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 [*]

August 10, 2023

Keywords: Isolation, quarantine, case counting, Gamma distribution.

Running title: Probability of infectiousness following exposure.

Abstract

We provide a method, programs, and web site for calculating the disease state probability of an individual immediately following an exposure to SARS-CoV-2 that may or may not have resulted in a transmission. We consider in particular the effect on these probabilities of the case where we subsequently observe that an individual is symptom free and/or has a negative SARS-CoV-2 test result. We illustrate the utility of these computations by calculating the time at which an exposed individual's risk of being infectious reaches an acceptable level, assessing the value of PCR and antigen testing for case counting, calculating the time at which the risk of an infected individual being still infectious reaches a level comparable with general background risk, and evaluating the benefit of a second test for a symptom free individual who has an initial negative test.

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

1 Introduction

Perhaps some more general introductory comments here to set the context and highlight significance.

We present a method for computing the SARS-CoV-2 disease state probabilities for an individual following a specific exposure event such as an occupational exposure for a health care worker or attendance at a gathering also attended by infected individuals. We then use these computations to address several relevant epidemiological questions.

Methods section 2 describes the SARS-CoV-2 disease progression model we have assumed, indicates sources in the literature for model parameter estimates, and outlines the nature of the computations. The details of, and justification for, the method are, however, deferred until appendix A, which also gives information regarding online resources for `R` source code (R Core Team 2015) implementing the methods and a `shiny` website that can be used by the reader to make these calculations and adjust the parameter values used.

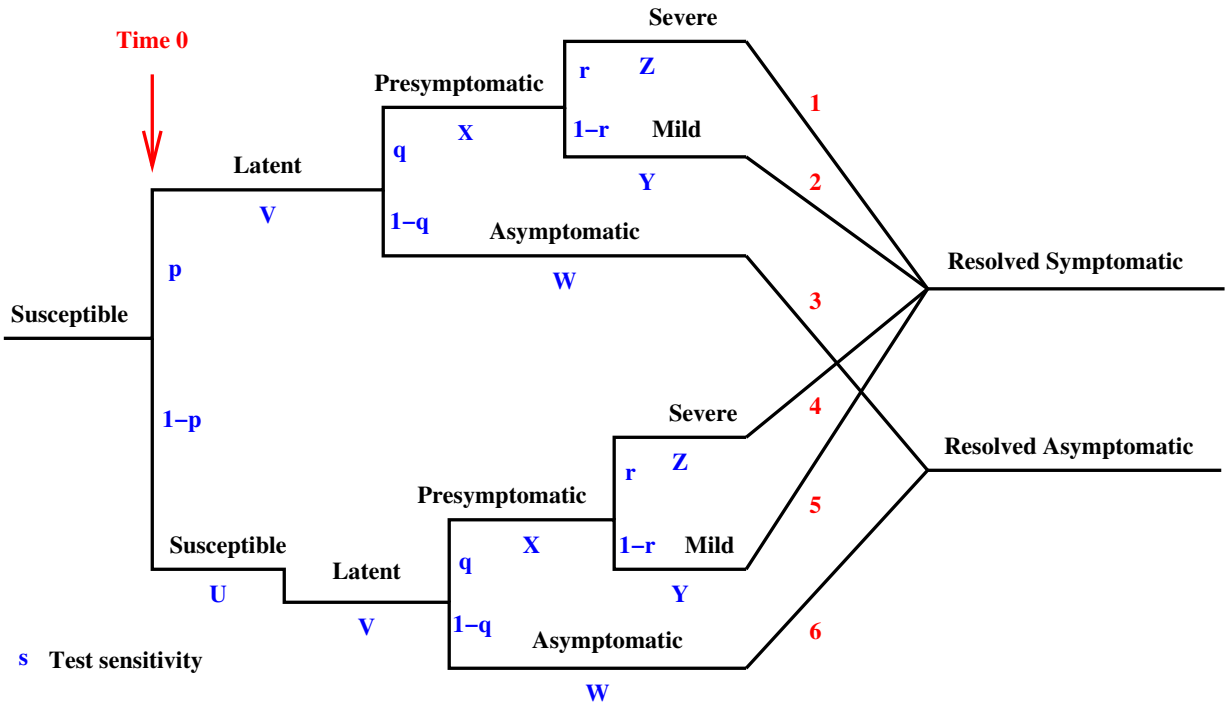
Results section 3.1 considers the disease state probabilities for an individual following an exposure event with a 30% probability of transmission. We additionally show the effect on these probabilities of observing the symptomatic state of the individual and of both PCR and antigen test results. Section 3.2 compares the value of PCR and antigen tests for case counting and otherwise inferring a transmission event. In section 3.3 we consider the time it takes for the risk of infectiousness of an infected individual to revert to that of a general member of the population. Finally, in section 3.4 we consider the added value of a second test for an individual who remains symptom free following an initial negative test.

