaccinating the high risk group first, or administering second doses without delay, can often incur higher death rates, hospitalizations, and symptomatic counts. Our framework can be easily adapted to future variants or pandemics through appropriate choice of the compartments of the disease and parameters.

2 Introduction

Since its beginning in December 2019, the COVID-19 pandemic has resulted in nearly 750 million infections and over 6 million deaths as of April 2023 [1]. Vaccines have proven to be the most effective countermeasure to the pandemic by limiting transmissions and protecting especially vulnerable populations [2]. During its early stages, the vaccination drive was heavily capacity constrained with demand far outstripping supply and administration capability – a challenge that continues to plague Low- and Middle-Income Countries (LMICs) [3]. This is bound to be the case for vaccines developed for every infectious disease.

The target population for COVID-19 is naturally heterogeneous with different individuals exhibiting different social contact patterns and different risks for developing a serious form of the disease and suffering hospitalization or death. Thus, under capacity constraints, governments and public health organizations must make the critical choice of whom to vaccinate first: 1) those who are at high risk of hospitalization or death 2) those who are likely to transmit the disease most, 3) a combination of the

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

first and second set. The first group, referred to as the "high risk group", is comprised of the elderly and the immunocompromised; vaccinating them protects them by significantly reducing the risk of severe symptoms, hospitalization, and death. The second, the "high contact group", are those who manifest high contact rates, perhaps because of their professions, and may spread the disease the most; thus, vaccinating them aims to prevent them from infecting others, significantly reducing the reach of the disease in the populace, including in the high risk group. For COVID-19, most public health bodies first vaccinated the high risk group[4][1], adopting the protection route. But was the choice optimal even if we consider the objective of minimizing only the death count? What if the prevention route minimizes the death count instead, by vaccinating first those who constitute the most effective vector of spread, which in turn reduces the spread of the disease to the most vulnerable? In particular, if a small fraction of the populace have much higher contact rates than the rest, then vaccinating them may inhibit in a short time the pathway of the disease to the most vulnerable.

Next, while some vaccines have one dose, others need two doses [5]. For the recipients who opted for the vaccines that needed two doses, the public health bodies, at least in the US, scheduled the second dose right after the mandated minimum time between the two doses, while others await their first doses. But do public health metrics improve if first everyone is administered the first dose and then the second doses are scheduled for those who need them? Or is the optimal strategy a hybrid between the two extremes, i.e. delay the second dose by a certain amount to administer the first dose sooner to larger number of individuals; the delay may be different for different contact and risk profiles as well. To answer these questions, we need a systematic methodology to determine the optimal vaccination strategy, for attaining well-defined public health objectives, considering both single dose and two-dose vaccines. Identifying the optimal vaccination strategy is imperative also because it provides a valuable benchmark for evaluation of any proposed vaccination policy. Comparing a proposed policy with the optimal informs by how much the former can be outperformed through a smarter design.

The challenge in determining the optimal vaccination strategy for COVID-19 is multi-fold. First, COVID-19 manifests some fundamentally different characteristics in susceptibility and fatality rates as compared to other contagious diseases such as various strains of influenza, rabies, which have been known for a while, and, more recently discovered ones such as Zika and Ebola [6]. Specifically, vulnerability to COVID-19 has shown to be much more lopsided with respect to age (presenting greater risk to the elderly) as well as other underlying health conditions such as diabetes, obesity, or hypertension [7]. Thus, given the nature of transmission and vulnerability, the extensive body of research that exists for vaccination for the other diseases mentioned above, could justifiably assume either a fully homogeneous target populace or limit heterogeneity to age-based compartments or geography [8, 9, 10]. For example, studies of rabies have rightly focused primarily on geographic movement of animal populations which are common hosts (eg. [10]). The age or geography based compartments were often considered to differ in contact rates, but not in risks, even when they consider heterogeneity. Thus, the literature on vaccinations for other diseases do not incorporate the heterogeneity in *both* contact rates and risk level which constitutes a defining characteristic for COVID-19. Yet, incorporating this multi-dimensional heterogeneity by considering each individual as a separate entity usually leads us to optimizations of an inordinate size, which are computationally intractable because of the curse of dimensionality and the scale of spread of COVID-19 [11]. On the other hand, homogeneous abstractions, though computationally tractable, lose the essential

[1] Health care workers were vaccinated before the elderly and immunocompromised, but they form of a negligible segment of the high contact group. High contact groups typically also include those with large and dynamic contact sets such as shopkeepers, bank tellers, receptionists, drivers of taxis, shared rides, waiters, bartenders, doormen, etc.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

characteristics of the spread and impact of COVID-19. Even the scale of spread of some of the previous diseases such as Zika or Ebola have been far more limited than COVID-19, with Ebola outbreaks being largely confined to West Africa [12], while COVID-19 afflicted all the continents. The much larger target populace amplifies the computation challenges that arise in determining the right prioritization based on individual characteristics such as contact rates and risk factors.

We redress the above challenges as follows. We model the multi-dimensional heterogeneity for COVID-19 by grouping different sections of the populace in accordance with their risk factors and contact rates. Individuals in each group are assumed to be statistically identical in terms of risk factors and contact rates among each other and across groups, but those in different groups can have different contact rates and risk factors. The contact rates between groups depend on pairs in question. This group-based modeling provides a tradeoff between retaining analytical and numerical tractability and capturing the inherent heterogeneity. The number of groups and the specific criteria for the classification is a design choice. Vaccination in different groups however remains coupled due to capacity constraints on overall rate of vaccination across all groups together. We subsequently formulate the determination of the optimal vaccination strategy that optimizes public health metrics of choice (eg, death, hospitalization, symptomatic counts) by allocating different vaccine capacities to different groups at different times subject to not exceeding the limit on the number of doses that can be administered (i.e. the vaccination capacity constraint).

There is a growing literature attempting to provide guidance on COVID-19 vaccine prioritization. Most papers compare a set of pre-identified policies using simulations e.g. [13, 14, 15, 16, 17, 18]. The most common among these policies prioritize based on age (e.g. [13]), others based on geographic locations etc [14, 15, 16]. Among these, a small set focus on LMIC settings (e.g. [17]) with the large majority confined to high income countries [18]. Incidentally, among prior pandemics, influenza may be the closest to COVID-19, and most vaccination strategies for influenza fall into similar categories, relying either on simulation of very few strategies (e.g. [19, 20]) or computationally costly optimization with limits on the feasible set of policies (e.g. [21, 22]). Such comparisons and evaluations often provide valuable insights, but the overall approach suffers from some fundamental limitations. First, in absence of the knowledge of the optimal vaccination strategy for attaining well-defined public health objectives, one does not know whether and by how much a future strategy, or even all past strategies, outperform the proposed strategies. Second, simulations require a large number of iterations to provide reliable performance estimates and each such iteration consumes significant time for even moderate size populaces, and the computation time and processor memory requirement for each iteration grows at least polynomially with the size of the populace. As a result, the works in this category which consider populations of the size that pandemic spread affects rely heavily on access to powerful computing resources like high performance computing clusters (e.g. [23, 24]). These simulations cannot therefore be conducted in settings which have limited computational resources such as local public health bodies in small towns of the US and throughout LMICs, which prevents the choice between strategies from being fine-tuned to parameters that evolve with time and can only be locally estimated. Barring few exceptions, most works consider single dose vaccines, though several COVID-19 vaccines require two doses. For example [25, 26] consider two dose vaccines, but both of these prioritize based only on age, as per a predetermined order, and choose between few predetermined delays for the second dose. The joint optimization of group prioritization and second dose delay has remained open even in the category that compares a set of predetermined strategies.

Another computational approach chooses the optimal vaccination strategy for COVID-19 among the policies that do not vary the vaccination rates to groups over

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

time (i.e. among static policies) (e.g. [27]), or among those that vary allocations over large predetermined time intervals (e.g. [6]). The limitation of these approaches is that they *a priori* rule out highly-dynamic vaccine allocations without considering if these can attain substantially better values of public health metrics. Besides, the appropriate time scale of optimization can not be determined *a priori* and universally, as different public health domains have different inherent flexibilities. The available approaches that optimize over predetermined fixed intervals can not easily generalize to different, specifically higher, numbers of intervals, i.e. they do not scale. For example Buckner et. al. deploy a two-step genetic algorithm with simulated annealing [6]. The computation time of genetic algorithms increases exponentially in the number of optimization variables. The number of optimization variables for Buckner et al. linearly increases in the number of intervals. Thus, the computation time significantly increases if the number of intervals increases (refer to Supporting Information 5 for more details). Again, these approaches may be prohibitively difficult in areas lacking advanced computing or the necessary expertise. All of the above only consider single dose vaccinations, and the generalization to two doses is not direct and is expected to increase the computation time even further and significantly so.

We adopt a different design approach altogether. We do not restrict to any preselected set of vaccination strategies, or any predetermined decision interval. We instead consider the problem in its most general form, optimizing over all potential vaccination strategies, including arbitrary functions of time which can vary over any time scale and arbitrary variations across the groups, towards the desired public health metric subject to the vaccination capacity constraint. We obtain a flexible framework that can accommodate different number of doses and different public health metrics (e.g. death, hospitalization, symptomatic counts), and tailor the vaccination strategy to any contact and risk heterogeneity of the target populace, population demography, disease parameters, and vaccination capacity constraints. We accomplish this in three broad steps: (1) capturing the heterogeneity of the populace in both contact rates and risk factors through a novel partition of the populace into three groups: high contact, high risk, and baseline (Section 3) (2) modeling the spread of the contagion and the progression of the stages of the disease within and across the groups, as a system of a small number of ordinary differential equations (ODEs) in a small number of variables, regardless of the size of the populace (Section 3) (3) posing the choice of the optimal capacity constrained vaccination strategy as an optimal control problem with the ODEs providing the state trajectories (Section 3). Optimal control has been deployed in designing dynamic vaccination prioritizations for several other contagious diseases (e.g. [28, 29]), but the deployment for COVID-19 will fundamentally differ on account of its distinctive characteristics. Overall, very few papers have applied optimal control to design of COVID-19 vaccine strategies [30, 31, 16]. [30] and [31] both deploy optimal control to determine the optimal budget allocation between several different types of mitigation strategies for COVID-19 (e.g. social distancing, vaccination, testing) assuming that the target populace is fully homogeneous. [16] focuses on allocation of available vaccines among large sub-regions (e.g. provinces of a nation, with Italy as their case study) of a geographic region so as to minimize cumulative infections. They consider the population in each sub-region to be homogeneous in that every individual has the same risk factor and contact rate therein. In the above respective scenarios, the fully homogeneous abstraction, or the homogeneity assumptions within each sub-region rules out design of vaccine prioritization among individuals based on their contact rates and risk factors. Thus, the problems they seek to solve are complementary to those in this paper. Also, note that in practice different individuals within large or even small sub-regions (e.g. neighborhoods) have widely diverging contact rates (e.g. due to their professions) and risk factors (e.g. due to their age and underlying health conditions).

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Thus, even the vaccine rollout processes deployed in practice have enacted different priorities among different individuals living even in proximity [4]. Thus,our consideration of heterogeneity captures a crucial element of reality.

We compute the optimal vaccination strategy for 911,250 instances of realistic parameters (see Supporting Information 4.2 for parameter selection), spanning different variants, disease parameters and population demographics of 139 countries and all states of US as well as specific case studies such as retirement homes, prisons, LMICs. The optimal solution, or a close approximation thereof, turns out to be easily deployable, and can mostly be computed within seconds regardless of the population size using standard numerical toolboxes and modest computation resources, which are usually accessible in local public health offices throughout US and in most LMICs. Owing to this computational tractability, we are able to evaluate the impact on the solution of 1) a range of disease and population parameters that arise in practice including in divergent case studies of interest, 2) restrictions on decision intervals, 3) error in estimating the parameters. Finally, we show how our framework can be used by public health authorities in designing vaccine rollout strategies for future pandemics (Section 5).

