Quantifying the direct and indirect components of COVID-19 vaccine effectiveness during the Delta variant era - Supplementary Methods

Anna Suomenrinne-Nordvik, Tuija Leino, Mikhail Shubin, Kari Auranen, Simopekka Vänskä

Contents

S1 Transmission model

S1.1 Age groups and infectious contacts

The population consists of N individuals, stratified into $J = 9$ non-overlapping age groups j,

$$j = \text{ 0–9, 10–19, } \ldots, \text{ 70–79, 80+ years,} \tag{1}$$

with N_j individuals each,

$$\sum_{j=1}^{J} N_j = N.$$

The sizes of age groups N_j are assumed constant during the analysis. Denote by $p = (p_j)$,

$$p_j = \frac{N_j}{N} \tag{2}$$

the proportion of individuals in age group j, $j = 1, \ldots, J$. The age of individuals is not updated and birth/mortality rates are not included in the model, i.e., the age groups remain of constant size.

If $n_{i,j}$ is the average number of daily contacts between age groups i and j, $i, j = 1, \ldots, J$, the contact matrix $C = (C_{i,j})$ is

$$C_{i,j} = \frac{n_{i,j}}{N_j}, \tag{3}$$

where $C_{i,j}$ is the average number of daily contacts [or, rate 1/day] from an individual in age group j into age group i. The contact matrix may depend on time, $C = C(t)$, e.g., by variable distancing measures.

To adjust the contact matrix C with the SARS-CoV-2 transmission, the contacts are adjusted both by the infectiousness ι, $\iota > 0$, of infectious individuals and the susceptibleness σ, $\sigma > 0$,of susceptible individuals (not to be confused with protection by immunity). Thus, for infectious individuals in age group j and susceptible individuals in age group i, the rate of infectious contacts $C^{inf}_{i,j}$ is

$$C^{inf}_{i,j} = \sigma_i C_{i,j} \iota_j. \tag{4}$$

The infectious contact matrix C^{inf} can further be factorized by

$$C^{inf} = \rho \widehat{C^{inf}}, \tag{5}$$

where ρ is the largest eigenvalue of C^{inf}, describing the general level of contacts, and

$$\widehat{C^{inf}} = C^{inf}/\rho \tag{6}$$

is the normalized contact matrix (with respect to the $\|\cdot\|_2$ norm), associated to the distribution of mutual infectious contacts between age groups.

In the Delta era, the changes in the COVID distancing measures considered generally population as a whole. Thus, in this analysis we assumed that the contact matrix is of form

$$C^{inf}(t) = \rho(t)\,\widetilde{C}, \tag{7}$$

where only the level of contacts $\rho = \rho(t)$ depends on time, and the normalized matrix of mutual infectious contacts,

$$\widetilde{C} = (\widehat{f_i C_{i,j} f_j}) \tag{8}$$

is unchanged over time. In addition, the factors ι and σ were assumed the same, $f_j = \iota_j = \sigma_j$, i.e., now the interpretation of the factor f_j is a general correction/adjustment factor for the age group j contributing transmission. Moreover, due to the normalization, the factors f_j can be assumed relative, and the age group j = 40–49 was chosen to be the reference with $f_{40-49} = 1$.

S1.2 SEIR model with differently vaccinated parts of population

Besides age, the population is stratified by the (protective) vaccination status s, indexed by the number of vaccine doses given ($s = 0, 1, 2, 3+$). In a group (age,vaccination status) = (j, s), the risk of acquisition of infection, or transmitting it if acquired, is reduced according to the corresponding components of vaccine efficacy (VE), by

$$r^{acq}_{j,s} = 1 - VE_{acq}(j, s), \quad r^{trm}_{j,s} = 1 - VE_{trm}(j, s). \tag{9}$$

The age-specific coverage of vaccination $cov^{VAC}_{j,s}$, for which national data exist, is assumed to be translated to the coverage of the vaccination status $cov_{j,s}$ (protection of at most status s) with a dose-specific delay

$$cov_{j,s}(t) = cov^{VAC}_{j,s}(t - \text{delay(s)}), \tag{10}$$

where the delay was set 21 days for the first dose, and 7 days for the subsequent doses.

The SEIR model compartments (S for susceptible, E for exposed, I for infectious, R for removed) are defined as proportions within the age group. This means that

$$\sum_s S_{j,s}(t) + E_{j,s}(t) + I_{j,s}(t) + R_{j,s}(t) = 1$$

for every age group j. To get the absolute numbers of individuals, the proportions need to be multiplied by the size of age group (N_j). For example, the absolute number of infectious individuals with age j and vaccine status s at time t is

$$N_j \cdot I_{j,s}(t).$$

Let $V_{j,s}(t)$ be the proportion of individuals with vaccination status s among age group j at time t, i.e.,

$$V_{j,s} = S_{j,s}(t) + E_{j,s}(t) + I_{j,s}(t) + R_{j,s}(t).$$

The proportion $V_{j,s}(t)$ can be expressed by the coverage of vaccination (status) by

