

# A within-host model of SARS-CoV-2 infection

Jonathan Carruthers<sup>1</sup>, Jingsi Xu<sup>2</sup>, Thomas Finnie<sup>1</sup>, Ian Hall<sup>2</sup>

1. Public Health Advice, Guidance and Expertise (PHAGE) Joint Modelling Team, UK Health Security Agency, 2. Department of Mathematics and School of Health Sciences, University of Manchester.

## Abstract

Within-host models have been used to successfully describe the dynamics of multiple viral infections, however, the dynamics of SARS-CoV-2 virus infection remain poorly understood. A greater understanding of how the virus interacts with the host can contribute to more realistic epidemiological models and help evaluate the effect of antiviral therapies and vaccines. Here, we present a within-host model to describe SARS-CoV-2 viral dynamics in the upper respiratory tract of individuals enrolled in the UK COVID-19 Human Challenge Study. Using this model, we investigate the viral dynamics and provide timescales of infection that independently verify key epidemiological parameters important in the management of an epidemic. In particular, we estimate that an infected individual is first capable of transmitting the virus after approximately 2.1 days, remains infectious for a further 8.3 days, but can continue to test positive using a PCR test for up to 27 days.

## Introduction

The SARS-CoV-2 virus has led to a global pandemic, causing millions of infections worldwide. Mathematical modelling has been used throughout the pandemic to predict numbers of cases, hospitalisations and deaths, as well as the impact of non-pharmaceutical interventions [1, 2, 3]. In these epidemiological models, the rate at which susceptible individuals become infected can be thought of as the product of a contact rate and a transmission probability. The contact rate is simply interpreted as the number of contacts an individual makes per unit time, with insights provided from social contact studies [4, 5]. The probability of transmission, however, is not as well-defined and depends on many factors such as the environment in which the contact takes place, the duration of the contact and the infectiousness of the individual. Intuitively, an individual is most infectious when viral loads in their respiratory tract peak [6, 7], therefore understanding how viral loads change over time is crucial to understanding infectiousness. Within-host models can be used to describe how interactions between virus particles, host cells and the immune system affect the viral load over time, with the link to infectiousness then provided by either dose-response relationships or the assumption that the transmission probability is an increasing function of viral load [8, 9].

Existing within-host models of SARS-CoV-2 have predominantly focused on describing the dynamics of the virus in the upper respiratory tract (URT), the primary site of infection, where data are readily available from nasopharyngeal swabs [10, 11, 12]. In some cases, these models are extended to include compartments such as the lower respiratory tract (LRT), however, observing LRT infection is difficult and is unlikely to play a

role in person-to-person transmission [13]. Often, swab data are presented in the form of polymerase chain reaction (PCR) cycle threshold values (Ct values) that are converted into RNA copies by means of a standard curve. However, since a PCR test captures all genetic material, it quantifies the total viral RNA and not the amount of viable virus that is capable of infecting cells [14]. This therefore makes it difficult for models parameterised using such data to predict the duration that an individual remains infectious; a previous study has shown that live virus cannot be isolated from respiratory samples 8-9 days after symptom onset notwithstanding persistently high viral RNA loads [15]. Despite this, some models have directly estimated the infectiousness of an individual from predictions of total viral load [16, 17], whilst others have first linked total and infectious viral loads using the probability that a sample is culture positive [18].

A second limitation of earlier within-host models is that swab samples were understandably only collected once an individual first displayed symptoms, so the exact date of infection is unknown. The majority of models account for this by assuming a fixed incubation period of between 3 and 7 days, however, since symptom onset approximately coincides with peak viral loads, the models can only reliably predict the decay in viral loads and not the initial growth [19]. Human challenge studies allow us to study the processes of infection and immunity from their inception and therefore provide a complete overview of the disease course [20]. Screening of participants can also ensure that samples are representative of primary infection, whilst a quarantine setting allows data to be obtained with a level of detail not achievable in a real-life setting [21].

In this paper, we develop a mathematical model to describe the dynamics of wild-type SARS-CoV-2 infection in the URT. We make use of infectious and total virus time courses from participants of a human challenge study, thereby addressing the above limitations of existing models. Bayesian inference methods are applied to estimate model parameters and the within-host basic reproduction number. The mathematical model is then used to predict timescales of infection, such as the duration that an infected individual remains infectious and PCR positive.

## Methods

### Model structure

The viral dynamics in the mid-turbinates of human challenge study participants are described using a target-cell limited model that has successfully been used to describe the dynamics of other viral infections, and makes use of measurements of both infectious and total viral loads [22, 23, 24]. In this model, target cells ( $T$ ) become infected with infectious virus ( $V_{\text{inf}}$ ) at rate  $\beta$  and enter a non-productive eclipse phase ( $E$ ) for  $1/k$  days, after which they enter a productively infectious phase lasting for a further  $1/\delta$  days. Whilst cells are in the infectious phase, they produce infectious virus at rate  $p_{\text{inf}}$  which in turn loses infectivity at rate  $c_{\text{inf}}$ . Viral RNA is produced by an infectious cell at rate  $p_{\text{tot}}$  and lose viability with rate  $c_{\text{tot}}$ . To represent an adaptive immune response, we note that effector cell responses ( $C$ ), such as those of  $\text{CD8}^+$  T cells, go through an antigen-dependent activation stage, followed by a virus-independent expansion phase. The growth rate of the effector cell population is therefore initially dependent on the total viral load ( $V_{\text{tot}}$ ), but once this becomes

large enough, exponential growth of the effector cell population follows [25]. Effector cells help to resolve infection through the mass action killing of infected cells with rate  $k_C$  and survive for an average of  $1/\delta_C$  days. The model is described by the system of ordinary differential equations:

$$\begin{aligned}
 \frac{dT}{dt} &= -\beta TV_{\text{inf}} \\
 \frac{dE}{dt} &= \beta TV_{\text{inf}} - kE \\
 \frac{dI}{dt} &= kE - \delta I - k_C CI \\
 \frac{dV_{\text{inf}}}{dt} &= p_{\text{inf}} I - c_{\text{inf}} V_{\text{inf}} \\
 \frac{dV_{\text{tot}}}{dt} &= p_{\text{tot}} I - c_{\text{tot}} V_{\text{tot}} \\
 \frac{dC}{dt} &= \frac{rCV_{\text{tot}}}{V_{\text{tot}} + s} - \delta_C C
 \end{aligned} \tag{M1}$$

where  $s$  determines the viral load at which effector cell growth becomes virus independent and  $r$  is the maximal growth rate of effector cells [25].