We summarize some specific findings from our work. We first list those that are particularly relevant for practitioners.

- For one dose vaccines, we could narrow down the set of optimal strategies to only two candidates: one which first vaccinates the high contact group, then the high risk group and finally the baseline group, the other reverses the order between the high contact and high risk groups (Section 4.2). One or other of these two policies optimizes important public health metrics among all policies for the bulk of the parameters considered, and nearly minimizes in the limited number of remaining instances. This is a significant reduction starting from the set of all possible vaccination strategies. Further, both candidate policies are easy to deploy. while the commonly recommended high risk prioritization was the optimal policy in some instances, in many realistic cases (e.g. nursing homes and LMICs), the high contact priority policy has a lower death count than the high risk priority policy which is currently the most widely deployed policy; for some realistic cases the high risk priority policy has higher death count than even a policy that accords uniform priority to all individuals i.e. randomly selects among them (Sections 4.3 and 4.4). This is somewhat counterintuitive because the high risk priority policy prioritizes the vaccination of those at greatest risk to die once infected, but has been explained through the insights the computations provide.

- For two dose vaccines, we find that the set of optimal vaccination strategies can be narrowed to three easy to deploy candidates, which vary in the order of selection of groups for vaccination and on whether the second dose is delayed until everyone receives the first dose or each group gets the two doses in succession while others wait for the first dose (Section 4.6). It is again not obvious *a priori* that the candidate set for two doses will increase only by a small number compared to that for one dose given the several additional decision variables in the former, including one decision variable, namely the delay for the second dose, assuming uncountably infinite number of possible values being a real number.

These two findings imply that the optimal vaccination strategy, or a close approximation thereof, can be determined by comparing a small set (i.e. either two or three) of vaccination strategies, for both one and two dose vaccines. The comparisons can again be done in a few seconds using our framework even when the computation resources are modest. This facilitates the universal deployment of the framework by enabling fine tuning the choice between the few candidate policies to specific parameters that can only be identified locally.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

We now highlight the findings that have significance from a methodological view.

- We show that despite segmenting the population into only three groups (baseline, high risk, and high contact), when the parameters are chosen correctly, our model is able to accurately predict both the spread of infection and death count of COVID-19 as observed over the course of the pandemic for all US states and 139 other countries including most LMICs (Section 4.1). Thus, our model is universally applicable. We also demonstrate that any decrease in the number of groups substantially undermines the accuracy of prediction. In this sense, our model has just enough complexity to characterize the state dynamics of COVID-19 so as to match actual data of spread and inflicted mortality.

- We demonstrate the use of optimal control as an alternative to simulation or other, more computationally expensive optimization techniques for determining vaccine prioritization strategies for pandemics in which the target populace manifests heterogeneity in both contact rates and risk factors.

Overall, our framework for the determination of the vaccination prioritization is practicable in (1) the ease of deployment of all the resulting candidate solutions, (2) the fast computation of the optimal solution even when the computation resources are modest as in local public health units in US and LMICs which allows such units to fine tune the solution they deploy in accordance with the local pandemic parameters, and (3) the ability of the underlying model to capture key public health metrics almost universally.

3 Methods

COVID-19 manifests heterogeneity both in the evolution of the disease in individuals and in the manner of the spread. Different sections of the populace constitute the most potent vector of spread while others have high risk for developing a severe form of disease once infected. We capture this heterogeneity through a simple classification of the populace. We divide the populace of N people into three groups: 1) baseline (X) 2) high risk (Y), and 3) high contact (Z). We describe the rationale and method underlying this classification. The *high risk* are those who are old or have underlying health conditions, which exacerbate the risk of developing severe forms of the disease, namely symptoms, hospitalization and death, once infected, e.g. retirees [7]. The *high contact* group are those who have a large set of potential contacts spanning entire neighborhoods or even cities, but establish actual contact with any individual in their set of potential contacts infrequently. The *baseline* is the rest of the populace, i.e. those who are neither high contact nor high risk.

The high contact group includes shopkeepers, bank tellers, receptionists at hotels, restaurants, waiters at restaurants, doctors, nurses, drivers of buses, cabs and shared ride, etc. For example the set of potential contacts of a shopkeeper or a driver includes entire neighborhoods or cities, but a shopkeeper or a driver meets any given member in their neighborhood or city infrequently. This group has overall high rates of contact with both of the other groups. Given their large and dynamic contact sets, individuals in this group are the most likely to imbibe an infectious disease early on and pass the same to a large number. Thus, this group constitutes the most potent vector of spread. The individuals who are not high contact, have small sets of "regular" contacts, (e.g. family members, colleagues, friends), with whom they connect frequently, and outside this set infrequently connect with members of the high contact group. The baseline group includes but is not limited to workers of the tech sector, financial sector, manufacturing sector, homemakers, and scientists. Note that a tech worker or a scientist

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

would regularly meet their friends, family, colleagues - all of whom constitute a small set; and infrequently meet shopkeepers, bank tellers, waiters, drivers in their neighborhood or city. Thus, members in this group largely, albeit indirectly, connect through those of the high contact group.

Considering publicly available data for the infection and death counts as a function of time in all states of US and 139 countries, we later show that the decomposition of the populace in these three groups helps capture the temporal variation of infection and death counts observed in reality (Section 4.1). The classification based on contact patterns and risk factors also allows us to capture the tension between focusing vaccination on high risk or high contact groups. Clearly the classification is comprehensive in that it includes all individuals. In principle there can be an overlap between high contact and high risk groups, but in practice this overlap is minimal because the individuals in high contact group are of working age and usually in good health. Thus, their risk of developing a severe form of the disease is generally low. In principle, the vaccination priority for the small fraction of individuals who are both high contact and high risk ought to be the highest. Thus, they can be excluded from the decision problem we consider, i.e. we consider the problem of allocation of vaccines to individuals after those who are both high contact and high risk are vaccinated. Thus, the classification consists of disjoint groups which together cover the populace.

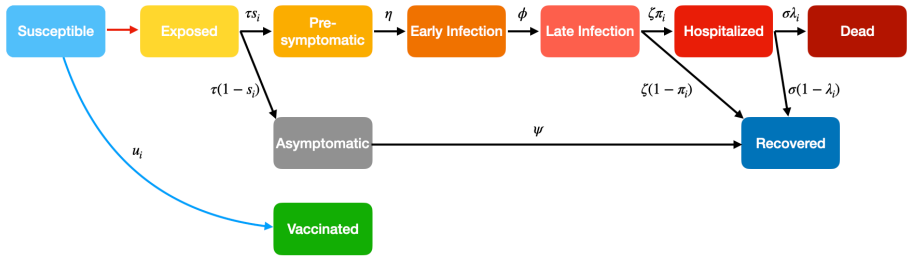
We also show, using the above mentioned data, that decomposition into fewer groups, namely one (i.e. entire population is one homogeneous group) or two (separating individuals based only on risk factor or only on contact patterns) can not capture the temporal variation of both infection and death counts (Section 4.1). Thus, our classification is the simplest possible for capturing the heterogeneity in the contact and risk patterns of the populace which determine the spread and evolution of COVID-19.

To describe state transitions modelling the contraction and evolution of COVID-19 in individuals, we employ an expanded compartmental model that accounts for the disease states associated with COVID-19 as well as our three group partition. We imbue each group with a compartmental model of the disease progression which includes 10 states: susceptible (S), exposed (E), pre-symptomatic (P), asymptomatic (A), early stage infected (I), late stage infection (L), hospitalized (H), recovered (R), vaccinated (V), and dead (D). In general, we refer to the fraction of individuals in a disease state with the group as a subscript and indexed by time e.g. $S_X(t) = \frac{N_X(t)}{N}$ where $N_X(t)$ is the number of susceptible, type X individuals at time t and N is the the total number of individuals.

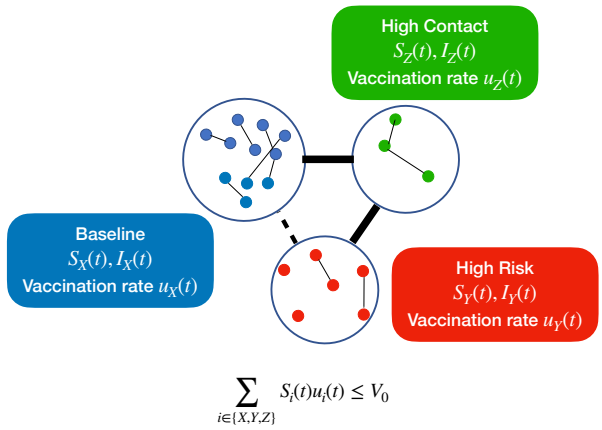
For ease of exposition, we first present a simplified model in which vaccines confer perfect immunity to infection (see Figure 1), an assumption relaxed in the model we actually use for all our numerical results (refer to details of this model in Supporting Information 2.1).

The vaccine policy is specified as the fraction of the susceptibles of each group who will receive the vaccine at a given time instant t $(u_X(t), u_Y(t), u_Z(t))$. Thus, $0 \leq u_i(t) \leq 1$ at any given time t. We are therefore assuming that only individuals who do not have active infection, ie, susceptibles are vaccinated. This assumption is in line with CDC recommendations which state that vaccine should be delayed at least until symptoms and isolation criterion subside and ideally until 3 months from the onset of symptoms or positive testing [33]. Furthermore, when vaccine capacity is severely limited (as is bound to be the case early on in vaccine availability or in low resource settings), it is especially important that only susceptible individuals who can reap full benefit receive vaccines [13]. Note that symptomatics can be easily excluded from vaccination based on manifestation of symptoms. The exposed, asymptomatics and presymptomatics can also be excluded if all individuals are tested before they are vaccinated. Usually by the time vaccines are available for any infectious disease, tests

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



(a) Simplified COVID state diagram



(b) Contact and vaccination model

Fig 1. In 1a we present a simplified state diagram in the case where vaccinated individuals and recovered individuals cannot become infected. Note that these assumptions are relaxed for our numerical computations (see Supporting Information 2.1 for further detail). Red arrow indicates exposure to a contagious individual, black arrows denote natural disease progression, and blue arrow denotes vaccination. Each transition (apart from exposure which depends on infected population) is labeled with the associated rate. Note that probability of becoming symptomatic (s_i), hospitalized (π_i), and dying (λ_i) are group-dependent as denoted by subscript i. Here τ, η, ϕ, ψ, ζ, and σ are the reciprocals of the expected durations of the exposed, pre-symptomatic, early infection, asymptomatic, late infection, and hospitalized states, respectively. Refer to Supporting Information 1 for notation tables. 1b depicts our three group model which dictates how susceptible and infectious individuals of different groups come in contact to spread the disease and how vaccines are allocated. Here I denotes a group's vector of all infectious disease states. Note that as vaccination rate $u_i(t)$ for a group increases the fraction of individuals that transitions to the path that takes them to hospitalization or dead states decreases. Given the respective vaccination rates and number of susceptibles, the vaccination capacity constraint is expressed as $\sum_{i \in \{X,Y,Z\}} S_i(t) u_i(t) \leq V_0$.

for the same are widely available and are also cheap, thus testing before vaccination can be easily realized. But if vaccination can not be preceded with tests, then the exposed, presymptomatic, and asymptomatic individuals will also receive vaccine doses, we show how our framework can be adapted to this setting in Supporting Information 2.

