Decisive Conditions for Strategic Vaccination against SARS-CoV-2

Lucas Böttcher[1,*] and Jan Nagler[2]

[1] *Dept. of Computational Medicine, University of California, Los Angeles, CA 90095-1766, United States of America*

[2] *Deep Dynamics Group, Centre for Human and Machine Intelligence, Frankfurt School of Finance & Management, Frankfurt, Germany*[†]

(Dated: March 5, 2021)

While vaccines that protect against SARS-CoV-2 are being approved, the number of available doses remains limited. As it may take months until the production of vaccines can meet the actual demand, the majority of available SARS-CoV-2 vaccines elicits strong immune responses when administered as prime-boost regimens. Since the immunological response to the first ("prime") injection may provide already a substantial degree of protection against infection, it may be more effective—under certain conditions—to vaccinate as many people as possible with only one shot, instead of administering a person a second ("boost") shot. Such a strategic vaccination campaign may help to more effectively slow down the spread of SARS-CoV-2, reduce hospitalizations, and reduce deaths. Yet, the conditions which make single-dose vaccination favourable over prime-boost administrations are not well understood. Here, we formulate a model that helps explore these decisive conditions as a function of the various time scales and epidemiological mechanisms at work. We study how these conditions arise from disease prevalence, vaccination rates, basic reproduction number, prime and prime-boost efficacies, prime-boost intervals, and waning rates. By combining epidemiological modeling, random sampling techniques, and decision tree learning, we find that prime-first vaccination is robustly favored over prime-boost vaccination campaigns, even for high vaccination rates, high disease prevalence, and a relatively low single-dose efficacy.

INTRODUCTION

After the initial identification of the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in Wuhan, China in December 2019, the virus quickly reached pandemic proportions and caused major public health and economic problems worldwide [1]. The disease associated with SARS-CoV-2 infections was termed coronavirus disease 2019 (COVID-19). As of March 1, 2021, the number of confirmed COVID-19 cases exceeded 115 million and more than 2.5 million COVID-19 deaths in more than 219 countries were reported [2]. Large differences between excess deaths and reported COVID-19 deaths across different countries suggest that the actual death toll associated with COVID-19 is even higher [3].

With the start of vaccination campaigns against SARS-CoV-2 in many countries [4], millions of people will receive partial immunization in the next months. The mRNA vaccines BNT162b2 (BioNTech-Pfizer) and mRNA-1273 (Moderna) received emergency use approval in the US and EU. When administered as prime-boost regimen, these vaccines have a reported protective efficacy of 95% [5] and 94.1% [6], respectively. A recent effectiveness evaluation of the BNT162b2 BioNTech-Pfizer vaccine shows that it may offer about 50% protection against SARS-CoV-2 infections about 2–3 weeks after receiving the first shot [7]. The adenovirus-based vaccine ChAdOx1 (Oxford-AstraZeneca) is being used in the UK with a reported single-shot regimen efficacy of 62% [8].

In addition to these vaccines, there are another 8 vaccines that received emergency use authorization in other countries including China, India, and Russia [9]. The majority of currently available SARS-CoV-2 vaccines elicits strong immune responses when administered as prime-boost regimens. Yet, given the current distribution and production constraints, it may take months until the production of COVID-19 vaccines can meet the actual global demand. Similar to vaccination campaigns in previous disease outbreaks, it may therefore be a favourable alternative to administer a single vaccination dose to twice as many people. In 2016, a single-dose vaccination campaign against cholera was implemented in Zambia because of the insufficient number of vaccination doses that were available to complete a standard two-dose campaign [10]. Other vaccines, like the oral cholera vaccines that require two doses, are highly effective after a single dose but their protection is short lived compared with the full dosage [11, 12].

Despite the clear advantages of single-dose vaccination campaigns, such as faster immunization of a larger number of people and lower vaccine-distribution infrastructure requirements and costs, different vaccination doses and protocols affect the level of vaccination-induced humoral and cell immunity. Recent clinical trial results [13] on the COVID-19

* lucasb@ucla.edu

† jan.nagler@gmail.com

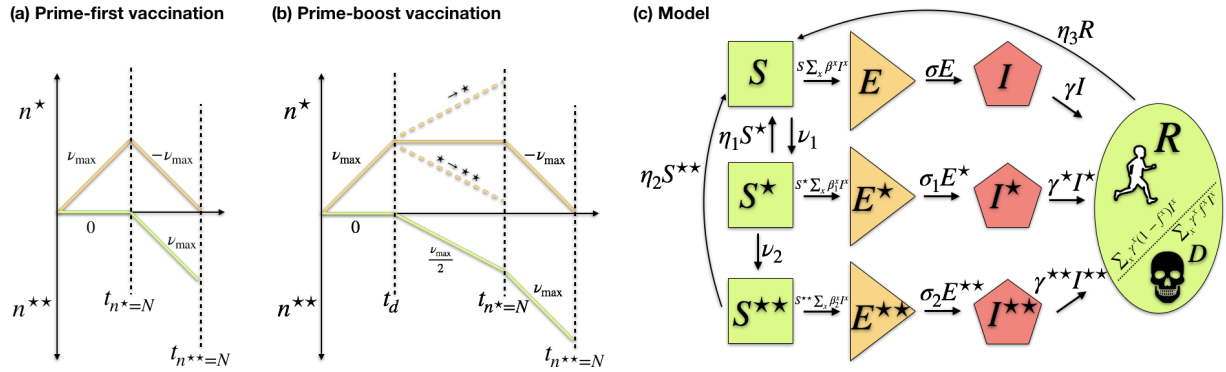

FIG. 1. **Schematic of campaigns and model.** (a,b) Evolution of the number of individuals that received prime ($n^\star$) and prime-boost ($n^{\star\star}$) shots in a population of size N. (a) Prime-first ($\star$): Prime shots are administered with maximal rate as long as unvaccinated susceptible individuals exist. Then, boost shots are given to prime-vaccinated individuals. (b) Prime-boost ($\star\star$) balances prime and boost shots equally. Boosting starts after t_d days. (c) In our model, susceptible individuals (S, $S^\star$, $S^{\star\star}$) become exposed (E, $E^\star$, $E^{\star\star}$) at rates $\beta, \beta^\star, \beta^{\star\star}$ and transition to an infected state (I, $I^\star$, $I^{\star\star}$) at rates $\sigma, \sigma_1, \sigma_2$. Infectious individuals either recover (R) or die (D) at rates $\sum_x (1-f^x)\gamma^x$ and $\sum_x f^x \gamma^x$, respectively. Prime vaccination doses ($\star$) are administered to susceptible individuals at rate ν_1 and prime-vaccinated individuals receive a boost shot ($\star\star$) at rate ν_2. Transitions from $S^\star$, $S^{\star\star}$, and R to S ("waning immunity") occur at rates ν_1, ν_2, and ν_3, respectively. Notation: $\sum_x \beta^x I^x = \beta I + \beta^\star I^\star + \beta^{\star\star} I^{\star\star}$.

vaccine BNT162b1 show that for prime-boost protocols antibody concentrations are about 5 to 20 times larger than those observed in patients who only received a single vaccination dose, highlighting the need for boosting. The vaccination-induced CD4$^+$ and CD8$^+$ T cell responses were also found to be significantly reduced if no boost shot was administered, indicating that boost doses are important for T-cell-mediated immunity against SARS-CoV-2. For type-1 inactivated poliovirus vaccine (IPV), clinical trial results [14] suggest that the level of neutralizing antibodies does not increase without a boost injection. Yet, for type-2 and 3 IPV, the first vaccination dose already elicits a neutralizing antibody response.

Here we ask when prime-boost vaccination is favorable over prime vaccination. As many countries are facing limited vaccination supply and fearing the increasingly wide-spread emergence of faster-spreading SARS-CoV-2 virus variants and the risk of collapsing health care systems [15], related critical issues are being controversially debated in the US [16], UK [17], and Germany [18]. The current controversy around prime and prime-boost vaccination strategies raises two connected questions, which we address in this paper: How do shortages in vaccine supplies and uncertainties in epidemiological parameters alter the possible advantage of prime-first over prime-boost vaccination? And how do possible differences in vaccine efficacy and loss of vaccine-induced immunity affect the decision boundary separating prime-first and prime-boost vaccination regimes in high-dimensional parameter space? By combining methods from epidemiological modeling, statistical mechanics, and decision tree learning, we explore position, extent, and sensitivity of the decision boundary and provide a characterization of criteria, sufficiently simple and immediately accessible to decision makers.

RESULTS

Prime-first versus prime-boost vaccination

Different vaccination campaigns may lead to different proportions of infected, recovered, and deceased individuals at a given time. We study the differences between prime-first and prime-boost by accounting for a vaccination-induced reduction in transmissibility in a susceptible-exposed-infected-recovered-deceased (SEIRD)-based model [19] (see Materials and Methods and Fig. 1). To quantify the effect vaccination protocols have on the overall disease-induced fatality we focus on two fatality measures. Let d_1 (prime-first) and d_2 (prime-boost) be the maximum (daily) changes in the total number of deaths within the time horizon $[0, T]$. As a measure of the relative difference between d_1 and d_2, we use the relative fatality change (RFC-δ),

$$\delta(d_1, d_2) = \frac{d_2 - d_1}{\max(d_1, d_2)} \,. \tag{1}$$

As a long-term measure, we study the relative change in the cumulative number of deaths (RFC-Δ),

$$\Delta(D_1, D_2) = \frac{D_2 - D_1}{\max(D_1, D_2)}\,, \tag{2}$$

defined within the same time horizon as RFC-δ.

For both measures (1) and (2), a positive sign indicates more fatalities for prime-boost vaccination than for prime-first, while a negative sign indicates to favor prime-boost over prime-first campaigns.

