# Estimating the impact of school closures on COVID-19 disease burden

Supplementary Appendix

——— Add Authors List here ———

## Contents

# 1 Model description

## 1.1 General approach

We use a semi-mechanistic compartmental model of COVID-19 transmission governed by ordinary differential equations (ODEs) to simulate country-specific COVID-19 epidemics during the first three years of the pandemic (2020-2022). Our model captures important factors of COVID-19 transmission and disease such as age-specific characteristics, heterogeneous mixing, vaccination and the emergence of different variants of concern. The ODE-based model is used to capture only states relevant to transmission, whereas hospitalisations and deaths are estimated through a convolution process applied to the ODE-based model's outputs. This process combines the model-estimated disease incidence with statistical distributions modelling the time to hospitalisation, the hospital stay duration and the time to death. This approach presents two main advantages. First it reduces the complexity of the dynamic system relying on numerical solving of ODEs, which is computationally expensive. Second, the convolution approach allows for more flexibility and produces more realistic assumptions regarding the timings of hospitalisations and deaths, compared to what could be achieved with a simple compartmental approach. The following sections describe the model in details and Figure S1 summarises the overall approach used in our analysis.

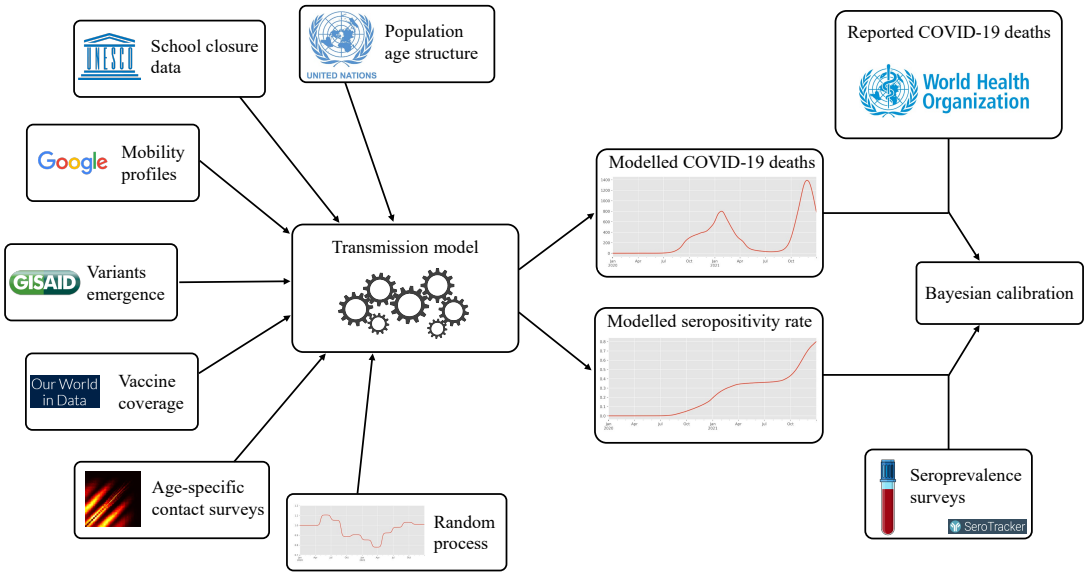


Figure S1: Conceptual approach used to build and calibrate the models

## 1.2 Transmission model

### 1.2.1 Compartment types and sequence

Model compartments represent sequential progressions through the processes of infection with, progression through, and recovery from the phases of SARS-CoV-2 infection and COVID-19 disease. The following types of compartments are implemented:

- Susceptible
  - Persons not previously infected with SARS-CoV-2 during or before the model simulation period
- Latent
  - Persons recently infected with SARS-CoV-2, but not in the active phase of the disease yet.
  - These individuals may still be infectious (see details in next paragraph).
- Active

  - Persons with active COVID-19 who are currently infectious.

- Recovered
  - Persons recovered from COVID-19 during the model simulation period
  - Reinfection from these compartments is permitted through exposure to a different strain than the one that most recently infected the individual (see strain stratification section for details).

The base model structure consists of a sequence of one susceptible compartment ($S$), four latent compartments ($E_1$, ..., $E_4$), four active disease compartments ($I^1$, ..., $I^4$) and one recovered compartment ($R$) (Figure S2).

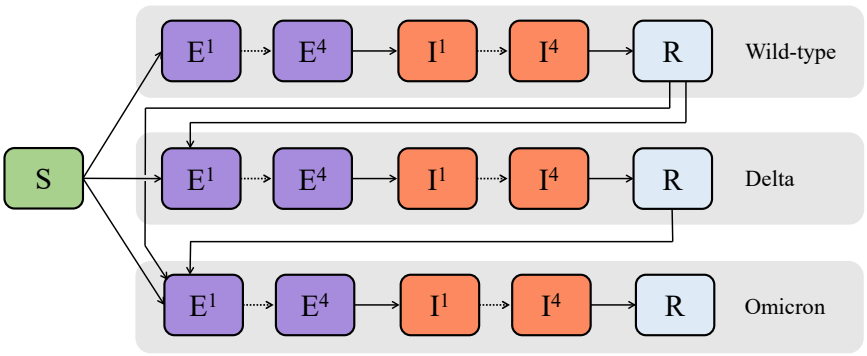


Figure S2: Compartmental model structure. S = Susceptible, E = Exposed / Latent, I = Active disease, R = Recovered. Stratification by age and vaccination status are not shown here.

The main rationale for using four serial compartments for both the latent and active states is to achieve an Erlang distribution for the time spent in each of these states. This distribution is more realistic than the exponential distribution which is the consequence of a single compartment assumption, because the Erlang distribution does not have a large density mass around 0 and is not heavy-tailed. Figure S3 illustrates the modelled distributions of the incubation period and the active disease period.

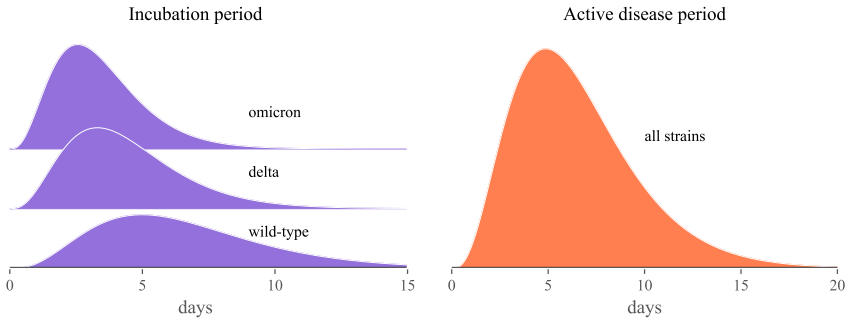


Figure S3: Modelled distributions of the incubation and active disease periods. Active disease period shown for an average duration of 8 days, but this parameter is varied during calibration.

The four active disease compartments all have identical characteristics. However, the last two latent compartments ($E^3$ and $E^4$) are infectious whereas the first two ($E^1$ and $E^2$) are not. We further assume that the infectious latent compartments are half as infectious as the active disease compartments.

### 1.2.2 Model stratification by age

All compartments of the base compartmental structure were stratified by age into the following age bands: zero to 14 years / 15 to 24 years / 25 to 49 years / 50 to 69 years / 70 years and above.

The initial population was distributed between the different age bands to reflect the age distributions reported by the United Nations Population Division. Demographic processes, including births, ageing and non-infection-related deaths are not simulated, given the timeframes considered in this simulation.

We assumed heterogeneous mixing between the different age groups to account for the assortative nature of social interactions by age (see Section 1.2.5). Age was assumed to affect:

- The susceptibility to infection
- The risk of COVID-19 hospitalisation
- The risk of COVID-19 death
- The vaccination rate (see Section 1.2.3)

### 1.2.3 Capturing the effects of vaccination

History of vaccination was captured by stratifying all model compartments by vaccination status. Two vaccination strata were included to represent those who have received at least two doses of a COVID-19 vaccine, and those who have not.

We used data from *Our World in Data* to inform the modelled dynamic vaccination coverage. In particular, we used the reported proportion of people "fully vaccinated" to specify the time-variant proportion of vaccinated people in our model. We assumed that older individuals are vaccinated first by prioritising the modelled age groups in descending order. That is, the oldest age group receives all available vaccines until a saturation coverage of 80% is reached for this group. Then the next oldest category starts receiving vaccines and we repeat this process until all available vaccines are allocated. Note that in the event that the population-level vaccine coverage exceeds 80%, the saturation coverage was set equal to the population-level coverage.

Let us consider two successive time points $t_i$ and $t_{i+1}$ for which vaccination data are available. Let us denote $r_{a,i}$ and $r_{a,i+1}$ the associated vaccine coverage for age group $a$. The time-variant and age-specific vaccination rate per capita $\omega_a(t)$ verifies:

$$1 - r_{a,i+1} = (1 - r_{a,i})e^{-w_a(t)(t_{i+1}-t_i)} \quad ,\forall t \in [t_i, t_{i+1}). \tag{1}$$

Then,

$$\omega_a(t) = \frac{\ln(1 - r_{a,i}) - \ln(1 - r_{a,i+1})}{t_{i+1} - t_i} \quad ,\forall t \in [t_i, t_{i+1}), \tag{2}$$

where $\ln(x)$ represents the natural logarithm of $x$.

Figure S4 shows the modelled vaccination coverage over time against the reported data for the analysed countries.

