Age-specific transmission dynamics of SARS-CoV-2 during the first two years of the pandemic

Otilia Boldea[a,1], **Amir Alipoor**[a], **Sen Pei**[b], **Jeffrey Shaman**[b,c], **and Ganna Rozhnova**[d,e,f,1]

[a]Department of Econometrics and OR, Tilburg School of Economics and Management, Tilburg University, Tilburg, The Netherlands; [b]Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, USA; [c]Columbia Climate School, Columbia University, New York, USA; [d]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands; [e]BioISI—Biosystems & Integrative Sciences Institute, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal; [f]Center for Complex Systems Studies (CCSS), Utrecht University, Utrecht, The Netherlands

This manuscript was compiled on November 23, 2023

During its first two years, the SARS-CoV-2 pandemic manifested as multiple waves shaped by complex interactions between variants of concern, non-pharmaceutical interventions, and the immunological landscape of the population. Understanding how the age-specific epidemiology of SARS-CoV-2 has evolved throughout the pandemic is crucial for informing policy decisions. In this paper, we aimed to develop an inference-based modeling approach to reconstruct the burden of true infections and hospital admissions in children, adolescents, and adults over the seven waves of four variants (wild-type, Alpha, Delta, Omicron BA.1) during the first two years of the pandemic, using the Netherlands as the motivating example. We find that reported cases are a considerable underestimate and a generally poor predictor of true infection burden, especially because case reporting differs by age. The contribution of children and adolescents to total infection and hospitalization burden increased with successive variants and was largest during the Omicron BA.1 period. However, the ratio of hospitalizations to infections decreased with each subsequent variant in all age categories. Before the Delta period, almost all infections were primary infections occurring in naive individuals. During the Delta and Omicron BA.1 periods, primary infections were common in children but relatively rare in adults who experienced either re-infections or breakthrough infections. Our approach can be used to understand age-specific epidemiology through successive waves in other countries where random community surveys uncovering true SARS-CoV-2 dynamics are absent but basic surveillance and statistics data are available.

COVID-19 | age-specific time-varying epidemiology | infection and hospitalization burden | stochastic metapopulation model | ensemble adjustment Kalman filter

During the pandemic, the dynamics of SARS-CoV-2 demonstrated a complex spatio-temporal pattern with multiple waves (1) and pronounced differences in the age-specific burden of confirmed cases and hospitalizations (2, 3). A notable example is the age distribution of reported cases in European countries, including the Netherlands, where a much lower number of cases in younger individuals were reported in the first wave of the wild-type variant, compared to the Omicron BA.1 wave (4).

Understanding how the age-specific epidemiology of SARS-CoV-2 has changed during the pandemic (5) is crucial for informing public health policy. For instance, information about how the true rather than the reported infection burden varies by age and time and the contribution of different age groups to transmission (6, 7) may inform non-pharmaceutical interventions (8–12) like school- and non-school-based measures (13), while understanding of the age-specific hospitalization burden (14) underpins prioritization of vaccination (15, 16). Some of this information can be provided by surveillance and serological surveys. However, full reconstruction of the age-specific burden of infections and hospitalizations in a country is complicated by several factors. Firstly, under-reporting of cases is age-specific and time-varying due to peculiarities of surveillance systems and testing policies, which varied across ages and time as the pandemic progressed. This fact coupled with the evidence of asymptomatic infection (17) undermines the ability of surveillance to capture the true burden of infections among different age groups. Few countries conduct random community PCR testing (18). Secondly, nationally representative serological surveys provide information on which subpopulations carry antibodies to SARS-CoV-2 and thus could help to characterize prior infection burden (19, 20). However, due to the waning of immunity after vaccination and infection, the immunological landscape of the population has become increasingly complex (21). Implementing representative serosurveys that estimate population immunity among different age groups

Significance Statement

To understand how to prevent and mitigate future pandemics of respiratory viruses, it is imperative that we understand the trajectory of SARS-CoV-2. Knowledge of how the age-specific burden of infections and hospitalizations has changed during the pandemic is crucial for designing public health interventions. Here, we investigate how SARS-CoV-2 dynamics unfolded in a complex spatio-temporal pattern of multiple waves shaped by interactions between variants, interventions, and accumulating immunity of the population. Using a computationally efficient model, we reconstruct total (reported and unreported) infection and hospitalization burdens over the first two years of the pandemic, stratified by variant (wild-type, Alpha, Delta, and Omicron BA.1), age (children, adolescents, and adults) and immune status (naïve and partially susceptible individuals after vaccination or waning of immunity after primary infection).

O.B. and G.R. designed the study, with important feedback from all other authors. O.B. and G.R. obtained the data. O.B. and A.A. designed and wrote the computer code, with substantial feedback from S.P. G.R. and O.B. wrote the manuscript, with substantial feedback from S.P. and J.S. O.B. wrote the Supplementary Materials, with substantial feedback from all other authors.

J.S. and Columbia University declare partial ownership of SK Analytics. The other authors declare no competing interests.

[1]To whom correspondence should be addressed. E-mails: g.rozhnova@umcutrecht.nl and o.boldea@tilburguniversity.edu.

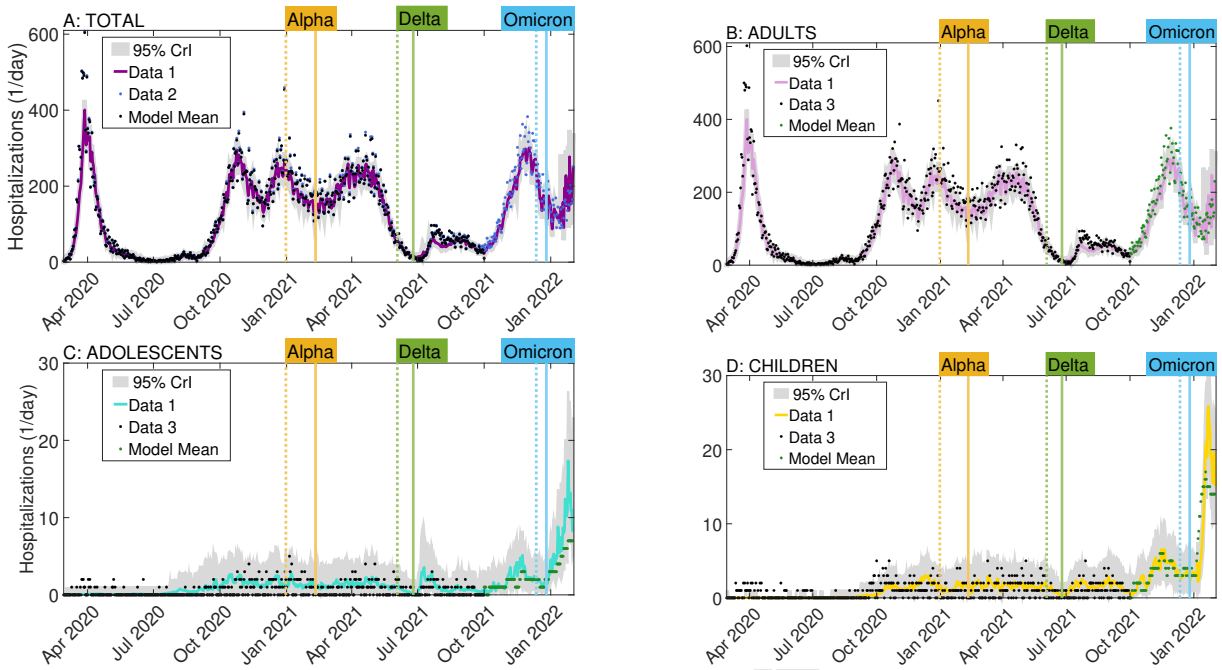


Fig. 1. Estimated hospital admissions. Total (**A**) and age-specific national daily hospital admissions in adults (**B**), adolescents (**C**) and children (**D**). The colored lines represent the estimated posterior means for adults (pink), adolescents (turquoise), children (yellow), and all ages (purple). The gray-shaded regions correspond to posterior 95% credible intervals defined as the 2.5% and 97.5% quantiles from 300 posterior ensemble values. The dots are daily hospital admission data used for fitting the model. Data 1 refers to data obtained directly from the National Institute for Public Health and the Environment (RIVM); Data 2 and 3 refer to the RIVM Dashboard data (see Table 1 and Supplementary Materials, Section 1.3). The dashed and solid vertical lines indicate when each VoC corresponded to 5% and 50% of samples in the genetic variant data described in Supplementary Materials, Section 1.8.

characterized by varying numbers of prior infections before and after vaccination is difficult, costly, and time-consuming. Thirdly, additional factors including variants of concern (VoCs) (22), non-pharmaceutical interventions (8–12) and changes in population immunity after vaccination or infection complicate estimation of the age-specific burden.

Modelling studies have provided important insights into temporal changes in the epidemiology of SARS-CoV-2 in specific countries (13, 20, 23–29). Some of these were hypothesis-generating and not rigorously validated against all real data evidence (25–27). Other studies did not reconstruct age-specific epidemiology (24) or were limited to specific periods of the pandemic such as the first wave (13, 30–33) or periods of dominance for the Alpha (20, 23) and Delta variants (28, 34). As data accumulate, formal evaluations based on mathematical models estimated based on different types of observational data (13, 24, 28) are crucial for reconstructing the burden of infections and hospitalizations over long time frames.