Current vaccination campaigns prioritize health care workers and vulnerable groups (e.g., elderly people with comorbidities) with a high risk of infection, leading to variations in vaccination rates. Further heterogeneity in model parameters may arise from infection rates that differ between age groups because of different degrees of susceptibility to infection [20] and different mobility characteristics. Our model accounts for these variations in epidemiological parameters through a large degree of parameterization. Nine different infection rates describe contacts between (susceptible and infectious) unvaccinated, prime-first vaccinated, and prime-boost vaccinated individuals. This large degree of parameterization accounts for possible correlations between age-group, transmissbility, and mobility. We therefore choose not to incorporate demographic compartmentalization [21] such as age-stratification in our model.

Vaccination-campaign-preference diagrams

To provide mechanistic insight into the population-level differences between prime and prime-boost vaccination campaigns, we study how RFC-δ and RFC-Δ are impacted by epidemiological parameters and epidemic state. As a function of two parameters, green domains as shown in Fig. 2 indicate excess deaths for prime-boost, while prime-boost is favored in red regions. The parameter ranges follow existing literature, or are chosen sufficiently broad to cover uncertainties. Empirical data [22] suggests an estimated range of the basic reproduction number $R_0 \in [1, 4]$. Differences in the waning rates η_1 and η_2 are not known at the present time, while clinical trials are still ongoing. Thus, we sample a broad parameter range, $\eta_1 - \eta_2 \in [10^{-4}, 10^{-1}]$ day^{-1} with $\eta_2 = 3 \times 10^{-3}$ day^{-1}, which includes waning time-scales that were reported earlier for SARS-CoV [23]. For the initial infection disease prevalence, we assume the range $I(0) \in [10^{-4}, 10^{-1}]$. This range includes up to 10% infected individuals but may lie outside current estimates of some places of very high prevalence such as Manaus, Brazil [24], and earlier estimates from New York City, USA [25]. For the range of the maximum vaccination rate $\nu_{\max}$ we use $[0, 10^{-1}]$ day^{-1} , which we inferred from current vaccination-campaign data [4].

We assume that the transmission rates β_1 and β_2 are proportional to the vaccine efficacies after single-dose and prime-boost vaccination, respectively. Thus, we identify the relative efficacy for prime-first immunization (RE) with the ratio β_2/β_1. Values close to one are favorable for prime-first campaigns, while a low RE disfavors prime-first.

In order to analyze the effect of RE on the effectiveness of prime and prime-boost vaccination campaigns, we study $\beta_2/\beta_1 \in [10^{-4}, 1]$ day^{-1}. We choose this rather broad range to account for the lack of reliable data, in particular regarding new variants of SARS-CoV-2 and possible adverse effects in vaccine protection [26]. Parameters that are held constant in our simulations are listed in Tab. I (see Materials and Methods).

The vaccination-campaign-preference diagrams (Fig. 2) suggest that prime vaccination campaigns are associated with a smaller death toll compared to prime-boost campaigns for a wide range of R_0, maximum vaccination rates, epidemic states, and relative efficacy ratios (green-shaded regions in Fig. 2).

As the main result of our study, we identify a two-parameter threshold combination that separates vaccination-campaign preferences (black dashed lines in Fig. 2). For a sufficiently small waning-rate difference $\eta_1 - \eta_2 \lesssim 0.02$ and a sufficiently low maximum vaccination rate $\nu_{\max} \lesssim 0.02$, we observe that prime-first vaccination outperforms prime-boost vaccination in all projections where parameters are held constant as specified in Tab. I. In the projections involving $\eta_2 - \eta_1$, prime-boost preference is observed if immunity wanes significantly faster for prime-vaccinated individuals than for prime-boost vaccinated individuals.

All projections in Fig. 2 combined suggest that prime-boost vaccination should only be favored for $\nu_{\max} \gtrsim 0.02$ day^{-1}, which largely exceeds SARS-CoV-2 immunization rates worldwide [4].

How a relatively low single-dose efficacy affects the preference for each campaign is shown in [Fig. 2(K–M)]. In Fig. 2(K,L) we assume a transmission reduction of only 10% after single-dose immunization, $\beta_1 = 0.9\beta$, together with a 40% reduction in fatality, $f^\star = 0.6 \times 10^{-2} = 0.6f$, and all other parameters as in Tab. I. This "low single-dose efficacy" domain is characterized by the occurrence of additional prime-boost preference regions in parameter space [red-shaded regions in Fig. 2(K,L)]. Yet, even if the fatality rates of prime-first and prime-boost deviate substantially, $f^{\star\star}/f^\star \lesssim 0.8$, only for low values of the relative prime-first efficacy RE $= \beta_2/\beta_1 \lesssim 0.1$, preference for prime-boost is observed [Fig. 2(K), shown range $0 \leq \beta_2/\beta_1 \leq 0.1$]. Given this range and current data in SARS-CoV-2 [7], the diagram suggests preference for prime-first. Regarding the waning rate difference, $\eta_1 - \eta_2$, a low single dose-efficacy does not suggest a threshold lower than 0.017 for prime-first preference [Fig. 2(L)]. Figure 2(M) shows the dependence of RFC-δ

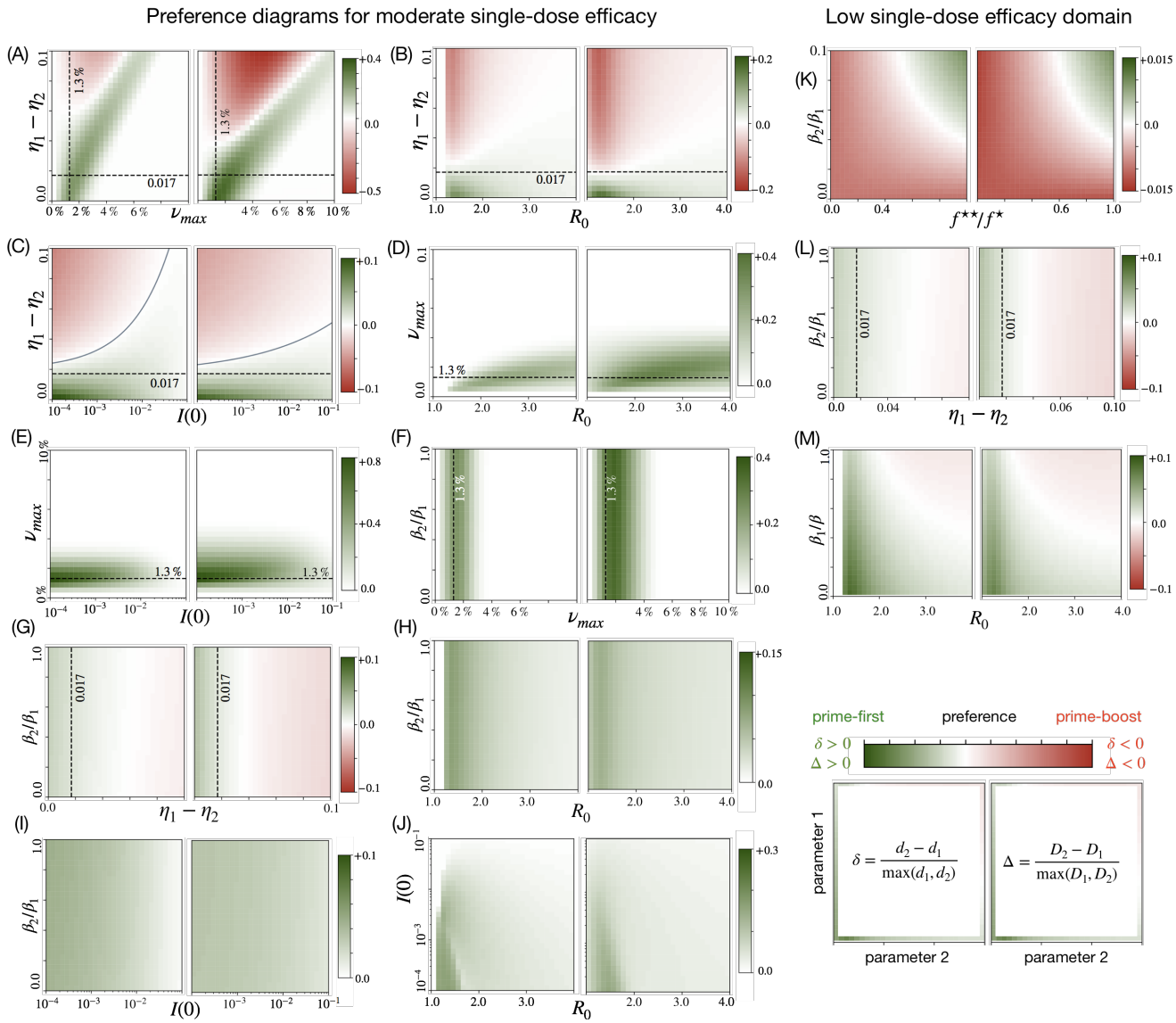

FIG. 2. **Vaccination-campaign-preference diagrams.** For combinations of basic reproduction number R_0, waning rate difference $\eta_1 - \eta_2$, initial disease prevalence $I(0)$, maximum vaccination rate $\nu_{\max}$, and relative efficacy for prime-first immunization (RE), β_2/β_1, we plot RFC-δ [Eq. (1)], and RFC-Δ [Eq. (2)]. Green-shaded regions indicate preference for prime (RFC-$\delta > 0$, RFC-$\Delta > 0$), red-shaded regions indicate preference for prime-boost (RFC-$\delta < 0$, RFC-$\Delta < 0$). (A-L): Parameter domain as in Tab. I, assuming a *moderate* single-dose efficacy, i.e. $\beta_1 = \beta/2$, $f^\star = f/10$. (K-M): *Low* single-dose efficacy domain: we set $\beta_1 = 0.9\beta$, $f^\star = 0.6 \times 10^{-2} = 0.6f$, and all remaining parameters as in Tab. I. For (M) we varied β_1/β and β (hence R_0) and set all other parameters according to Tab. I. The ratios $\beta_2/\beta_1 = \beta_2^\star/\beta_1^\star = \beta_2^{\star\star}/\beta_1^{\star\star} = 1/5$ are also as in Tab. I. Dashed lines: Decisive threshold $\eta_1 - \eta_2 = 0.017$, and Israel's vaccination rate as of Feb. 1, 2021 [4] ($\nu_{\max} = 0.013$). Solid line in (C): decision boundary as guide to the eye between $y = \eta_1 - \eta_2$ and $x = I(0)$ as given by the following nonlinear relations: $y = 0.4x^{1/2} + 0.02$ (RFC-δ), and $y = 0.06x^{1/4} + 0.017$ (RFC-Δ).

and RFC-Δ on β_1/β and R_0. Values of $\beta_1/\beta \approx 1$ indicate a very low single-shot efficacy, whereas $\beta_1/\beta \approx 0$ indicates an unrealistic high efficacy. For large R_0 and very low single-shot efficacies, $\beta_1/\beta \gtrsim 0.8$, prime-boost is preferred over prime-first [Fig. 2(M)]. This combination, however, is rather unanticipated [7].