The effect of vaccination on transmission is to reduce the rate of infection partially for all persons at-risk of infection in the vaccinated stratum. This includes both fully susceptible (never previously infected) persons, as well as recovered persons who are at risk of reinfection. The model allows for hybrid immunity in the sense that the vaccination-induced relative reduction of transmission risk is multiplied with that induced by previous infection. Vaccination is also assumed to reduce the risk of hospitalisation and death

Emerging variants of concern (VoCs) may partially escape vaccine-induced (as well as infection-induced) immunity, as described further below (Table S2 and Section 1.2.4).

### 1.2.4 Modelling multiple viral strains

The model was stratified by "strain" to simulate the emergence of multiple variants of concern (VoC). This approach explicitly represents multiple competing strains, each with an independent force of infection calculation. We assumed that VoCs can have different levels of transmissibility, incubation period and disease severity (hospitalisation and death risks) compared to the ancestral COVID-19 strain. In addition, VoCs were assumed to escape immunity partially for both vaccination- and infection-related immunity.

We assumed that individuals previously infected with the wild-type strain could only be reinfected with the delta or omicron strains. However, such individuals have a reduced risk of infection with these variants compared to infection-naive individuals (82% and 45% reduction for delta and omicron,

respectively) [34]. We assumed that individuals previously infected with the delta variant could only be reinfected with the omicron variant, with an infection risk reduced by 45% compared to infection-naive individuals [34]. The other parameters used to represent strain-specific characteristics are presented in Table S2.

Seeding of each new strain into the model was achieved through the importation of a small number (10 per million population) of new infectious persons with the relevant strain into the model. The seeding process was done over a ten-day period and the start of this period was set to the emergence date reported by GISAID for each country [10].

#### 1.2.5 Dynamic social mixing

The model captures changes in social interactions over time through a dynamic age-specific mixing matrix. The following sections describe how this matrix was defined and how it captures the different non-pharmaceutical interventions implemented in the analysed countries, including school closures. The overall approach is also illustrated by Figure S6.

##### 1.2.5.1 Reference mixing matrices

We extracted country-specific contact matrices using the *conmat* R package which derives social mixing matrices from contact survey data. These matrices provide the average numbers of contacts per day between different age groups, disaggregated by the following locations: home, school, work, other locations.

The overall contact matrix (before adjustments for mobility changes) results from the summation of the four location-specific contact matrices: $C_0 = C_H + m_S C_S + C_W + C_L$, where $C_H$, $C_S$, $C_W$ and $C_L$ are the age-specific contact matrices associated with households, schools, workplaces and other locations, respectively. Note that the school contribution $C_S$ is multiplied by the factor $m_S$ that was varied during model calibration (see Section 3), in order to account for uncertainty around the relative contribution of school contacts to overall mixing.

##### 1.2.5.2 Modifications of contact rates over time

To capture mobility changes over time, the contributions of the matrices $C_S$, $C_W$ and $C_L$ vary with time such that the input contact matrix can be written as:

$$C(t) = h(t)^2 C_H + s(t)^2 m_S C_S + w(t)^2 C_W + l(t)^2 C_L \tag{3}$$

The modifying functions $h$ (for households), $s$ (for schools), $w$ (for work) and $l$ (for other-locations) were each squared to capture the effect of the mobility changes on both the infector and the infectee in any given interaction that could potentially result in transmission.

**School closure/re-opening**

Reduced attendance at schools was represented through the function $s$, which represents the proportion of all school students currently attending on-site teaching. If schools are fully closed at time $t$, $s(t) = 0$ and $C_S$ does not contribute to the overall mixing matrix $C(t)$. The function $s$ was derived from the UNESCO database on school closures from the start of the COVID-19 pandemic [37]. This database provides school opening status over time as a categorical variable taking the following values: "Fully open", "Partially open", "Academic break", "Closed due to COVID-19". Table S1 indicates how the different categorical values were converted into the numerical function $s$.

Table S1: Assumed percentage of students on-site for the different UNESCO school closure categories.

| **UNESCO category** | **Assumed proportion of students on-site at national level ($s(t)$)** |
|---|---|
| Fully open | 100% |
| Partially open | 10-50% |
| Academic break | 0% |
| Closed due to COVID-19 | 0% |

We included uncertainty around the value associated with the partial closure category, as there was no quantitative data available to inform this parameter [37]. The partial closure periods are likely to be periods where only a small fraction of students such as children of "essential workers" were attending school. We assumed that between 10 and 30% of students attended on-site learning during these periods.

To model the counterfactual "no school closure" scenario, we assumed that the schools were "Fully open" during the periods reported as "Partially open" or "Closed due to COVID-19".

Figure S5 summarises the UNESCO data on school closure for the analysed countries.

**Dynamic mobility outside of schools and homes**

Changes to people's mobility in places other than schools and homes were modelled using Google Mobility data, after applying a seven-day moving average smoothing. We used the "Workplace" category of the Google data to scale the work-related matrix contribution $C_W$ to overall mixing over time, using the adjusting function $w$. The "other locations" matrix $C_L$ was scaled through the adjusting function $l$ which was defined as the average of the Google mobility indicators across the following Google categories: "Retail and recreation", "Grocery and pharmacy" and "Transit stations".

**Household contacts**

In the base case analysis, the contribution of household contacts to the overall mixing matrix was fixed over time (i.e. $h(t) = 1$ in Equation 3). Although Google provides mobility estimates for residential contacts, the nature of these data is different from that of each of the other Google mobility types. They represent the time spent in that location, as opposed to other categories, which measure a change in total visits rather than the duration. The daily frequency with which people attend their residence is likely to be close to one, and we considered that household members likely have a daily opportunity for infection with each other household member regardless of the background level of mobility. Therefore, we did not implement a function to scale the contribution of household contacts to the mixing matrix with time.

#### 1.2.5.3 Sensitivity analyses around dynamic social mixing

In addition to the Base Case analysis described above, we performed two sensitivity analyses considering different assumptions for the modelled social mixing.

**SA1: No Google mobility data** In a first sensitivity analysis (SA1), we removed the contribution of the Google mobility data to the modelled social mixing (see Figure S6). In this configuration, the calibrated random process ($W(t)$) was used to capture mobility changes implicitly.

**SA2: School closures increase household contact rates** In another sensitivity analysis (SA2), we considered an alternative assumption under which the effective contact rates within households were increased during periods of school closure. In that case, the household component of the mixing matrix is modified by the following function:

$$h(t) = 1 + 0.20(1 - s(t)) \quad ,$$

where $s$ is the function modifying school contacts as introduced in Section 1.2.5.2. This is equivalent to assuming that each individual has 20% more household contact potential when schools are fully closed.

### 1.2.6 Random transmission adjustment

The risk of SARS-CoV-2 transmission per contact was adjusted by a time-variant random process, making the model semi-mechanistic. This random process reflects the fact that all the variations observed in the transmission risk in the real world cannot be explained solely by the factors that are explicitly captured through model inputs such as vaccination, dynamic mobility or new variants' emergence. We therefore allowed for random perturbations to the risk of transmission over time, although these perturbations were highly auto-correlated to avoid unrealistic changes over a short period of time.

We used a random walk with Gaussian update defined by:

$$\begin{aligned} W(0) &= 0 \\ W(t+1) &\sim \mathcal{N}(W(t), 0.5) \quad , \end{aligned} \tag{4}$$

where $\mathcal{N}$ denotes the normal distribution. The random process $W$ was updated every two months and was transformed using the exponential function before being applied to the risk of transmission per

contact (see Equation 6). Finally, the contribution of the random process to the risk of transmission was squared in order to capture its effect on both the susceptible and the infectious individuals (see Equation 6).

### 1.2.7 Ordinary differential equations

Let us first introduce some new notations. Modelled age groups are indicated by the subscript $a$, and $\mathcal{A}$ represents the set of all modelled age groups (see Section 1.2.2). Vaccination status is represented by the subscript $v$, and $\mathcal{V}$ is the set of vaccination statuses (i.e. $\mathcal{V} = \{\text{“0”, “1”}\}$, where “0” represents unvaccinated people and “1” represents vaccinated people). The subscript $s$ is used to represent the different viral strains, and $\mathcal{S}$ is the set of all strains (i.e. $\mathcal{S} = \{\text{“wild-type”, “delta”, “omicron”}\}$). The average incubation period duration associated with strain $s$ is denoted $q_s$ and the average duration of active disease is denoted $w$. The relative susceptibility to infection with strain $s$ of individuals aged $a$ with vaccination status $v$ is denoted $\rho_{a,v,s}$. The term $b_{a,v,s}(t)$ designates the introduction of individuals of age $a$ and with vaccination status $v$ that are infected with strain $s$ (infection seeding). Vaccination is characterised by the age-specific and time-variant per-capita vaccination rate $\omega_a$. Finally, $\chi_{s,\sigma}$ represents the relative susceptibility to infection with strain $\sigma$ for individuals whose most recent infection episode was with strain $s$. Using this new notation combined with those previously introduced, we can describe the model with the following set of ordinary differential equations:

