Could Vaccine Dose Stretching Reduce COVID-19 Deaths?

Witold Więcek, Amrita Ahuja, Michael Kremer, Alexandre Simoes Gomes, Christopher M. Snyder, Alex Tabarrok, and Brandon Joel Tan

JUNE 2021
Could Vaccine Dose Stretching Reduce COVID-19 Deaths?
Witold Więcek, Amrita Ahuja, Michael Kremer, Alexandre Simoes Gomes, Christopher M. Snyder, Alex Tabarrok, and Brandon Joel Tan
June 2021

Abstract: We argue that alternative COVID-19 vaccine dosing regimens could potentially dramatically accelerate global COVID-19 vaccination and reduce mortality, and that the costs of testing these regimens are dwarfed by their potential benefits. We first use the high correlation between neutralizing antibody response and efficacy against disease (Khoury et. al. 2021) to show that half or even quarter doses of some vaccines generate immune responses associated with high vaccine efficacy. We then use an SEIR model to estimate that under these efficacy levels, doubling or quadrupling the rate of vaccination by using fractional doses would dramatically reduce infections and mortality. Since the correlation between immune response and efficacy may not be fully predictive of efficacy with fractional doses, we then use the SEIR model to show that fractional dosing would substantially reduce infections and mortality over a wide range of plausible efficacy levels. Further immunogenicity studies for a range of vaccine and dose combinations could deliver outcomes in weeks and could be conducted with a few hundred healthy volunteers. National regulatory authorities could also decide to test efficacy of fractional dosing in the context of vaccination campaigns based on existing immune response data, as some did for delayed second doses. If efficacy turned out to be high, the approach could be implemented broadly, while if it turned out to be low, downside risk could be limited by administering full doses to those who had received fractional doses. The SEIR model also suggests that delaying second vaccine doses will likely have substantial mortality benefits for multiple, but not all, vaccine-variant combinations, underscoring the importance of ongoing surveillance. Finally, we find that for countries choosing between approved but lower efficacy vaccines available immediately and waiting for mRNA vaccines, using immediately available vaccines typically reduces mortality.

Keywords: vaccine, pandemic, epidemiology, public health, supply

Acknowledgments: The authors are grateful to Nora Szech, Ezekiel Emanuel, Marc Lipsitch, Prashant Yadav, Patrick Smith, Frédéric Y. Bois, Virginia Schmit, Victor Dzau, Valmik Ahuja, Marcella Alsan, Garth Strohbehn and seminar participants at Harvard, M.I.T., the University of Chicago, and the NBER Entrepreneurship and Innovation Policy and the Economy conference for helpful comments and to Esha Chaudhuri for excellent research assistance.

Disclosures: This work was supported in part by the Wellspring Philanthropic Fund (Grant No: 15104) and Open Philanthropy. Więcek provides scientific consultancy for Certara, a drug development company and 1 Day Sooner, a COVID-19 human challenge trial advocacy group. Więcek reports no material conflicts of interest with regards to development of COVID-19 vaccines.

Contact Information: Więcek: Independent researcher <witold.wiecek@gmail.com>. Ahuja: Douglas B. Marshall Jr. Family Foundation <amrita.b.ahuja@gmail.com>. Kremer: Department of Economics, University of Chicago <kremermr@uchicago.edu>. Simoes Gomes: Department of Economics, University of Chicago <asimoes@uchicago.edu>. Snyder: Department of Economics, Dartmouth College <chris.snyder@dartmouth.edu>. Tabarrok: Economics Department, George Mason University <tabarrok@gmu.edu>. Tan: Economics Department, Harvard University <btan@g.harvard.edu>.
1. Introduction

Most of the world is facing a shortage of COVID-19 vaccines, and it is unclear how much production capacity can be added in time to materially affect the pandemic. In this paper, we investigate the public health impacts of dose stretching policies to vaccinate people more rapidly. In particular, we consider: (1) fractional doses, (2) longer delays between first and second doses, and (3) utilizing available vaccines rather than waiting for higher efficacy ones. We first discuss available evidence on the potential impact of such policies on protective efficacy for individuals receiving vaccines and then combine this evidence with an epidemiological model to assess overall impacts on public health. We find that vaccinating more people using dose stretching policies yields large reductions in mortality and infections if efficacy remains high, and potentially even if such policies entail considerable loss of efficacy.

For some dose stretching strategies, there is considerable evidence on efficacy from clinical trials and real world data that can inform epidemiological modeling. For example, we can readily examine the trade-offs that a country might face when choosing between a 70% effective vaccine available immediately, and a 95% effective vaccine available in two months. Longer delays between first and second doses—i.e., "first doses first" (FDF)—which allow more people to get at least one vaccine dose faster, have already been implemented by a number of countries. Clinical and observational data suggests high first-dose efficacy for certain vaccines against some variants of the virus, but lower efficacy for other vaccine-variant combinations.

There is no clinical evidence, to our knowledge, of the efficacy of alternative doses of COVID-19 vaccines. Therefore, to estimate the potential efficacy of fractional dosing, we combine existing data showing a high correlation between immune response and efficacy across vaccines (Khoury
et al. 2021) with early-stage clinical trial data on immune responses produced by different vaccine doses. While sample sizes are low, the results suggest that lower doses may be highly effective. For example, clinical trial data for mRNA-1273 (the vaccine developed by Moderna and manufactured by Moderna and GC Pharma) suggests that half doses produce an immune response associated with close to 95% efficacy, very similar to that of the standard dose. This evidence, while far from dispositive, suggests that half (or even quarter) doses of some vaccines could plausibly provide efficacy comparable to that of currently used doses.

We then use an SEIR model to assess the potential impact of dose stretching strategies when the immunization rate is constrained by available supply. The immunization rate in our base case scenario corresponds to the global vaccination rate as of early May 2021, which is roughly 0.25% of the population per day.\textsuperscript{1} At this baseline speed, a 95% effective vaccine averts 18–50% of infections compared to not vaccinating.\textsuperscript{2} Doubling the vaccination rate by moving to a half dose of a COVID-19 vaccine averts 41–70% of infections if there is no loss in efficacy. Even if half doses are only 70% effective, moving to a half dose and doubling the pace of vaccination would substantially reduce infections and deaths relative to the status quo policy.

If first-dose efficacy is 80%, as suggested by UK data on BNT162b2\textsuperscript{3} (the vaccine developed by BioNTech/Pfizer and manufactured by Pfizer and Fosun Biotech), then FDF reduces mortality by 29% and infections by 39% in a fast-growing epidemic scenario (\(R = 2\)) with vaccination speed of 0.75% of the population per day. The magnitude of benefits depends on epidemic and baseline

\textsuperscript{1}The number of doses administered daily per 100 people. See Mathieu et al. (2021).

\textsuperscript{2}The range of results correspond to the various scenarios simulated, ranging from a slowly decreasing epidemic (effective \(R = 0.99\)) to a fast-growing epidemic (\(R = 2\)) peaking three to four months after the start of the simulation. Under the maintained assumption that high-risk groups are prioritized for vaccination, we find that faster vaccination typically has more impact on preventing deaths than infections. We also vary model parameters to account for uncertainty about efficacy of fractional dosing and FDF.

\textsuperscript{3}This includes efficacy against the variants prevailing in the UK in the early part of 2021.
speed, but even if efficacy in the period between the first dose and second dose is as low as 50% (allowing time for immunity from the first dose to develop), we find that FDF reduces infections and mortality in all epidemic scenarios, as long as supply constraints limit the vaccination rate to 1% of the population per day or less. We also find that using a less effective vaccine available immediately instead of waiting for a more effective one reduces infections and mortality. If a 70% effective vaccine is available immediately while a 95% effective vaccine is available in two months, starting vaccinations with the immediately available vaccine would reduce infections by 11–22% and deaths by 20–37%.

Trials to assess the impact of fractional doses on immune response for a variety of vaccines and doses would involve a few hundred people each and could yield preliminary results within weeks. One such trial is currently underway in Belgium (ClinicalTrials.gov ID NCT04852861). Additional trials could be conducted in settings with little or no disease transmission, and among low risk populations. Armed with information from such trials, or even with the data currently available from early stage clinical trials, some jurisdictions might decide to try fractional dosing at scale accompanied by rigorous data gathering, just as the UK and a few other countries decided to implement FDF before complete data on first-dose efficacy was available. For example, a city in Brazil has recently begun an experimental roll-out of half-doses of ChAdOx1 nCoV-19 (the vaccine developed by Oxford and AstraZeneca, and manufactured by multiple organizations: AstraZeneca, FIOCRUZ, R-Pharm, SK Bioscience, and Serum Institute of India (Covishield)) to non-elderly adults (Governo ES, 2021). If fractional doses were found to have sufficiently high individual efficacy based on observational data, these pilots could be scaled up. If not, those who received

---

4Hybrid stretching policies, using status quo dosing regimens for seniors at greatest mortality risk and stretched dosing for younger people, can sometimes dominate the “pure” FDF approach, especially for averting deaths.
fractional doses could receive full doses, either of the original vaccine or of another vaccine (as was done for some viral vector vaccines, or vaccines found to be less effective).\textsuperscript{5} Because tests of fractional dosing are largely reversible and the potential benefits are so great, the costs of these tests would be tiny relative to the expected benefits if there is even a modest chance such policies could succeed. Additionally, fractional dosing can be implemented among some age groups only (e.g. in non-elderly populations), which would further reduce any risks.

Accelerating vaccination with dose-stretching policies would not only reduce infections and deaths, but also promote equity between groups or countries at opposite ends of the global queue for vaccines. For example, if people at the end of the queue would currently wait two years to be vaccinated, doubling supply would reduce their wait time by a full year. For someone near the front of the queue expecting to be vaccinated in a single month, doubling supply would move their vaccination forward by only two weeks. Additionally, rapidly vaccinating as many people as possible could reduce the likelihood of more dangerous variants emerging globally.

Finally, it is worth noting that switching to smaller doses might not just yield an equivalent vaccine, but could also potentially reduce side effects, yielding a superior vaccine.

Our paper adds to a recent literature studying optimal vaccine prioritization,\textsuperscript{6} which focuses on the order in which different groups of people should be served by a fixed supply of vaccine under different epidemic scenarios and efficacy assumptions. Our methods are similar, but our question is different: we focus on ways that a fixed supply can be stretched to protect more people

\textsuperscript{5}Effectiveness of mixing different vaccines has already been established in practice. For example, an ongoing trial in Spain found strong antibody response when combining ChAdOx1 nCoV-19 and BNT162b2 (ClinicalTrials.gov ID NCT04860739).

\textsuperscript{6}See Akbarpour \textit{et al.} 2021; Bubar \textit{et al.} 2021; Gallagher \textit{et al.} 2020; Hogan \textit{et al.} 2020; Matrajt \textit{et al.} 2020; Paltiel \textit{et al.} 2020. Our model only considers scenarios in which an older population cohort must finish vaccinations before the next cohort becomes eligible. In practice, however, logistical constraints often mean it is not possible to maintain strict prioritization. In some cases it may also be more beneficial to vaccinate age groups in a different order. We will examine this in future work.
while potentially sacrificing some individual efficacy. Under normal circumstances, when there is no constraint on vaccine supply, public health is well served by designing vaccine regimens to maximize individual health. In a pandemic, however, vaccination rates may be constrained by vaccine supply, and there may be a race to vaccinate people before they get infected. In this scenario, the vaccine regimen maximizing individual protection might differ from the socially optimal one. Individuals may benefit more from lower efficacy regimens if they allow for an increase in vaccination rates and, consequently, a decrease in overall infection risk. This paper uses simulations to provide quantitative results for impact of dose stretching in realistic epidemic scenarios calibrated to the COVID-19 pandemic. In a companion paper (in progress), we provide qualitative theoretical results on the optimality of dose stretching that hold generally for a more abstract epidemiological model.

