Estimating the distribution of COVID-19-susceptible, -recovered, and -vaccinated individuals in Germany up to April 2022

Benjamin F. Maier,[1,*] Annika H. Rose,[1] Angelique Burdinski,[1] Pascal Klamser,[1] Hannelore Neuhauser,[2] Ole Wichmann,[3] Lars Schaade,[1] Lothar H. Wieler,[1] and Dirk Brockmann[1]

[1]*Robert Koch Institute, Nordufer 20, D-13353 Berlin, Germany*

[2]*Robert Koch Institute, General-Pape-Straße 62-66, D-12101 Berlin, Germany*

[3]*Robert Koch Institute, Seestraße 10, D-13353 Berlin, Germany*

(Dated: April 19, 2022)

Abstract

After having affected the population for two years, the COVID-19 pandemic has reached a phase where a considerable number of people in Germany have been either infected with a SARS-CoV-2 variant, vaccinated, or both. Yet the full extent to which the population has been in contact with either virus or vaccine remains elusive, particularly on a regional level, because (a) infection counts suffer from under-reporting, and (b) the overlap between the vaccinated and recovered subpopulations is unknown. Since previous infection, vaccination, or especially a combination of both reduce the risk of severe disease, a high share of individuals with SARS-CoV-2 immunity lowers the probability of severe outbreaks that could potentially overburden the public health system once again, given that emerging variants do not escape this reduction in susceptibility. Here, we estimate the share of immunologically naïve individuals by age group for each of the 16 German federal states by integrating an infectious disease model based on weekly incidences of SARS-CoV-2 infections in the national surveillance system and vaccine uptake, as well as assumptions regarding under-ascertainment. We estimate a median share of 7.0% of individuals in the German population have neither been in contact with vaccine nor any variant as of March 31, 2022 (quartile range [3.6%–9.8%]). For the adult population at higher risk of severe disease, this figure is reduced to 3.5% [1.3%–5.5%] for ages 18–59 and 4.3% [2.7%–5.8%] for ages 60 and above. However, estimates vary between German states mostly due to heterogeneous vaccine uptake. Excluding Omicron infections from the analysis, 16.1% [14.0%–17.8%] of the population in Germany, across all ages, are estimated to be immunologically naïve, highlighting the large impact the Omicron wave had until the beginning of spring in 2022.

I. INTRODUCTION

The COVID-19 pandemic caused by the rapid global dissemination of the SARS-CoV-2 virus and its respective variants has led to a large number of infections worldwide [1]. In Germany, around 21.4 million infections have been reported as of the end of March 2022. Moreover, a large part of the population has received a primary vaccination series with one of the available COVID-19 vaccines (mRNA-vaccine by BioNTech or Moderna, or a vector-based vaccine by AstraZeneca or Janssen) [2]. The national COVID-19 vaccination campaign began at the end of 2020 by targeting older adults, residents of nursing homes, and healthcare workers, then shifting

* maierb@rki.de

focus to younger adults [3]. In August 2021, a recommendation to vaccinate adolescents aged 12-17 was issued [4] and since December 2021, children aged 5-11 years are recommended to receive a vaccination if underlying medical conditions put them at increased risk for severe disease [5]. In Germany, recovered individuals are advised not to receive a COVID-19 vaccination until 6 months [6], or 3 months [7] have passed after infection. At the time of analysis, booster vaccinations have been recommended for all persons aged 11 years and older [8, 9]. A central factor that will determine how the pandemic progresses in Germany in the near future is the number of people still immunologically naïve to infection, i.e. that have neither been in contact with the virus or any of its variants nor a vaccine against them. In Germany, several serological studies have been conducted [10, 11], but none that extend into the time of the Omicron waves, particularly with respect to children. Therefore, we choose a mathematical modeling approach here to estimate the number of immunologically naïve individuals in order to facilitate informed decisions with regard to the upcoming pandemic situation in the fall of 2022.

To estimate the number of people that have been in contact with either virus or vaccine, one might simply summate the number of vaccinations and the number of reported infections. However, doing so ignores the fact that (a) a considerable number of vaccinated people have suffered from additional breakthrough infections (taking into account both asymptomatic and symptomatic infections herein) [12], (b) a substantial number of previously infected people have chosen to be vaccinated in accordance with national recommendations [13–15], (c) some individuals have suffered from multiple infections [16], and (d) the exact extent of the total number of infections as compared to the reported number of infections is unknown because (i) asymptomatic infections are less likely to be identified and reported in the national surveillance system and (ii) underascertainment varies regionally [17, 18]. In order to estimate the overlap between the vaccinated and recovered subpopulations, one may assume that the probability of any recovered individual to be vaccinated is proportional to the probability of any individual to be vaccinated. However, this largely ignores (i) the heterogeneous dynamics of the spreading disease and vaccination campaigns, and (ii) that vaccinated individuals are less likely to suffer from an infection than unvaccinated individuals [19]. Here, we introduce modeling approaches that are devised to meet the aforementioned conditions and use them to estimate the distribution of immunologically naïve, (in the infectious disease modeling context called “fully susceptible” hereafter), recovered, and vaccinated individuals in Germany, taking into account regional and age differences. We find that although the percentage of the adult population in Germany that remains fully susceptible is ex-

pected to be in the single digits, the share of unaffected children may be considerably larger. Due to heterogeneities in vaccine uptake across German states, these values may differ by region. Our analysis cannot answer questions regarding the quality of achieved immunity against infection or disease, because we consider neither waning of immunity nor the emergence of variants with immune evasive properties, which is difficult to predict [20].

II. METHODS

We partition the population into $n_G = 16$ regions corresponding to the German states and $n_A = 5$ age groups corresponding to ages "00-04" (infants), "05-11" (children), "12-17" (adolescents), "18-59" (adults), "60+" (elderly), chosen in accordance with the population structure of publicly available vaccination data [2], i.e. into 80 subpopulations. To obtain nation-wide counts of individuals in age groups, we sum the respective results over all regions, to obtain counts of individuals for all ages, we sum over all age groups. To obtain an age-independent, nation-wide result, we sum over all ages and all regions.

As we are, first and foremost, interested in estimating the proportion of individuals $S_\infty \equiv S(t = t_{\max})$ that can be considered to be fully susceptible towards infection with any SARS-CoV-2 variant per region and age group, we report a simplified model here that captures the main ideas and gives the same results for $S(t)$ as the full model which is reported in the Appendix (see App. A 1).

We consider the population of size N (an age group in a region) to be composed of susceptible (S), infected/recovered (I), infected/recovered but eligible for reinfection or vaccination (Y), vaccinated (V), and boostered (B) individuals, assuming that the population count is constant over two years such that $N = S + I + Y + V + B = \text{const}$.

The central problem of estimating S_∞ is to determine the overlap between recovered and vaccinated subpopulations. Given that the cumulative number of unvaccinated infected R_∞ and the number of cumulative vaccinated individuals V_∞ is known, one may naively assume that the probability that an infected person that was initially unvaccinated is vaccinated later on is proportional to the probability that any person in the population is vaccinated, which is given as $p = V_\infty/N$. Then, the cohort size of unvaccinated and not yet infected individuals is $S_\infty = N - (1 - V_\infty/N)R_\infty - V_\infty$. However, this largely ignores the time course of infections and vaccinations, with incidence and daily vaccinations peaking at different time points, with a large number of infections occurring after the peak in vaccinations. Hence, one may assume instead that when a person becomes vacci-

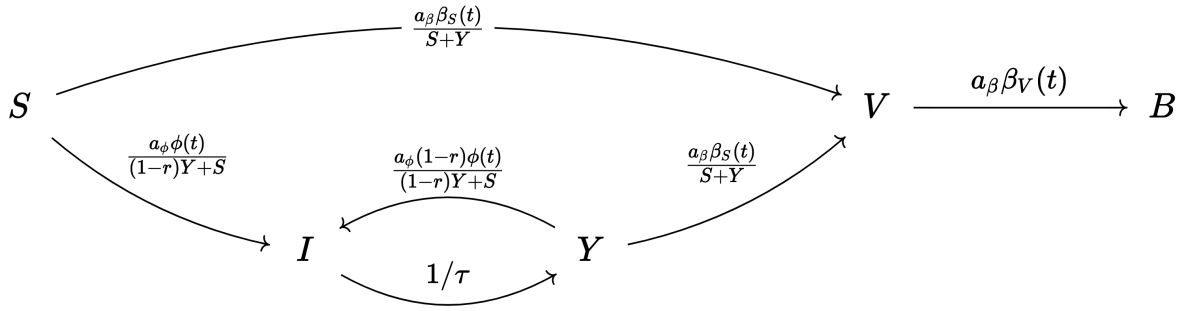

FIG. 1. Simplified model schema. On each day, $a_\beta \beta_S(t)\Delta t$ unvaccinated people become vaccinated, with under-ascertainment ratio a_β and $\Delta t = 1\text{d}$. The probability that a newly vaccinated person has been infected before is proportional to the respective size of the subpopulation of recovered people that are eligible for vaccination Y. Furthermore, on each day, $a_\phi \phi(t)\Delta t$ unvaccinated people become infected, with under-ascertainment ratio a_ϕ. The probability that a newly infected person has been infected before is proportional to the respective size of the subpopulation of recovered people that are eligible for reinfection $(1-r)Y$, where $1-r$ is the relative reinfection probability or "recovered immunity". Recovered individuals are expected to reach eligibility for reinfection/vaccination after an average duration of τ. (Note that in the full model breakthrough and reinfections of vaccinated individuals are possible (see App. A 1).)

nated at time t, the probability that this person was already infected is proportional to the number of infected/recovered individuals at time t that are eligible for vaccination as $p = Y/(S+Y)$. With incidence rates of $a_\phi \phi(t)$ (new unvaccinated cases per day) and vaccination rates of $a_\beta \beta_S(t)$ (new vaccinations per day) obtained from data, we assume that the count of individuals in the respective states evolves dynamically as

$$\partial_t S = -a_\phi \phi(t) \frac{S}{(1-r)Y+S} - a_\beta \beta_S(t) \frac{S}{Y+S} \tag{1}$$

$$\partial_t I = a_\phi \phi(t) - \frac{I}{\tau} \tag{2}$$

$$\partial_t Y = \frac{I}{\tau} - a_\beta \beta_S(t) \frac{Y}{Y+S} - a_\phi \phi(t) \frac{(1-r)Y}{(1-r)Y+S} \tag{3}$$

$$\partial_t V = a_\beta \beta_S(t) - a_\beta \beta_V(t) \tag{4}$$

$$\partial_t B = a_\beta \beta_V(t). \tag{5}$$

The last two equations are shown here for completeness, but note that the number of vaccinated and boostered individuals can simply be obtained from data, without integrating the dynamic equa-

tions, as their integrals can be evaluated analytically and are equal to the cumulative number of respective vaccinations. Above, a_ϕ and a_β are under-ascertainment ratios that account for infections and vaccinations that have not been reported. The time scale τ is equal to the average time after which an infected/recovered individual becomes eligible for reinfection or vaccination and $1-r$ is the relative probability that an unvaccinated recovered person is reinfected as compared to a fully susceptible individual.

