WP 2.1.1
Multi-compartment Within Host Model

Prepared for
The PROTECT COVID-19 National Core Study on transmission and environment

Report 2.1.1 (2023)
National Core Study Report

Prepared 2023
First published 2023

The PROTECT COVID-19 National Core Study on transmission and environment is a UK-wide research programme improving our understanding of how SARS-CoV-2 (the virus that causes COVID-19) is transmitted from person to person, and how this varies in different settings and environments. This improved understanding is enabling more effective measures to reduce transmission – saving lives and getting society back towards 'normal'.

This report is an extension of the work of PROTECT Phase 2 and covers the work done under phase 3 of PROTECT, in which researchers developed simplified models and introduced two different viral mechanisms and dose-response models. In this report, researchers extend these existing PROTECT models to account for location in body and infection site incorporating wider research.

This report documents research gaps in dose response and within host modelling. A framework is presented for the viral dynamics in different organs, which is left as a conceptual model due to the limitation of data. In ad-

dition, researchers adopt previously developed dose-response models and consider the infectiousness and infectious window. Good fits to the human challenge data from PROTECT can be seen in two-compartment within host models. For two-compartment models involving nose and throat, different scenarios have been developed. In each scenario, researchers observed the corresponding model can explain the available human challenge data. Whilst methods can be derived to explain data available, gaps remain in key processes such as transfer from host to contacts.

This report and the research it describes were funded by the PROTECT COVID- 19 National Core Study on transmission and environment, which is managed by the Health and Safety Executive (HSE) on behalf of HM Government. Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect UK Government or HSE policy.

Work Package 2.1.1 Multi-compartment Within Host Model

Jingsi Xu 1[1]
Jonathan Carruthers [2]
Thomas Finnie [2]
Ian Hall[1]
(Lead author: ian.hall@manchester.ac.uk**)**

[1] **University of Manchester**
Department of Mathematics
[2] **UK Health Security Agency**
PHAGE Joint Modelling Team

Executive Summary

Introduction

Within host model has been widely applied to study the viral mechanism of SARS-CoV-2 from the individual level. Studying viral dynamics in different body locations (i.e. organs) can be useful, especially in determining the mechanism of infection and evaluating infection severity.

Methodology

Based on the work of PROTECT Phase 2 (also see Xu et al. (2022)) and other literature, we develop meta-population models. These stochastic multi-compartment models include infection sites such as nose, throat, and lungs. Infection characteristics such as coughing and sneezing are simulated to see their impacts on viral load in each compartment. As with phase 2 work the data from the human challenge study Killingley et al. (2022) has been used to pursue model fitting. A framework is presented for the viral dynamics in different organs (eyes, nose, throat, lungs and stomach), which is left as a conceptual model due to the limitation of data. In addition we adopt the dose-response models developed in Xu et al. (2022) and consider the infectiousness and infectious window.

Conclusions

Good fits to the human challenge data from PROTECT can be seen in two-compartment within host models. Model selection cannot be decided on data calibration alone and requires other data and information. For multi-compartment models covering different organs, extra data is essential, hence, we only present it as a conceptual model. When standardised for 'distance' (i.e. for common scaling factor on shed viral load) the dose-response models under competing risk and logistic growth frameworks show different infectious windows. We also observe that in calibrating to the human challenge data a small number of participants shed majority of detectable virus suggesting these would be more efficient dis-

seminators of virus (though the potential spread is contingent on whether they would have met more people than others).

Key findings

- For two-compartment models involving nose and throat, different scenarios have been developed. In each scenario, we can see the corresponding model can explain the available human challenge data.
- Whilst methods can be derived to explain data available, gaps remain in key processes such as transfer from host to contacts requiring further integration of PROTECT research across themes.

PROTECT Phase 3 Final report of Work-package 2.1.1: Multi-compartment Within Host Model

Jingsi Xu 1[1]
Jonathan Carruthers [2]
Thomas Finnie [2]
Ian Hall[1]
(Lead author: ian.hall@manchester.ac.uk)

[1]Department of Mathematics, University of Manchester.
[2]PHAGE Joint Modelling Team, UK Health Security Agency

10 March 2023

1 Introduction

This report covers *Objective 2.1: Extension of in-host modelling and dose response*, incorporating emerging data and models from other studies. It is an extension of the work of PROTECT Phase 2 and covers the work done under phase 3 of PROTECT. In the previous year, we developed simplified models and introduced two different viral mechanisms Xu et al. (2022) based on the mechanism model developed in Goyal et al. (2021)(also see Goyal et al. (2020)). Also, we developed dose-response models under the competing risk framework (Haas et al. (2014)) and logistic framework used in literature such as Goyal et al. (2021) and Ke et al. (2021).

Here we extend these existing PROTECT models to account for location in body (local saturation of immune response) and infection site incorporating wider research. The report

documents research gaps in dose response and within host modelling and further work is part of existing project proposals and PhD studentships.

In recent years, a series of studies have adopted the within-host model to investigate the viral mechanism of SARS-CoV-2 infection. The work of Abuin et al. (2020) considers the simplest target cell limited model from the analytical perspective, which focuses on the equilibrium points with respect to virus extinction and its relations with the initial conditions. Similar analytical studies can also be seen in Nath et al. (2021) that pursue stability analysis and derive the within-host reproduction number for the simplest target cell limited model with the regeneration of targeted cells adopted in Li et al. (2020). In the mechanistic model of Challenger et al. (2022), the late immune response is triggered by a large magnitude of infected cells and reaches maturity via a series of equations, which does not wane once reaches maturity and provides accelerated clearance of infected cells. Gonçalves et al. (2021) assumes that the productively infected cells will produce infectious virus and non-infectious virus with probability. Gonçalves et al. (2021) further considers an additional model that incorporates the antigen-dependent immune response, however, this more complicated structure reduces the accuracy of parameter inference despite reduced BIC. Hernandez-Vargas & Velasco-Hernandez (2020) adopted several models to describe the viral load of hospital individuals from a study in Germany. These models include the simplest targeted cell limited model, its extension with an eclipse phase (also eg. Beauchemin et al. (2008), Holder et al. (2011), Madelain et al. (2018), Gonçalves et al. (2020), Wang et al. (2020), Baccam et al. (2006), and Smith & Perelson (2011)for similar application), and a model where the growth rate of virus depends on the virus density and maximum carry capacity of the virus. While adopting the targeted cell limited model with an eclipse, Néant et al. (2021) introduces an antigen-dependent response that enhances the clearance of productively infected cells and this allows the biphasic decay of viral load.