### Parameter inference and model selection

An Approximate Bayesian Computation Sequential Monte Carlo (ABC-SMC) algorithm is used to infer model parameters from the human challenge data [26, 27]. This method is an extension of the original ABC algorithm and uses a sequence of decreasing tolerances to obtain intermediate distributions that ultimately converge on the target posterior distribution. Parameter sets are sampled from the intermediate distribution of the previous generation and perturbed using a multivariate normal kernel, the covariance matrix of which is selected here to be the optimal local covariance matrix introduced in [28]. Choosing a suitable sequence of tolerances is important since tolerances that decrease too slowly will result in slower convergence, whilst a sequence that decreases too quickly will lead to lower acceptance rates. To avoid manually setting the tolerances beforehand, an adaptive approach is implemented where the tolerance for the current generation is equal to the 0.25 quantile of the distances from the previous generation [26]. For the first generation, the tolerance is equal to the  $N^{\text{th}}$  smallest distance from  $5N$  iterations of an ABC rejection sampling algorithm, where  $N = 5 \times 10^3$  is the desired sample size [29].

Solutions of the mathematical model (MM) for infectious virus and total virus are compared to the measurements of viral load from the human challenge (HC) nasal samples using the following distance metric:

$$d(\text{MM}, \text{HC})^2 = \sum_{t \in \mathcal{T}} \left[ \log_{10}(V_{\text{inf}}^{(\text{MM})}(t)) - \log_{10}(V_{\text{inf}}^{(\text{HC})}(t)) \right]^2 + \left[ \log_{10}(V_{\text{tot}}^{(\text{MM})}(t)) - \log_{10}(V_{\text{tot}}^{(\text{HC})}(t)) \right]^2,$$

where  $\mathcal{T}$  is the set of times at which samples were taken. At times when the model solution falls below the limit of quantification ( $6 \times 10^2$  RNA copies/mL and 5 PFU/mL), the solution is instead set equal to this limit. In doing so, no penalty is incurred when solutions drop far below the limit of quantification.

Given that the human challenge study collected samples from the first 14 days of infection, there is unlikely to be a significant decline in the adaptive immune response, therefore  $\delta_C = 0$ . For the length of the eclipse phase, *in vitro* studies have showed that virus is released from infected cells 8 hours after infection, suggesting that it is suitable to fix  $k = 3 \text{ day}^{-1}$  [30]. The viral load at which effector cell growth becomes independent of virus also remains constant at  $s = 10^4$  RNA copies/mL. For the remaining parameters, prior distributions are chosen as follows:

$$\begin{aligned} \log_{10} \beta &\sim U(-7, -2) & \delta &\sim U(0, 5) & \log_{10} p_{\text{inf}} &\sim U(-1, 3) & \log_{10} p_{\text{tot}} &\sim U(1, 6) \\ c_{\text{inf}} &\sim U(0, 50) & c_{\text{tot}} &\sim U(0, 10) & r &\sim U(0, 2) & \log_{10} k_C &\sim U(-5, -2) \end{aligned}$$

Previous studies have noted that only the product of  $T(0)$  and the viral production rate is identifiable from measurements of viral load [30]. Since it is more common to estimate viral production rates, the initial number of target and effector cells is set to  $T(0) = 10^6$  cells/mL and  $C(0) = 1$  cell/mL respectively, whilst all other cell populations are initially zero. Local sensitivity analysis indicated that  $V_{\text{tot}}(0)$  did not have a large impact on the viral load, likely because  $V_{\text{tot}}$  is not directly involved in the infection of target cells, and is therefore fixed at  $V_{\text{tot}}(0) = 10^2$  RNA copies/mL. Participants of the human challenge study received a challenge dose of 10 TCID<sub>50</sub> via intranasal drops, which approximately corresponds to 275 PFU/mL. However, it is unknown whether this level of virus reaches cells in the mid-turbinate. For this reason, the initial condition  $V_{\text{inf}}(0)$  is also inferred with prior distribution  $V_{\text{inf}}(0) \sim U(-2, 2)$ . The inference has been repeated using different prior distributions for  $V_{\text{inf}}(0)$ , but this did not result in significantly different approximate posterior distributions for the remaining model parameters (see Figure S1).

The ABC-SMC algorithm is used to infer parameters for each individual separately and therefore produces 16 separate approximate posterior distributions. To obtain a single distribution that is representative of all participants, a mixture distribution is constructed. Suppose that  $f(\theta)$  is the probability density function for the true posterior distribution, where  $\theta$  is the vector of model parameters. The density function for the mixture distribution is then given by:

$$m(\theta) = \sum_{i=1}^P w_i f_i(\theta),$$

where  $P = 16$  is the number of individuals that developed infection and  $w_i$  are the weights for each individual. Here,  $w_i = 1/P$  for all  $i$  since there is no evidence to suggest that any single participant is preferred over the others. Approximate samples from the mixture distribution can be obtained by first sampling an individual and then sampling from the corresponding approximate posterior sample.