For our analysis, we draw 1,000 pairs of a_ϕ and a_β from shifted Gamma distributions with means $\langle a_\phi \rangle = 2$, $\langle a_\beta \rangle = 1.03$, and standard deviations $\mathrm{Std}[a_\phi] = 1$, $\mathrm{Std}[a_\beta] = 0.02$ that are bounded below by $\min(a_\bullet) = 1$. Note that this distribution yields a median under-ascertainment ratio of $Q_2[a_\phi] = 1.7$, which is in line with results informed by seroprevalence data for Germany in 2020 [18]. Furthermore, with a 97.5th percentile of 4.7, the distribution is broad enough to account for occasional high under-ascertainment ratios that have been observed locally [10, 17, 18]. For infants, ascertainment is expected to be lower than for other age groups [21], which is why we double under-ascertainment ratios for this age group. We did not assume a higher under-ascertainment ratio for children older than 4 years, because regular screening via rapid antigen tests is mandatory in schools across the country [22]. We choose an eligibility time of $\tau = 90$d, which is approximately of the same order as the time for antibody concentrations to decay after an infection [23]. While it falls in the lower bound of officially recommended time for recovered individuals to wait before getting vaccinated, surveys indicate that people might not strictly follow the official recommendation but get vaccinated earlier. Further, people with asymptomatic courses might have no knowledge about their infection, likely leading to a bias towards shorter times between infection and vaccination in those cases. The influence of lower and higher values of τ is investigated in a sensitivity analysis. The "recovered immunity" parameter r quantifies the relative efficacy against reinfection. For the Alpha variant, this efficacy was observed to be lower than the vaccine efficacy against infection by mRNA- or vector-vaccines [24], but of similar order as the vaccine efficacy against infection with Delta, taking on values of $r \approx 0.65$ for both. As Omicron is considered to be a variant with partial immune escape, we set a lower default value of $r = 1/2$ for all variants, testing $r = 0$ (no protection against reinfection) and $r = 1$ (full immunity) in sensitivity analyses.

The daily vaccination rates $\beta_\bullet(t)$ are obtained from data [2] and averaged over calendar weeks to remove weekly modulations. Likewise, infection rates of unvaccinated individuals $\phi(t)$ are obtained from reported data in the German reporting system SurvStat [25], which is available in aggregated form upon request. While the vaccination status is unknown for a substantial number

of infections, we assume that for every day, the proportion of cases with unknown vaccination status that are, in fact, unvaccinated, is equal to the proportion of unvaccinated cases over the last seven days for which the vaccination status is known. This imputation method is performed for age- and region-stratified data.

For analyses disregarding infections with Omicron, we obtained the nation-wide and age-independent share of randomly sequenced samples in Germany [26] that the software framework "scorpio" identified as "Omicron" or "Probable Omicron" on a per-calendar-week basis by date of extraction ("Entnahmedatum") as $\sigma(t)$, assuming $\sigma(t) = 0$ for dates previous to Aug 1, 2021 and $\sigma(t) = 1$ for dates that exceed the last available date in the data. Then, all incidence rates were scaled as $\phi_{S,\text{pre-Omicron}}(t) = \phi_S(t)[1 - \sigma(t)]$. Note that vaccination rates are unaffected by this procedure.

Population sizes stratified by age and state were requested from destatis [27].

Eqs. (1)-(5) are integrated using Euler's method with Δt = 1d until the last day of available incidence/vaccination data. For dates where data is unavailable, we assume the respective rates are equal to zero.

III. RESULTS

We find an estimated nationwide median share of fully susceptible individuals of 7.0% (quartile range [3.6%–9.8%]). This result is, however, biased towards higher values due to a larger share of yet unaffected infants (44.6% [27.5%–56.8%]), children (22.5% [7.9%–34.3%]), and adolescents (5.0% [1.0%–10.1%]). For age groups that are associated with a higher probability of severe disease [28], we find a lower relative frequency of 3.5% [1.3%–5.5%] (adults), and 4.3% [2.7%–5.8%] (elderly).

These values are achieved largely due to the (at the time of analysis still ongoing) Omicron wave. Ignoring infections with the Omicron variant, the nationwide age-independent share of fully susceptibles increases to 16.1% [14.0%–17.8%], i.e. Omicron infections are expected to have caused a reduction in fully susceptible individuals on the order of 10 percentage points at the time of writing, though this number differs by age group. While the change in relative frequency of fully susceptibles in the "adult" and "elderly" age groups was only about a few percentage points (median decreases from 9.2% to 3.5% and from 6.6% to 4.3%, respectively), the three youngest age groups were affected much more strongly, with median values of fully susceptible

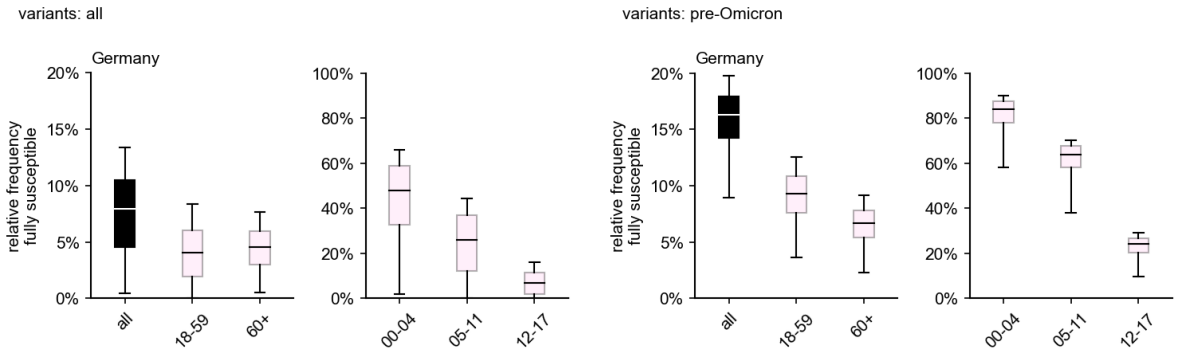

FIG. 2. Estimated nationwide relative frequency of fully susceptible individuals by age group, considering vaccinations and infections that took place up to and including March 2022. Boxes represent the area between quartiles Q_1, Q_3 and whiskers the 2.5th and 97.5th percentiles, respectively, the median is shown as a horizontal line. (Left) Considering infections with any variant. (Right) Considering infections with any variant other than Omicron and its sublineages.

individuals dropping from 83.3% to 44.6%, from 63.5% to 22.5%, and from 23.8% to 5.0% with increasing age (cf. Fig. 2). If all variants are considered, the median share of fully susceptible "adults" and "elderly" barely differ (absolute difference of 0.8% points), likely due to a larger fraction of Omicron-recovered "adults" (Fig. 2).

Although the relative frequency of fully susceptibles varies between federal state, certain commonalities are still shared. In all states, the frequency of fully susceptible individuals decreases with age, with a strong dependence on age for children. For ages 12-17, the frequency reaches values on the same order as those of the age groups "adults" and "elderly" (Fig. 3). Apart from the fact that adult and elderly age groups achieve relative frequencies of fully susceptible individuals below 10%, there are no other common patterns that stand out across all states regarding these age groups. In general, these age groups show overlapping quartile intervals, with the exception of Hamburg and Bremen, where "adults" show a comparatively lower relative frequency (Fig. 3). In fact, in Bremen virtually noone aged 18 and above is expected to not have been in contact with either virus or vaccine, according to the estimations.

In general, the above observations hold for the pre-Omicron analysis as well, except for the fact that, in the majority of states, the number of adults that were still unaffected decreased dramatically during the Omicron wave due to the large number of infections caused by the variant (comparing Figs. 4, 3). When excluding Omicron infections, the relative frequency of fully susceptibles differs

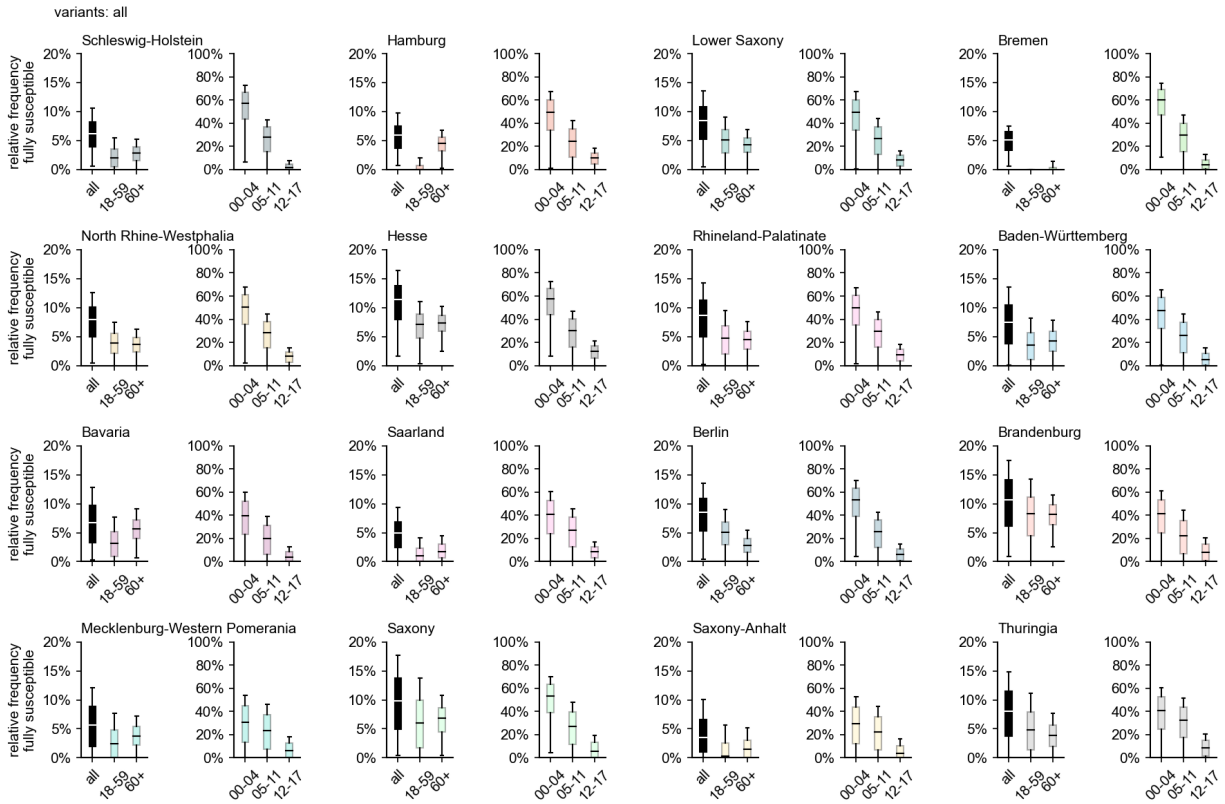

FIG. 3. Estimated relative frequency of fully susceptible individuals by age group and region considering infections with any variant and vaccinations up to and including the Omicron wave (as of March 31, 2022).

across states on the order of $\sim 10\%$, with Saxony and Bremen as the states with largest (20.3%) and smallest (10.0%) respective median values of fully susceptible individuals (Fig. 4). Including infections with Omicron, the median range between states is reduced to a difference of 6.0% points (median of 10.7% in Hesse and 4.7% in Bremen).

Our results are robust against changes in assumed eligibility time τ and recovered immunity r, varying by a few percentage points in the nationwide average for all ages. For the most at-risk age groups, i.e. adults and the elderly, these results vary even less, indicating that the influence of these parameters decreases with age (see Sec. B and Fig. 7).