Here we reconstruct the epidemiology of SARS-CoV-2 in the Netherlands over the first 23 months of the pandemic, a period including wild-type, Alpha, Delta, and Omicron BA.1 waves, using an age- and regionally stratified transmission model fitted to various data sources (see Table 1 and Supplementary Materials). Our fitting method is a quasi-Bayesian estimation (13, 23, 35–37) based on an ensemble adjustment Kalman filter (24, 38, 39), where in the fitting, we combined surveillance and national statistics data typically available from individual countries (hospital admissions, serological surveys, PCR testing data, genetic VoC data, vaccination coverage data, social contact matrices, demographic data, regional train, and Google mobility data) with the Netherlands used as the motivational example.

To account for control measures targeted at elementary and secondary schoolchildren versus the rest of the population (13) and for age-dependent transmission effects (3, 14, 17, 40–43), the population is stratified into young children (0 to 9 years old), adolescents (10 to 19 years old) and adults (above 19 years old) in twelve Dutch provinces. The model structure is rooted in current knowledge of SARS-CoV-2, which suggests waning of immunity after infection or vaccination (21) and potential changes in susceptibility, infectivity, and severity of re-infections and breakthrough infections (25, 26). The regional stratification is augmented with real-world mobility across the provinces under the assumption that infected individuals with undocumented infection may travel to and infect susceptible individuals in other provinces whereas infected individuals with confirmed infection stay in their province of origin. Previous studies indicate that epidemic models with transmission dynamics coupled across locations can identify the true number of infections in the populations and improve the identifiability of epidemiological parameters (24, 44, 45).

Results

Time-dependent burden of hospital admissions. The model is fitted to daily new reported cases and hospitalizations in adults, adolescents, and children across the twelve provinces of the Netherlands, as well as to four national seroprevalence surveys, allowing estimation of key time-varying epidemiological parameters compatible with national and regional epidemiology of SARS-CoV-2 in pre- and post-vaccination periods (Supplementary Tables A1—A6 and Figures A1—A4; see also Section "Parameter identifiability and sensitivity analyses").

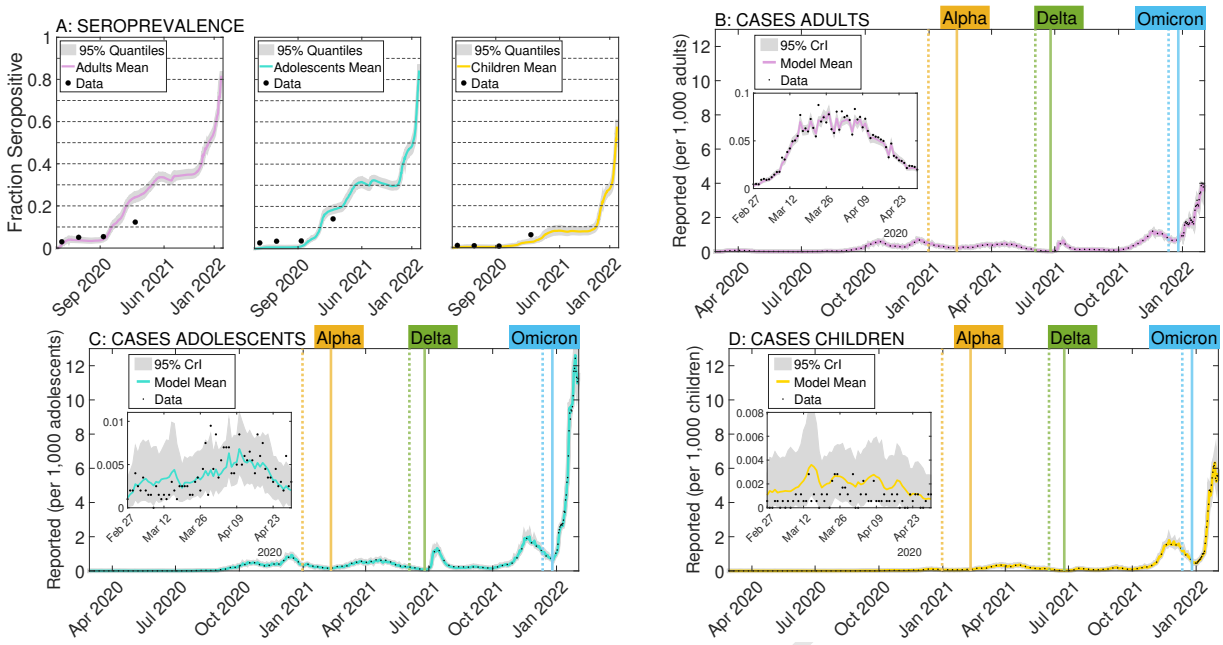


Fig. 2. Estimated seroprevalence and confirmed cases. Age-specific national seroprevalence due to infection (**A**) and national confirmed daily cases in adults (**B**), adolescents (**C**) and children (**D**). Inset: the first wave characterized by low case reporting. The colored lines represent the estimated means for adults (pink), adolescents (turquoise), and children (yellow). The gray-shaded regions correspond to posterior 95% credible intervals defined as the 2.5% and 97.5% quantiles from 300 posterior ensemble values. The black dots are seroprevalence (**A**) and daily confirmed cases (**B**, **C**, **D**) data used for fitting the model (see Table 1 and Supplementary Materials, Sections 1.2 and 1.4). The dashed and solid vertical lines indicate when each VoC corresponded to 5% and 50% of samples in the data, respectively. For comparison, the scale of the y-axis is the same in **B**, **C** and **D**.

Our estimates reproduce very well national admissions in young children and adolescents, and relatively well national and regional admissions in adults during all waves that occurred from the first official case on February 27, 2020 until January 31, 2022 (Figure 1, Supplementary Figures A5—A7). The total estimated mean number of hospital admissions is 83,589 (95% CrI 81,637—85,156), which amounts to 0.48% of the population (95% CrI 0.47%—0.49%) (Supplementary Table A7), compared to 0.52% of the population being reported hospitalized over seven waves in the wild-type, Alpha, Delta and Omicron BA.1 periods, with the last Omicron wave starting but not yet ending at the end of the study period, January 31, 2022 (Figure 1 **A**). Of these, two waves occurred before vaccination, one wave peaked in January 2021 around the onset of the vaccination program, and another four afterwards. As expected, the pattern of hospital admissions in adults largely mirrors that of total national admissions due to much higher probability of clinical disease and hospitalization in this subpopulation compared to adolescents and children (3, 13, 14, 23) (Figure 1 **B**). The burden of hospital admissions in adults progressively decreased over the periods of successive VoCs, in line with the expansion of the primary vaccination series and of the first booster campaign (Supplementary Table A7). The estimated cumulative mean number of hospital admissions for adults was thus largest for the wild-type (30,247, 95% CrI 29,422—31,029, or 37.19%, 95% CrI 36.43%—37.93%, of total estimated hospital admissions among adults) and smallest for the incomplete Omicron BA.1 period until January 31, 2022 (7,736, 95% CrI 7,386—8,187, or 9.51%, 95% CrI 9.07%—10.02%, of total estimated hospital admissions among adults) (Supplementary Table A7). The picture is different for adolescents and children for whom the hospitalization burden stayed lower than 5 hospital admissions per day for either group and did not demonstrate a pronounced pattern until October 2021 (Figure 1 **C** and **D**). Unlike for adults, for which 28.73% (95% CrI 27.97%—29.49%) of hospitalizations occurred during Delta and Omicron BA.1 periods, a much larger share of hospitalizations in children and adolescents occurred during the same periods: 56.73% (95% CrI 44.40%—66.45%) for adolescents and 68.69% (95% CrI 57.04%—80.75%) for children. While the number of hospital admissions in adolescents and children remained much lower than those of adults (for adolescents, 296, 95% CrI 252—352 hospitalizations during Omicron BA.1, and 248, 95% CrI 176—399, during Delta; for children, 469, 95% CrI 375—584 during Omicron and 411, 95% CrI 247—659 during Delta, relative to a population of approximately 2 million adolescents and 1.8 million children), and the hospitalization rates out of reported infected decreased with each variant period (Supplementary Table A3), we observe a shift in hospitalization burden towards children at the end of the sample, whose source we further document in the next three sections.