SARS-CoV-2 has been confirmed that can infect mucous membranes in direct or indirect ways Dawood (2021). The work of Wölfel et al. (2020) indicates that SARS-CoV-2 infects both the upper respiratory tract and lower respiratory tract. Hence, extending the within-host model to the multi-compartment case will help us further understand viral mechanisms in different organs and the migration between them. Dogra et al. (2020) develops a multi-compartment model to interpret the viral mechanism in different organs including the upper and lower respiratory tract, kidneys, brain, gastrointestinal tract, mononuclear phagocytic

system, heart, and plasma, in which the simplest SIV model is adopted to describe the viral dynamics in each compartment and plasma is different as it is only used to describe the viral migration between each compartment. Ke et al. (2020) develops a within-host model that describes the one-way transport of virus particles from upper respiratory tracts to lower respiratory tracts. Similar applications of multi-compartment within-host model can be seen in Li et al. (2022), and Afonyushkin et al. (2022).

In the work of PROTECT Phase 2 Xu et al. (2022), we have stated that in the model of Goyal et al. (2021), the infected cells are cleared by early immunity response and late T cell response, in which the late T cell response is driven by the SARS-CoV-2-specific effort cells raising from two stages of precursors cells. The SIV system fitted in Goyal et al. (2020) suggests viral load is quickly removed which suggests the viral load can be seen as a scaling of infected cells. Based on the observation, our previous work Xu et al. (2022) simplifies the model of Goyal et al. (2021) and extends it to two scenarios where the decay of the viral load is driven by the depletion of susceptible cells and adaptive response respectively, which reduces uncertainty in parameter inference and gives more insights into considering the viral shedding.

In Phase 3 of PROTECT, we aim to extend the simplified model in Xu et al. (2022), the one that interprets the decay of viral load by the depletion of susceptible cells, to different organs or regions of the body, including nasopharynx, oropharynx, lung, etc., and consider viral migration between each compartment. Viral load of the lung is correlated to the severity of lung damage and illness Williamson et al. (2020). Developing a multi-compartment model will help us better understand the viral load in each organ and evaluate the severity (i.e. if overwhelming infection locally, or moderate infection in a single organ, triggers adverse outcomes rather than global viral load).

In Section 2, we first extend the simplified model in Xu et al. (2022) to two compartments, which are nose and throat, and consider the two-way stochastic migration between nose and throat. This two-compartment model has been fitted to the nose and throat data from the Human Challenge Study (Killingley et al. (2022)). We introduce the compartment for the lung, in which we apply a branching process to present the viral load in the lung to avoid the unidentifiability issue due to the limited data. These model has been considered under different scenarios of coughing, inhalation, and sneezing. In Section 3, we present a conceptual model that describes the viral mechanism in the nose, throat, lungs, eyes,

and stomach. Since we only have data for the nose and throat, we can not make further progress on this model. In Section 4, we sample viral trajectory from the posterior prediction of Section 2 for each patient and recall the dose-response model developed in Xu et al. (2022) to investigate the infectiousness and infectious window.

2 Mathematical Analysis of Models

2.1 Multi-compartment within host model: nose and throat only

The model developed in Goyal et al. (2021) is based on the SIV equation system where the virus is cleared by innate immunity response and adaptive immunity response respectively, which has been studied and simplified in our previous study Xu et al. (2022). Hence, we have the simplified model:

$$\dot{S} = -\theta SI \tag{1}$$

$$\dot{I} = \theta SI - \delta I - \mu H(t-\tau)I \tag{2}$$

$$V = \omega I \tag{3}$$

where τ is the timescale for adaptive response to become effective and H is a Heaviside function (so zero when the argument is negative and 1 otherwise).

Note this is mathematically similar to a SIR compartment epidemic model Keeling & Rohani (2007) (but the R state is not relevant as the removal does not affect the replication process). V is then the free virus that may be emitted (and may be calibrated to the human challenge data). If we assume this virus is infectious material then the V state can directly be fitted to the PFU dataset.

Based upon Goyal et al. (2021) and Xu et al. (2022), we further develop a meta-population model to consider the viral shedding in the nose and throat and the migration between two compartments. In the following part, we will consider different scenarios to fully understand this model.

2.2 Scenario 1: Variant susceptibility model (different values of ρ)

According to the simplified model (1), we can develop a meta-population model, which is given by:

$$\begin{aligned} dS_1 &= -\theta_1 S_1 I_1 \\ dI_1 &= \theta_1 S_1 I_1 - \delta I_1 - \mu H(t-\tau) I_1 - \alpha_1 I_1 + \alpha_2 I_2 \\ V_2 &= \omega_1 I_1 \\ dS_2 &= -\theta_2 S_2 I_2 \\ dI_2 &= \theta_2 S_2 I_2 - \delta I_2 - \mu H(t-\tau) I_2 + \alpha_1 I_1 - \alpha_2 I_2 \\ V_2 &= \omega_2 I_2 \end{aligned} \tag{4}$$

In this model, S_1 and S_2 are the susceptible cells in nose and throat respectively, and $I_i, i = 1, 2$ are infected cells in the corresponding compartment. Parameters $\theta_i, i = 1, 2$ represents the infection rate of susceptible cells $S_i, i = 1, 2$ respectively. The infected cells are cleared with rate δ by early immunity response and by adaptive immunity response with rate $\mu H(t-\tau)$. In this situation, we assume that both early immunity respnse and adaptive immunity response are consistent for nose and throat. To non-dimensionalize the model, we assume $S_i = NX_i$, $I_i = NY_i$, and $V_i = \omega_i NY_i = \phi_i Y$ where $i = 1, 2$. Hence, after non-dimensionalisation, we have:

$$\begin{aligned} dX_1 &= -\frac{\delta \rho_1}{\phi_1} X_1 V_1 \\ dV_1 &= \delta \rho_1 X_1 V_1 - \delta V_1 - \mu H(t-\tau) V_1 - \alpha_1 V_1 + \alpha_2 \frac{\phi_1}{\phi_2} V_2 \\ dX_2 &= -\frac{\delta \rho_2}{\phi_2} X_2 V_2 \\ dV_2 &= \delta \rho_2 X_2 V_2 - \delta V_2 - \mu H(t-\tau) V_2 + \alpha_1 \frac{\phi_2}{\phi_1} V_1 - \alpha_2 V_2 \end{aligned} \tag{5}$$

where $\rho_i = \frac{\theta_i N}{\delta}$ for $i = 1, 2$. From equation (5), we note that $\alpha_1 V_1 = \alpha_1 \frac{\phi_2}{\phi_1} V_1$ if and only if $\phi_1 = \phi_2$, which leads to:

$$
\begin{aligned}
dX_1 &= -\frac{\delta \rho_1}{\phi} X_1 V_1 \\
dV_1 &= \delta \rho_1 X_1 V_1 - \delta V_1 - \mu H(t-\tau) V_1 - \alpha_1 V_1 + \alpha_2 V_2 \\
dX_2 &= -\frac{\delta \rho_2}{\phi} X_2 V_2 \\
dV_2 &= \delta \rho_2 X_2 V_2 - \delta V_2 - \mu H(t-\tau) V_2 + \alpha_1 V_1 - \alpha_2 V_2
\end{aligned}
\tag{6}
$$