$$\begin{cases} V_{j,0}(t) = 1 - cov_{j,1}(t), \\ V_{j,s}(t) = cov_{j,s}(t) - cov_{j,s-1}(t), \quad s \geq 1, \end{cases}$$

from which the evolution equations for the vaccination status follow by taking the time derivative,

$$\begin{cases} V'_{j,0}(t) = -v_{j,0}(t)V_{j,0}(t), \\ V'_{j,s}(t) = v_{j,s-1}(t)V_{j,s-1}(t) - v_{j,s}(t)V_{j,s}(t), \quad s \geq 1, \end{cases} \tag{11}$$

where the vaccination rates are

$$v_{j,s}(t) = \frac{cov'_{j,s+1}(t)}{V_{j,s}(t)}$$

for all s.

The number of new SARS-CoV-2 infections on time t (day) in the (j, s) population is

$$inc_{j,s}(t) = \sum_{j',s'=0}^{J,3} \rho(t)\, S_{j,s}(t) r_{j,s}^{acq}\, \widetilde{C}_{j,j'}\, r_{j',s'}^{trm}(t)\, I_{j',s'}(t) N_{j'}, \tag{12}$$

caused by infectious individuals at time t. In addition, an option for imported infections was implemented.

The SEIR model equations are obtained by combining the above equations (11) and (12),

$$\begin{cases} S'_{j,s}(t) &= -inc_{j,s}(t)/N_j - \phi_{j,s}(t) \\ & \quad +v_{j,s-1}(t)S_{j,s-1}(t) - v_{j,s}(t)S_{j,s}(t), \\ E'_{j,s}(t) &= inc_{j,s}(t)/N_j - \epsilon E_{j,s}(t) + \phi_{j,s}(t) \\ & \quad +v_{j,s-1}(t)E_{j,s-1}(t) - v_{j,s}(t)E_{j,s}(t), \\ I'_{j,s}(t) &= \epsilon E_{j,s}(t) - \gamma I_{j,s}(t) \\ & \quad +v_{j,s-1}(t)I_{j,s-1}(t) - v_{j,s}(t)I_{j,s}(t), \\ R'_{j,s}(t) &= \gamma I_{j,s}(t) \\ & \quad +v_{j,s-1}(t)R_{j,s-1}(t) - v_{j,s}(t)R_{j,s}(t), \end{cases} \tag{13}$$

where $\phi_{j,s}(t)$ is the flow for the imported infections, ϵ is the incubation rate from the exposed to infectious state, γ is the rate from the infectious to removed state. Here, ϵ^{-1}, and γ^{-1}, are the mean duration for the states exposed, and infectious, respectively.

The waning vaccine efficacy against acquisition of Delta infection was implemented by letting the efficacy, or the corresponding risk reduction (9), be time dependent, expressed by a logistic function (14),

$$\begin{cases} r_{j,s}^{acq}(t) = 1 - VE_{acq}(j,s), & t < t_j^w \\ r_{j,s}^{acq}(t) = (1 - VE_{acq}(j,s)) \left(\frac{1}{1+exp((t-t_j^c-t_0)/\nu)} + \frac{VR_{final}}{1+exp(-(t-t_j^c-t_0)/\nu)} \right), & t \geq t_j^w. \end{cases} \tag{14}$$

The age-specific timing of waning onset t_j^w was chosen as the date when 20% second dose coverage was reached in the age group, based on that maximum coverage was reached approximately three months later in all age groups. With values $\nu = 1/20$, $t_0 = 65$ days, and $VR_{final} = 0.65$ the age group specific efficacy reaches roughly 75% of initial efficacy 6 months after t_j^w, and 65% 9 months after, in line with observations [11].

S1.3 The reproduction numbers

By the expression (12) for new infections, and the mean duration γ^{-1} of the infectious state, the next generation matrix is

$$NGM_{i,i'} = \gamma^{-1} \rho(t) S(j,s)(t) r_{j,s}^{acq}(t) \widetilde{C}_{j,j'} r_{j',s'}^{trm}, \tag{15}$$

where all age/vaccination status groups ($j = j(i), s = s(i)$), in total $J \cdot 4$ groups, are ordered in some way with index i. The effective reproduction number, the average number of secondary infections at time t of an average infection under existing immunity in the population, is

$$R_{eff}(t) = \text{The largest eigenvalue of matrix } (NGM_{i,i'}). \tag{16}$$

The corresponding eigenvector, when normalized in $\| \cdot \|_1$ norm, provides the distribution of new/ongoing infections ("average infection") within the population at the time. Because the matrix $\widehat{C}$ here is normalized, the coefficient $\rho = \rho(t)$ in (15) is the daily average number of secondary infection caused by an average infectious individual in a totally susceptible population, i.e., the reproduction number under contacts at time t,

$$R(t) = \gamma^{-1} \rho(t). \tag{17}$$

The function $R(t)$, or $\rho(t)$, is a key parameter describing the strength of the contacts. Under the natural non-restricted contacts, the reproduction number R agrees with the basic reproduction number R_0, which can also depend on time (by season, for example).