Finally, the waning-rate threshold below which preference for prime-first is observed depends weakly on the initial infection prevalence. The preference threshold of the waning rate difference slightly increases as $I(0)$ decreases, $\eta_1 - \eta_2 \lesssim 0.010 - 0.017$ for $I(0) = 10^{-5} - 10^{-2}$, see Materials and Methods).

The presented campaign preference diagrams are only two-dimensional projections of a 25-dimensional parameter space, with the majority of parameters kept arbitrarily fixed (Tab. I). Hence, we examine next whether the preference

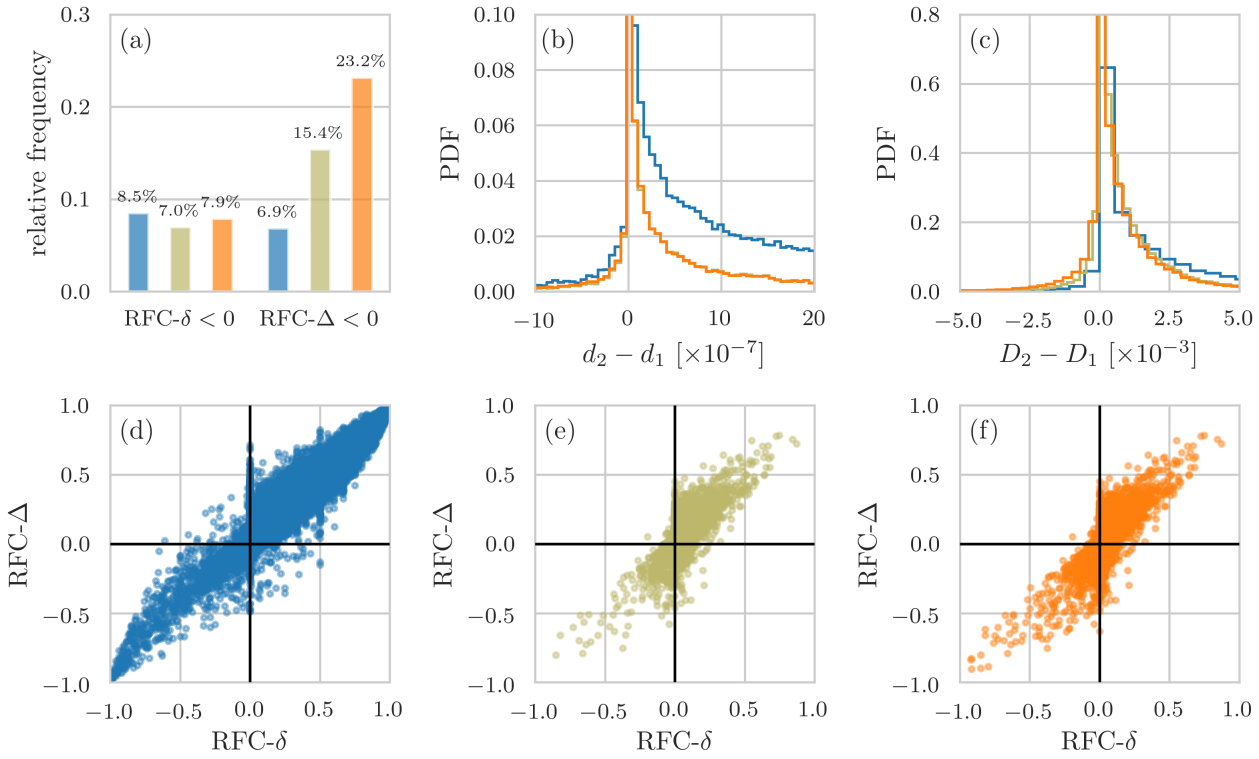

FIG. 3. **Monte Carlo sampling of entire and restricted high-dimensional parameter space.** Unconditioned data (orange) is compared with two conditioned, (i) data shown in beige: $\nu_{\max} \leq 0.047$ and $\eta_1 - \eta_2 \leq 0.056$, and (ii) data shown in blue: $\nu_{\max} \leq 0.013$ and $\eta_1 - \eta_2 \leq 0.017$. Thresholds (i) are suggested by our decision tree analysis (Materials and Methods). (a) Relative frequency of prime-boost-preference samples (RFC-$\delta < 0$, RFC-$\Delta < 0$) for the three datasets. Error bars are below 0.2% (not shown). (b) Probability density function (PDF) of the difference between death rates d_2 (prime-boost) and d_1 (prime). For the conditioned data, averages of $d_1 - d_2$ are about 3×10^{-6} (blue curve) and about 5×10^{-7} (beige curve), larger than the mean 4×10^{-7} of the unconditioned data. (c) PDF of the difference between the total number of deaths D_2 (prime-boost) and D_1 (prime). For the conditioned data, the means of $D_1 - D_2$ are about 2×10^{-3} (blue curve) and about 9×10^{-4} (beige curve), larger than the mean 7×10^{-4} of the unconditioned data. (d–f) Correlation plots of RFC-δ and RFC-Δ for all datasets. The presented data are based on 5×10^4 (blue and orange curves) and about 4.4×10^4 (beige curves) samples of the entire 25-dimensional parameter space.

for prime-first vaccination is supported by other independent methods.

High-dimensional parameter space Monte Carlo sampling

Thus far, our results suggest a pronounced preference for prime-first vaccination for a wide range of key epidemiological parameters. To further substantiate this conclusion, we performed Monte Carlo sampling of the entire 25-dimensional parameter space (see Materials and Methods). For the analyzed high-dimensional parameter space, our results support that prime-boost-preference occurs significantly less frequently than samples indicating an advantage of prime-first.

The relative frequencies of samples for which prime-boost vaccination outperforms prime-first vaccination, characterized by RFC-$\delta < 0$ and RFC-$\Delta < 0$, are estimated as 7.9% [standard error (SE): 0.2%] and 23.2% (SE: 0.4%), see orange bars in Fig. 3(a). For waning rate differences $\eta_1 - \eta_2 \leq 0.056$ day^{-1} and vaccination rates $\nu_{\max} \leq 0.047$ day^{-1}, we find that the proportions of prime-boost-preference samples are 6.9% (SE: 0.2%) for RFC-$\delta < 0$ and 15.4% (SE: 0.3%) for RFC-$\Delta < 0$ [beige bars in Fig. 3(a)]. Further restricting the parameter space using the condition $\eta_1 - \eta_2 < 0.017$ day^{-1} (black dashed lines in Fig. 2) and currently reported vaccination rates $\nu_{\max} < 0.013$ day^{-1} [4] leads to proportions of prime-boost-preference samples of 8.5% (SE: 0.2%) for RFC-$\delta < 0$ and 6.9% (SE: 0.2%) for RFC-$\Delta < 0$ [blue bars in Fig. 3(a)].

This means that constraining the studied parameter space by lowering $\nu_{\max}$ and $\eta_1 - \eta_2$ results in a substantially enhanced preference for prime-first in terms of reduced excess deaths, RFC$-\Delta$. In contrast, we find that the proportion

of prime-boost preference samples with RFC-$\delta < 0$ is almost unaffected by the chosen parameter restrictions, which is indicated by the observed narrow rage between 7 and 9% [Fig. 3(a)]. This supports the robustness of our results. Independent of the threshold combination, for randomly sampled parameters, prime-first is robustly and strongly preferred regarding the maximum daily rate of change of disease-induced deaths, RFC-$\delta > 0$. In addition, domination of prime-first preference is observed in the projections for two-parameter combinations with RFC-$\delta < 0$ and RFC-$\Delta < 0$ (Fig. 2).

RFC-δ and RFC-Δ are complementary fatality measures but are correlated [Fig. 3(d–f), $R^2 = 0.94$ (d), $R^2 = 0.68$ (e), and $R^2 = 0.68$ (f); corresponding p-values are smaller than machine precision]. The high correlation observed for the threshold combination $\nu_{\max} \leq 0.013$ and $\eta_1 - \eta_2 \leq 0.017$ confirms its discriminative power and the robustness of our results regarding the choice of both fatality measures.

Decision tree learning

As another independent method for determining decisive conditions for strategic vaccination campaigns, we performed binary decision tree learning with repeated stratified cross validation [27, 28]. This technique has proven useful to extract the most discriminative features in high-dimensional data. Our analysis suggests that $\nu_{\max}$ and $\eta_1 - \eta_2$ are the most discriminative parameters within the 25-dimensional parameter space in terms of their ability to separate prime-first and prime-boost-preference regimes (see Materials and Methods). For the samples that we generated according to the distributions listed in Tab. I (orange lines and markers in Fig. 3), we obtain accuracy scores and recalls above 70% for $\nu_{\max} \leq 0.047$ and $\eta_1 - \eta_2 \leq 0.056$ as a highly-discriminative prime-first preference threshold combination. Additionally constraining the parameter space with the thresholds that we used in the previous paragraph (beige and blue lines and markers in Fig. 3) results in prime-first preference for 93% of the parameter space volume.