in which ρ_i, $i = 1, 2$, includes the information of variant susceptibility, i.e. different growth rates, of two compartments respectively. Note that δ represents the viral clearance rate by innate immunity response and we assume that it is homogeneous for both compartments. Based on model (6), we further introduce stochasticity into the migration between compartments:

$$
\begin{aligned}
dX_1 &= -\frac{\delta \rho_1}{\phi} X_1 V_1 \\
dV_1 &= (\delta \rho_1 X_1 V_1 - \delta V_1 - \mu H(t-\tau) V_1 - \alpha_1 V_1 + \alpha_2 V_2) dt - \sqrt{\alpha_{12} V_1} dW_t^1 + \sqrt{\alpha_{21} V_2} dW_t^2 \\
dX_2 &= -\frac{\delta \rho_2}{\phi} X_2 V_2 \\
dV_2 &= (\delta \rho_2 X_2 V_2 - \delta V_2 - \mu H(t-\tau) V_2 + \alpha_1 V_1 - \alpha_2 V_2) dt + \sqrt{\alpha_{12} V_1} dW_t^1 - \sqrt{\alpha_{21} V_2} dW_t^2
\end{aligned}
\tag{7}
$$

We use the throat and nose data from the human challenge study Killingley et al. (2022) to pursue parameter inference. For model (7), Figure 1 shows the corresponding model fits for both nose and throat data for the 17 volunteers whilst Figure 2 shows the participant merged posterior distribution plots for each of the parameters of model (7).

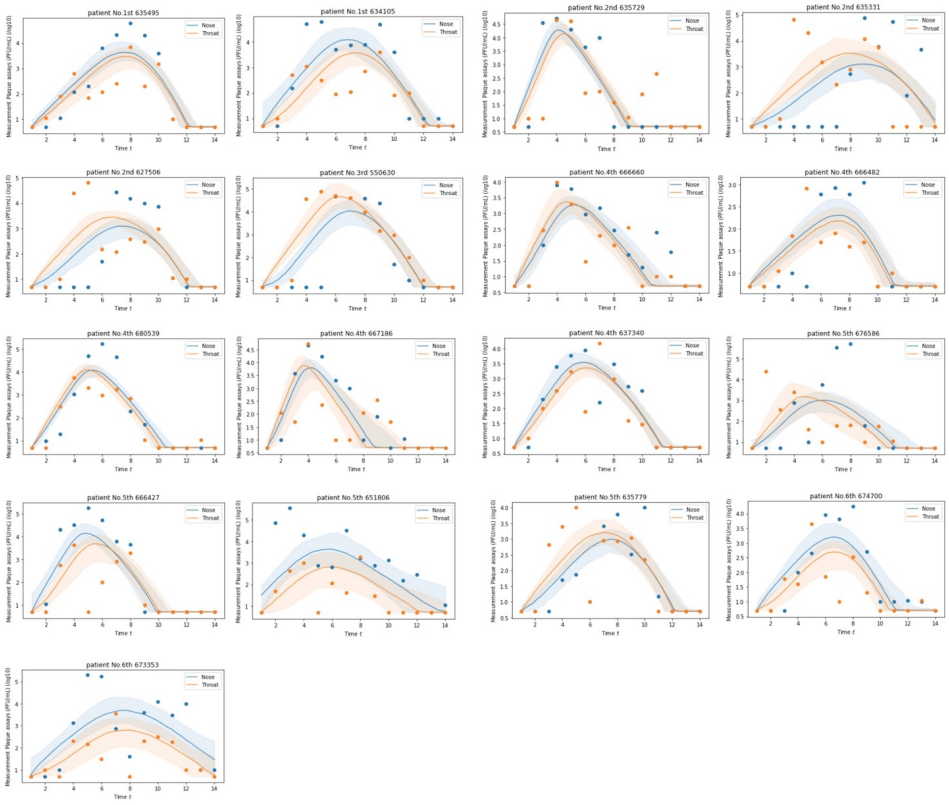

Figure 1: (Scenario 1) Posterior predictions of infectious virus (PFU/mL) from mid-turbinate and throat for 17 participants from the Human Challenge Study. Model is that derived in equation (7). Shaded regions represent the 95% credible interval.

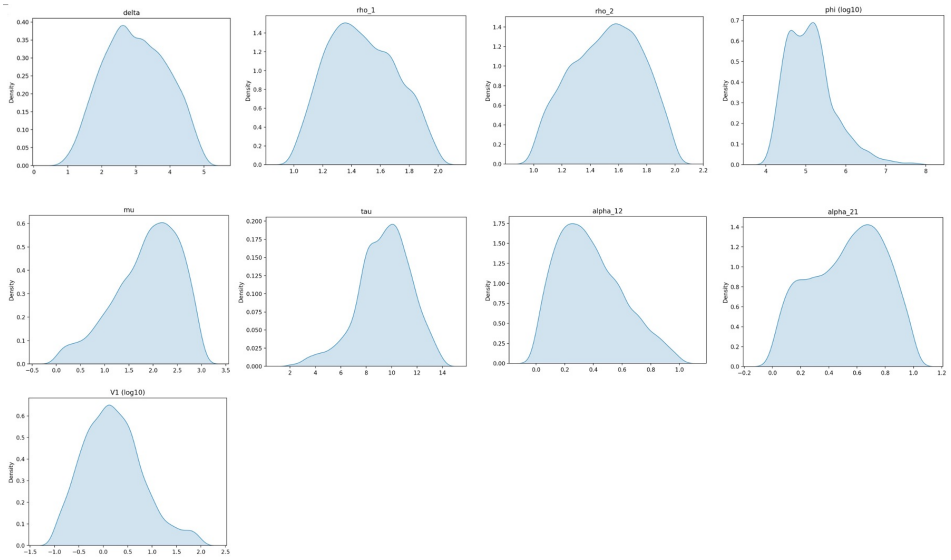

Figure 2: (Scenario 1) Approximate participant-merged Posterior distribution of parameter values from the result of ABC-SMC with model from equation (7) using both mid-turbinate and throat data

2.3 Scenario 2: dimensionalised migration between compartment (different values of ϕ)

In the simplified model, the super-parameter ,ϕ, represents the magnitude of viral load, and it is reasonable to consider that different compartments will have different values of ϕ. Hence, to further investigate this situation, we assume that the migration between two compartments are dimensionalised. Re-calling model (1), we can construct that:

$$
\begin{aligned}
dS_1 &= -\theta_1 S_1 I_1 \\
dI_1 &= \theta_1 S_1 I_1 - \delta I_1 - \mu H(t-\tau) I_1 \\
V_2 &= \omega_1 I_1 \\
dS_2 &= -\theta_2 S_2 I_2 \\
dI_2 &= \theta_2 S_2 I_2 - \delta I_2 - \mu H(t-\tau) I_2 \\
V_2 &= \omega_2 I_2
\end{aligned}
\tag{8}
$$