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

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 in appendix A. 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. **Is the following still appropriate?**
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, allow for under reporting of cases, and

model more intense stages of the epidemic. Note that the model used by Eikenberry et al. (2020) does not have a *presymptomatic* phase, so 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	6.53	0.933	7.00	7.50	2.71	13.30
X	Pre	9.00	3.000	3.00	1.00	1.37	5.25
Y	Mld	0.55	0.110	5.00	45.50	0.01	24.06
Z	Svr	2.66	0.242	11.00	45.50	1.97	27.62

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	Severe	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	6.53	1.037	6.30	6.08	2.43	11.97
X	Pre	9.00	3.333	2.70	0.81	1.23	4.73
Y	Mld	0.55	0.122	4.50	36.86	0.01	21.66
Z	Svr	2.66	0.269	9.90	36.86	1.78	24.86

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.

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.

The methods presented here have been implemented in the `R` statistical programming environment (R Core Team 2015), and can be downloaded from `github` at

`https://github.com/alun-thomas/covidisolation`.

We provide programs to calculate the required probabilities both by exact computation and by simulation, the latter being useful when we wish to condition on specific events by selecting the matching simulations.

An interactive `shiny` web 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/`.

3 Results

3.1 Disease state probabilities

Figure 2 shows the different state probabilities for an individual following an exposure with a 30% chance of resulting in an infection for the Delta SARS-CoV-2 variant. 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. Observing typical SARS-CoV-2 symptoms and/or a positive test result leads to relatively clear cut conclusions and actions, which is why we focus on the case of a symptom free individual with negative test results.

Figure 3 shows the same plots for the Omicron variant. Vertical lines indicate the first time at which the summed probabilities of the infected states decrease to 10%.

For the Delta variant, figure 2(a) shows that by 15 2/3 of the infected individuals will have passed through the disease and entered the resolved state. Observing that the individual has no symptoms, figure 2(b), shifts the 10% line to near day 12. It also provides posterior evidence that a transmission did not occur reflected in an almost 10% increase in the green block. Of the infected individuals, about 1/2 have progressed to resolved and 1/2 are likely to be in the asymptomatic disease state. If we additionally observe a negative test, figures 2(c) and (d), the 10% line moves to about 7 days. There is further posterior evidence that a transmission didn't occur. At this early stage, the probability is that any infected individual is still in the latent phase of the infection.

The results for the Omicron variant are qualitatively similar to those for the Delta, however, because of the shorter latent period, everything happens about 3 days earlier.

We note that the implications of a later negative test depends on whether it is antigen or PCR. For the antigen test, negativity can be evidence for there having been no initial transmission, or that the individual has passed through the infection to the resolved phase. A negative PCR test, on the other hand, is strong evidence for no transmission.

Can we comment further here about the clinical relevance of these observations, or should we leave that to the discussion?

Figure 2: State probabilities following a 30% chance of transmission at exposure to the SARS-CoV-2 Delta variant. This shows the unconditional probabilities and those when no symptoms and negative tests are observed. The vertical black lines show where the probability of the infected states decreases to 10%.