The rest of the paper is organized as follows: Section 2 discusses available data on the potential impact of dose stretching policies on efficacy; Section 3 and Section 4 respectively present the epidemiological model and simulation methods; Section 5 estimates the public health impact of speeding up vaccination; Section 6 analyzes a series of specific dose-stretching policies—fractional dosing, longer delays between first and second dose, and utilizing available vaccines rather than waiting for higher efficacy ones—under a variety of assumptions regarding efficacy. We discuss next steps for testing dose stretching policies and the option value of undertaking such testing in Section 7 and conclude in Section 8.
2. Evidence on Efficacy Under Dose-Stretching

In this section we consider evidence on the efficacy of various dose stretching policies, both from clinical trials and from real world vaccination roll-outs.

Fractional dosing for COVID-19 has not yet been tested outside of early stage clinical trials. However, we do have evidence from previous epidemics that very small doses can prove effective in some cases. For example, Brazil successfully used 1/5-doses of yellow fever vaccine to combat an epidemic in 2018 based on advice from the WHO (Pan-American Health Organization 2018). Fractional dosing has also been considered for seasonal influenza (Antony et al., 2020; Pan-American Health Organization, 2009).

Existing evidence from clinical trials suggests that fractional doses of some vaccines could yield high immune responses, comparable to those for standard doses of the same vaccines, and greater than those for some other, already approved vaccines. For example, a trial using ChAdOx1 nCoV-19 (Oxford/AstraZeneca) found that “differences in normalized titre levels [of neutralizing antibody (nAb)] for the standard dose and a low dose [of approximately 1/2 of the standard dose] within [18-55, 56-69, and 70+] age groups were not statistically significant” (Ramasamy et al. 2020). A trial with mRNA-1273 (Moderna) likewise found similar immune responses for both 50 and 100µg (standard) doses (Chu et al. 2021). This suggests significant scope for flexibility in vaccine dosage. This was acknowledged by the scientific advisor to the US’s Operation Warp Speed, Moncef Slaoui, who in January 2021 suggested giving half doses of mRNA-1273 to some adults (Wu, 2021). Recent comments from some vaccine manufacturers agree: Melissa Moore, Chief Scientific Officer for Moderna, discussing vaccine boosters, said that the high dose of mRNA-1273 was used to "guarantee effectiveness" but that she is confident doses would decrease in the
future, "reducing side effects without compromising protection" (Weintraub 2021).

Recent research also suggests that immune response, as measured by neutralizing antibody (NAb) levels, is highly predictive of protection from symptomatic infection, both among convalescent patients and among the vaccinated. Khoury et al. (2021) find a “remarkably predictive” logistic relationship between NAb levels and vaccine efficacy (Spearman $\rho$ of 0.905).\(^7\)

In Figure 1 we use Khoury et al.’s (2021) model (fitted to standard doses of vaccines) and then, assuming the relationship between immune response and efficacy holds, plot additional points based on immune response for different doses of currently used vaccines to add points for alternative doses. Underlying data are given in Table 2 in Appendix A.\(^8\) Despite the exploratory nature of this approach, the results strongly illustrate the potential benefits of adopting fractional doses. We observe that for some vaccines, immune responses associated with high efficacy can be obtained even with much smaller doses. For mRNA-1273 (Moderna), for example, doses 1/2 and 1/4 of the standard both have immune response levels associated with 90-95% efficacy. For BNT162b2 (Pfizer) there is no significant decrease in NAb level for a 2/3 dose in non-elderly populations (albeit with a very small sample size), while NAb levels are associated with efficacy between roughly 70% and 85% for other dose-age combinations. For other vaccines, we also sometimes observe unexpected trends, where lower doses lead to NAb levels associated with higher efficacy (e.g., NVX-CoV2373 (the vaccine developed by Novavax and not yet approved for distribution) and ChAdOx1 nCoV-19 (Oxford/AstraZeneca)). While these results are not necessarily unrealistic (AstraZeneca, 2020), they may be a consequence of limitations of the Khoury et al. modeling.

\(^7\)Data are based on phase 1-3 clinical trials for subsequently approved vaccines, which are publicly available. A similar relationship has also been reported by Earle et al. (2021).

\(^8\)Where available, we used the same studies referenced by Khoury et al. to derive ratios of fractional dose to standard dose immune response. The exceptions were: the 25µg dose of mRNA-1273 (Moderna), for which we used Chu et al. (2021); for BBV152 (the vaccine developed and manufactured by Bharat Biotech, sold as “Covaxin”), Ella et al. (2021); for ChAdOx1 nCoV-19 (Oxford/AstraZeneca), Ramasamy et al. (2020).
Figure 1: Efficacy Associated with Mean Neutralization Levels for Alternative Doses. The curve follows the model derived by Khoury et al. linking NAb levels to protection from symptomatic infection for standard doses of eight vaccines and in convalescents, with the shaded area corresponding to the 95% confidence interval of the model. Lighter data points represent the mean (normalized) immune response and clinical efficacy of specific vaccines (referred to by colors) at standard doses, following Khoury et al.; response in convalescents is also plotted. NAb levels for vaccines are normalized to those of convalescents using clinical trial data for each vaccine. We calculate the ratios of mean NAb responses for alternative versus standard doses using data from clinical trials that tested different doses. We then plot the alternative doses on Khoury et al.’s immunogenicity-efficacy curve as darker shapes. Doses for the elderly are represented by diamonds while doses for non-elderly adults (or all adults, where data is not available by age) are represented by circles. For consistency, if multiple age groups were compared, we use the immune response to the standard dose in younger adults to normalise mean NAb levels. We note small sample sizes, typical of early stage trials, and do not include measures of uncertainty.
approach, or of uncertainty inherent in early-stage clinical trials (including small sample sizes), or both.

As discussed by Khoury *et al.*, a comparison across clinical trials of different vaccines may be biased. Moreover, the relationship that holds across vaccines may not hold across doses of a given vaccine. The entire curve traced by Khoury *et al.* may be shifted downward for some viral variants. This could potentially change the relative benefits of using reduced versus standard doses for some vaccines, but would most likely not change that reduced doses of some vaccines may be more protective than full doses of others. More research is needed to establish the absolute level at which there is protection from infection or severe outcomes, how NAb levels affect efficacy against variants, and how decreases in NAb levels impact longevity of protection (Hall *et al.* 2021, Krammer 2021). On the other hand, another analysis in the same paper by Khoury *et al.* also suggests that even small NAb responses are protective against severe disease and death, potentially increasing the value of fractional dosing. Regardless, our analysis suggests that within-vaccine variations in efficacy are in many cases small compared to differences across vaccines. We return to the discussion of these results in light of our simulation analysis in Section 6.

In contrast to fractional doses, there is considerable evidence on the potential efficacy of a delayed dosing interval between first and second doses—or “first dose first” (FDF)—as such policies have already been enacted in several countries. For example, the default three- to four-week delay for BNT162b2, ChAdOx1 nCoV-19, and mRNA-1273 has been stretched to 12 weeks in the United Kingdom and 16 weeks in Canada. In addition, there are other determinants of vaccine effectiveness; for example, ChAdOx1 nCoV-19 has been shown to elicit a much stronger T cell response than BNT162b2 among the elderly (Parry *et al.* 2021).

In a joint statement the four UK Chief Medical Officers wrote: “We agree with [the Joint Committee for Vaccination and Immunisation] that at this stage of the pandemic prioritising the first doses of vaccine for as many people as possible on the priority list will protect the greatest number of at risk people overall in the shortest possible time and will have the greatest impact on reducing mortality, severe disease and hospitalisations and in protecting the NHS and...
first to 6-8 weeks and now to 12-16 weeks (The Times of India 2021).

The UK decision was initially motivated only by clinical trial evidence, including immune response data (see Appendix A for a non-comprehensive summary of clinical trial evidence on first doses). Subsequent large-scale observational studies, including from vaccination campaigns in the UK and Israel, estimate that a first dose of BNT162b2 is approximately 75% effective at preventing infections (Appendix A presents details of the evidence discussed here). Related studies which also include data for ChAdOx1 nCoV-19 (Oxford/AstraZeneca) find that first doses are 78–94% effective at preventing hospitalizations. Additionally, a recent study suggests that the second dose of ChAdOx1 nCoV-19 is more effective (82.4%) when it comes at 12 weeks, rather than at four weeks (54.9%) as tested in the initial clinical trial (Voysey, et al. 2021).

However, first dose efficacy is much lower for other vaccine-variant combinations. First doses of BNT162b2 and ChAdOx1 nCoV-19 have been significantly less effective against the variants of concern originating in South Africa (B.1.351), the UK (B.1.1.7), and India (B.1.617.2) (Burn-Murdoch et al. 2021). A study of CoronaVac (the vaccine developed by Sinovac and manufactured by multiple organizations: Sinovac, Instituto Butanta, Bio Farma, and Pharmaniaga) in Chile reported good efficacy against the standard variant after two doses, but very low efficacy after one dose (Dyer 2021, Taylor 2021).\footnote{We provide references and summarize information from these trials in Appendix A. We note that efficacy of first doses is also broadly consistent with the relationship between immune response and efficacy against symptomatic infection as laid out by Khoury et al. nAB levels from two doses are typically several times higher than for one dose, so while single doses may be sufficient for high efficacy against the original SARS-CoV2 virus, they may be less effective against variants with some degree of immune evasion.}

We also note, however, that as the number of natural infections in a community increases, the risk of providing just a single dose decreases, since some evidence suggests that a single shot provides as much protection for those who have been previously infected as two doses in those never infected (Ebinger et al. 2021; Stamatatos et al. 2021; Willyard 2021).
Some countries have also decided to use vaccines that were available immediately rather than waiting for more effective ones. For example, the UK, Europe, and Canada rolled out vaccinations with a combination ChAdOx1 nCoV-19 and BNT162b2, rather than waiting for more mRNA vaccines to be available. Chile and Seychelles have relied significantly on CoronaVac, with less robust efficacy data, to start vaccination campaigns earlier. Case rates remain high in both countries (Mathieu et al. 2021), although the vaccine does appear to be protective against hospitalization and death (as suggested in a recent study in Serrano, Brazil, where almost the entire adult population has been vaccinated (e.g., Pearson 2021)).

The epidemiological modeling that follows excludes the potential impact of dosing stretching policies on immune escape through mutation. Some have argued that dose stretching policies might exacerbate this risk due to a prolonged period of partial immunity increasing the risk of immune escape. However, many epidemiologists now believe that dose stretching may instead reduce the probability of immune escape (e.g., Cobey et al. 2021). We expand on the current evidence in Appendix A.

In summary, FDF approaches and using available vaccines early rather than waiting for more effective ones have been tested as vaccination campaigns rolled out. We therefore have good information on their effectiveness. We do not yet have efficacy data for fractional doses, but immune response data suggests that smaller doses of some vaccines may be very effective.

