The probability that a previously susceptible individual is infected as a function of time after exposure to SARS-CoV-2

Alun Thomas [*†] Karim Khader [*] Adam Hersh [‡]
Matthew Samore [*]

December 16, 2022

Keywords: Isolation, quarantine, Gamma distribution.

Running title: Probability of infectiousness following exposure.

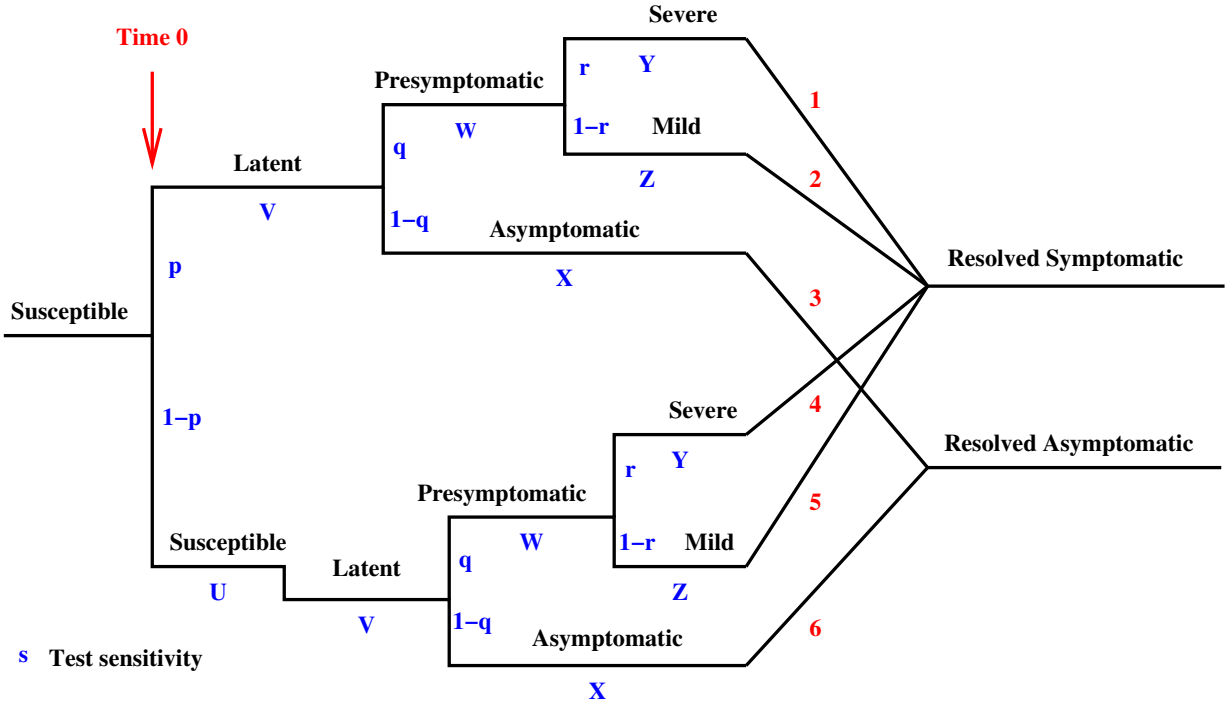
Abstract

We consider an individual who, following a specific exposure to SARS-CoV-2 , is symptom free and has a negative test and calculate the probability that they are in fact in the latent state or infectious at the time of the test. We parameterize the probability of actual infection as a consequence of the exposure and, in particular, compare the certainly infected and certainly uninfected cases. The results shed light on the questions of how long to isolate and when to test an individual following an exposure event.

[*]Division of Epidemiology, Department of Internal Medicine, University of Utah
[†]Corresponding author
[‡]Pediatric Infectious Diseases, University of Utah

1 Methods

Figure 1: Model for the course of infection. Letters representing the random variables specifying the time in each state and branching probabilities are shown in blue. The red numbers enumerate the possible paths.



We consider a susceptible individual who has a specific exposure to SARS-CoV-2 at time 0. They may or may not become infected as a consequence of the exposure. If they are infected, they enter the latent state and progress through the stages of the infection until it is resolved. If they are not infected they remain susceptible and are subject to a constant background risk of further exposure. At some time x after the exposure, the individual who has shown no symptoms up to that time has a negative SARS-CoV-2 test. We want the probability that at the time of the test, the individual is in fact infected.

Figure 1 shows the 6 possible paths for the progression of the disease with the states labeled. Also shown are the random variables representing the time in each state, and the branching probabilities.

In the latent state, we assume that any SARS-CoV-2 test will necessarily be negative. Otherwise, we assume complete specificity, but parameterize the sensitivity to allow for imperfect detection.