The time-dependent $R = R(t)$ is implemented by setting change points at observed inflection points in case numbers. At each change point t_k, $R(t)$ changes smoothly from value R_{k-1} to value R_k by a logistic function with scale δ_k,

$$R(t) = \sum_k \left(\frac{R_{k-1}}{1 + \exp((t - t_k)/\delta_k)} + \frac{R_k}{1 + \exp(-(t - t_k)/\delta_k)} \right). \quad (18)$$

S1.4 Imported infections

S1.4.1 Initial source term at the outbreak of COVID-19

A transmission model needs to be seeded to start emulating the epidemics. One way to initialize the model is to use an appropriate source term, corresponding cases that have been initially imported into the population. When using a source term, the initial state at starting time $t = t_0$ is set to $S(t_0) = 1$, and $E(t_0) = I(t_0) = R(t_0) = 0$.

Figure SM.1 presents the terms of initial imported infection flow $\phi_{j,0}(t)$ used in the SEIR equations (13) (individuals with vaccine doses $s > 0$ did not exist at that time). The shape is motivated by detected traveller cases in the country, increasing with time and reflecting the emerging epidemics in Europe. The age distribution follows the one of detected cases at that time. On the other hand, the travel restrictions were announced at March 16 2020, and a two-week quarantine was required for travellers arriving to the country. Indeed, because the Delta era is long after beginning of COVID-19 epidemics, the implementation of early outbreak is not so crucial for this analysis.

S1.4.2 Imported infections of Delta variant

When the Delta variant initially arrival to Finland, there was an observed uptick in the share of cases registered as acquired abroad. To avoid overestimation of the within population reproduction number, the Delta epidemic was seeded with the same imported infection flow function $\phi_{j,0}(t)$, applied between May 31 and July 11, 2021. The function was scaled so that

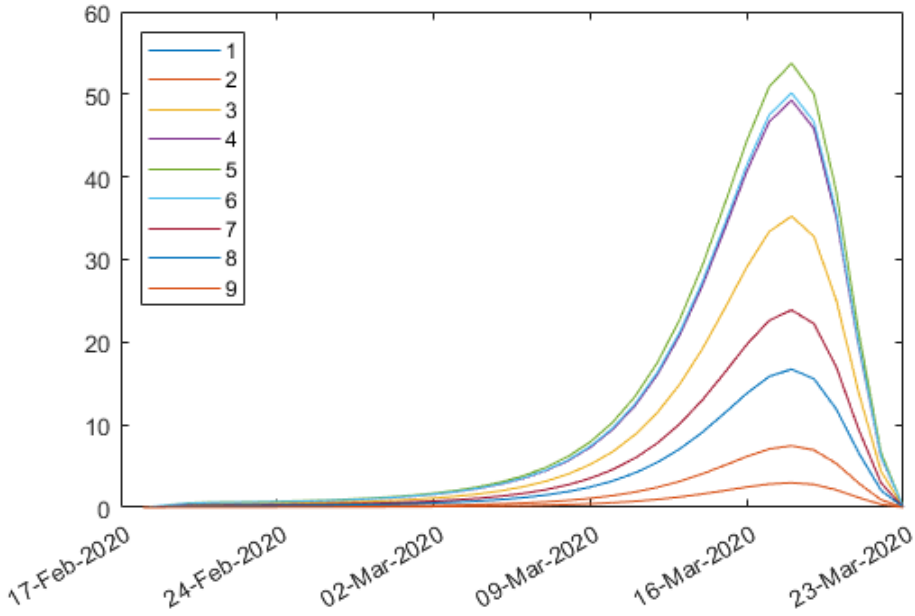

Figure SM.1. The flow of imported SARS-CoV-2 infections to seed the transmission model (terms $\phi_{j,0}(t)$ in the equation (13)).

it matched the observed share of imported cases, as well as the age and vaccination status distribution of all registered cases at the time.

S2 Disease model

S2.1 Risk of disease for an infection

An incident infection cause a disease outcome d among unvaccinated individuals with an age-group j specific probability q_j^d. For individuals with vaccinations status s, the probability is reduced by a factor $r_{j,s}^d$,

$$q_{j,s}^d = r_{j,s}^d q_j^d, \tag{19}$$

according to the end-point-specific vaccine efficacy $VE_d(j,s)$, which can also depend on age and vaccination status. Because the risk reduction for the probability q_j^d is conditional on infection, the compounded reduction of getting infected and subsequently getting the disease d has to match the vaccine efficacy against d,

$$1 - VE_d(j,s) = r_{j,s}^{acq} r_{j,s}^d,$$

or,

$$r_{j,s}^d = \frac{1 - VE_d(j,s)}{r_{j,s}^{acq}} = \frac{1 - VE_d(j,s)}{1 - VE_{acq}(j,s)}. \tag{20}$$

S2.2 Timing of disease

The timing of outcome d (t, time since acquisition of infection) is distributed by end-point-specific delay distribution $f_d^j(t)$,

$$\int_0^\infty f_d^j(t)dt = 1.$$

Thus, the number outcome d at calendar time t in the group (j, s) is, by (12),

$$d_{j,s}(t) = q_{j,s}^d \int_0^t inc_{j,s}(t')f_d^j(t-t')dt'. \tag{21}$$

Timing is in interest when the model is used for monitoring purposes. For the current study, timing has only a limited role. The delays for different disease end-points are listed in Table SM.1.

