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 [*]

January 10, 2023

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.

1 Introduction

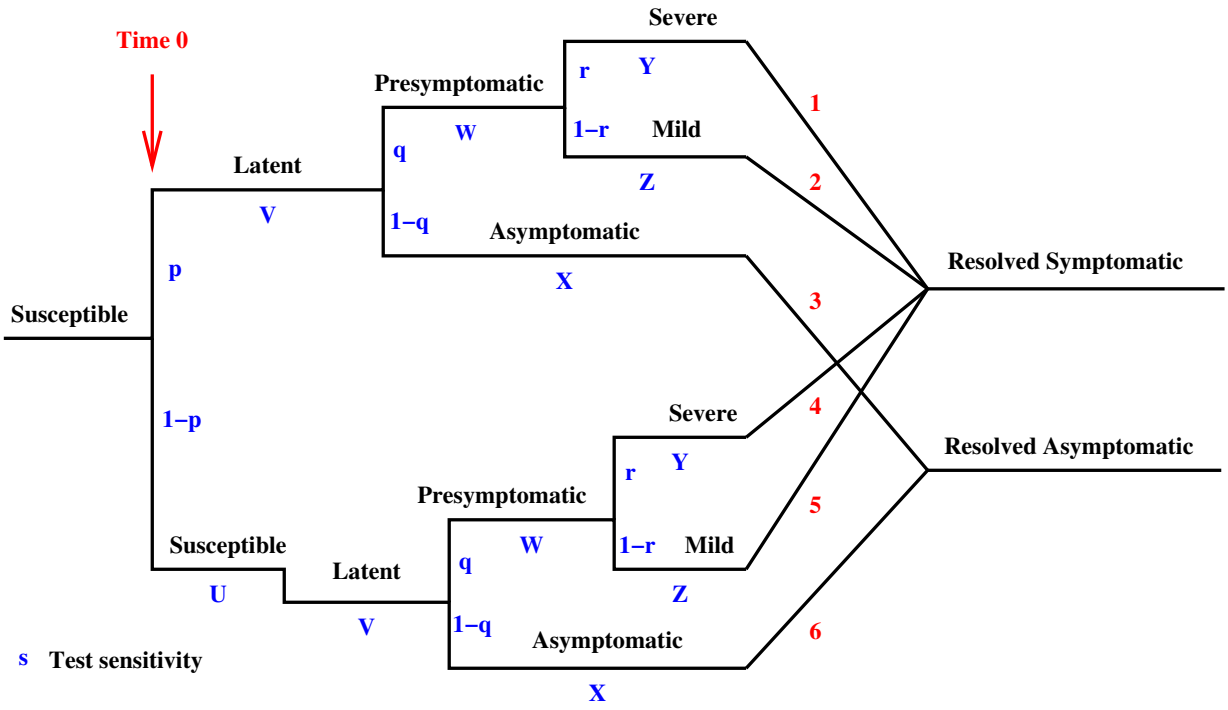
Throughout this work we use *infected* to mean that an individual is either in the latent state or one of the active states of the disease.

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

2 Methods

Figure 1 gives a schematic of the possible paths through a course of infection following an exposure to SARS-CoV-2 . With probability p the exposure results in an infection and the individual enters the latent state and progresses through subsequent phases. With probability $1 - p$ the individual remains susceptible until they become infected through general background exposure. Note that we distinguish two resolved states depending on whether or not the individual previously experienced symptoms. This is the sole distinction between these states and is introduced only as an artifice to simplify the computations described below. Probability parameters control the path taken, and the time in state variables are all assumed to be Gamma distributed.

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.



For the most part, to model the effects of the original variants of SARS-CoV-2 , we chose 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. 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 the 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: Probability parameter values specifying the path through the course of infection for the original SARS-CoV-2 variants.

Parameter	Name	Value
q	Symptomatic	0.50
r	Svrere	0.10

With regard to the later, omicron, SARS-CoV-2 variants, Jansen et al. (2021) estimate that the latent period is roughly 60% that for 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

Appendix A describes methods to compute the probabilities of being in each state at times following the exposure event given the model parameter values described here. The methods have been implemented in the `R` statistical programming environment (R Core Team 2015), and can be downloaded from `github` at

`https://github.com/alungwalia/covidisolation`.

A `shiny` application is also provided that allows the user to adjust the model and see the effects on state probabilities. This is available at

`https://alunthomas.shinyapps.io/covidisolation/`.

Observations about the individual following the exposure event will change the state probabilities. If they are observed not to have shown any symptoms, the mild and severe disease state are excluded as is the resolved symptomatic state.

In considering the effects of negative test results we distinguish between antigen tests and PCR tests. Since we assume that both test types are fully specific, they are both obligatorily negative for individuals in the susceptible and latent states, and subject to the test sensitivity, can be positive in the asymptomatic, presymptomatic, mild and severe disease states. Antigen tests are typically negative for resolved individuals, while PCR tests are typically positive for some time after resolution. PCR tests are also more sensitive. We set the sensitivity to 95% for PCR tests and 90% for antigen tests.

Appendix A also gives methods for computing state probabilities conditional on observations regarding symptoms and test results.