Note that *all* vaccination policies can be represented through appropriate choice of the functions $u_i(\cdot)$. For example, one can consider policies which sequentially vaccinate groups, e.g. $u_i(t)$ is as high as possible for a certain group initially, while $u_j(t) = 0$ for

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

$$\dot{S}_i(t) = -S_i(t)u_i(t) - S_i(t) \sum_{j \in \{X,Y,Z\}} \beta ij[P_j(t)\rho + A_j(t)mu + I_j(t)\omega] \quad (1)$$

$$\dot{E}_i(t) = -E_i(t)\tau + S_i(t) \sum_{j \in \{X,Y,Z\}} \beta ij[P_j(t)\rho + A_j(t)mu + I_j(t)\omega] \quad (2)$$

$$\dot{P}_i(t) = E_i(t)\tau s_i - P_i(t)\eta \quad (3)$$

$$\dot{A}_i(t) = E_i(t)\tau(1 - s_i) - A_i(t)\psi \quad (4)$$

$$\dot{I}_i(t) = P_i(t)\eta - I_i(t)\phi \quad (5)$$

$$\dot{L}_i(t) = I_i(t)\phi - L_i(t)\zeta \quad (6)$$

$$\dot{H}_i(t) = L_i(t)\zeta\pi_i - H_i(t)\sigma \quad (7)$$

$$\dot{V}_i(t) = S_i(t)u_i(t) \quad (8)$$

$$\dot{R}_i(t) = A_i(t)\psi + L_i(t)\zeta(1 - \pi_i) + H_i(t)\sigma(1 - \lambda_i) \quad (9)$$

$$\dot{D}_i(t) = H_i(t)\sigma\lambda_i \quad (10)$$

$$i \in \{X, Y, Z\}$$

Fig 2. The figure presents a set of ordinary differential equations (ODEs) which captures the dynamics of the states shown in Figure 1. All notation and associated parameters are defined and presented in full detail in Supporting Information 1. These ODEs have been obtained through an adaptation of the metapopulation model of epidemiological differential equations [32]. Each equation governs the rate of change of the fraction of individuals of a specific group and disease state. The equations are composed of both quadratic and linear terms. The quadratic terms (terms in red) correspond to the transmission of the virus, which involves interaction between two individuals, one contagious and another susceptible. The number of such interactions per unit time is linearly proportional to the number of such pairs, which is in turn linearly proportional to the product of the fraction of susceptibles and contagious individuals in the respective groups. The proportionality constant is the disease spread rate. Linear terms correspond to either (1) the progression of an individual through disease states (terms in black) or (2) the vaccination of susceptible individuals (terms in blue). The proportionality constants for (1) are the transition rates in and out of the states which are different for different groups. The proportionality constants for (2) are the vaccination rates which are functions of time and group. This system of ODEs consists of only 30 ODEs involving 30 variables, regardless of the population size. Note that this represents the simplified setting in which vaccinated and recovered individuals cannot become infected (corresponding to Figure 1).

other groups in the same period, then the focus changes to $u_j(t)$ for another group, etc. Or, $0 < u_i(t) < 1$ for all groups in a certain interval, which represents "mixed" policies which split vaccine capacity between multiple groups. In fact, $u_X(t), u_Y(t), u_Z(t)$ are allowed to be arbitrary functions of time, Thus, highly dynamic, complex policies are included among the vaccination strategies we consider.

We consider that the vaccination capacity is limited, i.e. V_0 fraction of the population (NV_0 individuals) can receive vaccines on any given day. Thus, $\sum_{i \in \{X,Y,Z\}} u_i(t)S_i(t) \leq V_0$ for each t. This vaccination capacity constraint is motivated by staffing and infrastructure limitations that arise in particular for emerging infectious diseases. Specifically the health care workers who administer the vaccines are limited. Also, only a certain number of doses can be stored on any given day in a health care unit due to limitations on cold chain storage space [34]. Thus, only a certain maximum number of doses can be administered on any given day leading to the hard constraint under consideration [34].

The dynamics of the system can be captured by a system of ODEs [2] provided in

[2]While the set of ODEs described is deterministic, the state transitions are usually stochastic. It can be shown, however, that, under certain mild regularity assumptions, the fraction of individuals in any given state and group in the stochastic system converges to the solution of the ODEs at each time t in the asymptotic limit that the size of the population increases to infinity ([35]; Supporting Information C, [32]). The assumption is that the duration of each state for every individual is exponentially

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 2. The state dynamics of the three groups are connected through (1) the infection across groups and (2) the dependence between the vaccination rate $\{u_i(\cdot)\}$ allocated to each group at time t. The dependence arises because of the vaccination capacity constraint. If the capacity allocated to one group is high at any given time, those allocated to the other groups must be low because of this capacity constraint.

We now seek to optimize public health metrics of our choice through a judicious selection of the vaccination strategy, namely the $\{u_i(t)\}$ for $i = X, Y, Z$, $t \in [0, T]$. Unless otherwise specified, we focus on minimizing the overall death count by the end of the time horizon under consideration, i.e. $N(D_X(T) + D_Y(T) + D_Z(T))$ (this is also the cumulative death count in $[0, T]$). We later describe how our framework can be adapted to minimize other public health metrics such as symptomatic counts, hospitalization counts, or socioeconomic costs related to the pandemic (Supporting Information 2.3).

The optimal vaccination strategy is the one that minimizes the overall death count among *all* vaccination strategies that satisfy the vaccination capacity constraint. Since the available choices include arbitrary functions of time, the optimal vaccination strategy can not be obtained by standard optimization, but needs to be characterized by solving an optimal control formulation, which we will shortly describe.

The first question however is if the vaccination strategy that minimizes the cumulative death count needs to be obtained by solving any optimization or optimal control problem at all. Wouldn't vaccinating all individuals in the high risk group first accomplish this goal always? To see why not consider the case that the high risk and high contact groups are respectively large and small in size, and the contact rates between the high contact group and other groups is much higher than the contact rates between other groups and within other groups. The initially infected individuals are all in the baseline group. Then, if the high risk group is vaccinated first, it will take some time to complete the vaccination strategy given the group's size. During this time the infection may be transmitted from the baseline group to the high risk group through the high contact group. Once the infection reaches the high risk group it kills many individuals in this group before they can be vaccinated. Instead if the high contact group would be vaccinated first, its vaccination would be completed in a short time because of its small size and the virus may not be able to reach the high risk group in this time. Once the high contact group is vaccinated in its entirety, the virus can not reach the high risk group as it lacks a path to it from the baseline group in which the infection had originated. Thus, the death count in the high risk group is low. The mortality risk is low in other groups, so very few individuals therein die even if they are infected. Thus, the overall death count is low. In other words, vaccinating the high contact group before the high risk group may in fact reduce the death count as compared to when this order is reversed. Thus, the optimal vaccination strategy is not *a priori* obvious. We therefore proceed to formulate the optimal control problem:

distributed; under this assumption the distribution of the number of individuals across the states and groups constitutes a continuous time Markov chain. Thus, under this assumption, the solution of the ODEs becomes a more and more accurate approximation of the temporal evolution of the actual state distribution as the population size increases to infinity. This constitutes an important computational strength as the system of ODEs can be readily solved regardless of the population size, while the time required to compute the probability of different population distributions across states and groups from the continuous time Markov chain representation increases exponentially in the population size. Thus, this latter computation is intractable for all practical purposes for populations of even moderate size.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i \in \{X,Y,Z\}} N \cdot D_i(T) \\
\text{subject to } & x \in \mathcal{S} \\
& x(0) = x_0, \\
& 0 \leq u_i(t) \leq 1 \quad \forall i \in \{X, Y, Z\}, t \in [0, T] \\
& \sum_{i \in \{X,Y,Z\}} u_i(t) S_i(t) \leq V_0 \quad \forall t \in [0, T]
\end{aligned} \tag{1}
$$

where x refers to a state trajectory. $x(0)$ is the state vector at the initial time, i.e. at $t = 0$, and its components are non-negative. $\mathcal{S}$ refers to the collection of state trajectories conforming to the system dynamics as specified by the ODEs. The last constraint in (1) represents the vaccination capacity constraint.

The optimal control problem can be solved using standard numerical toolboxes to yield the optimal vaccination strategy. We use the numerical toolboxes Yop [36] and CasADi [37] for this purpose.

Finally, we rule out the need to consider "mixed" policies which split vaccine capacity between multiple groups (until at least one group runs out of susceptibles to fully utilize the vaccine capacity). We prove that *there exists an optimal vaccination strategy that devotes full capacity to a single group while the susceptibles of each type exceed the vaccine capacity. Only after the number of susceptibles decreases below capacity will such a policy ever vaccinate two groups simultaneously* (see Supporting Information 3) [38]. This analytical result yields a significant reduction in the set of optimal strategies which must be considered. This reduction is also of practical importance: strategies which focus on one group at a time can be deployed more easily and align with the common practice of phased vaccine rollout to target groups.

4 Results

4.1 Model Validation

We first validate our model by demonstrating its applicability to COVID-19 infection and fatality data over a period of 9 months (1 April 2020 - 1 January, 2021, the period preceding the introduction of vaccines). We show that despite segmenting the population into only three groups (baseline, high risk, and high contact), when the parameters are chosen correctly, our model is able to accurately predict both the spread of infection and mortality of COVID-19 as observed over the course of the pandemic in all the US states and 139 countries. We also demonstrate that any decrease in the number of groups substantially undermines the accuracy of prediction.

We consider publicly available data on infection and death counts [39]. From census and survey data, we estimate the size of the three groups. Elderly populations are categorized as high risk while specific service industry and essential workers are high contact (see Supporting Information 4.1 for further detail) [40, 41, 42]. Disease parameters are obtained from WHO, CDC, and Johns Hopkins (see Supporting Info 1).

We select the contact rates within each group and across groups using regression, such that the mean squared normalized error (MSNE)[3] between the infection and death counts predicted by our model and the true values of the same is minimized. The

[3]The square normalized error for infection (death, respectively) on a given day is the square of the difference between the predicted and actual infection (death, respectively) counts divided by the actual infection (death, respectively) count on that day. The mean square normalized error for infection (death, respectively) is the average of the square normalized errors for infection (death, respectively)

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

minimum mean squared normalized error will be abbreviated as MMSNE. Different contact rates were chosen for periods of two months to account for changes in both government policy (e.g. start, relaxation, and end of lockdowns) and school openings which happen at low frequency.

We now compare the infection and death counts predicted by our model with the above choice of parameters and those that were actually recorded. Although the numbers that were actually observed were used to determine the contact rates in our model, the prediction using those need not match the reality, and the MMSNE may be high. This happens when the parameter space of the minimization is not large enough to capture the complexity of the pandemic progression, e.g. if the population needed to be segmented into a larger number of groups to capture the complexity. We show that this is not the case for different locations with different population demographics and government responses throughout US and worldwide. Specifically, in Figure 3 we consider 2 US states (California, Florida) and an LMIC (Bangladesh) to demonstrate that the predicted values closely match the values actually recorded, the average difference between the predicted and actually recorded values is less than 1% for both infections and death. From the US, we choose California and Florida because these two states have divergent age demographics and government responses. Florida has a significantly older population as well as among the most laissez-faire NPI restrictions while the opposite is true of California [43]. We choose Bangladesh as an example of an LMIC – in general LMICs have significantly younger population demographics and different government responses as compared to the US. The MMSNE for California, Florida, and Bangladesh are very small, 0.0005, 0.001, and 0.001, respectively. Thus, our reasonably simple model has distilled the essential heterogeneity innate to the system.

We now consider all the states of US and 139 different countries surveyed. Here, four countries for which death and infection counts are available are excluded due to apparent irregularities in the available data (see Supporting Information 4.1 for further detail). As seen in Figure 4, the MMSNE was very low, below 0.006, for all US states. Across the MMSNE for all countries, the median, mean, first and third quartile are 0.002, 0.005, 0.0006, 0.005, 0.05. While this error is slightly higher than that of the states, it is still quite low. The increased error is mostly due to low population sizes and abrupt, non-smooth changes in infection and death counts (further detail in Supporting Information 4.1). The low fitting error across the board shows that our model captures reality of geographical locations all over the world.