1.1 Parameter values for the original variants

For the most part, we chose default values for the path splitting parameters and the Gamma shape and rate parameters to approximately match the values given in table 1 of Eikenberry et al. (2020). Tables 1 and 2 give details of these parameters. The columns marked *Lower* and *Upper* in table 1 give the lower and upper 2.5% quantiles of the these Gamma distribution and roughly correspond to the *Likely range* column given by Eikenberry et al. (2020). The values for the time in the susceptible state were arbitrarily chosen so that the background risk of infection is 1 per 1000 per day. The current daily risk in the US is approximately 1 per 10000 per day, but we use the higher rate to better reflect the risks to active individuals, and allow for under reporting of cases. Although the default value of parameter p, the probability of infection following the exposure event is set by default to 100% corresponding to certain infection, we vary this in the analyses below.

Note that the model used by Eikenberry et al. (2020) does not have a *presymptomatic* phase. We assumed that this had a mean of 3 days with likely range of $(1.0, 5.0)$ and reduced the times in the symptomatic phases accordingly.

Table 1: Gamma time-in-state distribution parameters and summary values for original SARS-CoV-2 variants.

Parameter	Name	Shape	Rate	Mean	Variance	Lower	Upper
U	Sus	1.00	0.001	1000.00	1000000.00	25.32	3688.88
V	Lat	3.47	0.680	5.10	7.50	1.22	11.70
W	Asy	9.00	3.000	3.00	1.00	1.37	5.25
X	Pre	6.53	0.933	7.00	7.50	2.71	13.30
Y	Mld	2.66	0.242	11.00	45.50	1.97	27.62
Z	Svr	0.55	0.110	5.00	45.50	0.01	24.06

Table 2: Path probability parameters for original SARS-CoV-2 variants.

Parameter	Name	Value
q	Symptomatic	0.50
r	Svrere	0.10

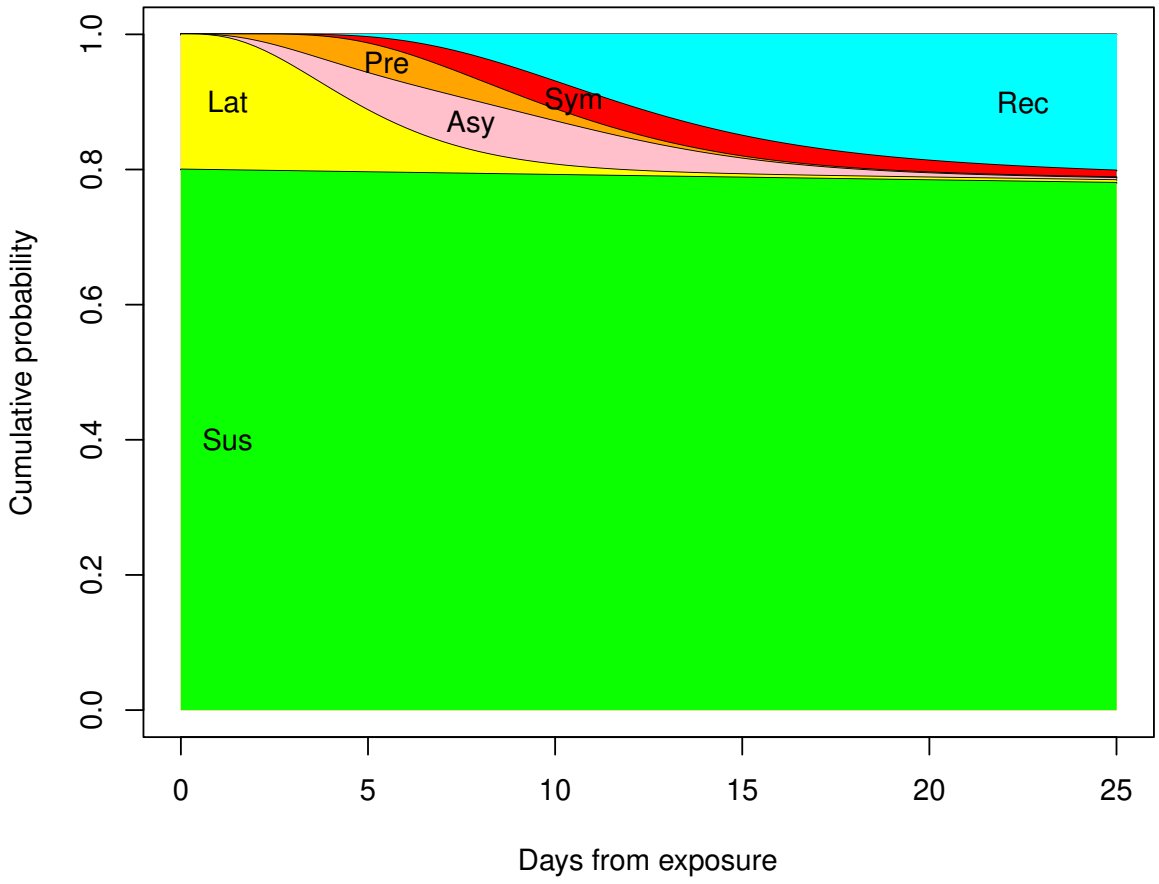
1.2 Modifying the parameters to match the Omicron variant