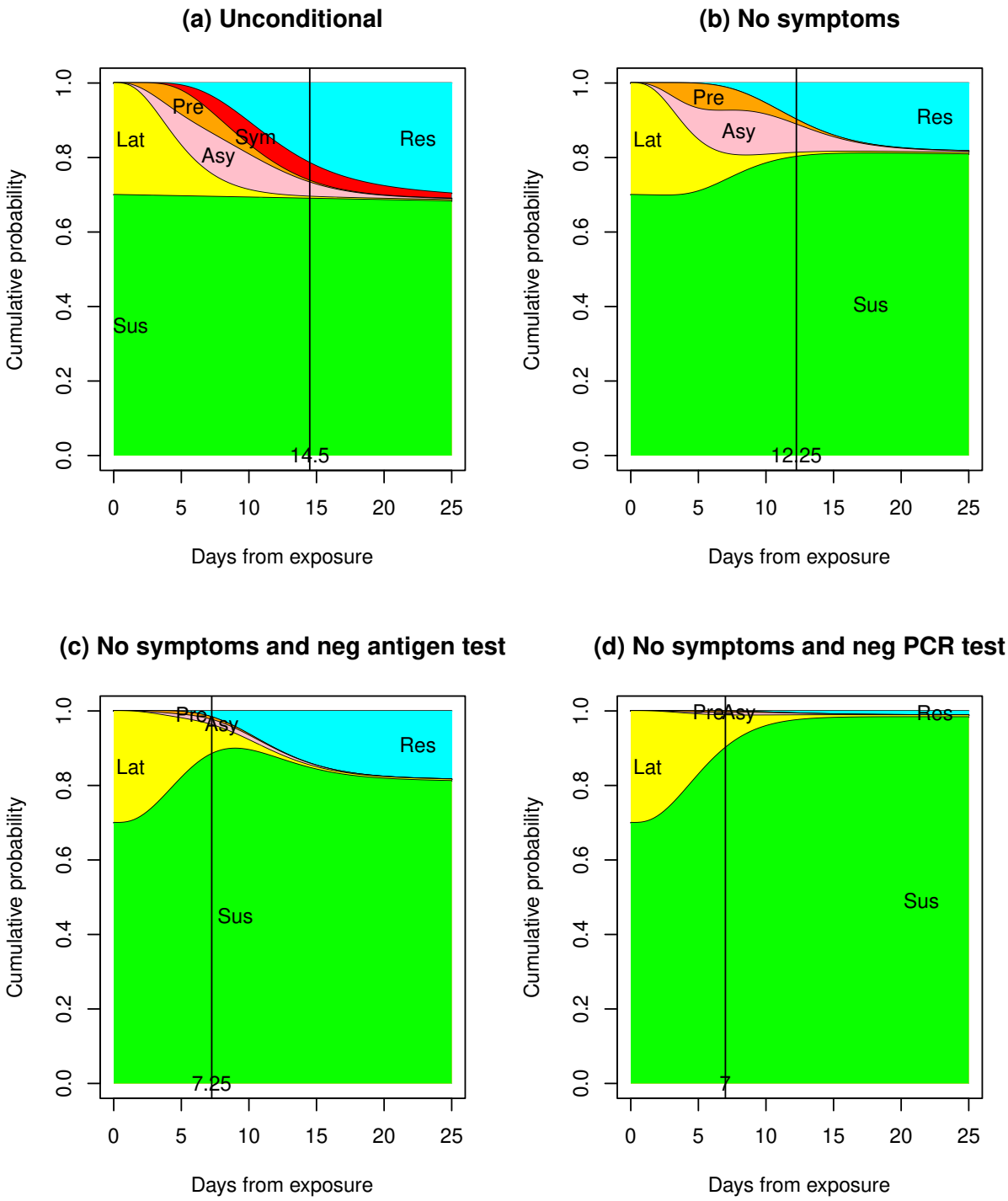