To compare the mathematical model in M1 to an equivalent model without an immune response, an ABC-SMC model selection algorithm is used [27]. Model M2 is defined using the same system as M1 but removes the final ODE describing the effector cell population and therefore also ignores the clearance of infected cells by effector cells. The model selection algorithm follows the same procedure as the ABC-SMC algorithm for parameter inference but introduces a step prior to sampling candidate parameter sets where a candidate model is sampled. At the end of each generation, the Bayes factor for

model M1 in favour of model M2 is calculated as follows:

$$B_{M1,M2} = \frac{P(M1|x)}{P(M2|x)},$$

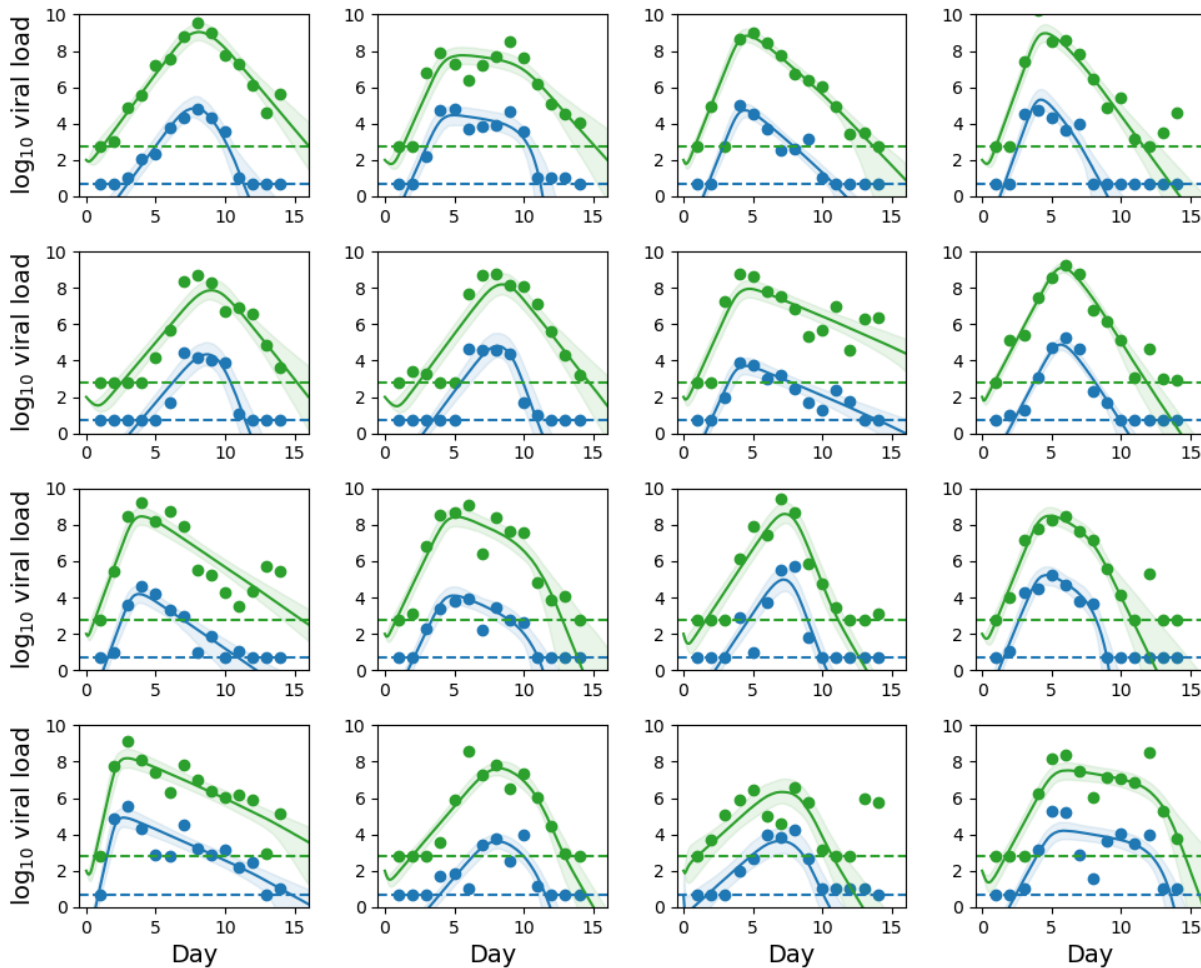
where the prior probabilities for model M1 and M2 are assumed to be equal. For each participant, the sequence of tolerances used in the model selection algorithm is the same as that obtained from the inference using model M1. A Bayes factor greater than 3 indicates positive evidence, whilst a value greater than 20 indicates strong evidence in favour of model M1 [27].

## Results

Measurements of viral load were obtained from the human challenge study described in [31]. In this study, 36 healthy, young participants were inoculated intranasally with a wild-type SARS-CoV-2 virus following a screening process that ruled out previous SARS-CoV-2 infection and recent respiratory infections. Two of these participants seroconverted between screening and inoculation and were removed from the study. Of the 34 individuals that continued in the study, 18 developed an infection. Six of these individuals were treated with remdesivir once two consecutive nose or throat swabs showed quantifiable SARS-CoV-2 detection by PCR, however, their viral loads were comparable to those of untreated participants and therefore no distinction was made between the treated and untreated participants in previous analyses [31]. Here, we continue to analyse the treated and untreated participants as a single cohort. Due to concentrations of viral RNA in the mid-turbinate remaining below the limit of quantification until day 6 and day 8, a further two participants are excluded since the model is not able to explain such long delays before viral growth. Although both nasal and throat swabs were collected, only measurements from the nasal swabs are used to parameterise our model.