Regarding the detailed distribution of individuals by vaccination/infection status, we find that the largest single compartment of the model population is the group of people that has received a booster vaccination and has never been in contact with the virus (see Sec. B and Fig. 6), with unvaccinated recovereds comprising the second largest group. When first excluding, then including Omicron infections, both the number of non-infected vaccinateds and non-infected booster vacci-

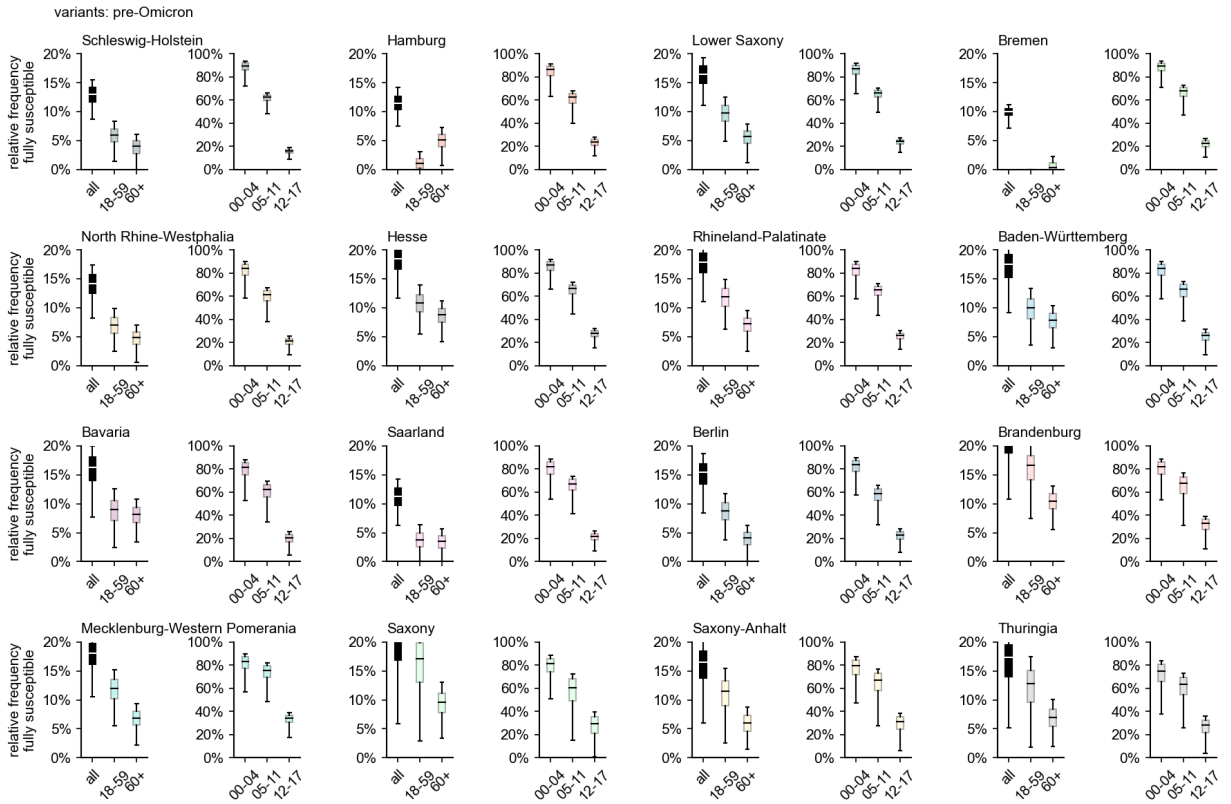

FIG. 4. Estimated relative frequency of fully susceptible individuals by age group and region, disregarding infections with Omicron and its sublineages, based on data available up to and including March 2022.

nateds decreases by about 10 percentage points, demonstrating the relative efficacy of the booster vaccination against infections with the Omicron variant. The prevalence of compartments that count infected individuals decreases with the number of (breakthrough) infections per individual, which is unsurprising given that the model probability to become infected decreases exponentially with every new infection. Note that our model cannot, however, track the number of reinfections per individual between achieving the different vaccination statuses.

As under-ascertainment is expected to be larger for infants than for other age groups, we scaled the respective under-ascertainment ratio to always assume twice the value of other age groups. Because most children below 5 years of age will remain unvaccinated as per official recommendations, only infections reduce the number of fully susceptible individuals, and, therefore, the under-ascertainment ratio has a large influence (see Sec. B and Fig. 8). With the degree of under-ascertainment in this age group comparatively unclear, the results must be considered relatively uncertain for this age group.

IV. CONCLUSION & DISCUSSION

As the pandemic progresses, a central quantity that will determine the upcoming dynamics is the population-wide susceptibility against infection with known or future variants of SARS-CoV-2. While protection from infection, either derived from vaccination or natural infection, wanes over time and depends on the circulating virus variant, an estimation of the respective subpopulation sizes of people that suffered from (one or more) infections or were vaccinated/boostered gives valuable information about the size of the population that is, as of yet, still fully susceptible to infection, because these individuals are more prone to infection and severe disease as compared to vaccinated or recovered individuals, given that future variants do not fully escape this immunity.

Here, we found that in Germany, a nationwide single-digit percentage of individuals have not been in contact with either a variant of SARS-CoV-2 nor a vaccine against them, yet these results vary between regions and age groups. Despite the high number of reported infections in infants, children, and adolescents, a considerably high percentage of these age groups may still be fully susceptible to infection. This may become problematic if a variant emerges that causes more severe disease in these age groups than previous variants. Yet, we cannot rule out the possibility that we underestimated the extent of under-ascertainment in these age groups, as the factors we used where informed by seroprevalence studies based on blood samples donated by adults (ages 18–74), while it has been reported that under-ascertainment ratios can assume values ranging from 2 up to 6 or 8 for children [29–31].

In comparison, the age groups of adults and elderly showed a relatively low share of fully susceptible individuals, considering infections with all variants, on the order of 5%. Only considering infections with pre-Omicron variants, however, around 7.4%–10.7% of the adult population and 5.3%–7.8% of the elderly population may still be at risk of infection with variants that have a higher probability of causing severe disease than Omicron, potentially causing large outbreaks that could put high pressure on the public health system once again (with these numbers representing quartile ranges).

Our results are subject to a number of limitations and biases. For instance, the reported uncertainties (quartile ranges) are heavily determined by the choice of distribution of a_ϕ. The distribution we chose has a median value of $Q_2[a_\phi] = 1.7$, which is slightly lower than what was observed in 2020 [18]. Moreover, the lower distribution bound of $\min(a_\phi) = 1$ might be rather low, as such a value would mean that every infection has been reported, which is unlikely. Hence, at least the

upper percentiles we report for S_∞ might be overestimations. Furthermore, we assume the same distribution of under-ascertainment ratios for all German states, which might not reflect potential heterogeneities in local ascertainment particularly well.

Regarding modeling choices for the eligibility time, a short average duration after infection to be eligible for vaccination leads to larger proportions of vaccine-eligible people and, hence, to a higher overlap between the vaccinated and recovered subpopulations, thus increasing the estimated number of fully susceptibles. While we chose a comparably low value of 90d for this parameter, lower values cannot be ruled out. However, (i) the value we chose lies below the official recommendation, and (ii) changes in this parameter are not expected to change our results drastically, as was shown in a sensitivity analysis.

Likewise, shorter durations of eligibility for reinfection and lower values of long-term immunity of recovered individuals increase the likelihood that a reported infection of an unvaccinated individual was, in fact, a reinfection event, thus leading to higher values of fully susceptible individuals over all. As above, our results are robust towards variations in these parameters.

Regarding results on a regional level, reported vaccinations and infections might be skewed regionally when a large number of people live in one state but traverse to others to seek medical help. These considerations might explain the extreme results observed for Hamburg and Bremen, which are city states enclosed by others.

The last German census took place in 2011 and population sizes per age group and region have been imputed for the year 2020 based on this data, thus potentially being subject to over- or under-counting. Uncertainties in population size may introduce systematic errors on the order of a few percentage points in relative frequencies. When such a relative frequency reaches low values, these absolute errors on the order of a few percentage points can lead to high relative errors in the results.

Considering incidence rates, we imputed the total number of unvaccinated cases per day from cases with undetermined vaccination status by assigning them the "unvaccinated" status with probability proportional to the share of unvaccinated cases in the set of cases with determined status. This procedure can introduce systematic errors when the ascertainment of vaccination status is biased towards any of the vaccination states, which may occur, for instance, when the probability of status ascertainment increases with severity of disease. In this case, people with breakthrough infections may be less likely to have their vaccination status reported in the reporting system, which would mean that we overestimated the number of unvaccinated cases per day, introducing a bias

241 towards lower values of the share of fully susceptible individuals.

242 For analyses regarding infections with variants prior to Omicron, we relied on the nationwide
243 share of Omicron sequences, multiplying all incidence rates (regardless of region, age, or vaccine
244 status) with this function. Since vaccines assume different efficacies against infection with differ-
245 ent variants and will likely vary across ages and regions, this assumption is expected to introduce
246 strong bias on a fine-grained population level, which may be expected to decrease when values are
247 aggregated over regions or ages.

248 Our results cannot be used to predict the future course of the pandemic directly. In fact, since
249 SARS-CoV-2 lacks phenotypical stability and neither infection nor vaccination elicit full long-
250 term protective immunity, especially with respect to the prevention of infection and transmission,
251 there are doubts that classical herd immunity can be reached for COVID-19 [32]. In several stud-
252 ies, hybrid immunity resulting from infection-acquired immunity boosted with vaccination con-
253 ferred the strongest, or longer-lasting protection, respectively [33, 34]. Similarly, Omicron break-
254 through infections in previously vaccinated individuals have been shown to drive cross-variant
255 neutralization and memory B cell formation [35], suggesting that a combination of both, natural
256 infection and vaccination, will have more impact on the future COVID-19 epidemiology than one
257 of the events alone.

258 To sum up, our study shows that, presumably, only a small part of the German population has
259 not yet been in contact with either a variant of SARS-CoV-2 or a respective vaccine against the
260 disease they cause, up to and including March 2022. We show important proportions of fully
261 susceptible elderly, who on average, by their age and age-associated morbidities, have a dispro-
262 portionately elevated risk of severe disease. These shares differ by region and could motivate
263 regionally targeted protection measures at the time of writing or in case of future outbreaks.

264 While the immunization campaign was successful in spring and summer 2021, in particular
265 reaching a large proportion of vulnerable people, it thereafter had difficulties to completely close
266 immunity gaps with vaccinations, albeit enhancing the protection of a large proportion of already
267 vaccinated people with a large booster vaccination campaign by the end of 2021. Our results
268 show that the Omicron wave had a high impact on naturally closing the aforementioned gaps. As
269 mentioned above, however, having been in contact with a variant of SARS-CoV-2 is not a robust
270 equivalent of immunity and may range from mild infection followed by rapid waning of antibodies
271 and a highly uncertain degree of immunity, to a fully vaccinated status including a booster and a
272 breakthrough infection, which confers a more long-lasting and robust degree of protection against

severe disease. At the lower end of this spectrum of presumed immunity, our analyses show that one in six persons was never vaccinated but infected once or more, in the majority of cases with Omicron. This group faces higher uncertainties for the upcoming fall and winter since protection against severe disease may be more short-lived and too narrowly targeted to this variant.

ACKNOWLEDGMENTS

The authors would like to thank Maria Waize and Matthias an der Heiden for helpful discussions. B.F.M. is supported as an *Add-on Fellow for Interdisciplinary Life Science* by the Joachim Herz Stiftung.