Time-dependent burden of confirmed cases and seroprevalence. The model further reproduces the age-specific seroprevalence (Figure 2 **A**) and confirmed cases nationally (Figure 2 **B**-**D**) and regionally (Supplementary Figures A8—A16). In the context of our study, the seroprevalence is the fraction of the population who have antibodies due to (any) infection. The level of seroprevalence is thus determined by how fast antibodies decay and how fast new infections happen. Confirmed cases refer to the part of infections captured by PCR testing surveillance and the model fit to these data is very good for all ages despite differential changes in surveillance eligibility

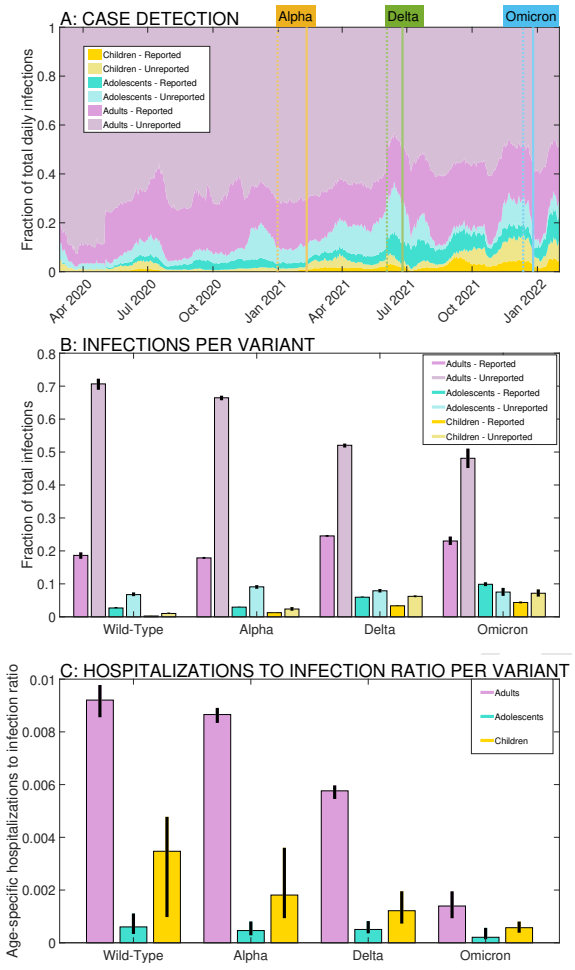


Fig. 3. Age distribution of estimated total infections and hospital admissions by VoC. (**A**) The time-varying estimated fraction of reported and unreported cases in the total national daily cases. (**B**) The estimated fraction of reported and unreported cases in total national cases, but now cumulative per VoC period; for each VoC, the bars sum up to 1. (**C**) Estimated age-specific hospital admissions as a fraction of total cumulative cases in that age group for each VoC period. The bars and the black lines show the means and 95% credible intervals obtained from 300 posterior ensemble values.

and recommendations across age categories. Our estimates show that the national seroprevalence steadily increased in all age groups reaching 82% (95% CrI 80%—84%) in adults, 84% (95% CrI 81%—87%) in adolescents but only 58% (95% CrI 54%—61%) in children by 31 January 2022 (Figure 2 **A** and Supplementary Table A8). This is in line with our susceptibility parameter estimates (Supplementary Table A3), which indicate that, even when controlling for variants, school closures, non-pharmaceutical interventions or holidays, the susceptibility of children remains much lower than that of adolescents and adults until the Omicron BA.1 period. In contrast with the magnitude of the first wave of hospital admissions in spring 2020, the case reporting in this period was overall low, and more so among children and adolescents (insets Figure 2 **B**, **C** and **D**), with estimated average case detection rates of 8% (95% CrI 5%—9%) in adults, 4% (95% CrI 3%—8%) in adolescents, and 4% (95% CrI 2%—9%) in children until March 30, 2020, when reported cases first peaked. Throughout most of the study period that covers the next five waves, less than 1 case per 1,000 adults was reported daily, but this sharply increased during the Omicron BA.1 wave to 4 cases per 1,000 adults by the end of January 2022 (Figure 2 **B**). The pattern of reported cases was similar in adolescents with less than 1 case per 1,000 adolescents daily during the first five waves; however, substantially more Omicron BA.1 cases were reported in this age group compared to adults during the sixth and seventh waves, skyrocketing from 1 to 12 daily cases per 1,000 adolescents in January 2022 (Figure 2 **C**). More cases were documented in children than in adults in the Omicron BA.1 wave too, i.e., about 6 versus 4 cases per 1,000 individuals per day at the peak, while case reporting in children was much lower than in adults during the rest of the study period (Figure 2 **D**). The observed increase in reported cases in adolescents and children compared to adults at a time when testing in schools was no longer compulsory may indicate a higher burden of infections in these subpopulations during the Omicron BA.1 period and a change in the age-specific distribution of total national cases which we verify further below, by presenting estimates of the share of age-specific reported and unreported cases in the total number of cases.

Age distribution of total infections and hospital admissions per VoC. Our model estimates temporal changes in the age-specific contact rates as a result of implemented control measures (Supplementary Materials, Section 2.4). These changes combined with using reported infected data and a model assumption that unreported infected travel across provinces and infect others, while reported infected stay in their province, allow identification of unreported cases. The estimated fraction of reported cases in total reported and unreported cases in the population (Figure 3 **A**) increased over time in all age categories, in line with the expansion of testing capacity. At the peak of each wave, the estimated fraction of reported to total cases plummeted, in line with hitting test capacity limits. For adults, we estimate that no more than 37% of new cases were detected at each point in time. Case detection rates in children and adolescents were initially very low and increased over time, likely due to changes in testing rules, recommendations, and contact tracing in schools in 2020 and 2021.

Figure 3 **B** shows again the age-specific distribution of reported and unreported cases in the total national infections

but now aggregated over each VoC period. The fraction of total infections in children among total estimated national infections (light and dark yellow bars) steadily increased from 1.3% (95% CrI 1.1%—1.5%) in the wild-type period to 11.5% (95% CrI 10.4%—12.6%) in the Omicron BA.1 period, of which about 36% (95% CrI 34.8%—37.3%) were reported in total (Supplementary Tables A9—A10). A similar increasing contribution to total infections per successive VoC periods is estimated for adolescents (light and dark turquoise bars), namely between 9.4% (95% CrI 8.9%—10.3%) and 17.4% (95% CrI 16%—18.7%) of all national infections for the wild-type and Omicron BA.1 periods, respectively, out of which 43.4% (95% CrI 42.5%—44.3%) were unreported in total. The contribution of adults to total infections across all ages (light and dark pink bars) decreased from 89.3% (95% CrI 88.2%—90.1%) in the wild-type period to 71.1% (95% CrI 69.1%—73.2%) during Omicron, while the fraction of estimated reported adult cases in total adult cases per variant increased from 20.9% (95% CrI 19.6%—22.1%) during the wild-type period to 32.3% (95% CrI 31.9%—32.7%) during the Omicron BA.1 period. Over the entire sample, an estimated 26.6% of adult cases were reported (95% CrI 26.1%—27%). The decrease in the contribution of adults to total infections confirms a shift in infection burden towards children and adolescents.

The hospitalizations to total infection ratio steadily decreased in all age categories with subsequent variants, but at a higher rate for adults, who experienced the highest burden in each wave and were vaccinated first (Figure 3 **C** and Supplementary Table A11). Figure 3 **C** also shows that the hospitalization burden per total infections in children did not increase during the Omicron BA.1 period, further supporting the conclusion that the increase in the number of hospitalized children is solely due to more infections in children in this period.

Age-specific burden stratification by immune status and VoC. We further analyzed the age-specific burden of total infections and hospital admissions stratified by immune status and period of VoC (Figure 4 and Supplementary Tables A12—A13). For this, we distinguished infections in fully susceptible (naive) individuals (primary infections) from infections in individuals who were vaccinated (breakthrough infections) or lost immunity after primary infection (re-infections) (Figure 4 **A**). Hospital admission burden was stratified into hospital admissions after primary infection versus after re-infection or breakthrough infection despite vaccination (Figure 4 **B**). An estimated 23.6% (95% CrI 22.4%—25.1%) and 20.3% (95% CrI 19.8%—21.4%) of adults (pink bars) had a primary infection during the wild-type and Alpha periods, respectively, while re-infections or breakthrough infections were experienced in only 3.9% of adults for both periods combined. This pattern reversed during the Delta and Omicron periods as re-infections or breakthrough infections occurred in 15.9% (95% CrI 15.7%—16.7%) and 23.4% (95% CrI 23.1%—23.9%) of adults, respectively, while primary infections were experienced by only about 4% of adults during each period (Supplementary Table A12). Hospital admissions as a ratio of total (reported and unreported) infections in adults that experienced a breakthrough/re-infection also substantially increased from the wild-type period to the Delta period, even though it stayed substantially lower than the hospital admissions after a primary infection. It is not until the incomplete Omicron BA.1 period that we see a decrease in the hospitalization ratio after primary and after breakthrough or re-infections in adults (Figure 4 **B**), owing to: a) delayed vaccination schedules that only took off massively during the Delta period; and b) lower virulence of Omicron. The burden of primary infections in adolescents (turquoise bars), as a share of their population, was only slightly lower than that of adults for the wild-type (17.2% of the age group, 95% CrI 16.2%—19.4%) and slightly higher for Alpha VoC (21.4% of the age group, 95% CrI 20%—24.2%). These findings are driven by our age-specific susceptibility parameter estimates for adolescents, which are not lower for most of the sample, and for parts of the sample are even higher, than those of adults (Supplementary Figure A1). However, unlike for adults, primary infections in adolescents were more common than breakthrough infections or re-infections during the Delta and Omicron BA.1 VoCs, too (i.e., 15.3% of the age group, 95% CrI 14.5%—16.7% for Delta, and 18.1% of the age group, 95% CrI 17.5%—18.8% for Omicron). Fewer breakthrough infections in adolescents compared to adults could be explained by fewer contacts in this group, and also by the delayed vaccination schedule for adolescents. The burden in adolescents was especially high during the Omicron BA.1 wave when an estimated 35.4% (95% CrI 34.2%—36.9%) of all adolescents were infected. The fraction of infections in children increased with time and was mainly due to primary infections. We estimated that approximately the same fraction of children were infected before Omicron and in the incomplete Omicron BA.1 period, 29.5% (95% CrI 27.7%—32.9%) and 29% (95% CrI 27.4%—31.3%), respectively. These results are mostly driven by the susceptibility of children, which is estimated to be lower than that in adults until the Omicron period (Supplementary Figure A1), and are in line with estimates and evidence from other studies (Figure 4 in (46) and (43, 47–50)). The large number of infections in children during Omicron did not result in a higher hospitalization rate per primary infections or per total infections for children during that period, underscoring that the main driver of more hospitalizations in younger children during Omicron compared to other periods is a shift towards more infections in this age category (yellow bars, Figure 4 **A** and **B**).