We now examine if our model can be simplified further while retaining the quality of the fit between the predicted values and reality. We again study California, Florida, and Bangladesh. We first consider the model that assumes a homogeneous populace – that is, there is only one group. There is therefore only one contact rate, that within the group. The mortality risk for the homogeneous populace is chosen as average over all age groups. We determine the contact rate in different time periods using the same regression technique. Figure 3 shows that the predicted counts do not closely match the counts observed in reality; the average error in predicted infections was 1.86% and the average error in predicted deaths was 82.91%. Thus, a homogeneous model can at best match only one count, and greater degrees of freedom are necessary to match both counts. We now examine if increasing the number of groups to two provides the

over all days. The mean square normalized error is the average of the mean square normalized error for infection and death. The division of the square error by the magnitude of the actual infection (death, respectively) ensures that discrepancies in magnitudes of infection and death counts do not imply that the errors of the quantity with significantly larger magnitude dominates that of the other. Error can be quantified in different forms, such as mean square error, mean absolute value of error etc. We chose the mean squared error for analytical convenience as the squared error is a smooth function, specifically it is smoother than the absolute value particularly when the square and absolute values are near zero. Thus, choosing contact rates that minimize the mean square normalized error is computationally simpler than choosing those that minimize the mean absolute value of the normalized error.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

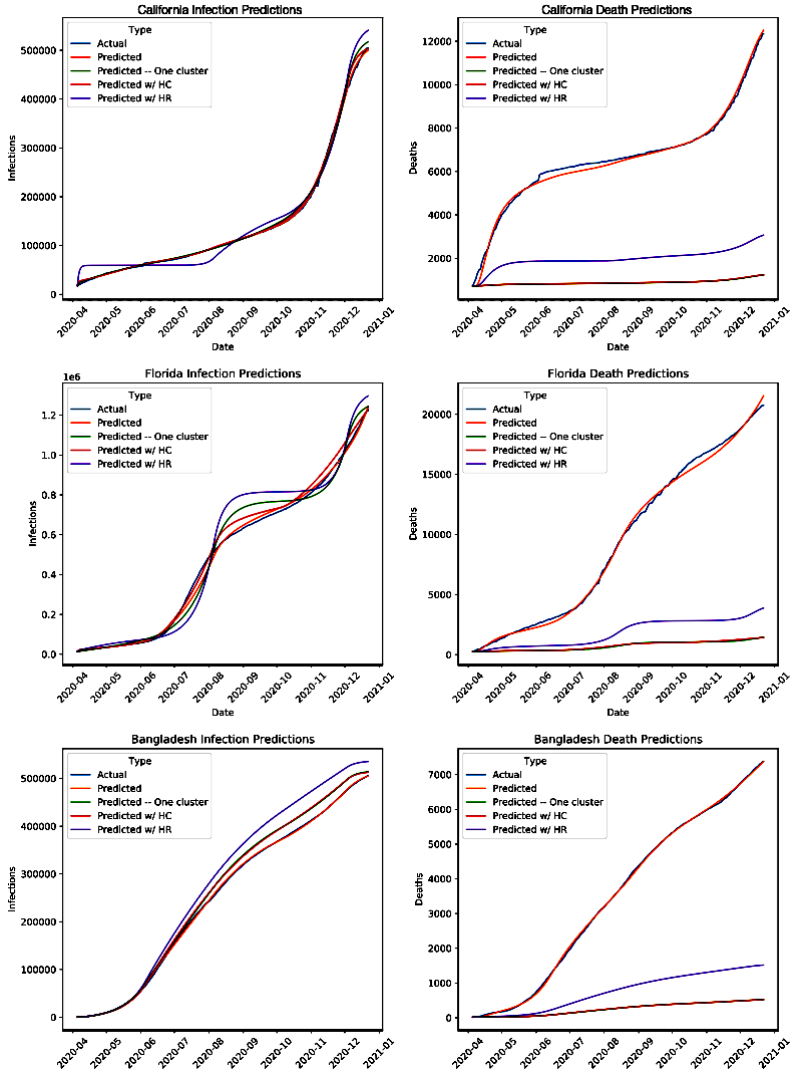


Fig 3. Here we show the accuracy of our model ('Predicted') when applied to the real infection and death counts of (a) California, (b) Florida, and (c) Bangladesh. The blue curves show the ground truth data [39]; orange and green show the predictions from our contact model and from a fully homogeneous model, respectively. Finally the red and purple show the predictions when only one type of heterogeneity is accounted for, contact rate and risk, respectively.

necessary degrees of freedom. We consider two groups with different contact rates but uniform mortality risk throughout the populace (which is average over the entire populace). The mismatch between predicted and recorded values remains similar, the average differences for the infection and death counts are respectively 1.93% and 82.85%. We now consider two groups with uniform contact rates throughout but one has higher mortality risk. Now, average error for the infection count increased to 6.95% while average death error decreased to 66.46%. Thus, all three groups are necessary for error values to be low, as three groups allows for heterogeneity in both mortality risks and contact rates. The heterogeneity innate to the evolution of the pandemic over a diverse population can not in general be captured by models simpler than ours. Having validated the model as above, we henceforth utilize the values of the parameters we

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

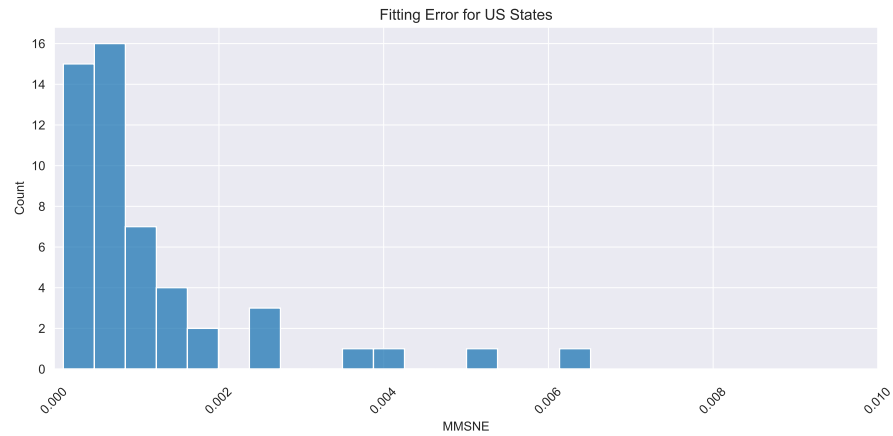


Fig 4. This figure shows the MMSNE histogram of the predictions from our fitted model when compared to actual case and death counts ([39]) for all US States.

extracted for three groups for all subsequent numerical computations using the model.

4.2 Structure of Optimal Vaccination Policies

Unless otherwise specified, we assume that our objective is to minimize the total death count. Recall that the challenge in determining optimal policies is that the space of potential vaccination strategies is very large – it can include straightforward strategies which vaccinate groups sequentially, complex time-dynamic policies which change focus day-to-day, and policies which split vaccine capacity among multiple groups at once. But our theoretical results summarized below in Theorem 4.1 rule out the need to consider the last class of policies. Using extensive numerical computations we show that the optimal policies lie in the first class in most of the cases that arise in practice, and the suboptimality of the strategies in the first class is limited even when the optimal strategy lies elsewhere.

For a policy $u(t)$, we will use $U(t)$ to denote the vaccine allocation in absolute quantity rather than fraction i.e.

Theorem 4.1. *There exists an optimal vaccination strategy u such that, for each t there exist $a(t), b(t), c(t) \in \{X, Y, Z\}$ all distinct and $U_{a(t)}(t) = min\left(V_0, S_{a(t)}(t)\right)$, $U_{b(t)}(t) = min\left(S_{b(t)}(t), V_0 - U_{a(t)}(t)\right)$, and $U_{c(t)}(t) = min\left(S_{c(t)}(t), V_0 - U_{a(t)}(t) - U_{b(t)}(t)\right)$.*

First, we clarify and interpret this result.

Note that for each vaccination strategy there exists a first time t_0 at which $S_i(t) < V_0$ for some i (until t_0 all groups have enough susceptible individuals to be vaccinated at full capacity if they are accorded that capacity), and a first time t_f at which $\sum_i S_i(t_f) \leq V_0$ (i.e. the total number of susceptible individuals is below vaccine capacity). These times are different for different strategies. Theorem 4.1 essentially says that there exists an optimal strategy u which follows the following structures before its t_0, between its t_0 and its t_f, and after its t_f.

1. For each $t < t_0$, there exists some $a(t) \in \{X, Y, Z\}$ such that $U_{a(t)}(t) = V_0, U_j(t) = 0$, for $j \neq a(t)$. That is, at any given time $a(t)$ is the highest priority group and it is accorded the full vaccine capacity and the rest of the groups get 0 vaccine capacity.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

2. For intermediate t ($t_0 \leq t \leq t_f$), there exists $a(t), b(t), c(t)$, such that either $U_{a(t)}(t) = V_0, U_j(t) = 0$, for $j \neq a(t)$, or $U_{a(t)}(t) = S_{a(t)}(t), U_{b(t)}(t) = V_0 - S_{a(t)}(t), U_{c(t)} = 0$, or $U_{a(t)}(t) = S_{a(t)}(t), U_{b(t)}(t) = S_{b(t)}(t), U_{c(t)} = V_0 - S_{a(t)} - S_{b(t)}$. That is, at a time t, $a(t), b(t), c(t)$ constitute groups in decreasing order of priority. The highest priority group $a(t)$ is given as much vaccine as possible, limited only by its availability of susceptibles. The residual capacity is passed to the next highest priority group, limited only by its availability of susceptibles, and this process continues.

3. For each $t > t_f$, $U_i(t) = S_i(t) \quad \forall i \in \{X, Y, Z\}$. That is, in the final phase, the total number of susceptibles of all groups together have fallen below the vaccine capacity, and therefore all groups are allocated the maximum possible vaccine capacity they can utilize, that is, vaccine capacity equaling their susceptibles. Note that this particular property holds for all optimal policies for times after their t_f (ie, even when the optimal policy is not unique).

Remark 4.1.1. *The salient point of this theorem is that at any given time the optimal policy described above accords different priorities to different groups; as long as the highest priority group has enough susceptibles to utilize the full vaccine capacity, it allocates the entire vaccine capacity to that group and the rest are not allocated any capacity. The policy starts splitting the capacity among different groups only when the highest priority group at the given time does not have enough susceptibles to utilize the capacity. Even then, the highest priority group is given the maximum capacity it can utilize, the residual is given to the next highest priority group limited only by its availability of susceptibles, and so on.*

Remark 4.1.2. *It is useful here to elaborate the extent of the above theorem. First, the theorem shows that such an optimal policy* exists*, but this optimal need not be unique. However, because such an optimal exists, the planner cannot achieve any further reduction in mortality by implementing more complex policies which instantaneously split vaccination capacity between two or more groups even when each group has enough susceptibles to utilize the entire vaccination capacity. Furthermore, while the specified optimal policy vaccinates one priority group at each instant (while it has enough susceptibles), our analytical result does not specify how frequently this priority group may change. We numerically study the number of switches in the priority accorded to the groups in the remainder of this section and find a maximum of 3 switches.*

To numerically study optimal vaccination policies, we vary disease parameters, initial population seroprevalence, and vaccine efficacy parameters over representative ranges based on best available estimates [44, 39, 45, 46]. Detailed derivations of all parameters and relevant sources are included in Supporting Information 1 and 4.2. We obtain the demographic data from census as in the previous section. The default assumption is that initial infections are seeded only in the baseline group; we explicitly specify when we deviate from the default assumptions. We used the contact rates obtained from the real evolution of the pandemic as in the previous section. We also use additional contact rates from ranges of contact rates within and across age groups in 139 countries obtained from surveys and the sizes of the different groups [47]. Overall, the large range of contact rates we consider capture varying degrees of implementation and compliance with non-pharmaceutical interventions (NPIs) such as social distancing and lockdowns in different US states and the 139 countries. Over all these parameter ranges, our model was instantiated and run on a fine grid of approximately $911,250$ settings (see Supporting Information 4.2). We also demonstrate that our model is robust to parameter estimation errors (see Supporting Information 4.3).

