ARTICLE

Check for updates

Germany's fourth COVID-19 wave was mainly driven by the unvaccinated

Benjamin F. Maier [1,2 ✉], Marc Wiedermann [1,2], Angelique Burdinski [1,2], Pascal P. Klamser [1,2], Mirjam A. Jenny [3,4,5,6], Cornelia Betsch [3,6] & Dirk Brockmann [1]

Abstract

Background While the majority of the German population was fully vaccinated at the time (about 65%), COVID-19 incidence started growing exponentially in October 2021 with about 41% of recorded new cases aged twelve or above being symptomatic breakthrough infections, presumably also contributing to the dynamics. So far, it remained elusive how significant this contribution was and whether targeted non-pharmaceutical interventions (NPIs) may have stopped the amplification of the crisis.

Methods We develop and introduce a contribution matrix approach based on the next-generation matrix of a population-structured compartmental infectious disease model to derive contributions of respective inter- and intragroup infection pathways of unvaccinated and vaccinated subpopulations to the effective reproduction number and new infections, considering empirical data of vaccine efficacies against infection and transmission.

Results Here we show that about 61%–76% of all new infections were caused by unvaccinated individuals and only 24%–39% were caused by the vaccinated. Furthermore, 32%–51% of new infections were likely caused by unvaccinated infecting other unvaccinated. Decreasing the transmissibility of the unvaccinated by, e. g. targeted NPIs, causes a steeper decrease in the effective reproduction number $\mathcal{R}$ than decreasing the transmissibility of vaccinated individuals, potentially leading to temporary epidemic control. Reducing contacts between vaccinated and unvaccinated individuals serves to decrease $\mathcal{R}$ in a similar manner as increasing vaccine uptake.

Conclusions A minority of the German population—the unvaccinated—is assumed to have caused the majority of new infections in the fall of 2021 in Germany. Our results highlight the importance of combined measures, such as vaccination campaigns and targeted contact reductions to achieve temporary epidemic control.

Plain language summary

With about 65% of its citizens vaccinated at the time, Germany experienced a large wave of COVID-19 in the fall of 2021, regionally overburdening the healthcare system. We are interested in how much this crisis was driven by infections in vaccinated versus unvaccinated people. We use a mathematical model to show that transmission of the disease during this period was largely driven by the unvaccinated population, despite representing a smaller proportion of the overall population. Our results suggest that higher vaccine uptake, reduced mixing between vaccinated and unvaccinated people, and targeted contact-reduction measures would have been effective measures to control spread at the time. These findings may have implications for how we manage future waves of COVID-19 or other diseases.

[1] Institute for Theoretical Biology and Integrated Research Institute for the Life-Sciences, Humboldt-University of Berlin, Philippstr. 13, 10115 Berlin, Germany. [2] Robert Koch Institute, Nordufer 20, 13353 Berlin, Germany. [3] University of Erfurt, Nordhäuserstr. 63, 99089 Erfurt, Germany. [4] Harding Center for Risk Literacy, University of Potsdam, Virchowstrasse 2-4, 14482 Potsdam, Germany. [5] Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. [6] Bernhard-Nocht-Institut, Bernhard-Nocht-Straße 74, 20359 Hamburg, Germany. ✉email: bfmaier@physik.hu-berlin.de

Vaccines are the most powerful pharmaceutical tool to prevent infections with SARS-CoV-2 and combat the COVID-19 pandemic. Fast vaccine uptake by as many individuals as possible saves lives, people's health, and livelihoods. Despite large-scale vaccine roll-out campaigns, many countries, most prominently in Europe, have experienced a rise in case numbers in the late summer and early fall of 2021 and reported effective reproduction numbers $\mathcal{R}$ above one for an extended period of time[1]. This means that on average, every infected person infected more than one other person, thus causing exponentially rising incidences[2]. Since the beginning of this pandemic, such resurgences have, in part, been mitigated by harsh non-pharmaceutical interventions (NPIs) such as lockdowns or curfews that limit the population's contacts, thereby decreasing the effective reproduction number and relieving overburdened public health systems[3,4]. Measures that affect large parts of the general population over a long period of time can have devastating effects, such as increasing social inequality and domestic violence, detrimental impacts on mental health, or economic disruptions[5–9]. Such harsh restrictions should therefore be considered a last resort of pandemic control.

During the onset of the fourth COVID-19 wave in Germany, many hospitals and intensive care units (ICUs) were operating at maximum capacity or were projected to do so at a later point[10]. In the four weeks between Oct 11, 2021, and Nov 7, 2021, Germany's central public health institute, the Robert Koch Institute (RKI) reported 250,552 new symptomatic infections in individuals with known vaccination status, 90,471 of which were attributed to vaccinated individuals, i.e. 36% were symptomatic breakthrough cases (41% in age groups eligible for vaccination)[11]. During this time, the average vaccination rate in different age groups [0,12), [12,18), [18,60), and 60+ were 0%, 40.1%, 72.4%, and 85.1%, respectively, leading to 0%, 4.8%, 41.6%, and 61.9% of new cases being classified as symptomatic breakthrough cases within the respective age groups[11], Table 1. Simultaneously, the effective reproduction number remained at a relatively stable value of $\mathcal{R} \approx 1.2$ (under the assumption of a generation time of four days)[12].

Given that breakthrough cases are a challenge both for communication and vaccine acceptance[13] and that harsh NPIs may be illegitimate for vaccinated individuals, the above situation raises two important questions: How much does the unvaccinated population contribute to the infection dynamics despite being in the minority? And could targeted NPIs aiming at reducing the contacts of unvaccinated individuals temporarily and sufficiently suppress the infection dynamics such that harsh, large-scale NPIs could be avoided?

To address these questions, we establish the contribution matrix approach, a theoretical concept derived from the next-generation matrix framework[14]. The contribution matrix quantifies the contributions to $\mathcal{R}$ caused by the infection pathways from un-/vaccinated individuals to other un-/vaccinated

individuals, considering the age and contact structure of the population, vaccination rates, as well as expected vaccine efficacies regarding susceptibility and transmission reductions, respectively. In its general form, it quantifies the contributions made by any combination of two subpopulations.

Based on this approach, we estimate that in October 2021, around 32%–51% (depending on vaccine efficacy) of the effective reproduction number was caused by unvaccinated individuals infecting other unvaccinated individuals (see Fig. 1). Since unvaccinated individuals have a higher probability of suffering from severe disease[15–17], this contribution is the major factor that drove the public health system into a crisis characterized by hospitals and ICUs reaching maximum capacity. In contrast, we estimate that only 15%–18% of the reproduction number were attributable to vaccinated individuals infecting unvaccinated individuals. In October 2021, about 65% of the German population was fully vaccinated, implying that the majority of the overall population contributed little to the amplification of the crisis. In total, we estimate that the vaccinated population contributed 24%–39% to $\mathcal{R}$ while the unvaccinated population contributed the remaining 61%–76%, despite the fact that unvaccinated individuals have been in the minority in Germany. 9%–21% of new infections would be caused by vaccinated individuals infecting other vaccinated people. In total, we estimate that unvaccinated individuals were involved in 8–9 out of 10 new infections, either as infecting, acquiring infection, or both.