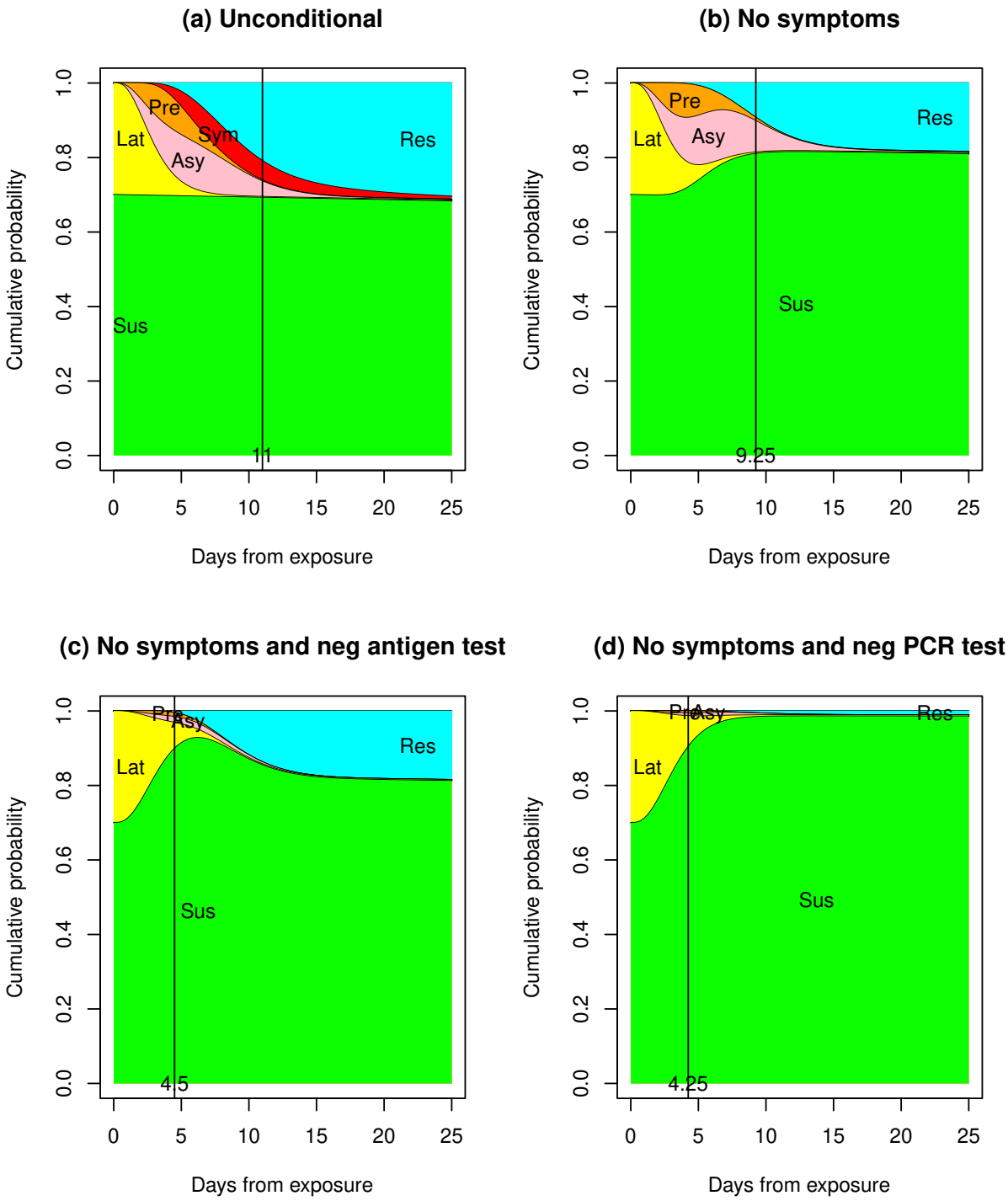