Table SM.1. Delays for the disease model.

End point	**Delay**
Infection to detected case	3 days
Infection to hospital	Gamma($\mu = 10, \sigma = 4$)
Infection to intensive care	Gamma($\mu = 11, \sigma = 4$)
Infection to death	Gamma($\mu = 10, \sigma = 4$)

S3 Realization and detection model

The transmission and disease model produces the daily values for the SEIR model compartments (13), for the incident infections (12), and for the disease outcomes (21). As the data set is used on the weekly basis, all daily model outcomes are translated into weekly numbers by simply summing up the daily outcomes of the week under consideration.

The deterministic differential equation based transmission and disease model outcomes are rigid compared to the real-life observations. To model variation of both the underlying epidemiological processes and of the detection of the process, the realizations are assumed to follow a Gamma-Poisson mixture model (over dispersion). If m is the (weekly) deterministic model outcome under consideration, the Poisson rate λ is assumed gamma distributed with the mean m and variance σ^2, i.e.,

$$\lambda \sim \text{Gamma}(k, \theta), \quad \text{with } \theta = \sigma^2/m, \quad k = m^2/\sigma^2. \tag{22}$$

The variance is modelled by

$$\sigma^2 = s_1^2 m^2 + s_2^2 m, \tag{23}$$

Table SM.2. Model parameters that are estimated from the monitoring data set by using the COVID-19 infectious disease model, and their prior distributions.

Parameter	Parameter name	Prior
R_i	Reproduction number at time t_i	$Gamma(\mu = 2, \sigma = 1.4)$
		$\lvert\frac{d^2R}{dt^2}(t_i)\rvert \sim \text{Exp}(10)$ for regularity
f_a	Adjustment factor for contact	$f_a > 0$ flat
q_a^H	Infection to hospitalization probability	$LogitNormal(\mu_a, \sigma_a^2 = 2.5^2)$,
		where $\mu_j = \frac{10\cdot(j-1)+5}{11} - 10$,
s_1, s_2	Realization hyperparameters	$s_j > 0$ flat

where s_1 and s_2 are the realization (hyper-)parameters. Using format (23), the probability distribution function of the compounded Gamma-Poisson mixture is

$$\begin{aligned} f_{GP}(n|m,s) &= \int_0^\infty f_{poisson}(n|\lambda) f_{gamma}(\lambda|k,\theta) d\lambda \\ &= (1-q)^n q^k \frac{\Gamma(n+k)}{\Gamma(k)\Gamma(n+1)}, \end{aligned} \tag{24}$$

where

$$k = \frac{m}{s_1^2 m + s_2^2}, \quad q = \frac{1}{1+\theta} = \frac{1}{1+s_1^2 m + s_2^2}.$$

Indeed, if $N \sim f_{GP}(\,\cdot\,|m,s)$, then $N|\lambda \sim Pois(\lambda)$, and so by the total expectation,

$$E(N) = E_\lambda(E(N|\lambda)) = E_\lambda(\lambda) \overset{(22)}{=} m, \tag{25}$$

and by the total variance,

$$\begin{aligned} Var(N) &= E_\lambda(Var(N|\lambda)) + Var_\lambda(E(N|\lambda)) \\ &= E_\lambda(\lambda) + Var_\lambda(\lambda) \\ &\overset{(22)}{=} m + \sigma^2. \end{aligned} \tag{26}$$

S4 Parameter estimation with the COVID-19 model

The model parameters that are estimated from the monitoring data by using the COVID-19 infectious disease model, and their prior distributions, are listed in Table SM.2. At this point, all other model parameters are considered to be given, either by assumptions, or by separate estimates. Let

$$X = (X_i) \tag{27}$$

be the vector of all model parameters to be estimated, including the hyper parameters s_1 and s_2 for the realization. Data

$$Z(d,k,a,t) = n(d,k,a,t) \tag{28}$$

about d =detected cases, and d =hospitalizations, by age group (a), vaccination status (k), and week (t), were used in estimating parameters X. By limited COVID-testing in the early epidemics, the use of detected cases was limited to the weeks from July 2020 on.