We further argue that regarding the situation in the fall of 2021, the unvaccinated would have had to reduce their transmissibility two to three times as strongly as the vaccinated in order for the system to reach $\mathcal{R} = 1$ (and hence containment of the infection wave), if the burden of contact reductions were to be distributed between the two subpopulations according to their respective contributions. Moreover, decreasing mixing between individuals of distinct vaccination status can decrease $\mathcal{R}$. Ultimately, a higher vaccine uptake would have led to less unvaccinated being involved in infections, which can not only decrease $\mathcal{R}$, but is critical for relieving an overburdened public health system, as they are more likely to suffer from severe disease. Combinations of these interventions that address mainly the unvaccinated might have rendered the dynamics subcritical.

Methods

Mathematical framework. We use a population-structured compartmental infectious disease model that captures a variety of aspects regarding vaccination against COVID-19 (see Supplementary Methods, Sec. 1.1.1). The model's dynamics are fully described by the next-generation matrix K_{ji} of small domain (see Supplementary Methods, Sec. 1.1.2), which quantifies the average number of offspring in group j caused by a single infectious individual in group i[14]. Here, the index i (or j, respectively) refers to the subpopulation that is determined by a respective age group and the vaccination status within that group, thus yielding two subpopulations per age group. In the regime of small outbreaks (relative to the total population size), the ordinary differential equations governing the epidemic growth can be linearized, with the dynamics being determined by K_{ji}, such that the generational growth of the number of infected individuals in group i follows

$$y_j(g+1) = \sum_i K_{ji} y_i(g), \qquad g = 0, 1, 2, \dots \qquad (1)$$

The incidence approaches the eigenstate y_i of K_{ji} that corresponds its spectral radius, which in turn is equal to the effective reproduction number[14]. Hence, the entries of the normalized

---

Table 1 Share of breakthrough infections in the age groups eligible for vaccination according to official estimates by the Robert Koch Institute (RKI)[11] and the model for "low efficacy", "medium efficacy", and "high efficacy" scenarios.

Age group	RKI report (symptomatic cases)	Model ("high eff.")	Model ("medium eff.")	Model ("low eff.")
adolescents	4.8%	5.1%	21.1%	25%
adults	41.6%	42.3%	51.2%	57%
elderly	61.9%	61.5%	74.1%	77.4%
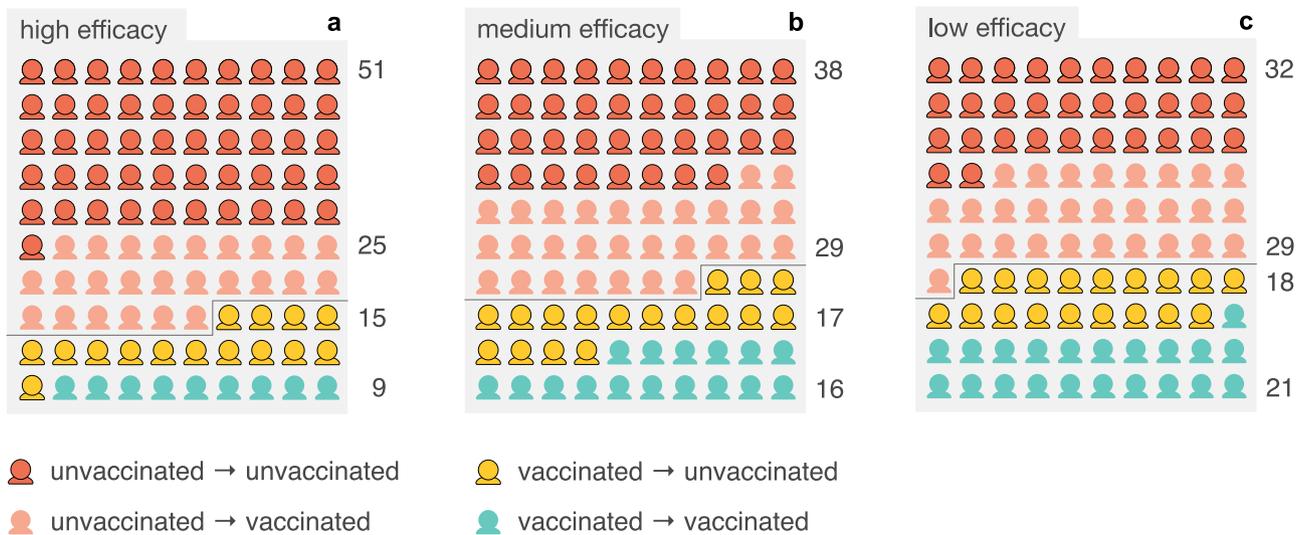
Fig. 1 Estimated contributions of infection pathways towards new cases within vaccinated and unvaccinated subpopulations. Estimated contributions of infection pathways to $\mathcal{R}$ in the (**a**) "high efficacy", (**b**) "medium efficacy", and (**c**) "low efficacy" scenarios as a graphical representation of Tabs. 2–4. The charts can be read as follows: Consider an infected population that caused a new generation of 100 new infecteds. Then for (**a**), 51 of those newly infected individuals will be unvaccinated people that have been infected by other unvaccinated people. Likewise, 25 newly infected individuals will be vaccinated people that have been infected by unvaccinated individuals. Hence, 76 new infections will have been caused by the unvaccinated. Along the same line, 15 newly infecteds will be unvaccinated people that have been infected by vaccinated individuals and 9 newly infecteds will be vaccinated people that have been infected by other vaccinated individuals, totaling 24 new infections that have been caused by vaccinated individuals.

eigenvector $\hat{y}_i = y_i/\sum_j y_j$ contain the relative frequency of newly infected individuals in age/vaccination group i.

Consequently, the number of j-offspring caused by i-individuals in a dynamical system defined by K_{ji} is given by the contribution matrix

$$C_{ji} = K_{ji}\hat{y}_i. \qquad (2)$$

Summing over all matrix elements of C_{ji} yields the effective reproduction number $\mathcal{R}$ (see Supplementary Methods, Sec. 1.1.1)). A single matrix element C_{ji} can thus be considered the contribution of the $i \rightarrow j$ infection pathway to the reproduction number (a derivation of the concept and an operational definition of C_{ji} can be found in the Supplementary Methods, Sec. 1.1.1–1.1.2 and Sec. 1.2.4), respectively). The normalized contribution matrix $C_{ji}/\mathcal{R}$ gives the relative contributions of $i \rightarrow j$ infections towards $\mathcal{R}$ (and consequently, towards the total number of new infections).