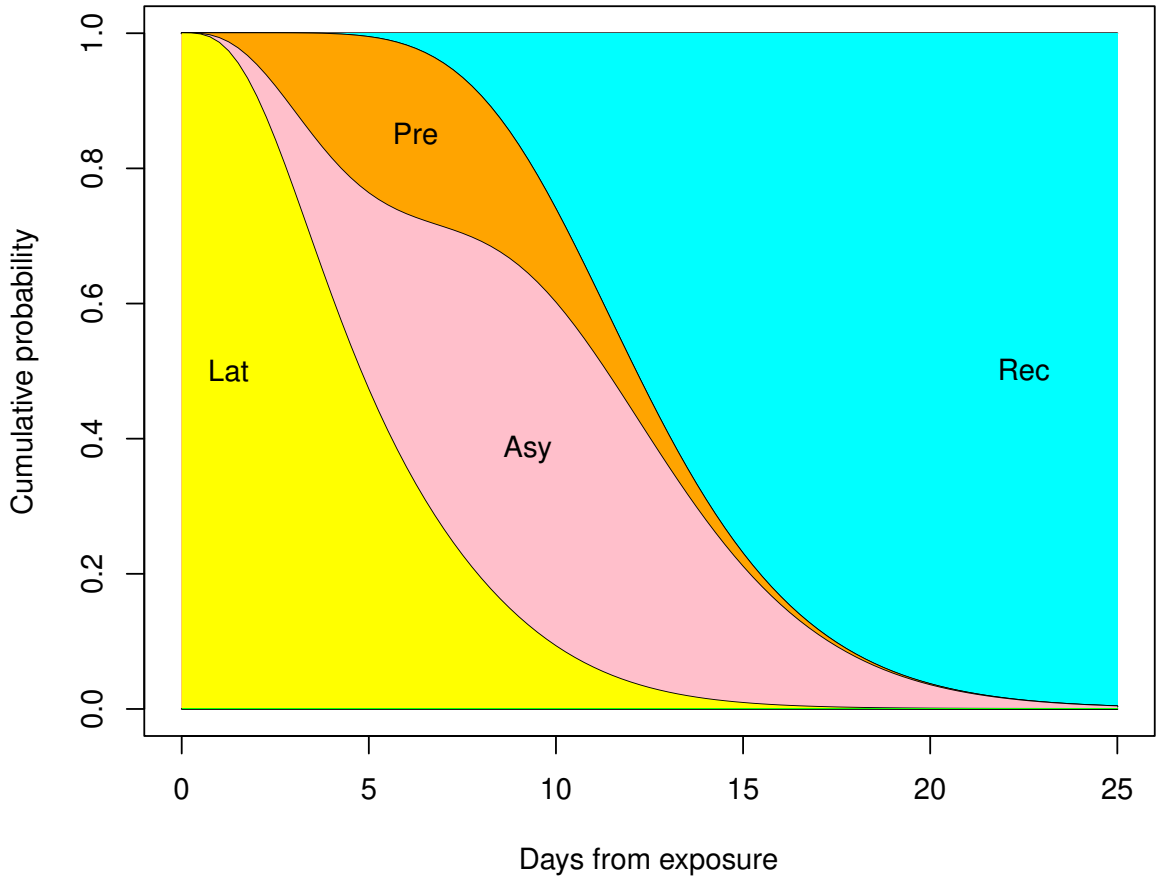
Jansen et al. (2021) estimate that the latent period for individuals infected by the later Omicron variants are roughly 60% of the those with the original variants. Symptomatic periods are estimated to be reduced by a factor of 90% (Menni et al. 2022), and we have assumed that both the presymptomatic and asymptomatic periods are similarly reduced. These changes have been made by reducing both the means and standard deviations of the Gamma distributions by the appropriate factor. Table 3 gives details of these Gamma distributions. The path probability parameters have been left unchanged.