Figure 3: State probabilities following a 30% chance of transmission at exposure to the SARS-CoV-2 Omicron variant. This shows the unconditional probabilities and those when no symptoms and negative tests are observed. The vertical black lines show where the probability of the infected states decreases to 10% .



3.2 Inferring transmission events

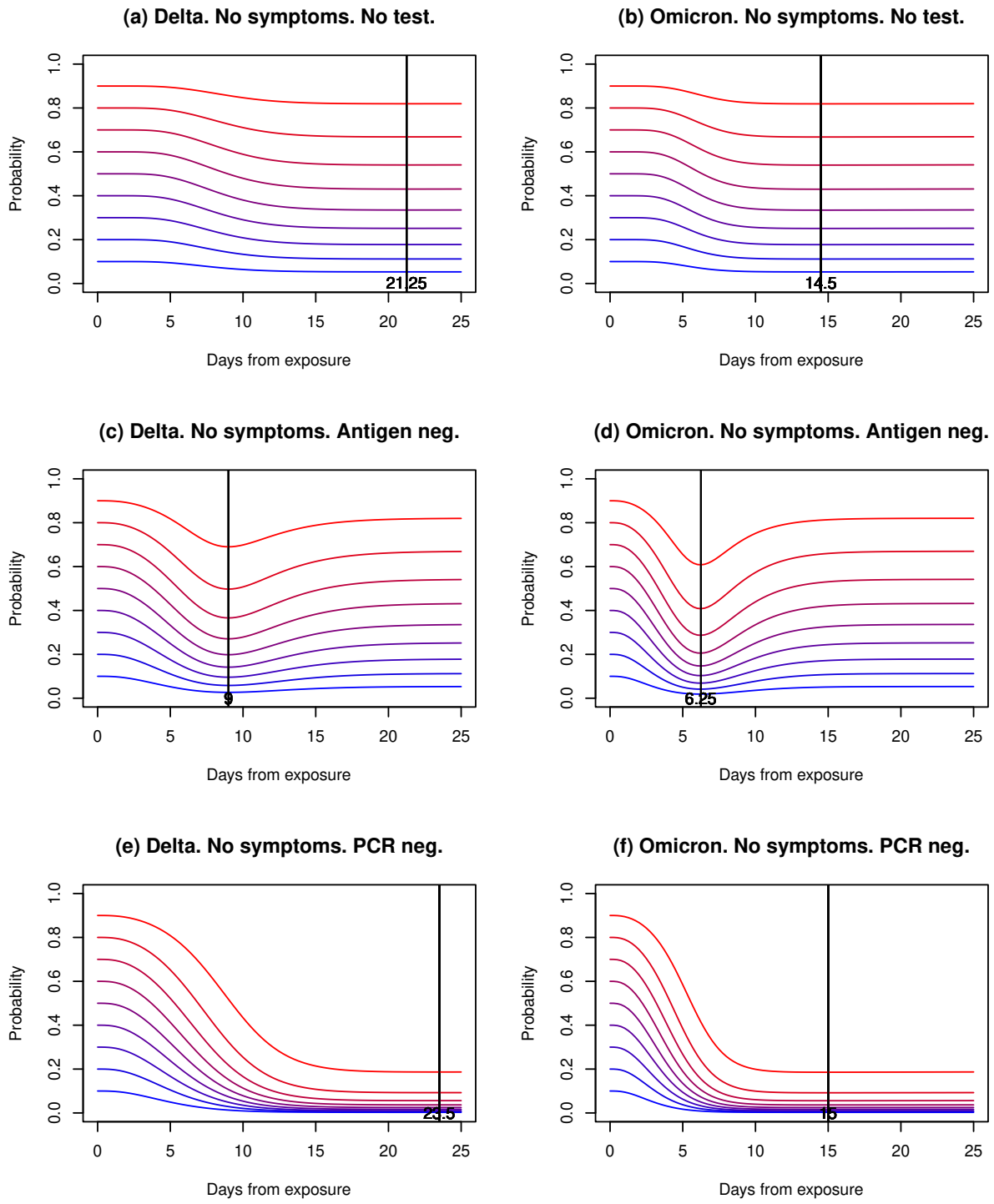
We now consider the value of testing in inferring whether or not not a transmission actually resulted from the exposure. We did this in a Bayesian analysis where we assumed a range of prior probabilities for transmission and considered how observing that the individual was symptom free and had negative test results affected the corresponding posterior probabilities. We again focused on negative evidence because the tests are highly specific and the consequences of positive results clear. Figure 4 shows the curves of the posterior probabilities as a function of time for a range of prior probabilities. This is done for no tests, antigen tests, and PCR tests and for both Delta and Omicron variants.

The broad impression given by these figures might well have been anticipated: because PCR tests can detect SARS-CoV-2 for some time after clinical infection is over, a negative PCR test is a far stronger indicator that the individual was never infected than a negative antigen test. This is shown by substantial reduction in levels in figures 4(e) and (f). However, it might be surprising just how uninformative a negative antigen test is as shown by the relatively small downward deflections in the curves in figures 4(c) and (d) relative to (a) and (b). Furthermore, there is a very narrow window in time, around 7 to 12 days for Delta, around 3 to 8 days for Omicron, for obtaining even this information with the curves reverting to the levels of the posteriors given only that the individual was symptom free shortly after as the individual passes into the resolved state.

For case counting purposes, a PCR test is far more informative and should be carried out after enough time has elapsed to avoid testing in the latent period. A rough rule might be to test around 15 days following the exposure.

Further observations regarding clinical significance?

Figure 4: The posterior probability that an infection occurred at time 0 given a negative test following and no symptoms. The curves are given for varying prior probabilities. The vertical lines show where the posteriors are lowest.