Let

$$Y(d,k,a,t) = n(d,k,a,t) \tag{29}$$

be the deterministic model outcomes with a parameter set X. Here, the values of $Y(d,k,a,t)$ are not necessary integers. By the realization model (24), the likelihood is

$$F_{likelihood}(Z|X) = \prod_{d,k,a,t} f_{GP}(Z(d,k,a,t)|Y(d,k,a,t),s), \tag{30}$$

where the realizations $Z(d,k,a,t)$ with the parameter set X are assumed mutually independent. The prior distribution is

$$F_{prior}(X) = \prod_i f_i(X_i), \tag{31}$$

where f_i is the prior distribution for the parameter X_i, Table SM.2. Eventually, the posterior density is

$$F_{posterior}(X|Z) = \frac{F(X,Z)}{F(Z)} \propto F_{prior}(X)F_{likelihood}(Z|X). \tag{32}$$

The posterior distribution was explored by drawing numerical samples

$$X(1), X(2), \ldots \tag{33}$$

with an adaptive Metropolis-Hastings algorithm. In the adaptive proposal density, the covariance matrix C,

$$C = C_n \cdot Cov(X(k), \ldots, X(k+K)) + \delta I,$$

was obtained by weighting the sample covariance of $K = 5000$ latest sample points by $C_n = 2.4/\sqrt{n}$, where n is the number of unknown parameters (dimension of X), and the regulating parameter was $\delta = 10^{-6}$. The sample size in (33) was 50 000, after the 90 000 burn-in.

S5 The posterior predictive distributions

Using the sample of the posterior distribution (33), two kinds of posterior predictive distributions (PPD) are computed. First, the PPD for the

deterministic model outcomes,

$$Y(d,k,a,t) \sim F(Y(d,k,a,t)|X)F_{posterior}(X), \tag{34}$$

is used to report the model outcomes for the unvaccinated Control Population. The PPD for the realization,

$$\begin{aligned} W(d,k,a,t) \sim & f_{GP}(W(d,k,a,t)|Y(d,k,a,t),s) \\ & \cdot F(Y(d,k,a,t)|X)F_{posterior}(X), \end{aligned} \tag{35}$$

is used to illustrate the model fit with the real-life monitoring data.

For both cases above, the dimension of the PPD is high ($\#d \cdot \#k \cdot \#a \cdot \#t = 10152$), so that it is not economic to keep all 50 000 simulation results of the weekly age-specific quantities in the computer memory. Instead, for each weekly age-specific quantity, the sample means

$$\hat{x}_n = \frac{1}{n}\sum_{k=1}^{n} x_k,$$

and variances

$$\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{k}(x_k - \hat{x}_n)^2$$

are updated recursively after each sample run by

$$\hat{x}_{n+1} = \frac{1}{n+1}(n\hat{x}_n + x_{n+1}), \tag{36}$$

$$\hat{\sigma}_{n+1}^2 = \frac{n-1}{n}\hat{\sigma}_n^2 + \frac{1}{n+1}(x_{n+1} - \hat{x}_n)^2. \tag{37}$$

The sample statistics can be directly used for the PPD of the deterministic model outcomes $Y(d,k,a,t)$, which do not depend on the hyperparameters $s = (s_1, s_2)$. For the PPD for the realization, by the total expectation, (in the following formulae, we use the shorthand notations $W = W(d,k,a,t)$ and $Y = Y(d,k,a,t)$)

$$E(W) = E_{Y,S}(E(W|Y,S)) \overset{(25)}{=} E_{Y,S}(Y) = E_Y(Y), \tag{38}$$

which is approximated by the sample mean $\hat{Y} = \hat{Y}(d,k,a,t)$. By the total variance

$$\begin{aligned} Var(W) &= E_{Y,S}(Var(W|Y,S)) + Var_{Y,S}(E(W|Y,S)) \\ \text{By (23), (25),(26)} \quad &= E_{Y,S}(Y + S_1^2Y^2 + S_2^2Y) + Var_{Y,S}(Y) \\ &= E_Y(Y) + E_{Y,S_1}(S_1^2Y^2) + E_{Y,S_2}(S_2^2Y) + Var_Y(Y), \end{aligned}$$

and because by the total expectation and the definition of variance,

$$\begin{aligned} E(S_1^2Y^2) &= E(E(S_1^2Y^2|Y)) = E(Y^2E(S_1^2|Y)) = E(Y^2)E(S_1^2) \\ &= (var(Y) + E(Y)^2)(var(S_1) + E(S_1)^2), \end{aligned}$$

and

$$\begin{aligned} E(S_2^2 Y) &= E(E(S_2^2 Y|Y)) = E(Y E(S_2^2|Y)) = E(Y)E(S_2^2) \\ &= E(Y)(var(S_2) + E(S_2)^2), \end{aligned}$$

and so,

$$\begin{aligned} Var(W) =& E(Y)(1 + var(S_2) + E(S_2)^2) \\ &+ var(Y)(1 + var(S_1) + E(S_1)^2) \\ &+ E(Y)^2(var(S_1) + E(S_1)^2). \end{aligned} \tag{39}$$

In this expression, all terms can be approximated by the sample statistics for $Y = Y(d, k, a, t)$, S_1, and S_2.