The Bayesian inference provides us with a quantitative description of the in-host dynamics of SARS-CoV-2 virus. Model predictions for the infectious and total viral loads for 16 participants of the human challenge study are provided in Figure 1. From the posterior mixture distribution (Figure 2), we estimate that during the early stages of infection, productively infected cells are lost at an average rate of  $1.17 \text{ day}^{-1}$ . By day 14, once an adaptive immune response is more established, this rate increases to  $16.12 \text{ day}^{-1}$ . Individual infected cells produce infectious virus at rate  $3.74 \text{ PFU/ml day}^{-1}$  which is considerably lower than the production of viral RNA at rate  $2.57 \times 10^3 \text{ RNA copies/mL day}^{-1}$ . The within-host basic reproduction number represents the number of secondary cellular infections a single infected cells causes in a population of fully susceptible cells, and is given here by  $R_0 = \beta T(0) p_{\text{inf}} / \delta c_{\text{inf}}$ . Using the mixture distribution to derive an estimate of the within-host basic reproduction number gives  $R_0 = 11$ . Note that this estimate will likely decrease as the infection progresses due to the shortened lifespan of productively infected cells.

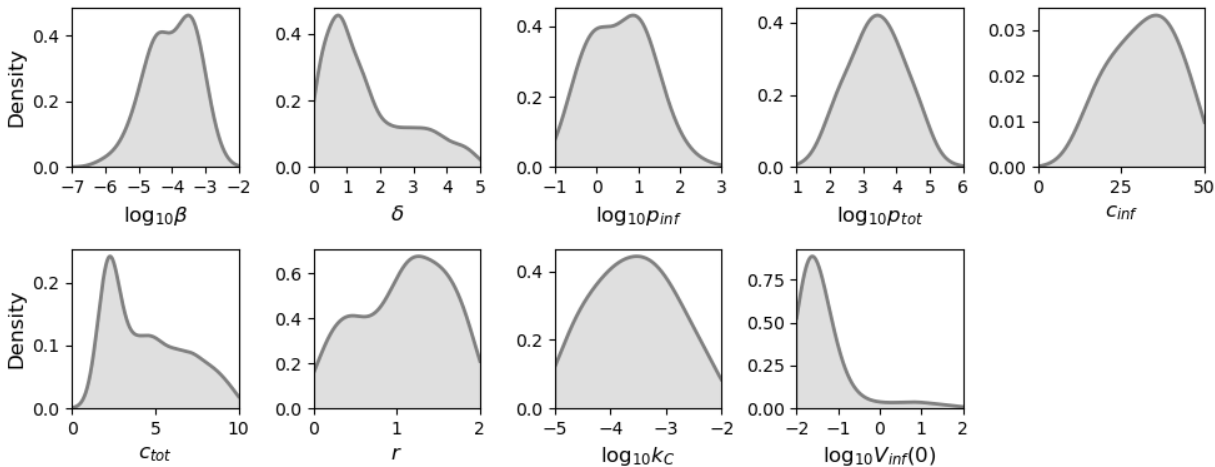
For many participants, the mathematical model predicts that infectious and total viral loads follow an initial exponential growth phase followed by an exponential decay phase. Since this behaviour can be captured by a model that excludes immune responses, a



**Figure 1: Posterior predictions:** Predicted viral dynamics of infectious virus (blue) and total virus (green) for participants of the human challenge study, constructed using approximate posterior samples from the ABC-SMC. Solid lines represent pointwise median predictions and shaded regions indicate 95% credible regions. Horizontal dashed lines indicate the limit of quantification.

model selection algorithm was used to determine if the inclusion of effector cells was necessary (see Methods). For six individuals, there was very strong evidence in favour of the model with the immune response, whilst for a further individual there was positive evidence (see Figure S2). For the remaining nine participants, there was not sufficient evidence to suggest that either model was preferable over the other. As a result of this, we believe that the inclusion of an adaptive immune response is appropriate.

Predictions from the mathematical model can also be used to estimate timescales of infection. Here, we define the time that an individual first tests positive by a PCR test as the time at which total virus ( $V_{\text{tot}}$ ) first exceeds the limit of quantification. Similarly, the first time a sample is culture positive is assumed to be equivalent to the first time the level of infectious virus ( $V_{\text{inf}}$ ) exceeds the limit of quantification, and can be thought of as the earliest time an individual is capable of transmitting the virus. The duration that an individual remains PCR (culture) positive is then defined as the number of days the