Parameter identifiability and sensitivity analyses. The posterior distribution of most parameters tightens after the first wave (Supplementary Materials, Section 4.1, Figures A1—A3), indicating that the data are informative in recovering model parameters, especially after the first wave. To further verify system identifiability in the first wave, we generated one synthetic outbreak and verified that the generated infection and hospitalization data follow similar trends to the original data. To assess parameter identifiability in the first wave, we considered higher, lower, and time-varying case detection rates. For each parameter combination, 100 synthetic outbreaks were generated, each of which was used as data to re-estimate the model parameters. Across the three model configurations, the "true" parameters were either within the range of the posterior mean densities, or the differences were not too large (Supplementary Materials, Section 6.2), although some parameters such as the mobility reporting error rate were less well identified.

Sensitivity analyses were conducted for model parameters that were not estimated but rather calibrated (see Methods and Supplementary Materials, Section 7), and the best-fitting

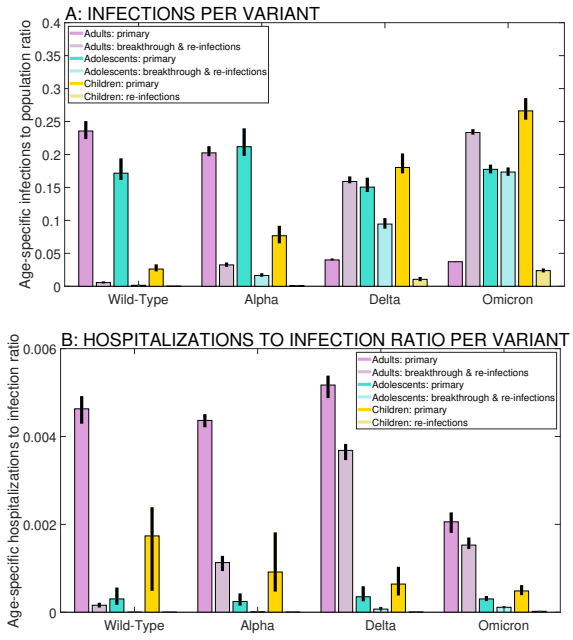


Fig. 4. Burden stratification by immune status and VoC. (**A**) Fraction of primary and breakthrough infections or re-infections, but now as a fraction of the population in each age group rather than in the total population. (**B**) Hospital admissions out of primary, respectively breakthrough or re-infections in each age category per variant. The bars and the black lines show the means and 95% credible intervals obtained from 300 posterior ensemble values.

model was selected based on maximizing the pseudo-likelihood of the model calculated by the ensemble adjustment Kalman filter.

Discussion

We developed an inference-based modeling approach to reconstruct how the age-specific epidemiology of SARS-CoV-2 changed through time and was shaped by the complex interaction between VoCs, non-pharmaceutical interventions, and the immunological landscape of the population. Our study provides several insights into changes in the age-specific burden of infections and hospitalizations over seven waves with different VoCs during the first two years of the pandemic in the Netherlands. Firstly, cases reported to surveillance were a considerable underestimate of total infections, as the majority of infections in all age groups remained unidentified despite overall improvement of case reporting through time. Reported cases are a generally poor predictor of true incidence of infection, especially because reporting differs substantially by age but also by variant periods and, within a variant period, by waves. Secondly, the contribution of children and adolescents to total infection increased with successive VoCs and was largest during the Omicron BA.1 period; the number of hospitalized children also increased, but we found that this is due to the steep increase in the number of infections in children during this period. Thirdly, we observed a shift in the pattern of infections and hospitalizations by immune status and age. Before Delta, almost all infections were primary infections occurring in naive individuals. However, during the Delta and Omicron BA.1 periods primary infections were common in children who were infected less frequently early on (due to their relative susceptibility being estimated to be substantially lower than those of adults), but were relatively rare in adults who experienced either re-infections or breakthrough infections.

A similar shift in the dynamics of infections was reported in a study based on reported case data only from the UK (51). However, unlike in (51) where re-infections dramatically increased during the Omicron wave, in our study the shift in primary infections versus re-infections and breakthrough infections occurred earlier, in the Delta period. Our findings can be attributed to multiple factors, including the fact that we estimated unreported cases as well, which led to more estimated infections in total and, with a waning delay, to more re-infections earlier on.

To our knowledge, this is the first and most comprehensive application of a computationally efficient inference-based modeling approach that provides estimates of infection and hospitalization burden by age, region, VoC, and immune status (compared to (13, 30, 52–57)). The estimation is conducted with daily data and applied over seven waves during the first two pandemic years in the Netherlands. While we focus on the Dutch case, our framework can be applied to other countries given the common surveillance and national statistics data typically available for most countries. Compared to other studies (24, 25, 27, 58), we separately model contacts in elementary schools, secondary schools, and the rest of the population to enable more reliable disentanglement of the roles of children, adolescents, and adults in transmission. We approximate the timing of contact changes due to non-pharmaceutical interventions outside the school environment based on changes in proxy mobility measures (59) — Google mobility at transit stations and train mobility: based on this proxy, most measures are taken up by the population ahead of official implementation. We estimate the speed of behavioral changes and the number of age-specific contacts after each intervention (13, 23, 60). This is in contrast to other inference-based studies (e.g. (28)) that assume that age-related mixing patterns remain constant over time. Retrospective analyses of the impact on transmission and burden in different age groups of vaccination and boosting and of non-pharmaceutical interventions targeted at elementary and secondary schoolchildren, or the rest of the population, will be a focus of future work.

Our model is fitted to regional reported cases, regional hospital admissions, and national seroprevalence. We find that the first two data sources, combined with modeling movement of unreported infected across regions using mobility data, are crucial for identifying unreported infections with our inference method. Unlike other studies (13, 23), the national seroprevalence data played a secondary role in reconstructing transmission dynamics due to the low frequency of serosurveys (i.e., only four surveys during 23 months in our case, which can only be implemented in the estimation as four data points for each age category) whose contribution to the likelihood is therefore minor (Supplementary Table A16). Our approach could therefore apply to countries where seroprevalence estimates are not available at a high enough frequency. Incorporation of other data into our model, particularly, publicly available SARS-CoV-2 measurements in sewage water, could help to address another current problem of lack of other

COVID-19 surveillance systems in many European countries including the Netherlands. This model extension could provide projections of transmission dynamics after large-scale PCR testing stopped. Our model could also be extended to quantify SARS-CoV-2 infection and COVID-19 hospitalization burdens in vulnerable populations with pre-existing chronic conditions (61) and to determine which interventions are required to protect them.

Our model has limitations. Firstly, rare testing in children and adolescents in the early stages of the pandemic makes the identification of their age-specific parameters (i.e., susceptibility) weaker in that initial period, and for some parameters such as mobility reporting errors, the simulated posterior distribution does not cover well an initial fixed value used in the data generating process in the first wave (Supplementary Figures A21—A23). Later on, as testing capacity was expanded to younger individuals, we observed better identification of all age-specific parameters. Secondly, we did not model older age groups separately, and this lack of heterogeneity may be the reason why our model underestimates the peak in hospitalizations in most waves. It is technically straightforward to stratify the adult population into smaller age groups relevant for estimating the burden in older ages. However, older individuals travel less which means that our metapopulation model with mobility may not capture their case detection rates, and other data might be needed to estimate their age-specific parameters such as differential case-detection rates. Thirdly, due to data available only as aggregate counts, the four serosurveys we used for fitting could only be implemented as four data points (at median time of sample collection minus 14 days to allow for seroconversion). We further assumed the data come from a Binomial distribution. If there was no uncertainty in the seroprevalence estimates from the four serosurveys, it may seem that we are underestimating seroprevalence in adults and adolescents, at least at the beginning of 2021. However, the real uncertainty in these data points estimates is not publicly available but seems to be large, varying between 2% and 10% depending on age group (Figure 4 in (46)), making it unclear whether our estimates are statistically distinguishable from the original survey estimates. Fourthly, the national recommendation was to get only one vaccine dose if individuals had an infection in the previous six months. If individuals followed this recommendation, they appear in the vaccination data along with those who received two doses as "fully vaccinated". Therefore, we could not model the effect of the first dose separately. Fifthly, our choice of stratification into primary versus breakthrough infections and re-infections is justified for the period when a large proportion of the population did not yet have any immunity to SARS-CoV-2. As SARS-CoV-2 transitions from pandemicity to endemicity, primary infections are experienced only by very young children born into the population (26, 62). For later periods, our model could be modified to differentiate between several immunity classes such as individuals with primary vaccination series and various boosters, prior infections, and hybrid immunity (63).

In conclusion, we developed an inference-based transmission model that estimates how the age-specific epidemiology of SARS-CoV-2 changes over time. This approach is relevant for countries in which random community surveys uncovering true SARS-CoV-2 dynamics are absent but basic surveillance and statistics data are available. The findings of our study on the burden of infections and hospitalizations in children, adolescents, and adults are important for informing public health policy on non-pharmaceutical interventions and vaccination.