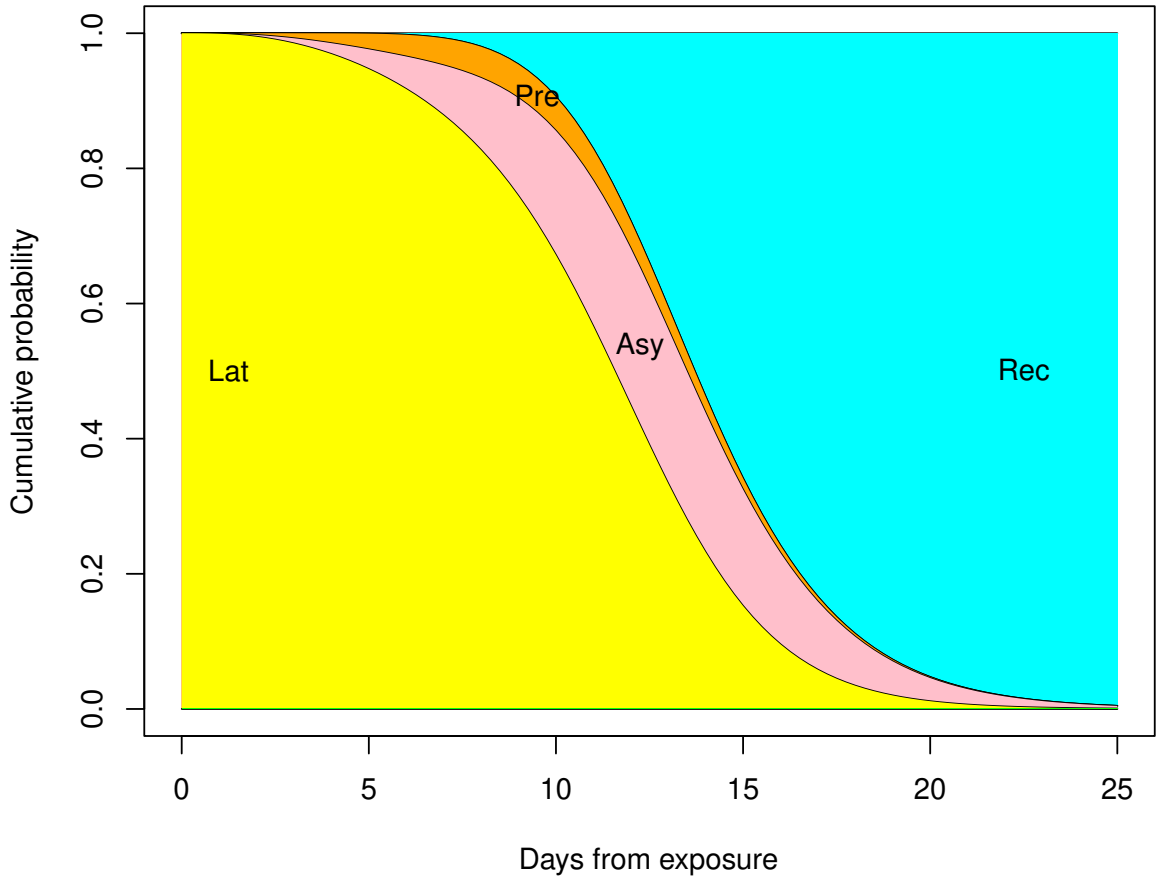
Table 3: Gamma time-in-state distribution parameters and summary values for more recent SARS-CoV-2 variants.

