Immunological heterogeneity informs estimation of the durability of COVID-19 vaccine protection

Matthieu Domenech de Cellès (domenech@mpiib-berlin.mpg.de)
Max Planck Institute for Infection Biology

Anabelle Wong
Max Planck Institute for Infection Biology

Laura Barrero Guevara
Max Planck Institute for Infection Biology

Pejman Rohani
University of Georgia

Keywords: COVID-19, vaccine protection, simulation study, vaccine immunity

DOI: https://doi.org/10.21203/rs.3.rs-566774/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Immunological heterogeneity informs estimation of the durability of COVID-19 vaccine protection

Matthieu Domenech de Cellès¹,* Anabelle Wong¹,², Laura Barrero Guevara¹,², Pejman Rohani³,⁴,⁵

1. Infectious Disease Epidemiology group, Max Planck Institute for Infection Biology, Berlin, 10117, Germany

2. Charité—Universitätsmedizin Berlin, 10117, Germany

3. Center for the Ecology of Infectious Diseases, University of Georgia, Athens, GA, 30602, USA

4. Odum School of Ecology, University of Georgia, Athens, GA, 30602, USA

5. Department of Infectious Diseases, University of Georgia, Athens, GA, 30602, USA

*Corresponding author: Dr. Matthieu Domenech de Cellès, domenech@mpiib-berlin.mpg.de. Address: Max Planck Institute for Infection Biology, Charitéplatz 1, Campus Charité Mitte, 10117 Berlin, Germany
Abstract

Deciphering the properties of vaccines against coronavirus disease 2019 (COVID-19) is essential to predict the future course of the pandemic caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). However, current uncertainties about COVID-19 vaccine immunity raise the question of how much time will be needed to estimate these properties, in particular the durability of vaccine protection. Here we designed a simulation study, based on empirically validated epidemiological models of SARS-CoV-2 transmission, to predict the impact of a breadth of vaccines with different mean duration (range: 2–5 years) and heterogeneity (coefficient of variation range: 50–100%) of protection against infection. We then assessed how confidently the duration of protection could be estimated under a range of epidemiological scenarios in the year following the start of mass immunization. We found that lower population mean and higher inter-individual variability facilitated estimation of the duration of vaccine protection. Across the vaccines tested, high waning and high heterogeneity permitted complete identification of the duration of protection; in contrast, low waning and low heterogeneity allowed only estimation of the fraction of vaccinees with rapid loss of immunity. These findings suggest that key aspects of COVID-19 vaccine immunity can be estimated with limited epidemiological data. More generally, they highlight that immunological heterogeneity can sensitively determine the impact of COVID-19 vaccines and, it is likely, of other vaccines.
The pandemic of coronavirus disease 2019 (COVID-19), caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has led to an unprecedented public health crisis worldwide. Following the first case reports in December 2019 in Wuhan, China [1], SARS-CoV-2 rapidly spread across the globe and was declared pandemic as early as March 2020 [2]. The high transmissibility of SARS-CoV-2 infection (facilitated by the frequent occurrence of paucisymptomatic transmission [3] and superspreading events [4]), in addition to the substantial severity and lethality of COVID-19 in some at-risk groups (like the elderly [5]), prompted the implementation of large-scale control measures, which have caused major disruptions globally. In response, intensive research efforts into deciphering the immune response to SARS-CoV-2 infection led to the development of multiple COVID-19 vaccines, some building on new technologies like viral vectors or mRNA [6, 7]. As of May 12, 2021, seven vaccines were used in >10 countries, including 2 mRNA and 3 viral vector vaccines [8]. Clinical trials demonstrated the high efficacy of COVID-19 vaccines, particularly of those based on mRNA [9, 10, 11, 12]. Early epidemiological studies conducted in the months that followed the roll-out of vaccination appear to confirm the effectiveness of COVID-19 vaccines in the field [13, 14], suggesting the feasibility of SARS-CoV-2 infection control by mass immunization.

Despite these encouraging findings, the foreseeable prospect of SARS-CoV-2 establishing itself as an endemic infection emphasizes the need to fully characterize the properties of COVID-19 vaccines [15]. Key among those is the duration of vaccine-derived protection, which remains undetermined. Schematically, two essential components that specify this duration are the population mean and the inter-individual variability. Indeed, a large body of epidemiological theory indicates that the mean duration of vaccine protection predicts essential metrics of long-term control, like critical vaccination coverage and vaccine impact [16, 17, 18]. Although less studied [19, 20], the inter-individual variability in the duration of protection may be equally important to predict the long-term dynamics of SARS-CoV-2 after vaccination, perhaps even more so because heterogeneity appears to be a defining feature of SARS-CoV-2 immunity [21]. Elucidating the duration of vaccine protection can be achieved by identifying immune correlates of protection, a central tenet of vaccinology [22]. Despite potential candidates—such as levels of neutralizing antibodies—identified in animal experimental models [23], the immune correlates of vaccine protection remain as yet unknown in humans [24]. In addition, the search for such correlates may be complicated by the intricate nature of immunity to SARS-CoV-2, characterized by substantial heterogeneity and multiple memory components with different kinetics of decay [21]. In complement to immunological and vaccine studies, mathematical models of transmission and vaccination—which allow translating hypotheses about individual-level mechanisms of vaccine protection into population-level dynamics—may be helpful to elucidate COVID-19 vaccine immunity [25, 26]. Earlier
applications demonstrated the ability of such models to accurately estimate relevant attributes of various vaccines, while capturing both their direct and indirect effects on infection dynamics [27, 28, 29]. However, as the epidemiology of COVID-19 changes rapidly (e.g., because of the emergence of new virus variants [30]), it remains to be determined how much longitudinal data will be needed to estimate the duration of vaccine protection using such models.