Appendix A: Main model

1. Model formulation

We partition the population into $n_G = 16$ regions corresponding to the German states and $n_A = 5$ age groups corresponding to ages "00-04" (infants), "05-11" (children), "12-17" (adolescents), "18-59" (adults), "60+" (elderly), chosen in accordance with the population structure of publicly available vaccination data [2]. Consequently, for any region- and age-specific compartment $X_{A,G}$, the nation-wide value is given as

$$X_A = \sum_{G=1}^{n_G} X_{A,G}, \tag{A1}$$

the corresponding value for all ages is given as

$$X_G = \sum_{A=1}^{n_A} X_{A,G}, \tag{A2}$$

and the total value is

$$X_{\text{tot}} = \sum_{A=1}^{n_A} \sum_{G=1}^{n_G} X_{A,G}. \tag{A3}$$

Because in the further analysis, none of the subpopulations are interacting, we will omit the region- and age-determining subscripts for simplicity.

For any population of size N, we are first and foremost interested in the number of susceptible individuals S, i.e. the number of individuals that have never been in contact with neither a variant

of SARS-CoV-2, nor a vaccine against it. We assume that previous to the pandemic, no individual has had contact with any variant of SARS-CoV-2 or a vaccine against them, i.e. $S(t=0)=N$. These susceptibles can then either (i) become infected (changing their status to I) or (ii) vaccinated (changing their status to V). The number of individuals changing their status per day is estimated from official data [2, 36], defining the number of reported newly infected unvaccinated individuals per day as ϕ_S and the number of newly vaccinated individuals per day as $\beta_S(t)$. We obtain these rates on a calendar-week basis in order to remove weekly modulations. Because the vaccination status of new infections is unknown for a considerable amount of people, we impute ϕ_S from incomplete incidence data in a procedure outlined further below. The rates are to be interpreted in a way such that

$$M_S = \int_0^{t_{\max}} dt\, \beta_S(t), \qquad \text{and} \tag{A4}$$

$$F_S = \int_0^{t_{\max}} dt\, \phi_S(t) \tag{A5}$$

give the cumulative number of vaccinated individuals and the cumulative number of reported infections of unvaccinated individuals, respectively, both up to time $t_{\max}$.

At any time t, the number of individuals eligible to receive a vaccine is proportional to (a) the number of susceptible individuals and (b) the number of recovered individuals . We assume that infected individuals become eligible for vaccination after an average amount of time τ passes. Hence, after obtaining an infection, we assume that individuals change their status with rate $1/\tau$ to become eligible (status Y). Then, the probability for a person that becomes vaccinated at time t to be of status S is given as $p_{V,S} = S/(S+Y)$ and for status Y as $p_{V,I} = Y/(S+Y)$. Consequently, the vaccination transition rate for both susceptibles and eligible recovereds to receive vaccination status is given as

$$\tilde{\beta}_S = \frac{a_\beta \beta_S}{S+Y}. \tag{A6}$$

Here, we further introduced the under-ascertainment ratio of vaccinations a_β. The corresponding transition processes are

$$S \xrightarrow{\tilde{\beta}_S} V \tag{A7}$$

$$Y \xrightarrow{\tilde{\beta}_S} C_{IV} \tag{A8}$$

where C_{IV} represents the compartment counting individuals who became infected at least once before receiving a vaccination.

$$I \xrightarrow{1/\tau} Y \tag{A9}$$

represents the process of recovered individuals becoming eligible for vaccination.

Similarly, the number of individuals eligible to transition to status "unvaccinated infected" is proportional to (a) the number of susceptible individuals and (b) the number of recovered individuals that are eligible for reinfection. We assume that individuals that recently suffered from an infection are fully immune, but may return to (partial) susceptibility after an average duration of τ, equating this to the average duration it takes to become eligible for vaccination for model parsimony and reasons outlined further below. Because reinfections are not registered in the German reporting system, we have to consider the relative probability for a recovered person to be reinfected by introducing an "immunity parameter" r that represents the relative probability of a recovered person to become infected after time τ since the last infection as compared to a fully susceptible person. Hence, the total number of people eligible to be counted as an infection of an unvaccinated individual at time t is given as $S+(1-r)Y$, the probability that an unvaccinated person that becomes infected at time t has been infected before is $p_{I,I} = (1-r)Y/(S+(1-r)Y)$, and $p_{I,S} = S/(S+(1-r)Y)$ that they have been fully susceptible. Consequently, the eligibility-corrected vaccination rate is given as

$$\tilde{\phi}_S = \frac{a_\phi \phi_S}{S+(1-r)Y}. \tag{A10}$$

Here, a_ϕ is the under-ascertainment ratio, accounting for infections that have not been reported. The corresponding transition processes are

$$S \xrightarrow{\tilde{\phi}_S} I \tag{A11}$$

$$Y \xrightarrow{(1-r)\tilde{\phi}_S} I. \tag{A12}$$

Again, Eq. (A9) represents the process of becoming eligible (both for vaccination after infection and reinfection).

Continuing with this line of argumentation, we further consider the adjusted rate of individuals that obtain a breakthrough infection as

$$\tilde{\phi}_V = \frac{\phi_V}{V+(1-r)C_{VY}+C_{IV}+(1-r)C_{IVY}}. \tag{A13}$$

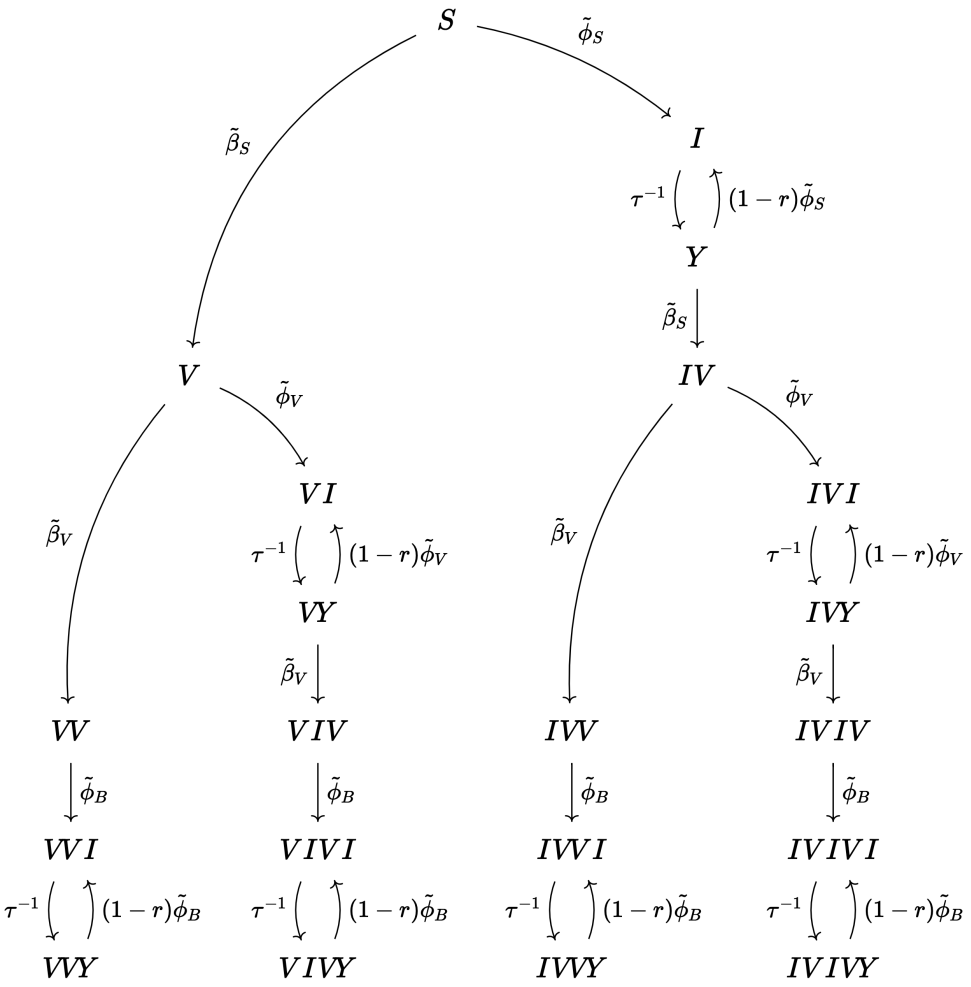

FIG. 5. Vaccination/infection model given by Eqs. (A6)-(A36). Individuals can become infected and recover (compartments ending in I), vaccinated (compartments ending in V), or eligible for reinfection/vaccination after a previous infection after an average duration of τ^{-1} (compartments ending in Y). Initially, all individuals are susceptible (S). Transition rates are determined by data and scaled by assumed under-ascertainment ratios (not shown here). Individuals that are eligible for reinfection are associated with a relative reduction in susceptibility r. The order of I and V in individual statuses represent the order in which infections and vaccinations happened to the respective individuals.

Here, C_{VY} are vaccinated individuals that suffered from a breakthrough infection before, and C_{IVY} counts individuals that, after recovery became vaccinated, then suffered from a breakthrough infection again. The respective transition processes are displayed in Fig. 5.

Similarly, the adjusted booster rate

$$\tilde{\beta}_V = \frac{\beta_V}{V + C_{VY} + C_{IV}} \tag{A14}$$

quantifies the rate with which previously vaccinated individuals receive a booster vaccination (processes shown in Fig. 5).

Finally, the adjusted booster breakthrough rate is

$$\tilde{\phi}_B = \frac{\phi_B}{C_{VV} + C_{VIV} + C_{IVV} + C_{IVIV} + (1-r)\left[C_{VVY} + C_{VIVY} + C_{IVVY} + C_{IVIVY}\right]}. \tag{A15}$$

For every compartment $C_\bullet$, the order of I and V in the subscript $\bullet$ represents the order in which infections and vaccinations happened to the individuals counted in the respective compartment.