To non-dimensionalize the model, we assume $S_i = NX_i$, $I_i = NY_i$, and $V_i = \omega_i NY_i$ where $i = 1, 2$. In this case, we assume that $\phi_i = \omega_i N, i = 1, 2$ and $\phi_1 \neq \phi_2$. Hence, after non-dimensionalisation, we have:

$$
\begin{aligned}
dX_1 &= -\frac{\delta\rho}{\phi_1} X_1 V_1 \\
dV_1 &= (\delta\rho X_1 V_1 - \delta V_1 - \mu H(t-\tau) V_1 \\
dX_2 &= -\frac{\delta\rho}{\phi_2} X_2 V_2 \\
dV_2 &= (\delta\rho X_2 V_2 - \delta V_2 - \mu H(t-\tau) V_2
\end{aligned} \tag{9}
$$

Based on equation (9), we further introduced viral migration between each compartment, which is dimensionalised in this case, and stochastic noise into this model. The model is given by:

$$
\begin{aligned}
dX_1 &= -\frac{\delta\rho}{\phi_1} X_1 V_1 \\
dV_1 &= (\delta\rho X_1 V_1 - \delta V_1 - \mu H(t-\tau) V_1 - \alpha_{12} V_1 + \alpha_{21} V_2) dt - \sqrt{\alpha_{12} V_1} dW_t + \sqrt{\alpha_{21} V_2} dW_t \\
dX_2 &= -\frac{\delta\rho}{\phi_2} X_2 V_2 \\
dV_2 &= (\delta\rho X_2 V_2 - \delta V_2 - \mu H(t-\tau) V_2 + \alpha_{12} V_1 - \alpha_{21} V_2) dt + \sqrt{\alpha_{12} V_1} dW_t - \sqrt{\alpha_{21} V_2} dW_t
\end{aligned} \tag{10}
$$

Different from model (7), we assume that ϕ takes different values for nose and throat compartments in this model, which determines the magnitude of the peak of viral load in different compartments.

2.4 Scenario 3: dimensionalised migration between compartment (different values of ϕ) with coughing and sneezing

Cough and sneeze are two of the major symptoms of SARS-CoV-2 infection and are playing a significant role in transmission. Studies have linked the viral dynamic to aerosol transmission Heitzman-Breen & Ciupe (2022). Hence, in this scenario, we introduce random events to

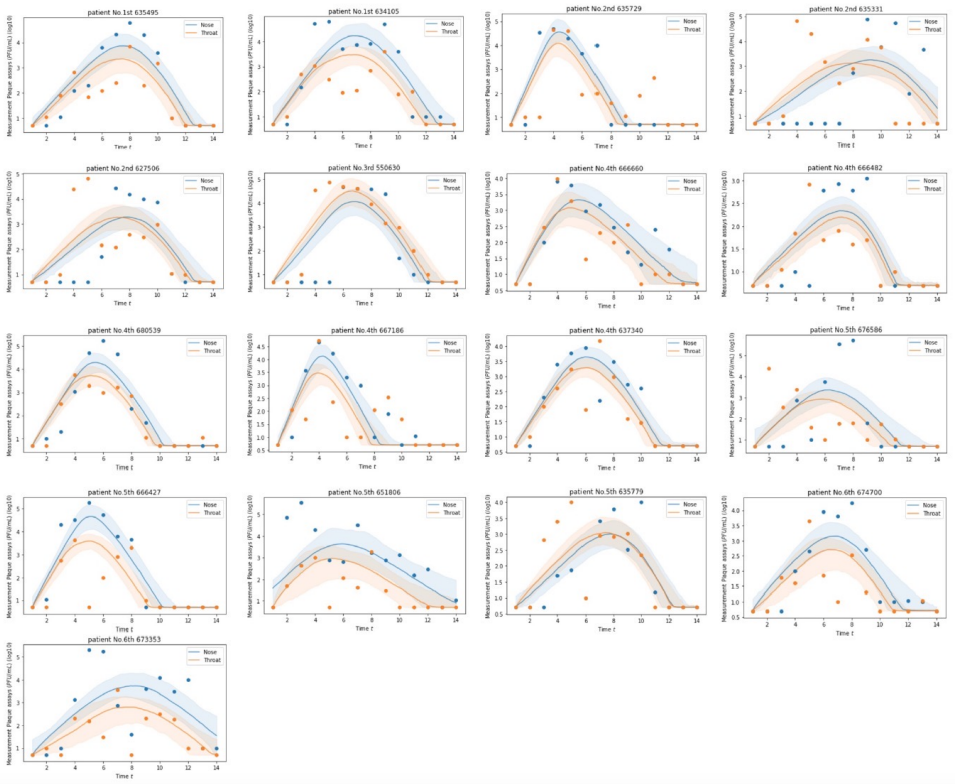

Figure 3: (Scenario 2) Posterior predictions of infectious virus (PFU/mL) from mid-turbinate and throat for 17 participants from the Human Challenge Study. Model is that derived in equation (10). Shaded regions represent the 95% credible interval. In this figure, the values of phi_1 and phi_2 are different.

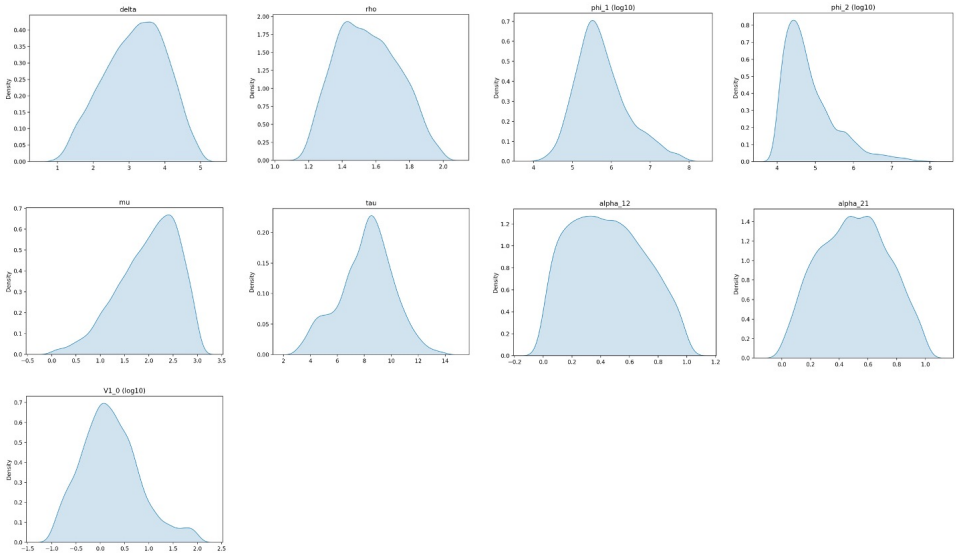

Figure 4: (Scenario 2) Approximate participant-merged Posterior distribution of parameter values from the result of ABC-SMC with model from equation (10) using both mid-turbinate and throat data