3. Epidemiological Model

Since the evidence on fractional dose efficacy is not dispositive, we model the potential impact on the pandemic of a range of efficacy levels using a standard epidemiological model. The model we
use extends the canonical susceptible-exposed-infectious-recovered (SEIR) model, which is widely used in mathematical epidemiology to characterize the spread of an infectious disease in a closed population (Kermack and McKendrick 1991, Anderson and May 1992). The SEIR model assumes individuals flow between disease and vaccination states over time, with sizes of population in each state changing according to a set of differential equations. We extend the canonical SEIR model to allow for death and vaccination (which is ineffective for some individuals), yielding the following equations:

\[ \dot{S}_i(t) = -\lambda_i(t)S_i(t) - v_i(t)\delta S_i(t) \]  
\[ \dot{E}_i(t) = \lambda_i(t)[S_i(t) + N_i(t)] - \gamma' E_i(t) \]  
\[ \dot{I}_i(t) = \gamma' E_i(t) - \gamma'' I_i(t) \]  
\[ \dot{D}_i(t) = p_i \gamma'' I_i(t) \]  
\[ \dot{R}_i(t) = (1 - p_i) \gamma'' I_i(t) - v_i(t)\delta R_i(t) \]  
\[ \dot{P}_i(t) = v_i(t)\delta [e\tilde{S}_i(t) + \tilde{R}_i(t)] \]  
\[ \dot{N}_i(t) = v_i(t)\delta (1 - e)\tilde{S}_i(t) - \lambda_i(t)N_i(t). \]

Dots denote derivatives with respect to time. Uppercase letters denote population compartments (i.e., the fraction of the population in a given state): \( S \) for susceptible, \( E \) for exposed (individuals carrying the virus, but who are not yet contagious), \( I \) for infectious, \( R \) for recovered, \( D \) for dead, \( P \) for protected by vaccine, and \( N \) for vaccinated but not protected. The population is divided into \( G \) age cohorts, indexed by \( i = 1, \ldots, G \), with respective sizes \( n_i \). Subscripting compartments by \( i \) allows for different epidemic evolution across age cohorts. Tildes denote the size of the
Figure 2: Compartment Flows in Epidemiological Model. Model described by Equations (1)–(7). Solid black lines reflect virus model and dashed lines vaccination. The full model has separate compartments for each age group $i$.

compartment in proportion to both compartments receiving vaccines (susceptible and recovered) i.e.,

\[
\tilde{S}_i(t) = \frac{S_i(t)}{S_i(t) + R_i(t)} \\
\tilde{R}_i(t) = \frac{R_i(t)}{S_i(t) + R_i(t)}.
\]

Figure 2 depicts the population flows between the compartments described in equations (1)–(7). Our initial model assumes a single vaccine, either with a single dose, or where efficacy does not change between two doses. Lowercase letters denote model parameters governing the evolution of compartments. All parameters except $e$ are age-specific, as denoted by subscript $i$. $\gamma'$ and $\gamma''$ are, respectively, the hazard rates of moving from exposed to infected and from infected to recovered or dead. These are estimated as the reciprocals of the durations of the virus’s incubation period and of the infectious period, respectively. The rate of new infections equals $\lambda_i(t)$, described in further
detail below. Parameter $p_i$ is the mortality risk. Vaccine efficacy, denoted $e$, is the probability the vaccine protects from infection. The model makes no distinction between the vaccine’s *efficacy* (performance measured in clinical trials) and *effectiveness* (performance in practice in the population); $e$ is used to denote both interchangeably. We assume recovered individuals (compartment $R$) are perfectly protected by vaccination and that exposed or infectious individuals (compartments $E$ and $I$) are not vaccinated.

To account for vaccine prioritization, we introduce an indicator variable $v_i(t)$, switching from 0 to 1 on the day age cohort $i$ becomes eligible for vaccination and to 0 again at the point where all willing members of the cohort have been vaccinated. Reflecting common practice, we assume older cohorts must finish vaccinations before the next cohort becomes eligible.\(^\text{12}\) When $v_i(t) = 1$, age cohort $i$ is vaccinated at a constant rate $\delta_i$, drawing on a continuous stream of vaccine production from a given capacity. To keep track of cumulative doses distributed, we introduce the auxiliary compartment $V$, where $\dot{V}_i(t) = \delta_i v_i(t)$, and where $V(t)$ is the proportion vaccinated in the entire population.

The rate of new infections, $\lambda_i(t)$, depends on the number of daily contacts a susceptible individual has with currently infectious individuals. To reflect differences in interaction across age cohorts, we use a contact matrix $C$, where entry $c(i, j) \geq 0$ denotes the number of contacts made by an individual in cohort $i$ with an individual in cohort $j$. To derive the proportion of each age group infected at time $t$, each contact is scaled by the probability of virus transmission on contact,

\(^{12}\)Most recent studies, e.g. by Bubar et al. (2021), find that vaccination of those aged 60 and older reduces the health burden more than for other age groups under a wide range of scenarios, especially in high-income countries. On the other hand, Matrajt et al. (2020) find that vaccinating high-transmission groups (including children) first can be optimal if the vaccine is highly effective and supply is very large, conditions that often apply to infectious diseases such as seasonal influenza outside of a pandemic. We will investigate different prioritization strategies in future versions of this work.
\( q \), and probability that the contacted person is infected, \( I_j(t) \), yielding

\[
\lambda_i(t) = q \sum_{j=1}^{G} c(i, j) I_j(t) \tag{10}
\]

For a given \( C \), \( q \) can be adjusted to match any desired reproductive number \( R \) for the virus (i.e., the number of secondary cases produced by a single infection).

The initial conditions of the system (1)–(7) require specifying the proportion of the population that is susceptible \( S(0) \), immune \( R(0) \), and infectious \( I(0) \) at the outset of the epidemic. We take \( I(0) \) to be small and for simplicity take \( E(0) = I(0) \). We assume that the proportion of each age cohort in each initial compartment is the same as in the overall population.

4. Simulation Methods and Parameters

This section discusses some of the main assumptions of the model and parametrization. A complete list of the parameters used in the simulations is provided in Table 3 in Appendix D. Here we highlight the most critical parametric assumptions. We discuss initial conditions in Section 4.1, the time horizon covered by the simulation in Section 4.2, assumptions related to disease spreading and burden in Section 4.3, assumptions concerning vaccination efficacy and constraints in Section 4.4, and the simulation methods in Section 4.5.

4.1. Initial Conditions

We run simulations for three illustrative epidemic scenarios. The slow decrease scenario sets the initial effective reproduction rate to \( R = 0.99 \) and initial infectious proportion to \( I(0) = 1\% \).\(^{13}\) This since we will assume 20\% of pre-existing protection in people aged 20 and over, the initial effective reproductive number \( R \) is lower than \( R_0 \), the basic reproductive number in a fully susceptible population.
scenario may capture a situation in which non-pharmaceutical interventions (NPIs) are introduced following an epidemic wave but are only effective enough to decrease cases slowly. The slow growth scenario sets $R = 1.1$ and $I(0) = 0.5\%$, perhaps reflecting a situation in which NPIs are not effective enough to prevent a subsequent wave of infections, such as the one experienced by the United States in late 2020. The fast growth scenario sets $R = 2$ and $I(0) = 0.1\%$, e.g., a case when a new virus strain suddenly emerges, thwarting previously effective NPIs (such the one as observed in the United Kingdom in December 2020, or the emergence of the P.1 variant in Brazil in late 2020\textsuperscript{14}). In both growth scenarios, $I(0)$ is adjusted so that the peak of infections occurs three to four months from the start of vaccinations.\textsuperscript{15}

We choose parameters for initial immunity that broadly reflect the state of the COVID-19 pandemic in early 2021. We assume 20\% of people aged 20 and over have immunity acquired from infection, leaving 80\% susceptible. To reflect the lower clinical case rate in the younger population (Davies \textit{et al.} 2020; Goldstein, Lipsitch, and Cevik 2020), we assume only 50\% of under 20s are susceptible.

### 4.2. Time Horizon

Each simulation runs for $T = 365$ days. This is sufficient time for the epidemic to die out in the scenarios considered but, we assume, not long enough for unmodeled factors to come into play, such as the alleviation of supply constraints with expanded capacity or the waning of vaccine protection from initial doses, perhaps warranting booster shots.\textsuperscript{16} Similarly, we assume that there is

\textsuperscript{14}See Sabino \textit{et al.} 2021.

\textsuperscript{15}Assuming that the epidemic peaks earlier or that vaccinations start in the declining phase of the epidemic would decrease measured vaccine benefits, as discussed in Section 6.1.

\textsuperscript{16}The choice of $T$ might play a bigger role in models where the reproductive number is close to 1, for example, due to behavioral responses to risk, as in Gans (2020).
no natural loss of immunity (no flow from recovered to susceptible) during the simulation period.\footnote{Muena \textit{et al.} (2021) show that neutralizing antibody responses can persist up to 12 months after infection. Hall \textit{et al.} (2021) show high levels of protection at six months after infection. In both cases, the upper limit is due to the lack of longer-term data at the time of publication.}

4.3. Infections and Deaths

We use a social contact matrix $c(i, j)$ based on a large cross-country study of contacts between different age groups, primarily in European countries (Mossong \textit{et al}. 2008). Our matrix is therefore more representative of high-income countries, but we are not aware of comparable data on social mixing in low-income countries. For most of the paper, cohort size ($n_i$) and mortality risk ($p_i$) for different age cohorts is based on data for high-income countries, although we also perform a separate analysis for low-income countries (see Appendix C). Throughout the age distribution, the risk of death from COVID-19 increases rapidly with age, about three-fold per decade (Manheim \textit{et al}. 2021).

The model assumes that contact frequencies are independent of infection risk, precluding behavioral changes in response to changes in infection risk as the epidemic progresses. Since we abstract from the emergence of viral variants, epidemics always have a single peak and fade out when the virus’s effective reproductive number satisfies $R \leq 1$, which happens when a sufficiently high fraction of population is protected, either by vaccination or recovery from natural infection.
4.4. Vaccination

The base case for vaccination is a 95% effective vaccine, when used as tested in Phase 3 trials (standard dosing, with a delay between two doses). We assume that those under 20 (constituting 22% of population in our base case simulations) receive no vaccination. To account for vaccine hesitancy, we assume 20% in each age group refuse vaccination. We assume that the vaccine becomes effective 10 days after it is administered.\textsuperscript{18}

As of early May 2021, the world is vaccinating at a rate of approximately 0.25% of the population per day (Mathieu et al. 2021), our base case immunization speed.\textsuperscript{19} At the high end, the United Kingdom, United States, Canada, Chile and Israel have all managed to vaccinate at rates well above 0.8% of the population per day;\textsuperscript{20} however, the current median global rate of vaccination (as of May) is only 0.27% of the population per day (Mathieu et al., 2021). Thus, at a global level, supply rather than delivery logistics or demand (e.g., vaccine hesitancy) seem likely to constrain full vaccination well into 2022, and perhaps for considerably longer.

Accordingly, our model is intended to apply to contexts in which vaccination rates are constrained primarily by the available supply.\textsuperscript{21} While this may not apply for some countries, this seems broadly to be the case globally. The model could be extended to consider other scenarios.

\textsuperscript{18}We achieve this by treating vaccinated compartments in the model as “effectively vaccinated”. Hence if vaccinations in a given age group start of day $t_1$ and end of $t_2$, we start the flow into vaccinated compartments on date $t_1 + 10$ and stop it on $t_2 + 10$.

\textsuperscript{19}The world rate of vaccination has been increasing since early May, driven in part by China.

\textsuperscript{20}Over two weeks in April and May, Mongolia was vaccinating 2% of its population per day.

\textsuperscript{21}In general, we are also agnostic as to whether shortages are due to shortages in drug substance or, for example, fill and finish capacity. With fractional dosing, the number of doses in each vial increases, so it alleviates constraints wherever they are in the production system. FDF and using lower efficacy vaccines early are more about the timing of vaccine use, and hence are orthogonal to where the constraints in vaccine supply are. If the constraints are in fill and finish capacity only, then another option - filling more doses in each vial - also becomes an option to alleviate constraints. However, this is also likely to increase waste to some extent, and force more centralized, camp-style vaccination approaches rather than vaccine delivery where people usually get health care, and so may eventually, in some contexts, make it hard to drive high take-up. This may also be true of fractional dosing - which effectively increases the number of doses per vial.
where, for example, delivery constraints might at some point be binding.

Additionally, while we treat efficacy as a scalar, in reality it is multidimensional: vaccines may differ in efficacy against different variants, in duration of protection, or in their protection against infection and disease.

### 4.5. Simulation Method

We generate a simulation run for each configuration of parameters by finding the deterministic solution of the differential-equation system consisting of these equations (1)–(7) using standard numerical methods.22

Figure 3 illustrates the evolution of vaccinations and infections for the various epidemic scenarios and vaccination rates analyzed. We will discuss the impact of vaccination below. With no vaccination, we find that from 8% (slow decrease scenario) to 55% (fast growth scenario) of the population get infected during the simulation period. Individuals aged 20 to 49 are responsible for between 55% and 59% (depending on the scenario) of all infections, assuming no vaccine. This is consistent with recent findings by Monod et al. (2021), who estimated that three quarters of infections in the US originated from individuals in that age bracket (albeit in a period with school closures). Figure 13 in Appendix D illustrates the age-specific dynamics of vaccination and infection in the model.