In total, the model is determined by the following set of ordinary differential equations (ODEs)

$$\partial_t S = -\tilde{\phi}_S S - \tilde{\beta}_S S \tag{A16}$$

$$\partial_t V = \tilde{\phi}_S S - \tilde{\beta}_V V - \tilde{\phi}_V \tag{A17}$$

$$\partial_t I = \tilde{\beta}_S S + (1-r)\tilde{\phi}_S Y - I/\tau \tag{A18}$$

$$\partial_t Y = I/\tau - (1-r)\tilde{\phi}_S Y - \tilde{\beta}_S Y \tag{A19}$$

$$\partial_t C_{IV} = \tilde{\beta}_S Y - \tilde{\beta}_V C_{IV} - \tilde{\phi}_V C_{IV} \tag{A20}$$

$$\partial_t C_{VI} = \tilde{\phi}_V V + (1-r)\tilde{\phi}_V C_{VY} - C_{VI}/\tau \tag{A21}$$

$$\partial_t C_{VY} = C_{VI}/\tau - (1-r)\tilde{\phi}_V C_{VY} - \tilde{\beta}_V C_{VY} \tag{A22}$$

$$\partial_t C_{IVI} = \tilde{\phi}_V C_{IV} + (1-r)\tilde{\phi}_V C_{IVY} - C_{IVI}/\tau \tag{A23}$$

$$\partial_t C_{IVY} = C_{IVI}/\tau - (1-r)\tilde{\phi}_V C_{IVY} - \tilde{\beta}_V C_{IVY} \tag{A24}$$

$$\partial_t C_{VV} = \tilde{\beta}_V V - \tilde{\phi}_B C_{VV} \tag{A25}$$

$$\partial_t C_{VIV} = \tilde{\beta}_V C_{VY} - \tilde{\phi}_B C_{VIV} \tag{A26}$$

$$\partial_t C_{IVV} = \tilde{\beta}_V C_{IV} - \tilde{\phi}_B C_{IVVI} \tag{A27}$$

$$\partial_t C_{IVIV} = \tilde{\beta}_V C_{IVY} - \tilde{\phi}_B C_{IVIV} \tag{A28}$$

$$\partial_t C_{VVI} = \tilde{\beta}_B C_{VV} + (1-r)\tilde{\phi}_B C_{VVY} - C_{VVI}/\tau \tag{A29}$$

$$\partial_t C_{VIVI} = \tilde{\beta}_B C_{VIV} + (1-r)\tilde{\phi}_B C_{VIVY} - C_{VIVI}/\tau \tag{A30}$$

$$\partial_t C_{IVVI} = \tilde{\beta}_B C_{IVVI} + (1-r)\tilde{\phi}_B C_{IVVY} - C_{IVVI}/\tau \tag{A31}$$

$$\partial_t C_{IVIVI} = \tilde{\beta}_B C_{IVIV} + (1-r)\tilde{\phi}_B C_{IVIVY} - C_{IVIVI}/\tau \tag{A32}$$

$$\partial_t C_{VVY} = -(1-r)\tilde{\phi}_B C_{VVY} + C_{VVI}/\tau \tag{A33}$$

$$\partial_t C_{VIVY} = -(1-r)\tilde{\phi}_B C_{VIVY} + C_{VIVI}/\tau \tag{A34}$$

$$\partial_t C_{IVVY} = -(1-r)\tilde{\phi}_B C_{IVVY} + C_{IVVI}/\tau \tag{A35}$$

$$\partial_t C_{IVIVY} = -(1-r)\tilde{\phi}_B C_{IVIVY} + C_{IVIVI}/\tau. \tag{A36}$$

2. Parameters and data

a. Incidence by vaccination status

For each combination of age group and region, we obtain the daily number of reported new cases in unvaccinated $\hat{n}_S(t)$ by "Meldedatum" (date of report), as well as the daily number of

reported breakthrough infections $\hat{n}_V(t)$, reported booster breakthrough infections $\hat{n}_B(t)$, as well as the daily number of infections where the vaccination status is unknown $\hat{n}_\varnothing(t)$ from the German reporting system SurvStat [25]. In order to assign vaccination statuses to cases where the status is originally unknown, we measure the proportion of infections per status in cases with known status in the last seven days and subsequently obtain the imputed number of daily cases as

$$n_X(t) = \hat{n}_X(t) + \hat{n}_\varnothing(t) \frac{\sum_{t'=t-6\mathrm{d}}^{t} \hat{n}_X(t')}{\sum_{t'=t-6\mathrm{d}}^{t} \left[\hat{n}_S(t') + \hat{n}_V(t') + \hat{n}_B(t')\right]}, \qquad \forall X \in \{S, V, B\}. \tag{A37}$$

This procedure removes weekly modulations for the imputation. It might be biased towards any of the statuses S, V, B due to different probabilities of severe disease by vaccination status and thus of being reported in a system of primarily symptom-based testing. Note that, for no region and age groups there were days for which $\aleph = \sum_{t'=t-6\mathrm{d}}^{t} \left[\hat{n}_S(t') + \hat{n}_V(t') + \hat{n}_B(t')\right] = 0$ and $\hat{n}_\varnothing(t) > 0$, which is why we set $n_X(t) = \hat{n}_X(t)$ on days where $\aleph = 0$. With the above definition, the infection rates are given as

$$\phi_X(t) = \frac{1}{|\mathcal{W}(t)|} \sum_{t' \in \mathcal{W}(t)} n_X(t'), \qquad \forall X \in \{S, V, B\} \tag{A38}$$

where $\mathcal{W}(t)$ is the set of days t' in calendar week of day t meeting $t' < t_{\max}$.

b. Vaccination rates

Similarly, weekly vaccination rates are given as

$$\beta_X(t) = \frac{1}{|\mathcal{W}(t)|} \sum_{t' \in \mathcal{W}(t)} \hat{v}_X(t'), \qquad \forall X \in \{S, V\} \tag{A39}$$

with $\hat{v}_S(t)$ and $\hat{v}_V(t)$ being the number of new vaccinations (new booster vaccinations, respectively) on day t. We define "new vaccinations" as entries in the data provided in [2] that have an "Impfschutz"-field value of "2", and as "new booster vaccinations" as entries that have an "Impfschutz"-field value of "3", ignoring single-shot vaccinations with value "1" (in the data, confirmed recovered individuals that received a single vector- or mRNA-vaccine dose are counted as being fully vaccinated with an "Impfschutz"-field value of "2"). The share of the population that received only one dose of an mRNA or the Vaxzevria vaccine is expected to be on the order of 1% of the German population up to and including March 2022 [2]. In the model, the infection of these individuals follows the same dynamics as the infection of fully susceptible individuals. Hence, ignoring this vaccination state will barely affect the results.

Note that we ignore the small number of vaccinations associated with the region "Bund" (region id "17").

c. *Under-ascertainment*

Based on seroprevalence data collected over the first waves in Germany, a nation-wide under-ascertainment ratio of $a_\phi \approx 2$ was found, with regional variations that went up to a factor of $a_\phi \approx 5$ in regions of large outbreaks [10, 18]. In absence of more fine-grained and temporally resolved estimations, we assume an under-ascertainment of $a_\phi = 1 + \hat{a}_\phi$ with $\hat{a}_\phi$ being a Gamma-distributed random variable such that $\langle a_\phi \rangle = 2$ and $\mathrm{Std}[a_\phi] = 1$.

It has further been reported that there might be low under-ascertainment in vaccinations [37]. We assume an under-ascertainment of $a_\beta = 1 + \hat{a}_\beta$ with $\hat{a}_\beta$ being a Gamma-distributed random variable such that $\langle a_\beta \rangle = 1.03$ and $\mathrm{Std}[a_\beta] = 0.02$.

Infants are less likely to display symptoms when infected and are not subject to the strict testing strategies applied in schools [38]. A lower ascertainment in this age group is, therefore, a plausible assumption. We hence assume double the value of the under-ascertainment ratio for this age group.

d. *Eligibility time and immunity of recovered individuals*

We assume an average eligibility time of $\tau = 90$d for vaccination after infection or reinfection. Regarding reinfection, this is a reasonable time scale, as it is of the order of the mean duration neutralising antibodies can be found after an infection. For vaccinations, the official assumption for receiving a vaccine after infection has been 3–6 months. In non-representative survey data, it was found that participants generally followed these recommendations, with a large number of participants waiting less and became vaccinated about 3 months after a confirmed infection. While the cohort of this study is assumed to be composed of highly compliant individuals, the average time to receive a vaccination is also lowered assuming a large number of asymptomatic infections, where the date of the infection might be unknown to recovered individuals themselves. Note, however, that we test the influence of this parameter on our results in a sensitivity analysis (see App. B).

We recognize that recovered individuals might still have a lowered susceptibility for reinfection even after transitioning to the eligibility state. The "recovered immunity" parameter r quantifies

the relative efficacy against reinfection. For the Alpha variant, this efficacy was observed to be lower than the vaccine efficacy against infection by mRNA- or vector-vaccines [24], but of similar order as the vaccine efficacy against Infection with Delta, taking on values of $r \approx 0.65$ for both. As Omicron is considered to be a variant with partial immune escape, we set a lower default value of $r = 1/2$ for all variants, testing $r = 0$ (no protection against reinfection) and $r = 1$ (full immunity) in sensitivity analyses.

e. Variant share

For analyses disregarding infections with Omicron, we obtained sequences that were sampled randomly nation-wide and independent of age [26]. For each calendar week w we obtained the total number $m(w)$ of randomly sampled sequences with date of extraction t that lie in w. We further aggregated the number $m_o(w)$ of randomly sampled sequences that the software framework "scorpio" identified as "Omicron" or "Probable Omicron". Then, the share of Omicron on day t is given as

$$\sigma(t) = \begin{cases} 0, & t < \text{Aug 1, 2021} \\ 1, & w(t) > w_{\max} \\ m_o(w(t))/m(w(t)) & \text{otherwise,} \end{cases} \tag{A40}$$

with $w_{\max}$ being the last week for which data was available.

For analyses labeled "pre-Omicron" we analyzed the model with all incidence rates being scaled as $\phi_{\bullet,\text{pre-Omicron}}(t) = \phi_\bullet(t)[1 - \sigma(t)]$.

f. Simulations

We draw 1,000 pairs of (a_ϕ, a_β) as described above and assume those under-ascertainment ratios to be constant across all respective ages and regions (bar infants, whose under-ascertainment ratio is set as $a_{\phi,\text{infants}} = \omega a_\phi$ with $\omega = 2$ to account for the fact that under-ascertainment is expected to be higher in this age group). Then, Eqs. (A16)–(A36) are integrated with Euler's method using a time step of $\Delta t = 1$d, starting on Jan 6, 2020 until March 31, 2022. We then obtain the final state

of the compartments, and additionally aggregated states as

$$C_{I^*} = I + Y \tag{A41}$$

$$C_{VI^*} = C_{VI} + C_{VY} \tag{A42}$$

$$C_{IVI^*} = C_{IVI} + C_{IVY} \tag{A43}$$

$$C_{VVI^*} = C_{VVI} + C_{VVY} \tag{A44}$$

$$C_{VIVI^*} = C_{VIVI} + C_{VIVY} \tag{A45}$$

$$C_{IVVI^*} = C_{IVVI} + C_{IVVY} \tag{A46}$$

$$C_{IVIVI^*} = C_{IVIVI} + C_{IVIVY} \tag{A47}$$

$$C_{0V1I} = I + Y \tag{A48}$$

$$C_{1V0I} = V \tag{A49}$$

$$C_{1V1I} = C_{IV} + C_{VI} + C_{VY} \tag{A50}$$

$$C_{1V2I} = C_{IVI} + C_{IVY} \tag{A51}$$

$$C_{2V0I} = C_{VV} \tag{A52}$$

$$C_{2V1I} = C_{VIV} + C_{IVV} + C_{VVI} + C_{VVY} \tag{A53}$$

$$C_{2V2I} = C_{IVIV} + C_{VIVI} + C_{IVVI} + C_{VIVY} + C_{IVVY} \tag{A54}$$

$$C_{2V3I} = C_{IVIVI} + C_{IVIVY} \tag{A55}$$

$$C_{1V} = V + C_{IV} + C_{VI} + C_{IVI} + C_{VY} + C_{IVY} \tag{A56}$$

$$C_{2V} = C_{VV} + C_{VIV} + C_{IVV} + C_{IVIV} + C_{VVI} + C_{VIVI} + \tag{A57}$$

$$+ C_{IVVI} + C_{IVIVI} + C_{VVY} + C_{VIVY} + C_{IVVY} + C_{IVIVY} \tag{A58}$$

$$C_{1I} = I + Y + C_{IV} + C_{VI} + C_{VY} + C_{VIV} + C_{IVV} + C_{VVI} + C_{VVY} \tag{A59}$$

$$C_{2I} = C_{IVI} + C_{IVY} + C_{IVIV} + C_{VIVI} + C_{IVVI} + C_{VIVY} + C_{IVVY} \tag{A60}$$

$$C_{3I} = C_{IVIVI} + C_{IVIVY}. \tag{A61}$$

These states combine compartments that have certain commonalities, e.g. compartments C_{nVmI} is the number of individuals that were vaccinated n times and infected m times (re-infections excluded), C_{nV} is the number of individuals that were vaccinated n times, and C_{mI} is the number of individuals that were infected m times (re-infections excluded, which means that if an individual was infected $m = 3$ times, they must have been infected before, between, and after the respective inoculations.