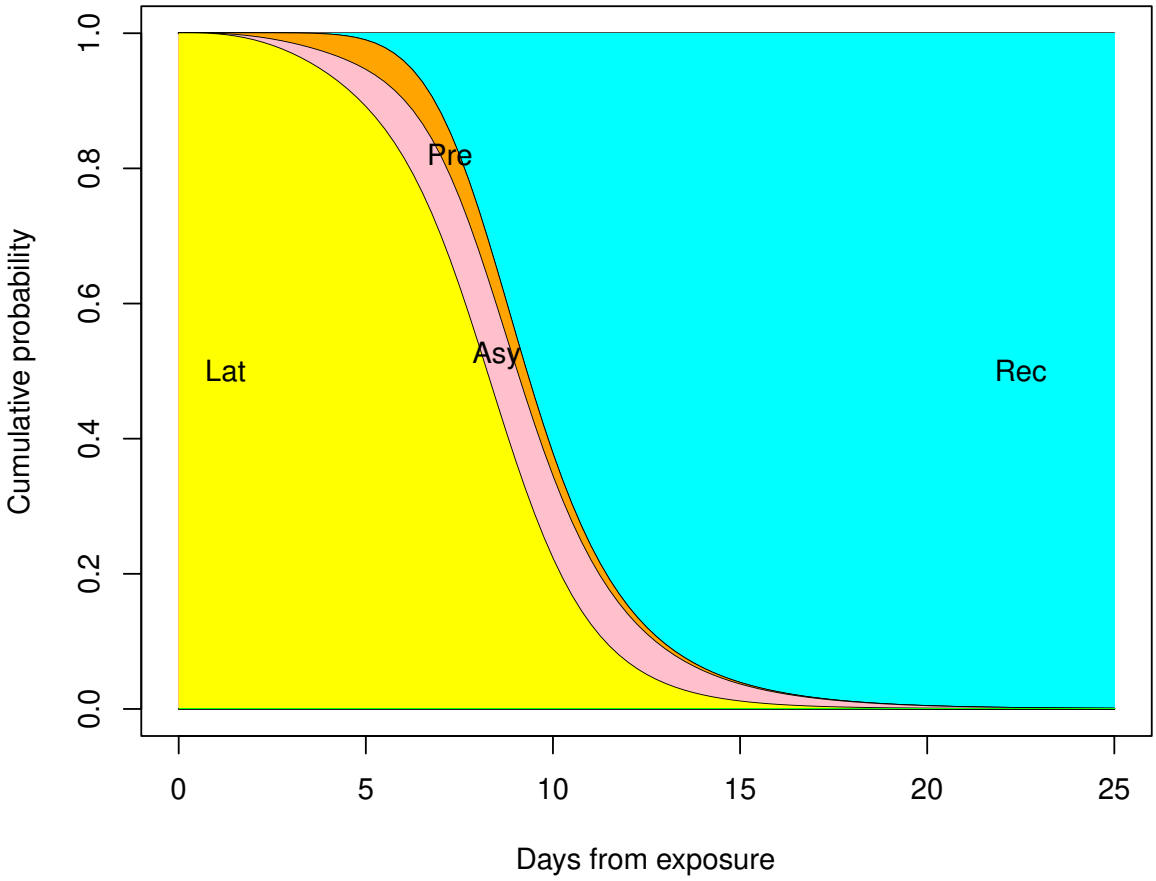
Parameter	Name	Shape	Rate	Mean	Variance	Lower	Upper
U	Sus	1.00	0.001	1000.00	1000000.00	25.32	3688.88
V	Lat	3.47	1.133	3.06	2.70	0.73	7.02
W	Asy	9.00	3.333	2.70	0.81	1.23	4.73
X	Pre	6.53	1.037	6.30	6.08	2.43	11.97
Y	Mld	2.66	0.269	9.90	36.86	1.78	24.86
Z	Svr	0.55	0.122	4.50	36.86	0.01	21.66