The outcome variables for our simulations are the burden of infection, defined as the proportion of the total population that develop new infections during the simulation period, and the burden of death, defined as the proportion of the total population that die during the simulation period. We

22We solve all differential-equation systems using the odin package, version 1.0.8, and generate exhibits using R, version 4.0.2. All code used in this project is available at https://github.com/wwiecek/covstretch.
focus on these rather than alternative outcomes to present a reasonable number of results.\footnote{Additional outcome variables could include hospitalizations and severe infections, both of which are correlated with age. Our focus on counts of infections and deaths allows us to abstract from the delay between infection and illness or death, which would have to be parametrized in an analysis of healthcare use.}

5. Benefits of Speeding Vaccination, Holding Efficacy Constant

In this section we examine the potential impact of accelerating vaccination through a dose stretching policy if it entails no loss of efficacy. Section 5.1 considers increasing the rate of vaccination, while Section 5.2 considers starting the vaccination campaign earlier. Subsequent sections consider the case when dose stretching policies come at the cost of reduced efficacy.

5.1. Faster Vaccination Rate

Increasing vaccination rates without sacrificing efficacy would dramatically reduce disease burdens, as shown in Figure 4. For example, the bottom left panel indicates that vaccinating 0.25% of the population daily averts 18–50% of infections and 46–71% of deaths across epidemic scenar-
**Epidemic scenario**

- Slow-decrease
- Slow-growth
- Fast-growth

<table>
<thead>
<tr>
<th>Infections</th>
<th>Deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.00</td>
<td>1.00</td>
</tr>
<tr>
<td>0.75</td>
<td>0.50</td>
</tr>
<tr>
<td>0.25</td>
<td>0.10</td>
</tr>
<tr>
<td>0.10</td>
<td>0.01</td>
</tr>
<tr>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>0.10</td>
<td>0.30</td>
</tr>
</tbody>
</table>

**Percentage of pop. vaccinated daily**

<table>
<thead>
<tr>
<th>% burden averted</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
</tr>
<tr>
<td>75</td>
</tr>
<tr>
<td>50</td>
</tr>
<tr>
<td>25</td>
</tr>
<tr>
<td>0</td>
</tr>
</tbody>
</table>

**Burden (log scale)**

Infections

- Slow-decrease (gold)
- Slow-growth (orange)
- Fast-growth (red)

Deaths

- Slow-decrease (gold)
- Slow-growth (orange)
- Fast-growth (red)

**Percentage of pop. vaccinated daily**

- 0.01
- 0.03
- 0.10
- 0.30
- 1.00
- 2.00

**Reductions**

Infections

- Slow-decrease (gold)
- Slow-growth (orange)
- Fast-growth (red)

Deaths

- Slow-decrease (gold)
- Slow-growth (orange)
- Fast-growth (red)

**Figure 4: Benefits of Faster Vaccination.** The top panels show simulation results for burdens under various vaccination rates. The bottom panels show simulation results for the percentage reduction in burdens relative to no vaccination. In the left panels, we refer to the burden from infection, while in the right panels we refer to the burden from death. In all scenarios we assume a 95% effective vaccine and sequential age prioritization.

In contrast, if the vaccination rate was doubled to 0.5%, 41–70% of infections and 58–80% of deaths could be averted. The top panels show that increasing vaccination rates reduces disease burdens considerably more in absolute terms in the fast-growth scenario than in other epidemic scenarios. This is not surprising since disease burdens are highest in this scenario.

Typically, faster vaccination reduces burden at a decreasing rate. This pattern is depicted by the concavity of most of the burden reduction curves, especially for deaths. This effect is a conse-
quence of age prioritization: vaccinating the highest mortality risk people first allows more deaths to be averted with relatively fewer vaccines. Higher vaccination rates mainly speed the vaccination of lower mortality risk people. The biggest departure from decreasing returns to faster vaccination is observed with infections in the fast-growth scenario, where some convexity can be seen in the bottom panel. The pattern of the results across infections and deaths is roughly similar to that found in other recent models of COVID-19 vaccination (Bubar et al. 2021; Hogan et al. 2020; Matrajt et al. 2020).

5.2. Earlier Vaccination Start

Another way of speeding vaccinations is to start the vaccination campaign earlier. Figure 5 presents results from simulations in which the start of the campaign is varied from four months before the peak of the epidemic to three months after. We restrict attention to a fast-growth epidemic.

Consistent with the results in the previous section, for any start date, the higher the vaccination rate, the greater the burden averted. Additionally, for any vaccination rate, the earlier the campaign is started, the greater the burden averted. Faster and earlier vaccination campaigns allow more people to be protected before they are exposed to infection.

If the campaign starts early enough in the epidemic, the disease burden can be almost entirely averted, even for low daily vaccination rates. Such behaviour is typical in deterministic compartmental models such as SEIR. At the other extreme, if the campaign does not start until the population is close to herd immunity, the campaign hardly reduces disease burden, regardless of the vaccination rate, since most infections and deaths will have already occurred.
6. Impact of dose stretching policies with efficacy trade-offs

This section provides simulation results on the health impact of a series of alternative dosing policies relative to the status quo, under a range of plausible assumptions regarding potential trade-offs between these policies and individual vaccine efficacy. We analyze fractional dosing in Section 6.1, FDF in Section 6.2, and using an inferior vaccine that is available sooner in Section 6.3.

6.1. Fractional Dosing

Fractional dosing involves giving less than the standard dose of a vaccine to obtain more doses from limited capacity.

Because we consider the case in which the pace of immunization is subject only to a supply constraint, the vaccination rate is proportional to the reciprocal of dose size. So, for example, using...
half rather than full doses would double the vaccination rate. In this section, we analyze the impact of fractional dosing on the burden of infections and deaths while varying three variables: dose fractions, efficacy reductions associated with moving to fractional dosing, and epidemic scenarios.

We return to the baseline assumption that the vaccine is available from $t = 0$ and maintain the other baseline assumptions, including that a full dose has 95% efficacy and can be supplied at a rate of 0.25% of the population per day. We now assume that a dose fraction $\phi \in (0, 1]$ can be used, increasing the vaccination rate to $0.25%/\phi$. We initially consider dose size to be uniform across age groups, but we discuss potential alternatives at the end of this section.

Figure 6 presents the simulation results, comparing the impact of fractional versus full dosing. Using the estimated efficacies suggested by our exploration in Figure 1, we can estimate the potential benefits of fractional dosing. For illustration, we consider the cases of three of the vaccines with largest production: mRNA-1273 (Moderna), BNT162b2 (Pfizer), and ChAdOx1 nCoV-19 (Oxford/AstraZeneca). For mRNA-1273, a half dose leads to NAb levels associated with an efficacy of approximately 95%. In this case, moving to half doses would avert 24–29% of infections and 22–47% of deaths. For BNT162b2, a 1/3 dose produces NAb levels associated with approximately 80% efficacy, so moving to 1/3 doses would reduce 34–52% of infections and 25–49% of deaths. Although not shown in our results (efficacy of a standard dose is not 95%), halving the dose of ChAdOx1 nCoV-19 would also lead to reductions in infections and deaths, as there is no observed impact on immune response and therefore no associated loss of vaccine efficacy in comparison to the standard dose.

Significant gains are possible even if the expected efficacies of fractional doses are not as high as those presented above. For the epidemic scenarios that we consider, moving from a full to a half dose (thereby doubling vaccination speed to 0.5% per day) reduces the burden of infections and
death as long as efficacy is at least 70%. Moving from a full to a quarter dose (thereby increasing speed four-fold, to 1% per day) decreases the burden of infections and death even if efficacy drops to 50%. The health benefits can be substantial: for example, moving from a full to a half dose that is 80% as effective reduces the burden of infection by 21–36% relative to that of using the full dose regimen, depending on the epidemic scenario, while reducing deaths by 14–25%.

We find that fractional doses generally reduce infections more than deaths. For example, moving from a full to a half dose that is 50% as effective reduces infections across all epidemic scenarios, but increases deaths. The reason is that the baseline vaccination rate of 0.25% per day is already fast enough to protect older cohorts with high mortality risks, given they receive priority in
vaccination. Therefore, further speed gains do not offset the decrease in efficacy for older groups. The potential for increasing the burden of deaths by reducing efficacy is particularly apparent in the fast-growth scenario, where a half dose increases death burden unless its efficacy is at least 70%. Hence, trade-offs between efficacy and speed are sensitive to both epidemic context and the burden measure prioritized by the policymaker.

Our base case assumed a daily vaccination speed of 0.25% per day. We replicate our analysis for vaccination speeds between 0.1% and 2%, with results presented in Figure 14 in Appendix D. Predictably, we find that adjusting dose size is sensitive to all factors that we investigate (epidemic, vaccine supply, loss of efficacy from fractional dosing). However, fractional dosing is preferred in most situations, with the exception of fast-growing epidemics combined with a large vaccine supply.24

The fractional dosing policies we have discussed here are uniform across age cohorts. In simulations, we found that switching from uniform to age-varying doses produces only small gains relative to switching from full doses to the optimal uniform fractional dose.25 However, as evidenced by clinical trial data, it is likely that for some vaccines, efficacy will vary across age groups. In such situations, the additional benefits from using age-varying an fractional dosing strategy will be larger.

The three vaccines discussed above (mRNA-1273, BNT162b2, ChAdOx1 nCoV-19) are among those with the largest production levels, with total manufacturing of approximately 5-6 billion

\[\text{For example, if a half dose is 80\% effective, then it would reduce deaths and infections relative to a 95\% effective dose under all scenarios, except if the epidemic is growing fast (R = 2) and daily vaccinations with full dose are at 1\% of the population or more.}\]

\[\text{For example, in a fast-growing epidemic, vaccinating 30\% of the population (with age prioritization) averts 82\% of deaths suffered under no vaccination. The optimal uniform fractional dose is 3/10 of a full dose, averting 93.3\% of deaths under no vaccination. The optimal age-varying fractional dosing scheme averts 93.6\% of deaths under no vaccination, only 0.3\% more.}\]
doses projected for 2021 (Burger 2021, Moderna 2021, AstraZeneca 2021). The majority of these
doses have not yet been delivered to people’s arms, and so could potentially be available for frac-
tional dosing. Moderna has also announced plans to increase production to 3 billion doses in 2022.
Although it is hard to accurately predict the number of extra doses that would be generated from
fractional dosing, these figures give us an idea of the huge potential for extra supply that could
come at no or little cost to efficacy. These benefits could also be available quickly, unlike the
benefits of expanding production, which occur with some delay.

6.2. First Doses First

Next, we consider the effect of delaying the second dose of a two-dose vaccination course so that
more people can receive their first dose sooner. Here too we vary assumed efficacy of the first dose
to understand under what range of efficacies the alternative policies are preferred.

The basic model in equations (1)–(7) needs to be extended to allow one or two doses to have
differing efficacy and to connect the delay between doses to the rate of vaccination. We maintain
the assumption that a single vaccine candidate is available. Figure 7 provides a schematic diagram
of the compartment flows in the extended two-dose model. Differential equations are provided in
Appendix B.

Subscripts 1 and 2 indicate variables associated with the first and second doses. Thus, compart-
ments $P_1$ and $N_1$ contain people protected or not, respectively, following their first vaccine dose.
First doses are administered at a rate of $\delta_{1i}$ in age group $i$, and efficacy for those vaccinated with
a single dose is $e_1$. Individuals in compartments $P_1$ and $N_1$ receive a second vaccine dose at a rate
of $\delta_{2i}$. With probability $e_2$, the second dose is effective and the recipient flows into compartment
$P_2$; otherwise, the second dose is ineffective and the individual flows into $N_2$. Let the auxiliary
variable $V_i(t) = P_{1i}(t) + N_{1i}(t) + P_{2i}(t) + N_{2i}(t)$ denote cumulative vaccinations, counting both first and second doses (so a cohort is double counted if all members receive both doses).