We test how robust our results are if per region and age group, individual pairs (a_ϕ, a_β) were drawn from their respective distribution, i.e. assuming heterogeneous under-ascertainment in ages and regions per simulation run, which could potentially change the width of the distribution of respective aggregated values, finding that it does not have a substantial effect.

The results of these simulations can be obtained from [39].

Appendix B: Sensitivity and other analyses

Nation-wide results for all compartments as well as Eqs. ((A41)–(A61)) can be found in Fig. 6. The compartment with the largest share of the population is C_{VV}, i.e. boostered and never infected, assuming a value of 45.8% [41.1%–49.0%]. Considering all variants, the second largest value can be found for individuals that have never been vaccinated but infected once or more with C_{I^*} assuming 14.9% [12.0%–18.1%]. This value is considerably lower (5.6% [4.3%–7.5%]) when infections with Omicron are excluded. Likewise, the share of vaccinated, yet non-infected individuals V is estimated to assume 14.6% [13.4%–15.3%] with Omicron infections excluded, but 8.6% [5.8%–10.6%] considering all variants. With Omicron infections excluded, the boostered and non-infected population assumes an estimated size of 54.9% [53.1%–56.2%], demonstrating the increased efficacy of the booster vaccination against infection with Omicron as compared to individuals who only finished the first vaccination series.

Regarding the influence of eligibility time, higher values lead to a lower probability of reinfections and vaccinations of recovereds during the most active period of the vaccination campaign, implying the estimated number of fully susceptible individuals decreases with increasing τ. Likewise, the assumed immunity of recovereds r leads to a decreasing value of fully susceptible individuals. The results we reported above lie central within the range of results for extreme value pairs of $\tau = 30$d, $r = 0$ (low), as well as $\tau = 150$d, $r = 1$ (high). For instance for all ages, the results vary between median values of 9.5% (low) and 4.6% (high) with our reported result in the main text ($\tau = 90$d, $r = 0.5$) being equal to 7.0%. The influence of these parameters are higher for the younger population with a "low"-to-"high" variation leading to respective median ranges of 51.0% to 38.8% (infants), 31.5% to 14.0% (children), and 10.3% to 1.3% (adolescents). In the older population, the influence of these parameters is rather small, leading to median ranges of 5.4% to 1.5% (adults) and 5.1% to 3.3% (elderly). These results are displayed in Fig. 7.

In the main text, we assumed that the relative under-ascertainment factor in infants assume a

value of $a_{\phi,\text{infants}}/a_\phi = \omega = 2$. For $\omega = 1$, fully susceptible infants is higher than what we reported in the main text (see Fig. 8. Since empirical values for ω are difficult to obtain, we are probably underestimating the uncertainty in our results for infants.

Appendix C: Additional, sophisticated Model

We further want to develop a model that allows waning to be included in the analyses and could therefore potentially be used to estimate seroprevalence in future studies.

We hypothesize that exposure to either the pathogen or a vaccine results in an initial immune response that then decays over a period of time and account for this by introducing intermediate compartments representing different gradations of immunity.

We define as S susceptibles, I infected, V vaccinated, Y breakthroughs from vaccinated V and U as breakthroughs from boostered B. For each compartment X, we consider $n_X + 1$ gradations, i.e. we assume that individuals who reach the status X pass through intermediate compartments in the form of a chain from initial X_0 to final $X_{n,X}$, per transition $X_i \longrightarrow X_{i+1}$ with transition rate $1/\tau_{X,i+1}$. This means that for each individual, each of these transitions is subject to a random delay

$$T_{X,i} \sim \text{Exp}(1/\tau_{X,i+1}) \tag{C1}$$

where $\text{Exp}(\lambda_X)$ is an exponential distribution with mean λ_X^{-1}. This approach allows us to more accurately model both waning of immunity and the timing of vaccination or breakthrough infection. For susceptibles, we set $n_S = 0$, i.e. no transitions and exactly one gradation.

We denote $\hat{X}$ as the total number of individuals in status X that are susceptible to infection. That is, we define

$$\hat{X} = \sum_{i=0}^{n_X} (1 - e_{X,i}) X_i, \tag{C2}$$

where $e_{X,i}$ is the susceptibility reduction of a person in status X_i (due to previous infection or vaccination).

We define $\tilde{X}$ as the total number of individuals in status X who can receive one or the next vaccination. Usually, this is the case after a defined time Θ_X has passed since the last infection or the last receipt of a vaccine dose (comparable to the 'eligibility time' used in the main analyses of this study). The total time it takes for an individual in status X_i to reach status X_{i+1} is given by the

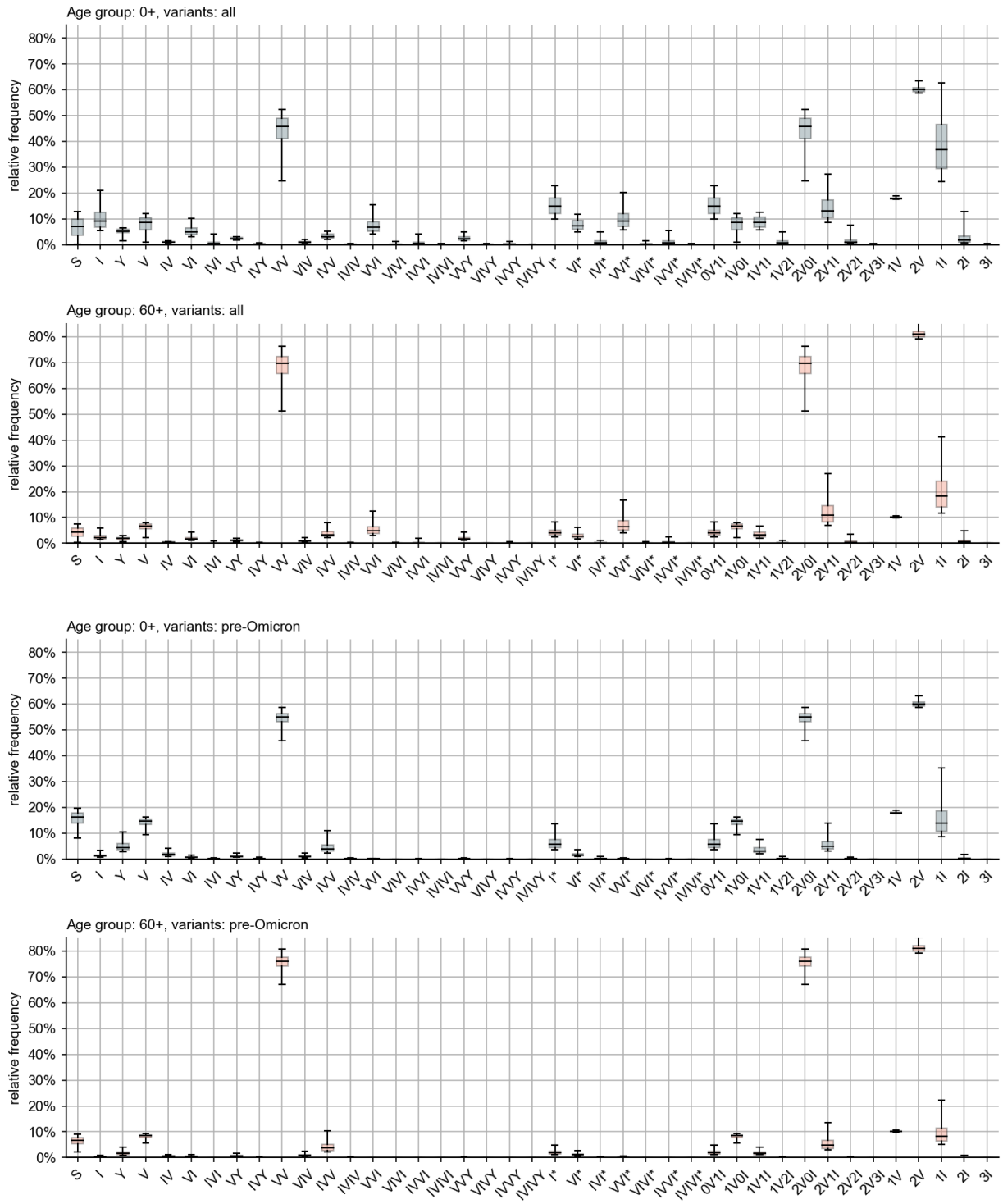

FIG. 6. Relative frequency of all compartments given by vaccination and infection status across Germany, for all age groups and variants as well as for the elderly and pre-Omicron variants. Some compartments shown are aggregates, e.g. labels "nVmI" represent the number of individuals that were vaccinated n times and infected m times (re-infections excluded), labels "nV" give the number of individuals that were vaccinated n times, and labels "mI" are the number of individuals that were infected m times (re-infections excluded), see Eqs. ((A41))-((A61))

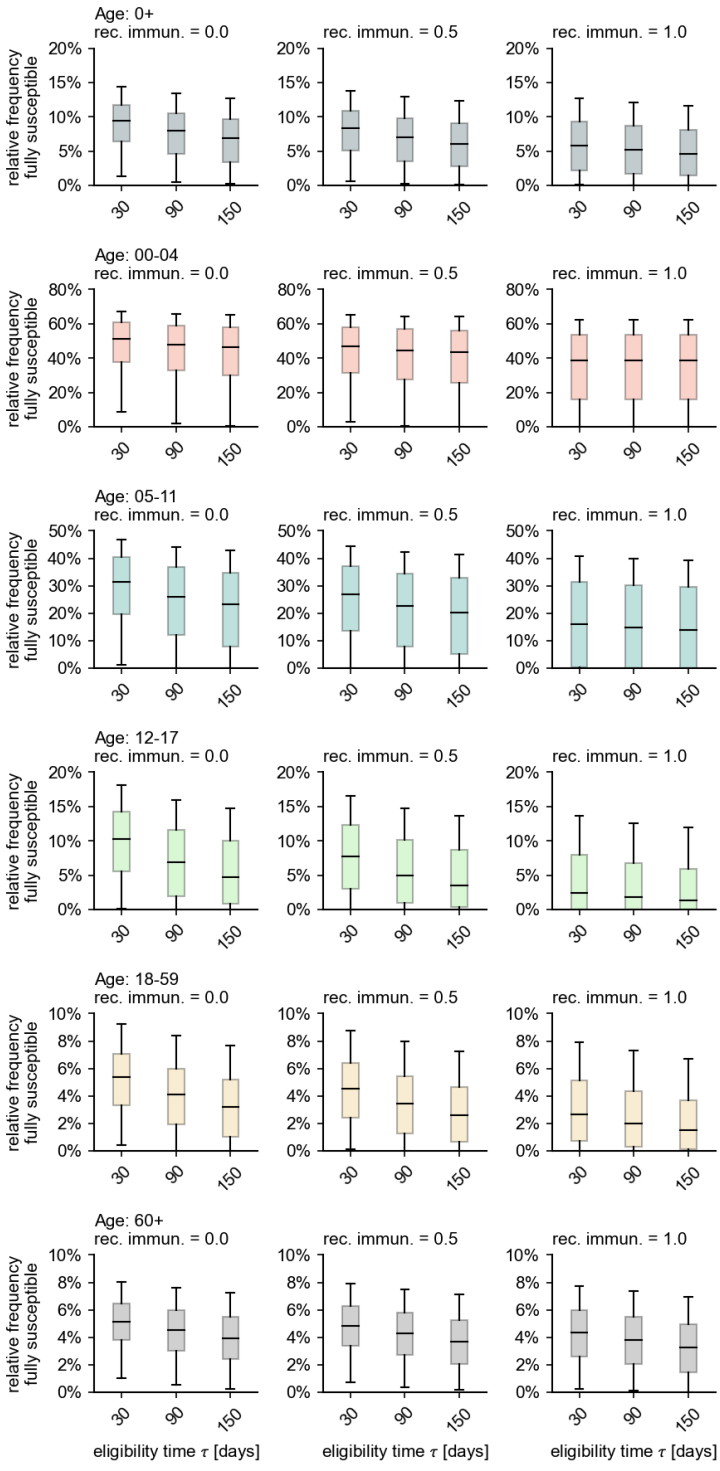

FIG. 7. The influence of the assumed average eligibility duration as well as the long-term immunity of recovered individuals.