To address this gap, we designed a simulation study to determine how reliably the duration of COVID-19 vaccine immunity could be estimated with at most one year of daily, age-specific case report data. We used a previously described, deterministic model of SARS-CoV-2 transmission, empirically validated with data from six countries during December 2019–March 2020 [31], and here adapted to model the post-vaccination dynamics of SARS-CoV-2 in Germany. The model distinguished between subclinical and clinical infections and incorporated age-related variations in susceptibility to SARS-CoV-2 infection and in the fraction with clinical infection, reflecting known features of COVID-19 epidemiology. To assess the impact of immunization campaigns, we extended that model to simulate the effect of vaccines that conferred imperfect and temporary immunity against infection. Specifically, we assumed that COVID-19 vaccines conferred leaky protection against clinical and subclinical infections, as evidenced by earlier clinical and epidemiological studies [9, 10, 11, 12, 13]. To capture waning immunity against either type of infection, we modeled the duration of protection as a Gamma distribution, specified by its mean and its coefficient of variation (CV, a dimensionless and easily interpretable measure of statistical variability defined as the ratio of the standard deviation to the mean). Despite current uncertainties surrounding the degree of waning vaccine immunity, it has been speculated that vaccine protection, though not lifelong, may last at least a few years [32]. The variability in the duration of vaccine protection is similarly unknown, but the large heterogeneity in the immune response—in particular neutralizing antibodies—observed after vaccination in phase 1/2 trials suggests it may be quite large [33, 34, 35]. In numerical applications, we therefore assumed a mean duration of protection in the range 2–5 years and a variability in the range 50–100%.

For different sets of fixed vaccine properties, we then simulated a range of 10 plausible epidemiological scenarios in the year following the start of vaccination in Germany. In all scenarios, we assumed an effective reproduction number of 1.1 [36], a prevalence of past infections of 10% [37], and a prevalence of active infections of $10^{-4}$ at the start of vaccination (Table S1). We modeled case under-reporting using a Negative Binomial observation model, under the conservative assumptions that 50% of clinical infections and 0% of subclinical infections were detected [38, 39]. All the above parameters were assumed known with reasonable accuracy, for example from external epidemiological studies or hindcasts of transmission models around the time of vaccine introduction. In contrast, multiple factors—including, non-exhaustively, non-pharmaceutical interventions [40], climate-induced seasonality, or emergence of viral variants [30]—may cause unpredictable
changes in SARS-CoV-2 transmission, requiring specific effect estimation in practical applications. For simplicity and generality, we modeled such variations using a flexible time-varying function, constrained to result in epidemiological time series consistent with an a priori hypothesis of persistent, but low circulation of SARS-CoV-2 (see Methods and Fig. S5). We used maximum likelihood estimation to estimate this function, as well as the over-dispersion in case reporting and the duration of vaccine protection. In sum, by combining a previously validated model of SARS-CoV-2 transmission and a plausible model of vaccination against COVID-19, we aimed to generate realistic epidemiological data to use as a test bed for estimating the duration of vaccine protection, while controlling for noise in case reporting and confounding trends in transmission.

In our base model, we assumed vaccine effectiveness broadly consistent with those of mRNA vaccines (95% effectiveness against clinical infections [9, 10] and 90% effectiveness against subclinical infections [13]), a mean duration of protection of 5 years, and we contrasted 3 different levels of variability (50%, 71%, and 100%, Fig. S4). The resulting estimates of the average duration of protection—under the assumption that the degree of variability was known a priori—are displayed in Fig. 1A and the full likelihood profiles in Fig. S6. For high variability in the duration of protection (CV = 1, bottom panel), we found that estimation based on 6 months of daily data was typically inaccurate and imprecise. In most scenarios, the limited amount of data resulted in practical non-identifiabilities, such that an upper confidence bound could not be derived. In contrast, a lower confidence bound could be derived in all scenarios tested (range of lower 99% CI bounds: 0.8–2.5 years). Comparable, but more severe estimation problems were found for lower variability, with near-complete lack of identifiability in some epidemiological scenarios (Fig. 1A, upper and middle panels). Increasing the length of data to 1 year substantially improved estimation of the mean duration in all scenarios tested. For high variability, precise and accurate estimates of the mean duration of protection could be derived in all scenarios tested (mean absolute bias: 0.2 years, mean 99% CI: 4.2–6.1 years). Although non-identifiability typically persisted for lower levels of variability, a precise lower confidence bound could be calculated in every scenario (range of lower bounds: 1.7–2.6 years for low variability [CV = 0.50] and 2.7–4.6 years for medium variability [CV = 0.71]). Hence, we found evidence that, all else being equal, higher inter-individual variability facilitated estimation of the mean duration of vaccine protection.

These results suggest that, in most cases, information present in one year of epidemiological data permits reliable estimation of only the lower but not the upper confidence limit for the mean duration of protection. However, such information is highly relevant from a public health perspective because the lower confidence limit relates to the degree of early loss of vaccine protection. It can therefore caution public health actors in decision making when it comes to infection prevention and surveillance. To illustrate, we present in Fig. 1B
the estimates of the fraction of vaccinees losing immunity within 1 year after vaccination—henceforth referred
to as the fraction with short-term immunity. Despite the practical non-identifiabilities reported above, we
found that relatively accurate and precise estimation of this key metric was possible with only 1 year of data
(mean absolute bias ranging from 0.7 to 2.7% across the levels of variability). Importantly, despite identical
mean fixed in all simulations, the fraction with short-term immunity increased with the variability in the
duration of vaccine protection. This observation likely explains the more efficient estimation for vaccines
that confer more heterogeneous protection: more vaccinated individuals lose immunity quickly and can
subsequently get infected and transmit, leaving a more pronounced dynamical fingerprint in epidemiological
data. These findings emphasize the need to pay careful attention to the full distribution of the duration of
vaccine protection to understand the post-vaccination dynamics of SARS-CoV-2.