$$
\begin{aligned}
\frac{dS_{a,v}}{dt} &= -\sum_{s\in\mathcal{S}} \lambda_{a,s}(t)\rho_{a,v,s}S_{a,v} + \Phi_v\omega_a(t)S_{a,v=0} \quad , \\
\frac{dE^1_{a,v,s}}{dt} &= \lambda_{a,s}(t)\rho_{a,v,s}\Big(S_{a,v} + \sum_{\sigma\in\mathcal{S}} \chi_{\sigma,s}R_{a,v,\sigma}\Big) - \frac{4}{q_s}E^1_{a,v,s} + \Phi_v\omega_a(t)E^1_{a,v=0,s} \quad , \\
\frac{dE^k_{a,v,s}}{dt} &= \frac{4}{q_s}E^{k-1}_{a,v,s} - \frac{4}{q_s}E^k_{a,v,s} + \Phi_v\omega_a(t)E^k_{a,v=0,s} \quad , \forall k \in \{2,3,4\}, \\
\frac{dI^1_{a,v,s}}{dt} &= \frac{4}{q_s}E^4_{a,v,s} - \frac{4}{w}I^1_{a,v,s} + b_{a,v,s}(t) + \Phi_v\omega_a(t)I^1_{a,v=0,s} \quad , \\
\frac{dI^k_{a,v,s}}{dt} &= \frac{4}{w}I^{k-1}_{a,v,s} - \frac{4}{w}I^k_{a,v,s} + \Phi_v\omega_a(t)I^k_{a,v=0,s} \quad , \forall k \in \{2,3,4\}, \\
\frac{dR_{a,v,s}}{dt} &= \frac{4}{w}I^4_{a,v,s} - \sum_{\sigma\in\mathcal{S}} \lambda_{a,\sigma}(t)\rho_{a,v,\sigma}\chi_{s,\sigma}R_{a,v,s} + \Phi_v\omega_a(t)R_{a,v=0,s} \quad ,
\end{aligned}
\tag{5}
$$

where $\lambda_{a,s}$ represents the force of infection of strain $s$ affecting individuals of age $a$. The quantity $\Phi_v$ is a binary variable used to switch between a positive and negative multiplier depending on vaccination status. It is equal to 1 when $v =$“1” and $-1$ when $v =$“0”. In other words, $\Phi_v = 2\mathbb{1}_{v=\text{“1”}} - 1$.

The force of infection was calculated as:

$$
\lambda_{a,s}(t) = \beta e^{2W(t)}\psi_s \sum_{\alpha\in\mathcal{A}}\sum_{v\in\mathcal{V}} \frac{c_{a,\alpha}(t)}{N_{\alpha,v}}\Big(0.5\sum_{k=3}^{4} E^k_{\alpha,v,s} + \sum_{k=1}^{4} I^k_{\alpha,v,s}\Big) \quad . \tag{6}
$$

In the previous equation, $\beta$ represents the unadjusted risk of transmission per contact, $\psi_s$ is the relative infectiousness of strain $s$, and $W(t)$ is the random process introduced in Section 1.2.6. The size of the population of age $\alpha$ with vaccination status $v$ is denoted $N_{\alpha,v}$. The term $c_{a,\alpha}(t)$ is a single element of the contact matrix $C(t)$ introduced in Equation 3. It represents the average numbers of contacts per day that a individual of age $a$ has with individuals of age $\alpha$.

## 1.3 Estimation of COVID-19-related hospital pressure and deaths

The transmission model described in Section 1.2 provides estimates of COVID-19 incidence over time, disaggregated by age, vaccination status and strain. We combine these incidence estimates with the age-, vaccination- and strain-specific risks of hospitalisation and deaths, as well as statistical distributions of time to events to compute COVID-19-related hospital pressure and deaths over time.

### 1.3.1 COVID-19-related hospital pressure

The risk of hospitalisation given infection was expected to vary markedly by setting. For example, different countries may have different criteria for whether or not a COVID-19 patient should be admitted to a hospital. This makes it difficult to provide accurate estimates of hospitalisation rates for multiple countries.

For this reason we introduced a universal indicator named "hospital pressure" in our analysis. This indicator was obtained by considering the age-specific risk of hospitalisation given infection observed in the first year of the pandemic in the Netherlands, adjusted for vaccination status and for the infecting strain (Table S2). The "hospital pressure" indicator can therefore be interpreted as the level of hospital occupancy that would be observed in the analysed country if the rates of hospitalisation given infection in this country were the same as for the Netherlands. This quantity is expected to vary proportionately with occupancy over time, providing an indicator of hospital pressure. Note that this indicator was used in order to make comparisons between scenarios, such that one should interpret the relative differences between scenarios rather than the absolute values of the indicator.

Let us denote $i_{a,v,s}(t)$ the number of new disease episodes estimated to start at time $t$ for people aged $a$ with vaccination status $v$ and infected with strain $s$. The number of new hospital admissions occurring at time $t$ was calculated using the following convolution product:

$$\eta(t) = \sum_{a,v,s} \kappa_{a,v,s} \int_{u \geq 0} i_{a,v,s}(t-u) g_h(u) du \quad , \tag{7}$$

where $\kappa_{a,v,s}$ is the risk of hospitalisation given infection for age $a$, vaccination status $v$ and strain $s$ based on the Netherlands data, and $g_h$ is the probability density function of the statistical distribution chosen to represent the time from symptom onset to hospitalisation (Table S2).

We then computed the "hospital pressure" quantity $h$, which is an indicator of hospital occupancy level, by combining the number of new hospital admissions $\eta$ with the statistical distribution used to model hospital stay duration:

$$h(t) = \int_{u \geq 0} \eta(t-u)(1-\tau(u)) du \quad , \tag{8}$$

where $\tau$ is the cumulative density function of the statistical distribution chosen to represent the hospital stay duration (Table S2).

### 1.3.2 COVID-19 deaths

We estimated the number of COVID-19 deaths over time using a similar approach as for the hospital pressure indicator. We used the age-specific infection fatality rates reported in O'Driscoll et al. [27], adjusted for vaccination status and for the infecting strain to estimate COVID-19 mortality. Using the same notations as in Section 1.3.1, the number of COVID-19 deaths observed at time $t$ was obtained by:

$$\mu(t) = m_C \sum_{a,v,s} ifr_{a,v,s} \int_{u \geq 0} i_{a,v,s}(t-u) g_d(u) du \quad , \tag{9}$$

where $ifr_{a,v,s}$ is the risk of death given infection for age $a$, vaccination status $v$ and strain $s$, and $g_d$ is the probability density function of the statistical distribution chosen to represent the time from symptom onset to death (Table S2). We used the country-specific adjuster $m_C$ to capture the fact that the infection fatality ratio is expected to vary by country, in part due to differences in COVID-19 death definition and reporting standards. This adjustment was automatically calibrated by the MCMC (Section 3).

## 1.4 Model parameters

### 1.4.1 Parameter tables

The model parameters and their values (or associated prior distributions) are listed in Table S2 and Table S3. The following sections include discussion of the evidence and reasoning used to inform the parameters.

Table S2: Model parameters

| Parameter | Value/Distribution | Evidence |
|---|---|---|
| Transmission probability per contact [C] | Uniform (0.01, 0.06) | Calibrated |
| Mean active disease period (days) | 4.5 | [33, 12, 19] |
| Country-specific IFR multiplier ($m_C$) [C] | Uniform (0.5, 1.5) | Calibrated |
| Uncertainty multiplier for school contacts ($m_S$) [C] | Uniform (0.8, 1.2) | Calibrated |
| Prop. students on-site during "Partially open" periods [C] | Uniform (0.1, 0.5) | Calibrated |
| VE against infection | 0.7 | [3, 29, 13, 4, 17] |
| VE against hospitalisation | 0.9 | [7, 4, 17] |
| VE against death | 0.9 | [7, 17] |
| Time from symptom onset to hospitalisation (days) | Gamma (shape=5, mean=3) | [11] |
| Hospital stay duration (days) | Gamma (shape=5, mean=9) | [11] |
| Time from symptom onset to death (days) | Gamma (shape=10, mean=15.93) | [18] |
| **Strain-specific parameters** | | |
| *Wild-type strain* | | |
| Mean incubation period (days) | 6.65 | [39] |
| *Delta variant* | | |
| Mean incubation period (days) | 4.41 | [39] |
| Relative intrinsic transmissibility (ref. wild-type) | 1.5 | [20, 25] |
| Relative risk of hospitalisation (ref. wild-type) | 2.0 | [9] |
| Relative risk of death (ref. wild-type) | 2.3 | [9] |
| Prop. escaping vaccine immunity against infection | 0.3 | [15, 23] |
| *Omicron variant* | | |
| Mean incubation period (days) | 3.42 | [39] |
| Relative intrinsic transmissibility (ref. wild-type) | 2.0 | [15, 23] |
| Relative risk of hospitalisation (ref. wild-type) | 0.82 | [9, 26] |
| Relative risk of death (ref. wild-type) | 0.71 | [9, 26] |
| Prop. escaping vaccine immunity against infection | 0.6 | [15, 23] |

Table S3: Age-specific parameters for wild-type COVID-19

| Age group | Rel. suscepitbility to infection (ref. 15-69 y.o.) [40] | Proportion symptomatic [31] | Proportion of symptomatic patients hospitalised [30] | Infection fatality rate [27] |
|---|---|---|---|---|
| 0-4 | 0.36 | 0.533 | 0.0777 | 0.00003 |
| 5-9 | 0.36 | 0.533 | 0.0069 | 0.00001 |
| 10-14 | 0.36 | 0.533 | 0.0034 | 0.00001 |
| 15-19 | 1.00 | 0.533 | 0.0051 | 0.00003 |
| 20-24 | 1.00 | 0.679 | 0.0068 | 0.00006 |
| 25-29 | 1.00 | 0.679 | 0.0080 | 0.00013 |
| 30-34 | 1.00 | 0.679 | 0.0124 | 0.00024 |
| 35-39 | 1.00 | 0.679 | 0.0129 | 0.00040 |
| 40-44 | 1.00 | 0.679 | 0.0190 | 0.00075 |
| 45-49 | 1.00 | 0.679 | 0.0331 | 0.00121 |
| 50-54 | 1.00 | 0.679 | 0.0383 | 0.00207 |
| 55-59 | 1.00 | 0.679 | 0.0579 | 0.00323 |
| 60-64 | 1.00 | 0.803 | 0.0617 | 0.00456 |
| 65-69 | 1.00 | 0.803 | 0.1030 | 0.01075 |
| 70-74 | 1.41 | 0.803 | 0.1072 | 0.01674 |
| 75-79 | 1.41 | 0.803 | 0.0703 | 0.03203 |
| 80 and above | 1.41 | 0.803 | 0.0703 | 0.08292 |