model (10) aiming to consider viral dynamics when there is viral shedding caused by cough and sneeze. Simulating cough and sneeze can be challenging as there is insufficient evidence to inform how much virus will release when the patient sneezes or coughs and how frequent the cough and sneeze can be. In striving for a reasonable simulation, we consider the work of Kelsall et al. (2009) and Lee et al. (2013) which studies chronic cough. Both studies give 10 to 20 coughs per hour. However, it should be noticed that coughs are unlikely to be evenly distributed as patients tend to cough multiple times in a short period of time. Also, Kelsall et al. (2009) indicates that males and females have different cough frequencies when the cough is chronic. The latter paper Lee et al. (2013) suggests that patient would have 85 coughing bouts per day, which may be more evenly distributed over waking hours, and so a Poisson distribution with a variable rate of 40 to 200 per day might be a reasonable place to start. Compared with cough, there is no study to suggest a proper sneeze rate.

Hence, to simulate the random events of cough and sneeze, we assume that the frequency of cough and sneezing are driven by Poisson distribution with different values of mean, and when the patient will shed a certain percentage of viral load from nose and throat respectively. The fitting results of this scenario are illustrated in Figure 5 and 6. Remember that this

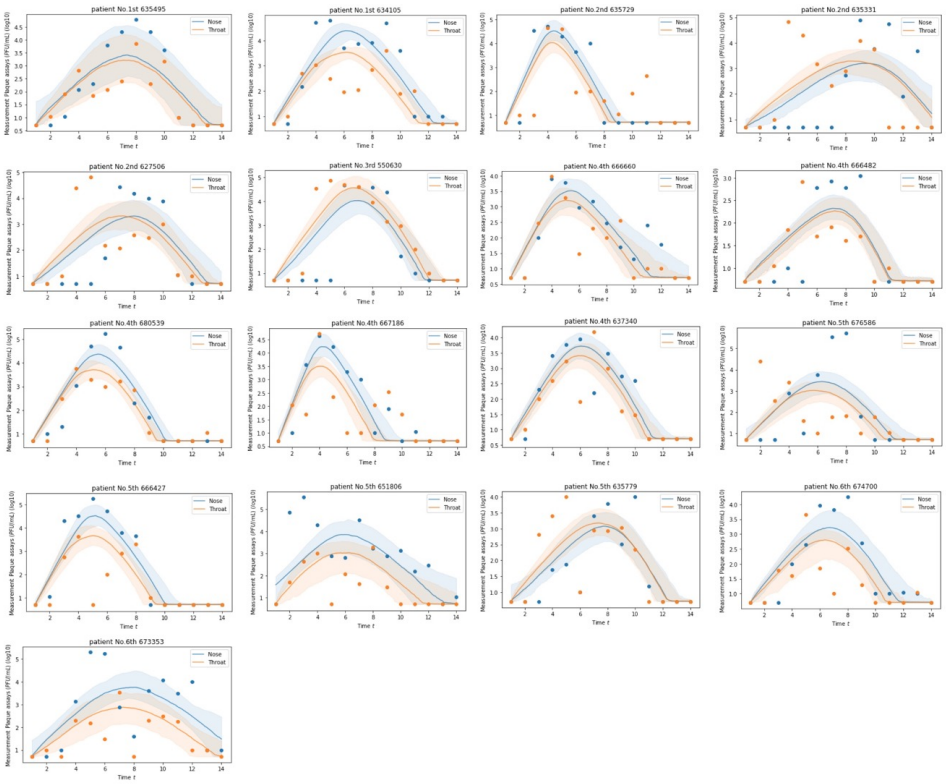

Figure 5: (Scenario 3) Posterior predictions of infectious virus (PFU/mL) from mid-turbinate and throat for 17 participants from the Human Challenge Study. Model is that derived in equation (10) by adding random events: cough and sneeze during the time. Shaded regions represent the 95% credible interval. In this figure, the values of ϕ_1 and ϕ_2 are different.

is based on studies investigating chronic cough which may be different from the acute respiratory cough of SARS-CoV-2, there is no data on sneezing rate nor emmission volume. Also note that some cases are largely asymptotmatic and so coughing and sneezing is not universally reported in clinical symptoms. As such this is presented as a conceptual framework in the absence of further data.

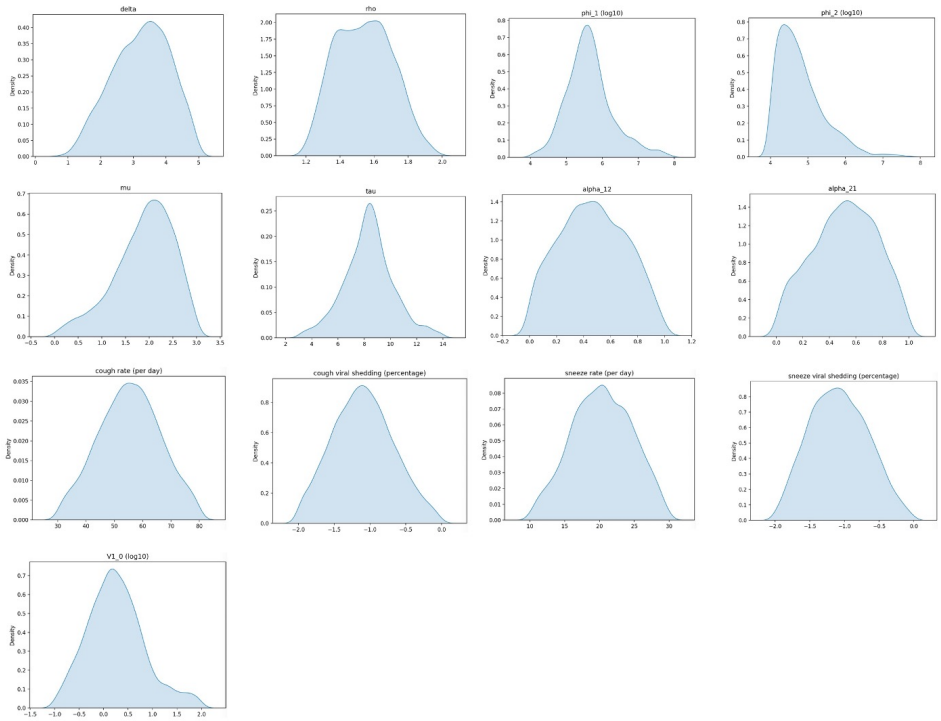

Figure 6: (Scenario 3) Approximate participant-merged Posterior distribution of parameter values from the result of ABC-SMC with model from equation (10) by adding random events: cough and sneeze during the time and using both mid-turbinate and throat data