In practice, current uncertainties about the immune response to COVID-19 vaccination imply that quanti-
tative estimates of the heterogeneity of protection are not available. Next, we therefore proceeded to estimate
both the mean of, and the variability in, the duration of vaccine protection from our set of epidemiological
scenarios. To ensure comparability across the various levels of variability, and because the corresponding
distributions had a markedly different shape (but the same mean of 5 years, cf. Fig. S4), we focused on esti-
mating the fraction with short-term immunity. In simulations of lowly heterogeneous vaccines (CV = 0.50,
true fraction with short-term immunity of 0.9%, Fig. 2 top panel), we found that all models resulted in
equally good fit to the data, such that the degree of variability could not be identified. For mis-specified
levels of variability, however, the fraction with short-term immunity remained well estimated in all scenarios
(mean absolute bias of 1% and 0.5% for medium and high variability, respectively). In simulations of highly
heterogeneous vaccines (CV = 1.00, true fraction with short-term immunity of 18%, Fig. 2 bottom panel), we
found that more information on the degree of variability was present in the simulated data. Specifically, the
hypothesis of low variability could be rejected in all scenarios and the true variability recovered in three. In
the other 7 scenarios, mis-specification of the level of variability caused moderate over-estimation of the true
fraction with short-term immunity (range of point estimates: 0.30–0.34, mean absolute bias: 0.14). These
results confirm that higher inter-individual variability helps estimate the full duration of vaccine protection.
They further suggest that relatively accurate—though partial—information on this duration can be obtained
from limited epidemiological data.

To test the robustness of our results, we conducted three sensitivity analyses. First, we verified that
another set of 10 epidemiological scenarios with the same vaccine properties led to comparable estimates
of the durability of protection (Fig. S7). Second, to take into account the range of efficacy estimates from
clinical trials [9, 10, 11, 12], we tested a vaccine with the same duration of protection but a lower effectiveness
of 60% against clinical infections and of 50% against subclinical infections. We found our main results to
hold in this case, with broadly similar patterns in parameter identifiability (Fig. S8). Finally, we tested a hypothetical vaccine that conferred an average duration of protection against infection of 2 years, instead of 5 years in our base model. Assuming the degree of variability to be known a priori (as in Fig. 1), we found that estimates were substantially more accurate and precise, with finite confidence bounds in most (27 out of 30) scenarios (Fig. S9). Assuming the degree of variability to be unknown a priori, we found that simulations of a highly heterogeneous vaccine permitted complete identification of the distribution of protection in all scenarios tested (Fig. S10). In contrast, simulations of a lowly heterogeneous vaccine did not generally allow identification of the degree of variability, but relatively correct, though slightly biased, estimation of the fraction with short-term immunity (mean absolute bias of 6% for medium variability, of 10% for low variability). We conclude from these experiments that both the mean of, and the variability in, the duration of vaccine protection sensitively determine the impact of COVID-19 vaccines and the ability to estimate their properties.

Our study has important limitations, which result from our deliberate choice to use a simple model of COVID-19 vaccination. First, we did not consider the possibility of a concomitant roll-out of multiple vaccines, yet a common situation in the field. Our simple model could still be applied to ≥ 2 vaccines deemed to confer comparable immunity (e.g., mRNA vaccines), but will require modifications otherwise. Second, as for other RNA viruses, the high mutation rate of SARS-CoV-2 may cause the emergence of vaccine-escape variants that progressively modify, or even nullify, vaccine protection. In this case, our assumption of fixed vaccine properties would need to be revised; we believe, however, it is reasonable because of the short time period (1 year) considered in our study. Third, we assumed that post-vaccine infections (i.e., infections after immunity had waned) in vaccinees had the same characteristics as primary infections in immunologically naïve individuals. In reality, more complex immunological mechanisms may be at play, such that partial protection against infection or transmission is maintained via recall B or T cell memory response, even after disappearance of residual sterilizing immunity mediated (for example) by neutralizing antibodies [24]. Differences between post-vaccine and primary infections—e.g., in terms of infectiousness, duration, and severity—may sensitively shape the future dynamics of COVID-19 [15], but may be hard to resolve in practice, even with years of epidemiological data [29]. Alternatively, specific vaccine studies could help elucidate these differences, for example studies designed to estimate vaccine effectiveness on infectiousness or on disease progression in vaccinated breakthrough cases [41, 17]. Finally, for simplicity and interpretability, we assumed that the duration of vaccine protection could be modeled by a simple statistical distribution, fully specified by its first two moments (mean and variance). Nevertheless, more complex distributions may be required to capture the heterogeneity of vaccine protection and its consequences on the long-term dynamics of SARS-CoV-2.
In conclusion, our study provides reassuring evidence that only a short period of time may be required to estimate some key aspects of COVID-19 immunity, even in the face of uncertainties about changes in SARS-CoV-2 transmission. Echoing earlier findings [19], our results also highlight the need to accurately quantify heterogeneity of vaccine protection to predict the impact of mass immunization. More generally, a more systematic examination of the consequences of heterogeneity—the norm in biology [42]—may be warranted, not only for the study of COVID-19 vaccines but also of other vaccines.

Methods

Study population and study period In this simulation study, we considered Germany as our study population and we simulated a study period of at most 1 year following the introduction of COVID-19 vaccines. As vaccination against COVID-19 started in late 2020/early 2021 in Germany, the study period approximately covered year 2021 and initial conditions were assumed to represent the epidemiological situation at the beginning of that year. The endpoints considered were case reports of clinical or subclinical SARS-CoV-2 infections.

Model formulation and parametrization We used a previously described deterministic model of SARS-CoV-2 transmission, empirically estimated and validated with data in six countries (China, Italy, Japan, Singapore, Canada, and South Korea) during the first wave (December 2019–March 2020) of COVID-19 [31]. Briefly, the model is an extension of the standard Susceptible–Exposed–Infected–Recovered (SEIR) model, stratified according to age (\( I = 8 \) age groups considered here: 0–9, 10–19, . . . , 60–69, and ≥70yr) and according to type of infection, either subclinical (defined as asymptomatic or paucisymptomatic) or clinical. To add biological realism, the latent and the infectious periods were assumed Gamma-distributed with a shape parameter of 2 and a mean of \( 1/\sigma = 3 \) days and \( 1/\gamma = 5 \) days, respectively [31]. The age-specific susceptibility to SARS-CoV-2 infection (denoted by \( u_{i=1,...,I} \)) and clinical fraction (\( y_{i=1,...,I} \)) were fixed from the estimates in Ref. [31] and are displayed in Fig. S2. As in [31], we also assumed that subclinical infections were \( \theta = 50\% \) as transmissible as clinical infections. The rates of contacts between age groups were fixed using data from the POLYMOD study in the UK [43], corrected for reciprocity with 2019 age-specific demographic data in Germany. Although Germany was also part of the POLYMOD study, to keep our model general we did not use the corresponding data because they were found to differ markedly from those in other European countries [44].