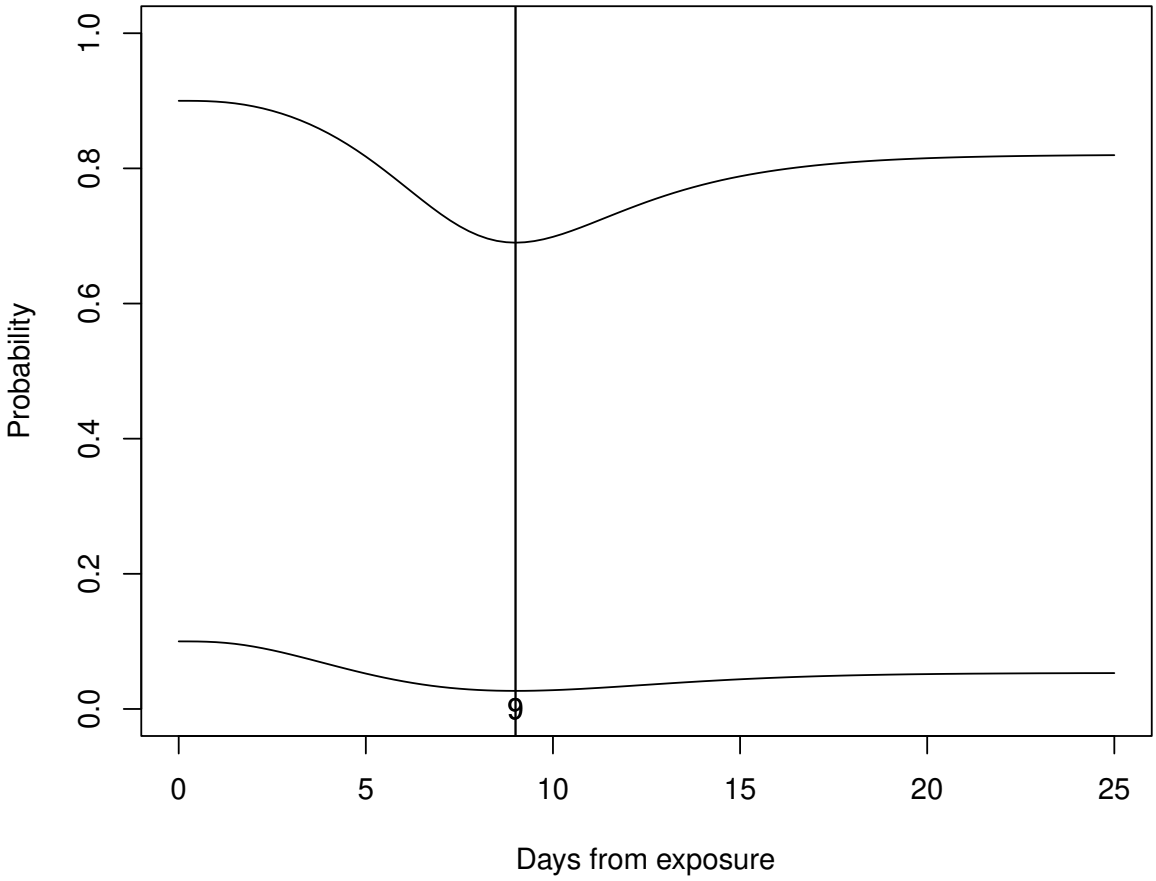
2 Results

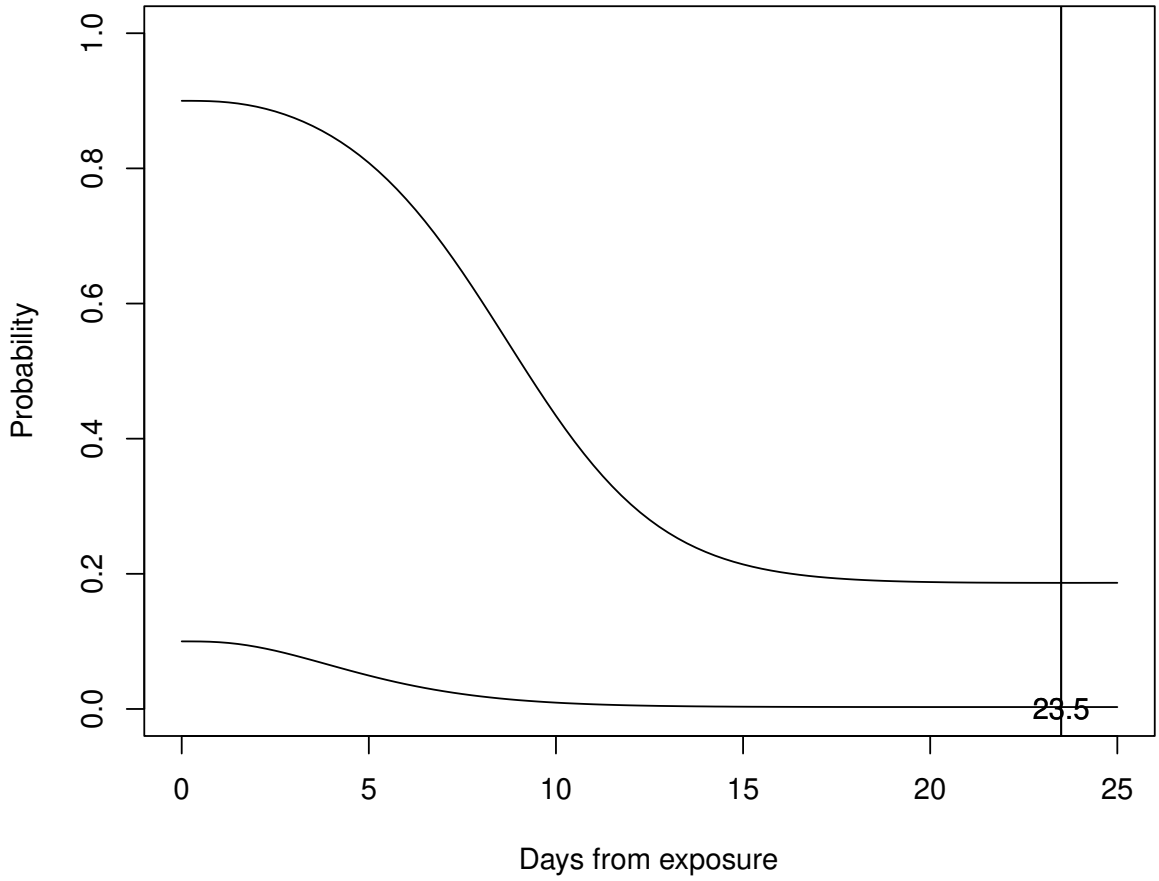

Cumulative probability
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
Days from exposure
Lat
Pre
Asy
Sym
Rec
Sus