Materials and Methods

Overview. The transmission model was calibrated using surveillance and national statistics data (PCR testing data, hospital admissions, serological surveys, demographic data, regional train and Google mobility data, vaccination coverage data, genetic VoC data, social contact matrices) for the Netherlands. Parameter estimates were obtained from the model fit to (i) age- and province-stratified SARS-CoV-2 case notification data during the period from February 27, 2020 until January 31, 2022; (ii) age- and province-stratified COVID-19 hospital admission data for the same period; and (iii) cross-sectional age-stratified national seroprevalence data from four serosurveys assessed on April 3, June 4 and September 20, 2020, and February 11, 2021 (median dates). Additional data for the model input were: (iv) population by age and province on January 1, 2020; (v) daily commuters across twelve provinces computed from Dutch national train data during the period from February 1, 2020 until September 30, 2022, and extrapolated using Google mobility data from February 5, 2020 until January 31, 2022; (vi) daily full vaccinations per province and age group during the period from January 31, 2021 until January 31, 2022; (vii) weekly boosters administered during the period from November 21, 2021 until January 31, 2022; (viii) weekly genetic VoC data during the period from December 1, 2020 until January 31, 2022; and (ix) school and non-school contact matrices from three surveys: before the pandemic, April 2020 and June 2020. All model analyses were performed in MATLAB 2022A. Data cleaning was performed in STATA SE v16. The code for fitting the model over the entire sample period runs in approximately 23 minutes on a Windows 10 Dell laptop with an Intel Core i5 processor and without parallelization.

Data. Table 1 gives an overview of the data, notation and sources. More details on how each dataset was constructed for use in the model fitting are given in the Supplementary Materials, Section 1.

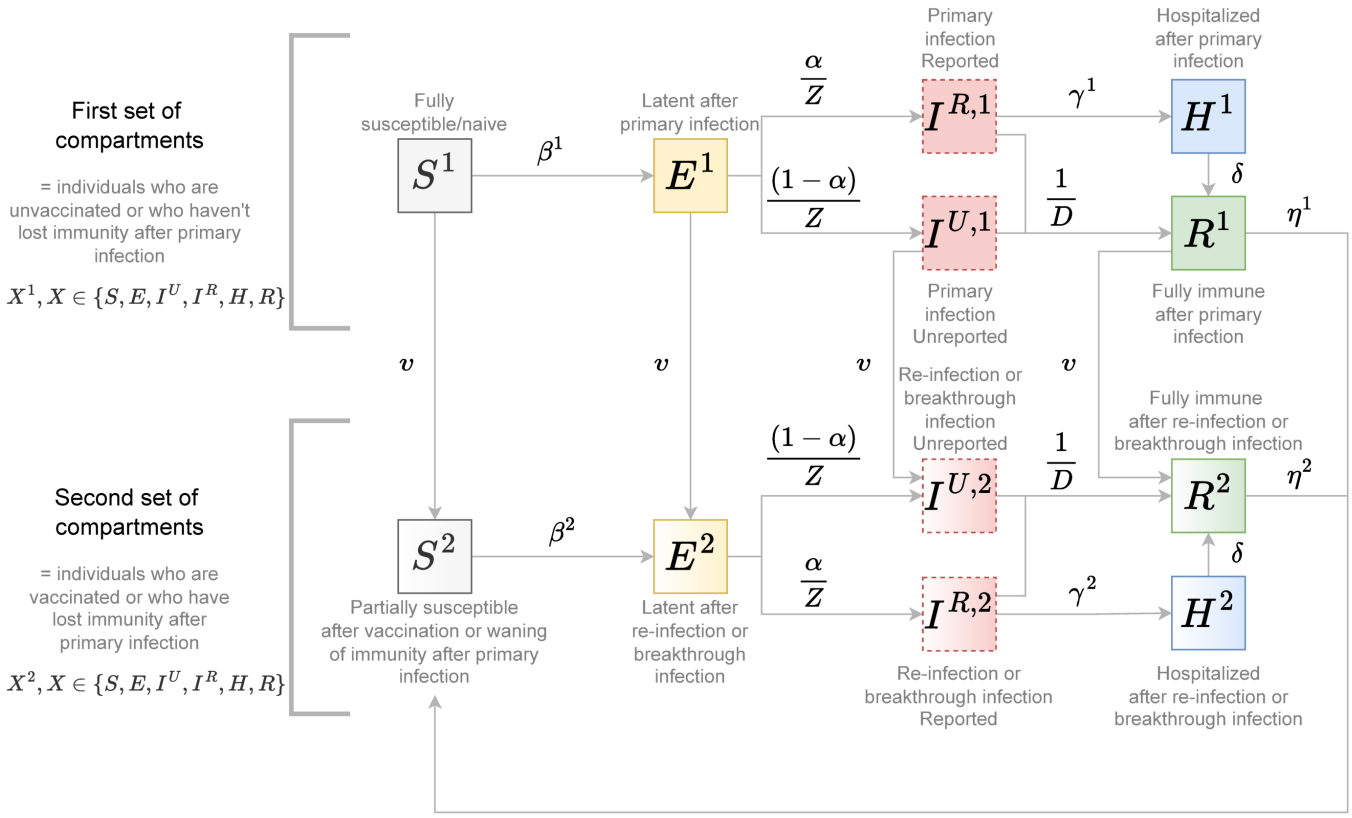
Transmission model. We developed a stochastic compartmental metapopulation model describing SARS-CoV-2 transmission in the population of the Netherlands stratified by province, disease status, and age. A schematic depicting disease and mobility dynamics, as well as an overview of the main model parameters are shown in Figure 5. The choice for the population stratification by age into children (0 to 9 years old), adolescents (10 to 19 years old), and adults (above 19 years old) was motivated by the surveillance data that was available for the model fitting. Contact matrices were aggregated to accommodate this stratification.

Disease dynamics. The fully naive susceptible individuals (S^1) in each age group can become latently infected but are not infectious (E^1) with the age-specific force of infection β^1. After an average latent period Z, the latently infected individuals become infectious (primary infection, I^1). Out of the total daily new infectious cases, an age-specific fraction α is reported to surveillance ($I^{R,1}$), and the $(1-\alpha)$ fraction is unreported ($I^{U,1}$). Both reported and unreported infectious individuals may recover without hospitalization (R^1) after an average infectious period D, but only reported infectious individuals may be hospitalized (H^1) with the age-specific rate γ^1. We assume that hospitalized persons have contacts with the personnel and visitors but are not infectious because of the use of individual protective measures. Hospitalized individuals are discharged after an average age-specific hospitalization period $1/\delta$. After a primary infection, individuals lose immunity and become partially susceptible (S^2) with the rate η^1. Disease progression for partially susceptible individuals is similar to that for fully susceptible individuals. However, partially susceptible individuals are reinfected with the age-specific force of infection β^2, individuals reported with re-infection are hospitalized with the age-specific rate γ^2, and the immunity after re-infection is lost with the rate η^2. Upon losing immunity after re-infection, the individuals return to the same partially susceptible compartment (S^2). Vaccination of individuals with the age-specific rate v occurs in all disease stages

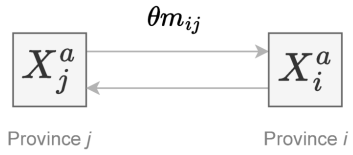
Table 1. Overview of the data used in the model fitting.

Data	Supplement	Source
Population by province-age	Section 1.1	CBS, January 1, 2020 https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=37259eng&_theme=1135
Daily reported cases	Section 1.2	RIVM Dashboard https://data.rivm.nl/covid-19/COVID-19_casus_landelijk.csv
Daily hospital admissions	Section 1.3	
Dataset 1		RIVM data, from February 27, 2020 until September 30, 2021
Dataset 2		RIVM Dashboard, from October 1, 2021 onward https://data.rivm.nl/covid-19/COVID-19_ziekenhuisopnames.csv https://data.rivm.nl/covid-19/COVID-19_ziekenhuisopnames_tm_03102021.csv
Dataset 3		RIVM Dashboard, from October 1, 2021 onward https://data.rivm.nl/covid-19/COVID-19_ziekenhuis_ic_opnames_per_leeftijdsgroep.csv https://data.rivm.nl/covid-19/COVID-19_ziekenhuis_ic_opnames_per_leeftijdsgroep_tm_03102021.csv
Seroprevalence data 4 Serosurvey rounds Median inclusion time Fraction seropositive Sample size	Section 1.4	obtained from RIVM PIENTER Corona Study (13, 64) https://www.rivm.nl/en/pienter-corona-study/results
Train mobility	Section 1.5	computed from the Dutch Railways (NS) data, from February 27, 2020 until September 30, 2021; extrapolated until January 31, 2022 using Google mobility data
Fraction commuters		CBS, 2019 https://www.cbs.nl/nl-nl/cijfers/detail/83628NED
Google mobility		from February 7, 2020 onward https://www.google.com/covid19/mobility/
Daily vaccinations	Section 1.6	constructed from RIVM Dashboard and RIVM data https://data.rivm.nl/covid-19/COVID-19_vaccinatiegraad_per_gemeente_per_week_leeftijd.csv https://www.rivm.nl/en/covid-19-vaccination/archive-covid-19-vaccination-figures-2021
Booster transition function	Section 1.7	fitted to RIVM data https://www.rivm.nl/en/covid-19-vaccination/archive-covid-19-vaccination-figures-2022
Variants transition function	Section 1.8	fitted to RIVM Dashboard data https://data.rivm.nl/covid-19/COVID-19_varianten.csv
Contact matrices Elementary school contacts Secondary school contacts All-setting contacts Non-school contacts	Section 1.9	computed from (41, 65) and additional information on school closures and vacations computed from (65) computed from (41, 65)