DISCUSSION

Effective vaccination protocols are crucial to achieve a high immunization coverage, especially if vaccination supplies are limited. The ongoing debate on the most effective way of distributing prime-boost regimens against SARS-CoV-2 has been sparked by arguments suggesting that single-dose vaccination protocols may be more effective than immediate prime-boost administration given the current supply shortages [16–18, 29, 30]. By combining epidemiological modeling, methods from statistical mechanics, and decision tree learning, we have studied the effect of relevant epidemiological parameters (e.g., vaccine efficacy and immunity waning) on a possible advantage of prime-first over prime-boost vaccination. We have identified and studied decision boundaries separating the parameter regimes in which one or the other vaccination protocol is preferable. Our results suggest that prime-first campaigns are associated with a lower death toll compared to prime-boost vaccination campaigns, even for relatively high vaccination rates and low single-dose efficacies.

Our methodology and findings complement those of other recent works, such as that of Matrajt *et al.* [29], who compare single-dose and prime-boost vaccination campaigns against SARS-CoV-2, yet without accounting for immunity waning. They report that single-dose vaccination campaigns make optimal use of resources in the short term, given a sufficiently large single-dose efficacy that they identify as the main discriminative factor. In contrast, our study calls attention to immunity waning together with the vaccination rate as the highly-discriminative factors, while we find that vaccine efficacies are less discriminative.

Matrajt *et al.* [31] also argue that due to the complexity of the model and limitations from available data, a recommendation can only be given, once the precision in epidemiological parameters becomes sufficiently high. A similar assessment is reported in [11] for Avian influenza A (H5N1 & H7N9). These studies emphasize to prioritize data collection such as further field efficacy data [7].

Saad-Roy *et al.* [32] focus on the long-term effects of waning and evolutionary immune response in a highly-parameterized model. Their results underscore the importance of studying viral phylodynamics, from within host to global scales. Certain scenarios they analyze suggest that single-dose campaigns may be favorable for some time scales but not for others, depending on a combination of parameters, waning rates included. Based on our analyses, we provide a robust preference criterion based on only two highly-discriminative parameters, vaccination rate and waning rate difference.

Preference for prime-first vaccination is not unexpected. For the initial inter-dose interval time, both vaccination strategies are identical since, regardless of the chosen strategy, booster jabs are not yet administered. In the subsequent time interval twice as many susceptible individuals can be immunized with a prime-first protocol compared to prime-boost vaccination. This means, about 50% of individuals who could have received a shot will actually remain unvaccinated. Let us refer to this unvaccinated group as *group A* and denote with *group B* those that receive both

shots in the prime-boost campaign. One can assume that the infection rates of individuals in group A are larger than those of individuals in group B, who benefit from a more effective immune response. As a result, higher transmission in group A is the expected dominating differential adverse effect. As expected for effective prime-boost vaccines, one may assume that the prime-boost infection rate, β_2, and the fatality rate, $f^{\star\star}$, are significantly lower than their counterparts for prime-first. Thus, the effective transmission rate for group A and B combined is dominated by group A's rate but not critically dependent on β_2, which intuitively explains why the relative efficacy ratio $\mathrm{RE} = \beta_2/\beta_1$ is not a highly discriminative factor.

For very low single-shot efficacies, or very high single-shot disease-induced fatality rates, the single-dose efficacy β_1/β, the relative prime-first efficacy RE, and R_0, may be discriminative, depending on the circumstances. Current data on SARS-CoV-2 vaccination campaigns [7], however, suggest that those parameter combinations are unlikely to occur. Furthermore, if immunity wanes substantially faster after the first shot than after the additional booster jab, prime-boost vaccination may become favorable over prime-first, depending on R_0. Unvaccinated and susceptible individuals should also receive both vaccination shots if a few percent of a jurisdiction's total population can be vaccinated daily. However, even for the relatively large vaccination rate of about $\sim 1\%$ per day, as realized in Israel [4], our analyses suggest that prime-first vaccination is still favorable over prime-boost campaigns.

In summary, our results align with existing literature [11, 29, 31, 32] in the sense that reliable epidemiological data are required to be collected to identify most effective vaccination protocols. However, our analysis suggests that even for a large degree of uncertainty in key epidemiological data, prime-first vaccination is robustly preferred over prime-boost vaccination—if the waning rate difference is sufficiently small.

Although clinical studies of the approved SARS-CoV-2 vaccines may strongly suggest that these vaccines are safe and effective, only little is known about their possible long-term adverse effects [33]. Clearly, for the comparison of different vaccination strategies we assume that negative long-term effects are negligible. Adverse effects and different levels of protection can be incorporated in existing models accounting for different subpopulations (e.g., young and old) [34]. Since our results are independent of the actual fatality ratio for unvaccinated individuals, age-group stratification does not show discriminative power. For predicting non-discriminative features such as campaign-specific death tolls, it may however be useful to account for different age groups and heterogeneous contact patterns.

To conclude, while current vaccine supplies are not keeping up with demand, especially in low and middle income countries, and newly-emerging variants of SARS-CoV-2 may reduce the effectiveness of currently available vaccines [15], it is desirable to provide decision makers with transparent tools that supports them in assessing different vaccination protocols. The selection criteria that we found in this study are a step in this direction and may be of immediate help to healthcare officials and decision makers. More generally, our framework establishes how vaccine dosing strategies [35, 36] can be integrated into effective pandemic control plans.

ACKNOWLEDGEMENTS

LB acknowledges financial support from the SNF (P2EZP2_191888). Parts of the simulations were performed on the ETH Euler cluster.

DATA AVAILABILITY

Our source codes are publicly available at `https://github.com/lubo93/vaccination`.

---

[1] World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. `https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020` (2020). Accessed: 2020-04-18.

[2] COVID-19 statistics. `https://www.worldometers.info/coronavirus/` (2020). Accessed: 2021-03-01.

[3] Böttcher, L., D'Orsogna, M. & Chou, T. Using excess deaths and testing statistics to improve estimates of COVID-19 mortalities. *medRxiv:2021.01.10.21249524,arXiv:2101.03467* (2021).

[4] Coronavirus (COVID-19) Vaccinations (2020, last accessed January 4, 2021). URL `https://ourworldindata.org/covid-vaccinations`.

[5] Polack, F. P. *et al.* Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *New England Journal of Medicine* (2020).

[6] Vaccines and Related Biological Products Advisory Committee Meeting, FDA Briefing Document, Moderna COVID-19 Vaccine. `https://www.fda.gov/media/144434/download/` (2020). Accessed: 2021-01-13.

[7] Chodcik, G. *et al.* The effectiveness of the first dose of BNT162b2 vaccine in reducing SARS-CoV-2 infection 13-24 days after immunization: real-world evidence. *Medrxiv* (2021).

[8] Voysey, M. *et al.* Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *The Lancet* (2020).

[9] Coronavirus Vaccine Tracker. `https://www.nytimes.com/interactive/2020/science/coronavirus-vaccine-tracker.html` (2021). Accessed: 2021-03-05.

[10] Ferreras, E. *et al.* Single-dose cholera vaccine in response to an outbreak in Zambia. *New England Journal of Medicine* **378**, 577–579 (2018).

[11] Azman, A. S. *et al.* The impact of a one-dose versus two-dose oral cholera vaccine regimen in outbreak settings: a modeling study. *PLoS Medicine* **12**, e1001867 (2015).

[12] Qadri, F. *et al.* Efficacy of a single-dose regimen of inactivated whole-cell oral cholera vaccine: results from 2 years of follow-up of a randomised trial. *The Lancet Infectious Diseases* **18**, 666–674 (2018).

[13] Sahin, U. *et al.* COVID-19 vaccine BNT162b1 elicits human antibody and TH 1 T cell responses. *Nature* **586**, 594–599 (2020).

[14] Tzeng, S. Y. *et al.* Stabilized single-injection inactivated polio vaccine elicits a strong neutralizing immune response. *Proceedings of the National Academy of Sciences* **115**, E5269–E5278 (2018).

[15] CDC. New COVID-19 Variants (2021, last accessed January 15, 2021). URL `https://www.cdc.gov/coronavirus/2019-ncov/transmission/variant.html`.

[16] Post, T. W. Coronavirus cases among lawmakers who sheltered in lockdown show one vaccine dose may not immediately protect against infection (2021, last accessed January 15, 2021). URL `https://www.washingtonpost.com/health/2021/01/13/capitol-riot-lockdown-coronavirus-vaccine/`.

[17] Iacobucci, G. & Mahase, E. Covid-19 vaccination: What's the evidence for extending the dosing interval? (2021).

[18] Guardian, T. Germany and Denmark consider delaying second Covid vaccine dose (2021, last accessed January 15, 2021). URL `https://www.theguardian.com/world/2021/jan/04/germany-denmark-second-dose-covid-vaccine`.

[19] Keeling, M. J. & Rohani, P. *Modeling infectious diseases in humans and animals* (Princeton University Press, 2011).

[20] Davies, N. G. *et al.* Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine* **26**, 1205–1211 (2020).

[21] Aleta, A. *et al.* Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour* **4**, 964–971 (2020).

[22] Lai, C.-C., Shih, T.-P., Ko, W.-C., Tang, H.-J. & Hsueh, P.-R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and corona virus disease-2019 (COVID-19): the epidemic and the challenges. *International Journal of Antimicrobial Agents* 105924 (2020).