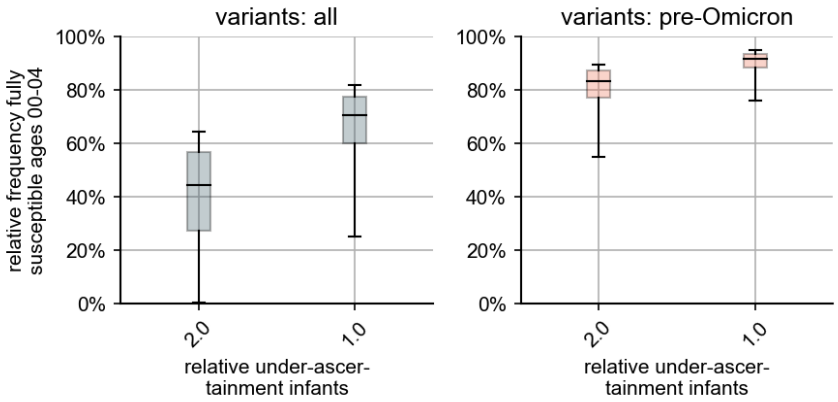

FIG. 8. Influence of relative under-ascertainment for infants. For the main results, we assumed that the relative under-ascertainment factor assumes, for infants, a value of $a_{\phi,\text{infants}}/a_\phi = \omega = 2$. For $\omega = 1$, the number of yet fully susceptible infants is higher than what we reported in the main text.

random variable

$$Z_{X,i} = \sum_{j=0}^{i} T_{X,j}. \tag{C3}$$

Let $F_{X,i}(z)$ be the cumulative distribution function of the random variable $Z_{X,i}$. Then, the probability $w_{X,i}$ that a given individual in status X_i has been in status X for longer than Θ_X is given by

$$w_{X,i} = P(Z_{X,i} > \Theta_X) = 1 - F_{X,i}(\Theta_X). \tag{C4}$$

We find such

$$\tilde{X} = \sum_{i=0}^{n_X} \left[1 - F_{X,i}(\Theta_X)\right] X_i. \tag{C5}$$

The probabilities $w_{X,i} = 1 - F_{X,i}(\Theta_X)$ are constant and can thus be determined numerically after defining the times $\{\tau_{X,i}\}$ and Θ_X. For susceptibles, let $S = \hat{S} = \tilde{S}$.

Let $\mathcal{I}(X)$ be the compartment to which an individual in status X transitions after infection and $\mathcal{V}(X)$ be the compartment to which an individual in status X transitions after vaccination. We define the following transitions

$$\mathcal{I}(S) = \mathcal{I}(I) = I \tag{C6}$$

$$\mathcal{V}(S) = \mathcal{V}(I) = V, \tag{C7}$$

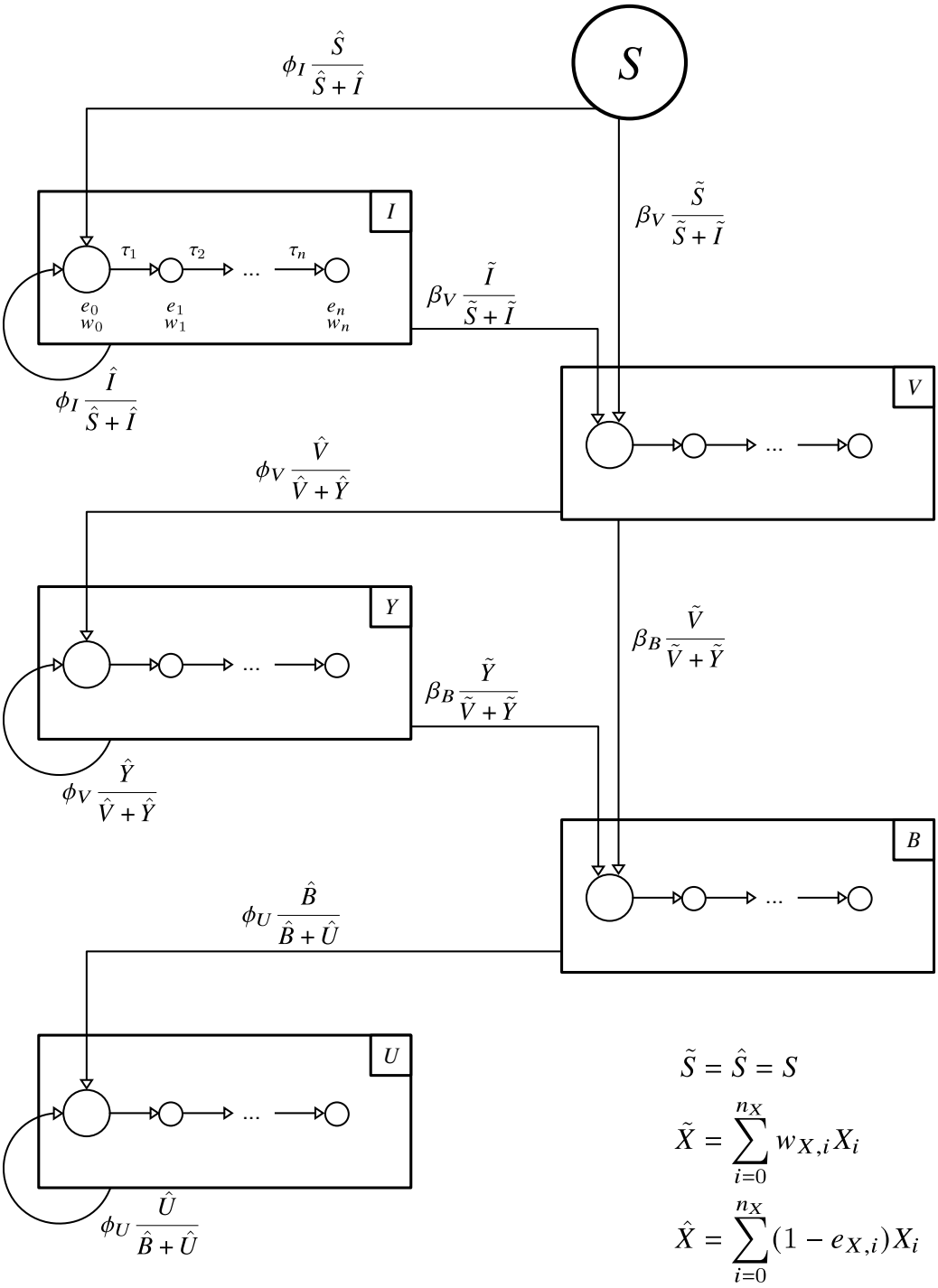

FIG. 9. Detailed model that includes waning.

i.e. susceptibles S who become infected transition to status I and susceptibles who are vaccinated transition to status V. Recovered I who become infected again transition to status I and recovered people who get vaccinated transition to status V. Furthermore,

$$\mathcal{I}(V) = \mathcal{I}(Y) = Y \tag{C8}$$

$$\mathcal{V}(V) = \mathcal{V}(Y) = B, \tag{C9}$$

i.e. vaccinated individuals V who become infected transition to status Y and those vaccinated that receive a third dose transition to status B. Breakthrough-recovereds Y who become reinfected again transition to status Y and breakthrough-recovered individuals who become vaccinated transition to status B. Last,

$$\mathcal{I}(B) = \mathcal{I}(U) = U \tag{C10}$$

$$\mathcal{V}(B) = \mathcal{V}(U) = \varnothing, \tag{C11}$$

i.e. boostered persons B who become infected transition to status U but further vaccination is not provided. Recovered booster vaccinated persons U who become infected again will again transition to status U. The dynamics of all states X_i follows

$$\partial_t X_i = \underbrace{\phi_X \delta_{i,0} - \phi_{\mathcal{I}(X)} \left(1 - e_{X,i}\right) X_i}_{\text{infections}} + \underbrace{\beta_X \delta_{i,0} - \beta_{\mathcal{V}(X)} \left(1 - F_{X,i}(\Theta_X)\right) X_i}_{\text{vaccinations}} + \underbrace{\frac{X_{i-1}}{\tau_{X,i}} - \frac{X_i}{\tau_{X,i+1}}}_{\text{waning}}. \tag{C12}$$

By definition, we have $X_j = 0 \forall j < 0 \wedge j > n_X + 1$, as well as $\phi_\varnothing = 0$ and $\beta_\varnothing = 0$. Furthermore, we set $\beta_S = \beta_I = \beta_Y = \beta_U = 0$ and $\phi_S = \phi_V = \phi_B = 0$, that is, there are no infections ending in vaccination compartments and no vaccinations ending in infection compartments and no transitions ending in S. Additionally, susceptibles are maximally susceptible (i.e. $e_S = 0$) and from $n_S = 0$ follows $w_S = 1$. To ensure the validity of transition terms in intermediate compartments, we additionally define $\tau_{X,j} \neq 0 \forall X, j \leq 0 \wedge j > n_X + 1$.

With regard to under-reporting, we assume that under-ascertainment ratios are already included in the respective rates $\phi_\bullet$ and $\beta_\bullet$.

Finally, the aim of this analysis is to estimate seroprevalence at time t. For each state $X_i \neq S$, we denote by $p_{X,i}$ the probability that antibodies are found in a person in state X_i. Then, the seroprevalence P of the age group/population of consideration is given as

$$P(t) = \sum_{X \neq S} \sum_{i=0}^{n_X} p_{X,i} X_i(t). \tag{C13}$$

The model is illustrated in Fig. 9.

A large number of parameters are required to calibrate the model. For each state $X \in \{I, V, Y, B, U\}$ the number of transitions n_X have to be defined, then n_X mean transition times as well as $n_X + 1$ susceptibility reductions. For compartments I, V, Y and B, eligibility times $\Theta_\bullet$ for receiving a vaccination are to be determined. From reporting data, we obtain the daily number of new infections of unvaccinated $\phi_I(t)$, vaccinated $\phi_V(t)$ and boostered $\phi_U(t)$ individuals. From the vaccination archive, we obtain the daily number of completed initial vaccination series $\beta_V(t)$ and booster vaccinations $\beta_B(t)$. Under-reporting of infections and booster vaccinations must be estimated and accounted for in the respective rates. For each state $X_i \neq S$, the probability $p_{X,i}$ of finding antibodies in a person in state X_i must also be defined.