We take the status quo policy, second doses first (SDF), as requiring a four-week gap between doses. Setting aside the ten days required to develop immunity, the effective gap is $\delta_{2i} = 18$ days for all age cohorts under a SDF policy. For the first doses first (FDF) policy, we assume a delay of 12 weeks until second dose ($\delta_{2i} = 74$). We also consider a class of hybrid policies (HDF-$a$), with a 12-week delay for those under age $a$ and a 4-week delay for those $a$ and over. Hybrid policies maintain high levels of protection for older cohorts while prioritizing speed over efficacy in younger cohorts. All simulations fix $e_2 = 0.95$ but vary $e_1$. As in previous sections, we consider a range of vaccination speeds. Here, however, delicate calculations are required to ensure

---

26 For BNT162b2 (Pfizer) and mRNA-1273 (Moderna), gaps of three and four weeks, respectively, are recommended. Using the longer of the two gaps for the SDF baseline makes SDF look more like FDF and thus leads to a more conservative measure of the difference between the policies.

27 Same as in the earlier model, to account for this delay we set indicators $v_i(t)$ to correspond to effective vaccination times, not time of injection.
the consistency of speed across alternative policies.\textsuperscript{28}

Figure 8 shows the evolution of vaccine protection provided by one and two doses under SDF, FDF, and HDF-60 policies for baseline parameters. FDF quickly provides protection to a greater share of the population with the first dose than SDF; HDF-60 does not diverge from SDF until all individuals aged 60 and over receive their first dose. The third panel verifies that the overall speed of vaccinations is consistent across policies, as the paths of cumulative doses used virtually overlap.

Figure 9 presents results on the ratio of benefits under the status quo SDF policy to benefits under the best alternative policy, whether FDF or HDF-\(a\). Benefits in the top panels are averted infections and in the bottom panels are averted deaths. Each tile represents a different combination of epidemic growth, level of efficacy of first vaccine dose, and speed of vaccination campaign.

For most entries, a policy involving some stretching, either “pure” FDF or a hybrid (HDF) policy, averts more infections and deaths than SDF. SDF does better only when the vaccination rate is extremely high (2\% of population per day receiving first doses under SDF) and the efficacy of a single dose low (\(e_1 \leq 0.7\) for the fast growth scenario and \(e_1 \leq 0.5\) for slow growth). SDF does better under those limited circumstances because such a rapid vaccination rate reduces scarcity and

\textsuperscript{28}Fixing the target speed under examination, we first compute the \(\delta_1\) needed to generate that speed under SDF as a baseline. We then compute the \(\delta_1\) under FDF so that the time series of doses used matches that under SDF as closely as possible. Since an exact match is impossible, we use the value of \(\delta_1\) for FDF that minimizes the mean squared error.

\[
\sum_{t=1}^{T'} [V_{FDF}(t) - V_{SDF}(t)]^2
\]

between the cumulative number of vaccine doses administered under FDF in each period, \(V_{FDF}(t)\), and under SDF, \(V_{SDF}(t)\), over the entire time period (\(T' = T = 365\)) or while vaccinations are happening (in case all those eligible and willing to get vaccinated do so in less than one year). We compute \(V_{FDF}(t)\) via simulation for a grid of integer values of \(1/\delta_1\) and select the value minimizing the expression above. A similar procedure is used for HDF-\(a\), with the exception that the status quo vaccination rate is maintained for those aged \(a\) and over. For example, setting \(\delta_1 = 1/180\) under SDF, the mean-squared-error minimization leads to \(\delta_1 = 1/145\) under FDF. Under HDF-60, we maintain \(\delta_1 = 1/180\) for cohorts age 60 and over. The mean-squared-error minimization leads to \(\delta_1 = 1/155\) for cohorts under age 60.
Figure 8: Vaccine Coverage under Various Delays in Second Dose. Proportion of population protected by one dose (first panel) or two doses (second panel) under various policies, assuming $e_1 = 0.8$ and $e_2 = 0.95$. Note that the first two panels do not show the total number of people that received one or two doses but rather the number of people that received one or two doses and became protected. SDF refers to second doses first, the status quo policy of delaying four weeks until the second dose; FDF refers to first doses first, delaying 12 weeks until the second dose; HDF-60 refers to a hybrid policy of delaying four weeks for age 60 and over and 12 weeks for under 60. For SDF, we set $1/\delta_1 = 0.25\%$ vaccinated each day with the first dose and $\delta_2 = 1/18$. Then $\delta_1$ is adjusted for FDF and HDF-60 so that the time series of total vaccinations across policies match (third panel). Note cumulative doses can exceed 100% as all individuals eventually receive two doses.

the need to trade off efficacy versus speed.

FDF for all age groups is nearly always better at reducing infections than other policies. One or another HDF policy is better at reducing mortality than FDF at the lowest levels of $e_1$. FDF increases protection for younger individuals, achieving reductions in community transmission, but at the cost of reducing efficacy in the elderly, who face the greatest mortality risk.

The relative benefit of delaying first doses varies with the vaccination rate in an inverted U-shape. When the vaccine is rolled out very slowly, say, 0.1% first doses per day under SDF, FDF averts no more than a few percentage points of disease burden, even if delay results in little to no loss of efficacy. For intermediate vaccinate rates, FDF can avert substantially more burden than SDF. For example, averted infections are 46% higher and averted deaths 48% higher for a vaccination rate of 0.75% first doses per day under SDF under the fast-growth scenario. At very
Figure 9: Burden Averted under SDF Relative to Best Alternative Policy. Entries present simulation results for ratio of burden averted under status quo SDF policy to best alternative. Greater than 1 favors SDF, and less than 1 favors an alternative. Color indicates optimal policy, ranging from dark green when FDF is optimal to light purple when SDF is optimal, with an HDF policy optimal for colors in between. Each tile represents a different combination of epidemic growth, level of efficacy of first vaccine dose, and speed of vaccination campaign. For a comparison of the burden averted under various vaccine policies relative to no vaccination rather than relative to each other, see Figure 16 in Appendix D.

In practice, FDF may be better than SDF or hybrid policies in reducing the burden of infection and death in the population, provided the efficacy from the first dose is sufficiently high for some COVID-19 vaccines. For vaccines for which efficacy following first dose drops below 80%, a hybrid policy may be best. Conclusions about the relative benefits of FDF policies are sensitive to assumptions about efficacy, epidemic growth, and vaccination rate.
6.3. Using Available Vaccines Immediately Versus Waiting for More Effective Vaccines

Some countries may face a choice between using a vaccine available immediately or waiting in a queue for a more effective vaccine.

If the planner has to choose one vaccine or the other and cannot switch later when the more effective vaccine becomes available, analysis of the optimal country decision is straightforward using simulations based on the basic model. The trade-offs are similar to those involved in fractional dosing: speeding supply at the cost of lower efficacy. Formally, suppose there are two possible vaccines. Vaccine 1 with efficacy $e_1$ is available starting at time $t_1 = 0$. Vaccine 2 with efficacy $e_2$ becomes available at time $t_2 > 0$. Once either vaccine becomes available, it can be produced at constant rate of $\delta_1 = \delta_2 = 0.25\%$ of the population daily. We will fix $e_2 = 95\%$ and analyze a range of efficacies for vaccine 1, $e_1 \leq 95\%$.

Figure 10 presents results comparing the ratio of burdens averted by the vaccines in simulations. Predictably, the longer the delay in availability of the more-effective vaccine, starting with the less effective vaccine becomes relatively more beneficial. We find that a delay of two months or more tips the balance toward immediate use of the less effective vaccine. Although their exact magnitudes depend on assumed parameter values and epidemic scenarios, the results can be stark. If the immediately available vaccine is 70% effective (comparable to ChAdOx1 nCoV-19 (Oxford/AstraZeneca)) and a 95% effective vaccine is available two months later, then starting vaccinations with the immediately available vaccine would reduce infections by 11–22% and deaths by 20–37%.

If the planner can start with the immediately available vaccine and freely switch to the more
effective vaccine when it becomes available later, switching dominates continuing to use the first vaccine. Anyone who would have received the immediately available vaccine after time $t_2$ can be given the more effective vaccine instead, providing more protection. Comparing switching to waiting for the more effective vaccine requires analysis; being vaccinated with an immediately available vaccine that is less effective may do little to quell the epidemic and prevents the individuals from receiving more protection later (unless re-vaccination with a booster shot of a more effective vaccine is feasible).

Formally, we need to modify the single-vaccine model given by equations (1)–(7) to account for sequential use of two vaccines. We can do this by adding protected compartments, $P_i$, and

**Figure 10: Burden Averted Using Immediately Available Vaccine Relative to More Effective Vaccine Available with Delay.** Entries are simulation results for ratio of burden averted using a 95% effective vaccine available after delay $t_2$ indicated on horizontal axis, to that from an immediately available vaccine with efficacy $e_1$ indicated on the vertical axis. Entries greater than 1 favor waiting for the 95% effective vaccine, while entries less than 1 favor use of the less effective vaccine. Only the situation where it is not feasible to switch to the more effective vaccine later, when it becomes available, is considered here.
Figure 11: Compartment Flows in Model with Two Vaccines. Differential equations provided in equations (B10)–(B18) in Appendix B. Additional notes from Figure 2 apply.

$P_{2i}$, not-protected compartments, $N_{1i}$ and $N_{2i}$, vaccination rates, $\delta_{1i}$ and $\delta_{2i}$, and vaccine access indicators $v_{1i}(t)$ and $v_{2i}(t)$ for each of the two vaccines. Differential equations for the model with two vaccines are provided in equations (B10)–(B18) in Appendix B. Figure 11 provides a schematic diagram of the compartment flows for the two-vaccine model.

The results are relegated to Figure 15 in Appendix D because they are easy to summarize in words. The few cases in which delay dominated immediate vaccination in Figure 10 have disappeared. Starting vaccination immediately and switching later strictly dominates waiting in all cases with any nontrivial delay period.

It is important to acknowledge important limitations. Our analysis has a $T = 365$ day horizon and is restricted to a single epidemic outbreak, thus abstracting from subsequent epidemic waves and benefits of reaching herd immunity. We omit these factors in the belief that beyond this period, it is reasonably likely that additional supply could come online, and policy could be adjusted accordingly. Moreover, uncertainty involving booster shots or seasonal vaccination makes a
calculation of longer-term benefits too speculative at this point.


As discussed in Section 2, existing immunogenicity data from clinical trials suggest that effectiveness for fractional doses of at least some vaccines could potentially be high (see Figure 1). Simulations in Section 6.1 show that fractional dosing policies would dramatically reduce mortality and infections if efficacy is at the levels associated with the current data on immunogenicity, and would still substantially reduce mortality even if efficacy fell significantly, for example to the level of some existing approved vaccines. However, the evidence of immune response is limited, and the projected impacts of such policies are subject to model uncertainty and may vary across vaccines, variants, and epidemic settings.

The high potential benefits, combined with remaining uncertainty, make information on the impact of fractional dosing very valuable. However, despite promising clinical trial data on immunogenicity of different doses of COVID-19 vaccines being available since autumn of 2020, we are aware of only one ongoing immunogenicity trial testing fractional doses: a study of half doses of mRNA-1273 (Moderna) with 200 participants in Belgium (ClinicalTrials.gov ID NCT04852861), and one real-world evaluation of the impact of fractional dosing on disease burden, recently announced in Brazil (Governo ES, 2021), which will evaluate immunogenicity and effectiveness of half a dose of ChAdOx1 nCoV-19 (Oxford/AstraZeneca). For much of this section we discuss practical options for evaluating fractional dosing policies (which also apply to other dose stretching policies), most of which could be implemented in a few weeks.
Given recent work in establishing correlates of protection against SARS-CoV-2 infection, further immunogenicity trials could provide very valuable information safely and at low cost. While clinical trial data on a few candidates already exists, much of this is with small sample sizes, and many more doses could be tested. New immunogenicity trials can be conducted in a matter of weeks and require only a few hundred participants. Since the outcome measured is not infection, there is no need to conduct trials in places where the risk of COVID-19 is high—in fact, they can be conducted in places where the virus is not circulating and which have strong health care systems, to minimize risk to participants. Clinical pharmacology modeling, routinely used during vaccine development, should also be used at this stage, allowing us to combine data from multiple trials and to extrapolate a dose-response function to new doses and across age groups.