Since we are primarily concerned with modeling the short term following exposure, we do not model a transition from the resolved state back to susceptible, and we do not model a decrease in PCR sensitivity with time in the resolved state.

3 Results

Figure 2 shows the different state probabilities for an individual following an exposure with a 20% chance of resulting in an infection. This also shows the effects on the these probabilities of observing that the individual shows no symptoms, and, in addition, that they have either a negative PCR or antigen test.

Figure 3 addresses the inference problem of whether an infection occurred following exposure given that the individual remains asymptomatic and has a negative test. The curves show the posterior probabilities in a Bayesian analysis assuming a range of prior probabilities. These show that regardless of the prior probability assumed, the lowest posterior occurs at 9 days for an antigen test and 23.5 days for a PCR test indicating the most informative time to use the test for inferring an infection following exposure.

Figure 2: State probabilities following 70% chance of infection at exposure. This shows the unconditional probabilities and those when no symptoms and a negative test are observed.

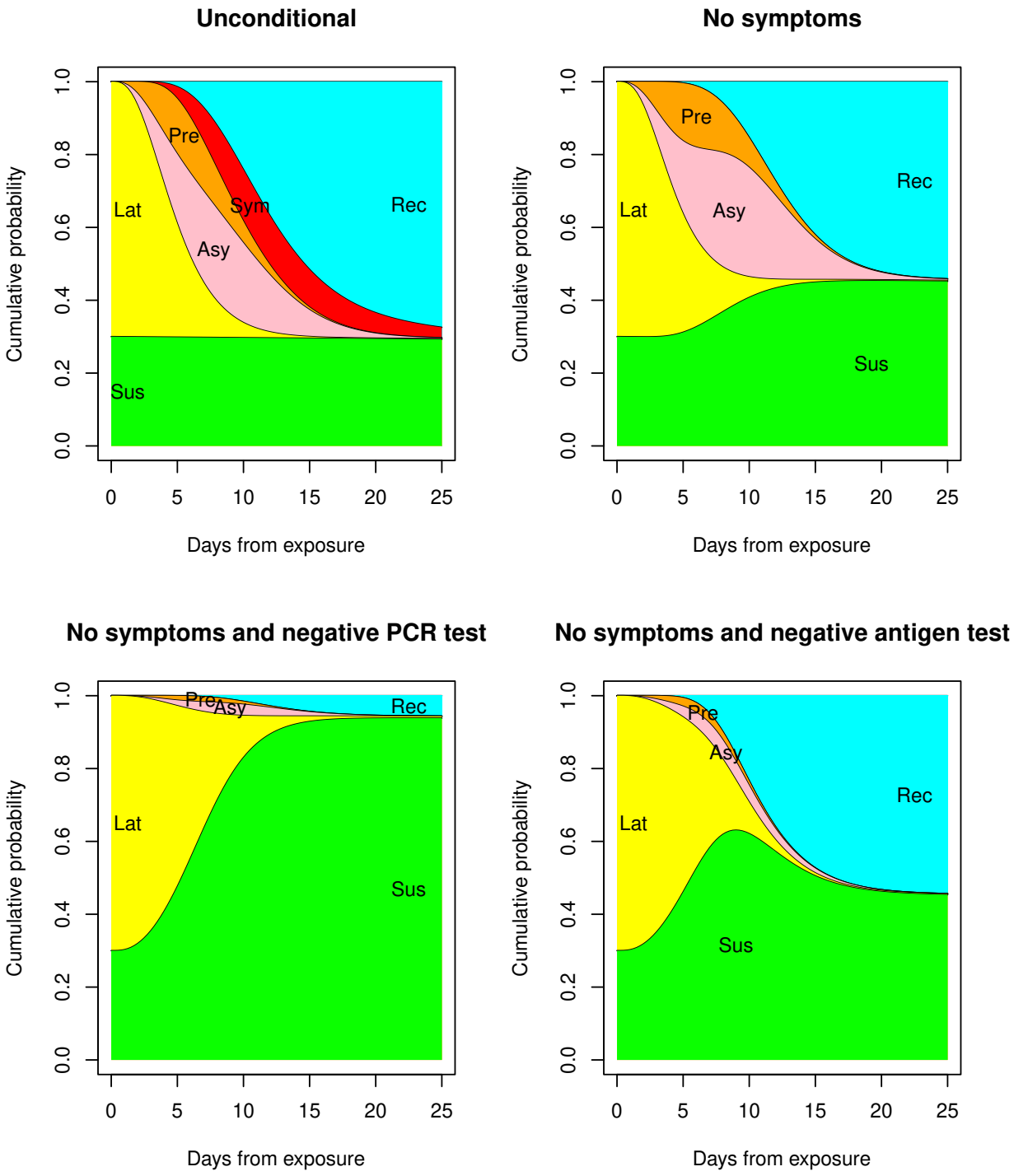
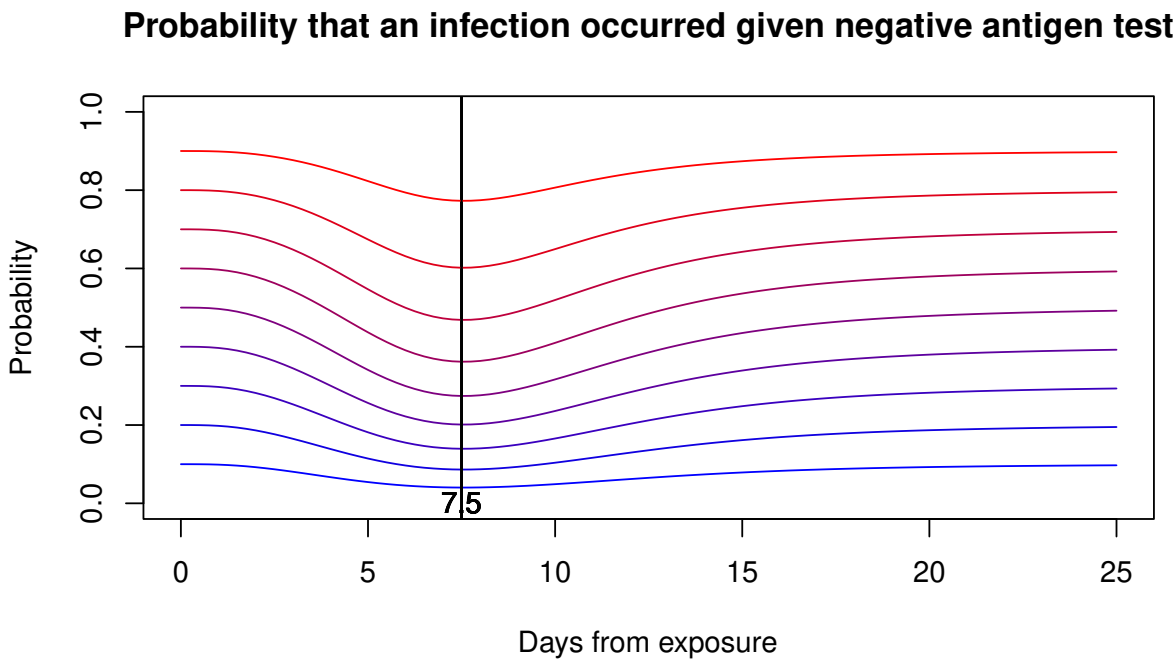
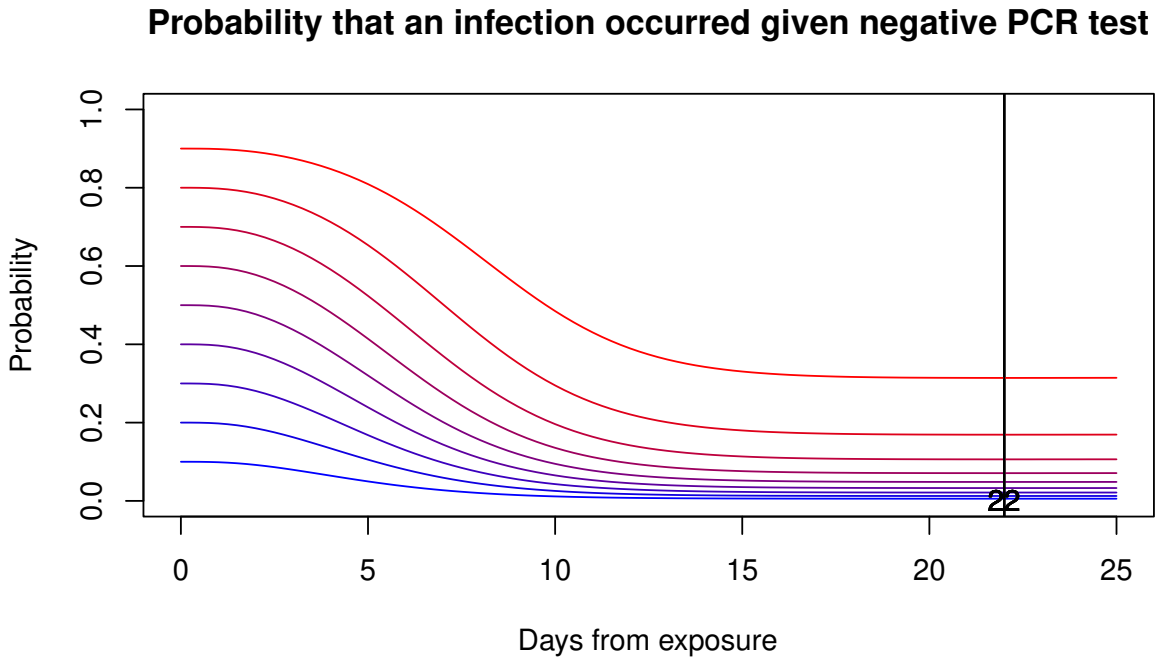


Figure 3: The posterior probability that an infection occurred at time 0 given a negative test following. The curves are given for varying prior probabilities. The vertical lines show where the posteriors are lowest.





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 variables. 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 individuals 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 occurred (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 implement 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 parameters 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 installed 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 application allows the user to input parameters and change the conditioning events.