To model changes in SARS-CoV-2 transmission over the year following the start of vaccination, we used a flexible time-varying function \( \beta(t) \) :
\[
\log \beta(t) = \log \beta_0 + \beta_1 (\cos \omega t - \cos 2\omega t) + \beta_2 \sin \omega t + \beta_3 \sin 2\omega t + \beta_4 \sin 3\omega t
\]

where \( \omega = \frac{2\pi}{365} \) day\(^{-1} \) and \( \beta(0) = \beta_0 \). Although arbitrary, this function was chosen to allow multiple (at most 3) peaks over the year, as observed in time series of the effective reproduction number during year 2020 in European countries. At start of vaccination, we assumed an initial effective reproduction number \( R_e(0) = 1.1 \) \cite{36}; the corresponding value of \( \beta_0 \) was calculated by equating \( R_e(0) \) to the leading eigenvalue of the next-generation matrix \cite{31}:

\[
\text{NGM}_{ij} = \beta_0 u_i c_{ij} \frac{y_j}{\gamma} (1 - y_j) \lambda_i(t) \frac{N_i}{N_j}
\]

where \( r_i(0) \) is the initial fraction recovered and \( N_i \) the population size in age group \( i \). The other trend parameters \( \beta_i=1,...,4 \) were estimated from the simulated data, as explained below. In every age group, we assumed an initial prevalence of past infections of 10\% \cite{37}, of exposed infections of \( 10^{-4} \) and of active infections of \( 10^{-4} \). We supposed all initial conditions to be known with reasonable accuracy, for example from epidemiological studies or from hindcasts of epidemiological models around the time of vaccine introduction.

Next, we extended the model described above to incorporate vaccination against COVID-19. Specifically, we assumed that COVID-19 vaccines conferred imperfect immunity against clinical infections (vaccine effectiveness against clinical infections \( \text{VE}_C = 1 - \epsilon_C \), where \( \epsilon_C \) is the leakiness to clinical infections) and against subclinical infections (vaccine effectiveness against subclinical infections: \( \text{VE}_S = 1 - \epsilon_S \)). Vaccine protection was assumed leaky—as opposed to “all-or-nothing” \cite{18}—based on the observations from phase 1/2 clinical trials, which detected a neutralizing antibody response in all vaccine recipients after a booster dose \cite{45,33,46,34}. Finally, vaccine protection was assumed to wane over time, at rate \( \alpha \). Using the method of stages \cite{47}, the distribution of the duration of protection was varied by subdividing the vaccinated class \( V \) into \( n_V \) sub-classes. The resulting distribution \( D_V \) was Gamma with shape parameter \( n_V \), mean \( \text{E}(D_V) = \frac{1}{\alpha} \), and coefficient of variation \( \text{CV}(D_V) = \frac{\sqrt{\text{V}(D_V)}}{\text{E}(D_V)} = \frac{1}{\sqrt{n_V}} \). Vaccination was assumed to vary over time and to occur continuously, at rate \( v_i(t) \) in age group \( i \).

Given these assumptions, the model was given by the following set of differential equations:

\[
\begin{align*}
\dot{V}_{i,1} &= v_{i}(t)S_{i} - n_{V} \alpha V_{i,1} - [\epsilon_C y_i + \epsilon_S (1 - y_i)] \lambda_i(t) V_{i,1} \\
\dot{V}_{i,j} = 2,...,n_{V} &= n_{V} \alpha (V_{i,j-1} - V_{i,j}) - [\epsilon_C y_i + \epsilon_S (1 - y_i)] \lambda_i(t) V_{i,j}
\end{align*}
\]
\[
\begin{align*}
\dot{S}_i &= -[v_i(t) + \lambda_i(t)]S_i \\
\dot{S}_{V,i} &= n_V \alpha V_{i,n_V} - \lambda_i(t)S_{V,i} \\
\dot{E}_{C,i,1} &= y_i\lambda_i(t)(S_i + S_{V,i} + \epsilon_C V_i) - 2\sigma E_{C,i,1} \\
\dot{E}_{C,i,2} &= 2\sigma (E_{C,i,1} - E_{C,i,2}) \\
\dot{I}_{C,i,1} &= 2\sigma E_{C,i,2} - 2\gamma I_{C,i,1} \\
\dot{I}_{C,i,2} &= 2\gamma (I_{C,i,1} - I_{C,i,2}) \\
\dot{E}_{S,i,1} &= (1 - y_i)\lambda_i(t)(S_i + S_{V,i} + \epsilon_S V_i) - 2\sigma E_{S,i,1} \\
\dot{E}_{S,i,2} &= 2\sigma (E_{S,i,1} - E_{S,i,2}) \\
\dot{I}_{S,i,1} &= 2\sigma E_{S,i,2} - 2\gamma I_{S,i,1} \\
\dot{I}_{S,i,2} &= 2\gamma (I_{S,i,1} - I_{S,i,2}) \\
\dot{R}_i &= 2\gamma (I_{C,i,2} + I_{S,i,2})
\end{align*}
\]

where \(V_i = \sum_j V_{i,j}\) is the total number of individuals vaccinated and:

\[
\lambda_i(t) = \beta(t)u_i \sum_j c_{ij} \frac{I_{C,j,1} + I_{C,j,2} + \theta(I_{S,j,1} + I_{S,j,2})}{N_j}
\]

the (per capita) force of infection in age group \(i\).