Beyond immunogenicity trials, non-randomized approaches can also be used to establish effectiveness of different dosing strategies against infection or mortality as a vaccination campaign is rolled out. Observational designs, such as test-negative case-control studies, are recommended by the WHO to study vaccine effectiveness, and the same methods can be used to quickly determine the effect of fractional dosing in real-world settings. However, unlike immunogenicity studies, they need to be conducted where infections are occurring.\textsuperscript{29} Model-based synthesis of immunogenicity data, pre-existing clinical trial data (on full doses), and new data on fractional doses can also be used to extrapolate outcomes between doses, to safely generate estimates for immunogenicity in high risk individuals, and also to increase statistical power.\textsuperscript{30}

\textsuperscript{29}Both practical and modeling assessments (in influenza) suggest that the risk of bias of these study designs is acceptable (Shi \textit{et al.} 2017). The number of COVID-19 cases required to detect (with a precision of 10\%) 70\% efficacy with a test-negative design is 1,345, assuming that vaccine coverage in the underlying population is 50\%. This is according to the WHO’s sample size calculator, included in its official guidance document. The number would likely be lower under a non-inferiority design needed for testing fractional doses. Matching techniques can also be used to correct for bias in effectiveness studies (Dagan \textit{et al.} 2021).

\textsuperscript{30}It is important to note that ChAdOx1 nCoV-19 (Oxford/AstraZeneca) was approved for emergency use in all adults in the UK and Europe despite (at the time) lack of clear evidence on efficacy in the elderly. Extrapolating
Testing, however, can take time. Even a rapid study of immune response requires, in the case of 2-dose vaccines, time for administration of two doses (for fractional doses, a minimum of 3 to 4 weeks apart, depending on the vaccine; for FDF, several months), and at least several days for immune responses to develop following a second dose. During a period of epidemic growth, the delay associated with running experiments could be costly in terms of infections and deaths, and more lives could be saved in expectation by implementing whichever policy has higher expected value and simultaneously evaluating outcomes. Thus, many countries decided to delay second doses for some vaccines early on, without waiting for more complete data. In the case of the UK, the decision was made only weeks into the vaccination campaign. Similarly, data on fractional dose efficacy was not yet available when the WHO’s Strategic Advisory Group of Experts on Immunization (SAGE) approved the use of fractional doses of yellow fever vaccine during epidemics (World Health Organization 2016).

In some cases, randomized trials that explicitly test the impact of dose stretching on infection and disease risk can also be part of roll-out plans (Kominers and Tabarrok 2020; Bach 2020). Roll-out of modified plans need not wait until trial results are known. Instead, plans can be adjusted as new information emerges. The vaccinated can be comprehensively monitored, allowing decision-makers to modify policy based on infection rates and immune response data.

An important feature of the dose stretching policies we have analyzed is that they are, to a large degree, reversible. In the case of FDF, if it is found to be ineffective, the dosing interval can be shortened. In the case of fractional dosing, reversibility can be understood as the ability to increase dosing or provide later booster shots of the same or a different vaccine. The costs of reversing response across ages for a smaller dose of an already widely used vaccine with comprehensive real-world data is much less risky than for a new vaccine.
fractional dosing may be higher than those of reversing FDF, since it may require additional vaccine supply and use of healthcare staff to deliver vaccines. Additionally, if not effective, either policy will lead to a higher number of infections and deaths in the period until it is reversed.

Even if policymakers are pessimistic about the odds that fractional dosing will be effective, testing the policy is likely to save lives in expectation as long as there is even a relatively small chance of success. Suppose there is a 75% chance that half-dose efficacy is 30% and a 25% chance that it is 90%. If the policy is reversible, trying fractional dosing with a limited population to learn whether efficacy was truly 30% or 90% would save lives in expectation. If the trial was done in a healthy volunteer population in a region with access to good care, health risks would be minimal, and tiny relative to the potentially enormous gains from rapidly increasing vaccination rates with fractional dosing, as long as efficacy could be learned quickly enough.

Ultimately, regulators and public health officials in each jurisdiction have to decide whether to wait for additional immunogenicity studies before testing efficacy of fractional doses, and, if they do test fractional dosing, whether to do so as part of vaccine roll-outs or through formal trials.

8. Conclusion

The COVID-19 vaccination rate in most countries is limited by scarce vaccine supplies. We analyze policies that stretch scarce supplies, speeding protection to more people. In particular, we explore the impacts of three dose stretching policies: fractional dosing; delaying second doses; and using available vaccines early rather than waiting for more efficacious vaccines. Delayed second doses has been tested as part of vaccination roll-outs in some countries, with indications of substantial benefits for at least some vaccine-variant combinations. However, fractional dosing has
not been tested yet, despite high immunogenicity from smaller doses of vaccines observed in early stage clinical trials. For fractional doses, we use the high correlation (across vaccines) between neutralizing antibody response and efficacy against disease to provide suggestive evidence that half or even quarter doses of some vaccines generate immune responses associated with high efficacy. Next, we used an SEIR model to explore the impact of the three dose stretching approaches on disease burden, and show that all three substantially reduce infections and mortality over a wide range of plausible efficacy levels. These results derive from the value of speed in vaccinations — both starting vaccinations early and vaccinating at high daily rates. The value is so high that under supply constraints, the gains from enhanced speed enabled by dose stretching outstrip the costs of potential efficacy loss for a wide range of efficacies for stretched doses.

Our simulations used averted burden of infection and death as outcome variables, abstracting from other important economic and social costs. Faster vaccination could substantially reduce the time until relaxation of lockdowns and other costly non-pharmaceutical interventions. Fractional dosing, in particular, can also bring long-term benefits, even if supply constraints are eventually alleviated, such as a decrease in production costs for vaccine manufacturers and lower prices for consumers. Additionally, the trials and testing required to evaluate and potentially implement fractional dosing could also be helpful in the continued development of newer vaccine platforms, such as viral vector and mRNA vaccines, since little is currently known about optimal dosing for such vaccines. Our simulations thus likely underestimate the wider benefits of accelerating vaccination with dose-stretching policies.

To accommodate uncertainty about the efficacy of untested alternative dosing policies, our analytical approach was to simulate a range of potential values for the efficacy loss entailed by the policy. In practice, policymakers may not know which of these efficacy values applies to their
situation. However, it is possible to learn more safely, quickly, and inexpensively: the reversibility of the dose-stretching policies strengthens the case for trying them. If efficacy falls too much, the standard dose policy can be reinstated; if not, the dose-stretching policy can be retained and expanded.

Dose stretching may be particularly beneficial for low-income countries (LICs), not necessarily through their use in LICs themselves but dose-stretching in other countries could free up supplies to start widespread vaccination in LICs earlier. LICs have signed fewer bilateral deals with vaccine manufacturers and have received small allocations from COVAX and other international arrangements. If the present capacity is used without stretching measures, it may take years before LICs reach 70% coverage of COVID-19 vaccines (Castillo et al. 2021). Estimating the benefit to LICs of having other countries undertake stretching measures is beyond the scope of this paper. Instead, we discuss dose stretching in LICs in Appendix C, which repeats the main simulations, parametrized for LICs.

Dose stretching may also be particularly beneficial for non-elderly populations, who often have stronger immune responses to vaccines. Decision-makers might consider rolling out dose stretching policies only for younger adults, the majority of whom remain unvaccinated by the end of May 2021, including in many high income countries. For COVID-19 vaccines, this may also reduce side effects (including mild to moderate ones that can still be a barrier to uptake), which have generally been higher in non-elderly populations.

Our SEIR model is based on a vaccine with 95% efficacy and does not allow for partial protection (e.g., from mortality but not infection). It uses an age structure similar to those in many high income countries and a disease burden appropriate for SARS-CoV2 variants most prevalent in 2020. In further research, we will modify some of the modeling assumptions. In particular, we
will examine the case of a vaccine with efficacy in the 70% range (such as for ChAdOx1 nCoV-19 (Oxford/AstraZeneca), or mRNA vaccines against recently emerged variants), explore the case of low-income countries in more detail, allow for differences in effectiveness against different types of disease burden, and check for robustness of our conclusions to the emergence of viral variants. Further work should also investigate the effect that simultaneously implementing fractional dosing and increased delays between two doses would have on health benefits. Additionally, in future work we plan to consider longer time horizons, accommodating boosters, loss of immunity, and endogenous behavioral response to risk, such that $R = 1$ for an extended period of time (Gans 2020). We will also explicitly model a global production constraint, not only the problems of individual countries, and will analyze scenarios where it is not possible to maintain strict age prioritization. Last, we will extend the analysis to allow for the vaccination of children.

To conclude, we argue that alternative COVID-19 vaccine dosing regimens could potentially dramatically accelerate global vaccination and reduce mortality, and that these potential benefits dwarf the costs of testing these regimens. While our paper focuses on the COVID-19 context, its conclusions are broadly applicable to vaccine shortages during an epidemic.
Appendix A: Survey of Evidence on Dose Stretching

This appendix surveys existing evidence from medical literature on the effects on individual vaccine efficacy (VE) of dose stretching policies. Table 1 lists clinical trials for several important vaccines. Note that in some cases, while early trials tested multiple doses, later trials used only a single dose (such as for BNT162b2), while in others (e.g., ChAdOx1 nCoV-19 (Oxford/AstraZeneca)), multiple dose sizes were tested at all stages.

Fractional dosing

In early stage clinical trials, lower dosages of COVID-19 vaccines were often found to stimulate a strong NAb response, at least in non-elderly patients. Evidence on the immunogenicity of a range of dose sizes of each vaccine is summarized in Table 2. Note that in some later trials, such as those for JNJ-78436735 (Johnson & Johnson) and NVX-CoV2373 (Novavax), Phase 3 clinical trials proceeded with the smaller of two dose options tested in early trials after those trials found no statistically significant difference in immune response between the doses.
<table>
<thead>
<tr>
<th>Vaccine</th>
<th>Phase</th>
<th>Date posted</th>
<th>1st dose results*</th>
<th>Doses</th>
<th>Dose interval (days)</th>
<th>Treatment arms</th>
<th>Age group</th>
<th>N</th>
<th>Clinical trials.gov number</th>
</tr>
</thead>
<tbody>
<tr>
<td>ChAdOx1 nCoV-19</td>
<td>1/2</td>
<td>3/27/2020</td>
<td>Yes</td>
<td>2</td>
<td>28, 56</td>
<td>1 \times 5e10 v.p.</td>
<td>18-55</td>
<td>1090</td>
<td>NCT04324606</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 \times 5e10</td>
<td>18-65</td>
<td>2130</td>
<td>NCT04444674</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5e10, 2.2e10</td>
<td>18-70+</td>
<td>12390</td>
<td>NCT04400838</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5e10, 3.5-6.5e10</td>
<td>18+</td>
<td>10300</td>
<td>NCT04536051</td>
</tr>
</tbody>
</table>

| JNJ-78436735 | 1     | 6/18/2020   | Yes               | 1     | n/a                  | 1 \times 5e10 v.p. | 18-55     | 25  | NCT04436276             |
|             | 1/2a  | 6/18/2020   | Yes               | 2     | 56                   | 1 \times 5e10 v.p. | 18-55, 65+| 1085 | NCT04436276             |
|             |       |             |                   |       |                      | 2 \times 5e10    | 18+       | 44325 | NCT04505722             |

| mRNA-1273   | 1     | 2/25/2020   | No                | 2     | 28                   | 2 \times 25µg    | 18+       | 120  | NCT04283461             |
|             | 2a    | 5/28/2020   | No                | 2     | 28                   | 2 \times 50µg    | 18+       | 660  | NCT04405076             |

| NVX-CoV2373** | 1/2  | 4/30/2020   | Yes               | 2     | 21                   | 2 \times 5µg     | 18-59/18-84| 131/1500 | NCT04368988         |
|             |      |             |                   |       |                      | 2 \times 25      | 18-65-85  | 195   | NCT04368728             |

| BNT162b2    | 1     | 4/20/2020   | No                | 2     | 21                   | 2 \times 10µg    | 18-55, 65-85| 43548 | NCT04368728             |
|             | 2/3   | 4/20/2020   | No                | 2     | 21                   | 2 \times 20      | 12-15, 16-55, 55+| 195   | NCT04368728             |