3.3 Effect of testing on assessing infectiousness

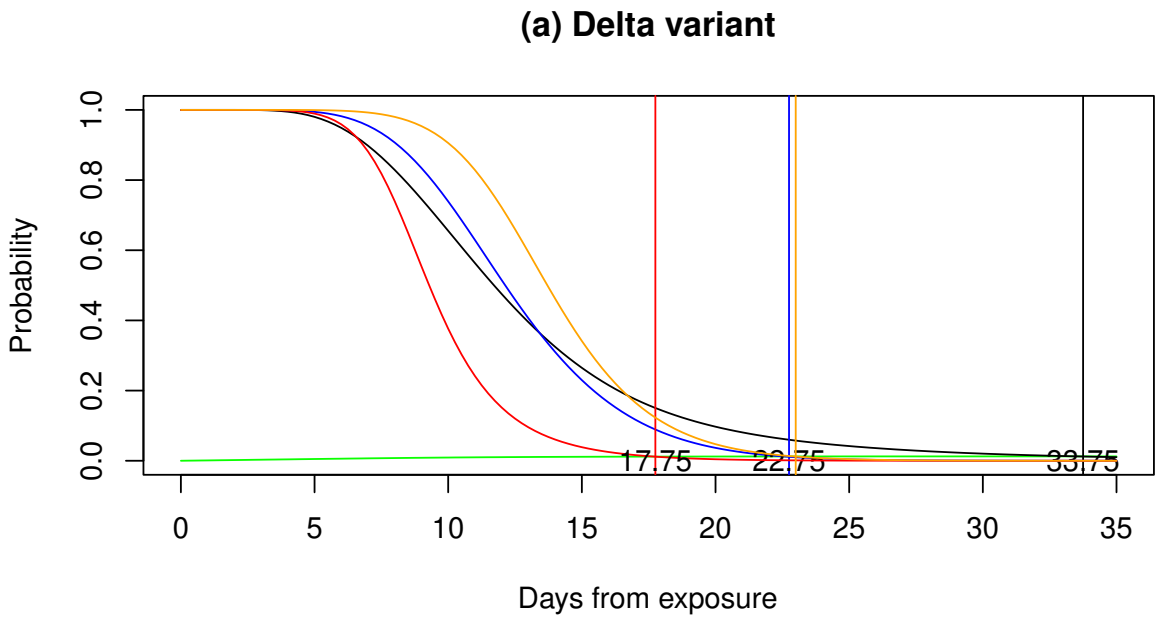
Perhaps the most pressing clinical issue for an individual with suspected exposure is when to consider their risk of infecting other susceptible individuals has reduced to an acceptable level. To address this, we considered the two extreme cases following exposure, that is, the case of an individual who we know for certain was infected with the case of one we who we know for certain was not. Infected individuals will progress through to the resolved state with survivors becoming resistant at least in the medium term and within the assumptions of our model. Conversely, uninfected individuals will remain susceptible to background risk of infection. If we, therefore, plot the probabilities of being infected for these two cases as a function of time, the curves will eventually cross and this crossing time would be an indicator that the infected individual presents no risk above that of a general member of the population. While this might be a very conservative criterion for ending isolation, examining it can still inform the value of testing. The curves for the cases where we assume an intermediate probability of transmission will be numerical averages of the two extreme cases and, therefore, will share the same crossing points. Figure 5 shows plots for these probabilities, and also how they are affected by observing that an individual shows no symptoms and has a negative SARS-CoV-2 test.

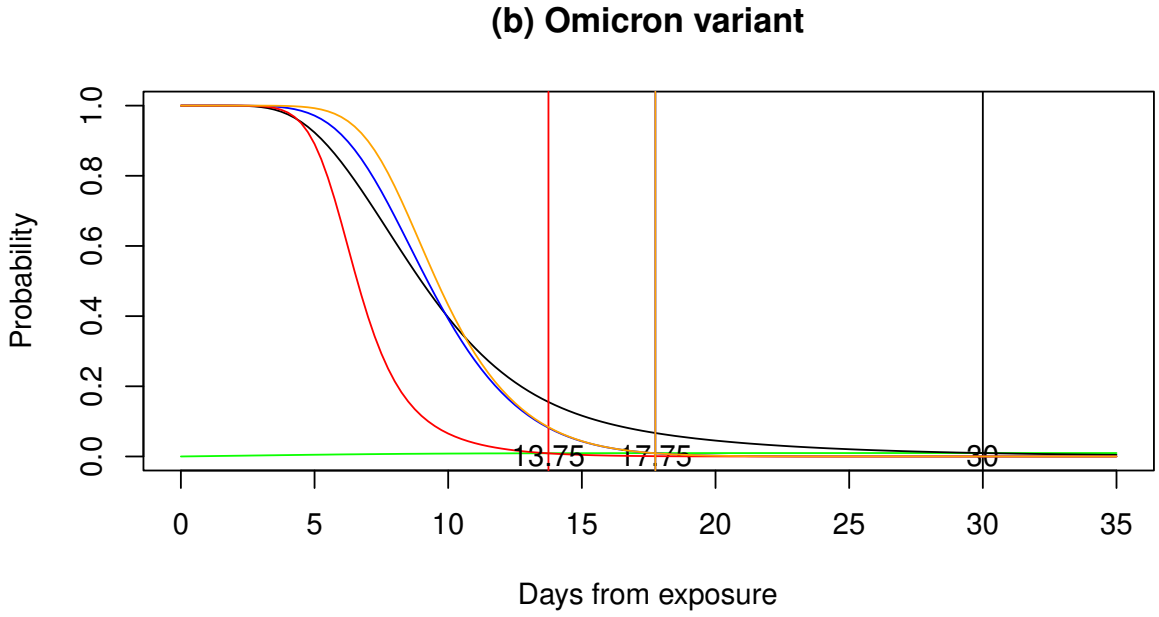
The black lines in figures 5(a) and (b) show that it takes on average almost 5 weeks, 33.75 days for Delta, 30 days for Omicron, for an individual to revert to presenting background levels of risk. Monitoring symptoms and testing can, however, detect those who have progressed quickly and, hence, present lower risk sooner, at a much earlier time. Symptom free individuals reach background risk at 22.75 days for Delta and 17.75 days for Omicron as shown by the blue lines Note that a negative PCR test does not give much further additional information as shown by the closeness of the blue and orange lines in the figure. For the Omicron variant, figure 5(b), the vertical lines showing the crossing points are indistinguishable.