We derive explicit equations for the contributions of un-/vaccinated individuals in the homogeneous case, i.e. ignoring age structure (see Supplementary Methods, Sec. 1.2.3). These contributions are

$$C_{u \leftarrow u} = \frac{(1-v)^2}{1-vs}\mathcal{R}_u \qquad (3)$$

$$C_{u \leftarrow v} = \frac{v(1-v)(1-s)(1-r')}{1-vs}\mathcal{R}_v \qquad (4)$$

$$C_{v \leftarrow u} = \frac{v(1-v)(1-s)}{1-vs}\mathcal{R}_u \qquad (5)$$

$$C_{v \leftarrow v} = \frac{v^2(1-s)^2(1-r')}{1-vs}\mathcal{R}_v, \qquad (6)$$

where v is the vaccine uptake, s is the susceptibility reduction after vaccination, $r' = 1 - (1-r)/b$ is the adjusted transmissibility reduction (i.e. it contains the relative increase of the recovery rate after a breakthrough infection b and viral shedding reduction r), $\mathcal{R}_u$ is the base transmissibility of unvaccinated infecteds, and $\mathcal{R}_v$

is the base transmissibility of vaccinated infecteds (both of which quantify differences in behavior in the respective groups). The total effective reproduction number is given by

$$\begin{aligned} \mathcal{R} &= C_{u \leftarrow u} + C_{v \leftarrow u} + C_{u \leftarrow v} + C_{v \leftarrow v} \\ &= (1-v)\mathcal{R}_u + v(1-s)(1-r')\mathcal{R}_v. \end{aligned} \qquad (7)$$

Model structure, parameters, and scenarios. In the full model, we construct the next-generation matrix of small domain (see Supplementary Methods, Eq. (S3)) based on the following observations, assumptions, and estimates: We structure the population into four age groups [0,12) (children), [12,18) (adolescents), [18,60) (adults), and 60+ (elderly). Contact numbers between those age groups and subpopulation sizes were constructed based on the POLYMOD (2005) data set[18,19] using the 'socialmixr' software package[20] (see Supplementary Methods, Sec. 1.2.1). Since vaccine efficacy was, at the time of writing, estimated only for the status "fully vaccinated" in Germany without distinguishing between different vaccines, we solely distinguish between "unvaccinated" and "vaccinated" individuals in the model, regardless of the make of the received doses (note that by the fall of 2021, a total number of four vaccine types was available in Germany, i.e., Spikevax (Moderna), Ad26.COV2.S (Janssen), Vaxzevria (AstraZeneca), and Comirnaty (BioNTech/Pfizer) with the latter being by far the most used[21]). Following the example of Scholz et al.[22], we further assume that children and adolescents have reduced susceptibility to the virus and a reduced base transmissibility if infected, as was observed in Germany, Israel, and Greece[23–25]. In the discussed time frame, 14.7%, 9.4%, 60.2%, and 15.7% of new cases can be attributed to the respective age groups [0,12), [12,18), [18,60), and 60+[26]. In order to match this distribution approximately, we calibrate the base susceptibility (i. e. susceptibility without vaccination) and infectiousness of our model by assuming that children are 72% as susceptible and 63% as infectious as adults (72% and 81% for adolescents), which is larger than what was observed for the wild type[24,25], see Supplementary Methods, Sec. 1.2. However, since the B.1.617.2 variant (Delta) that was predominant in Germany in October/

November 2021 was generally observed to be more infectious than the wild type[27], such an increase is plausible. Note that in principle, heterogeneous ascertainment may lead to a distribution of detected cases that is skewed towards the adult population, as children and adolescents may have higher probability of suffering from an asymptomatic infection[28] and thus are less likely to be detected via symptom-based testing strategies. Yet, by the fall of 2021, Germany made regular screening via rapid antigen tests mandatory in schools across the country, potentially lowering the level of under-ascertainment in these age groups[29]. Nevertheless, we test how our results change by assuming children and adolescents are as susceptible as adults in a sensitivity analysis (see Supplementary Methods, Sec. 1.3.3). Additionally, note that we ignore the number of recovered individuals. Until Oct 10, 2021, about 4.3 million infections were reported in Germany[12], 74% of which likely received a vaccination[30–32] and are therefore considered as vaccinated in our analysis. With an under-ascertainment ratio of about 1.8[33], we estimate that the total number of non-vaccinated recovered individuals was on the order of 2.4% of the population in Germany at the time, and therefore negligible in our analysis.

In Germany, an estimated average vaccine efficacy of 72% against symptomatic COVID-19 in adults and the elderly was found for cases reported between Oct 11, 2021 and Nov 7, 2021[11]. Vaccine efficacy for adolescents was not reported due to the respective data being potentially unreliable (low number of cases). Because these efficacies were computed for symptomatic cases, we use their values as a "high efficacy" scenario regarding vaccine efficacy in our analysis, because unreported and/or asymptomatic breakthrough infections might lower the estimated efficacies (see Supplementary Methods, Sec. 1.2.5). However, note that these observed 72% vaccine efficacy are in line with an estimated population-wide vaccine efficacy against infection based on vaccination time series and waning immunity data that was published in a meta-review by the WHO[34,35]. In order to obtain breakthrough infection rates in adolescents on the order of observed symptomatic breakthrough cases we assume a vaccine efficacy of $s = 92\%$ for adolescents. Despite being comparably large, this value seems justified considering that adolescents have been made eligible to receive a vaccine in Germany only shortly prior to the study period, and a high vaccine efficacy against infection with the Delta variant has been reported for this age group[36]. Regarding the infectiousness of individuals suffering from breakthrough infections, viral load of vaccinated patients suffering from symptomatic COVID-19 was reported to be at the same level as of those unvaccinated[37,38]. Another study from the UK found decreased infectiousness in breakthrough infections[39]. Considering both these results, we assume a conservative transmission reduction of $r = 10\%$ for breakthrough infections. In agreement with the literature[37,40] we further consider that the average infectious period of breakthrough infections is shorter than for unvaccinated individuals and assume a 50% increase in recovery rate for the vaccinated, amounting to an average infectious period that is 2/3 as long as that of unvaccinated infecteds ($b = 3/2$) (see Supplementary Methods, Sec. 1.2.2). Such an increased recovery rate can also be caused by deliberate behavior. As individuals that are not opposed to vaccination typically adhere to protection measures more consistently[41], behavioral changes following a breakthrough infection might further decrease the effective infectious period. Note that together with a decreased duration of infection $b = 3/2$, the adjusted transmission reduction reads $r' = 1 - (1 - r)/b = 40\%$, which is lower than a 63% reduction that was observed for household transmissions of the Delta variant between infected vaccinated and susceptible unvaccinated individuals in the Netherlands in August and September 2021, close to our observational period[42].