Table 1: *Trial reports immunogenicity at least three weeks after dose administration, before a second dose (if planned) has been administered, and has a comparable outcome (in terms of age group, dose, and day measured) for second doses. **Treatment arms also included groups in which each vaccine dose arm was administered with and without 50µg of Matrix-M adjuvant.
<table>
<thead>
<tr>
<th>Vaccine</th>
<th>nAb assay</th>
<th>Day of 2nd dose</th>
<th>Age group</th>
<th>N</th>
<th>Measured on day</th>
<th>Standard dose*</th>
<th>Dose fraction</th>
<th>NAb response</th>
<th>Standard dose NAb response</th>
<th>NAb response as fraction of standard dose</th>
<th>Difference is significant (at 5% level)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>ChAdOx1 nCoV-19</td>
<td>MN&lt;sub&gt;80&lt;/sub&gt;</td>
<td>14</td>
<td>18-55</td>
<td>41</td>
<td>42</td>
<td>5e10 v.p.</td>
<td>0.44</td>
<td>161</td>
<td>193</td>
<td>0.83</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>56-69</td>
<td>28</td>
<td>42</td>
<td>2.2e10 v.p.</td>
<td>0.44</td>
<td>143</td>
<td>144</td>
<td>0.99</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>70+</td>
<td>34</td>
<td>42</td>
<td></td>
<td>0.44</td>
<td>150</td>
<td>161</td>
<td>0.93</td>
<td>no</td>
</tr>
<tr>
<td>BBV152</td>
<td>PRNT&lt;sub&gt;50&lt;/sub&gt;</td>
<td>28</td>
<td>12-65</td>
<td>190</td>
<td>56</td>
<td>6µg</td>
<td>0.5</td>
<td>100</td>
<td>197</td>
<td>0.51</td>
<td>yes</td>
</tr>
<tr>
<td>JNJ-78436735</td>
<td>PRNT IC&lt;sub&gt;50&lt;/sub&gt;</td>
<td>n/a</td>
<td>18-55</td>
<td>24</td>
<td>56</td>
<td>5e10 v.p.</td>
<td>2</td>
<td>310</td>
<td>288</td>
<td>1.08</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>65+</td>
<td>50</td>
<td>28</td>
<td>1e11 v.p.</td>
<td>2</td>
<td>212</td>
<td>277</td>
<td>0.77</td>
<td>no</td>
</tr>
<tr>
<td>mRNA-1273</td>
<td>PRNT&lt;sub&gt;80&lt;/sub&gt;</td>
<td>28</td>
<td>18-55</td>
<td>15</td>
<td>42</td>
<td>25µg</td>
<td>0.25</td>
<td>340</td>
<td>640</td>
<td>0.52</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>MN&lt;sub&gt;50&lt;/sub&gt;</td>
<td></td>
<td>18-55</td>
<td>78</td>
<td>42</td>
<td>50µg</td>
<td>0.5</td>
<td>1733</td>
<td>1909</td>
<td>0.91</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>55+</td>
<td>63</td>
<td>42</td>
<td>50µg</td>
<td>0.5</td>
<td>1827</td>
<td>1686</td>
<td>1.08</td>
<td>no</td>
</tr>
<tr>
<td>NVX-CoV2373**</td>
<td>MN IC&lt;sub&gt;99&lt;/sub&gt;</td>
<td>21</td>
<td>18-59</td>
<td>50</td>
<td>35</td>
<td>5µg</td>
<td>5</td>
<td>3305</td>
<td>3906</td>
<td>0.85</td>
<td>no</td>
</tr>
<tr>
<td>BNT162b2***</td>
<td>PRNT&lt;sub&gt;50&lt;/sub&gt;</td>
<td>21</td>
<td>18-55</td>
<td>12</td>
<td>28</td>
<td>10µg</td>
<td>0.33</td>
<td>157</td>
<td>361</td>
<td>0.43</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>65-85</td>
<td>12</td>
<td>35</td>
<td>10µg</td>
<td>0.33</td>
<td>111</td>
<td>206</td>
<td>0.54</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18-55</td>
<td>12</td>
<td>28</td>
<td>20µg</td>
<td>0.67</td>
<td>363</td>
<td>361</td>
<td>1.01</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>65-85</td>
<td>12</td>
<td>28</td>
<td>20µg</td>
<td>0.67</td>
<td>84</td>
<td>206</td>
<td>0.41</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Table 2: Neutralizing antibody (nAb) responses listed are the peak levels recorded in published trial data. Sources: Ramasamy et al. 2020 (ChAdOx1 nCoV-19, Oxford/AstraZeneca), Ella et al. 2021 (BBV152, Covaxin), Sadoff et al. 2021a (JNJ-78436735, Johnson & Johnson), Jackson et al. 2021 (mRNA-1273, Moderna - 25µg), Chu et al. 2021 (mRNA-1273, Moderna - 50µg), Keech et al. 2020 (NVX-CoV2373, Novavax), Walsh et al. 2020 (BNT162b2, Pfizer). *Standard dose* is what is being used in vaccine roll-outs thus far (v.p.=viral particles). **Both doses were administered together with 50µg of adjuvant. ***Sample size was too small to determine significance.
First Doses First

There is already some evidence both from clinical trials and real-world data on the efficacy of first doses. Efficacy varies by vaccine, by dose size, and by the virus variant. In some cases, first doses have been found to be highly effective.

- Efficacy of ChAdOx1 nCoV-19 after two doses was 62-90 percent (Voysey et al. 2021). Efficacy at preventing symptomatic infection was 73 percent starting 22 days after the first dose (UK Joint Committee on Vaccination and Immunisation 2020a). A recent clinical trial demonstrated that delaying the second dose of ChAdOx1 nCoV-19 (Oxford/AstraZeneca) to 12 weeks or later improved long-term vaccine efficacy (Voysey, et al. 2021). Hence the authors concluded, “ChAdOx1 nCoV-19 vaccination programs aimed at vaccinating a large proportion of the population with a single dose, with a second dose given after a 3 month period is an effective strategy for reducing disease, and may be the optimal for rollout of a pandemic vaccine when supplies are limited in the short term.”

- In Phase 3 clinical trials, VE of BNT162b2 (Pfizer) between 15 and 28 days after the first dose (before immune response from the second dose, delivered at 14 days, would have kicked in) was 90% (CITE). 31

- In a trial that administered first and second doses of mRNA-1273 (Moderna) four weeks apart, VE two weeks after the first dose was 95.2% (Baden et al. 2020).

While we have limited information on how long these immune responses last, we do know that the second doses of mRNA-1273 (Moderna) were given as late as 42 days after the first shot (Infectious Disease Society of America 2021), that the UK’s Joint Committee on Vaccination and Immunisation’s analysis suggests that protection from ChAdOx1 nCoV-19 (Oxford/AstraZeneca) lasts up to 3 months (JCVI 2020a), and that immunity from natural infection seems to last at least 6 months.

Most vaccine schedules offer some flexibility on when booster shots may be taken, and typically require a minimum interval of over four weeks between shots.

Real-world data on ongoing vaccinations has produced results similar to those found in clinical trials.

- In Israel, BNT162b2 (Pfizer) was estimated to be 78% effective against hospitalizations and 80% effective against disease 21-27 days following the first dose (Dagan et al. 2021).

- A study of data covering the entire population of Scotland found that BNT162b2 and ChAdOx1 nCoV-19 (Oxford/AstraZeneca) reduced hospitalization after four weeks by 85 and 94 percent, respectively, in real-time vaccination campaigns (University of Edinburgh 2021).

- A study of older adults in the UK found that a single dose of BNT162b2 or ChAdOx1 nCoV-19 was 80% effective at preventing hospitalization (Lopez Bernal et al. 2021a).

31 One trial found that first dose efficacy was only 54%, but most infections in the treatment group occurred in the first two weeks, before the immune response is likely to have kicked in.
• In the UK, a study of healthcare workers who received the first dose of BNT162b2 found that it was 75% effective at preventing positive cases and reduced asymptomatic infections by half (Hall et al. 2021).

However, first doses might not be as effective for all vaccines. For example, the first dose of CoronaVac was found to be only 3% effective in a study of Chile’s vaccine rollout (although trials have also found CoronaVac to be less effective than other vaccines even after two doses) (Dyer 2021). It is possible that the coronavirus outbreak in Chile in April was amplified by CoronaVac’s low first-dose efficacy, which made up 93% of administered doses (over a third of adults had received at least the first dose of the vaccine at the beginning of the month) (Taylor 2021).

First doses have also been less effective against new variants of the COVID-19 virus.  

• First doses of BNT162b2 (Pfizer) were only 17% effective against the South African variant, B.1.351 (Burn-Murdoch et al. 2021).

• Data from the UK suggest that first doses of BNT162b2 and ChAdOx1 nCoV-19 (Oxford/AstraZeneca) are only 51% effective against the B.1.17 variant (Lopez Bernal et al. 2021b).

• New data from India, aggregated across BNT162b2 and ChAdOx1 nCoV-19, suggests first dose efficacy against the B.1.617.2 variant is just 33% (Lopez Bernal et al. 2021b).

Immune escape risk

One serious concern about modified vaccination approaches is that they might lead to weak immune responses and immune escape through mutation (Wadman 2021). In considering such risks it is important to also consider the risks of the status quo. Without a modified vaccination approach there is a higher probability of more infections. A non-vaccinated person who becomes infected goes through a period of “partial immunity” and thus also increases the risk of immune escape. Indeed, variants of the SARS-CoV-2 virus are already circulating that are more transmissible and might be less vulnerable to vaccines but these arose before widespread vaccination (Kupferschmidt 2021). Thus, it isn’t clear whether the balance of probabilities on immune escape favors or dis-favors the modified approach. The risk of immune escape should also be evaluated in the larger context of vaccine approval. In June 2020, the FDA was willing to accept a vaccine with an efficacy rate as low as 50% (U.S. Food and Drug Administration 2020a).
70%\(^{37}\) is very good. If the efficacy rate of a dosage regime, such as offering a second dose of an mRNA vaccine at 12 weeks never falls below 70%, then the dosing regime should be considered no more risky than the approvable vaccine. The calculus should also take into account the fact that we have some indication that vaccines may be protective against asymptomatic infection even with one dose (e.g. (Voysey et al. 2021)) and that milder and less symptomatic infections lead to less transmission (Kampen et al. 2020).

The New and Emerging Respiratory Virus Threats Advisory Group (NERVTAG), a scientific advisory group to the British government, recently considered the risk of immune escape and concluded that although the risk of immune escape from delaying a second dose is real it is likely small, especially in comparison to other sources of immune escape such as therapeutics and natural infection. Moreover, the risk is outweighed by the measurable benefits of getting more does out quickly.

It is not currently possible to quantify the probability of emergence of vaccine resistance as a result of the delayed second dose, but it is likely to be small. The UK currently has more than 1,000 COVID-19 related deaths each day and has limited supplies of vaccine. In the current UK circumstances the unquantifiable but likely small probability of the delayed second dose generating a vaccine escape mutant must be weighed against the measurable benefits of doubling the speed with which the most vulnerable can be given vaccine-induced protection.

...a single dose of vaccine does not generate a new/novel risk. Given what we have observed recently with the variants B.1.1.7 and B1.351, it is a realistic possibility that over time immune escape variants will emerge, most likely driven by increasing population immunity following natural infection.

---

\(^{37}\)A recent study (Bartsch et al. 2020) suggested that a vaccine of efficacy 70 percent could prevent a COVID-19 epidemic from taking off, and one of 80 percent could extinguish an epidemic without any other measures such as social distancing.
Appendix B: Extensions of Epidemiological Model

Simulating several policies required extending the basic epidemiological model in (1)–(7). This appendix presents the systems of differential equations behind those extensions.