A Disease dynamics



B Mobility dynamics



$$X \in \{S, E, I^U, R\},\ a \in \{1, 2\},\ j \neq i$$

C Overview of the parameters

Age-specific parameters

β^1	- force of infection for fully susceptible
β^2	- force of infection for partially susceptible
α	- case detection rate
γ^1	- hospitalization rate after primary infection
γ^2	- hospitalization rate after re-infection or breakthrough infection
v	- vaccination rate
$1/\delta$	- hospitalization period
θ	- mobility reporting error
m_{ij}	- mobility between province *j* and province *i*

Constant parameters

Z	- latent period
D	- infectious period
$1/\eta^1$	- duration of immunity after primary infection
$1/\eta^2$	- duration of immunity after re-infection or breakthrough infection

Fig. 5. Schematic of the metapopulation transmission model. (**A**) Disease dynamics. (**B**) Mobility dynamics. (**C**) Overview of the main model parameters (see Supplementary Materials for the full description). For simplicity of presentation, the age- and region-specific indices are not shown in the schematic.

except for those with primary reported infection or hospitalization. We used a simplified approach whereby susceptible individuals after vaccination and individuals recovered after primary infection whose immunity has waned are grouped together in one class of partially susceptible individuals (S^2). Therefore, re-infections and breakthrough infections are grouped, too.

Vaccination or prior infection has three effects: (i) lower susceptibility to re-infection and breakthrough infection that affects β^2, (ii) lower infectivity of re-infection and breakthrough infection (not shown in the diagram), and (iii) lower hospitalization rate after re-infection and breakthrough infection γ^2. The model assumes a constant population size during the study period.

Our choice for the stratification of the infectious compartments into reported and unreported infectious is driven by the data used for the model fitting. We use daily data on reported positive PCR tests (Table 1), and most of the reported cases are by the self-reported first date of symptoms. Therefore, we assume that the reported infectious compartments contain mostly symptomatic infections, and the unreported infectious compartments contain mostly sub-clinical asymptomatic or mildly symptomatic infections. Based on this fact, coupled with the evidence of lower infectivity of asymptomatic infections relative to symptomatic infections (66), we assume that the relative infectivity of the unreported infectious compartments compared to the reported infectious compartments is $0 < \mu \leq 1$ (Figure 5).

Mobility dynamics. The model assumes the unreported infected travel and infect other individuals in other regions, and that susceptible individuals can travel and become exposed in another region, while reported infected do not travel. Due to heterogeneity in population size and mobility across regions, this assumption equips the model with additional dynamics for the unreported infected compared to reported infected. It was shown by simulations (24, 67) and proven mathematically (45)) that the identification of unreported cases is possible as long as the case detection rate is the same across regions, and there are sufficient observations available across regions or time. These two conditions are met in our study. Firstly, the case detection rate is assumed common across regions due to the ability to test at any location available across the country. Secondly, we estimate the model at a daily frequency for all twelve Dutch provinces.

The number of people commuting from province j to province i is constructed as described in Supplementary Materials, Section 1.6. It involves multiplying the total time-varying number of commuters into i (the Dutch Railways data based on outward movement) by the pre-pandemic fixed percentage of commuters from j into i (CBS — Statistics Netherlands — data on all employees in 2018). The CBS data indicate that about 4% of all employees were commuting for work across municipalities. Out of these, the percentage of individuals commuting from each province j to another province i is shown in Figure fig6, along with the size of each province, as a percentage of the population. As expected, most people commute within provinces, and more people commute to larger provinces than to smaller ones. There is a sizable and heterogeneous percentage of commuters from most provinces to other provinces. Supplementary Figure S10 also indicates that the percentage of inward commuters is higher from neighboring regions. All these suggest that not accounting for mobility across regions is likely to underestimate the speed at which SARS-CoV-2 spreads in the population.

Some studies find that mobility data such as SafeGraph for the United States cannot alone predict the evolution of cases later in the pandemic, due to erratic policies and behavior that differed across US counties (68). However, this criticism does not apply to our study. First, in the Netherlands, all measures were national, and we modeled the non-pharmaceutical interventions directly, including the speed at which they were adopted in the population. Second, the Netherlands is a small and well-connected country, with a large share of the population commuting for work outside their municipality. Third, we do not use only mobility data to infer cases, we use them in conjunction with an elaborate model that is fitted to regional case data, hospitalizations, and seroprevalence surveys.

Force of infection. The force of infection, $\beta^a_{ik}(t)$, is time-dependent age- and province-specific and depends on whether individuals are fully susceptible or partially susceptible. The subscript i denotes the province of the Netherlands ($i = 1, \ldots, 12$). The subscript k denotes the age group ($k = 1, 2, 3$), namely $k = 1$ — adults (> 19 years old), $k = 2$ — adolescents ($10 - 19$ years old), and $k = 3$ — children ($0 - 9$ years old). The susceptibility class is denoted by the superscript a, so that $a = 1$ (fully susceptible/naive individuals), and $a = 2$ (individuals partially susceptible after vaccination or waning of immunity after primary infection).

The force of infection is a multiplicative function

$$\beta^a_{ik}(t) = \epsilon \times f_{\epsilon,k} \times voc(t) \times \lambda^a_{ik}(t), \qquad [1]$$

where ϵ is the probability of transmission per contact for the wild-type variant, $f_{\epsilon,k}$ is the susceptibility of age group $k = 2, 3$ relative to the reference age group $k = 1$ (i.e., $f_{\epsilon,1} = 1$), $voc(t)$ is the increase in the probability of transmission per contact due to VoCs, and $\lambda^a_{ik}(t)$ is the time-dependent average number of daily close transmission-relevant contacts in all settings for one individual in age group k and province i with all individuals in other age groups multiplied by the proportion of infectious individuals in those age groups, weighted by their infectivity.

For the fully susceptible individuals,

$$\lambda^1_{ik}(t) = \sum_{k^*=1}^{3} c_{i,kk^*}(t) \sum_{j=1}^{12} \left[\frac{I^{R,1}_{ijk^*}(t) + [1 - p_{TR,k^*}(t)] I^{R,2}_{ijk^*}(t)}{N_{ik^*}} + \mu \frac{I^{U,1}_{ijk^*}(t) + [1 - p_{TU,k^*}(t)] I^{U,2}_{ijk^*}(t)}{N_{ik^*}} \right], \qquad [2]$$

where $k, k^* = 1, 2, 3$ and $j = 1, \ldots, 12$. Here $c_{i,kk^*}(t)$ is the average number of daily close transmission-relevant contacts in all settings one individual in age group k has with all individuals in age group k^* in province i at time t. $I^{b,a}_{ijk}(t)$ is the number of infectious individuals in age group k from province j present in province i at time t who have $a = 1$ (primary infection), 2 (re-infection or breakthrough infection) and are $b = R$ (reported), U (unreported). N_{ik} is the total population in age group k present in province i, including those individuals who moved to province i from all other provinces. $0 < \mu \leq 1$ is the relative infectivity of the unreported infectious individuals compared to the reported infectious compartments. $p_{Tb,k}$ is the reduction in infectivity of re-infection and breakthrough infection relative to primary infection in age group k for $b = R$ (reported), and U (unreported) infectious individuals.

For the partially susceptible individuals,

$$\lambda^2_{ik}(t) = [1 - p_{I,k}(t)]\, \lambda^1_{ik}(t), \qquad [3]$$

where $p_{I,k}(t)$ is the reduction in susceptibility to re-infection or breakthrough infection relative to primary infection in age group k.

VoC. The model did not account for VoCs separately using, e.g., a multi-strain approach. Instead, we approximated the spread of a new VoC by increasing the probability of transmission per contact for the wild-type variant by a factor $voc(t)$ as the proportion of the new VoC increased from 0% to 100% using a logistic transition function. We further assumed that infection with one VoC gives the same protection against all other VoCs but Omicron. This assumption is based on the strong evidence that Omicron BA.1 was not only more transmissible but also provided a partial immune escape from the previous VoCs (see, e.g., (69) and references therein). We modeled the immune evasion by Omicron by increasing the immunity waning rates η^1 and η^2 as the proportion of Omicron in the population increased (see Supplementary Materials, Section 2.3). In a model without immunity waning and constant parameters, not accounting for a two-strain model but increasing the probability of transmission per contact using a logistic transition function amounts to assuming that the susceptible population is constant during the transition from one strain to the next (70, 71). Therefore, our estimates of the susceptible population during those transitions may be biased, however, as (70) illustrates, the differences to a multi-strain model tend to be minor.

A full description of the model and the model equations are reported in Supplementary Materials, Section 2.