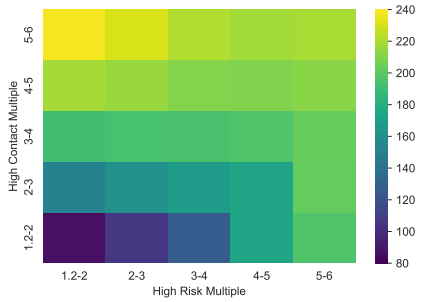
medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

In 93.4% of the cases we considered, the optimal solution was either to (1) vaccinate the high contact group fully, next vaccinate the high risk group fully (2) vaccinate the high risk group fully, next vaccinate the high contact group fully. In both these cases, the baseline group is vaccinated at the end. We refer to the first as *high contact prioritization* and the second as *high risk prioritization*. In the rest of the cases there is an extra step where the optimal policy switches from high contact vaccination to high risk before returning to the high contact. In these few cases, the better performer of the above two simple policies had only 2.2% more deaths than the optimal policy. Even including these cases, across all numerical experiments, the priority group for vaccination changed at most three times over the course of the optimal policy, and the number of switches was three in only 6.4%. Coupled with our theoretical result, this simplifies our potential policy space down to two easily deployable policies even without any assumed restrictions on policy complexity.

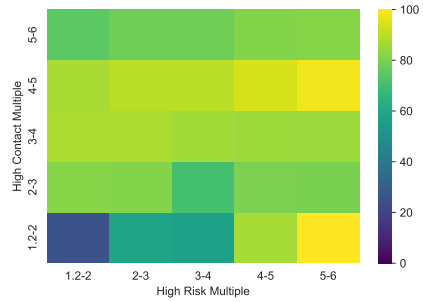
Seeing that the aforementioned simple policies perform well, one may question whether all simple policies, even static ones, perform similarly well in reducing mortality. To answer this question, we compare the death counts of the overall optimal policy with two more restricted policies: (1) the optimal policy among those whose vaccine capacity allocations to the groups are constant over the entire time horizon (*static policy*) and (2) the optimal policy among those whose vaccine capacity allocations to the groups are constant in periods of one month (*monthly policy*). The first policy is completely static while the vaccination capacities allocated by the second can vary with time but over large time scale. In contrast the overall optimal policy has been obtained by optimizing among policies that can vary their vaccination capacity allocation to groups as frequently as deemed necessary. We find that even at their best, the less dynamic classes of policies under consideration perform significantly worse than the general optimal. In particular, Figures 5a and 5b, show the additional mortality incurred by each of these restricted policies when compared to the optimal policy mortality as a percentage of the latter. The axes are measures of the two types of heterogeneity: the ratio of risk parameters of the high risk and baseline groups (x-axis) and the ratio of daily average contact of the high contact and baseline groups (y-axis). We see that as heterogeneity increases, the difference in mortality between the optimal and the two static variations is magnified. Furthermore, the suboptimality of the more static policies is especially sensitive to the high contact group. Intuitively, a very active high contact group significantly accelerates the spread of the epidemic, thus necessitating a more dynamic vaccine allocation strategy. Over all instances, the average increase in death counts under the two restricted policies as compared to the overall optimal vaccination strategy were 206.3% and 72.16%, respectively. In the extreme, the mortality of the static policy is over thrice that of the optimal. We note that the monthly policy has both lower death count and is more dynamic than the static policy. The most extreme difference between the two restricted policies occurs when the high contact multiple is high i.e. when the high contact group is very active. In this regime the monthly policy far outperforms the static policy (while both remain significantly less effective than the true optimal). Intuitively, a very active high contact group significantly accelerates the spread of the epidemic, thus necessitating a more dynamic vaccine allocation strategy.

Motivated by the above observations, we investigate the importance of policy dynamism for mortality reduction (Figure 5c). We consider one specific instance for this purpose: the US baseline setting in which contact rates and population demographics are averaged over the data reported in different states [41, 48, 47]. Figure 5c depicts how mortality substantially decreases as we increases the number of decision intervals, expanding our optimization space to allow for more dynamic policies. The dynamism of our optimal policies is integral to their performance in reducing mortality. If such dynamic policies are ignored, we do not have a proper foundation for reasoning about

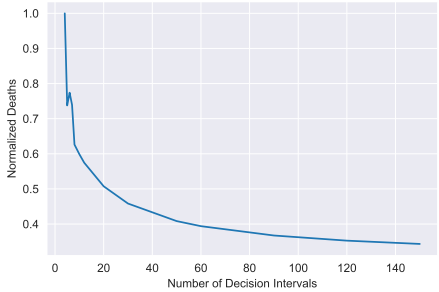
medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



(a) The increase in mortality caused by static vaccination policies over the optimal vaccination policy



(b) The increase in mortality caused by policies that change allocations only monthly over the optimal vaccination policy



(c) The benefit of more dynamic vaccination policies.

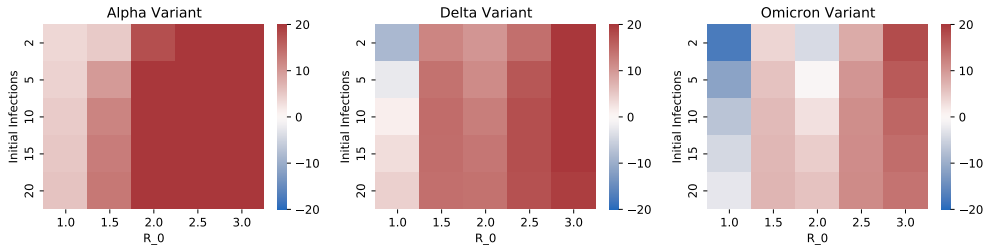
Fig 5. Figure (a) shows the increase in the death counts of the static vaccination policy over that of the optimal policy as a percentage of the latter. The x-axis is the ratio of death rates of the high risk and baseline groups (this is also the ratio of the symptomatic rates and hospitalization rates of the two groups). The y-axis is the ratio of daily average contacts of the high contact and baseline groups. As these ratios increase, our contact network becomes more heterogeneous and the suboptimality of the static policy increases. Figure (b) shows the same for the monthly policy, i.e. when the vaccine rate allocation to groups is allowed to change only on a monthly basis. Figure (c) shows the decrease in the number of deaths in the US baseline instance as one increases the number of decision intervals. The allocation of the vaccination capacity to groups remain constant over each interval. Number of days in each interval equals 365 divided by the number of intervals. Thus, more decision intervals allow for more dynamic policies. The death count is normalized by the maximum number of deaths across all data points.

the efficacy of vaccination policies because such restrictions preclude identifying the minimum possible death count.

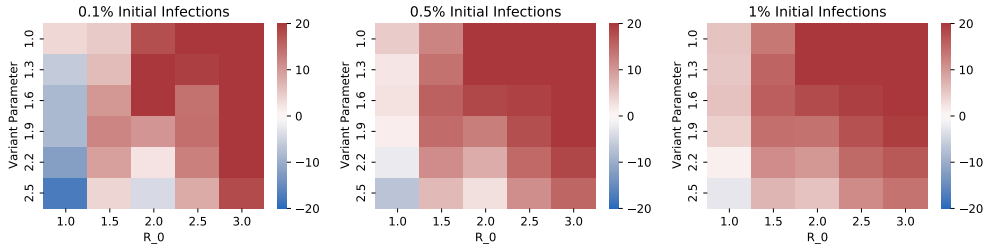
A strength of our framework is its computational tractability. The mean computation time needed to determine the optimal vaccination strategy over all aforementioned parameter settings was just 10.86 seconds with a standard deviation of 39.88 seconds. Excluding outliers[4], the mean computation time was just 8.04 seconds with a standard deviation of 14.45 seconds and a maximum of 142.42 seconds (a histogram of runtimes can be found in the Supporting Information 5). These computation times were obtained using modest computational resources, namely an

[4]In less than 0.6% of instances, the computation time exceeded the maximum allotted time of 500 seconds

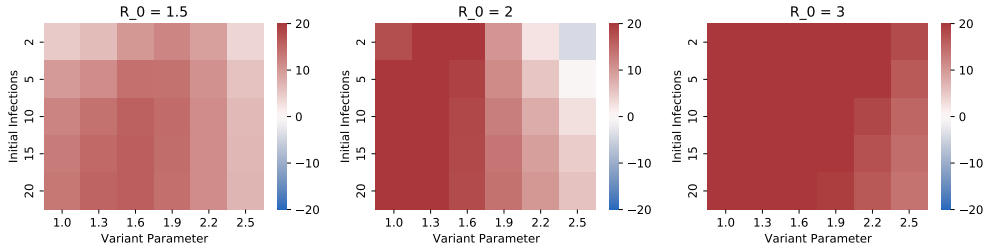
medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



(a) The effect of different COVID-19 variants on optimal vaccination policy



(b) The effect of initial seroprevalence on optimal vaccination policy



(c) The effect of R_0 on optimal vaccination policy

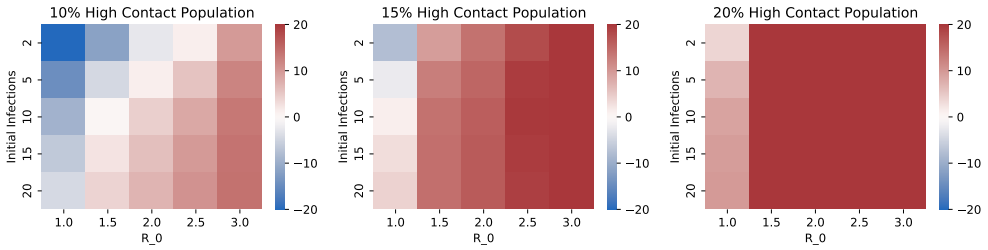
Fig 6. Each heatmap displays the difference in resultant mortality between high contact prioritization vaccination and high risk prioritization as a percentage of the mortality of high risk prioritization policy. (a), (b), (c) respectively demonstrate the impact of 1) variants with different transmissibility and mortality characteristics 2) different values of initial seroprevalence 3) different values of R_0.

Intel i7 2.7 GHz machine with 16 GB RAM, which could also be readily available in local public health bodies and resource-constrained settings such as LMICs.

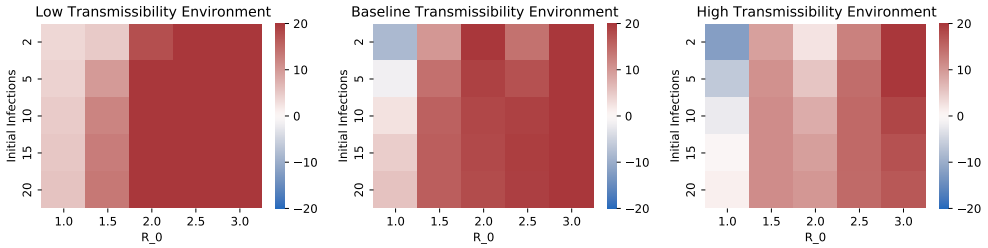
4.3 Comparison of mortalities under high contact and high risk prioritization across different parameter ranges

It is noteworthy that most public health authorities implemented a strategy closest to high risk prioritization with the caveat of early vaccination for health care workers. High contact prioritization, on the other hand, was not often recommended, especially with the broader definition of high contact which we use. Yet, in the previous section we found that the high contact prioritization was optimal in several realistic settings. In this section we compare the mortalities under high contact and high risk prioritization policies by varying different combinations of parameters. We find that the high contact prioritization has substantially lower mortality than the high risk prioritization in 42% of instances; thus the prevalent norm has been suboptimal in many cases.