To approximate the predictive intervals, after the simulation is finished, the sample statistics for $Y = Y(d, k, a, t)$, or for $W = W(d, k, a, t)$, are used to determine the corresponding Gamma distribution from which the predictive intervals are determined.

S6 Parameters of the COVID-19 model, summary

A summary of the COVID-19 model parameters and their sources are given in Table SM.3. The means and 90% credible intervals of the posterior distributions of the parameters estimated with the model are given Tables SM.4 and SM.5. Table SM.6 contains the separately estimated age-dependent risks for ICU admission per hospitalization and death per case.

Table SM.3. A summary of the COVID-19 infectious disease model parameters and their sources.

Notation	Explanation	Source
Transmission model parameters		
$C(a, a')$	Contact matrix	[2]
ϵ	$\epsilon^{-1} = 3$ days, the average duration of E state Time in E is incubation time (5) minus the time of presymptomatic transmission (2)	[3–7]
γ	$\gamma^{-1} = 7$ days, the average duration of I state Time in I is time of presymptomatic transmission +5 days after symptoms onset (shortened from 7-10 days due to self-isolation)	[3, 4]
f_a	Age-specific adjustment factor for contacts $f_{40-49} = 1$ fixed	Estimated
R_i	The reproduction number R after the change point t_i Two first reproduction numbers are kept fixed: $R_1 = 1.9$, t_1 =16.2.2020; $R_2 = 0.745$, t_2 =20.3.2020	Estimated
$\mathrm{VE}_{acq}(a, k)$	Age-specific vaccine efficacy against infection in age group a with k doses $\mathrm{VE}_{acq}(a, 1) = 0.5/0.5/0.5$ (baseline/higher/lower) $\mathrm{VE}_{acq}(a, 2+) = 0.7/0.8/0.6$ (baseline/higher/lower) A reduction in efficacy against infection of 2+ doses was applied for ages 70-79 (25%) and 80+ (50%).	[8–13]
$\mathrm{VE}_{trm}(a, k)$	Age-specific vaccine efficacy against infectiousness given infection in age group a with k doses $\mathrm{VE}_{trm}(a, 1) = 0.5/0.5/0.5$ (baseline/higher/lower) $\mathrm{VE}_{trm}(a, 2+) = 0.5/0.6/0.625$ (baseline/higher/lower)	[8–13]
$\mathrm{VE}_{sd}(a, k)$	Age-specific vaccine efficacy against severe disease (hospitalization, ICU admission, death) given infection in age group a with k doses $\mathrm{VE}_{sd}(a, 1) = 0.5/0.5/0.5$ (baseline/higher/lower) $\mathrm{VE}_{sd}(a, 2+) = 0.5/0.7/0.625$ (baseline/higher/lower)	[8–13]
Disease model parameters		
r_{covid}	Detection probability (infection-to-case rate) Baseline $r_{covid} = 0.75$ Sensitivity analysis 0.50/0.90 for low/high detection	Unknown
q_a^H	Infection to hospitalization probability	Estimated
q_a^{ICU}	Infection to ICU probability $q_a^{ICU} = q_a^H \, c_a^{ICU}$, c_a^{ICU} estimated hospitalization to ICU probability	Separate estimate
q_a^{death}	Infection to death probability $q_a^{death} = r_{covid} \, c_a^{death}$ c_a^{death} estimated case to death probability	Separate estimate
Observation model parameters		
s_1, s_2	Scale parameters for the scale s of Gamma distribution Scale $s = s_1 m + s_2 \sqrt{m}$ (m is the incidence rate)	Estimated

Table SM.4. Posterior mean and 90% credible intervals for the reproduction numbers and observation model parameters estimated from the monitoring data set by using the COVID-19 infectious disease model.