### 1.4.2 Incubation period and active disease duration

A systematic review and meta-analysis was conducted in 2022 attempting to estimate the mean incubation period for different SARS-CoV-2 variants [39]. Their pooled estimate for the mean incubation period for the wild-type variant was 6.65 days (95% CI 6.30-6.99), leveraging estimates from 119 individual studies [39]. The mean incubation period for the Delta variant was found to be 4.41 days (3.76-5.05). The pooled estimate for the mean incubation period for the Omicron variant was 3.42 days (2.88-3.96). These estimates are consistent with evidence from a rapid review that identified shortening of the serial interval of the Delta variant compared to wild-type, and further shortening of the serial interval for the Omicron variant compared to the Delta and wild-type variants [24].

Active disease duration is difficult to estimate for a number of reasons. Observing the period of time over which transmission events occur from a primary case to their secondary cases can be biased towards shorter time periods due to interventions. Particularly case isolation truncates the period at which secondary infections can occur. Proxies of infectiousness such as positive rapid antigen positivity, PCR or viral culture may not perfectly reflect true infectiousness. Duration of live virus shedding by viable viral culture is a better proxy indicator than PCR positivity and duration of live virus shedding is likely the best proxy coupled with epidemiological evidence to indicate the upper limit for duration of infection based on observed secondary cases. In studies that recruited people infected in the community, the median duration of viable viral shedding has been shown to be 4-5 days from symptom onset for the ancestral/wild-type strain [33, 12]. This is coupled with evidence of the possibility of transmission occurring 1-3 days prior to the onset of symptoms places the duration of infection at around 6-7 days on average for the ancestral/wild-type strain [14]. This is supported by evidence of the duration of variable virus detectable in a small number of healthy young adults in a human challenge study being 6.5 days [19].

In our model, the active disease duration parameter defines the time spent in the compartments $I_1 ... I_4$ (Figure S2), so we assumed a mean value of 4.5 days for this duration. Moreover, our assumption that infected individuals are infectious during the second half of the incubation period (compartments $E_3$ and

$E_4$) is consistent with the evidence above.

### 1.4.3 Times to hospitalisation and death / hospital stay duration

This section describes the evidence used to inform the statistical distributions listed in Table S2 regarding times to disease outcomes (hospitalisation and death) and hospital stay duration.

*Time from symptom onset to hospitalisation*

To inform the mean time from symptom onset to hospitalisation we used the estimate published in the March 2022 International Severe Acute Respiratory and emerging Infections Consortium (ISARIC) Clinical Data Report [11]. ISARIC developed a standardised reporting system to systematically collect, analyse and report on COVID-19 clinical data over the pandemic. Since the start of the pandemic, they have collected clinical data on over 800,000 individuals across 60 countries. The published estimate provided in this latest dated report and used in our model represents the mean time across all COVID-19 variants and so we believe it is the best representative estimate to use in this analysis, as we assumed this distribution is the same across all SARS-CoV-2 variants modelled.

*Hospital stay duration*

To inform the mean duration of hospital stay for COVID-19 cases we use the estimate published in the ISARIC Clinical Data Report published March 2022 [11]. The ISARIC data is briefly described above. The published estimate provided in this latest dated report and used in our model represents the mean time across all COVID-19 variants.

*Time from symptom onset to death*

The mean time from symptom onset to death was informed by a systematic review and meta-analysis published in 2020 [18]. This pooled estimation was based on three studies that examined the clinical characteristics and outcomes of cases from early outbreaks in China, in 2020.

### 1.4.4 Vaccine efficacy (against wild-type virus)

We aimed to generate an approximate average estimate for vaccine efficacy (VE) against the outcomes below, that best represents the protection from two-doses of the most common vaccine types used globally. First, we were interested in the protection of vaccines against the wild-type virus for parameterisation of our model, and we then considered the different strains' characteristics to determine how vaccine protection should be adjusted for the different viral strains (Section 1.4.5). We considered studies evaluating the efficacy of adenoviral vector vaccines (namely AZD1222/ChAdOx1 nCoV-19), the mRNA vaccines (BNT162b2 and mRNA-1273) and the inactivated whole-virion vaccines (namely, Coronavac, BIBP-CorV and BBV152).

*Vaccine efficacy against infection*

The results of ChAdOx1 nCoV-19 trial (adenoviral vector vaccine) demonstrated an efficacy against infection of 64.3% (95% CI, 56.1 to 71.0) [8]. Whilst, data from the mRNA-1273 (mRNA vaccine) trial suggested a VE against infection of 82% (79.5-84.2) against any infection [5]. The BNT162b2 (mRNA vaccine) trial only evaluated the efficacy against symptomatic infection, with a vaccine efficacy of 91.3% (89.0-93.2) [36]. This is similar to the VE of mRNA-1273 against symptomatic infection and so we assume BNT162b2 likely has a similar VE against infection as mRNA-1273 [5, 36]. Inactivated vaccine trials estimated a two-dose VE against infection of 68.8% (46.7–82.5) (BBV152), 64.0% (48.8-74.7) (BIBP-CorV, WIV04 strain) and 73.5% (60.6-82.2) (BIBP-CorV, HB02 strain) [6, 1]. For Coronavac, clinical trials focused on VE against symptomatic infection, with a VE of 83.5% (65.4-92.1), which is similar to the VE against symptomatic infection seen for BBV152 and BIBP-CorV and so we assume that Coronavac has a similar VE against infection [35, 6, 1]. Combining this evidence suggests that 2-dose protection against infection ranges from approximately 60-80% and so we assume that pooled effectiveness would fall in the mid-point 70%.

*Vaccine efficacy against hospitalisation*

This parameter represents the VE against hospitalisation once infected with the wild-type virus. The early trials for two doses (and one dose) of COVID-19 vaccines conducted in 2020 during wild-type circulation were either: 1. Underpowered to precisely estimate the efficacy on the outcome of hospitalisation [8, 35]. Or 2. Evaluated efficacy against severe COVID-19 as a secondary end-point and not explicitly hospitalisation [5, 36, 6, 1]. However, these trials still suggested that the vaccines significantly protect against hospitalisation and severe disease (a possible proxy of hospitalisation, with point-estimates ranging from 93-100% efficacy [8, 35, 5, 36, 6, 1].

Observational studies conducted slightly later in the pandemic, late 2020 into 2021 during periods where variants such as alpha, beta and gamma were circulating, but before the emergence of the Delta variant, however, were able to estimate the VE against hospitalisation of 87% (55-100) for BNT162b2 and 87.5% (86.7–88.2) for Coronavac [4, 17]. We assumed protection against hospitalisation is higher against the wild-type virus than these estimates against variants of concern and so we assumed the modelled VE of 2-doses against hospitalisation is 90% based on the combined evidence.

*Vaccine efficacy against death*

This parameter represents the VE against death once infected with wild-type virus. The early vaccine trials were underpowered to precisely evaluate the outcome of death, however as discussed earlier some demonstrated an effect on severe disease, which precedes COVID-19 death [5, 36, 6, 1].

An observational study in older adults (who have a greater risk death compared to younger ages) in the UK estimated a VE against death for two doses of BNT162b2 of 69% (31-86) following infection [16]. In broader population studies estimated a VE of 84% (44-100) following one dose of BNT162b2 and 86.3% (84.5 to 87.9) following two doses of Coronavac [4, 17]. Furthermore, evaluation of VE against death in Scotland across a broad population following infection with the more severe delta variant estimated a 2-dose VE of 90% (83-94) for BNT162b2 and 91% (86-94) for AZD1222 [32]. Therefore we assume that VE against death following infection with the wild-type virus as modelled in this analysis to be 90% based on the combined evidence.

### 1.4.5 Variant-specific adjustments

#### 1.4.5.1 Delta variant

*Relative intrinsic transmissibility*

Part of the Delta variant's transmission advantage is likely conferred by it being intrinsically more transmissible than prior circulating variants. Estimating how much more intrinsically transmissible it is is difficult in practice because of additional properties around immune evasion of both vaccination and prior infection derived immunity that also give the Delta variant a transmission advantage.

Our assumption is that the relative intrinsic transmissibility as implemented in the model can be approximated by taking the ratio of the estimates of $R_0$ for each variant. $R_0$ is increasingly difficult to estimate for new emerging variants such as the Delta variant due to accumulation of population immunity through vaccination and infection. One study from an outbreak of the Delta variant in a relatively immune naive population in Guangdong Province, China, estimated an R0 of 3.2 [25]. This is approximately 1.5 times that of wild-type estimates from early in the pandemic, $R_0 = 2.2$ (95% CI, 1.4 to 3.9) [20].

Other estimates of the relative effective reproduction number for the Delta variant compared to wild-type reports a two-fold increase in transmissibility [2]. However, this did not account for the immune escape advantage that Delta may have, so is still consistent with our assumptions.