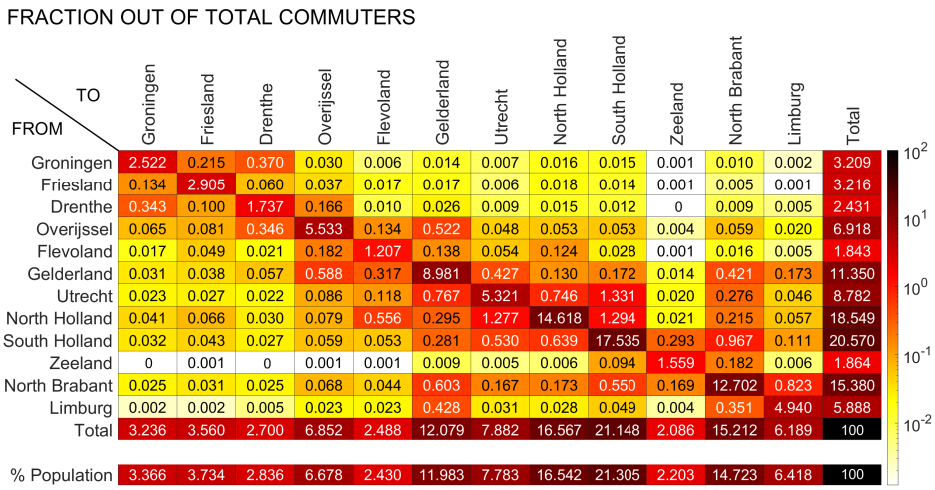


Fig. 6. Commuting data used in the model. Shown is the percentage out of total commuters commuting within and across provinces before the pandemic by the Statistics Netherlands (CBS).

Parameter inference. The model is fitted to infection, hospitalization, and seroprevalence data using the ensemble adjustment Kalman filter (38). This method allows inference on a high-dimensional system of observed and unobserved variables and parameters in a computationally efficient way. First, a large number of ensembles are drawn from priors on all parameters and state variables. The latter then evolve according to a stochastic version of the metapopulation model (Figure 5), and the stochasticity is needed because the populations per age-region can be small (the smallest being 37,264 children in the province Groningen). For the filter updates, each ensemble member is individually propagated forward by means of a closed form approximation akin to the normal distribution but with an additional bias correction. The updates are computed sequentially based on the data in each province and its neighbours, allowing for a large number of state variables and parameters to be updated. Further details on the estimation and the algorithm can be found in Sections 3.3 and 3.4 of the Supplementary Materials.

A small number of parameters that could not be identified were calibrated. These calibrations are described and motivated in Supplementary Materials, Section 3.1. The remaining parameters were estimated, and the choice of priors is detailed and motivated in the Supplementary Materials, Section 3.2. The parameter posteriors and their time evolution over variant periods are discussed in Section 4 of the Supplementary Materials.

Model outcomes. The exact calculations behind Figures 1—4 are explained in Supplementary Materials, Sections 5.1—5.4. The data plotted in Figures 3 and 4 ad further numbers reported in the text when discussing these Figures can be found in Supplementary Tables A9—A13.

Parameter identifiability and sensitivity analyses. To verify system identifiability for the first wave, we fixed the parameters at values similar to their estimated posteriors in the original sample (observations from February 27, 2020 until March 30, 2020, see Supplementary Table A14), generated one synthetic outbreak and re-estimated the model parameters on the synthetic data. In addition to showing that the synthetic data are similar to the original data and are fitted well by our method, we also verified that the fitted and model implied seroprevalence match, by not using the model implied seroprevalence in the fitting procedure (Supplementary Materials, Section 6.1, Table A15 and Figures A17—A20). To assess parameter identifiability for the first wave, we considered three parameter configurations: (i) as described above; (ii) larger case detection rates; and (iii) case detection rates as in (i) until March 30, 2020, and as in (ii) until April 30, 2020. For each parameter combination, 100 synthetic outbreaks were generated, each of which was used as data to re-estimate the model parameters. Across the three model configurations, the "true" parameters were either within the range of the posterior mean densities or the differences were not too large (Supplementary Materials, Section 6.2, Figures A21—A23), although some parameters such as the mobility reporting error rate were less well identified. In case (iii), the filter also approximated well the increase of case detection rates, even though this increase was not modelled. This suggests that changes in testing capacity or recommendations, and other unmodelled parameter changes can be captured by our model inference technique even in periods of high uncertainty, when data are less informative (Supplementary Figure A24).

For our sensitivity analyses, we re-estimated the model over the entire sample period: (i) with seasonality included in transmission; (ii) without seroprevalence of including it at different median date; (iii) without mobility; (iv) with lower and higher immunity waning rates after primary infection; (iii) with lower and higher immunity waning rates at the start of Omicron BA.1 period after both primary and breakthrough infection/re-infection; (iv) with higher vaccine efficacies against infection, upon their waning. In all these cases, the likelihood fit is poorer (Supplementary Table A16).

Data availability. All output datasets generated during this study are publicly available in the designated repository (at https://dataverse.nl/dataset.xhtml?persistentId=doi:10.34894/J5UHNL or, if no institutional login available, also at https://drive.google.com/file/d/1n1bR2jaKQD3zEcKLQWiOeTlPra138mG_/view?usp=sharing). The datasets used as input in this study are listed in Table 1. All input datasets except for train mobility data are publicly available designated GitHub repository (https://dataverse.nl/dataset.xhtml?persistentId=doi:10.34894/J5UHNL or https://drive.google.com/file/d/1n1bR2jaKQD3zEcKLQWiOeTlPra138mG_/view?usp=sharing). Requests for train mobility data should be addressed to Jan Banninga (jan.banninga@ns.nl) from the Dutch Railways (Nederlandse Spoorwegen — NS). The train mobility dataset is available to reviewers and editors upon request.

Code availability. All codes reproducing the results of this study are available at the designated GitHub repository (https://dataverse.nl/dataset.xhtml?persistentId=doi:10.34894/J5UHNL or https://drive.google.com/file/d/1n1bR2jaKQD3zEcKLQWiOeTlPra138mG_/view?usp=sharing).

Additional information. Supplementary Materials contain details of

the data, the model, inference, system identifiability, sensitivity analyses, Supplementary Figures, and Tables.

ACKNOWLEDGMENTS. G.R. was supported by the VERDI project (101045989), funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Health and Digital Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. G.R. and S.P. were supported by the Fundação para a Ciência e a Tecnologia project 2022.01448.PTDC. O.B. and A.A. were supported by the Tilburg University Fund. J.S. and S.P. were supported by US NIAID grant R01AI163023 and CDC contract 75D30122C14289. We thank Prof. Marc Bonten (University Medical Center Utrecht), Dr. Susan van den Hof (National Institute for Public Health and the Environment), Prof. Dennis Huisman (Erasmus University Rotterdam), and Jan Banninga (Nederlandse Spoorwegen) for their substantial help in obtaining the data for this study. We also thank Dr. Bettina Siflinger (Tilburg University) for the Tilburg CentER LISS Panel Survey questions added to support our study, Dr. Marino van Zelst (Wageningen University), and Prof. Jaap Abbring (Tilburg University) for useful discussions on data availability and modeling.