[23] Jeyanathan, M. *et al.* Immunological considerations for COVID-19 vaccine strategies. *Nature Reviews Immunology* **20**, 615–632 (2020).

[24] Buss, L. F. *et al.* COVID-19 herd immunity in the Brazilian Amazon. *Science* (2020).

[25] 1 in 5 New Yorkers May Have Had Covid-19, Antibody Tests Suggest. `https://www.nytimes.com/2020/04/23/nyregion/coronavirus-antibodies-test-ny.html` (2020, last accessed January 31, 2021).

[26] Callaway, E. & Mallapaty, S. Novavax offers first evidence that COVID vaccines protect people against variants. *Nature* (2021).

[27] Breiman, L., Friedman, J., Stone, C. J. & Olshen, R. A. *Classification and regression trees* (CRC press, 1984).

[28] Hastie, T., Tibshirani, R. & Friedman, J. *The elements of statistical learning: data mining, inference, and prediction* (Springer Science & Business Media, 2009).

[29] Matrajt, L. & Eaton, J. Optimizing vaccine allocation for COVID-19 vaccines: critical role of single-dose vaccination. *medRxiv* 2020–12.

[30] Lipsitch, M. & Kahn, R. Interpreting vaccine efficacy trial results for infection and transmission. *medRxiv* (2021). URL `https://www.medrxiv.org/content/early/2021/02/28/2021.02.25.21252415`. https://www.medrxiv.org/content/early/2021/02/28/2021.02.25.21252415.full.pdf.

[31] Matrajt, L., Britton, T., Halloran, M. E. & Longini Jr, I. M. One versus two doses: What is the best use of vaccine in an influenza pandemic? *Epidemics* **13**, 17–27 (2015).

[32] Saad-Roy, C. M. *et al.* Epidemiological and evolutionary considerations of SARS-CoV-2 vaccine dosing regimes. *medRxiv* (2021).

[33] de Vrieze, J. Pfizer's vaccine raises allergy concerns. *Science* (2021).

[34] Böttcher, L., Xia, M. & Chou, T. Why case fatality ratios can be misleading: individual-and population-based mortality estimates and factors influencing them. *Physical Biology* **17**, 065003 (2020).

[35] Riley, S., Wu, J. T. & Leung, G. M. Optimizing the dose of pre-pandemic influenza vaccines to reduce the infection attack rate. *PLoS Medicine* **4**, e218 (2007).

[36] Anderson, R. M., Vegvari, C., Truscott, J. & Collyer, B. S. Challenges in creating herd immunity to SARS-CoV-2 infection by mass vaccination. *The Lancet* **396**, 1614–1616 (2020).

[37] Stollmeier, F. & Nagler, J. Unfair and anomalous evolutionary dynamics from fluctuating payoffs. *Physical Review Letters* **120**, 058101 (2018).

[38] Van den Driessche, P. & Watmough, J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences* **180**, 29–48 (2002).

[39] He, X. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine* **26**, 672–675 (2020).

[40] Duration of Isolation and Precautions for Adults with COVID-19 (2020, last accessed January 4, 2021). URL `https://www.cdc.gov/coronavirus/2019-ncov/hcp/duration-isolation.html`.

[41] Salje, H. *et al.* Estimating the burden of SARS-CoV-2 in France. *Science* (2020).

[42] Dormand, J. R. & Prince, P. J. A family of embedded Runge-Kutta formulae. *Journal of Computational and Applied Mathematics* **6**, 19–26 (1980).

MATERIALS AND METHODS

Modeling prime and prime-boost vaccination

We adapt the SEIRD model [19] to account for immunity waning and a vaccination-induced reduction in transmissibility [Fig. 1]. The fractions of susceptible, exposed, infected, recovered, and deceased individuals at time t are denoted by $S(t)$, $E(t)$, $I(t)$, $R(t)$, and $D(t)$ respectively. Moreover, we denote the fractions of prime and prime-boost vaccinated susceptible individuals by $S^\star(t)$ and $S^{\star\star}(t)$, respectively. With rate ν_1, susceptible individuals get vaccinated with prime shots and with rate ν_2 prime-vaccinated susceptible individuals get vaccinated with boost shots. The time dependence in the vaccination rates reflects temporal variations in the availability of vaccination doses, as explained below. The corresponding fractions of vaccinated exposed and infected individuals are denoted by $E^\star(t)$ and $E^{\star\star}(t)$ and $I^\star(t)$ and $I^{\star\star}(t)$. We use three constant rates η_1, η_2, η_3 to model immunity waning (i.e., transitions from $S^\star$, $S^{\star\star}$, and R to S). Characteristic time scales of waning immunity [23], defined by the inverse of the corresponding rates, are much longer than those associated with entering and leaving exposed and infected compartments, so we do not explicitly model waning immunity in these compartments. For long time horizons, additional birth and death processes may be employed to model birth and age-related death.

The resulting dynamics of the susceptible and exposed classes is described by the following rate equations:

$$\begin{aligned}
\frac{\mathrm{d}S}{\mathrm{d}t} &= -\beta SI - \beta^\star SI^\star - \beta^{\star\star} SI^{\star\star} - \nu_1 + \eta_1 S^\star + \eta_2 S^{\star\star} + \eta_3 R\,,\\
\frac{\mathrm{d}S^\star}{\mathrm{d}t} &= \nu_1 - \beta_1 S^\star I - \beta_1^\star S^\star I^\star - \beta_1^{\star\star} S^\star I^{\star\star} - \nu_2 - \eta_1 S^\star\,,\\
\frac{\mathrm{d}S^{\star\star}}{\mathrm{d}t} &= \nu_2 - \beta_2 S^{\star\star} I - \beta_2^\star S^{\star\star} I^\star - \beta_2^{\star\star} S^{\star\star} I^{\star\star} - \eta_2 S^{\star\star}\,,\\
\frac{\mathrm{d}E}{\mathrm{d}t} &= \beta SI + \beta^\star SI^\star + \beta^{\star\star} SI^{\star\star} - \sigma E\,,\\
\frac{\mathrm{d}E^\star}{\mathrm{d}t} &= \beta_1 S^\star I + \beta_1^\star S^\star I^\star + \beta_1^{\star\star} S^\star I^{\star\star} - \sigma_1 E^\star\,,\\
\frac{\mathrm{d}E^{\star\star}}{\mathrm{d}t} &= \beta_2 S^{\star\star} I + \beta_2^\star S^{\star\star} I^\star + \beta_2^{\star\star} S^{\star\star} I^{\star\star} - \sigma_2 E^{\star\star}\,.
\end{aligned} \tag{3}$$

The maximum proportion of susceptible individuals that can be prime and prime-boost vaccinated is $S(t)$ and $S^\star(t)$, respectively. In our model, we use the vaccination rates

$$\nu_1(\mu_1, \mu_2, S, S^\star, t_d, t) = (\mu_1 + \mu_2)\, H[S(t)] H[t_\mathrm{d} - t] + \mu_1 H[S(t)] H[t - t_\mathrm{d}] \tag{4}$$

and

$$\nu_2(\mu_1, \mu_2, S, S^\star, t_d, t) = \mu_2 H[S^\star(t)] H[t - t_\mathrm{d}] + \mu_1 \left(1 - H[S(t)]\right) H[S^\star(t)] H[t - t_d]\,, \tag{5}$$

where $\mu_1 = \nu_\mathrm{max}$ and $\mu_2 = 0$ for prime-first vaccination and $\mu_1 = \mu_2 = \nu_\mathrm{max}/2$ for prime-boost vaccination. Here, $H[x]$ denotes the Heaviside step function, which is zero for $x < 0$ and one for $x \geq 0$. The function $H[t - t_\mathrm{d}]$ describes the delay t_d of about 2–3 weeks [9] between prime and boost shots. Up to time t_d, susceptible individuals get vaccinated with rate $\mu_1 + \mu_2$. If no susceptible individuals are left, prime-vaccinated individuals get vaccinated with rate μ_1 too, leading to the term $\mu_1(1 - H[S(t)])H[S^\star(t)]H[t - t_d]$ in Eq. (5). In our model, only susceptible individuals are vaccinated. This can be justified by the assumption that susceptible individuals outnumber those in other disease states.

Exposed individuals transition to infected state at rates σ, σ_1, and σ_2. The evolution of the infected, recovered, and deceased compartments is described by:

$$\begin{aligned}
\frac{\mathrm{d}I}{\mathrm{d}t} &= \sigma E - \gamma I\,,\\
\frac{\mathrm{d}I^\star}{\mathrm{d}t} &= \sigma_1 E^\star - \gamma^\star I^\star\,,\\
\frac{\mathrm{d}I^{\star\star}}{\mathrm{d}t} &= \sigma_2 E^{\star\star} - \gamma^{\star\star} I^{\star\star}\,,\\
\frac{\mathrm{d}R}{\mathrm{d}t} &= \gamma(1-f)I + \gamma^\star(1-f^\star)I^\star + \gamma^{\star\star}(1-f^{\star\star})I^{\star\star} - \eta_3 R,\,,\\
\frac{\mathrm{d}D}{\mathrm{d}t} &= \gamma f I + \gamma^\star f^\star I^\star + \gamma^{\star\star} f^{\star\star} I^{\star\star}\,.
\end{aligned} \tag{6}$$

Only 10 of equations (3) and (6) are independent since we employ the normalization condition $S + S^\star + S^{\star\star} + E + E^\star + E^{\star\star} + I + I^\star + I^{\star\star} + R + D = 1$. Different transmissibilities $\beta, \beta^\star, \beta^{\star\star}$, $\beta_1, \beta_1^\star, \beta_1^{\star\star}$, and $\beta_2, \beta_2^\star, \beta_2^{\star\star}$ describe interactions between susceptible and infected individuals with different immunity levels.