*Relative intrinsic risk of hospitalisation*

A large population-level cohort study in Ontario, Canada found that the adjusted odds of hospitalisation following infection with the Delta variant were 2.08 (1.78-2.40) that of wild-type virus infections [9]. We used the approach set out in [41] to crudely convert the point-estimate of the adjusted odds ratio to an approximate estimate of the relative risk of 2 in the model.

*Relative intrinsic risk of death*

The same cohort study conducted in Ontario found that the adjusted odds of death following infection with the Delta variant were 2.33 (1.54-3.31) that of wild-type virus infections [9]. Employing the same conversion approach as mentioned in the previous paragraph, we used a relative risk of 2.3 in the model.

*Immune escape property*

Part of the Delta variant's transmission advantage is likely conferred by its ability to evade vaccine-induced immunity. In particular there is evidence for reduced activity in immune correlates for protection against infection, such as neutralising antibody titres [28]. However, it is less clear how this translates into VE against infection. One population level observational study in the UK estimated two doses of either an mRNA or adenoviral vector vaccine at being approximately 10% less effective at preventing symptomatic COVID-19 infection for the Delta variant compared to the Alpha variant [21]. However in highly-vaccinated populations this is likely an estimate of the combined effect that includes the protection of the vaccine against onward transmission in infected, which we do not capture in this model. The best estimate we have found for VE against Delta variant infection from unvaccinated source, comes a Danish household transmission study and estimates a VE of 46% (40-52) for two doses of an mRNA or adenoviral vector vaccine [22]. Similarly, a household transmission study in Norway estimated the VE against infection for 2-dose-mRNA-vaccinated individuals following contact with a Delta household contact was 42% (23-55) [15]. These estimates are about 60-66% of our assumed modelled VE of 70% based on prior evidence from vaccine trials (Section 1.4.4). Given the uncertainty we assume a conservative estimate that the Delta variant is able to escape 30% of prior vaccine immunity against infection.

#### 1.4.5.2 Omicron variant

*Relative intrinsic transmissibility*

As we described above for the Delta variant, the Omicron variant likely has a combined transmission advantage due to increased intrinsic transmissibility and its ability to evade prior immunity from infection and vaccination. $R_0$ estimates for the Omicron variant are very difficult due to the high levels of population immunity from infection and vaccination in nearly every setting at the time of its emergence. However, estimates of the risk of infection in unvaccinated close contacts of individuals infected with the Omicron variant compared to other variants should also allow us to estimate the change in relative intrinsic transmissibility. These estimates largely only exist comparing the first Omicron subvariant BA.1 against the Delta variant and do not exist relative to the wild-type variant. We can, however, consider the product of our relative intrinsic transmissibility for the Delta variant compared to wild-type and our estimate of the relative intrinsic transmissibility for the Omicron variant compared to Delta. A household transmission study in Norway found that unvaccinated household contacts of primary cases infected with Omicron BA.1 had a risk of being infected that is 1.27 times that of unvaccinated household contacts of a Delta infected primary case [15]. However, susceptibility is just one component of transmissibility as infectiousness also plays a role. The same study estimated that household close contacts of unvaccinated Omicron BA.1 primary cases had a relative risk of being infected 1.5 times that of households close contacts of unvaccinated Delta primary cases [15]. This suggests that the Omicron BA.1 variant was between 1.2-1.5 times as transmissible as the Delta variant in unvaccinated individuals. Therefore, in this model we assumed the modelled Omicron variant is 2 times (1.3, (mid-point of 1.2-1.5) multiplied by 1.5 (the relative intrinsic transmissibility of the Delta variant)) as intrinsically transmissible as the wild-type virus.

*Relative intrinsic risk of hospitalisation*

We could not find clear estimates of the relative intrinsic risk of hospitalisation for Omicron variant infection compared to the wild-type virus. However, we could find estimates for the relative risk of Omicron compared to Delta and thus can use the product of this estimate with our estimate of the relative risk of hospitalisation with the Delta variant compared to the wild-type virus. A large population-level study in England found an adjusted hazard ratio for hospitalisation of 0.41 (0.39-0.43) for infection with the Omicron variant compared to the Delta variant [26]. We assumed that the hazard ratio can be used to estimate the relative risk and thus our estimate for the relative intrinsic risk of hospitalisation for Omicron variant infection compared to the wild-type was 0.82 ($0.41 \times 2$).

*Relative intrinsic risk of death*

To estimate the relative risk of death for Omicron compared to wild-type virus, we used the same approach as described above for relative risk intrinsic risk of hospitalisation. We leveraged estimates relative to the Delta variant to find an approximate estimate for this parameter in the model. A large population-level study in England found an adjusted hazard ratio for hospitalisation of 0.31 (0.26-0.37) for infection with the Omicron variant compared to the Delta variant [26]. We assumed that the hazard ratio can be used

to estimate the relative risk and thus our estimate for the relative intrinsic risk of hospitalisation for Omicron variant infection compared to the wild-type is 0.71 ($0.31 \times 2.3$).

*Immune escape property*
In vitro neutralisation studies leveraging sera from vaccinated individuals demonstrated substantial immune escape of Omicron BA.1 compared to the Delta variant [38]. The best estimate of VE against infection we could find was from a household transmission study in Norway which estimated a VE against infection with the Omicron variant of 27% (6-49) for 2-dose-vaccinated individuals [15]. This is about 65% of their estimated VE against the Delta variant and about 38% of our modelled VE for the wild-type virus of 70% [15]. We therefore assumed that the Omicron variant in the model is able to escape 60% of vaccine immunity.

# 2 Software and code used to conduct the analyses

## 2.1 Code

The Python code and data used to perform the analyses is fully available on Github at the following link: https://github.com/monash-emu/AuTuMN. In particular, the code associated with the implementation of the model is available here. The suite of functions used to run the model, perform calibration and analyse the outputs is available here. Finally, the model can be run for the included countries via an online Google Colab notebook, allowing the user to conduct their own analyses without any local software installation requirement. NEED TO CREATE NOTEBOOK

## 2.2 Software implementation

We used the *summer2* Python package (v 1.2.5) to implement the model. This is a domain specific library for compartmental epidemiological models that addresses a few of the key concerns as follows.

### 2.2.1 Application Programming Interface (API)

The model specification was done via a simple yet expressive Python API, while the numerical implementation was largely autogenerated by the *summer2* package at runtime. This specification is composable via stratification classes and other reusable components, thus the complexity of the software is kept to a minimum, reducing cognitive overhead and greatly reducing the possibility for error.

### 2.2.2 Optimising compiler

The *summer2* package uses the jax library as its computational backend, meaning that while the specification of models is done largely in Python, the model execution itself is transformed via an optimising compiler into fast native code. This brings the model runtime from several seconds (for a naive implementation) to under 50ms per iteration, which was necessary to perform the computationally-intensive calibration tasks described in Section 3.

# 3 Model calibration and uncertainty propagation

The model was calibrated using a Bayesian approach. In particular, we used the Adaptive Differential Evolution Metropolis (DEMetropolisZ) algorithm implemented with the *PyMC* Python package (v.5.2.0) to sample parameters from their posterior distributions. For each country, we ran 8 independent DEMetropolisZ chains of 35,000 iterations, each starting from a different starting point and using the first 5,000 draws for algorithm tuning. To determine the 8 starting points of the DEMetropolisZ chains, we conducted 8 independent optimisation searches using the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) method, implemented with the *nevergrad* Python package (v.0.6.0) and with a budget of 10,000 model evaluations per search. The 8 optimisation searches' initial points were randomly drawn from Latin Hypercube Sampling (LHS) based on the parameter priors shown in Table S2. The parameters sampled with LHS were the transmission probability per contact, the IFR multiplier, the uncertainty multiplier for school contacts, and the proportion of students on-site during "Partially open"

periods. The random process variables $W(t)$ (Section 1.2.6) were all set to 0 for each optimisation starting point.

Our calibration approach required a total of 360,000 ($8 \times (10000 + 35000)$) model evaluations per country analysis, which were completed in about two hours on a machine with 8 cpus and 32-GiB memory.

For each country, the results presented in the manuscript are associated with 1000 parameter sets randomly sampled from the posterior distributions obtained from DEMetropolisZ sampling (after discarding the first 25,000 iterations for each chain). The definitions of the prior distributions and the likelihood are detailed in the following sections.

## 3.1 Parameters varied during calibration

The parameters varied during calibration along with their associated prior distributions are listed in Table S2 and indicated with the superscript $^{C}$. We used uniform prior distributions for all calibrated parameters. The primary parameters varied during calibration are the unadjusted risk of transmission per contact ($\beta$), the IFR multiplier ($m_C$), the original infection seeding time, the proportion of students on-site during "Partially open" periods, and the uncertainty multiplier modifying the school contacts' contribution ($m_S$).

Note that the values of the random process $W_t$ described in Section 1.2.6 are also treated as calibrated parameters by the MCMC. The Gaussian auto-regressive component described in Equation 4 is incorporated in the posterior likelihood computation (Section 3.3).

## 3.2 Calibration targets

Model calibration was performed independently for each country. All models were fitted to the reported number of COVID-19 deaths over time. We used the daily number of COVID-19 deaths reported by WHO and applied a 7-day moving average to the observed data.