As this reduction was observed to wane over time[43], $r' = 40\%$ is a reasonable assumption.

In a second, "medium efficacy" scenario, we consider that vaccine efficacies against infection are in the range of 50%–60%, i.e. lower than the observed value against symptomatic COVID-19, and lower than vaccine efficacies reported in the UK for the Comirnaty (BioNTech/Pfizer) vaccine[44], considering that partial immunity might have waned over time[45]. Since vaccine efficacy is expected to decrease with age[45,46], we assume an efficacy against infection of $s = 60\%$ for adolescents and adults as well as $s = 50\%$ for the elderly (see Supplementary Methods, Sec. 1.2.4).

Finally, we also discuss a "low efficacy" scenario where the susceptibility reduction is assumed to be much lower than the observed efficacy against symptomatic COVID-19, namely 50% for adolescents and adults, and 40% for the elderly (see Supplementary Methods, Sec. 1.2.5).

To summarize the main scenarios, for the "high efficacy" the vaccination efficacy s for adolescents, adults, and elderly is assumed to be 92%, 72%, 72%, in the "medium efficacy" scenario 60%, 60%, 50%, and in the "low efficacy" scenario 50%, 50%, 40%, respectively.

Based on these considerations we compute the respective full-model next generation matrices K_{ji} and numerically find the normalized population eigenvectors $\hat{y}_i$ corresponding to the respective and the contribution matrices C_{ji}, which we further reduce to the two-dimensional vaccination status space by summing over the respective contributions of age groups (see Supplementary Methods, Eq. (S2)).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Results

As a first model validation we find that for the high efficacy scenario the relative size of breakthrough infections within age groups eligible for vaccination is in good agreement with the share of reported symptomatic breakthrough cases (Table 1), albeit being slightly larger than reported values, mirroring the fact that the official number of breakthrough infections is likely affected by underreporting[11] and that the number of infections will be larger than the number of symptomatic breakthrough cases.

For all scenarios, we find that the largest entry in the contribution matrix is given by the unvaccinated → unvaccinated infection pathway, with a 51.4% (high efficacy), 38.1% (medium efficacy) and 31.6% (low efficacy) contribution respectively, see Tables 2, 3, 4 and Fig. 1. Most noteworthy, these numbers represent the largest contributions although the unvaccinated population is smaller than the vaccinated one. Moreover, the total contribution of the unvaccinated population to the effective reproduction number is 75.9%, 66.6%, and 61.1% for the high, medium, and low efficacy scenarios, respectively. In total, the unvaccinated population plays a role in 91.1% (high), 84%

Table 2 Contribution to $\mathcal{R}$ from infections between vaccinated and unvaccinated populations for the upper parameter bounds.

	← (u)nvaccinated	← (v)accinated
u ←	51.4%	15.0%
v ←	24.5%	9.1%
total	75.9%	24.1%

Table 3 Relative contributions to $\mathcal{R}$ from infections between vaccinated and unvaccinated groups for the "medium efficacy" scenario.

	← (u)nvaccinated	← (v)accinated
u ←	38.1%	17.4%
v ←	28.5%	16.0%
total	66.6%	33.4%

Table 4 Relative contributions to $\mathcal{R}$ from infections between vaccinated and unvaccinated groups for the "low efficacy" scenario.

	← (u)nvaccinated	← (v)accinated
u ←	31.6%	18.2%
v ←	29.5%	20.7%
total	61.1%	38.9%

(medium), or 79.3% (low efficacy) of cases—either as infecting, acquiring infection, or both.

Since vaccine efficacy is expected to decrease with age and time passed after vaccination[45], we test how our results for the "medium efficacy" scenario change when assuming a more pessimistic susceptibility reduction of 40% for the elderly while keeping 60% for all other age groups (see Supplementary Methods, Sec. 1.3.5). We find that our results do not change substantially (see Supplementary Table 3), which can be attributed to the fact that the elderly generally have a lower contact behavior than other age groups.

In order to test the validity of the homogeneous approach, we further use Eqs. (3)–(6) to compute the contribution matrix with $v = 65\%$, $s = 72\%$, $r = 10\%$, and $b = 3/2$, assuming $\mathcal{R}_u = \mathcal{R}_v$. We find relative contributions of $C_{u \leftarrow u}/\mathcal{R} = 50.1\%$, $C_{v \leftarrow u}/\mathcal{R} = 26.1\%$, $C_{u \leftarrow v}/\mathcal{R} = 15.7\%$, $C_{v \leftarrow v}/\mathcal{R} = 8.1\%$, hence being in good agreement with the results of the age-structured model (cf. Tab. 2), showing that Eqs. (3)–(6)) can be used to estimate the order of magnitude of the contributions by the respective infection pathways. We expect this approximation to lose its validity for situations in which model assumptions become even more heterogeneous (e.g. strong differences in contact structure between age groups, vaccine uptakes per age group, or vaccine efficacy per age group).

During the period of time when vaccine efficacies were measured[11], the reproduction number in Germany was reported to be at a relatively stable value of $\mathcal{R} = 1.2$[12]. In order to achieve temporary epidemic control, it is necessary to reach a value of $\mathcal{R} < 1$ for a substantial amount of time[2]. We therefore study how the effective reproduction number would change if the transmissibility of unvaccinated individuals would be reduced. This could, for instance, be achieved by strict enforcement of contact rules regarding unvaccinated individuals at private and public gatherings that were partially in place in Germany[47]. For our analysis we gauge K_{ji} such that $\sum_{ji} C_{ji} = \mathcal{R} = 1.2$ for either of the base scenarios and then individually scale the transmissibility of the vaccinated and unvaccinated to obtain those values at which the critical value $\mathcal{R} = 1$ is attained, Fig. 2a. We find that a transmission reduction of 22%–27% in the unvaccinated population would suffice to reach $\mathcal{R} = 1$ without the need for any further restrictions. In contrast, NPIs that would affect both, vaccinated and unvaccinated to the same degree, would need to cause more than 17% of transmissibility reduction across the entire population to achieve epidemic control. For completeness and to put numbers in perspective one may also consider the

unlikely scenario where NPIs are only in place for the vaccinated population yielding a required transmissibility reduction of 43%–73% in that group to achieve epidemic control, highlighting that vaccinated individuals would have to decrease their transmissibility less strongly than unvaccinated individuals for a distribution of the burden of contact reductions that corresponds to their respective contributions. The way to reach $\mathcal{R} = 1$ in the plane spanned by NPI-based transmissibility reductions in both respective subpopulations that acknowledges these contributions with appropriate weighting is given by the linear function that is perpendicular to the isoclines shown in Fig. 2a. Using the fact that the homogeneous model given by Eqs. (3)–(6) yields acceptable approximations to the full model, we use Eq. (7) to derive the slope $\chi = v(1 - s)(1 - r')/(1 - v)$ of this function (see Supplementary Methods, Sec. 1.3.2). This quantity has to be read as "if the unvaccinated population reduces its transmissibility by 10%, the vaccinated population has to reduce its transmissibility by $\chi \times 10\%$ in order for the system to quickly approach $\mathcal{R} = 1$". With $v = 65\%$, $s = 72\%$, $r' = 40\%$ for the "high efficacy" scenario, as well as $s = 60\%$ for the "medium" and $s = 50\%$ for the "low efficacy" scenario, we find $\chi = 0.31$, $\chi = 0.45$, and $\chi = 0.55$, respectively, which suggests that in order to adequately distribute the burden of further transmissibility reductions between the respective subpopulations, unvaccinated individuals would have to reduce their transmissibility two to three times as strongly as the vaccinated population.