Parameter	Baseline VE, baseline detection[a]	Baseline VE, lower detection	Baseline VE, higher detection	Lower VE, baseline detection	Higher VE, baseline detection
R_3	0.232	0.395	0.214	0.234	0.227
t_3 =13.5.2020	(0.136 - 0.337)	(0.276 - 0.511)	(0.139 - 0.3)	(0.146 - 0.323)	(0.145 - 0.312)
R_4	0.486	0.47	0.457	0.491	0.48
t_4 =1.6.2020	(0.41 - 0.564)	(0.395 - 0.554)	(0.393 - 0.518)	(0.417 - 0.566)	(0.411 - 0.549)
R_5	1.322	1.322	1.321	1.315	1.324
t_5 =4.7.2020	(1.297 - 1.346)	(1.298 - 1.345)	(1.302 - 1.342)	(1.291 - 1.339)	(1.302 - 1.347)
R_6	1.19	1.192	1.192	1.192	1.19
t_6 =1.10.2020	(1.171 - 1.21)	(1.172 - 1.212)	(1.171 - 1.212)	(1.172 - 1.212)	(1.169 - 1.211)
R_7	0.881	0.885	0.879	0.882	0.881
t_7 =27.11.2020	(0.865 - 0.897)	(0.868 - 0.901)	(0.861 - 0.897)	(0.865 - 0.899)	(0.863 - 0.9)
R_8	1.165	1.174	1.161	1.162	1.166
t_8 =25.1.2021	(1.141 - 1.188)	(1.152 - 1.196)	(1.139 - 1.183)	(1.138 - 1.187)	(1.141 - 1.189)
R_9	0.828	0.836	0.827	0.835	0.827
t_9 =15.3.2021	(0.803 - 0.855)	(0.81 - 0.862)	(0.802 - 0.851)	(0.807 - 0.863)	(0.803 - 0.853)
R_{10}	0.75	0.763	0.746	0.748	0.757
t_{10} =22.4.2021	(0.735 - 0.764)	(0.747 - 0.779)	(0.73 - 0.761)	(0.733 - 0.764)	(0.743 - 0.772)
R_{11}	1.88	1.902	1.872	1.849	1.937
t_{11} =25.6.2021	(1.843 - 1.917)	(1.863 - 1.943)	(1.833 - 1.913)	(1.812 - 1.887)	(1.897 - 1.977)
R_{12}	1.784	1.82	1.777	1.808	1.882
t_{12} =1.8.2021	(1.75 - 1.817)	(1.786 - 1.853)	(1.745 - 1.808)	(1.77 - 1.847)	(1.849 - 1.915)
R_{13}	2.413	2.476	2.394	2.428	2.673
t_{13} =20.9.2021	(2.368 - 2.459)	(2.428 - 2.523)	(2.353 - 2.437)	(2.379 - 2.476)	(2.624 - 2.724)
R_{14}	3.024	3.128	3.005	3.015	3.469
t_{14} =20.10.2021	(2.972 - 3.074)	(3.073 - 3.182)	(2.955 - 3.053)	(2.955 - 3.075)	(3.406 - 3.532)
s_1	0.343	0.344	0.345	0.378	0.341
	(0.332 - 0.354)	(0.333 - 0.356)	(0.333 - 0.358)	(0.366 - 0.391)	(0.329 - 0.353)
s_2	0.899	0.903	0.913	0.865	0.868
	(0.824 - 0.957)	(0.842 - 0.949)	(0.842 - 0.961)	(0.809 - 0.908)	(0.808 - 0.917)

[a]Baseline scenario: 0.75 detection probability, vaccine efficacy of 2+ doses (VE) 0.7 against infection, 0.85 against infectiousness, 0.85 against severe disease. Sensitivity analysis scenarios: Lower detection probability 0.5, higher detection 0.9. Lower VE: 0.6 against infection, 0.85 against infectiousness, 0.85 against severe disease. Higher VE: 0.8 against infection, 0.94 against infectiousness, 0.92 against severe disease.

Table SM.5. Posterior mean and 90% credible intervals for age-specific parameters that are estimated from the monitoring data set by using the COVID-19 infectious disease model. Note that $f_{40-49} = 1$ is fixed.

Parameter	Baseline VE, baseline detection[a]	Baseline VE, lower detection	Baseline VE, higher detection	Lower VE, baseline detection	Higher VE, baseline detection
f_{0-9}	0.606 (0.541 - 0.676)	0.605 (0.54 - 0.676)	0.59 (0.532 - 0.655)	0.613 (0.537 - 0.696)	0.538 (0.48 - 0.601)
f_{10-19}	1.532 (1.403 - 1.669)	1.537 (1.414 - 1.672)	1.496 (1.378 - 1.626)	1.507 (1.382 - 1.646)	1.375 (1.268 - 1.489)
f_{20-29}	1.237 (1.152 - 1.325)	1.245 (1.164 - 1.331)	1.223 (1.148 - 1.304)	1.226 (1.148 - 1.313)	1.172 (1.099 - 1.248)
f_{30-39}	1.159 (1.058 - 1.273)	1.16 (1.06 - 1.269)	1.145 (1.056 - 1.241)	1.13 (1.032 - 1.236)	1.128 (1.035 - 1.224)
f_{50-59}	0.957 (0.861 - 1.058)	0.954 (0.859 - 1.056)	0.925 (0.837 - 1.021)	0.923 (0.824 - 1.033)	0.88 (0.793 - 0.976)
f_{60-69}	0.767 (0.693 - 0.846)	0.774 (0.703 - 0.851)	0.762 (0.698 - 0.834)	0.744 (0.67 - 0.824)	0.735 (0.668 - 0.808)
f_{70-79}	2.097 (1.876 - 2.341)	2.097 (1.884 - 2.328)	2.066 (1.876 - 2.279)	2.099 (1.881 - 2.344)	1.951 (1.755 - 2.171)
f_{80+}	1.83 (1.648 - 2.043)	1.77 (1.597 - 1.965)	1.779 (1.62 - 1.953)	1.821 (1.619 - 2.058)	1.649 (1.468 - 1.841)
q^H_{0-9}	0.004 (0.004 - 0.005)	0.003 (0.003 - 0.004)	0.005 (0.005 - 0.006)	0.004 (0.004 - 0.005)	0.004 (0.003 - 0.005)
q^H_{10-19}	0.003 (0.002 - 0.003)	0.002 (0.002 - 0.002)	0.003 (0.003 - 0.004)	0.003 (0.002 - 0.003)	0.003 (0.002 - 0.003)
q^H_{20-29}	0.008 (0.007 - 0.009)	0.005 (0.005 - 0.006)	0.009 (0.008 - 0.01)	0.008 (0.007 - 0.009)	0.008 (0.007 - 0.009)
q^H_{30-39}	0.016 (0.014 - 0.017)	0.011 (0.01 - 0.012)	0.018 (0.017 - 0.02)	0.016 (0.015 - 0.018)	0.016 (0.014 - 0.017)
q^H_{40-49}	0.028 (0.025 - 0.03)	0.019 (0.018 - 0.021)	0.032 (0.03 - 0.035)	0.028 (0.026 - 0.031)	0.028 (0.026 - 0.03)
q^H_{50-59}	0.057 (0.052 - 0.061)	0.04 (0.037 - 0.043)	0.066 (0.061 - 0.072)	0.059 (0.054 - 0.064)	0.061 (0.057 - 0.066)
q^H_{60-69}	0.097 (0.089 - 0.105)	0.068 (0.063 - 0.074)	0.113 (0.104 - 0.123)	0.101 (0.092 - 0.111)	0.105 (0.097 - 0.113)
q^H_{70-79}	0.174 (0.159 - 0.189)	0.122 (0.112 - 0.132)	0.203 (0.185 - 0.222)	0.184 (0.169 - 0.201)	0.2 (0.181 - 0.22)
q^H_{80+}	0.18 (0.164 - 0.197)	0.131 (0.119 - 0.144)	0.215 (0.196 - 0.237)	0.202 (0.182 - 0.223)	0.212 (0.194 - 0.231)