Cumulative probability
0.0
0.2
0.4
0.6
0.8
1.0
Lat
Pre
Asy
Rec
0
5
10
15
20
25
Days from exposure


Cumulative probability
0.0
0.2
0.4
0.6
0.8
1.0
Lat
Pre
Asy
Rec
0
5
10
15
20
25
Days from exposure


Cumulative probability
0.0
0.2
0.4
0.6
0.8
1.0
Lat
Pre
Asy
Rec
0
5
10
15
20
25
Days from exposure


Probability
0.0
0.2
0.4
0.6
0.8
1.0
9
0
5
10
15
20
25
Days from exposure


Probability
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
23.5
Days from exposure

3 Discussion

- For individuals who were definitely infected on day 0 the probability of being in the latent or infectious state, given that they are asymptomatic and have a negative test, is less than the equivalent probability for individuals who were definitely not infected at day 0 **after about 14 days**.
- This is because most of the former are then recovered, and hence resistant, and the latter are still exposed to background risk of infection.
- Risk at, or possibly above, the baseline should be acceptable.
- Limitation: This analysis is limited to the short or medium term, before the possibility of reinfection becomes a factor.

References

Eikenberry, S. E., Mancuso, M., Iboi, E., Phan, T., Eikenberry, K., Kuang, Y., Kostelich, E. & Gumel, A. B. (2020), To mask or not to mask: Modeling the epotential for face mask use by the general public to curtail the COVID-19 pandemic, *Infectious Disease Modelling* **5**, 293–308.

Hu, C., Pozdnyakov, V. & Yan, J. (2020), Density and distribution evaluation for convolution of independent gamma variables, *Computational Statistics* **35**, 327–342.

Hu, C., Pozdnyakov, V. & Yan, J. (2021), *coga: Convolution of Gamma Distributions*. R package version 1.1.1.

Jansen, L., Tegomoh, B., Lange, K., Showalter, K., Figliomeni, J., BAbdalhamid, Iwen, P. C., Fauver, J., Buss, B. & Donahue, M. (2021), Investication of a SARS-CoV-2 B.1.1.529 (Omicron) variant cluster – Nebraska, November–December 2021, Morbidity and Mortality Weekly Report, CDC.

Menni, C., Valdes, A. M., Polidori, L., Antonelli, M., Penamakuri, S., Nogal, A., Louca, P., May, A., Fiueirede, J. C., Hu, C., Molteni, E., Canas, L., Osterdahl, M. F., Modat, M., Sudre, C. H., Fox, B., Hammers, A., Wolf, J., Capdevila, J., Chan, A. T., David, S. P., Steves, C. J., Ourselin, S. & Spector, T. D. (2022), Symptom prevalence, duration, and risk of hospital admission in individuals infected with SARS-CoV-2 during periods of omicron and delta variant dominance: a prospective observational study from the ZOE COVID Study, *Lancet* pp. 1618–1624.

R Core Team (2015), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

A Methods and programs for probability calculations

We calculate the state probabilities following an exposure event by conditioning on the paths taken through the course of the infection. The paths, states, variable and parameter names used here are those illustrated in figure 1.

We let $H \in \{1, \ldots 6\}$ index the path and let
$S_x \in \{\text{Sus, Lat, Asy, Pre, Mld, Svr, ResS, ResA}\}$ be the state the individual is in at time x days following the exposure event. Table 4 gives $P(H)$, the path probabilities derived from the splitting probabilities, and $P(S_x|H)$, the conditional probabilities of being in each state S_x given that path H is followed. We use $F_?$ as shorthand for the distribution function of sums of arbitrary Gammma distributed random varaibles. Thus, for instance,

$$\begin{aligned} F_{UVW} &= F_{U+V+W}(x) \\ &= P(U+V+W \leq x) \\ &\quad \text{where } U, V, W \sim \text{Gamma}(\bullet, \bullet) \end{aligned} \tag{1}$$

The state probabilities are then given by

$$P(S_x) = \sum_H P(S_x|H)P(H). \tag{2}$$

The specific path splitting probabilities and Gamma shape and rate parameters used here are given in tables 1, 2 and 3, but our implementation allows the user to change these values.

Table 4: The probability, $P(H)$, of following each possible path of progression through the infection, and the probability of being in each state at time x given the path, $P(S_x|H)$.