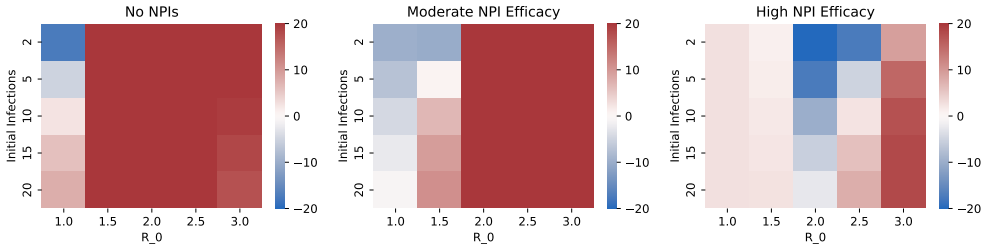
medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



(a) The effect of demography on optimal vaccination policy



(b) The effect of transmissibility on optimal vaccination policy



(c) The effect of NPI Efficacy on optimal vaccination policy

Fig 7. Each heatmap displays the difference in resultant mortality between high contact prioritization vaccination and high risk prioritization as a percentage of the mortality of high risk prioritization policy. (a), (b), (c) respectively demonstrate the impact of 1) different population distributions, i.e. different values of the percentage of the total population that is in the high contact cluster 2) transmissibility of the virus 3) NPIs (e.g. social distancing, masking, lockdowns).

To compare the two strategies, in each case, we vary two parameters as shown in the two axes of Figures 6 and 7, for given choice of one variable and include the relative decrease of mortality of one policy over another for all values of other parameters as considered in Section 4.2[5]. The default assumption is that the variant is Alpha; we explicitly specify when we deviate from the default assumptions. After all other parameters are set, R_0 is computed via the next generation matrix method (see [49]) and contact matrices are normalized to achieve the desired R_0, whenever we choose values of R_0. Refer to Supporting Information 4.2 for further details.

Fig. 6 shows the effect of COVID variants, level of initial infection and R_0 on the mortalities under the two policies. We considered the three major variants thus far (Alpha, Delta, and Omicron) by choosing the transmissibility (probability of

[5]We evaluate the death count of any given policy by utilizing the system of ODEs in Figure 2 with vaccination rates $\{u_i(\cdot)\}$ that correspond to the policy.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

transmission upon contact with an infective individual) and mortality characteristics according to recent clinical data for each strain [50]. In particular, relative to the Alpha variant, Delta was nearly twice as contagious and with double the hospitalization risk while Omicron was more contagious still but with lower mortality rates [51]. The variant parameter in Figures 6b and 6c is applied as a multiplicative factor on transmissibility – this range allows us to consider potential future variants of concern. Figure 6 shows that, among the three variants, Omicron has the highest relative benefit of high contact prioritization and Alpha the least. This happens because Omicron and Alpha respectively have the highest and least transmissibilities. And, vaccinating high contact individuals reduces the overall transmission significantly which is particularly effective in decreasing the overall mortality when the virus has higher transmissibility. We also note that the relative benefit of high contact prioritization decreases with increase in the level of initial infection. This happens because if the initial infection is high the virus spreads in the high risk group during the initial time in which the high contact group is being vaccinated, which in turn substantially increases the mortality of the high risk group because of the innate high mortality rates of the infected in this group. The effect of R_0, however, was non-monotonic with high risk prioritization attaining lower mortality than the high contact prioritization at extreme values of R_0. We will explain the non-monotonic behavior in the next paragraph.

Fig. 7 shows the impact of the high contact population size, viral transmissibility, and NPI (non-pharmaceutical intervention, i.e. social distancing, masking, lockdowns) efficacy on the two policies under consideration. Note that while the variants considered in the previous paragraph have different transmissibilities and mortality risks, here we study the impact of both these factors in isolation. We choose multiplicative factors of 1.0, 1.5, and 2.0 respectively on the baseline transmissibility to obtain the low and high transmissibility environments. NPI efficacy is modeled as a multiplicative factor on contact rates to capture different extents of social distancing, masking, and lockdowns. The low, moderate and high values of NPI efficacy respectively correspond to multiplicative factors of 1.0, 0.7, and 0.4, that is, no reduction in contacts, 30% reduction in contacts, and 60% reduction in contacts. Higher viral transmissibility increases the relative benefit of high contact prioritization due to larger reductions in potential infections. However, the relative benefit of the early vaccination of the high contact group is decreasing in the group's size. This happens because larger groups need more time to be vaccinated. Thus, according vaccine priority to large sized high contact groups significantly delay vaccination to high risk groups; the disease may spread in the high risk group during this delay leading to high death counts owing to higher mortality rates therein. Note that R_0 is increasing in both transmissibility and size of high contact groups. Thus, an increase in R_0 can shift the relative benefits of high contact prioritization policy in either direction.

Finally, the range in Figures 6 and 7 which exceed 20% difference or fall below -20% difference between the two policies includes values considerably higher than 20% or lower than -20%. Figure 8 shows the histogram of percent differences in mortality between the two strategies. Specifically, for certain parameter values in the above range, the percentage difference was as high as 71.71% and as low as -81.85%. Overall, the numerical computations reveal that, while the common high risk priority policy is optimal in the majority, there is a considerable range of realistic parameters for which high contact prioritization more effectively lowers mortality.

4.4 Case Studies

We now compare the death counts of the high contact and high risk prioritization vaccination strategies and a benchmark vaccination strategy for three important case studies which arise in practice. These cases: (1) LMICs, (2) prisons, (3) long-term care

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

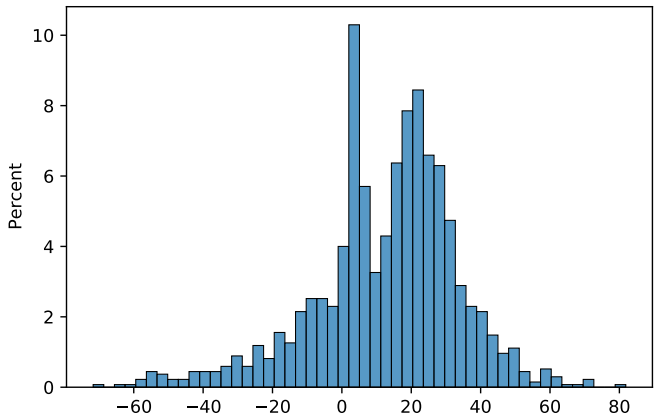


Fig 8. Histogram of the percent difference in mortality between high contact prioritization and high risk prioritization vaccination strategies over all instances.

facilities, have different characteristics, suffer from high vulnerability to the disease, and are hotbeds for the generation of new variants [52, 53, 54]. In particular, LMICs have a smaller elderly population, but, due to the greater prevalence of inter-generational households, high risk individuals are less able to isolate. They also have generally lower vaccine capacity [3]. Prisons also have a smaller elderly population, but have high levels of transmissibility due to unhealthy living conditions and presence of a large number of individuals in a relatively small confined space. Long-term care facilities have a large elderly population alongside a small core group of employees that have high contact rates with those at high risk. These differences were modeled by setting appropriate vaccination capacity constraints, contact rates between groups and sizes of the different clusters according to vaccine records, census data, government statistics, and contact surveys (for further detail see Supporting Information 1, 4.2) [3, 39, 55, 56] . We also consider the US baseline as described in Section 4.2. The benchmark vaccination strategy, referred to as Uniform, distributes vaccines among groups proportionally to each group's size. Uniform is similar to vaccinating randomly chosen individuals subject to vaccination capacity constraints.

Fig. 9(a-b) shows the differing efficacy of vaccine policies between the US baseline model and LMIC. Unlike the US baseline, the high contact prioritization strategy minimized mortality in the LMIC setting among the policies considered. This is because high risk individuals are in more frequent contact with the high contact group in LMICs than in the US. Thus, the high contact group constitutes a greater vector of spread to the high risk group in LMICs than in the US, and vaccinating the former on a priority basis helps protect the latter. Fig. 9(a) and (b) also show how vaccine scarcity enhances the impact of right vaccine prioritization. Note that $0.5\%, 0.2\%$ of overall populace can be vaccinated every day in US baseline and LMIC respectively. We see that the difference in infections, hospitalizations, and YLLs between the best performing strategy and the alternatives is much greater in the LMIC than in the US Baseline. Even when considering deaths, the difference between uniform and high contact prioritization is much greater in the LMIC.

Next, due to the large high risk population, long-term care facilities have relatively high mortality rates [57]. Even with high individual mortality, high contact prioritization more effectively minimizes mortality (Figure 9(d)). Intuitively, because the high risk group is large in this setting, vaccinating the group fully would be a

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

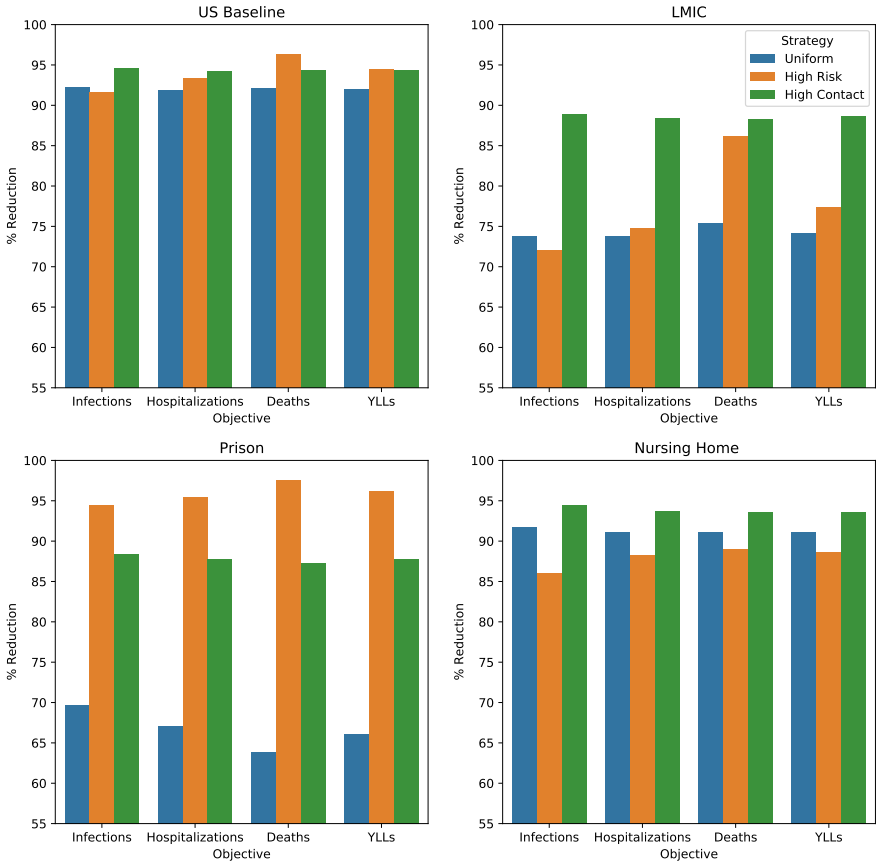


Fig 9. Each figure compares the infections, hospitalizations, deaths, and years of life lost (YLLs) under uniform (blue), high risk (orange), and high contact (green) prioritization vaccination policies. We depict the percentage reduction of these attributes under each policy relative to the no vaccination scenario. (a) (top left) is instantiated on US demography, NPI efficacy, and topology. (b) (top right) is instantiated on an LMIC model with lower elderly population, greater high contact population, and higher transmissibility. (c) (bottom left) is instantiated on a prison model comprised of guards (high contact), facility employees (baseline), and prisoners (baseline, high risk). This environment corresponds to low elderly population and high transmissibility. (d) (bottom right) is instantiated on a nursing home model with a large elderly population (high risk), medical staff (high contact), and administrative staff (baseline).

lengthy process during which the infection can spread amongst the group. Thus, cutting off the high contact group as a vector of the spread more effectively reduces mortality. The opposite is true in the prison setting because the high contact group is large and viral transmissibility is high due to living conditions [58]. Both of these ensure widespread infections. The best recourse is to vaccinate the small size high risk group first, this can be accomplished in a short time (Figure 9(c)).