2.5 Scenario 4: three compartment models under different values of ϕ

Studies report symptoms of SARS-CoV-2 beyond coughing and sneezing, such as fatigue, fever and following the invasion of virus into the lung space difficulty of breathing, pneumonia and even death of Canada (2020). Hence, understanding the viral dynamics in lungs is likely important. Since the human challenge study does not contain data from lower respiratory tract, we aim to develop a simple model with minimal new parameters to avoid unidentifibility issue. Assume that there are always substantial susceptible cells in lungs before the adaptive immunity kicks in and there is no susceptible cells after activating the adaptive immunity response, which means $X_t = 1$ for $0 \leq t \leq \tau$ and $X_t = 0$ for $\tau < t$. We construct a branching process to describe the viral dynamics in lungs, and the model is given by:

$$
\begin{aligned}
dX_1 &= -\frac{\delta\rho}{\phi_1} X_1 V_1 \\
dV_1 &= (\delta\rho X_1 V_1 - \delta V_1 - \mu H(t-\tau) V_1 - \alpha_{12} V_1 + \alpha_{21} V_2)dt - \sqrt{\alpha_{12} V_1}\, dW_t + \sqrt{\alpha_{21} V_2}\, dW_t \\
dX_2 &= -\frac{\delta\rho}{\phi_2} X_2 V_2 \\
dV_2 &= (\delta\rho X_2 V_2 - \delta V_2 - \mu H(t-\tau) V_2 + \alpha_{12} V_1 - \alpha_{21} V_2)dt + \sqrt{\alpha_{12} V_1}\, dW_t - \sqrt{\alpha_{21} V_2}\, dW_t \\
dV_3 &= \begin{cases} (\delta\rho V_3 - \delta V_3 + \alpha_{23} V_2 - \alpha_{32} V_3)dt + \sqrt{\alpha_{23} V_2}\, dW_t - \sqrt{\alpha_{32} V_3}\, dW_t & \text{if } t < \tau \\ (-\delta V_3 - \mu V_3 + \alpha_{23} V_2 - \alpha_{32} V_3)dt + \sqrt{\alpha_{23} V_2}\, dW_t - \sqrt{\alpha_{32} V_3}\, dW_t & \text{if } t \geq \tau \end{cases}
\end{aligned}
\tag{11}
$$

One advantage of model (11) is that it does not involves unidentifiable parameters. The fitting results are shown in Figure 7 and 8

2.6 Model calibration method

Approximate Bayesian Computation Sequential Monte Carlo (ABC-SMC) algorithm is adopted to conduct parameter inference (cf. Toni et al. (2009), and Minter & Retkute (2019)), in which a multivariate normal distribution with optimal local covariance matrix is used as the perturbation kernel Filippi et al. (2013). For ABC-SMC algorithm, the distance function is defined

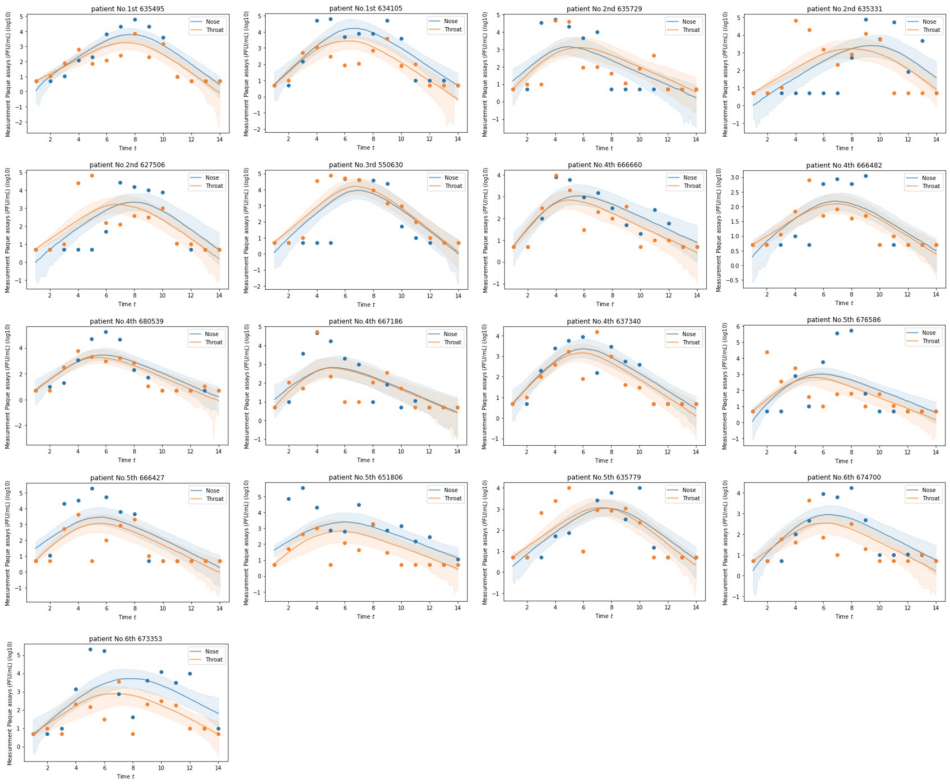

Figure 7: (Scenario 4) Posterior predictions of infectious virus (PFU/mL) from mid-turbinate and throat for 17 participants from the Human Challenge Study. Model is that derived in equation (11). Shaded regions represent the 95% credible interval. In this figure, the values of ϕ_1 and ϕ_2 are different.

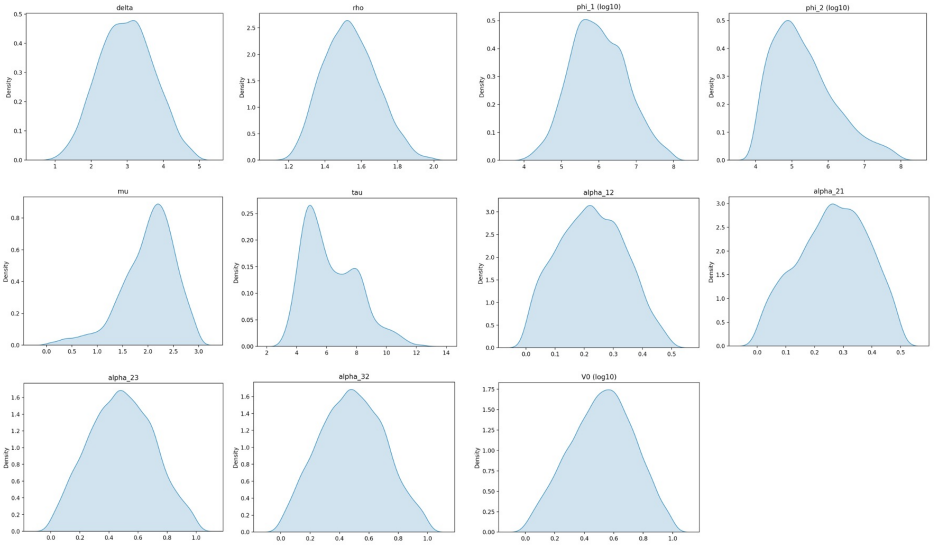

Figure 8: (Scenario 4) Approximate participant-merged Posterior distribution of parameter values from the result of ABC-SMC with model from equation (11) by adding random events: cough and sneeze during the time and using both mid-turbinate and throat data