In addition, for countries where a nationally representative seroprevalence survey had been conducted (n=36), we include seroprevalence data in the calibration likelihood. We used the online platform SeroTracker to extract country-specific seroprevalence estimates. The estimates had to verify the following conditions to be included in the analysis as calibration targets:

- Be aligned with the World Health Organization's Unity protocol (WHO Unity) for general population seroepidemiological studies.
- Have a sample size greater than 599 (minimum sample size recommended in WHO sero-surveys protocol)
- Have a sampling start date later than 1 May 2020 (to avoid very early surveys that may be less accurate)
- Be nationally representative (as classified by SeroTracker)

We further excluded studies that focused on specific population subgroups presenting a risk of selection bias for seroprevalence (e.g. pregnant women, slum population, healthcare workers, quarantine workers). Finally, to minimise the risk of interference with vaccination, we only included studies for which the vaccination coverage at the time of the survey (midpoint date) was lower than 10% of the measured seroprevalence.

All the criteria listed above were first verified systematically using the SeroTracker database, and the extracted studies were then analysed individually by one author (RR) to check that all inclusion criteria were verified.

When a seroprevalence study was restricted to a specific age-group, we matched the survey estimate to the modelled seroprevalence measured in the closest modelled age-group. For example, as the age-group reported in the Kenya survey was 16-64 years old, we used the modelled seroprevalence in the age-group 15-69 years-old to inform model calibration.

When multiple estimates were available for a country, we selected the highest ranked estimate after ordering by the following preference criteria (applied in the presented order):

1. Lowest risk of bias (according to SeroTracker),

2. Latest sampling start date (when a greater number of infections have occurred, and to avoid bias due to early geographic heterogeneity),

The estimates used to inform the models are summarised in Table S4, with the associated original reports accessible by clicking on the countries' names.

Table S4: Seroprevalence data extracted from SeroTracker (national surveys).

| country | sampling start date | sampling end date | age min | age max | denom. value | serum pos prevalence | estimate grade | overall risk of bias |
|---|---|---|---|---|---|---|---|---|
| Australia | 2020-11-03 | 2021-03-12 | | 19 | 1685 | 0.23% | National | High |
| Austria | 2020-06-05 | 2020-12-04 | 18 | 72 | 20228 | 2.5% | National | High |
| Belgium | 2020-10-12 | 2020-10-17 | | 101 | 2966 | 4.18% | National | Low |
| Brazil | 2020-05-14 | 2020-06-23 | | | 89362 | 2.3% | National | Moderate |
| Canada | 2021-04-13 | 2021-04-30 | 17 | | 16931 | 26.92% | National | Moderate |
| Chile | 2020-09-25 | 2020-11-25 | 7 | 94 | 2493 | 10.4% | National | Low |
| Colombia | 2020-09-21 | 2020-12-11 | 5 | 80 | 17863 | 32.53% | National | Moderate |
| Croatia | 2020-12-15 | 2021-02-15 | | | 1436 | 25.1% | National | High |
| Czechia | 2021-02-01 | 2021-03-31 | 18 | | 19548 | 51.0% | National | High |
| Denmark | 2020-12-01 | 2020-12-31 | 12 | | 4044 | 4.3% | National | Low |
| Ecuador | 2020-10-12 | 2020-10-19 | | | 1250 | 11.68% | National | Moderate |
| Egypt | 2021-01-15 | 2021-06-15 | | | 2360 | 46.3% | National | Moderate |
| France | 2020-05-11 | 2020-05-17 | | | 3592 | 4.93% | National | Low |
| Germany | 2020-10-26 | 2020-11-18 | 18 | | 9929 | 1.1% | National | Moderate |
| Honduras | 2020-06-16 | 2020-06-23 | 5 | | 792 | 6.2% | National | Moderate |
| Hungary | 2020-05-01 | 2020-05-16 | 14 | | 10474 | 0.68% | National | Low |
| India | 2020-12-18 | 2021-01-06 | 10 | | 28598 | 24.1% | National | Low |
| Israel | 2020-06-28 | 2020-09-14 | | | 54357 | 4.6% | National | Moderate |
| Italy | 2020-05-25 | 2020-07-15 | | | 64660 | 2.5% | National | Moderate |
| Japan | 2020-06-01 | 2020-06-07 | 20 | | 7950 | 0.1% | National | Moderate |
| Jordan | 2020-12-27 | 2021-01-06 | | | 5044 | 34.2% | National | Moderate |
| Kazakhstan | 2020-07-16 | 2021-07-07 | | | 85346 | 63.0% | National | High |
| Kenya | 2021-01-03 | 2021-03-15 | 16 | 64 | 3018 | 48.5% | National | Moderate |
| Lebanon | 2020-12-07 | 2021-01-15 | | | 2058 | 18.5% | National | Low |
| Lithuania | 2020-08-10 | 2020-09-10 | 18 | 92 | 3089 | 1.4% | National | Moderate |
| Mexico | 2020-08-15 | 2020-11-15 | 3 | 12 | 944 | 18.7% | National | Low |
| Nepal | 2020-10-09 | 2020-10-22 | | | 3040 | 14.4% | National | Low |
| Pakistan | 2020-10-21 | 2020-11-08 | | | 4998 | 7.02% | National | Moderate |
| Portugal | 2020-09-08 | 2020-10-14 | | | 13398 | 2.2% | National | Moderate |
| Slovenia | 2020-10-17 | 2020-11-10 | | 99 | 1211 | 4.29% | National | Low |
| South Africa | 2021-01-15 | 2021-05-15 | 15 | 69 | 16762 | 47.4% | National | Moderate |
| Rep. of Korea | 2020-09-24 | 2020-12-09 | 18 | 86 | 4085 | 0.39% | National | Moderate |
| Spain | 2020-06-08 | 2020-06-22 | | | 62167 | 5.2% | National | Low |
| Sweden | 2020-11-23 | 2020-12-04 | | | 3183 | 7.0% | National | Moderate |
| USA | 2020-08-09 | 2020-12-08 | 18 | | 4654 | 4.71% | National | Low |
| UK | 2020-08-24 | 2020-09-18 | 17 | | 8230 | 6.1% | National | Low |

## 3.3 Likelihood definition

Let $d_w$ denote the rounded average daily number of COVID-19 deaths during week $w$, and $\hat{d_w}^{\theta}$ the associated predicted number of deaths according to the model with parameter set $\theta$.

For countries with seroprevalence data, let us denote $\pi$ the measured seroprevalence proportion extracted from SeroTracker (Section 3.2). Let $\hat{\pi}^{\theta}$ denote the modelled age-matched proportion ever infected

by the time the survey was conducted (using the midpoint date) associated with the parameter set $\theta$. The likelihood was defined as follows for countries with seroprevalence data:

$$\mathcal{L}(\theta) := f_{\sigma}(\pi|\hat{\pi}^{\theta}) \times \prod_{w} g_r(d_w|\ \hat{d_w}^{\theta}) \quad , \tag{10}$$

where $f_{\sigma}(.|\mu)$ is the probability density function of a $[0,1]$-truncated normal distribution with mean $\mu$ and standard deviation $\sigma$; and $g_r(.|\mu)$ is the probability mass function of a negative binomial distribution with mean $\mu$ and overdispersion parameter $r$. The overdispersion parameter $r$ was automatically estimated by the MCMC algorithm, while the standard deviation $\sigma$ was set to different values depending on the SeroTracker-reported risk of bias associated with the seroprevalence estimate ($\sigma = 0.05$ if "Low", $\sigma = 0.1$ if "Moderate", $\sigma = 0.2$ if "High").

For countries without seroprevalence data, the likelihood equation reduces to:

$$\mathcal{L}(\theta) := \prod_{w} g_r(d_w|\ \hat{d_w}^{\theta}) \quad . \tag{11}$$

The likelihood functions described above represent the goodness of fit of a particular model parameterisation with regards to the targeted data. This quantity needs to be adjusted for the prior likelihood of the parameter set in order to compute the MCMC acceptance quantity $\mathcal{Q}(\theta)$. As we used uniform priors for all the parameters, the inclusion of the individual parameters' priors in the acceptance quantity is not necessary. Indeed, their respective contributions would cancel out as the same quantity would appear in the numerator and the denominator of the MCMC acceptance quantity ratio. However, the auto-regressive relationship described in Equation 4 must be accounted for as part of the combined prior likelihood of a parameter set. This prevents unrealistically large fluctuations of the random process. If $W^{\theta}$ represents the random process associated with the parameter set $\theta$, the overall MCMC acceptance quantity is obtained by:

$$\mathcal{Q}(\theta) = \mathcal{L}(\theta) \times \prod_{i=1}^{n} z_{W_{i-1}^{\theta},\epsilon}(W_i^{\theta}) \quad , \tag{12}$$

where $z_{\mu,\epsilon}(.)$ represents the probability density function of the normal distribution $\mathcal{N}(\mu, \epsilon)$, and $n$ is the number of random process updates.

# References


[1] Nawal Al Kaabi et al. "Effect of 2 Inactivated SARS-CoV-2 Vaccines on Symptomatic COVID-19 Infection in Adults: A Randomized Clinical Trial". In: *JAMA* 326.1 (2021), pp. 35–45. DOI: 10.1001/jama.2021.8565. URL: https://doi.org/10.1001/jama.2021.8565.

[2] Finlay Campbell et al. "Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021". In: *Eurosurveillance* 26.24 (2021), p. 2100509. DOI: doi:https://doi.org/10.2807/1560-7917.ES.2021.26.24.2100509. URL: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2021.26.24.2100509.