We further test the robustness of our results regarding vaccine efficacy by varying an age-independent vaccine efficacy against infection that ranges from $s = 100\%$ to $s = 0\%$, (i) leaving $r = 10\%$ and $b = 3/2$ unchanged as an optimistic estimation and (ii) proportionally scaling $r = s/10$ and $b = s/2 + 1$ as a pessimistic estimation, while assuming vaccine uptake as reported in the Methods section (see Supplementary Methods, Sec. 1.3.1). We find a monotonic decrease of breakthrough infections from nonzero values for $s = 0\%$ to zero for $s = 100\%$. Notably, we find that as long as vaccine efficacies do not drop below 22% (optimistic) or 41% (pessimistic), the majority of new cases remains to be caused by the minority of the population, which are the unvaccinated (see Fig. 3 and the results for an additional "very low efficacy" scenario in Supplementary Methods Sec. 1.3.7 as well as Supplementary Table 6).

Next, we also account for the fact that the infectiousness of children and adolescents has been a matter of debate[23–25,48,49]. While for all analyses presented above we assumed reduced infectiousness for those respective age groups compared to adults and elderly, we now assume (as an upper limit) that children and adolescents are as infectious as adults (see Supplementary Methods, Sec. 1.3.3). This generally leads to higher contributions by unvaccinated individuals to the overall share of infections since they represent by far the majority in these age groups. We find that the unvaccinated in this scenario cause 76%–85% of all new infections for the "medium" and "high" scenario, respectively (see Supplementary Tables 1 and 2) which is substantially larger than the 67%–76% obtained when susceptibility and infectiousness in children and adolescents is reduced (see again Fig. 1a, b and Tabs. 2, 3).

Moreover, we test how our results change if the assumption of homogeneous mixing between vaccinated and unvaccinated individuals is no longer met. This captures the likely scenario that vaccinated and unvaccinated populations are more prone to meet individuals of similar vaccination status rather than opposing vaccination status either due to homophily[50–52] or deliberate non-pharmaceutical interventions, such as limiting access to public gatherings, immune shielding[53], or social distancing informed by serological testing[54]. We conceptualize this process by scaling the off-diagonal matrix elements indicating offspring
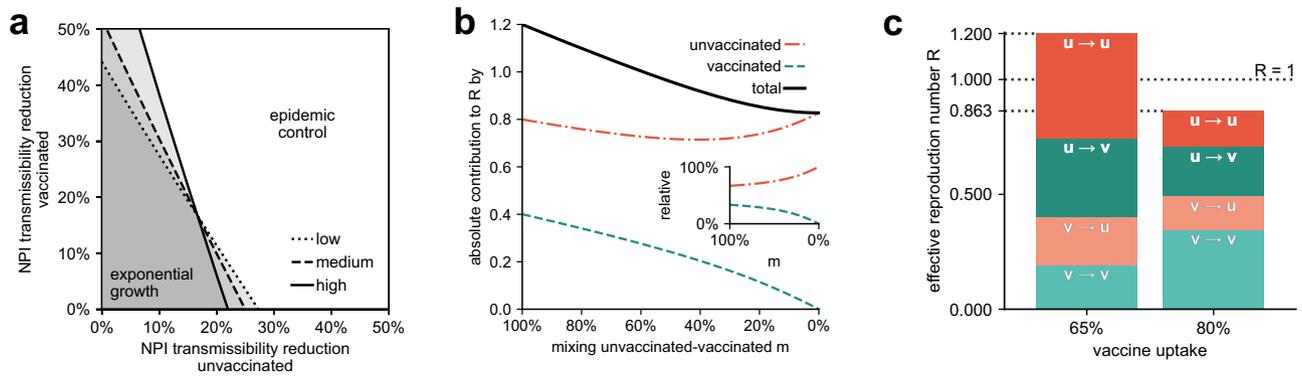
Fig. 2 Efficacy of potential interventions to achieve temporary epidemic control. a Required additional transmissibility reduction for the unvaccinated (horizontal axis) and vaccinated (vertical axis) population to lower $\mathcal{R}$ to values below one, based on the assumption that the initial effective reproduction number is equal to $\mathcal{R} = 1.2$. **b** The absolute contributions to $\mathcal{R}$ of the unvaccinated (orange) and vaccinated population (green) as well as their sum (black) with decreasing mixing m between both groups, based on the "medium efficacy" scenario. The inset shows the respective relative contributions. Note that if heterogeneous mixing was already present during our observational period, the monotonically increasing contribution of the unvaccinated displayed in the inset implies that our results of Fig. 1 are actually lower bounds of the true contribution. **c** Absolute contributions to $\mathcal{R}$ for infections between and across groups of vaccinated and unvaccinated individuals at the vaccine uptake during the observational period (left bar) and a hypothetical vaccine uptake of 80% in the total population, i.e., 90% in the age groups that were, at the time, eligible for vaccination (right bar), based on the "medium efficacy" scenario. The latter would have sufficed to suppress $\mathcal{R}$ sufficiently below one, assuming that other factors determining the base transmissibility remained on the same level.
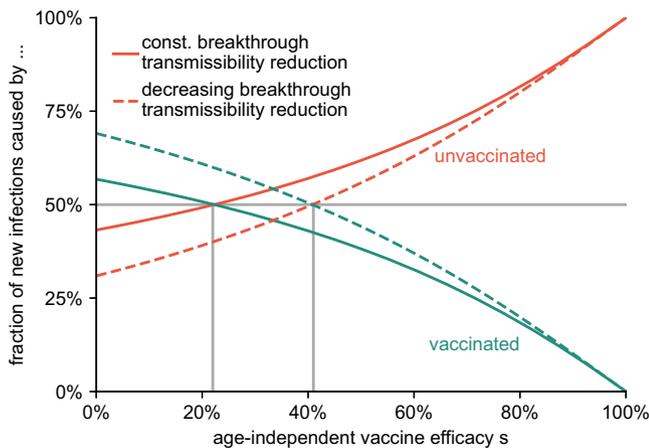
Fig. 3 Fraction of new cases caused by the unvaccinated and vaccinated population for varying age-independent vaccine efficacy s. We consider an optimistic scenario with constant $r = 0.1$ and $b = 3/2$ (solid lines), and a pessimistic estimation in which r and b decrease according to $r = s/10$ and $b = s/2 + 1$ (dashed lines). As long as s remains larger than approximately 22% (optimistic, $r' = 40\%'$) or 41% (pessimistic, $r' = 20\%$), the unvaccinated minority still causes the majority of infections, see also Supplementary Methods, Sec. 1.3.1 and Sec. 1.3.7.