To complete the model specification and to generate the simulated data, we used a Negative Binomial observation model. Specifically, we assumed that \(\rho_C = 50\%\) of clinical infections were reported to the surveillance system, in keeping with previous estimates during March–June 2020 in Germany [38]. The reporting fidelity of subclinical infections (denoted by \(\rho_S\)) is less well known, but can be evaluated by comparing the burden estimated from case reports and sero-epidemiological surveys. Such comparison indicated an overall (i.e., including clinical and subclinical infections) reporting probability of 15–20% in Germany [39], suggesting low reporting of subclinical infections. Here we made the most conservative assumption that all subclinical infections were unreported (\(\rho_S = 0\)). The over-dispersion in case reporting was fixed to \(k = 0.04\) (approximate coefficient of variation of the observation model: 20%) and estimated from the data. Given current surveillance efforts, cases were assumed to be reported daily in every age group.

**Vaccine properties and epidemiological scenarios considered**  
In accordance with vaccination strategies defined by Germany and most other European countries [48], vaccination was assumed continuous and staggered, starting with individuals \(\geq 70\) yr, followed by \(60–69\) yr, and \(20–59\) yr. According to polls on vaccination willingness conducted in January 2021 in Germany [49], we assumed that 60% of \(\geq 70\) yr, \(60–69\) yr, and
20–59 yr would get vaccinated 0 to <4, 4 to <8, and 8 to <12 months after start of vaccination, respectively. In our base model, we assumed a mean duration of protection of 5 years and we contrasted 3 different levels of variability fixed to 50, 71, and 100% ($n_V \in \{4, 2, 1\}$, respectively). Other vaccine attributes were assumed known and fixed from the results of clinical trials and early epidemiological studies: vaccine effectiveness of 95% against clinical infections [10, 9] and of 90% against subclinical infections [13].

To generate alternative epidemiological scenarios for the spread of SARS-CoV-2 in the year following start of vaccination, we randomly sampled values for the trend parameters ($\beta_1, \ldots, 4$, simulation range: –0.3 to 0.3). To add realism, we selected 10 scenarios that resulted in simulated data meeting the following criteria:

- Overall cumulative incidence of total cases ≤10%.
- Maximum incidence of ≤50/100,000 daily case reports.
- No peak in overall case reports during summer.
- Maximum of 5% of days with no case reports.

These criteria were chosen to reflect the a priori hypothesis of persistent, but low circulation of SARS-CoV-2 in the year following start of vaccination.

**Parameters’ estimation**  For every scenario, the following parameters were estimated from the data: rate of waning vaccine protection ($\alpha$), trend parameters ($\beta_1, \ldots, 4$) and over-dispersion in case reporting ($k_C$). Estimation was conducted using maximum likelihood estimation via trajectory matching [50]. In a first analysis, we assumed that the variability in the duration of vaccine protection was known, and we calculated the profile log-likelihood to estimate the maximum likelihood estimate and the 99% confidence interval (CI) of the average duration of vaccine protection ($1/\alpha$) [51]. We evaluated the profile log-likelihood on a grid of 100 values in the range 0.5–50 years, uniformly distributed on a logarithmic scale. For every grid point, estimation was conducted 10 times, with starting parameter values sampled over broad ranges using a Latin hypercube design (Table S1). In a second analysis, we assumed that the variability in the duration of protection was unknown, and we estimated both the mean and the coefficient of variation of the duration of protection. The estimation proceeded similarly, except that the profile log-likelihood was evaluated on a grid of 100 values of the fraction with short-term immunity (here defined as the proportion of individuals losing immunity within 1 year after vaccination), uniformly arranged in the interval 0.005–0.995. We focused on this fraction in the second analysis, because, even though all the distributions considered had a similar mean, their overall shape differed markedly (see Fig. S4). All likelihood profiles were inspected visually and subsequently smoothed using local quadratic regression. The degree of smoothing was selected by visual inspection of the resulting...
smoothed profile; the resulting span parameter varied across scenarios, but all values were either 0.1 or 0.5. An approximate 99% confidence interval was calculated as the range of smoothed parameter values within $\chi^2_{p=0.99, df=1/2} \approx 3.32$ log-likelihood units of the maximum log-likelihood.

**Assessment of estimation performance**  To assess estimation efficiency, we quantified both the accuracy and the precision of the estimates of average duration of protection. Specifically, we measured estimation accuracy by the mean absolute bias, defined as the mean absolute difference between the estimated values and the true value. To evaluate estimation precision, we examined the range of lower and upper bounds of the 99% confidence intervals. If the lower (upper) confidence bound equalled the lower (upper) limit of the estimation interval (0.5–50 years) for a given epidemiological scenario, the corresponding parameter estimates were considered practically non-identifiable for decreasing (increasing) parameter values [51].

**Numerical implementation**  All models were implemented, simulated, and estimated using the pomp package [52, 50], operating in R version 3.6.3 [53]. The socialmixr package was used to calculate the contact matrix [54], based on the data from the POLYMOD survey included therein. The renv package was used to keep track of all packages’ version and to increase reproducibility [55].
Figure 1: Estimates of the average duration of protection (A) and of the fraction with short-term immunity (B), assuming known variability in the duration of protection conferred by COVID-19 vaccines. The black dashed line indicates the true values used in numerical simulations, corresponding to a mean duration of protection of 5 years (full distribution displayed in Fig. S4). For each of 10 simulations (ordered by increasing simulation number from bottom to top and displayed, for CV=1.00, in Fig. S5), light points (intervals) represent the maximum likelihood estimate (99% confidence interval), calculated using the profile log-likelihood for a mean duration of protection in the range 0.5–50 years. Solid points (intervals) represent the corresponding quantities, averaged across simulations. The data length refers to the duration of simulated data used for estimation. In panel A, the x-axis values are \( \log_{10} \)-transformed for visual clarity. CV: coefficient of variation, quantifying the variability in the duration of vaccine protection.
Figure 2: Estimates of the fraction with short-term immunity, assuming unknown variability in the duration of protection conferred by covid vaccines. Represented are the maximum likelihood estimates of the fraction with short-term immunity (y-axis), for different values of the variability in the duration of protection (x-axis). In each panel, the dashed black line indicates the true fraction with short-term immunity, and the title the true variability in the duration of vaccine protection that were fixed in model simulations. As in Fig. 1, the mean duration of vaccine protection was fixed to 5 years in all scenarios tested. The grey lines connect points that belong to the same simulation (10 simulations in total). Each point’s color indicates the scaled log-likelihood (calculated, for each simulation, as the difference with the maximum log-likelihood); points colored in blue are within $\frac{1}{2} \chi^2_{p=0.99,df=1} \approx 3.32$ log-likelihood units from the maximum log-likelihood and therefore in the 99% confidence interval for the corresponding simulation.
References

[1] Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med. 2020 02;382(8):727–733.