1. World Health Organization. Coronavirus (COVID-19) Dashboard (2022) Accessed July 25, 2022.
2. Variation in the COVID-19 infection-fatality ratio by age, time, and geography during the pre-vaccine era: a systematic analysis. *The Lancet* **399**, 1469–1488 (2022).
3. M O'Driscoll, et al., Age-specific mortality and immunity patterns of sars-cov-2. *Nature* **590**, 140–145 (2021).
4. Coronavirus dashboard, (2020) Accessed July 25, 2022.
5. K Koelle, MA Martin, R Antia, B Lopman, NE Dean, The changing epidemiology of SARS-CoV-2. *Science* **375**, 1116–1121 (2022).
6. M Monod, et al., Age groups that sustain resurging COVID-19 epidemics in the united states. *Science* **371**, eabe8372 (2021).
7. NG Davies, et al., Age-dependent effects in the transmission and control of covid-19 epidemics. *Nat. Medicine* **26**, 1205–1211 (2020).
8. N Perra, Non-pharmaceutical interventions during the COVID-19 pandemic: A review. *Phys. Reports* **913**, 1–52 (2021) Non-pharmaceutical interventions during the COVID-19 pandemic: a review.
9. Y Liu, et al., The impact of non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories. *BMC medicine* **19**, 40–40 (2021) 33541353[pmid].
10. M Sharma, et al., Understanding the effectiveness of government interventions against the resurgence of COVID-19 in europe. *Nat. communications* **12**, 5820–5820 (2021) 34611158[pmid].
11. JM Brauner, et al., Inferring the effectiveness of government interventions against covid-19. *Science* **371**, eabd9338 (2021).
12. Y Li, et al., The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (r) of SARS-CoV-2: a modelling study across 131 countries. *The Lancet Infect. Dis.* **21**, 193–202 (2021) 33729915[pmid].
13. G Rozhnova, et al., Model-based evaluation of school and non-school-related measures to control the COVID-19 pandemic. *Nat. Commun.* **12**, 1614 (2021).
14. M Flook, et al., Informing the public health response to COVID-19: a systematic review of risk factors for disease, severity, and mortality. *BMC infectious diseases* **21**, 342–342 (2021) 33845766[pmid].
15. KM Bubar, et al., Model-informed covid-19 vaccine prioritization strategies by age and serostatus. *Science* **371**, 916–921 (2021).
16. L Matrajt, J Eaton, T Leung, ER Brown, Vaccine optimization for COVID-19: Who to vaccinate first? *Sci. Adv.* **7**, eabf1374 (2021).
17. P Sah, et al., Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis. *Proc. Natl. Acad. Sci.* **118**, e2109229118 (2021).
18. Office for National Statistics. Coronavirus (COVID-19) Infection Survey, UK (2022) Accessed 8 December, 2022.
19. K Shioda, et al., Estimating the cumulative incidence of SARS-CoV-2 infection and the infection fatality ratio in light of waning antibodies. *Epidemiology* **32** (2021).
20. N Hozé, et al., Monitoring the proportion of the population infected by SARS-CoV-2 using age-stratified hospitalisation and serological data: a modelling study. *The Lancet Public Heal.* **6**, e408–e415 (2021).
21. G Milne, et al., Does infection with or vaccination against SARS-CoV-2 lead to lasting immunity? *The Lancet Respir. Medicine* **9**, 1450–1466 (2021).
22. World Health Organization. Tracking SARS-CoV-2 variants (2022) Accessed July 25, 2022.
23. J Viana, et al., Controlling the pandemic during the SARS-CoV-2 vaccination rollout. *Nat. Commun.* **12**, 3674 (2021).
24. S Pei, TK Yamana, S Kandula, M Galanti, J Shaman, Burden and characteristics of COVID-19 in the United States during 2020. *Nature* **598**, 338–341 (2021).
25. CM Saad-Roy, et al., Immune life history, vaccination, and the dynamics of SARS-CoV-2 over the next 5 years. *Science* **370**, 811–818 (2020).
26. JS Lavine, ON Bjornstad, R Antia, Immunological characteristics govern the transition of COVID-19 to endemicity. *Science* **371**, 741–745 (2021).
27. SM Kissler, C Tedijanto, E Goldstein, YH Grad, M Lipsitch, Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**, 860–868 (2020).
28. R Sonabend, et al., Non-pharmaceutical interventions, vaccination, and the SARS-CoV-2 delta variant in England: a mathematical modelling study. *The Lancet* **398**, 1825–1835 (2021).
29. PT de Boer, et al., Age-specific severity of severe acute respiratory syndrome coronavirus 2 in february 2020 to june 2021 in the netherlands. *Influ. Other Respir. Viruses* **17**, e13174 (2023).
30. MM Dekker, LE Coffeng, FP Pijpers, D Panja, SJ de Vlas, Reducing societal impacts of sars-cov-2 interventions through subnational implementation. *medRxiv* (2022).
31. M Gatto, et al., Spread and dynamics of the COVID-19 epidemic in italy: Effects of emergency containment measures. *Proc. Natl. Acad. Sci.* **117**, 10484–10491 (2020).
32. E Bertuzzo, et al., The geography of covid-19 spread in italy and implications for the relaxation of confinement measures. *Nat. Commun.* **11**, 4264 (2020).
33. G Giordano, et al., Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nat. Med* **26**, 855–860 (2020).
34. P Bosetti, et al., Epidemiology and control of sars-cov-2 epidemics in partially vaccinated populations: a modeling study applied to france. *BMC Medicine* **20**, 33 (2022).
35. AM Presanis, D De Angelis, A Goubar, ON Gill, AE Ades, Bayesian evidence synthesis for a transmission dynamic model for HIV among men who have sex with men. *Biostatistics* **12**, 666–681 (2011).
36. PJ Birrell, et al., Bayesian modeling to unmask and predict influenza a/h1n1pdm dynamics in london. *Proc. Natl. Acad. Sci.* **108**, 18238–18243 (2011).
37. G Rozhnova, et al., Short- and long-term impact of vaccination against cytomegalovirus: a modeling study. *BMC Medicine* **18** (2020).
38. JL Anderson, An ensemble adjustment kalman filter for data assimilation. *Mon. Weather. Rev.* **129**, 2884 – 2903 (2001).
39. S Pei, S Kandula, J Shaman, Differential effects of intervention timing on COVID-19 spread in the United States. *Sci. Adv.* **6**, eabd6370 (2020).
40. J Mossong, et al., Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLOS Medicine* **5**, 1–1 (2008).
41. D Mistry, et al., Inferring high-resolution human mixing patterns for disease modeling. *Nat. Commun.* **12**, 323–323 (2021) 33436609[pmid].
42. E Goldstein, M Lipsitch, M Cevik, On the effect of age on the transmission of SARS-CoV-2 in households, schools and the community. *The J. Infect. Dis.* (2020) jiaa691.
43. RM Viner, et al., Susceptibility to SARS-CoV-2 Infection Among Children and Adolescents Compared With Adults: A Systematic Review and Meta-analysis. *JAMA Pediatr.* **175**, 143–156 (2021).
44. R Li, et al., Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).
45. A Hortaçsu, J Liu, T Schwieg, Estimating the fraction of unreported infections in epidemics with a known epicenter: An application to COVID-19. *J. Econom.* **220**, 106–129 (2021).
46. The National Institute for Public Health and the Environment (RIVM), Pienter Corona Study (2022) Accessed October 30, 2023.
47. Q Bi, et al., Insights into household transmission of sars-cov-2 from a population-based serological survey. *Nat. communications* **12**, 3643 (2021).
48. Q Jing, et al., Household secondary attack rate of COVID-19 and associated determinants in Guangzhou, China: a retrospective cohort study. *Lancet Infect Dis* **20**, 1141–1150 (2020).
49. K Mizumoto, R Omori, H Nishiura, Age specificity of cases and attack rate of novel coronavirus disease (COVID-19). *MedRxiv* pp. 2020–03 (2020).
50. J Zhang, et al., Changes in contact patterns shape the dynamics of the COVID-19 outbreak in china. *Science* **368**, 1481–1486 (2020).
51. MJ Keeling, Patterns of reported infection and reinfection of SARS-CoV-2 in England. *J. Theor. Biol.* **556**, 111299 (2023).
52. F Miura, KY Leung, D Klinkenberg, KEC Ainslie, J Wallinga, Optimal vaccine allocation for COVID-19 in the netherlands: A data-driven prioritization. *PLOS Comput. Biol.* **17**, 1–13 (2021).
53. KEC Ainslie, et al., The impact of vaccinating adolescents and children on COVID-19 disease outcomes. *medRxiv* (2021).
54. M van Boven, et al., Modelling patterns of SARS-CoV-2 circulation in the Netherlands, August 2020-February 2022, revealed by a nationwide sewage surveillance program. *medRxiv* (2022).
55. MHH Schoot Uiterkamp, M Gösgens, H Heesterbeek, R van der Hofstad, N Litvak, The role of inter-regional mobility in forecasting SARS-CoV-2 transmission. *J. The Royal Soc. Interface* **19**, 20220486 (2022).
56. M Gösgens, et al., Trade-offs between mobility restrictions and transmission of SARS-CoV-2. *J. The Royal Soc. Interface* **18**, 20200936 (2021).
57. SJ de Vlas, LE Coffeng, Achieving herd immunity against COVID-19 at the country level by the exit strategy of a phased lift of control. *Sci. Reports* **11**, 4445 (2021).
58. SPC Brand, et al., COVID-19 transmission dynamics underlying epidemic waves in Kenya. *Science* **374**, 989–994 (2021).
59. N Kishore, et al., Evaluating the reliability of mobility metrics from aggregated mobile phone data as proxies for SARS-CoV-2 transmission in the USA: a population-based study. *The Lancet Digit. Heal.* **4**, e27–e36 (2022).
60. O Boldea, A Cornea-Madeira, J Madeira, Disentangling the effect of measures, variants and vaccines on sars-cov-2 infections in england: A dynamic intensity model. *The Econom. J.* **26**, 444–466 (2023) https://doi.org/10.1093/ectj/utad004.
61. European Centre for Disease Prevention and Control. Core protocol for ECDC studies of COVID-19 vaccine effectiveness against hospitalisation with Severe Acute Respiratory Infection laboratory confirmed with SARS-CoV-2, version 1.0. (2021) Accessed December 8, 2022.
62. LE Cohen, DJ Spiro, C Viboud, Projecting the sars-cov-2 transition from pandemicity to endemicity: Epidemiological and immunological considerations. *PLOS Pathog.* **18**, 1–21 (2022).
63. N Bobrovitz, et al., Protective effectiveness of prior sars-cov-2 infection and hybrid immunity against omicron infection and severe disease: a systematic review and meta-regression. *medRxiv* (2022).
64. ERA Vos, et al., Nationwide seroprevalence of SARS-CoV-2 and identification of risk factors in the general population of the Netherlands during the first epidemic wave. *J. Epidemiol. &*

Community Heal. (2020).
65. JA Backer, et al., Impact of physical distancing measures against COVID-19 on contacts and mixing patterns: repeated cross-sectional surveys, the Netherlands, 2016-17, April 2020 and June 2020. *Eurosurveillance* **26** (2021).
66. D Buitrago-Garcia, et al., Occurrence and transmission potential of asymptomatic and presymptomatic sars-cov-2 infections: Update of a living systematic review and meta-analysis. *PLOS Medicine* **19**, 1–30 (2022).
67. Q Li, et al., Early transmission dynamics in wuhan, china, of novel coronavirus-infected pneumonia. *New Engl. J. Medicine* (2020).
68. HS Badr, LM Gardner, Limitations of using mobile phone data to model covid-19 transmission in the usa. *The Lancet Infect. Dis.* **21**, e113 (2021).
69. M Hoffmann, P Arora, S Pöhlmann, Understanding omicron: Transmissibility, immune evasion and antiviral intervention. *Clin. Transl. Medicine* **12**, e839 (2022).
70. T Gö, W Bock, R Rockenfeller, M Schäfer, A two-strain SARS-COV-2 model for Germany - evidence from a linearization. (2021).
71. P Hansen, Relative contagiousness of emerging virus variants: an analysis of the Alpha, Delta, and Omicron SARS-CoV-2 variants. *The Econom. J.* **25**, 739–761 (2022).