[a]Baseline scenario: 0.75 detection probability, vaccine efficacy of 2+ doses (VE) 0.7 against infection, 0.85 against infectiousness, 0.85 against severe disease. Sensitivity analysis scenarios: Lower detection probability 0.5, higher detection 0.9. Lower VE: 0.6 against infection, 0.85 against infectiousness, 0.85 against severe disease. Higher VE: 0.8 against infection, 0.94 against infectiousness, 0.92 against severe disease.

Table SM.6. The age-specific disease model parameters estimated separately using register data from the beginning of 2021 to the end of the study period.

Age group a	**Hospitalization to ICU probability** c_a^{ICU}	**Case to death probability** c_a^{death}
$0-9$	0.0194	0
$10-19$	0.0714	0.0001
$20-29$	0.133	0
$30-39$	0.1827	0.0004
$40-49$	0.2071	0.0012
$50-59$	0.2618	0.0043
$60-69$	0.3206	0.0164
$70-79$	0.2650	0.0812
$80+$	0.0625	0.2269

References

1. Poukka E, *et al.* (2022) Cohort study of Covid-19 vaccine effectiveness among healthcare workers in Finland, December 2020-October 2021. *Vaccine*, 40:701–705.

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

3. He X, *et al.* (2020) Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature medicine*, 26:672–675.

4. Ashcroft P, *et al.* (2020) COVID-19 infectivity profile correction. *arXiv preprint arXiv:2007.06602*.

5. Lauer SA, *et al.* (2020) The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine*, 172:577–582.

6. Li Q, *et al.* (2020) Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England journal of medicine*, 382:1199–1207.

7. Nishiura H, *et al.* (2020) Serial interval of novel coronavirus (COVID-19) infections. *International journal of infectious diseases*, 93:284–286.

8. Bernal JL, *et al.* (2021) Effectiveness of Covid-19 vaccines against the B. 1.617. 2 (Delta) variant. *The New England journal of medicine*.

9. Baum U, *et al.* (2021) Effectiveness of vaccination against SARS-CoV-2 infection and Covid-19 hospitalisation among Finnish elderly and chronically ill—An interim analysis of a nationwide cohort study. *PLoS One*, 16:e0258 704.

10. Andrews N, *et al.* (2022) Duration of protection against mild and severe disease by Covid-19 vaccines. *New England Journal of Medicine*, 386:340–350.

11. Poukka E, *et al.* (2022) Cohort study of Covid-19 vaccine effectiveness among healthcare workers in Finland, December 2020-October 2021. *Vaccine*, 40:701–705.

12. Baum U., *et al.* (2022) High vaccine effectiveness against severe COVID-19 in the elderly in Finland before and after the emergence of Omicron. *BMC infectious diseases*, 22:816.

13. Hall VJ., *et al.* (2021) COVID-19 vaccine coverage in health-care workers in England and effectiveness of BNT162b2 mRNA vaccine against infection (SIREN): a prospective, multicentre, cohort study. *The Lancet*, 397:1725–1735.