[3] Tiffany Charmet et al. "Impact of original, B.1.1.7, and B.1.351/P.1 SARS-CoV-2 lineages on vaccine effectiveness of two doses of COVID-19 mRNA vaccines: Results from a nationwide case-control study in France". In: *The Lancet Regional Health – Europe* 8 (2021). DOI: 10.1016/j.lanepe.2021.100171. URL: https://doi.org/10.1016/j.lanepe.2021.100171.

[4] Noa Dagan et al. "BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting". In: *New England Journal of Medicine* 384.15 (2021), pp. 1412–1423. DOI: 10.1056/NEJMoa2101765. URL: https://doi.org/10.1056/NEJMoa2101765.

[5] Hana M. El Sahly et al. "Efficacy of the mRNA-1273 SARS-CoV-2 Vaccine at Completion of Blinded Phase". In: *New England Journal of Medicine* 385.19 (2021), pp. 1774–1785. DOI: 10.1056/NEJMoa2113017. URL: https://doi.org/10.1056/NEJMoa2113017.

[6] Raches Ella et al. "Efficacy, safety, and lot-to-lot immunogenicity of an inactivated SARS-CoV-2 vaccine (BBV152): interim results of a randomised, double-blind, controlled, phase 3 trial". In: *The Lancet* 398.10317 (2021), pp. 2173–2184. DOI: 10.1016/S0140-6736(21)02000-6. URL: https://doi.org/10.1016/S0140-6736(21)02000-6.

[7] Public Health England. *COVID-19 vaccine surveillance report - Week 26*. Report. Jan. 2021. URL: https://www.gov.uk/government/publications/covid-19-vaccine-surveillance-report.

[8] Ann R. Falsey et al. "Phase 3 Safety and Efficacy of AZD1222 (ChAdOx1 nCoV-19) Covid-19 Vaccine". In: *New England Journal of Medicine* 385.25 (2021), pp. 2348–2360. DOI: 10.1056/NEJMoa2105290. URL: https://doi.org/10.1056/NEJMoa2105290.

[9] David N. Fisman and Ashleigh R. Tuite. "Evaluation of the relative virulence of novel SARS-CoV-2 variants: a retrospective cohort study in Ontario, Canada". In: *Canadian Medical Association Journal* 193.42 (2021), E1619. DOI: 10.1503/cmaj.211248. URL: http://www.cmaj.ca/content/193/42/E1619.abstract.

[10] GISAID. *GISAID database*. Web Page. 2023. URL: https://gisaid.org/.

[11] Isaric Clinical Characterisation Group et al. "ISARIC COVID-19 Clinical Data Report issued: 27 March 2022". In: *medRxiv* (2022), p. 2020.07.17.20155218. DOI: 10.1101/2020.07.17.20155218. URL: http://medrxiv.org/content/early/2022/04/13/2020.07.17.20155218.abstract.

[12] Seran Hakki et al. "Onset and window of SARS-CoV-2 infectiousness and temporal correlation with symptom onset: a prospective, longitudinal, community cohort study". In: *The Lancet Respiratory Medicine* 10.11 (2022), pp. 1061–1073. DOI: 10.1016/S2213-2600(22)00226-0. URL: https://doi.org/10.1016/S2213-2600(22)00226-0.

[13] Victoria Jane Hall et al. "COVID-19 vaccine coverage in health-care workers in England and effectiveness of BNT162b2 mRNA vaccine against infection (SIREN): a prospective, multicentre, cohort study". In: *The Lancet* 397.10286 (2021), pp. 1725–1735. DOI: 10.1016/S0140-6736(21)00790-X. URL: https://doi.org/10.1016/S0140-6736(21)00790-X.

[14] X. He et al. "Temporal dynamics in viral shedding and transmissibility of COVID-19". In: *Nat Med* 26.5 (2020), pp. 672–675. DOI: 10.1038/s41591-020-0869-5. URL: https://www.ncbi.nlm.nih.gov/pubmed/32296168.

[15] Neda Jalali et al. "Increased household transmission and immune escape of the SARS-CoV-2 Omicron compared to Delta variants". In: *Nature Communications* 13.1 (2022), p. 5706. DOI: 10.1038/s41467-022-33233-9. URL: https://doi.org/10.1038/s41467-022-33233-9.

[16] Bernal Jamie Lopez et al. “Effectiveness of BNT162b2 mRNA vaccine and ChAdOx1 adenovirus vector vaccine on mortality following COVID-19”. In: *medRxiv* (2021), p. 2021.05.14.21257218. DOI: 10.1101/2021.05.14.21257218. URL: http://medrxiv.org/content/early/2021/05/18/2021.05.14.21257218.abstract.

[17] Alejandro Jara et al. “Effectiveness of an Inactivated SARS-CoV-2 Vaccine in Chile”. In: *New England Journal of Medicine* 385.10 (2021), pp. 875–884. DOI: 10.1056/NEJMoa2107715. URL: https://doi.org/10.1056/NEJMoa2107715.

[18] Malahat Khalili et al. “Epidemiological characteristics of COVID-19: a systematic review and meta-analysis”. In: *Epidemiology and Infection* 148 (2020), e130. DOI: 10.1017/S0950268820001430. URL: https://www.cambridge.org/core/article/epidemiological-characteristics-of-covid19-a-systematic-review-and-metaanalysis/8B565B2FE5A97054E8B2564FB2CE6D3E.

[19] Ben Killingley et al. “Safety, tolerability and viral kinetics during SARS-CoV-2 human challenge in young adults”. In: *Nature Medicine* 28.5 (2022), pp. 1031–1041. DOI: 10.1038/s41591-022-01780-9. URL: https://doi.org/10.1038/s41591-022-01780-9.

[20] Q. Li et al. “Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia”. In: *N Engl J Med* 382.13 (2020), pp. 1199–1207. DOI: 10.1056/NEJMoa2001316. URL: https://www.ncbi.nlm.nih.gov/pubmed/31995857.

[21] Jamie Lopez Bernal et al. “Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant”. In: *New England Journal of Medicine* 385.7 (2021), pp. 585–594. DOI: 10.1056/NEJMoa2108891. URL: https://doi.org/10.1056/NEJMoa2108891.

[22] Frederik Plesner Lyngse et al. “Effect of vaccination on household transmission of SARS-CoV-2 Delta variant of concern”. In: *Nature Communications* 13.1 (2022), p. 3764. DOI: 10.1038/s41467-022-31494-y. URL: https://doi.org/10.1038/s41467-022-31494-y.

[23] Frederik Plesner Lyngse et al. “Household transmission of the SARS-CoV-2 Omicron variant in Denmark”. In: *Nature Communications* 13.1 (2022), p. 5573. DOI: 10.1038/s41467-022-33328-3. URL: https://doi.org/10.1038/s41467-022-33328-3.

[24] Zachary J. Madewell et al. “Rapid review and meta-analysis of serial intervals for SARS-CoV-2 Delta and Omicron variants”. In: *BMC Infectious Diseases* 23.1 (2023), p. 429. DOI: 10.1186/s12879-023-08407-5. URL: https://doi.org/10.1186/s12879-023-08407-5.

[25] Zhang Meng et al. “Transmission Dynamics of an Outbreak of the COVID-19 Delta Variant B.1.617.2 — Guangdong Province, China, May–June 2021”. In: *China CDC Weekly* 3.27 (2021), pp. 584–586. DOI: 10.46234/ccdcw2021.148. URL: https://weekly.chinacdc.cn//article/id/eb772589-1584-4ef9-beac-cac3ab2fbb12.

[26] T. Nyberg et al. “Comparative analysis of the risks of hospitalisation and death associated with SARS-CoV-2 omicron (B.1.1.529) and delta (B.1.617.2) variants in England: a cohort study”. In: *Lancet* 399.10332 (2022), pp. 1303–1312. DOI: 10.1016/s0140-6736(22)00462-7.

[27] M. O’Driscoll et al. “Age-specific mortality and immunity patterns of SARS-CoV-2”. In: *Nature* 590.7844 (2021), pp. 140–145. DOI: 10.1038/s41586-020-2918-0. URL: https://www.ncbi.nlm.nih.gov/pubmed/33137809.

[28] Eddy Pérez-Then et al. “Neutralizing antibodies against the SARS-CoV-2 Delta and Omicron variants following heterologous CoronaVac plus BNT162b2 booster vaccination”. In: *Nature Medicine* 28.3 (2022), pp. 481–485. DOI: 10.1038/s41591-022-01705-6. URL: https://doi.org/10.1038/s41591-022-01705-6.

[29] Emma Pritchard et al. “Impact of vaccination on new SARS-CoV-2 infections in the United Kingdom”. In: *Nature Medicine* 27.8 (2021), pp. 1370–1378. DOI: 10.1038/s41591-021-01410-w. URL: https://doi.org/10.1038/s41591-021-01410-w.

[30] Netherlands RIVM. *Epidemiologische situatie COVID-19 in Nederland*. Government Document. 2020. URL: https://www.rivm.nl/sites/default/files/2020-08/COVID-19_WebSite_rapport_wekelijks_20200804_1306.pdf.

[31] Pratha Sah et al. “Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis”. In: *Proceedings of the National Academy of Sciences* 118.34 (2021), e2109229118. DOI: 10.1073/pnas.2109229118. URL: http://www.pnas.org/content/118/34/e2109229118.abstract.

[32] Aziz Sheikh, Chris Robertson, and Bob Taylor. "BNT162b2 and ChAdOx1 nCoV-19 Vaccine Effectiveness against Death from the Delta Variant". In: *New England Journal of Medicine* 385.23 (2021), pp. 2195–2197. DOI: 10.1056/NEJMc2113864. URL: https://doi.org/10.1056/NEJMc2113864.