as:

$$d(M,O)^2 = \sum_{t\in T}\left[\left(log_{10}V_1^{(O)}(t) - log_{10}V_1^{(M)}(t)\right)^2 + \left(log_{10}V_2^{(O)}(t) - log_{10}V_2^{(M)}(t)\right)^2\right]$$

in which T represents the set of time points of the data, and $V_i^{(O)}$ and $V_i^{(M)}$, represent the observations and simulated outputs respectively, and subscript $i = 1, 2$ represent the nose and throat data. For each generation of iterations, 250 particles are collected and the number of generations is 5 for improving accuracy. For each generation, we arrange the values from the cost function from smallest to largest and set the value of the 1st quartile as the tolerance level for the next generation.

In this section, we considered four scenarios, and for most of parameters, the corresponding prior distributions do not change in different scenarios. The prior distributions are given by:

$$\delta \sim U(0,5) \qquad \rho_{1,2} \sim U(1,2) \qquad log_{10}\phi_{1,2} \sim U(4,8) \qquad \mu \sim U(0,3)$$
$$\tau \sim U(\min(\tau_{nose}^{peak}, \tau_{throat}^{peak}), 14) \qquad \alpha_i \sim U(0,1) \qquad \log_{10} V_0 \sim U(-1,2)$$

Note that τ^{peak} represents the point of time of peak viral load and α_i represents the migration between each compartment. For cough and sneeze in Scenario 3, we assume they are driven by Poisson distribution with different values of mean, which are 80 and 30 respectively. The data from Human Challenge Study has been used to carry out parameter inference Killingley et al. (2022). Posterior distributions and predictions for these four scenarios can be seen from Figure 1 to 8.

2.7 Conceptual model for multi-compartments: nose, throat, eyes, lung, and stomach

As we mentioned in above, a within-host model that can predict the viral trajectory of different organs such as lungs can help us efficiently evaluate the severity. Hence, based on Scenario 2 in the previous section, we further develop a model that contains five compartments, eyes, nose, throat, lungs and stomach, and assume there exist stochastic viral migration between each compartment. The conceptual model is given by:

$$
\begin{aligned}
dX_n &= -\frac{\delta\rho}{\phi_n} X_n V_n \\
dV_n &= \delta\rho X_n V_n - \delta V_n - \mu H(t-\tau) V_n + M_1(V_n, V_e) + M_2(V_n, V_{th}) \\
dX_{th} &= -\frac{\delta\rho}{\phi_{th}} X_{th} V_{th} \\
dV_{th} &= \delta\rho X_{th} V_{th} - \delta V_{th} - \mu H(t-\tau) V_{th} - M_2(V_n, V_{th}) + M_3(V_{th}, V_l) - M_{sto}(V_{th}) \\
dX_e &= -\frac{\delta\rho}{\phi_e} X_e V_e \\
dV_e &= \delta\rho X_e V_e - \delta V_e - \mu H(t-\tau) V_e - M_1(V_n, V_e) \\
dX_l &= -\frac{\delta\rho}{\phi_l} X_l V_l \\
dV_l &= \delta\rho X_l V_l - \delta V_l - \mu H(t-\tau) V_l - M_3(V_{th}, V_l) \\
dX_{sto} &= -\frac{\delta\rho}{\phi_{sto}} X_{sto} V_{sto} \\
dV_{sto} &= \delta\rho X_{sto} V_{sto} - \delta V_{sto} - \mu H(t-\tau) V_{sto}
\end{aligned}
\tag{12}
$$

M_i functions represent migration between different compartments, which are given by:

$$
\begin{aligned}
M_1(V_n, V_e) &= -\alpha_{n\hookrightarrow e}V_n + \alpha_{e\hookrightarrow n}V_e \\
M_2(V_n, V_{th}) &= -\alpha_{n\hookrightarrow th}V_n + \alpha_{th\hookrightarrow n}V_{th} \\
M_3(V_{th}, V_l) &= -\alpha_{th\hookrightarrow l}V_{th} + \alpha_{l\hookrightarrow th}V_l \\
M_{sto}(Vth) &= -\alpha_{th\hookrightarrow sto}V_{th}
\end{aligned}
$$

In this model, the subscript $i = n, th, e, l,$ and sto of X and V represents the susceptible cells and viral load in nose, throat, eyes, lungs, and stomach. In this model, parameters μ, δ and τ retain their biological interpretation of viral kill rates and timescale for adaptive response. ρ and ϕ are super parameters, which contain implicitly information of the virus. We assume that for different compartments, ρ, δ,μ,τ are consistent and ϕ will vary. Note that we assume there is only one way migration between throat and stomach, and there is no migration between stomach and lungs. Since we only have the throat and nose data from human challenge study Killingley et al. (2022), it is impossible for us to verify this model as most of the parameters will be unidentifiable. Hence, we only present equation 12 as an conceptual model for future study.

3 Dose-response model and Probability of Infection

Dose-response models have been playing an important role in mapping viral load over time to infectiousness. The work of Protect Phase 2 Xu et al. (2022) presents three dose-response models under two mechanisms, the competing risk framework based on Haas et al. (2014) and the logistic growth framework used in viral load models such as Goyal et al. (2021), Heitzman-Breen & Ciupe (2022), and Ke et al. (2021). In the following part, we will first introduce the dose-response models from Xu et al. (2022) briefly and consider the infectiousness.