caused by vaccinated infecting unvaccinated individuals and vice versa with a constant factor $m \in [0, 1]$ such that $m = 1$ refers to our base scenario of homogeneous mixing between the two groups, Fig. 2b and Supplementary Methods Sec. 1.3.4. As expected, we find that the relative contribution to $\mathcal{R}$ made by the unvaccinated increases monotonically with decreasing m (inset of Fig. 2b). In case the system was, in fact, already in a state of heterogeneous mixing during the observational period, this implies that our main results shown in Fig. 1 present lower bounds of the contribution made by unvaccinated individuals. If mixing was decreased by additional NPIs that reduce contacts between unvaccinated and vaccinated individuals, the absolute value of $\mathcal{R}$ decreases with decreasing mixing m. This illustrates the efficacy such NPIs would have towards mitigation, assuming that the reduced inter-group contacts are not balanced by

increased intra-group contact numbers. In the latter case, an increased number of contacts between unvaccinated individuals could even lead to an increase in $\mathcal{R}$, potentially worsening the situation.

Ultimately, we investigate how different the situation would have been if vaccine uptake was higher than 65% in the fall of 2021. To this end, we choose the "medium efficacy" scenario, but increase the respective vaccine uptake for adolescents, adults, and elderly to 90% each, leading to an 80% uptake in the total population, Fig. 2c and Supplementary Methods Sec. 1.3.6. In this case, the effective reproduction number would be lowered to a value of $\mathcal{R} = 0.86$ instead of $\mathcal{R} = 1.2$, implying epidemic control. Because more people would be vaccinated, both the relative and absolute contributions of vaccinated individuals to $\mathcal{R}$ would increase. Yet, the most important differences to the base scenario of $v = 65\%$ are the respective reductions of the absolute contributions of unvaccinated individuals, which would decrease from (i) $C_{u \leftarrow u} + C_{v \leftarrow u} = 0.8$ to $C_{u \leftarrow u} + C_{v \leftarrow u} = 0.37$ for infections caused and (ii) from $C_{u \leftarrow u} + C_{u \leftarrow v} = 0.67$ to $C_{u \leftarrow u} + C_{u \leftarrow v} = 0.3$ for becoming infected, both more than halved (see Supplementary Tables 4 and 5). Because unvaccinated infecteds have a much higher probability of suffering from severe disease and being hospitalized, such a reduction can be substantial for relieving an overburdened public health system.

Discussion

After vaccine rollout programs in Germany have slowed down in late summer, incidences were rising to unprecedented levels in the fall of 2021, with hospitals and ICUs having reached maximum capacity. As about 41% of reported cases aged 12 or above were recorded as breakthrough infections in October 2021, two questions naturally arise: (i) How much were the vaccinated still contributing to the infection dynamics and (ii) how need NPIs to be targeted and calibrated to help achieve temporal epidemic control?

Here, we developed a model-based framework that allows for quantifying the contributions of different infection pathways between and across vaccinated and unvaccinated groups towards the effective reproduction number $\mathcal{R}$. Based on this framework and reasonable assumptions regarding vaccine efficacy, we conclude that about 61%–76% percent of the effective reproduction number were caused

by unvaccinated individuals, with 32%–51% of its value determined by unvaccinated individuals infecting other unvaccinated individuals. Depending on the assumed efficacy scenario, 34%–50% of the infections are expected to be breakthrough infections. Although these numbers might seem comparatively large at first glance, such results that focus solely on the presence or absence of an infection (not the severity) are to be expected[55]. Our study highlights the importance of analyzing the limited contribution these breakthrough cases make towards the overall infection dynamics, especially in relation to the size of the respective vaccinated/unvaccinated subpopulations. Additionally, such proportions of breakthrough infections are not necessarily indicative of a potential burden to the public health system, as all vaccines against COVID-19 have been reported to substantially reduce the risk of a severe course of the disease[11,15–17].

We further showed that targeted NPIs that would decrease the transmissibility of unvaccinated individuals by 22%–27% could have suppressed epidemic growth reaching $\mathcal{R}<1$, under the assumption that vaccinated individuals would continue to behave as before, i.e., with no additional NPIs in place for this respective group. Yet, it is questionable how well NPIs can be targeted towards single subpopulations, both for ethical and pragmatic reasons. We found that for NPIs that would affect both unvaccinated and vaccinated individuals, those that reduce the transmissibility of the unvaccinated two to three times as strongly as the vaccinated population would reduce $\mathcal{R}$ in the most efficient manner.

Our assumptions regarding vaccine efficacy against transmission (effective transmissibility reduction) were lower than values observed in the Netherlands[52]. Assuming that the efficacy is of larger value would further increase the contributions of unvaccinated individuals towards the infection dynamics. Similarly, if children and adolescents were found to be as susceptible and infectious as adults, the contributions made by the unvaccinated subpopulation would be of larger value as well.

The analyses performed here represent model-based estimations that are limited by data quality and a large number of parameters that have to be estimated based on available empirical results. This includes epidemiological data as well as contact data from the POLYMOD study, which is already over 15 years old and might therefore inaccurately portray the mixing behavior of the German population at the time of writing. A further limiting factor is that the under-ascertainment of breakthrough infections might be larger than accounted for, as vaccinated infecteds experiencing mild symptoms might not be as likely to have their infection reported, thus leading to a potential overestimation of vaccine efficacy. Yet, vaccinated individuals might have increased their contact behavior compared to unvaccinated individuals, a behavioral change that compensates for the vaccine-induced lowered individual risk of infection. Because vaccine efficacies were estimated using Farrington's method[11,55], such a relative increase in contact behavior of vaccinated individuals could lead to an underestimation of the true vaccine efficacy, thus potentially balancing a hypothetical inequality in ascertainment. Due to such uncertainties, future empirical studies, e.g. using contact tracing data, will be necessary to confirm or refute our claims.