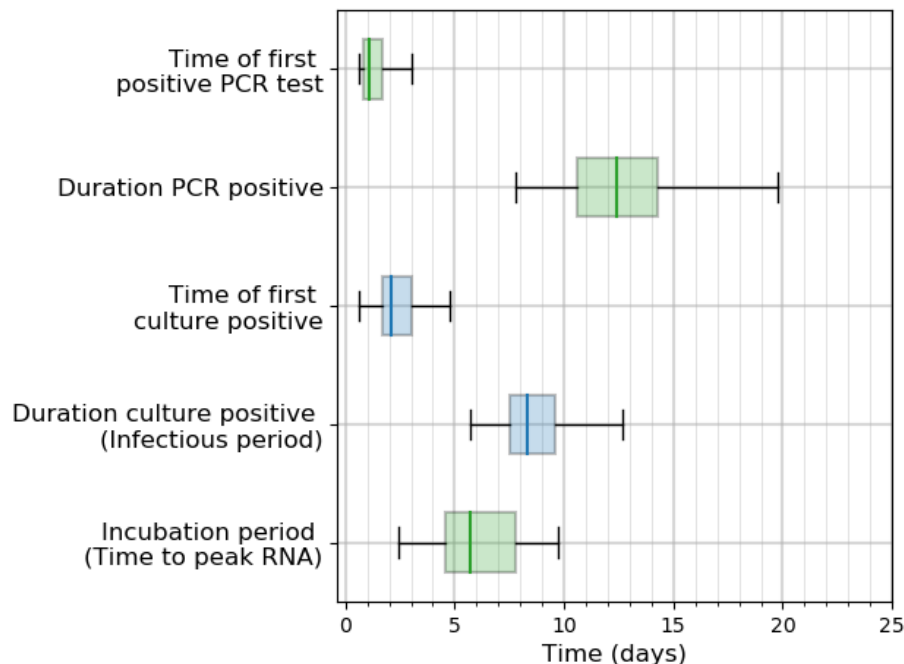


**Figure 2: Inferred parameter distributions for human SARS-CoV-2 infection:** Approximate posterior distributions for model parameters inferred using an ABC-SMC algorithm from mid-turbinate viral loads of human challenge study participants.

total (infectious) viral load continuously exceeds the limit of quantification. Estimates of these timescales are provided in Figure 3, along with an estimate of the incubation period, defined as the time to reach peak RNA copies/mL. Individuals are first PCR positive 1.1 days (interquartile range: 0.8–1.7 days) after infection and become infectious after 2.1 days (1.7–3 days). The incubation period lasts for 5.7 days (4.6–7.8 days) and is therefore consistent with the idea of pre-symptomatic transmission. Individuals remain infectious for 8.3 days (7.5–9.6 days) but would continue to test positive using a PCR test for 12.4 days (10.6–14.3 days). Some individuals remain PCR positive for up to 27 days following infection.

Whilst human challenge studies have many advantages, the strict screening process results in a relatively homogeneous population and so predictions from our model may not be representative of a more general population. For this reason, model predictions are validated against the Assessment of Transmission and Contagiousness of COVID-19 in Contacts (ATACCC) study, a longitudinal cohort study of community contacts of SARS-CoV-2 cases [32]. Whilst this study spanned multiple waves to capture the viral dynamics of unvaccinated and vaccinated individuals infected with different variants of SARS-CoV-2, we only consider those individuals who are unvaccinated and infected with a pre-alpha variant, since these will be the most similar to participants of the human challenge study.

To determine if our mathematical model can describe the viral dynamics of ATACCC participants, a brute force approach is used that initially involves sampling  $10^4$  parameter sets from the mixture distribution provided in Figure 2, in order to evaluate the model. Since participants are first contacted following symptom onset of the index case, the exact date of infection is unknown. Therefore, delays of up to 10 days between infection and enrolment are introduced to the model solutions in increments of 0.1 days. Solutions for  $\log_{10}(V_{\text{tot}})$  are then compared to log-RNA copies/mL from URT samples using the Euclidean distance, where only pairs of parameter sets and delays that yield the lowest 1% of distances are retained. These retained values are used to create the predictions in Figure



**Figure 3: Timescales of infection:** Estimates for the time of first positive test and the duration an individual remains positive. The timescales are derived from model predictions of both total (green) and infectious (blue) virus.

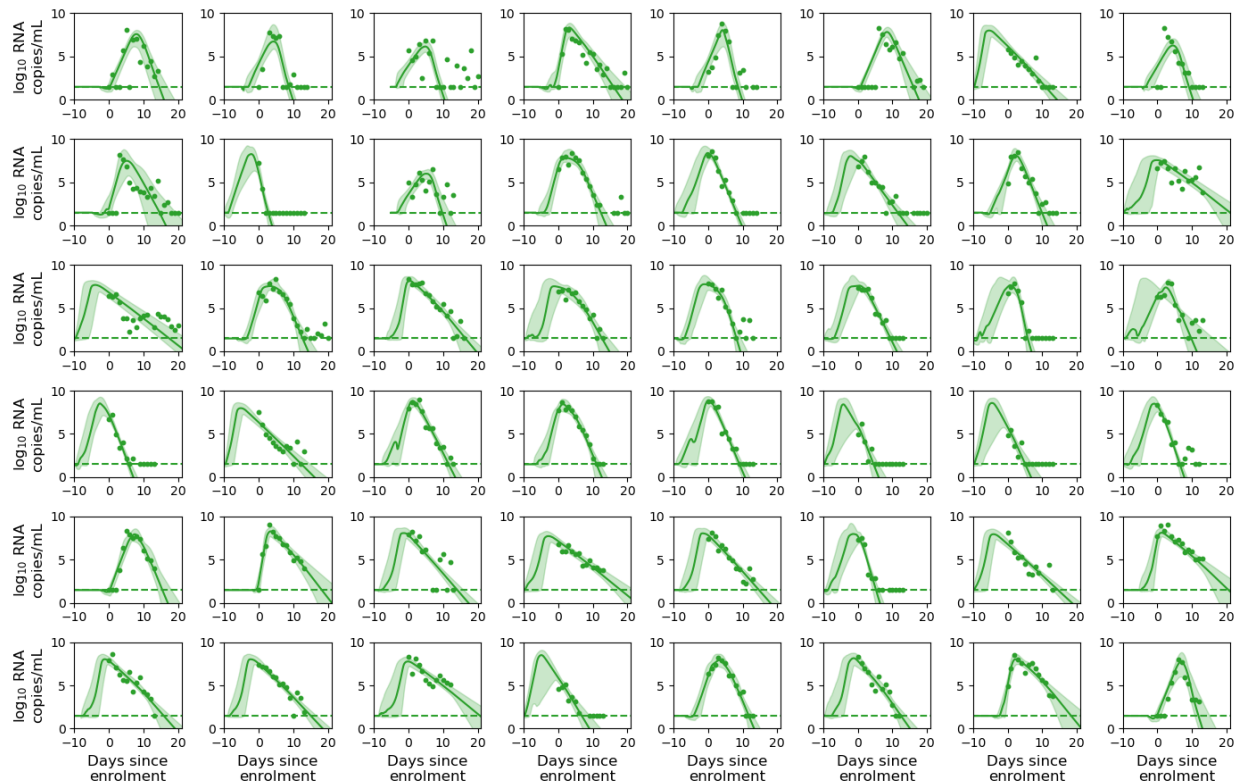
4. For the majority of ATACCC participants, the model and its existing parameterisation is able to explain the data well, suggesting that results derived from the human challenge study may be applicable to a general population. Furthermore, the median delay between infection and enrolment is 5.4 days which is plausible given that in most individuals we only observe the decay in the viral load and this decay begins around the time of symptom onset.