For each infected compartment I, $I^\star$, and $I^{\star\star}$, we calculate the infection fatality ratios (IFRs) [3] by dividing the associated cumulative number of deaths by the total number of infections in the unvaccinated, prime-vaccinated, and prime-boost-vaccinated compartments, respectively. The IFR of the unvaccinated pool of individuals is

$$\mathrm{IFR}(t) = \frac{\int_0^t \gamma f I(t')\,\mathrm{d}t'}{I(t) + \int_0^t \gamma f I(t')\,\mathrm{d}t' + \int_0^t \gamma(1-f) I(t')\,\mathrm{d}t'}\,. \tag{7}$$

For constant γ, f, we obtain

$$\mathrm{IFR}(t) = \frac{\gamma f \int_0^t I(t')\,\mathrm{d}t'}{I(t) + \gamma \int_0^t I(t')\,\mathrm{d}t'}\,. \tag{8}$$

As the number of infected individuals approaches zero for long time horizons (i.e., $\lim_{t\to\infty} I(t) = 0$), the IFR satisfies $\lim_{t\to\infty} \mathrm{IFR}(t) = f$. Similarly, $\lim_{t\to\infty} \mathrm{IFR}^\star(t) = f^\star$ and $\lim_{t\to\infty} \mathrm{IFR}^{\star\star}(t) = f^{\star\star}$ if $\gamma^\star, f^\star$ and $\gamma^{\star\star}, f^{\star\star}$ are time-independent.

Due to ergodicity breaking effects from multiplicative noise [37] deterministic models tend to overestimate infection and fatality. However, it is realistic to assume that the effects from noise are not discriminative as they do not differ for either vaccination campaign.

Basic reproduction number

We calculate the basic reproduction number R_0 of the epidemic model (3) and (6) using the next-generation matrix method [38]. As a first step, we rewrite the rate equations (3) and (6) of the infected compartments in matrix form

$$\dot{x}(t) = F(x,y) - V(x,y)\,, \tag{9}$$

where $x = (E, E^\star, E^{\star\star}, I, I^\star, I^{\star\star})^\top$, F represents new infections, and V describes transitions from one compartment to another. We thus find

$$F = \begin{pmatrix} 0 & 0 & 0 & \beta S & \beta^\star S & \beta^{\star\star} S \\ 0 & 0 & 0 & \beta_1 S^\star & \beta_1^\star S^\star & \beta_1^{\star\star} S^\star \\ 0 & 0 & 0 & \beta_2 S^{\star\star} & \beta_2^\star S^{\star\star} & \beta_2^{\star\star} S^{\star\star} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} \sigma & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_2 & 0 & 0 & 0 \\ -\sigma & 0 & 0 & \gamma & 0 & 0 \\ 0 & -\sigma_1 & 0 & 0 & \gamma^\star & 0 \\ 0 & 0 & -\sigma_2 & 0 & 0 & \gamma^{\star\star} \end{pmatrix}. \tag{10}$$

The basic reproduction number R_0, the expected number of infections generated by an infectious individual in an otherwise completely susceptible population, is the spectral radius of the next-generation matrix FV^{-1} [38]. Finding R_0 for the general system (10) involves the analytically cumbersome task of finding roots of a cubic equation, which can be avoided by using numerical methods (e.g., the power method). For very effective vaccines, however, one may assume that the transmissibility of prime-boost vaccinated individuals is much lower than the transmissibility of unvaccinated individuals. That is, $\beta^{\star\star} \ll \beta$, $\beta_1^{\star\star} \ll \beta$, and $\beta_2^{\star\star} \ll \beta$. In this approximation, we obtain

$$R_0 = \frac{\beta S \gamma^\star + \beta_1^\star S^\star \gamma + \sqrt{\beta^2 \gamma^{\star 2} S^2 + 2\gamma\gamma^\star S S^\star \left(2\beta_1\beta^\star - \beta\beta_1^\star\right) + \gamma^2 \beta_1^{\star 2} S^{\star 2}}}{2\gamma\gamma^\star}\,. \tag{11}$$

For $S(0) = 1$ and $S^\star(0) = 0$, the basic reproduction number is

$$R_0 = \frac{\beta}{\gamma}\,. \tag{12}$$

Numerical solution and model parameters

To solve Eqs. (3) and (6) numerically, we use the Dormand–Prince method [42] with a maximum time step of 10^{-1} and simulate the evolution of different epidemics in the time interval $[0, T]$ where $T = 300$ days. For the simulation

TABLE I. **Overview of model parameters.** The listed parameter values are used when the associated model parameters are held constant in the parameter-space plots that we show in the results section. As initial fractions of infected and susceptible individuals, we use $I(0) = 10^{-2}$ and $S(0) = 1 - I(0)$.

Parameter	Symbol	Value	Units	Comments/References
infection rates S	β, $\beta^\star$, $\beta^{\star\star}$	3/14, $\beta/10$, $\beta/20$	[day^{-1}]	β inferred from $R_0 = \beta/\gamma$ [22]
infection rates $S^\star$	β_1, $\beta_1^\star$, $\beta_1^{\star\star}$	$\beta/2$, $\beta^\star/2$, $\beta^{\star\star}/2$	[day^{-1}]	estimate
infection rates $S^{\star\star}$	β_2, $\beta_2^\star$, $\beta_2^{\star\star}$	$\beta/10$, $\beta^\star/10$, $\beta^{\star\star}/10$	[day^{-1}]	estimate
incubation rate E	σ	1/5	[day^{-1}]	[22, 39]
incubation rate $E^\star$	σ_1	1/5	[day^{-1}]	estimate
incubation rate $E^{\star\star}$	σ_2	1/5	[day^{-1}]	estimate
vaccination rate	$\nu_{\max}$	10^{-3}	[day^{-1}]	[4]
waning rate (prime)	η_1	10^{-2}	[day^{-1}]	estimate
waning rate (prime-boost)	η_2	3×10^{-3}	[day^{-1}]	estimate
waning rate (recovered)	η_3	0	[day^{-1}]	estimate
resolution rate I	γ	1/14	[day^{-1}]	[34, 40]
resolution rate $I^\star$	$\gamma^\star$	2γ	[day^{-1}]	estimate
resolution rate $I^{\star\star}$	$\gamma^{\star\star}$	4γ	[day^{-1}]	estimate
fatality ratio I	f	10^{-2}	-	[3, 41]
fatality ratio $I^\star$	$f^\star$	10^{-3}	-	estimated from reported efficacy [9]
fatality ratio $I^{\star\star}$	$f^{\star\star}$	10^{-3}	-	estimated from reported efficacy [9]
prime-boost delay	t_{d}	21	[day]	[9]

results that we show in Fig. 2, we set $I(0) = 10^{-2}$ and $S(0) = 1 - I(0)$. If model parameters are held constant in Fig. 2, we use the parameters that are listed in Tab. I. Transmissibilities of infection events that involve at least one vaccinated individual are smaller than or equal to the baseline transmissibility β as long as mobility and distancing characteristics of vaccinated individuals do not differ significantly from those who are unvaccinated. In our model, this means that $\beta \geq \beta^\star \geq \beta^{\star\star}$, $\beta \geq \beta_1 \geq \beta_2$, $\beta_1 \geq \beta_1^\star \geq \beta_1^{\star\star}$, and $\beta_2 \geq \beta_2^\star \geq \beta_2^{\star\star}$ (equality holds for very ineffective vaccines). As vaccination campaigns and vaccine effectiveness analyses are ongoing, we used estimates for $\beta^\star$, $\beta^{\star\star}$, β_1, $\beta_1^\star$, $\beta_1^{\star\star}$, β_2, $\beta_2^\star$, and $\beta_2^{\star\star}$ as reported in Tab. I. We also model the effect of small incidence rates and broader parameter ranges in a random-sampling analysis as reported in the next sections.

There are two more constraints that our model parameters have to satisfy to describe the impact of vaccination campaigns on disease transmission. First, the fatality ratio in the unvaccinated compartment is larger than the fatality ratios in the vaccinated compartments (i.e., $f \geq f^\star \geq f^{\star\star}$). We assume that differences in $f^\star$ and $f^{\star\star}$ are negligible. Second, the waning rate in the prime-boost vaccinated compartment is smaller than the waning rate in the prime-vaccinated compartment (i.e., $\eta_2 \leq \eta_1$).

Influence of small incidence rates

To study the effect of small incidence rates on the location and extent of prime-first and prime-boost preference regions, we set $I(0) = 10^{-5}$, which is three orders of magnitude smaller than the value $I(0) = 10^{-2}$ we used in Fig. 2, and show vaccination preference diagrams for $\eta_1 - \eta_2$ vs. $\nu_{\max}$ and $\eta_1 - \eta_2$ vs. R_0 in Fig. 4. We observe that a threshold $\eta_1 - \eta_2 = 0.01$ separates prime-first and prime-boost preference regions in both diagrams. This value is smaller than the threshold of $\eta_1 - \eta_2 = 0.017$, which we used in Fig. 2.