In all settings, the better of the high risk and high contact prioritizations, reduces mortality more than uniform vaccination. But the worse of the two prioritizations is sometimes worse than uniform vaccination. This is the case in the US baseline and

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

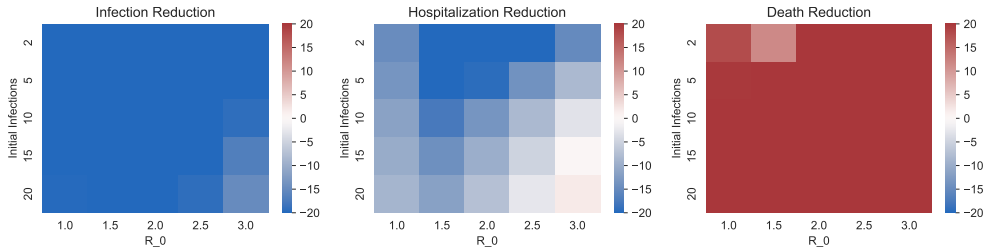


Fig 10. Each heatmap displays the difference in resultant objective function between high contact prioritization and high risk prioritization vaccination strategies as a percentage of the objective value of high risk prioritization policy. (a), (b), (c) respectively demonstrate the policy landscape when the objective is to minimize time average of symptomatic counts, time average of hospitalization counts, and overall death counts.

nursing home as seen in Figure 9(a) and (d). Simply using one of the two prioritizations in all cases, without verifying which one has lower mortality count, incurs higher mortality even compared to uniform (i.e. random) vaccination in some cases. Thus, one must tailor the prioritization to the specific context.

4.5 Different Objective Functions

We now consider optimal vaccination strategies that minimize the time averages of symptomatic and hospitalization counts as the public health objectives. The optimal control formulations that minimize these objectives can be found in Supporting Information 2.3. Again we found that in most cases (90.91%) the optimal vaccination strategies was either the high contact or the high risk prioritization strategies. In the remaining cases, the average suboptimality of the better of the two was 2.26% and 5.99% for the two objectives respectively.

Figure 10 shows how the change in objective function affects the relative performances of the high contact and high risk prioritization strategies. The high contact prioritization strategy outperforms the high risk prioritization strategy for greater number of instances when minimizing symptomatics rather than deaths is the public health objective because the increased rates of hospitalization and death of the high risk individuals do not affect the symptomatic count. Similarly, minimizing hospitalizations reveals an intermediate result where high risk prioritization is more important than when minimizing symptomatics but less important than when minimizing deaths. This aligns with our intuition that the high risk group drives the largest share of the death count while the high contact group is predominantly responsible for spreading the disease.

4.6 Two Dose Vaccines

We formulate the problem of determining the optimal two dose vaccination strategy in Supporting Information 2.2. When considering two dose vaccination, our decision space becomes richer as the second dose can be delayed by any given amount to accommodate the first doses for all or some groups. The decision problem therefore consists of the amount of delay for each group and the ordering of both doses. Different countries have implemented different delays and orderings for the two doses [59, 60]. There is an additional constraint now: the first dose must precede the second dose for every individual. In addition to the contributory factors considered so far, the optimal decision also depends on the degree of immunity provided by the first dose.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

We ran our two dose model on the same parameter combination detailed in the first subsection. The range of single dose efficacy was 40% to 80% in intervals of 10% in line with clinical data from multiple available vaccines [61]. In the 4,556,250 instances we considered, the three most common optimal policies were those that administer (1) both doses to high risk group, then both doses to high contact group, then both doses to the baseline group (60.91% of instances); (2) both doses to high contact group, then both doses to high risk group, then both doses to the baseline group (7.27% of instances); and (3) first dose to high contact group, then first dose to high risk group, then second dose to high contact group, then second dose to high risk group, then both doses to baseline group (30% of instances). Together these cover 98.18% cases. Clearly these three candidates are easy to deploy[6]. Overall, 68.18% (policies (1) and (2)) did not delay second doses to accommodate first doses to multiple groups; in these *"non-delay"* cases the optimal strategy follows either the high contact or high risk prioritization policies and provides both doses to these entire groups in one go. In the remaining cases, the optimal policy opted to delay second doses for one or more groups. Thus, there are regimes where delaying second doses can provide significant additional mortality reduction.

Here it is useful to characterize when delaying second doses is beneficial in relation with our single dose studies. We will refer to instances where high risk prioritization was optimal in the single dose model as high risk prioritization regime and the same for high contact prioritization. We find that delaying the second dose is never beneficial in high risk prioritization regime. In this regime the optimal strategy focuses on those most vulnerable to the pandemic, therefore, interim efficacy after just one dose is not sufficient for this strategy. In high contact prioritization regimes, for sufficiently high interim vaccine efficacy ($> 70\%$), we see that delaying the second dose of the high risk group and accommodating first doses to others within this delay can lead to significant mortality reduction (31.62% on average) over the best alternative non-delay policy. Put simply, if the protection of one dose is sufficiently high, then delaying second doses and accommodating first doses to others can help reduce mortality. It is also effective in reducing the number of infections and, thus, may effectively reduce the burden on health systems or economic impact of the pandemic in some regions.

5 Discussion

We now summarize how our contributions have advanced the state of the art on research on COVID-19 vaccination strategy. We have provided a framework for obtaining vaccination strategies that attain several desired public health objectives (e.g. minimizing symptomatic events, hospitalization, death counts) subject to vaccination capacity constraints. The framework incorporates both single dose and multidose vaccines, arbitrary combination of parameters and attributes that occur in practice such as reinfection, breakthrough infections. Our framework relies on the following methodological innovations: 1) capturing the innate heterogeneity of COVID-19 in contact and risk profiles through classification of the populace into three groups 2) formulating the optimization of a variety of public health objectives subject to vaccination capacity constraints as an optimal control problem with the state space evolution modeled as ODEs. We optimize vaccination strategy among literally all possible strategies including highly dynamic ones and without any constraint on how

[6]Two other easily deployable policies constitute the remaining 2%: administer 1) both doses to high risk group, first dose to high contact group, first dose to baseline group, first dose to high contact group, then second dose to baseline group; 2) first dose to high contact group, both doses to high risk groups, second dose to high contact group, then both doses to baseline group. Each constitutes about 1% of instances.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

often the capacity allocations can be changed among groups. Such an optimal strategy, even when difficult to deploy, provides a valuable benchmark for comparisons of any proposed policy. Our work provides this benchmark, and goes beyond by providing a few (two to three depending on the number of doses of vaccines) easy-to-deploy strategies, at least one of which is optimal or near-optimal in a wide array of parameter combination. Our framework typically identifies optimal vaccination strategies within seconds for single dose vaccine using modest computation resources which are usually available in public health centres including in the LMICs. For multi-dose vaccines the computation time merely increases to minutes. Owing to this computational tractability we could present results for large landscapes of 911,250 instances involving variations of parameters in large realistic ranges (Section 4). Large landscapes allow for more reliable conclusions on how the optimal strategy and the optimal value of the public health objective changes with variations in parameter values; large landscapes also help identify near-optimal, easy-to-deploy vaccination strategies which we accomplish. In contrast, even for the simplest scenario of single dose vaccines, no breakthrough or reinfection, the previous works are largely computationally intensive. Most of the previous published work has been limited to evaluations and comparisons of a handful of vaccination strategies by simulations; a smaller class of previous work optimized among a restricted class of strategies which either do not change capacity allocation to groups at all (i.e. fully static strategies) or change once in a large predetermined decision intervals (i.e. limited dynamism). We show that allowing fully dynamic allocation of vaccine capacities among different population groups (i.e. not restricting the size of decision intervals) substantially enhances public health metrics; the performance in fact improves with increase in the allowed dynamism. Our vaccination strategy is therefore able to significantly outperform the static or limited-dynamism strategies.

Our framework does assume the knowledge of relevant disease parameters, namely the contact rates of different groups and disease parameters (transmissibility and duration of different disease states and rates of symptomatic infection, hospitalization, and death). Our robustness study (see Supporting Information 4.3) shows that the optimal policies are generally robust to imperfect knowledge of disease parameters and contact rates. The disease parameters are also consistent over time for a given strain of virus. The contact rates, however, may well change over the course of the pandemic even for a single strain. As detailed in Section 4.1, we allow contact rates to change every two months to account for infrequent shifts (i.e. lockdown changes and school openings); however, the controller is assumed to have knowledge of these contact rates over the entire course of the pandemic.

COVID-19 is not over yet and new variants are emerging in different parts of the world. It is unclear if the same vaccines will be effective for future variants. If not, then the vaccination process will need to restart from scratch. In addition, given how frequently pandemics have recurred in recorded human history [62], an entirely new pandemic is likely to emerge at some point in future, and it is imperative to be prepared for countering the same with a well-founded vaccination strategy. We describe how this paper can provide a vaccination strategy for a future variant or pandemic.

The stages of the disease we consider appear with some variation in different contagious diseases which spread through contact. The variations are likely to be in the transition rates between the stages and in risk profiles. The distinguishing aspect of our model is that the risks of hospitalization and death significantly increase with age and poor health. Our framework will provide the optimal, or near-optimal vaccination strategy, whenever these general principles hold. This is likely to be the case for future variants and possibly for one or more future pandemics. In such instances, the framework will proceed as follows:

1. Following the classification strategy we proposed for COVID-19 (Section 3), divide

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

the populace into the same three groups: 1) high risk (e.g. retirees) 2) high contact (those with large and dynamic contact sets owing to professional requirements) 3) baseline (rest).

2. For single dose vaccine, compare the value of the objective function of interest (death, hospitalization, symptomatic counts) for high contact and high risk priority vaccination strategies. Deploy the better of the two.

3. For two dose vaccines, compare as above the three candidate optimal vaccination strategies of Section 4.6 and deploy the best. The comparisons can be performed using the heat maps of Figures 6, 7, 10 or from the ODEs of Figure 2 and the ODEs associated with the generalized model (see Supporting Information 2).

For possible future pandemics, and current evolving pandemic response strategies, our framework's flexibility, computational efficiency, and usability present a new practicable pathway for systematically reasoning about optimal vaccination policies.

References

[1] *WHO Coronavirus (COVID-19) Dashboard*. URL: https://covid19.who.int/info/.

[2] Gang Lv et al. "Mortality rate and characteristics of deaths following COVID-19 vaccination". In: *Frontiers in Medicine* 8 (2021).

[3] Edouard Mathieu et al. "A global database of COVID-19 vaccinations". In: *Nature human behaviour* 5.7 (2021), pp. 947–953.

[4] "WHO SAGE Roadmap for Prioritizing Use of COVID-19 Vaccines". In: (2022).

[5] *Comparing the differences between covid-19 vaccines*. URL: https://www.mayoclinic.org/coronavirus-covid-19/vaccine/comparing-vaccines.

[6] Jack H Buckner, Gerardo Chowell, and Michael R Springborn. "Dynamic prioritization of COVID-19 vaccines when social distancing is limited for essential workers". In: *Proceedings of the National Academy of Sciences* 118.16 (2021).

[7] Amir Emami et al. "The role of comorbidities on mortality of COVID-19 in patients with diabetes". In: *Obesity Medicine* 25 (2021), p. 100352.

[8] Erika Asano et al. "Optimal control of vaccine distribution in a rabies metapopulation model". In: *Mathematical Biosciences and Engineering* 5.2 (2008), pp. 219–238.

[9] Naveen Sharma et al. "Modeling assumptions, optimal control strategies and mitigation through vaccination to Zika virus". In: *Chaos, Solitons Fractals* 150 (2021), p. 111137. ISSN: 0960-0779. DOI: https://doi.org/10.1016/j.chaos.2021.111137. URL: https://www.sciencedirect.com/science/article/pii/S0960077921004914.