## Discussion

In this paper, we have used viral load measurements from a human challenge study to parameterise a mathematical model of SARS-CoV-2 infection. Whilst existing models have distinguished between infectious and non-infectious virus, to the best of our knowledge this is the first modelling study that makes use of infectious virus titres during parameter inference. Our estimate of the within-host basic reproduction number ( $R_0=11$ ) is similar to estimates from previous modelling studies (8.5–14.2), as is our estimate of 0.85 days for the lifespan of productively infected cells during the early stages of infection (0.53–1.66 days) [11, 17, 33].

A review of existing studies determined that the median duration of wild-type SARS-CoV-2 virus detection in the URT is 14.5 days following symptom onset [19]. Whilst this estimate falls within the range of our PCR-positive distribution, it is based off studies from early on in the pandemic with small sample sizes and is therefore unlikely to be robust.





**Figure 4: ATACCC validation:** Predicted URT viral loads for participants in the ATACCC study. Predictions are constructed using samples from the mixture distribution presented in Figure 2 and allowing for a delay between infection and enrolment of up to 10 days. Solid lines represent point-wise median predictions and shaded regions indicate 95% credible regions. Horizontal dashed lines indicate the limit of detection.

A more comparable study involved measuring the viral RNA trajectories of basketball players and estimated that acute infection lasted for 11.2 days in asymptomatic cases and 14.3 days in symptomatic cases, both of which are consistent with our predictions [34]. One limitation of our current analysis is that in some participants, RNA copies/mL remained high after the end of the 14-day study period. As a result, there is greater uncertainty in our predictions after this time which leads to greater uncertainty in the tail of the PCR positive distribution. In order to obtain accurate estimates for the duration of infection, regular sampling must be extended past 14 days.

Previous estimates of an individual's infectiousness have been based on the ability to successfully culture virus samples, however, such studies identify the presence of replicative competent virus and do not quantify the viral load [14]. Infectious virus may therefore be detectable but not at sufficient levels for someone to be infectious. As a result, estimates of the infectious period based on these studies are generally longer than that predicted here [14, 35]. A recent within-host model defines the infectious period to be the time when an individual's transmission probability exceeds 10% of its maximum value, obtaining estimates ranging between 1.9 days and 7.9 days that lie towards the lower end of our predicted distribution [18]. Ultimately, the duration of the infectious period depends

on how infectiousness is defined. Approaches based on viral load are preferable because they can be used to represent infectiousness on a continuous scale (such as a probability), whereas culture positivity only provides a binary outcome. Furthermore, viral load curves for infectious virus are beneficial as they provide a measure of the amount of virus capable of causing infection. Therefore, whilst our assumption that an individual is infectious if their infectious viral load exceeds 5 PFU/mL may be too simple, the predictions of infectious virus obtained here can be used alongside existing dose-response methods to provide more accurate estimates of infectiousness [8, 36].

Future work can use infectiousness curves as a link between within-host and population dynamics to develop more realistic multi-scale models of SARS-CoV-2 [9, 37]. With further information regarding T-cell responses and antibody levels following vaccination, such multi-scale models could be extended to consider the impact of waning immunity on susceptibility and transmission [38]. The results presented here are specific to wild-type SARS-CoV-2 virus, however, there is wide interest in variants of the virus since these have been associated with increased transmission and can rapidly spread throughout a population. One advantage of mechanistic models is that the parameters have simple interpretations, therefore, when the Delta variant reportedly showed a 10-fold increase in spike-mediated entry compared to a wild-type strain, this can be explicitly accounted for in the cellular infection rate [39]. The viral dynamics framework described here has also been extended to account for the genetic drift of viruses caused by mutations that occur during replication [40]. Applying this to SARS-CoV-2 virus may help understand the potential for new variants to arise within-host, with a recent report suggesting that the spike protein evolves to resist antibodies in immunocompromised individuals that have persistent infections [41]. Altogether, within-host models offer important insights into the interaction between SARS-CoV-2 virus and the host, however, it is the subsequent applications of these models that are most relevant for informing policy and decision making.

## Funding

This work was undertaken as part of the PROTECT COVID-19 National Core Studies, Theme 2, funded by HM Treasury.