	H					
	1	2	3	4	5	6
$P(H)$	pqr	$pq(1-r)$	$p(1-q)$	$(1-p)qr$	$(1-p)q(1-r)$	$(1-p)(1-q)$
S_x	$P(S_x\|H)$					
Sus	0	0	0	$1-F_U$	$1-F_U$	$1-F_U$
Lat	$1-F_V$	$1-F_V$	$1-F_V$	F_U-F_{UV}	F_U-F_{UV}	F_U-F_{UV}
Asy	0	0	F_V-F_{VX}	0	0	$F_{UV}-F_{UVX}$
Pre	F_V-F_{VW}	F_V-F_{VW}	0	$F_{UV}-F_{UVW}$	$F_{UV}-F_{UVW}$	0
Mld	0	$F_{VW}-F_{VWZ}$	0	0	$F_{UVW}-F_{UVWZ}$	0
Svr	$F_{VW}-F_{VWY}$	0	0	$F_{UVW}-F_{UVWY}$	0	0
ResS	F_{VWY}	F_{VWZ}	0	F_{UVWY}	F_{UVWZ}	0
ResA	0	0	F_{VX}	0	0	F_{UVX}

To calculate the conditional probabilities given in the results section, we consider the probability of some observation O_x at time x and specify $P(O_x|S_x)$, the probability of the observation given the underlying state at that time. Then, using Bayes rule, we have

$$P(S_x|O_x) = \frac{P(O_x|S_x)P(S_x)}{\sum_{S_x} P(O_x|S_x)P(S_x)}. \tag{3}$$

Table 5 specifies the conditional probabilities required for the following events:

- I_x indicates that the individual is uninfected (0), or infected (1), at time x,
- M_x indicates that the individuals has not (0) or has (1) exhibited symptoms up to time x,
- T_x indicates that a test at time x is negative (0) or positive (1).

We note that the states ResS and ResA differ only in the path taken to get to them and we use this convention only because it allows us to distinguish between recovered individuls who have or have not shown symptoms in the past without explicit reference to the path, thus simplifying the first column of table 5.

To address the inferential problem of whether an exposure at time 0 resulted in an infection we let E_0 be and indicator for the event that an infection occured (1) or didn't (0). We can then calculate the required likelihoods given observations O_x as

$$\begin{aligned} P(O_x|E_0 = 0) &= \textstyle\sum_{S_x} P(O_x|S_x)P(S_x|p = 0) \\ P(O_x|E_0 = 1) &= \textstyle\sum_{S_x} P(O_x|S_x)P(S_x|p = 1) \end{aligned} \tag{4}$$

and if $P(E_0 = 1)$ is the prior probability that an infection occurred, the posterior given O_x is

$$P(E_0 = 1|O_x) = \frac{P(O_x|E_0 = 1)P(E_0 = 1)}{P(O_x|E_0 = 1)P(E_0 = 1) + P(O_x|E_0 = 0)P(E_0 = 0)}. \tag{5}$$

Table 5: For each state we give $P(M_x = 0|S_x)$, the probability of no symptoms up to time x given the state, $P(I_x = 1|S_x)$, the probability of being infected given the state, $P(T_x = 0|S_x, \text{Antigen})$, the probability of having a negative antigen test given the state, and $P(T_x = 0|S_x, \text{PCR})$, the probability of having a negative PCR test given the state.

S_x	$P(M_x = 0\|S_x)$	$P(I_x = 1\|S_x)$	$P(T_x = 0\|S_x, \text{Antigen})$	$P(T_x = 0\|S_x, \text{PCR})$
Sus	1	0	1	1
Lat	1	1	1	1
Asy	1	1	$1 - s$	$1 - s$
Pre	1	1	$1 - s$	$1 - s$
Mld	0	1	$1 - s$	$1 - s$
Svr	0	1	$1 - s$	$1 - s$
ResS	0	0	1	$1 - s$
ResA	1	0	1	$1 - s$

We wrote programs to implment the above methods using the `R` statistical programming environment (R Core Team 2015). These allow specifying arbitrary path splitting probabilities and arbitrary shape and rate paramiters for the Gamma time in state variables. These calculations require the `pcoga()` function in the `coga` package (Hu et al. 2021, Hu et al. 2020) to calculate the density functions of sums of independent Gamma random variables. codecoga can be intalled using the standard `R` command `install.packages("coga")`, and

is also available from the `CRAN` repository of `R` packages at
`https://cran.r-project.org/web/packages/coga/index.html`.

The `R` source code for these general programs is available from github at
`https://github.com/alungwalia/covidisolation` in the file `covid.R`. The specific code necessary to produce the tables and figures in this paper are also provided in the file `isol.R` and may be a useful example for those wishing to use or extend the methods described here.

A `shiny` application for calculating the state probabilities following an exposure using these programs is available at
`https://alunthomas.shinyapps.io/covidisolation/`. The app allows the user to input parameters and change the conditioning events.