Conversely, a negative antigen test moves the crossing point 5 days earlier for Delta, and 4 days earlier for Omicron as shown by the red lines. This is because a negative PCR test can select for individuals with an unusually long latent period, while a negative antigen test detects those individuals who have progressed quickly through the whole course of the infection. Thus, for purposes of assessing the risk an infected individual poses to other, an antigen test is far more valuable.

Further comments?

Figure 5: The probability that an individual who was certainly infected at time zero is still infectious at a given time following is shown in black. The blue line shows the probability for and individual who is asymptomatic up to the given time, the red line for in asymptomatic individual who has a negative Antigen test at the given time, the orange line for in asymptomatic individual who has a negative PCR test at the given time. The green line shows the probability for an individual who was certainly not infected at time zero, but was subsequently exposed to background risk of infection. Vertical lines in the appropriate colour show where the probabilities for certainly infected individuals first fall below those for certainly uninfected.





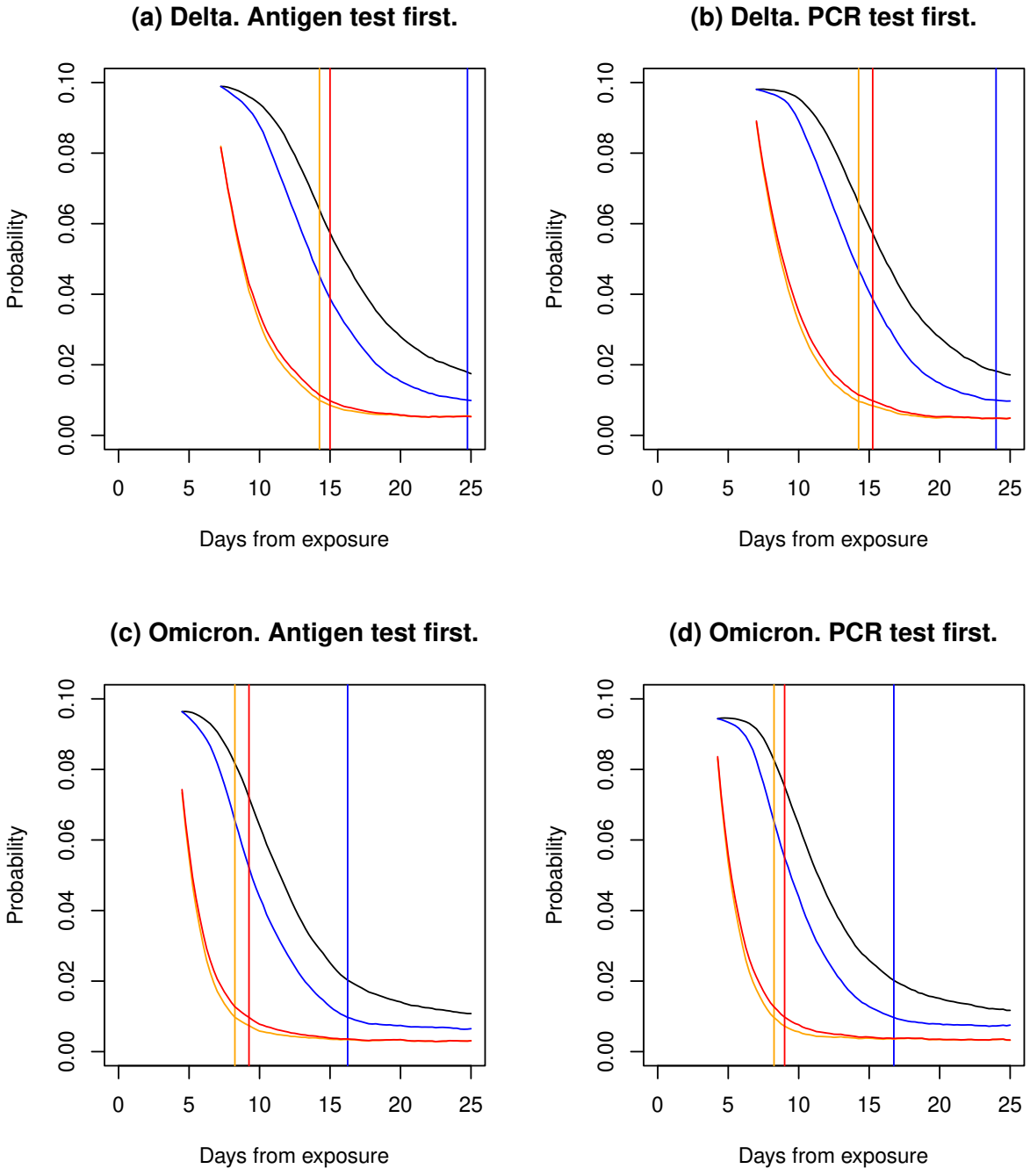
3.4 Follow up testing

To investigate the value of a second test in the case of a symptom free individual, we considered applying one to an individual with an initial negative test. The time for the initial test varied by the variant being considered, Delta or Omicron, and the type of the first test, antigen or PCR, and was chosen to be at the time established in section 3.1 where the posterior probability that an individual, with a 30% prior probability of infection, is infected first drops below 10%. We calculated the ongoing probability of infection for unmonitored individuals, those monitored and observed to remain symptom free, and those who in addition had a second test.

These probabilities of infection are presented in figure 6. They were calculated by simulation where the simulations were weighted by the probability that an individual remains symptom free and tests negative at the initial test time.

The blue lines in the figures show that the probability that a symptom free individual is infected decays to 1% at about 25 days for the Delta variant, and 16 days for Omicron. A negative follow up test, however, can indicate an individual whose probability has fallen to this level about 10 days earlier for Delta, and 8 days earlier for Omicron. Thus, for instance, for Omicron negative tests at day 5 and at day 10 can detect a such a low risk individual. While the results are different for the two variants, they do not depend substantially on which types of tests are used or on the order if two different types are used.