[33] Anika Singanayagam et al. "Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020". In: *Eurosurveillance* 25.32 (2020), p. 2001483. DOI: doi:https://doi.org/10.2807/1560-7917.ES.2020.25.32.2001483. URL: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.32.2001483.

[34] Caroline Stein et al. "Past SARS-CoV-2 infection protection against re-infection: a systematic review and meta-analysis". In: *The Lancet* 401.10379 (2023), pp. 833–842. DOI: 10.1016/S0140-6736(22)02465-5. URL: https://doi.org/10.1016/S0140-6736(22)02465-5.

[35] Mine Durusu Tanriover et al. "Efficacy and safety of an inactivated whole-virion SARS-CoV-2 vaccine (CoronaVac): interim results of a double-blind, randomised, placebo-controlled, phase 3 trial in Turkey". In: *The Lancet* 398.10296 (2021), pp. 213–222. DOI: 10.1016/S0140-6736(21)01429-X. URL: https://doi.org/10.1016/S0140-6736(21)01429-X.

[36] Stephen J. Thomas et al. "Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine through 6 Months". In: *New England Journal of Medicine* 385.19 (2021), pp. 1761–1773. DOI: 10.1056/NEJMoa2110345. URL: https://doi.org/10.1056/NEJMoa2110345.

[37] UNESCO. *UNESCO Global Monitoring of School Closures Caused by the COVID-19 Pandemic*. Web Page. 2023. URL: https://covid19.uis.unesco.org/global-monitoring-school-closures-covid19/.

[38] Brian J. Willett et al. "SARS-CoV-2 Omicron is an immune escape variant with an altered cell entry pathway". In: *Nature Microbiology* 7.8 (2022), pp. 1161–1179. DOI: 10.1038/s41564-022-01143-7. URL: https://doi.org/10.1038/s41564-022-01143-7.

[39] Yu Wu et al. "Incubation Period of COVID-19 Caused by Unique SARS-CoV-2 Strains: A Systematic Review and Meta-analysis". In: *JAMA Network Open* 5.8 (2022), e2228008–e2228008. DOI: 10.1001/jamanetworkopen.2022.28008. URL: https://doi.org/10.1001/jamanetworkopen.2022.28008.

[40] J. Zhang et al. "Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China". In: *Science* 368.6498 (2020), pp. 1481–1486. DOI: 10.1126/science.abb8001. URL: https://www.ncbi.nlm.nih.gov/pubmed/32350060.

[41] Jun Zhang and Kai F. Yu. "What's the Relative Risk?A Method of Correcting the Odds Ratio in Cohort Studies of Common Outcomes". In: *JAMA* 280.19 (1998), pp. 1690–1691. DOI: 10.1001/jama.280.19.1690. URL: https://doi.org/10.1001/jama.280.19.1690.


Figure S4: Modelled vaccine coverage (lines) against data (dots).

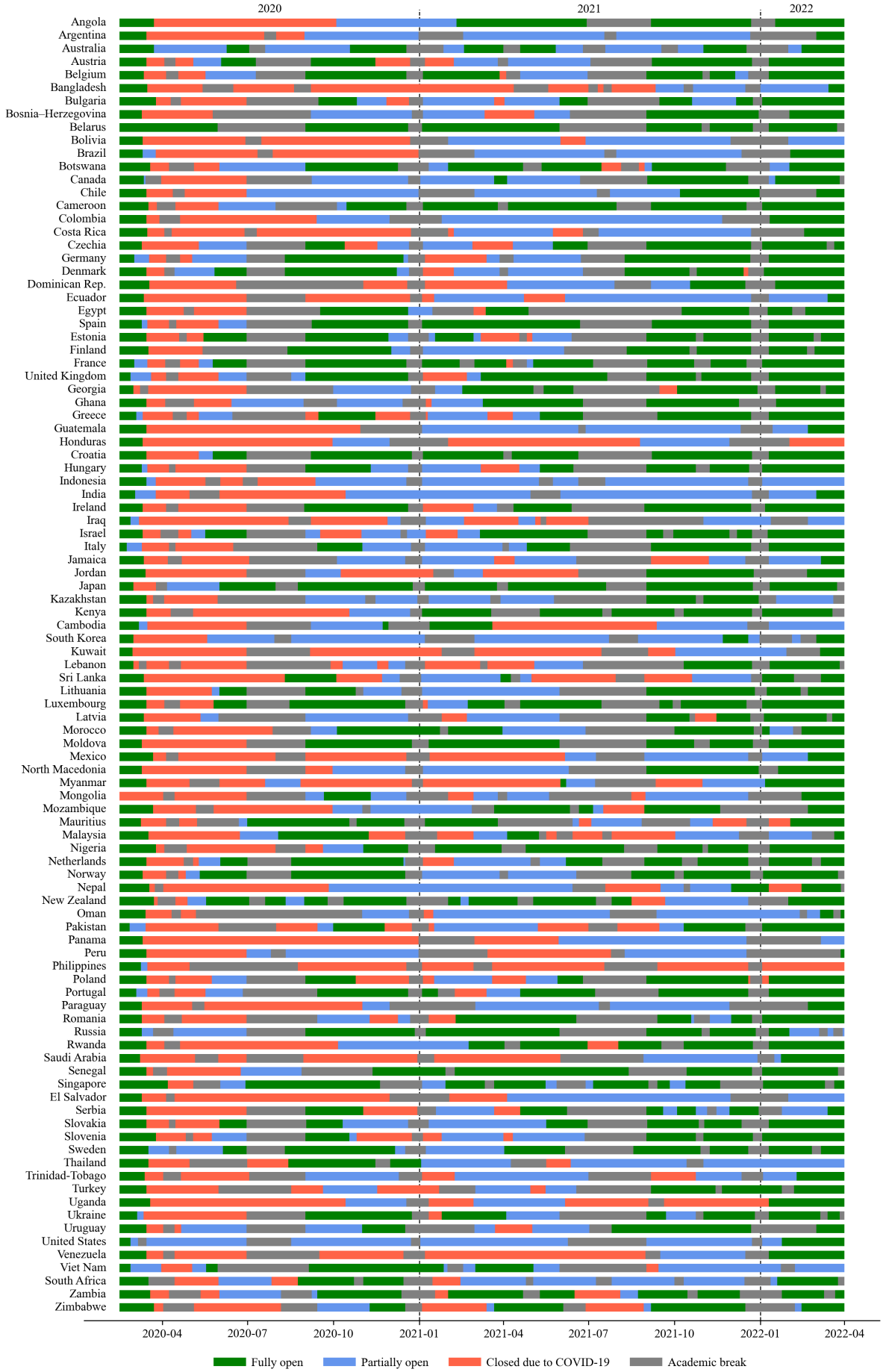




Figure S5: UNESCO data on school closures during the COVID-19 pandemic.

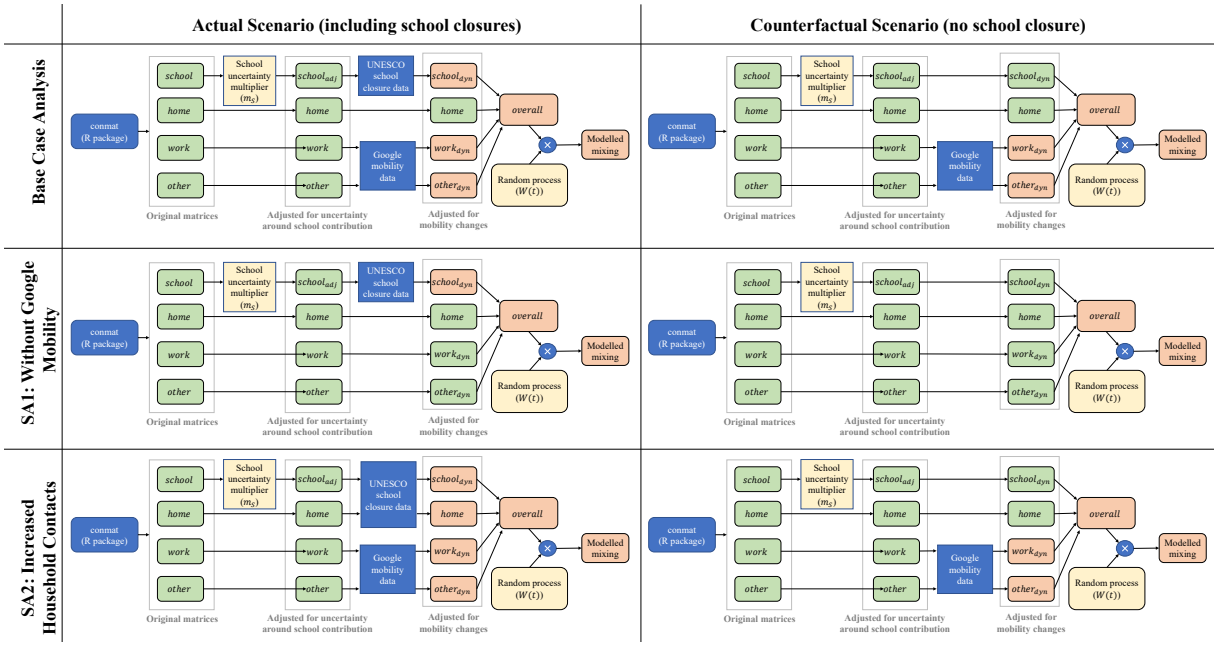


Figure S6: Modelled social mixing under Base Case and Sensitivity Analyses.