## References

- [1] Davies NG, Kucharski AJ, Eggo RM, Gimma A, Edmunds WJ et al.. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. *The Lancet Public Health*. 2020; 5(7): e375–e385. [https://doi.org/10.1016/S2468-2667\(20\)30133-X](https://doi.org/10.1016/S2468-2667(20)30133-X)
- [2] Birrell P, Blake J, Van Leeuwen E, Gent N, De Angelis D. Real-time nowcasting and forecasting of COVID-19 dynamics in England: the first wave. *Philosophical Transactions of the Royal Society B*. 2021; 376(1829): 20200279. <https://doi.org/10.1098/rstb.2020.0279>

- [3] Giordano G, Blanchini F, Bruno R, Colaneri P, Di Filippo A, Di Matteo A, Colaneri M. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*. 2020; 26(6): 855–860. <https://doi.org/10.1038/s41591-020-0883-7>
- [4] Mossong K, Hens N, Jit M, Beutels P, Auranen K, et al.. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine*. 2008; 5(3): e74. <https://doi.org/10.1371/journal.pmed.0050074>
- [5] Jarvis CI, Van Zandvoort K, Gimma A, Prem K, Klepac P, Rubin GJ, Edmunds WJ. Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. *BMC Medicine*. 2020; 18(1): 1–10. <https://doi.org/10.1186/s12916-020-01597-8>
- [6] Handel A, Rohani P. Crossing the scale from within-host infection dynamics to between-host transmission fitness: a discussion of current assumptions and knowledge. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2015; 370(1675): 20140302. <https://doi.org/10.1098/rstb.2014.0302>
- [7] Marks M, Millat-Martinez P, Ouchi D, Roberts CH, Alemany A, Corbacho-Monné M et al.. Transmission of COVID-19 in 282 clusters in Catalonia, Spain: a cohort study. *The Lancet Infectious Diseases*. 2021; 21(5): 629–636. [https://doi.org/10.1016/S1473-3099\(20\)30985-3](https://doi.org/10.1016/S1473-3099(20)30985-3)
- [8] Chen SC, Chio CP, Jou LJ, Liao CM et al.. Viral kinetics and exhaled droplet size affect indoor transmission dynamics of influenza infection. *Indoor Air*. 2009; 19(5): 401. <https://doi.org/10.1111/j.1600-0668.2009.00603.x>
- [9] Handel A, Brown J, Stallknecht D, Rohani P. A multi-scale analysis of influenza A virus fitness trade-offs due to temperature-dependent virus persistence. *PLoS Computational Biology*. 2013; 9(3): e1002989. <https://doi.org/10.1371/journal.pcbi.1002989>
- [10] Challenger JD, Foo CY, Wu Y, Yan AWC, Marjaneh M, Liew F, Thwaites RS, Okell LC, Cunnington AJ. Modelling upper respiratory viral load dynamics of SARS-CoV-2. *BMC medicine*. 2022; 20(1): 1–20. <https://doi.org/10.1186/s12916-021-02220-0>
- [11] Gonçalves A, Bertrand J, Ke R, Comets E, De Lamballerie X, Malvy D, Pizzorno A et al.. Timing of antiviral treatment initiation is critical to reduce SARS-CoV-2 viral load. *CPT: Pharmacometrics & Systems Pharmacology*. 2020; 9(9): 509–514. <https://doi.org/10.1002/psp4.12543>
- [12] Hernandez-Vargas EA, Velasco-Hernandez JX. In-host mathematical modelling of COVID-19 in humans. *Annual Reviews in Control*. 2020; 50: 448–456. <https://doi.org/10.1016/j.arcontrol.2020.09.006>
- [13] Morawska L, Droplet fate in indoor environments, or can we prevent the spread of infection? *Indoor Air 2005: Proceedings of the 10th International Conference on*

- Indoor Air Quality and Climate. 2005; 16(5): 335–347. <https://doi.org/10.1111/j.1600-0668.2006.00432.x>
- [14] Walsh KA, Spillane S, Comber L, Cardwell K, Harrington P, Connell J et al.. The duration of infectiousness of individuals infected with SARS-CoV-2. *Journal of Infection*. 2020; 81(6): 847–856. <https://doi.org/10.1016/j.jinf.2020.10.009>
- [15] Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A. SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. *The Lancet Microbe*. 2021; 2(1): e13–e22. [https://doi.org/10.1016/S2666-5247\(20\)30172-5](https://doi.org/10.1016/S2666-5247(20)30172-5)
- [16] Goyal A, Reeves DB, Cardozo-Ojeda EF, Schiffer JT, Mayer BT. Viral load and contact heterogeneity predict SARS-CoV-2 transmission and super-spreading events. *ELife*. 2021; 10: e63537. <https://doi.org/10.7554/eLife.63537>
- [17] Ke R, Zitzmann C, Ribeiro RM, Perelson AS. Kinetics of SARS-CoV-2 infection in the human upper and lower respiratory tracts and their relationship with infectiousness. *MedRxiv*. 2020. <https://doi.org/10.1101/2020.09.25.20201772>
- [18] Ke R, Zitzmann C, Ho DD, Ribeiro RM, Perelson AS. In vivo kinetics of SARS-CoV-2 infection and its relationship with a persons infectiousness. *Proceedings of the National Academy of Sciences*. 2021; 118(49): e2111477118. <https://doi.org/10.1073/pnas.2111477118>
- [19] Walsh KA, Jordan K, Clyne B, Rohde D, Drummond L et al.. SARS-CoV-2 detection, viral load and infectivity over the course of an infection. *Journal of Infection*. 2020; 81(3): 357–371. <https://doi.org/10.1016/j.jinf.2020.06.067>
- [20] World Health Organization. Key criteria for the ethical acceptability of COVID-19 human challenge studies. 2020.
- [21] Darton TC, Blohmke CJ, Moorthy VS, Altmann DM, Hayden FG et al.. Design, recruitment, and microbiological considerations in human challenge studies. *The Lancet Infectious Diseases*. 2015; 15(7): 840–851. [http://dx.doi.org/10.1016/S1473-3099\(15\)00068-7](http://dx.doi.org/10.1016/S1473-3099(15)00068-7)
- [22] Liao LE, Carruthers J, Smither SJ, CL4 Virology Team, Weller SA, Williamson D, et al.. Quantification of Ebola virus replication kinetics in vitro. *PLoS Computational Biology*. 2020; 16(11): e1008375. <https://doi.org/10.1371/journal.pcbi.1008375>
- [23] Iwami S, Holder BP, Beauchemin CAA, Morita S, Tada T, Sato K, Igarashi T, Miura T. Quantification system for the viral dynamics of a highly pathogenic simian/human immunodeficiency virus based on an in vitro experiment and a mathematical model. *Retrovirology*. 2012; 9(1): 1–12. <https://doi.org/10.1186/PREACCEPT-8502330256154242>