While we consider population mixing across age groups, we also implicitly assume homogeneous mixing between vaccinated and unvaccinated individuals in our base scenarios. Yet, the intention to vaccinate has been shown to follow rules of social contagion, rendering it likely that vaccinated and unvaccinated individuals mix less across groups[56]. We showed that, in this case, the contribution of unvaccinated individuals to $\mathcal{R}$ would be of even larger magnitude. NPIs that reduced contacts between both subpopulations (i.e. reduced mixing) would lead to a decrease in $\mathcal{R}$, as long as these reductions are not balanced by an increase in contacts among unvaccinated individuals, in which case $\mathcal{R}$ might even increase, highlighting the necessity for well-targeted measures. We want to stress that one should be careful, however, not to overinterpret this result as explicit advice for future NPIs to increase segregation between the vaccinated and unvaccinated. Indeed, other research shows that after measures that restricted access to shopping and leisure activities only for the unvaccinated, societal polarization was high[56]. While this may reduce mixing, it creates other, potentially worse societal problems. Our analysis does not account for any psychological or socio-cultural consequences of such policies or recommendations[57] and, as always, recommendations should be weighed against potential risks.

Finally, an increased vaccine uptake would increase both the relative and absolute contributions that the vaccinated population makes towards $\mathcal{R}$ while similarly decreasing the effective reproduction number's absolute value, potentially leading to temporary epidemic control under the assumption of unchanged behavior. In light of the slow growth of vaccine uptake in Germany after the summer 2021[21] and low intention to vaccinate among those that are unvaccinated[41], such an increase in uptake, however, seems unlikely to be achieved.

We furthermore stress that our results are estimations made for the comparatively short period between October 11, 2021 and November 7, 2021. As vaccine efficacy against infection has been reported to decrease with time, fast and wide-spread booster vaccination is a crucial measure to avoid an increasing reproduction number and a potentially worsening public health crisis. Also, the spread of immune escape variants may change the situation.

In summary, our results suggest that a minority of the population (i.e., the unvaccinated) contributed a substantial part to the infection dynamics, thus making them the primary driver of the public health crisis in Germany during the fourth wave of the COVID-19 pandemic and presumably also in other countries that were experiencing similar dynamics. We also show that this effect can be compensated through targeted NPIs that effectively lower the transmissibility of infected, yet unvaccinated, individuals. Hence, our study further underlines the importance of vaccines as a pharmaceutical intervention regarding epidemic control and highlights the importance of increasing vaccine uptake, e.g. through campaigning or low-threshold offers, wherever possible, in order to achieve efficient and long-term epidemic control and preventing an overload of public health systems.

Data availability

Code availability

References
1. World Health Organisation. WHO Coronavirus (COVID-19) Dashboard (2021). https://covid19.who.int. Accessed: 2021-11-19.
2. Keeling, M. J. & Rohani, P. Modeling infectious diseases in humans and animals (Princeton University Press, 2011). https://doi.org/10.2307/j.ctvcm4gk0.

3. Haug, N. et al. Ranking the effectiveness of worldwide COVID-19 government interventions. *Nat. Human Behav.* **4**, 1303–1312 (2020).

4. Maier, B. F. & Brockmann, D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science* **368**, 742–746 (2020).

5. Habersaat, K. B. et al. Ten considerations for effectively managing the COVID-19 transition. *Nat. Human Behav.* **4**, 677–687 (2020).

6. Bajos, N. et al. When lockdown policies amplify social inequalities in COVID-19 infections: evidence from a cross-sectional population-based survey in France. *BMC Public Health* **21**, 705 (2021).

7. Elmer, T., Mepham, K. & Stadtfeld, C. Students under lockdown: Comparisons of students' social networks and mental health before and during the COVID-19 crisis in Switzerland. *PLOS ONE* **15**, 1–22 (2020).

8. Das, M., Das, A. & Mandal, A. Examining the impact of lockdown (due to COVID-19) on Domestic Violence (DV): An evidences from India. *Asian J. Psychiatry* **54**, 102335 (2020).

9. Atalan, A. Is the lockdown important to prevent the COVID-19 pandemic? Effects on psychology, environment and economy-perspective. *Ann. Med. Surg.* **56**, 38–42 (2020).

10. Intensivregister-Team am RKI. Tagesreport aus dem DIVI-Intensivregister https://doi.org/10.25646/9290 (2021).

11. Robert Koch Institute. Wöchentlicher Lagebericht des RKI zur Coronavirus-Krankheit COVID-19 – Wochenbericht vom 11.11.2021 https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/Wochenbericht/Wochenbericht_2021-11-11.pdf (2021). Accessed: 2021-11-19.

12. an der Heiden, M. SARS-CoV-2-Nowcasting und -R-Schaetzung https://doi.org/10.5281/zenodo.5712013 (2021).

13. Dzinamarira, T. et al. Unpacking the implications of SARS-CoV-2 breakthrough infections on COVID-19 vaccination programs. *Vaccines* **10**, 252 (2022).

14. Diekmann, O., Heesterbeek, J. A. P. & Roberts, M. G. The construction of next-generation matrices for compartmental epidemic models. *J. Royal Society, Interf.* **7**, 873–885 (2010).

15. Voysey, M. et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet* **397**, 99–111 (2021).

16. Polack, F. P. et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N. Eng. J. Med.* **383**, 2603–2615 (2020).

17. Baden, L. R. et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N. Eng. J. Med.* **384**, 403–416 (2020).

18. Mossong, J. et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74 (2008).

19. Mossong, J. et al. Polymod social contact data https://doi.org/10.5281/zenodo.1157934 (2017). Version 1.1.

20. Funk, S. socialmixr: R package for deriving social mixing matrices from survey data, github.com/epiforecasts/socialmixr (2020).

21. Robert Koch Institute, Fachgebiet 33. COVID-19-Impfungen in Deutschland https://doi.org/10.5281/zenodo.5126652 (2021).

22. Scholz, S. et al. Einfluss von Impfungen und Kontaktreduktionen auf die dritte Welle der SARS-CoV-2-Pandemie und perspektivische Rückkehr zu prä-pandemischem Kontaktverhalten. *Epidemiologisches Bull.* **13**, 3–22 (2021).

23. Otte im Kampe, E., Lehfeld, A.-S., Buda, S., Buchholz, U. & Haas, W. Surveillance of COVID-19 school outbreaks, Germany, March to August 2020. *Euro Surveill.* **25**, 2001645 (2020).

24. Dattner, I. et al. The role of children in the spread of COVID-19: Using household data from Bnei Brak, Israel, to estimate the relative susceptibility and infectivity of children. *PLOS Comput. Biol.* **17**, 1–19 (2021).

25. Bistaraki, A., Roussos, S., Tsiodras, S. & Sypsa, V. Age-dependent effects on infectivity and susceptibility to SARS-CoV-2 infection: results from nationwide contact tracing data in Greece. *Infect. Dis.* **54**, 186–195 (2022).

26. Robert Koch Institute. SurvStat@RKI 2.0 – Web-basierte Abfrage der Meldedaten gemäß Infektionsschutzgesetz (IfSG) (2021). Accessed: 2021-11-19.

27. Ong, S. W. X. et al. Clinical and virological features of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants of concern: A retrospective cohort study comparing B.1.1.7 (Alpha), B.1.351 (Beta), and B.1.617.2 (Delta). Clinical Infectious Diseasesciab721 https://doi.org/10.1093/cid/ciab721 (2021).