B1: Model with Two Doses of a Single Vaccine

The analysis of a first-doses-first (FDF) policy in Section 6.2 requires extending the basic model to allow one or two doses to have differing efficacy and to connect the delay between doses to the rate of vaccination. The model maintain the assumption that a single vaccine candidate is available. In this subsection, let subscripts 1 and 2, respectively indicate variables associated with the first and second doses. The following differential equations govern the extended model:

\[ S_i(t) = -\lambda_i(t)S_i(t) - v_i(t)\delta_1S_i(t) \]
\[ E_i(t) = \lambda_i(t)[S_i(t) + N_{1i}(t) + N_{2i}(t)] - \gamma E_i(t) \]
\[ I_i(t) = \gamma E_i(t) - \gamma'I_i(t) \]
\[ D_i(t) = p_i\gamma'I_i(t) \]
\[ R_i(t) = (1 - p_i)\gamma'I_i(t) - v_i(t)\delta_1R_i(t) \]
\[ P_{1i}(t) = v_i(t)\delta_1[e_1S_i(t) + R_i(t)] - v_i(t)\delta_2P_{1i}(t) \]
\[ P_{2i}(t) = v_i(t)\delta_2[e_2P_{1i}(t) + N_{1i}(t)] \]
\[ N_{1i}(t) = v_i(t)\delta_1(1 - e_1)S_i(t) - v_i(t)\delta_2N_{1i}(t) - \lambda_i(t)N_{1i}(t) \]
\[ N_{2i}(t) = v_i(t)\delta_2(1 - e_2)[P_{1i}(t) + N_{1i}(t)] - \lambda_i(t)N_{2i}(t). \]

B2: Model with Two Vaccines

The basic model must be extended if the planner is allowed to switch to the more effective vaccine when it becomes available because the first vaccine continues to provide protection to the successfully immunized when the new vaccine is introduced, governed by different parameters. In this subsection, we reinterpret subscript 1 as indicating variables associated with the less effective vaccine available earlier and 2 as indicating prime variables associated with the more effective vaccine available later. Compartment flows follow these differential equations:

\[ \dot{S}_i(t) = -\lambda_i(t)S_i(t) - [v_{1i}(t)\delta_{1i} + v_{2i}(t)\delta_{2i}]S_i(t) \]
\[ \dot{E}_i(t) = \lambda_i(t)[S_i(t) + N_{1i}(t) + N_{2i}(t)] - \gamma E_i(t) \]
\[ \dot{I}_i(t) = \gamma E_i(t) - \gamma'I_i(t) \]
\[ \dot{D}_i(t) = p_i\gamma'I_i(t) \]
\[ \dot{R}_i(t) = (1 - p_i)\gamma'I_i(t) - [v_{1i}(t)\delta_{1i} + v_{2i}(t)\delta_{2i}]R_i(t) \]
\[ \dot{P}_{1i}(t) = v_{1i}(t)\delta_{1i}[e_1S_i(t) + R_i(t)] \]
\[ \dot{P}_{2i}(t) = v_{2i}(t)\delta_{2i}[e_2S_i(t) + \tilde{R}_i(t)] \]
\[ \dot{N}_{1i}(t) = v_{1i}(t)\delta_{1i}(1 - e_1)S_i(t) - \lambda_i(t)N_{2i}(t) \]
\[ \dot{N}_{2i}(t) = v_{2i}(t)\delta_{2i}(1 - e_2)S_i(t) - \lambda_i(t)N_{1i}(t). \]
Appendix C: Low-Income Countries

The simulations so far have been calibrated based on data from a subset of countries with the most extensive data, typically high-income country (HIC) parameters. These differ in several respects from parameters for low-income countries. LICs generally have much younger populations. Additionally, mortality risk rises less steeply with age in LICs than HICs, close to a two-fold increase per decade (Demombynes 2020) compared to the three-fold increase per decade in HICs. Other factors also differ, including the virus variants that are most frequently found in both areas and the behavioral aspect of the population (for example, for a significant portion of the population in LICs, self-isolation might be infeasible). Patterns may also be changing over time, as exemplified by India’s terrible recent (April 2021) wave of infections. Our current analysis abstracts from this and should be considered preliminary.

![Supply constraint - HIC - LIC Epidemic scenario - Slow-decrease - Slow-growth - Fast-growth](image)

**Figure 12: Comparing Averted Burden from Faster Vaccination in LICs to HICs.** Vertical axis shows percentage reduction in burdens relative to no vaccination, in infections (left panel) and deaths (right panel), at various vaccination speeds (horizontal axis) and epidemic scenarios (colors). The plot is analogous to Figure 4 but with low-income country added (dashed lines) to high-income country case (solid lines).

Figure 12 reprises the results for burden averted by faster vaccination for LICs that were provided for HICs in Figure 4. The broad pattern of results is similar for LICs and HICs. For both sets of countries, faster vaccination has a large and nonlinear effect on the proportion of disease burden without a vaccine averted. At high vaccination rates (1–2% per day), then burden reductions in LICs are very similar to HICs. At very low vaccination rates, the percentage of burden averted is substantially larger in LICs than HICs. For example, at a vaccination rate of 0.1% of
the population per day under the slow-growth scenario, 19% of cases are averted in HICs but 54% in LICs. The difference is likely due to population structure. Only 5% of the LIC population is above age 60 compared to 25% of the population in HICs. An increase in a very low vaccination rate allows LICs to quickly move beyond their senior populations to younger cohorts with higher transmission rates. The quantitative results depend on our assumption, in the absence of good data on the contact matrix for LICs, that social mixing is the same in LICs as HICs.

We reprized all of the previous HIC results for LICs but for brevity present just the results on the relative benefits of fractional dosing, in Figure 17 in Appendix D. The relative benefit of a given fractional-dosing policy shown for LICs is slightly lower than that for HICs in Figure 6, but the magnitudes are similar.

**Appendix D: Supplementary Exhibits**
Table 3: Parameters Specified in Epidemiological Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Interpretation</th>
<th>Value</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>Length of each simulation (in days)</td>
<td>365</td>
<td></td>
</tr>
<tr>
<td>$n_i$</td>
<td>Age distribution over different age groups: high income and low income countries</td>
<td>[0,10): 0.11, 0.29 [10,20): 0.11, 0.23 [20,30): 0.13, 0.17 [30,40): 0.14, 0.12 [40,50): 0.14, 0.08 [50,60): 0.13, 0.05 [60,70): 0.11, 0.03 [70,80): 0.08, 0.02 [80, $\infty$): 0.05, 0.004</td>
<td>World Population Prospects 2019 distribution for high and low income country categories, based on R package wpp2019.</td>
</tr>
<tr>
<td>$c(i,j)$</td>
<td>Contacts from age cohort $i$ to age cohort $j$</td>
<td>9x9 matrix based on large scale survey of age-structured contacts</td>
<td>Mossong et al. (2008). Dominant eigenvalue is $\gamma = 14.17$.</td>
</tr>
<tr>
<td>$q$</td>
<td>Probability of transmission on contact</td>
<td>0.0203 for $R=0.99$, 0.02264 for $R=1.1$, 0.0411 for $R=2$</td>
<td>Adjusted to match desired $R_0$.</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Reciprocal of average length of incubation period</td>
<td>$1/5 = 0.2$</td>
<td>Incubation period was initially estimated to be about five days (Verity et al. 2020; Cevik et al. 2021). Some have suggested a longer period, but we opted for five days to account for shorter serial interval for COVID-19 (five to six days), as we do not model pre-symptomatic transmission.</td>
</tr>
<tr>
<td>$\gamma'$</td>
<td>Reciprocal of average length of infectious period</td>
<td>$1/5 = 0.2$</td>
<td>While duration of viral shedding is longer, infectiousness decreases quickly after onset of symptoms (Cevik et al. 2021).</td>
</tr>
<tr>
<td>$p_i$</td>
<td>Age-specific infection fatality rate (IFR)</td>
<td>[0,10): 2e-05 [10,20): 6e-05 [20,30): 3e-04 [30,40): 8e-04 [40,50): 0.0015 [50,60): 0.006 [60,70): 0.022 [70,80): 0.051 [80, $\infty$): 0.093</td>
<td>Based on a recent meta-analysis of IFRs (Manheim et al. 2020).</td>
</tr>
</tbody>
</table>
Figure 13: Age-specific Evolution of Vaccinated, Infected, and Susceptible Compartments. Colors indicate compartment percentages for various age cohorts. Assumes slow-growth epidemic scenario and baseline values of parameters (95% efficacy, vaccination rate of 0.25% of population per day, target 80% vaccination rate in cohorts age 20 and over). The assumed policy of prioritizing older cohorts is reflected in vaccination panel. Because they are vaccinated earlier, older cohorts experience an earlier downturn from their peak infection rate.
Figure 14: Threshold Speed-up at which Fractional Dominates Full Dosing. Entries present threshold speed-up in the vaccination rate when moving from fractional to full dosing such that fractional dosing averts greater disease burden. Shading intensity increases with the needed threshold. Entries marked “Inf,” denoting infinity, are cases in which full dosing dominates for all speed-ups simulated, even as high as 32. Reciprocal of entry can be interpreted as the size of fractional dose, holding constant efficacy on vertical axis, below which fractional dominates full dosing. Each tile represents a different combination of epidemic growth, level of efficacy of fractional dose, and vaccination rate assumed for the full dose. Maintains baseline assumption that full dose is 95% effective.
Figure 15: Burden Averted Using Delayed Vaccine Relative to Switching Policy. Entries are ratios of burden averted using vaccine 2, a 95% effective vaccine available after delay $t_2$ indicated on horizontal axis, to that from starting with vaccine 1, with efficacy $e_1$ indicated on the vertical axis, immediately and switching to vaccine 2 at $t_2$. Entries greater than 1 favor vaccine 2 and less than 1 favor the switching policy. The policies are no different when there is no delay, both equivalent to using the more-effective vaccine only.
Figure 16: Burden Averted Relative to No Vaccination under Various Delays in Second Dose. Results related to the same policies studied in Figure 9, but here burden that each policy averts is compared to no vaccination, while Figure 9 compares SDF to the optimal alternative. SDF refers to second doses first, the status-quo policy of delaying four weeks until the second dose; FDF refers to first doses first, delaying 12 weeks until the second dose; and HDF-60 refers to a hybrid policy of delaying four weeks for age 60 and over and 12 weeks for under 60.
Figure 17: Comparing Full to Fractional Dosing in LICs. Repeats simulation results from Figure 6 on the relative burden of infections (top panel) and deaths (bottom panel) under fractional dosing vs full dose but uses LIC rather than HIC parameters for the size ($n_i$) and mortality rate ($p_i$) in age cohorts. Values lower than 1 favour fractional dose. Contact matrix and other parameters are the same as in the HIC case. See Figure 6 for additional notes.
References


Demombynes, Gabriel. (2020) “COVID-19 Age-Mortality Curves are Flatter in Developing Countries.” World Bank working paper no. 9313 [Online only]. Available at http://hdl.handle.net/10986/34028.


Goldberg, Yair et al. (2021) “Protection of previous SARS-CoV-2 infection is similar to that of BNT162b2 vaccine protection: A three-month nationwide experience from Israel,” medRxiv Preprint [Online only]. Available at https://www.medrxiv.org/content/early/2021/04/24/2021.04.20.21255670.


Lopez Bernal, Jamie, et al. (2021a) “COVID-19 Vaccine Coverage in Health-Care Workers in England and Effectiveness of BNT162b2 mRNA Vaccine Against Infection (SIREN): A Prospective, Multicentre, Cohort Study.” Lancet, April 23 [Online only]. doi: 10.1016/S0140-6736(21)00790-X.


Parry, Helen Marie et al. “Single Vaccination with BNT162b2 or ChAdOx1 in Older People Induces Equivalent Antibody Generation but Enhanced Cellular Responses after ChAdOx1.” SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, April 13, 2021.