[2] World Health Organization. WHO Director-General’s opening remarks at the media briefing on COVID-19 - 11 March 2020;. Available from: https://s.gwdg.de/rQpZ03.

[3] Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science. 2020 05;368(6490):489–493.

[4] Althouse BM, Wenger EA, Miller JC, Scarpino SV, Allard A, Hébert-Dufresne L, et al. Superspreading events in the transmission dynamics of SARS-CoV-2: Opportunities for interventions and control. PLoS Biol. 2020 11;18(11):e3000897.

[5] Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. Factors associated with COVID-19-related death using OpenSAFELY. Nature. 2020 08;584(7821):430–436.

[6] Carvalho T, Krammer F, Iwasaki A. The first 12 months of COVID-19: a timeline of immunological insights. Nat Rev Immunol. 2021 04;21(4):245–256.

[7] Kyriakidis NC, López-Cortés A, González EV, Grimaldos AB, Prado EO. SARS-CoV-2 vaccines strategies: a comprehensive review of phase 3 candidates. NPJ Vaccines. 2021 Feb;6(1):28.

[8] The New York Times. Tracking Coronavirus Vaccinations Around the World;. Available from: https://www.nytimes.com/interactive/2021/world/covid-vaccinations-tracker.html.

[9] Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, et al. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. N Engl J Med. 2021 02;384(5):403–416.

[10] Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. N Engl J Med. 2020 Dec;383(27):2603–2615.

[11] Ramasamy MN, Minassian AM, Ewer KJ, Flaxman AL, Folegatti PM, Owens DR, et al. Safety and immunogenicity of ChAdOx1 nCoV-19 vaccine administered in a prime-boost regimen in young and old adults (COV002): a single-blind, randomised, controlled, phase 2/3 trial. Lancet. 2021 12;396(10267):1979–1993.

[12] Voysey M, Clemens SAC, Madhi SA, Weckx LY, Folegatti PM, Aley PK, 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. Lancet. 2021 01;397(10269):99–111.
[13] Dagan N, Barda N, Kepten E, Miron O, Perchik S, Katz MA, et al. BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting. N Engl J Med. 2021 Feb;

[14] Hall VJ, Foulkes S, Saei A, Andrews N, Oguti B, Charlett A, et al. 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. 2021 Apr;

[15] Saad-Roy CM, Wagner CE, Baker RE, Morris SE, Farrar J, Graham AL, et al. Immune life history, vaccination, and the dynamics of SARS-CoV-2 over the next 5 years. Science. 2020 11;370(6518):811–818.

[16] Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton: Princeton University Press; 2008. Available from: http://www.loc.gov/catdir/toc/fy0805/2006939548.html.

[17] Halloran ME, Longini IM, Struchiner CJ, Longini IM. Design and analysis of vaccine studies. vol. 18. Springer; 2010.

[18] Magpantay FMG, Riolo MA, DE Cellès MD, King AA, Rohani P. Epidemiological consequences of imperfect vaccines for immunizing infections. SIAM J Appl Math. 2014;74(6):1810–1830.

[19] Gomes MGM, Lipsitch M, Wargo AR, Kurath G, Rebelo C, Medley GF, et al. A missing dimension in measures of vaccination impacts. PLoS Pathog. 2014 Mar;10(3):e1003849.

[20] Woolthuis RG, Wallinga J, van Boven M. Variation in loss of immunity shapes influenza epidemics and the impact of vaccination. BMC Infect Dis. 2017 09;17(1):632.

[21] Dan JM, Mateus J, Kato Y, Hastie KM, Yu ED, Faliti CE, et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. Science. 2021 02;371(6529).

[22] Pollard AJ, Bijker EM. A guide to vaccinology: from basic principles to new developments. Nat Rev Immunol. 2021 02;21(2):83–100.

[23] McMahan K, Yu J, Mercado NB, Loos C, Tostanoski LH, Chandrashekar A, et al. Correlates of protection against SARS-CoV-2 in rhesus macaques. Nature. 2021 02;590(7847):630–634.

[24] Cromer D, Juno JA, Khoury D, Reynaldi A, Wheatley AK, Kent SJ, et al. Prospects for durable immune control of SARS-CoV-2 and prevention of reinfection. Nat Rev Immunol. 2021 Apr;

[25] Lavine JS, Rohani P. Resolving pertussis immunity and vaccine effectiveness using incidence time series. Expert Rev Vaccines. 2012 Nov;11(11):1319–29.
[26] Domenech de Cellès M, King AA, Rohani P. Commentary: resolving pertussis resurgence and vaccine immunity using mathematical transmission models. Hum Vaccin Immunother. 2019;15(3):683–686.

[27] Blackwood JC, Cummings DAT, Broutin H, Iamsirithaworn S, Rohani P. Deciphering the impacts of vaccination and immunity on pertussis epidemiology in Thailand. Proc Natl Acad Sci U S A. 2013 Jun;110(23):9595–600.

[28] Argante L, Tizzoni M, Medini D. Fast and accurate dynamic estimation of field effectiveness of meningococcal vaccines. BMC Med. 2016 Jun;14:98.

[29] Domenech de Cellès M, Magpantay FMG, King AA, Rohani P. The impact of past vaccination coverage and immunity on pertussis resurgence. Sci Transl Med. 2018 03;10(434).