28. Li, B. et al. Epidemiological and clinical characteristics of COVID-19 in children: A systematic review and meta-analysis. *Frontiers in Pediatrics* **8**, 2296–2360 (2020).

29. Berger, U., Fritz, C. & Kauermann, G. Eine statistische Analyse des Effekts von verpflichtenden Tests an Schulen mit Präsenzunterricht im Vergleich zum Distanzunterricht238 https://doi.org/10.5282/ubm/epub.76005 (2021).

30. Betsch, C. et al. COSMO – Zusammenfassung und Empfehlungen Wellen 52 und 53 https://projekte.uni-erfurt.de/cosmo2020/web/summary/52-53/ (2021).

31. Betsch, C. et al. COVID-19 Snapshot Monitoring (COSMO Germany) – Wave 52 https://doi.org/10.23668/PSYCHARCHIVES.5145 (2021). Dataset.

32. Betsch, C. et al. COVID-19 Snapshot Monitoring (COSMO Germany) – Wave 53 https://doi.org/10.23668/PSYCHARCHIVES.5152 (2021). Dataset.

33. Neuhauser, H. et al. Germany's low SARS-CoV-2 seroprevalence confirms effective containment in 2020: Results of the nationwide RKI-SOEP study. medRxiv https://doi.org/10.1101/2021.11.22.21266711 (2021). Preprint.

34. Maier, B. F. et al. Modeling the impact of the Omicron infection wave in Germany. medRxiv https://doi.org/10.1101/2022.07.07.22277391 (2022). Preprint.

35. Feikin, D. R. et al. Duration of effectiveness of vaccines against SARS-CoV-2 infection and COVID-19 disease: results of a systematic review and meta-regression. *Lancet* **399**, 924–944 (2022).

36. Reis, B. Y. et al. Effectiveness of BNT162b2 vaccine against Delta variant in adolescents. *N. Eng. J. Med.* **385**, 2101–2103 (2021).

37. Chia, P. Y. et al. Virological and serological kinetics of SARS-CoV-2 Delta variant vaccine breakthrough infections: a multicentre cohort study. *Clin. Microbiol. Infect.* **28**, 612.e1–612.e7 (2022).

38. Riemersma, K. K. et al. Shedding of infectious SARS-CoV-2 despite vaccination. medRxiv https://doi.org/10.1101/2021.07.31.21261387 (2022). Preprint.

39. Elliott, P. et al. Exponential growth, high prevalence of SARS-CoV-2, and vaccine effectiveness associated with the Delta variant. *Science* **374**, eabl9551 (2021).

40. Thompson, M. G. et al. Prevention and attenuation of Covid-19 with the BNT162b2 and mRNA-1273 vaccines. *New England Journal of Medicine* **385**, 320–329 (2021).

41. Betsch, C. et al. COVID-19 Snapshot Monitoring (COSMO Germany) – Wave 56. PsychArchives https://doi.org/10.23668/psycharchives.5235 (2021).

42. de Gier, B. et al. Vaccine effectiveness against SARS-CoV-2 transmission to household contacts during dominance of Delta variant (B.1.617.2), the Netherlands,August-September 2021. *Euro Surveill.* **26**, 2100977 (2021).

43. Eyre, D. W. et al. Effect of Covid-19 vaccination on transmission of Alpha and Delta variants. *N. Eng. J. Med.* **386**, 744–756 (2022).

44. UK Health Security Agency. COVID-19 vaccine surveillance report – Week 46 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1034383/Vaccine-surveillance-report-week-46.pdf (2021).

45. Pouwels, K. B. et al. Effect of Delta variant on viral burden and vaccine effectiveness against new SARS-CoV-2 infections in the UK. *Nat. Med.* **27**, 2127–2135 (2021).

46. Collier, D. A. et al. Age-related immune response heterogeneity to SARS-CoV-2 vaccine BNT162b2. *Nature* **596**, 417–422 (2021).

47. Sächsisches Staatsministerium für Soziales und Gesellschaftlichen Zusammenhalt. Sächsische Corona-Notfall-Verordnung – SächsCoronaNotVO – Vom 19. November 2021 https://www.coronavirus.sachsen.de/download/SaechsCoronaNotVO-2021-11-19.pdf (2021).

48. Han, M. S. et al. Clinical Characteristics and Viral RNA Detection in Children With Coronavirus Disease 2019 in the Republic of Korea. *JAMA Pediat.* **175**, 73–80 (2021).

49. Cai, J. et al. Comparison of clinical and epidemiological characteristics of asymptomatic and symptomatic SARS-CoV-2 infection in children. *Virolog. Sinica* **35**, 803–810 (2020).

50. Ernst, E. Anthroposophy: A risk factor for noncompliance with measles immunization. *Pediatric Infect. Dis. J.* **30**, 187–189 (2011).

51. Bier, M. & Brak, B. A simple model to quantitatively account for periodic outbreaks of the measles in the Dutch Bible Belt. *Eur. Phys. J. B* **88**, 107 (2015).

52. Knol, M. J. et al. Large ongoing measles outbreak in a religious community in the Netherlands since May 2013. *Eurosurveillance* **18**, 20580 (2013).

53. Weitz, J. S. et al. Modeling shield immunity to reduce COVID-19 epidemic spread. *Nat. Med.* **26**, 849–854 (2020).

54. Kraay, A. N. et al. Modeling serological testing to inform relaxation of social distancing for COVID-19 control. *Nat. Commun.* **12**, 1–10 (2021).

55. Farrington, C. P. Estimation of vaccine effectiveness using the screening method. *Int. J. Epidemiol.* **22**, 742–746 (1993).

56. Henkel, L., Sprengholz, P., Korn, L., Betsch, C. & Böhm, R. Understanding the trouble spot: Does vaccination status identification fuel societal polarization? Preprint https://doi.org/10.31234/osf.io/mgqk5 (2022).

57. Norheim, O. F. Protecting the population with immune individuals. *Nat. Med.* **26**, 823–824 (2020).

58. Maier, B. F. Vaccontrib: Quantifying the contributions unvaccinated and vaccinated subpopulations make towards the infection dynamics. Zenodo https://doi.org/10.5281/zenodo.6811013 (2022). Software.

Acknowledgements

Author contributions

B.F.M. developed the initial research idea. Analyses were performed by B.F.M, following discussions with M.W., A.B., P.P.K., M.A.J., C.B., and D.B. A.B. performed additional literature research. B.F.M. and M.W. wrote the first manuscript draft, which was subsequently edited by B.F.M., M.W., A.B., P.P.K., M.A.J., C.B., and D.B.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43856-022-00176-7.

Correspondence and requests for materials should be addressed to Benjamin F. Maier.

Peer review information *Communications Medicine* thanks Peter Klimek and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.