- [24] Yan AWC, Zhou Jie, Beauchemin CAA, Russell CA, Barclay WS, Riley S. Quantifying mechanistic traits of influenza viral dynamics using in vitro data. *Epidemics*. 2020; 33: 100406. <https://doi.org/10.1016/j.epidem.2020.100406>
- [25] Li Y, Handel A. Modeling inoculum dose dependent patterns of acute virus infections. *Journal of Theoretical Biology*. 2014; 347: 63–73. <http://dx.doi.org/10.1016/j.jtbi.2014.01.008>
- [26] Minter A, Retkute R. Approximate Bayesian Computation for infectious disease modelling. *Epidemics*. 2019; 29: 100368. <https://doi.org/10.1016/j.epidem.2019.100368>.
- [27] Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface*. 2009; 6(31): 187–202. <https://doi.org/10.1098/rsif.2008.0172>
- [28] Filippi S, Barnes CP, Cornebise J, Stumpf MPH. On optimality of kernels for approximate Bayesian computation using sequential Monte Carlo. *Statistical applications in genetics and molecular biology*. 2013; 12(1): 87–107. <https://doi.org/10.1515/sagmb-2012-0069>
- [29] Simola U, Cisewski-Kehe, J, Gutman MU, Corander J. Adaptive approximate Bayesian computation tolerance selection. *Bayesian analysis*. 2021; 16(2): 397–423. <https://doi.org/10.1214/20-BA1211>
- [30] Gonçalves A, Maisonnasse P, Donati F, Albert M, Behillil S, Contreras V, et al.. SARS-CoV-2 viral dynamics in non-human primates. *PLoS Computational Biology*. 2021; 17(3): e1008785. <https://doi.org/10.1371/journal.pcbi.1008785>
- [31] Killingley B, Mann A, Kalinova M, Boyers A, Goonawardane N, Zhou J, Lindsell K et al.. Safety, tolerability and viral kinetics during SARS CoV-2 human challenge. *Research Square*. 2022. <https://doi.org/10.21203/rs.3.rs-1121993/v1>
- [32] Singanayagam A, Hakki S, Dunning J, Madon KJ, Crone MA, Koycheva A, Derqui-Fernandez N, Barnett JL, Whitfield MG, Varro R et al.. Community transmission and viral load kinetics of the SARS-CoV-2 delta (B. 1.617. 2) variant in vaccinated and unvaccinated individuals in the UK: a prospective, longitudinal, cohort study. *The Lancet Infectious Diseases*. 2021; 22: 183–195. [https://doi.org/10.1016/S1473-3099\(21\)00648-4](https://doi.org/10.1016/S1473-3099(21)00648-4)
- [33] Marc A, Kerioui M, Blanquart F, Bertrand J, Mitjà O, Corbacho-Monné M, Marks M, Guedj J. Quantifying the relationship between SARS-CoV-2 viral load and infectiousness. *Elife*. 2021; 10: e69302. <https://doi.org/10.7554/eLife.69302>
- [34] Kissler SM, Fauver JR, Mack C, Olesen SW, Tai C, Shiue KY et al.. Viral dynamics of acute SARS-CoV-2 infection and applications to diagnostic and public health strategies. *PLoS Biology*. 2021; 19(7): e3001333. <https://doi.org/10.1371/journal.pbio.3001333>

- [35] Singanayagam A, Patel M, Charlett A, Lopez Bernal J, Saliba V, Ellis J, Ladhani S, Zambon M, Gopal Robin. Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020. *Eurosurveillance*. 2020; 25(32): 2001483. <https://doi.org/10.2807/1560-7917.ES.2020.25.32.2001483>
- [36] Pratt A, Bennett E, Gillard J, Leach S, Hall I. Dose–Response Modeling: Extrapolating From Experimental Data to Real-World Populations. *Risk Analysis*. 2021; 41(1): 67–78. <https://doi.org/10.1111/risa.13597>
- [37] Forde JE, Ciupe SM. Quantification of the tradeoff between test sensitivity and test frequency in a COVID-19 epidemic—a multi-scale modeling approach. *Viruses*. 2021; 13(3): 457. <https://doi.org/10.3390/viruses13030457>
- [38] Handel A, Li Y, McKay B, Pawelek KA, Zarnitsyna V, Antia R. Exploring the impact of inoculum dose on host immunity and morbidity to inform model-based vaccine design. *PLoS Computational Biology*. 2018; 14(10): e1006505. <https://doi.org/10.1371/journal.pcbi.1006505>
- [39] Wang Y, Liu C, Zhang C, Wang Y, Hong Q, Xu S, Li Z, Yang Y, Huang Z, Cong Y. Structural basis for SARS-CoV-2 Delta variant recognition of ACE2 receptor and broadly neutralizing antibodies. *Nature Communications*. 2022; 13(1): 1–12. <https://doi.org/10.1038/s41467-022-28528-w>
- [40] Korobeinikov A. Immune response and within-host viral evolution: Immune response can accelerate evolution. *Journal of Theoretical Biology*. 2018; 456: 74–83. <https://doi.org/10.1016/j.jtbi.2018.08.003>
- [41] Clark SA, Clark LE, Pan J, Coscia A, McKay LGA, Shankar S et al.. SARS-CoV-2 evolution in an immunocompromised host reveals shared neutralization escape mechanisms. *Cell*. 2021; 184(10): 2605–2617. <https://doi.org/10.1016/j.cell.2021.03.027>