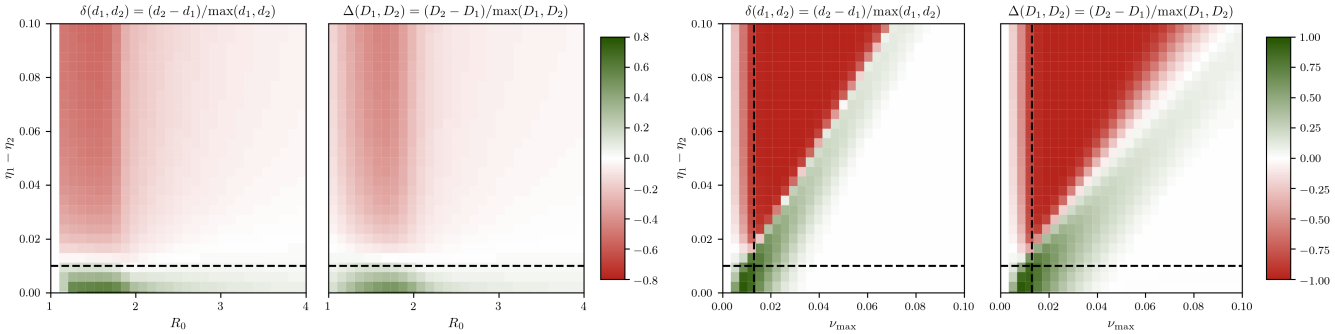

FIG. 4. **Selection of vaccination-campaign-preference diagrams for small incidence rates.** For selected combinations of basic reproduction number R_0, waning rate difference $\eta_1 - \eta_2$, and maximum vaccination rate $\nu_{\max}$, we plot the RFC-δ [Eq. (1)], and RFC-Δ [Eq. (2)]. Green-shaded regions indicate preference for prime (RFC-$\delta > 0$, RFC-$\Delta > 0$), red-shaded regions indicate preference for prime-boost (RFC-$\delta < 0$, RFC-$\Delta < 0$). In all simulations as presented here, we set $I(0) = 10^{-5}$. The remaining parameters are as in Tab. I. Dashed lines are guides to the eye: Thresholds $\eta_1 - \eta_2 = 0.01$, and Israel's vaccination rate as of Feb. 1, 2021 [4] ($\nu_{\max} = 0.013$).

Monte Carlo sampling

The parameter distributions that we use in our random sampling and decision tree analysis are summarized in Tab. II. We generate two datasets, which are further described in the next section, with $N = 50000$ samples each and analyze the influence of different combinations of model parameters and initial conditions on RFC-$\delta(d_1, d_2)$ [Eq. (1)] and RFC-$\Delta(D_1, D_2)$ [Eq. (2)].

Decision Tree Analysis

A binary decision tree consists of a root condition and branches, where the left branch refers to the "yes"-branch while the rights branch refers to the "no"-branch.

We employed binary decision tree learning with repeated stratified cross validation ($k = 5$ folds, $n = 10$ repeats). The algorithm `RepeatedStratifiedKFold` (available in the Python library scikit-learn[1]) optimizes for split purity using Gini as loss function (split criterion).

Gini impurity is a standard measure in tree learning that quantifies how often a randomly chosen sample from the training dataset would be incorrectly labeled if it was entirely randomly labeled, given the distribution of (binary) labels in the subset. In our analysis, labels are e "prime-first" and "prime-boost".

Stratified cross-validation is based on splitting the data into folds such that each fold has the same proportion of observations with a given categorical value. It is particularly useful for imbalanced datasets. Overall, here, we have more prime-first samples than prime-boost ones.

Following standard procedure, we split the dataset (randomly) into **training** and **test** datasets, 70% and 30%, respectively. Learning is performed using the training dataset while the test dataset is cross-validated.

The training accuracy score (for a binary classification task) is defined as the relative number of correctly predicted labels, that is,

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}\,, \tag{13}$$

where TP are true positives, TN are true negatives, FP are false positives, and FN are false negatives.

In our binary classification problem, positives are prime-first labeled samples, negatives are prime-boost labeled samples. Class prime-boost is defined by $\Delta D < 0$ (red-shaded regions in Fig. 2). Class prime-first is defined by $\Delta D > 0$ (green-shaded regions in Fig. 2).

An n-times repeated stratified cross-validation is based on the following iteration: (i) Shuffle the test dataset randomly, (ii) split the dataset into k folds, (iii) for each fold: take the fold as test dataset and take the remaining folds as training dataset, (iv) fit the tree on the training dataset and evaluate it on the test dataset.

---

[1] `https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.RepeatedStratifiedKFold.html`, accessed: 02-28-2021

TABLE II. **Overview of sampling distributions.** The listed parameter values and distributions are used in our random-sampling analysis. As initial fraction of susceptible individuals, we use $S(0) = 1 - I(0) - S^\star(0) - S^{\star\star}(0)$. We set $E(0) = 0$, $E^\star(0) = 0$, $E^{\star\star}(0) = 0$, $I^\star(0) = 0$, $I^{\star\star}(0) = 0$, $R(0) = 0$, $D(0) = 0$. The minimum and maximum values of β are $\beta_{\min} = \gamma$ and $\beta_{\max} = 4\gamma$, respectively. A uniform distributions with boundaries a and b is indicated by $\mathcal{U}(a, b)$.

Parameter	Symbol	Value/Distribution	Units
infection rates S	β, $\beta^\star$, $\beta^{\star\star}$	$\mathcal{U}(\beta_{\min}, \beta_{\max})$, $\mathcal{U}(0, \beta)$, $\mathcal{U}(0, \beta^\star)$	[day^{-1}]
infection rates $S^\star$	β_1, $\beta_1^\star$, $\beta_1^{\star\star}$	$\mathcal{U}(0, \beta)$, $\mathcal{U}(0, \beta_1)$, $\mathcal{U}(0, \beta_1^\star)$	[day^{-1}]
infection rates $S^{\star\star}$	β_2, $\beta_2^\star$, $\beta_2^{\star\star}$	$\mathcal{U}(0, \beta_1)$, $\mathcal{U}(0, \beta_1^\star)$, $\mathcal{U}(0, \beta_1^{\star\star})$	[day^{-1}]
incubation rate E	σ	$\mathcal{U}(0.2, 0.5)$	[day^{-1}]
incubation rate $E^\star$	σ_1	$\mathcal{U}(0.2, 0.5)$	[day^{-1}]
incubation rate $E^{\star\star}$	σ_2	$\mathcal{U}(0.2, 0.5)$	[day^{-1}]
vaccination rate	$\nu_{\max}$	$\mathcal{U}(0, 0.1)$	[day^{-1}]
waning rate (prime)	η_1	$\mathcal{U}(0, 0.1)$	[day^{-1}]
waning rate (prime-boost)	η_2	$\mathcal{U}(0, \eta_1)$	[day^{-1}]
waning rate (recovered)	η_3	0	[day^{-1}]
resolution rate I	γ	1/14	[day^{-1}]
resolution rate $I^\star$	$\gamma^\star$	$\mathcal{U}(\gamma, 2\gamma)$	[day^{-1}]
resolution rate $I^{\star\star}$	$\gamma^{\star\star}$	$\mathcal{U}(\gamma^\star, 2\gamma^\star)$	[day^{-1}]
fatality ratio I	f	$\mathcal{U}(10^{-3}, 10^{-1})$	-
fatality ratio $I^\star$	$f^\star$	$\mathcal{U}(10^{-3}, f)$	-
fatality ratio $I^{\star\star}$	$f^{\star\star}$	$f^\star$	-
prime-boost delay	t_{d}	$\mathcal{U}(7, 35)$	[day]
initially infected individuals	$I(0)$	$\mathcal{U}(10^{-4}, 3 \times 10^{-1})$	-
initially prime-vaccinated individuals	$S^\star(0)$	$\mathcal{U}(10^{-4}, 10^{-1})$	-
initially prime-boost vaccinated individuals	$S^{\star\star}(0)$	$\mathcal{U}(10^{-4}, 10^{-1})$	-

Accuracy and balanced accuracy are monitored as main cross-validation scores. We also monitored precision, F1-score based metrics, ROC AUC, and recall. Balanced accuracy is warranted for imbalanced datasets and defined as the arithmetic mean of *sensitivity*$=\frac{\mathrm{TP}}{\mathrm{TP+FN}}$ (true positive rate) and *specificity*$=\frac{\mathrm{TN}}{\mathrm{TN+FP}}$ (true negative rate).

Dataset A Dataset A comprises of 50000 randomly sampled data points for parameter ranges and disease stages as described in Tab. II. In Fig. 3 dataset A is called "unconditioned data" (displayed in orange).

Here, we analyze dataset A, see Fig. 5. Training performance is excellent and reaches 100% for large depths due to overfitting. Learning performance is satisfactory, as seen from similar behaviors for test accuracy and balanced accuracy, around 70% for depth = 3. The resulting tree reveals two highly-discriminative conditions for prime-boost preference, $\nu_{\max} \leq 0.047$ and $\eta_1 - \eta_2 \leq 0.056$. This threshold combination is used in Fig. 3, referred to as the conditioned data, displayed in beige.

Dataset B Here we study the conditioned data constrained by $\eta_1 - \eta_2 \leq 0.017$ and $\nu_{\max} < 0.013$, as analyzed in Fig. 3 (blue), here called dataset B. For this dataset, we uniformly sampled initial proportions of infected individuals, $I(0)$, on a logarithmic scale from 10^{-7} to 3×10^{-1}. This way of sampling allows us to study the robustness of decision boundary thresholds for a large range of initial disease prevalences.

Dataset B comprises of 50000 randomly sampled data points where samples simultaneously satisfy $\eta_1 - \eta_2 \leq 0.017$ and $\nu_{\max} < 0.013$. Class prime-boost is defined by $\Delta D < 0$ (red-shaded regions in Fig. 2). Class prime-first is defined by $\Delta D > 0$ (green-shaded regions in Fig. 2).