[10] Rachael Miller Neilan and Suzanne Lenhart. "Optimal vaccine distribution in a spatiotemporal epidemic model with an application to rabies and raccoons". In: *Journal of mathematical analysis and applications* 378.2 (2011), pp. 603–619.

[11] James Aspnes, Kevin Chang, and Aleksandr Yampolskiy. "Inoculation Strategies for Victims of Viruses and the Sum-of-Squares Partition Problem". In: *J. Comput. Syst. Sci.* 72.6 (Sept. 2006), pp. 1077–1093. DOI: 10.1016/j.jcss.2006.02.003.

[12] Jean-Jacques Muyembe-Tamfum et al. "Ebola virus outbreaks in Africa: past and present". In: *Onderstepoort Journal of Veterinary Research* 79.2 (2012), pp. 06–13.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[13] Kate M Bubar et al. "Model-informed COVID-19 vaccine prioritization strategies by age and serostatus". In: *Science* 371.6532 (2021), pp. 916–921.

[14] Bjorn Goldenbogen et al. "Optimality in COVID-19 vaccination strategies determined by heterogeneity in human-human interaction networks". In: (2020).

[15] Shuli Zhou et al. "Optimizing spatial allocation of COVID-19 vaccine by agent-based spatiotemporal simulations". In: *GeoHealth* 5.6 (2021), e2021GH000427.

[16] Joseph Chadi Lemaitre et al. "Optimizing the spatio-temporal allocation of COVID-19 vaccines: Italy as a case study". In: *medRxiv* (2021).

[17] Prashanth Selvaraj et al. "Rural prioritization may increase the impact of COVID-19 vaccines in a representative COVAX AMC country setting due to ongoing internal migration: A modeling study". In: *PLOS Global Public Health* 2 (1 Jan. 2022), e0000053. DOI: 10.1371/JOURNAL.PGPH.0000053.

[18] Nuru Saadi et al. "Models of COVID-19 vaccine prioritisation: A systematic literature search and Narrative Review". In: *BMC Medicine* 19.1 (2021). DOI: 10.1186/s12916-021-02190-3.

[19] Ozgur M Araz, Alison Galvani, and Lauren A Meyers. "Geographic prioritization of distributing pandemic influenza vaccines". In: *Health Care Management Science* 15 (2012), pp. 175–187.

[20] Ashleigh R. Tuite et al. "Optimal pandemic influenza vaccine allocation strategies for the Canadian population". In: *PLoS ONE* 5 (5 2010). ISSN: 19326203. DOI: 10.1371/JOURNAL.PONE.0010520.

[21] Rajan Patel, Ira M Longini Jr, and M Elizabeth Halloran. "Finding optimal vaccination strategies for pandemic influenza using genetic algorithms". In: *Journal of theoretical biology* 234.2 (2005), pp. 201–212.

[22] Laura Matrajt, M. Elizabeth Halloran, and Ira M. Longini. "Optimal Vaccine Allocation for the Early Mitigation of Pandemic Influenza". In: *PLoS Computational Biology* 9 (3 2013). ISSN: 15537358. DOI: 10.1371/JOURNAL.PCBI.1002964.

[23] Benjamin Faucher et al. "Agent-based modelling of reactive vaccination of workplaces and schools against COVID-19". In: *Nature Communications 2022 13:1* 13 (1 Mar. 2022), pp. 1–11. ISSN: 2041-1723. DOI: 10.1038/s41467-022-29015-y. URL: https://www.nature.com/articles/s41467-022-29015-y.

[24] Qingfeng Li and Yajing Huang. "Optimizing global COVID-19 vaccine allocation: An agent-based computational model of 148 countries". In: *PLOS Computational Biology* 18 (9 Sept. 2022), e1010463. ISSN: 1553-7358. DOI: 10.1371/JOURNAL.PCBI.1010463. URL: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1010463.

[25] Leonardo Souto Ferreira et al. "Assessing the best time interval between doses in a two-dose vaccination regimen to reduce the number of deaths in an ongoing epidemic of SARS-CoV-2". In: *PLoS Computational Biology* 18 (3 Mar. 2022). ISSN: 15537358. DOI: 10.1371/JOURNAL.PCBI.1009978.

[26] L Matrajt et al. "Optimizing vaccine allocation for COVID-19 vaccines shows the potential role of single-dose vaccination". In: *nature.com* (). URL: https://www.nature.com/articles/s41467-021-23761-1.

[27] Laura Matrajt et al. "Vaccine optimization for COVID-19: Who to vaccinate first?" In: *Science Advances* 7.6 (2021), eabf1374.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[28] Suresh P. Sethi and Preston W. Staats. "Optimal Control of Some Simple Deterministic Epidemic Models". In: *The Journal of the Operational Research Society* 29.2 (1978), pp. 129–136. ISSN: 01605682, 14769360. URL: http://www.jstor.org/stable/3009792.

[29] Rachael Miller Neilan and Suzanne Lenhart. "An Introduction to Optimal Control with an Application in Disease Modeling." In: *Modeling paradigms and analysis of disease trasmission models*. 2010, pp. 67–81.

[30] Zhong-Hua Shen et al. "Mathematical modeling and optimal control of the COVID-19 dynamics". In: *Results in Physics* 31 (2021), p. 105028.

[31] Tingting Li and Youming Guo. "Modeling and optimal control of mutated COVID-19 (Delta strain) with imperfect vaccination". In: *Chaos, Solitons & Fractals* 156 (2022), p. 111825.

[32] Rex N Ali, Harvey Rubin, and Saswati Sarkar. "Countering the potential re-emergence of a deadly infectious disease—Information warfare, identifying strategic threats, launching countermeasures". In: *Plos one* 16.8 (2021), e0256014.

[33] Apr. 2024. URL: https://www.cdc.gov/vaccines/covid-19/clinical-considerations/interim-considerations-us.html.

[34] World Health Organization et al. *COVID-19 vaccination: supply and logistics guidance: interim guidance, 12 February 2021*. Tech. rep. World Health Organization, 2021.

[35] Thomas G. Kurtz. "Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes". In: *Journal of Applied Probability* 7.1 (1970), pp. 49–58. ISSN: 00219002. URL: http://www.jstor.org/stable/3212147.

[36] Viktor Leek. *An optimal control toolbox for MATLAB based on CasADi*. 2016.

[37] Joel A E Andersson et al. "CasADi – A software framework for nonlinear optimization and optimal control". In: *Mathematical Programming Computation* 11.1 (2019), pp. 1–36. DOI: 10.1007/s12532-018-0139-4.

[38] Raghu Arghal, Shirin Saeedi Bidokhti, and Saswati Sarkar. "Optimal Capacity-Constrained COVID-19 Vaccination for Heterogeneous Populations". In: *2022 IEEE 61st Conference on Decision and Control (CDC)*. 2022, pp. 5620–5626. DOI: 10.1109/CDC51059.2022.9992682.

[39] O. Wahltinez et al. "COVID-19 Open-Data: curating a fine-grained, global-scale data repository for SARS-CoV-2". In: (2020). Work in progress. URL: https://goo.gle/covid-19-open-data.

[40] *Report COVID-19: Essential Workers in the States*. URL: https://www.ncsl.org/labor-and-employment/covid-19-essential-workers-in-the-states.

[41] *US states with the most essential workers*. Dec. 2021. URL: https://unitedwaynca.org/blog/us-states-with-the-most-ssential-workers/.

[42] URL: https://bbs.portal.gov.bd/sites/default/files/files/bbs.portal.gov.bd/page/057b0f3b_a9e8_4fde_b3a6_6daec3853586/2021-12-02-10-01-a5b3adcd2ea20db89d4bae0c90bd86cf.pdf.

[43] Greg Allen. *Florida's governor lifts all covid-19 restrictions on businesses statewide*. Sept. 2020. URL: https://www.npr.org/sections/coronavirus-live-updates/2020/09/25/916969969/floridas-governor-lifts-all-covid-19-restrictions-on-businesses-statewide.

[44] *Covid-19 pandemic planning scenarios*. Mar. 2021. URL: https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[45] Laith J Abu-Raddad, Hiam Chemaitelly, and Adeel A Butt. "Effectiveness of the BNT162b2 Covid-19 Vaccine against the B. 1.1. 7 and B. 1.351 Variants". In: *New England Journal of Medicine* 385.2 (2021), pp. 187–189.

[46] Srinivas Nanduri et al. "Effectiveness of Pfizer-BioNTech and Moderna vaccines in preventing SARS-CoV-2 infection among nursing home residents before and during widespread circulation of the SARS-CoV-2 B. 1.617. 2 (Delta) variant—National Healthcare Safety Network, March 1–August 1, 2021". In: *Morbidity and Mortality Weekly Report* 70.34 (2021), p. 1163.

[47] Kiesha Prem, Alex R Cook, and Mark Jit. "Projecting social contact matrices in 152 countries using contact surveys and demographic data". In: *PLoS computational biology* 13.9 (2017), e1005697.

[48] Emma Rubin. *Elderly population in U.S. by State*. Feb. 2022. URL: `https://www.consumeraffairs.com/homeowners/elderly-population-by-state.html`.

[49] Hyun Mo Yang. "The basic reproduction number obtained from Jacobian and next generation matrices–A case study of dengue transmission modelling". In: *Biosystems* 126 (2014), pp. 52–75.

[50] Yusha Araf et al. "Omicron variant of SARS-CoV-2: genomics, transmissibility, and responses to current COVID-19 vaccines". In: *Journal of medical virology* 94.5 (2022), pp. 1825–1832.

[51] Kathy Katella. *Omicron, Delta, Alpha, and more: What to know about the coronavirusnbsp;variants*. Feb. 2023. URL: `https://www.yalemedicine.org/news/covid-19-variants-of-concern-omicron`.

[52] Cristina Mesa Vieira et al. "COVID-19: The forgotten priorities of the pandemic". In: *Maturitas* 136 (2020), pp. 38–41.

[53] Courtney H Van Houtven, Nathan A Boucher, and Walter D Dawson. "Impact of the COVID-19 outbreak on long-term care in the United States". In: *International Long-Term Care Policy Network* (2020).

[54] Laura Hawks, Steffie Woolhandler, and Danny McCormick. "COVID-19 in prisons and jails in the United States". In: *JAMA Internal Medicine* 180.8 (2020), pp. 1041–1042.

[55] *Federal Bureau of Prisons*. URL: `https://www.bop.gov/about/statistics/population_statistics.jsp`.

[56] *FASTSTATS - Residential Care Community*. Dec. 2022. URL: `https://www.cdc.gov/nchs/fastats/residential-care-communities.htm`.

[57] Andrew T Levin et al. "COVID-19 prevalence and mortality in longer-term care facilities". In: *European Journal of Epidemiology* (2022), pp. 1–8.

[58] J O'grady et al. *Tuberculosis in prisons: anatomy of global neglect*. 2011.

[59] Nisha Nambiar. *Experts suggest delaying 2nd vaccine dose to improve efficacy: Mumbai News - Times of India*. URL: `https://timesofindia.indiatimes.com/city/mumbai/experts-suggest-delaying-2nd-vax-dose-to-improve-efficacy/articleshow/81456760.cms`.

[60] Seyed M Moghadas et al. "Evaluation of COVID-19 vaccination strategies with a delayed second dose". In: *PLoS biology* 19.4 (2021), e3001211.

[61] Jamie Lopez Bernal et al. "Effectiveness of Covid-19 vaccines against the B. 1.617. 2 (Delta) variant". In: *New England Journal of Medicine* 385.7 (2021), pp. 585–594.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.10.23299100; this version posted July 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[62] Jocelyne Piret and Guy Boivin. "Pandemics Throughout History". In: *Frontiers in Microbiology* 11 (2021). DOI: 10.3389/fmicb.2020.631736. URL: https://www.frontiersin.org/articles/10.3389/fmicb.2020.631736.