[30] Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. Science. 2021 04;372(6538).

[31] Davies NG, Klepac P, Liu Y, Prem K, Jit M, CMMID COVID-19 working group, et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. Nat Med. 2020 08;26(8):1205–1211.

[32] Kim DS, Rowland-Jones S, Gea-Mallorquí E. Will SARS-CoV-2 Infection Elicit Long-Lasting Protective or Sterilising Immunity? Implications for Vaccine Strategies (2020). Front Immunol. 2020;11:571481.

[33] Anderson EJ, Rouphael NG, Widge AT, Jackson LA, Roberts PC, Makhene M, et al. Safety and Immunogenicity of SARS-CoV-2 mRNA-1273 Vaccine in Older Adults. N Engl J Med. 2020 12;383(25):2427–2438.

[34] Folegatti PM, Ewer KJ, Aley PK, Angus B, Becker S, Belij-Rammerstorfer S, et al. Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. Lancet. 2020 08;396(10249):467–478.

[35] Mulligan MJ, Lyke KE, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Phase I/II study of COVID-19 RNA vaccine BNT162b1 in adults. Nature. 2020 10;586(7830):589–593.

[36] Abbott S, Bennett C, Hickson J, Allen J, Sherratt K, Funk S. National and Subnational Estimates of the Covid 19 Reproduction Number (R) for Germany Based on Test Results. Harvard Dataverse; 2020. Available from: https://doi.org/10.7910/DVN/LNMJYJ.

[37] Arora RK, Joseph A, Van Wyk J, Rocco S, Atmaja A, May E, et al. SeroTracker: a global SARS-CoV-2 seroprevalence dashboard. Lancet Infect Dis. 2021 04;21(4):e75–e76.
[38] Russell TW, Golding N, Hellewell J, Abbott S, Wright L, Pearson CAB, et al. Reconstructing the early global dynamics of under-ascertained COVID-19 cases and infections. BMC Med. 2020 10;18(1):332.

[39] Santos-Hövener C, Neuhauser HK, Rosario AS, Busch M, Schlaud M, Hoffmann R, et al. Serology- and PCR-based cumulative incidence of SARS-CoV-2 infection in adults in a successfully contained early hotspot (CoMoLo study), Germany, May to June 2020. Euro Surveill. 2020 11;25(47).

[40] Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature. 2020 08;584(7820):257–261.

[41] Halloran ME, Struchiner CJ, Longini IM Jr. Study designs for evaluating different efficacy and effectiveness aspects of vaccines. Am J Epidemiol. 1997 Nov;146(10):789–803.

[42] Schwartz Daniel. Le jeu de la science et du hasard: la statistique et le vivant. Paris: Flammarion; 1994 [In French].

[43] Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med. 2008 Mar;5(3):e74.

[44] Rohani P, Zhong X, King AA. Contact network structure explains the changing epidemiology of pertussis. Science. 2010 Nov;330(6006):982–5.

[45] Walsh EE, Frenck RW Jr, Falsey AR, Kitchin N, Absalon J, Gurtman A, et al. Safety and Immunogenicity of Two RNA-Based Covid-19 Vaccine Candidates. N Engl J Med. 2020 12;383(25):2439–2450.

[46] Jackson LA, Anderson EJ, Rouphael NG, Roberts PC, Makhene M, Coler RN, et al. An mRNA Vaccine against SARS-CoV-2 - Preliminary Report. N Engl J Med. 2020 11;383(20):1920–1931.

[47] Lloyd AL. Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods. Proc Biol Sci. 2001 May;268(1470):985–93.

[48] European Centre for Disease Prevention and Control. Overview of COVID-19 vaccination strategies and vaccine deployment plans in the EU/EEA and the UK. European Centre for Disease Prevention and Control; 2 December 2020.

[49] YouGov. International study: How many people will take the COVID vaccine?. Available from: https://s.gwdg.de/r7g7qT.

[50] King AA, Nguyen D, Ionides EL. Statistical Inference for Partially Observed Markov Processes via the R Package pomp. Journal of Statistical Software. 2016;69(1):1–43.
[51] Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, et al. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics. 2009;25(15):1923–1929.

[52] Aaron A King and Edward L Ionides and Carles Martinez Bretó and Stephen P Ellner and Matthew J Ferrari and Bruce E Kendall and Michael Lavine and Dao Nguyen and Daniel C Reuman and Helen Wearing and Simon N Wood. pomp: Statistical Inference for Partially Observed Markov Processes; 2020. R package, version 2.7. Available from: https://kingaa.github.io/pomp/.

[53] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2020. Https://www.R-project.org/. Available from: https://www.R-project.org/.

[54] Sebastian Funk. socialmixr: Social Mixing Matrices for Infectious Disease Modelling; 2020. R package version 0.1.8. Available from: https://CRAN.R-project.org/package=socialmixr.

[55] Kevin Ushey. renv: Project Environments; 2021. R package version 0.12.5. Available from: https: //CRAN.R-project.org/package=renv.
Declarations

Availability of data and materials  All R programming codes can be found at https://transfer.mpiib-berlin.mpg.de/s/xnMr3dHDCmTMaC and will be made available via an open-access digital repository upon publication.

Competing interests  MDdC received postdoctoral funding (2017–2019) from Pfizer and consulting fees from GSK. All other authors declare no competing interests.

Funding  No specific funding was used for this study.

Authors Contributions  MDdC conceived of the study design and performed the analysis. AW and LGB conducted bibliographical research and provided content expertise. PR conceived of the study design and oversaw the analysis. All authors helped draft, and approved the final version of, the manuscript.