All these parameters have to be determined for each of the subpopulations (age groups, regions).

---

[1] World Health Organisation. WHO Coronavirus (COVID-19) Dashboard; 2021. Accessed: 2021-11-19. Available from: `https://covid19.who.int`.

[2] Robert Koch-Institut, Fachgebiet 33. COVID-19-Impfungen in Deutschland. Zenodo; 2021. Available from: `https://doi.org/10.5281/zenodo.5126652`.

[3] Vygen-Bonnet S, Koch J, Bogdan C, Harder T, Heininger U, Kling K, et al. Beschluss und Wissenschaftliche Begründung der Ständigen Impfkommission (STIKO) für die COVID-19-Impfempfehlung. 2020;(2):3-63.

[4] Vygen-Bonnet S, Koch J, Armann J, Berner R, Bogdan C, Harder T, et al. Beschluss der STIKO zur 9. Aktualisierung der COVID-19-Impfempfehlung und die dazugehörige wissenschaftliche Begründung. 2021;(33):3-46.

[5] Vygen-Bonnet S, Koch J, Armann J, Berner R, Bogdan C, Burchard G, et al. STIKO-Empfehlung zur COVID-19-Impfung von Kindern im Alter von 5–11 Jahren mit dem mRNA-Impfstoff Comirnaty und die dazugehörige wissenschaftliche Begründung. 2021;(1):16-47.

[6] Ständige Impfkommission (STIKO). Beschluss der STIKO zur 10. Aktualisierung der COVID-19-Impfempfehlung. 2021;(38):3-9.

[7] Ständige Impfkommission (STIKO). Ständige Impfkommission: Beschluss der STIKO zur 18. Aktualisierung der COVID-19-Impfempfehlung. Epidemiologisches Bulletin. 2022;(7):3-18.

[8] Koch J, Vygen-Bonnet S, Bogdan C, Burchard G, Falman A, Garbe E, et al. Wissenschaftliche Begründung der STIKO zur Aktualisierung der Empfehlung der COVID-19-Auffrischimpfung mit einem mRNA-Impfstoff. 2021;(48):15-41.

[9] Vygen-Bonnet S, Koch J, Bogdan C, Burchard G, Garbe E, Heininger U, et al. STIKO-Empfehlung zur COVID-19-Auffrischimpfung mit einem mRNA-Impfstoff für 12 – 17-jährige Kinder und Jugendliche mit dem mRNA-Impfstoff Comirnaty und die dazugehörige wissenschaftliche Begründung. Epidemiologisches Bulletin. 2022;(3):21-31.

[10] Neuhauser H, Buttmann-Schweiger N, Ellert U, Fiebig J, Hövener C, Offergeld R, et al. Seroepidemiological studies on SARS-CoV-2 in samples from the general population and blood donors in Germany – findings up to August 2021. Robert Koch-Institut, Epid Bull. 2021;37:3-12.

[11] Robert Koch Institute. Seroepidemiological Studies in Germany; 2022. Available from: www.rki.de/covid-19-serostudies-germany.

[12] Robert Koch Institute. Wöchentlicher Lagebericht des RKI zur Coronavirus-Krankheit-2019 (COVID-19) - 31.03.2022. 2022 Mar. Available from: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/Wochenbericht/Wochenbericht_2022-03-31.pdf?__blob=publicationFile.

[13] Betsch C, Korn L, Felgendreff L, Eitze S, Schmid P, Sprengholz P, et al. COSMO – Zusammenfassung und Empfehlungen Wellen 52 und 53. 2021 Oct. Available from: https://projekte.uni-erfurt.de/cosmo2020/web/summary/52-53/.

[14] Betsch C, Korn L, Felgendreff L, Eitze S, Schmid P, Sprengholz P, et al.. COVID-19 Snapshot Monitoring (COSMO Germany) - Wave 52. PsychArchives; 2021. Type: dataset. Available from: https://www.psycharchives.org/handle/20.500.12034/4564.

[15] Betsch C, Korn L, Felgendreff L, Eitze S, Schmid P, Sprengholz P, et al.. COVID-19 Snapshot Monitoring (COSMO Germany) - Wave 53. PsychArchives; 2021. Type: dataset. Available from: https://www.psycharchives.org/handle/20.500.12034/4570.

[16] Sotoodeh Ghorbani S, Taherpour N, Bayat S, Ghajari H, Mohseni P, Hashemi Nazari SS. Epidemiologic characteristics of cases with reinfection, recurrence, and hospital readmission due to COVID-19: A systematic review and meta-analysis. Journal of Medical Virology. 2022 Jan;94(1):44-53. Available from: https://onlinelibrary.wiley.com/doi/10.1002/jmv.27281.

[17] 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 Surveillance: Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin. 2020 Nov;25(47).

[18] Neuhauser H, Rosario AS, Butschalowsky H, Haller S, Hoebel J, Michel J, et al. Germany's low SARS-CoV-2 seroprevalence confirms effective containment in 2020: Results of the nationwide RKI-SOEP study. medRxiv. 2021. Available from: https://www.medrxiv.org/content/early/2021/11/24/2021.11.22.21266711.

[19] Feikin DR, Higdon MM, Abu-Raddad LJ, Andrews N, Araos R, Goldberg Y, et al. Duration of effectiveness of vaccines against SARS-CoV-2 infection and COVID-19 disease: results of a systematic review and meta-regression. The Lancet. 2022 Mar;399(10328):924-44. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673622001520.

[20] Krueger T, Gogolewski K, Bodych M, Gambin A, Giordano G, Cuschieri S, et al. Risk assessment of COVID-19 epidemic resurgence in relation to SARS-CoV-2 variants and vaccination passes. Communications Medicine. 2022 Mar;2(1):23. Available from: https://doi.org/10.1038/s43856-022-00084-w.

[21] Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet Infectious Diseases. 2020 Jun;20(6):669-77. Available from: https://doi.org/10.1016/S1473-3099(20)30243-7.

[22] Berger U, Fritz C, Kauermann G. Eine statistische Analyse des Effekts von verpflichtenden Tests an Schulen mit Präsenzunterricht im Vergleich zum Distanzunterricht. 2021;238. Available from: https://epub.ub.uni-muenchen.de/76005/.

[23] Iyer AS, Jones FK, Nodoushani A, Kelly M, Becker M, Slater D, et al. Persistence and decay of human antibody responses to the receptor binding domain of SARS-CoV-2 spike protein in COVID-19 patients. Science Immunology. 2020;5(52):eabe0367. Available from: https://www.science.org/doi/abs/10.1126/sciimmunol.abe0367.

[24] Office for National Statistics (ONS) - UK. Coronavirus (COVID-19) Infection Survey Technical Article: Impact of vaccination of testing positive in the UK: October 2021. 2021. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19infectionsurveytechnicalarticleimpactofvaccinationontestingposi
october2021.

[25] Faensen, D and Krause, G. SurvStat@RKI – a web-based solution to query surveillance data in

Germany. Weekly releases (1997–2007) Euro Surveill. 2004;8(22). Available from: https://www.eurosurveillance.org/content/10.2807/esw.08.22.02477-en.

[26] Robert Koch-Institut. SARS-CoV-2 Sequenzdaten aus Deutschland. Zenodo; 2022. Available from: https://doi.org/10.5281/zenodo.6381314.

[27] Statistisches Bundesamt (Destatis). GENESIS-Tabelle: 12411-0012, Bevölkerung: Bundesländer, Stichtag, Altersjahre. Statistisches Bundesamt (Destatis); 2020. Available from: https://www-genesis.destatis.de/genesis/online?sequenz=statistikTabellen&selectionname=12411#abreadcrumb.

[28] Barek MA, Aziz MA, Islam MS. Impact of age, sex, comorbidities and clinical symptoms on the severity of COVID-19 cases: A meta-analysis with 55 studies and 10014 cases. Heliyon. 2020;6(12):e05684. Available from: https://www.sciencedirect.com/science/article/pii/S2405844020325275.

[29] Brinkmann F, Diebner HH, Matenar C, Schlegtendal A, Spiecker J, Eitner L, et al. Longitudinal Rise in Seroprevalence of SARS-CoV-2 Infections in Children in Western Germany—A Blind Spot in Epidemiology? Infect Dis Rep. 2021 Nov;13:957 964.

[30] Hippich M, Holthaus L, Assfalg R, Zapardiel-Gonzalo J, Kapfelsperger H, Heigermoser M, et al. A Public Health Antibody Screening Indicates a 6-Fold Higher SARS-CoV-2 Exposure Rate than Reported Cases in Children. Med (NY). 2021 Feb;2:149 163.

[31] Wachter F, Regensburger AP, Peter AS, Knieling F, Wagner AL, Simon D, et al. Continuous monitoring of SARS-CoV-2 seroprevalence in children using residual blood samples from routine clinical chemistry. Clinical Chemistry and Laboratory Medicine (CCLM). 2022. Available from: https://doi.org/10.1515/cclm-2022-0037.

[32] Morens DM, Folkers GK, Fauci AS. The Concept of Classical Herd Immunity May Not Apply to COVID-19. The Journal of Infectious Diseases. 2022 Mar. Available from: https://doi.org/10.1093/infdis/jiac109.

[33] Hall V, Foulkes S, Insalata F, Kirwan P, Saei A, Atti A, et al. Protection against SARS-CoV-2 after Covid-19 Vaccination and Previous Infection. New England Journal of Medicine. 2022 Mar;386(13):1207-20. Available from: http://www.nejm.org/doi/10.1056/NEJMoa2118691.

[34] Altarawneh HN, Chemaitelly H, Ayoub H, Tang P, Hasan MR, Yassine HM, et al. Effect of prior infection, vaccination, and hybrid immunity against symptomatic BA.1 and BA.2 Omicron infections and severe COVID-19 in Qatar. Epidemiology; 2022. Available from: http://medrxiv.org/

lookup/doi/10.1101/2022.03.22.22272745.

[35] Quandt J, Muik A, Salisch N, Lui BG, Lutz S, Krüger K, et al. Omicron breakthrough infection drives cross-variant neutralization and memory B cell formation. Immunology; 2022. Available from: http://biorxiv.org/lookup/doi/10.1101/2022.04.01.486695.

[36] Robert Koch-Institut. SARS-CoV-2 Infektionen in Deutschland. Zenodo; 2022. Available from: https://doi.org/10.5281/zenodo.5989727.

[37] Robert Koch-Institut. Wöchentlicher Lagebericht des RKI zur Coronavirus-Krankheit-2019 (COVID-19). 2021 dec. Available from: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/Wochenbericht/Wochenbericht_2021-12-23.pdf?__blob=publicationFile.

[38] Autorengruppe Corona-KiTa-Studie (2021). Quartalsbericht der Corona-KiTa-Studie 6. Quartalsbericht (IV/2021). DJI. 2021. Available from: https://corona-kita-studie.de/quartalsberichte-der-corona-kita-studie.

[39] Maier B. serocov: Estimating the distribution of COVID-19-susceptible, -recovered, and -vaccinated individuals in Germany up to April 2022. Zenodo; 2022. Available from: https://doi.org/10.5281/zenodo.6470799.