Figure 6: Plots of the probability that an individual is infected given that they are symptom free and have an initial negative test. These are presented for both Delta and Omicron variants and for the case when the initial test is an antigen test and a PCR test. The black lines are for individuals with no follow up, blue for individuals who are observed to remain symptom free, red for symptom free individuals with a further negative antigen test and orange for symptom free individuals with a further negative PCR test. The vertical lines indicate where the curves first cross the 1% threshold.



4 Discussion

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 Gamma 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_{VW}	0	0	$F_{UV}-F_{UVW}$
Pre	F_V-F_{VX}	F_V-F_{VX}	0	$F_{UV}-F_{UVX}$	$F_{UV}-F_{UVX}$	0
Mld	0	$F_{VX}-F_{VXY}$	0	0	$F_{UVX}-F_{UVXY}$	0
Svr	$F_{VX}-F_{VXZ}$	0	0	$F_{UVX}-F_{UVXZ}$	0	0
ResS	F_{VXZ}	F_{VXY}	0	F_{UVXZ}	F_{UVXY}	0
ResA	0	0	F_{VW}	0	0	F_{UVW}

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 resolved 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. `coga` 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`.

We also wrote programs to calculate these probabilities by simulating infection histories for individuals over a period of time following exposure. This was done in the straightforward way suggested by figure 1 using the standard `R` functions `runif()` to generate Uniform random variates for the path splitting events. and `rgamma()` to generate the time-in-state values. We checked that the exact numerical and the simulation methods gave the same results within the sampling error of the simulation approach. The simulations are quick and the results are visually almost indistinguishable from the exact values when using 10000 simulated individual histories. More simulations may be needed if rejecting histories to condition on specific events.

The `R` source code for these general programs is available from github at `https://github.com/alun-thomas/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. A sctreenshot of the application is shown in figure 7.

Figure 7: A screen shot of the `shiny` application for calculating SARS-CoV-2 state probabilities.