Acknowledgements  Computations underlying the analysis were performed at the Max Planck Computing and Data Facility (MPCDF).
Figure S1: **Model schematic.** For simplicity, only a given age group $i = 1, \ldots, I$ is represented. Accumulator variables used for the observation model are represented by circles, the other state variables by squares. To add biological realism, the compartments $E$, $I$, and $V$ were divided into multiple sub-classes, not represented here for simplicity. See table S1 for the list of model parameters.
Figure S2: **Age-specific clinical fraction and susceptibility to infection with SARS-CoV-2.** Mean estimates from Davies et al., Figure S4 [31].
Figure S3: **Contact matrix.** Data from POLYMOD study in the UK [43], corrected for reciprocity using 2019 demographic data in Germany.
Figure S4: Assumed distribution of the duration of vaccine protection. Using the method of stages, the duration of protection $D_V$ was assumed Gamma distributed with shape parameter $n_V \in \{1, 2, 4\}$ and a mean of 5 years (depicted by the vertical dashed line). The resulting distributions had a coefficient of variation $\text{CV}(D_V) = \frac{1}{\sqrt{n_V}} \in \{0.5, 0.71, 1.00\}$. 
| Parameter  | Meaning                                           | Fixed value (starting parameter estimation range) | Source/comment |
|------------|--------------------------------------------------|--------------------------------------------------|----------------|
| $1/\sigma$ | Average latent period                            | 3 days                                           | [31]           |
| $1/\gamma$ | Average infectious period                         | 5 days                                           | [31]           |
| $u_i$      | Susceptibility to SARS-CoV-2 infection            | Fig. S2                                          | [31], Fig. S4  |
| $y_i$      | Clinical fraction                                 | Fig. S2                                          | [31], Fig. S4  |
| $\theta$  | Relative infectiousness of subclinical infections | 0.5                                              | [31]           |
| $c_{ij}$   | Age-specific contact rates                        | Fig. S3                                          | [43, 54]       |
| $R_e(0)$   | Initial reproduction number                       | 1.1                                              | [36]           |
| $\beta_0$ | Initial transmission coefficient                  | 0.045                                            | Calculated so that $R_e(0) = 1.1$ |
| $\beta_1,\ldots,4$ | Trend in transmission | $(-1, 1)$                                       |               |
| $E_i(0)$   | Initial fraction exposed                          | $10^{-4}$                                        | Assumption     |
| $I_i(0)$   | Initial fraction infected                         | $10^{-4}$                                        | Assumption     |
| $R_i(0)$   | Initial fraction recovered                        | 0.1                                              | Seroprevalence studies [37] |
| $\rho_C$  | Reporting probability of clinical infections      | 0.5                                              | [38]           |
| $\rho_S$  | Reporting probability of subclinical infections   | 0                                                | Assumption (cf. [39] and main text) |
| $k_C$      | Over-dispersion in case reporting                 | 0.04 ($0, 0.25$)                                 | Coefficient of variation of observation model $\approx 20\%$ |
| $v_i(t)$   | COVID-19 vaccination rate                         | Calculated so that 60% of $\geq 70$, 60–69, and 20–59 yr vaccinated during months 1–4, 5–8, 9–12, respectively | [48, 49] |
| $1 - \epsilon_C$ | Vaccine effectiveness against clinical infections | 0.95                                              | [9, 10]. SA: 60% |
| $1 - \epsilon_S$ | Vaccine effectiveness against subclinical infections | 0.90                                              | [13]. SA: 50% |
| $\bar{\tau}(D_V) = 1/\alpha$ | Mean duration of vaccine protection | 5 years                                           | Assumption. SA: 2 years |
| $CV(D_V) = \frac{1}{\sqrt{n_V}}$ | Variability in duration of protection | 0.50, 0.71, 1.00                                   | Assumption     |

Table S1: **List of model parameters.** SA: sensitivity analysis.
Figure S5: **Simulated data.** Simulated daily incidence of reported cases for a highly heterogeneous vaccine ($n_V = 1$, $CV = 1.00$) with the following properties: average duration of protection of 5 years, effectiveness of 95% against clinical infections and of 90% against subclinical infections (base model). NB: the model and the data used for statistical inference consisted of 8 age groups; for visual clarity, some age groups have been merged in this figure. The $y$-axis values are $\log_{10}$-transformed for visual clarity.
Supplementary results

Figure S6: Full likelihood profiles of the average duration of vaccine protection. Vaccine properties assumed (base model, as in Fig. 1): average duration of vaccine protection of 5 years, effectiveness of 95% against clinical infections and of 90% against subclinical infections. Represented are the smoothed profiles, calculated using local quadratic regression; each line represents one of 10 distinct epidemiological scenarios. The vertical black dashed line indicates the true value used in numerical simulations (5 years). For each scenario, the horizontal black dashed line is at $\frac{1}{2} \chi^2_{p=0.99, df=1} \approx 3.32$ log-likelihood units from the maximum scaled log-likelihood.
Figure S7: Estimates of the average duration of protection (A) and of the fraction with short-term immunity (B), assuming known variability in the duration of protection conferred by COVID-19 vaccines. Results of a sensitivity analysis for a vaccine with mean duration of protection of 5 years, effectiveness of 95% against clinical infections and of 90% against subclinical infections (base model, different set of 10 epidemiological scenarios). Please refer to the legend of Fig. 1 for further explanation of this figure.
Figure S8: Estimates of the average duration of protection (A) and of the fraction with short-term immunity (B), assuming known variability in the duration of protection conferred by COVID-19 vaccines. Results of a sensitivity analysis for a vaccine with mean duration of protection of 5 years, effectiveness of 60% against clinical infections and of 50% against subclinical infections. Please refer to the legend of Fig. 1 for further explanation of this figure.
Figure S9: Estimates of the average duration of protection (A) and of the fraction with short-term immunity (B), assuming known variability in the duration of protection conferred by COVID-19 vaccines. Results of a sensitivity analysis for a vaccine with mean duration of protection of 2 years, effectiveness of 95% against clinical infections and of 90% against subclinical infections. Please refer to the legend of Fig. 1 for further explanation of this figure.
Figure S10: Estimates of the fraction with short-term immunity, assuming unknown variability in the duration of protection conferred by covid vaccines. Results of a sensitivity analysis for a vaccine with mean duration of protection of 2 years, effectiveness of 95% against clinical infections and of 90% against subclinical infections. Please refer to the legend of Fig. 2 for further explanation of this figure.