Results are presented in Fig. 6. The training accuracy curve (green curve) shows high accuracy levels from overfitting

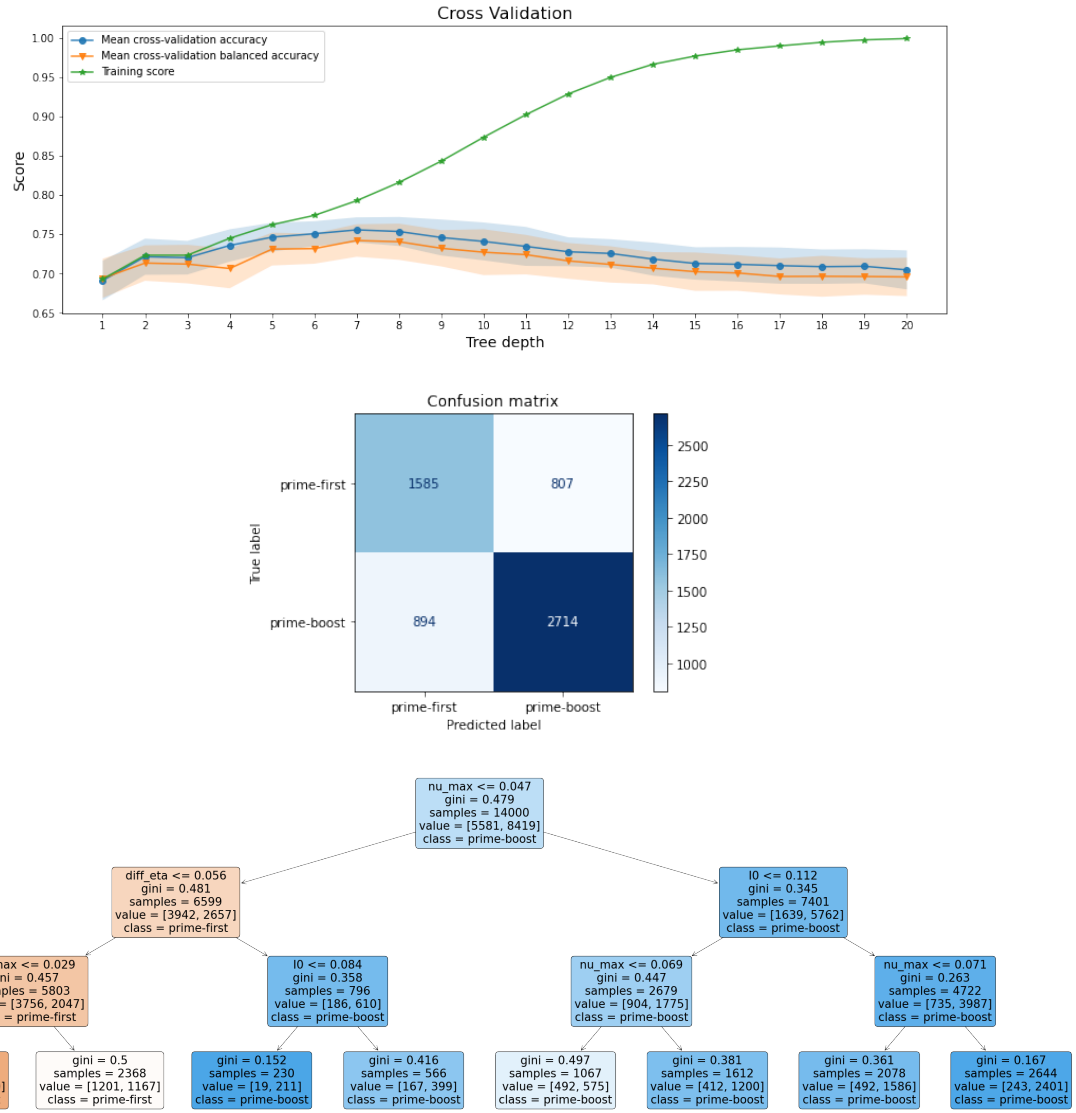

FIG. 5. **Decision tree analysis for dataset A.** Cross-validation: Training accuracy (green curve) reaches 100% for large tree size. This indicates that fitting is satisfactory, while test training data is overfitted for large tree sizes, as expected. Accuracy (blue curve and shades) and balanced accuracy (orange curve and shades) show a fair and also similar performance for all depths. This indicates satisfactory class prediction. Confusion matrix: Number of instances for predicted and true (ground truth) labels, for prime-first and prime-boost. Upper left: True positives, Upper right: False positives, Lower left: False negatives, Lower right: True negatives. Notation: Positives are prime-first, negatives prime-boost. Decision tree: (*Value*) denotes the number of samples in class "prime-first" (left part) and class "prime-boost" (right part), respectively, for the given branch (brackets), while *samples* denote the total sum of samples at the given branch. Left branches satisfy the displayed condition ("yes" branch), right descendants are "no" branches. Notation: diff_eta $= \eta_1 - \eta_2$, nu_max $= \nu_{\max}$, and $I0 = I(0)$. Shown tree: depth $= 3$.

of the **training** set that reaches 100% for large tree depths. The accuracy curve (blue) shows the mean of the cross-validation of the accuracy for the **test** dataset.

The results confirm that no conditions other than the constraints $\eta_1 - \eta_2 \leq 0.017$ and $\nu_{\max} < 0.013$ robustly characterize prime-first preference domains.

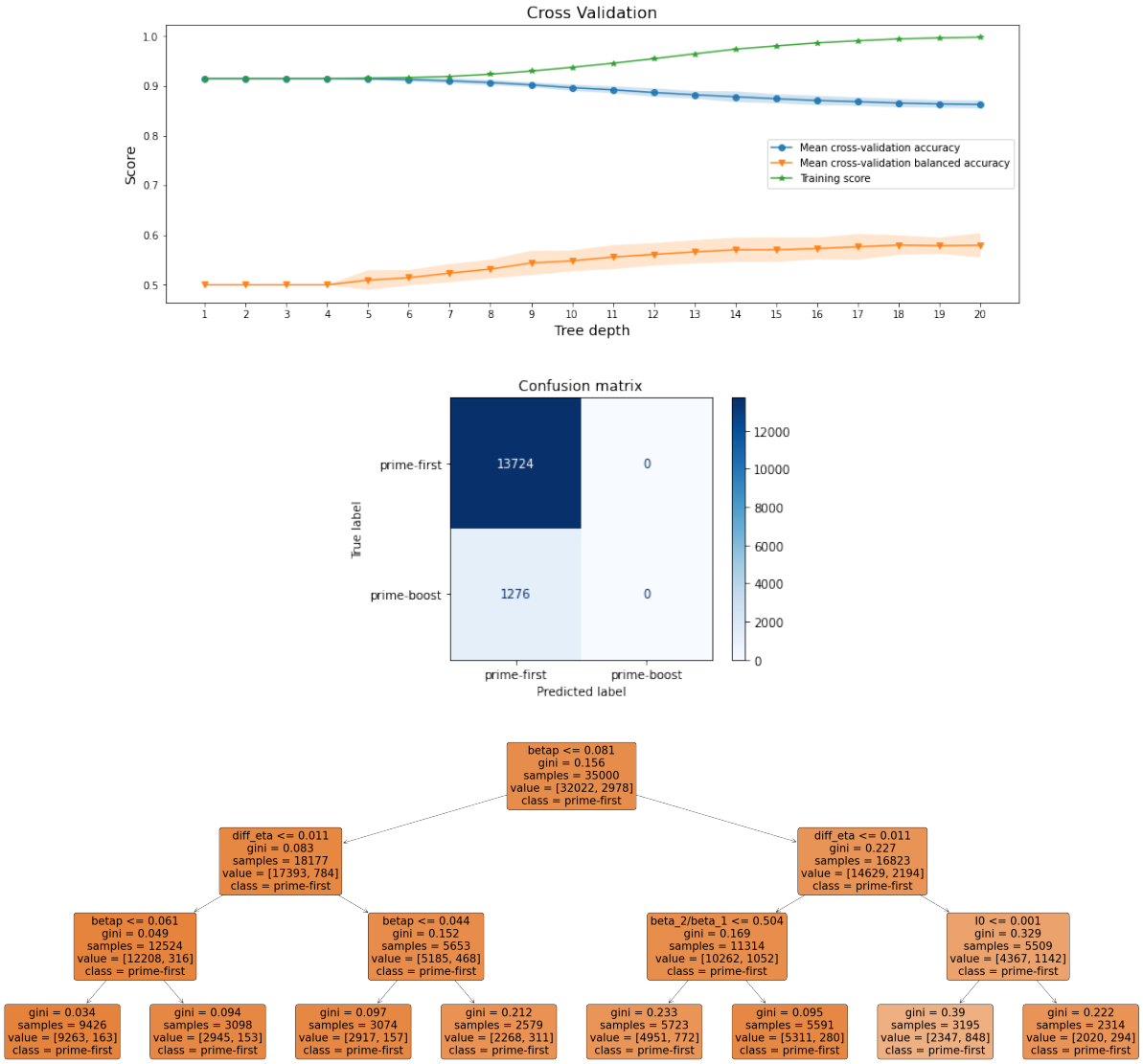

FIG. 6. **Decision tree analysis for dataset B.** Cross-validation: Large spread between training accuracy (green curve) and balanced accuracy (orange) for all depths indicate poor learning performance regarding a possible subtree structure of prime-boost samples but excellent class prediction. High values of accuracy results from excellent prediction for prime-first samples (majority 13724) while prediction of prime-boost is poor (1276 false negatives, 0 true negatives in confusion matrix). Confusion matrix: Number of instances for predicted and true (ground truth) labels, for prime-first and prime-boost. Upper left: True positives, Upper right: False positives, Lower left: False negatives, Lower right: True negatives. Notation: Positives are prime-first, negatives prime-boost. Decision tree: (*Value*) denotes the number of samples in class "prime-first" (left part) and class "prime-boost" (right part), respectively, for the given branch (brackets), while *samples* denote the total sum of samples at the given branch. Left branches satisfy the displayed condition ("yes" branch), right descendants are "no" branches. Notation: betap $= \beta^\star$, diff_eta $= \eta_1 - \eta_2$, beta_2/beta_1 $= \beta_2/\beta_1$, $I0 = I(0)$. Shown depth $= 3$. Resulting tree of depth $= 3$ is essentially equivalent to *always prime-first*, leading to an accuracy of 92%. With increasing depth balanced accuracy (and recall) increase only slightly.