Assume that the expected inhaled dose of an exposed contact is $x = \beta v$, in which β is a scaling factor measuring the reduction in viral load due to environmental factors and v is the viral load estimated from the within-host model representing the viral load from a transmitter. In this situation, the chance of the exposed contact to getting infected driven by the exponential

dose-response model is given by:

$$p(x; d_{50}) = 1 - Q = 1 - \exp\left(-\ln(2)x/d_{50}\right) \tag{13}$$

in which we assume that d_{50}, the median infecting dose, equals to 10TCID50=50PFU. Moreover, for the approximate beta-Poisson dose-response model, the probability of infection is defined by:

$$p_{\mathrm{ABP}}(x; \nu, d_{50}) = 1 - \left(1 + \left(2^{1/\nu} - 1\right)\frac{x}{d_{50}}\right)^{-\nu}. \tag{14}$$

Note that in Equation (14), the new parameter ν acts to model between host variation in response. If $\nu \gg 1$ then $p_{\mathrm{ABP}}(x; \nu, d_{50}) \to p(x; d_{50})$. This is used to motivate further thinking on dose and we prefer the use of alternate functions such as in Pratt et al. (2020).

Aside from the competing risk framework, the logistic dose-response model has been used in literature. Recalling Xu et al. (2022), the logistic dose-response model is given by:

$$p_{\mathrm{L}}(x) = \frac{1}{1 + (d_{50}/x)^{\eta}} \tag{15}$$

so as $x \to \infty$ then $p_L(x) \to 1$ and when $x = 0$ we have $p_L(0) = 0$. In the case when $\eta = \nu = 1$, we see that equations (14) and (15) are identical. Further details can be seen in Xu et al. (2022).

3.1 Results of linking in host models and dose-response models

To achieve results from the dose-response models, the key step is choosing a reasonable proxy for ID50. In the work of Killingley et al. (2022), 53% of patients developed PCR-confirmed infection under the inoculation of 10TCID50 (TCID50 is the median tissue culture infectious dose) of SARS-CoV-2 with 95% confidence interval is (35,70). Hence, in the following part, we will set 10TCID50=50PFU. To illustrate the result, we randomly sample one viral trajectory from the posterior predictions generated from fit to mid-turbinate data of each of the 17 patients (using model described in Scenario 2 of Section 2). We further assume that the scaling factor $\beta = 0.001$ in the following part. Also, we set $\nu = 0.1$ for the approximate beta-Poisson dose-response model 14 and $\eta = 1$ for the logistic dose-response model 15 as illustrative values noting that $\nu = 1$ choice means this is equivalent to and ABP model with

$\eta = 1$.

In Figure 9, we see that for most of patients, their maximum probability of infection during the entire 14 days will be less than or equal to 20%, however, 2 of 17 patients show a much higher probability to infect exposed contacts. Figure 10 demonstrates the impact of host variation in the model; compared with Figure 9, with patients showing higher infectiousness. The maximum probability of infection of most patients is over 30%. For each patient, they show a longer infectious window and most patients still have more than 10% of chance to infect exposed contact. This observation indicates that the variation at a host level in deposition, germination and removal rates does have impacts on infectiousness. The result from Figure 15 is very similar to Figure 13, suggesting that the ν value would have to be made smaller and we will omit further details in this case.

Studies He et al. (2020) and Wölfel et al. (2020) have shown that SARS-CoV-2 can reach high viral load and transmit widely, which is consistent with results from our does-response model. In the dose-response model, viral load peak naturally leads to the highest probability of infection, which may lead to substantial secondary infection. Also, dose-response models can help us to estimate the infectious window, however, as mentioned in Carruthers et al. (2022), the infectious period relies on how we define infectiousness. By setting the threshold as 10% of the maximum probability of infection, Ke et al. (2021) observes that the estimated infectious window is from 1.9 days to 7.9 days with a mean of 5.5 days. Under the same assumption, we notice that the exponential dose-response model shows estimated infectious windows from 3.95 days to 8.43 days with a mean of 5.46 days. The approximate beta-Poisson dose-response model shows estimated infectious windows from 5.13 days to 11.13 days with a mean of 9.00 days, which is significantly larger than the results from the exponential dose-response model. The logistic dose-response model shows an estimated infectious window from 4.17 days to 8.43 days with a mean of 5.78 days. Whilst comparable with the literature this definition of infectiousness as 10% of peak probability of infection is variable by individual and relative to viral load so may need refinement in future study.

These dose response models have been applied to the Test and Trace data under an Alan Turing Institute PhD studentship, which suggests that the models are feasible but that the received dose responds as a square root rather than linear (i.e. the realized dose at higher doses may be less than expected). This need further consideration to explain (it may be real

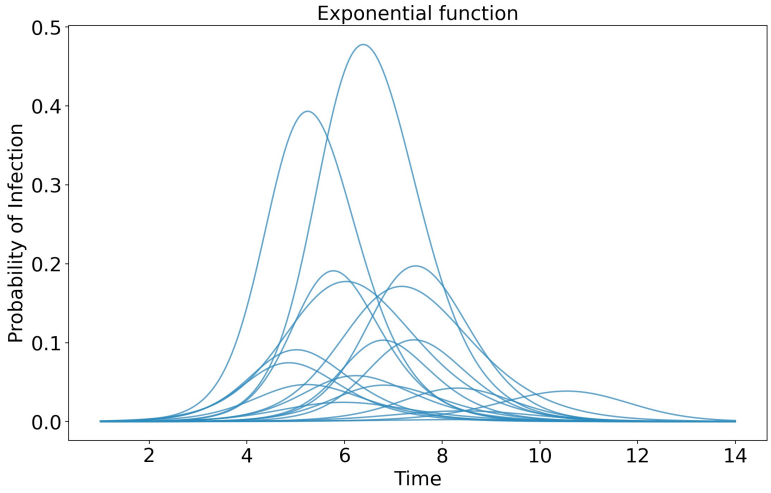

Figure 9: Viral load dependent probability of infection driven by exponential function 13 for 17 patients.

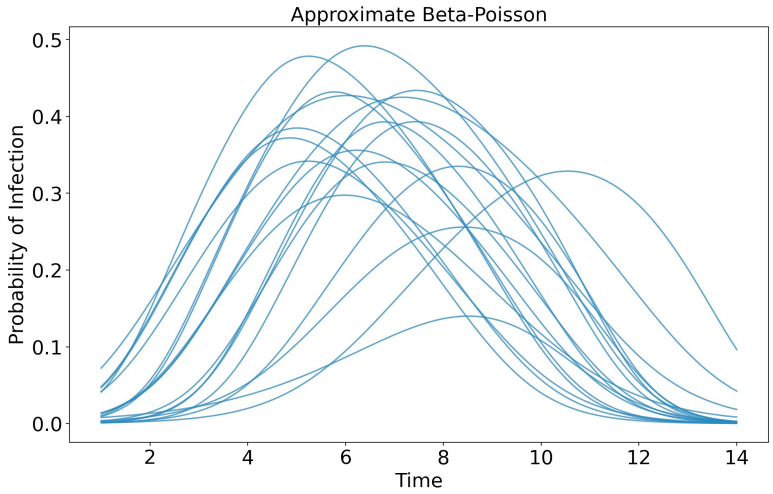

Figure 10: Viral load dependent probability of infection driven by exponential function 14 for 17 patients.

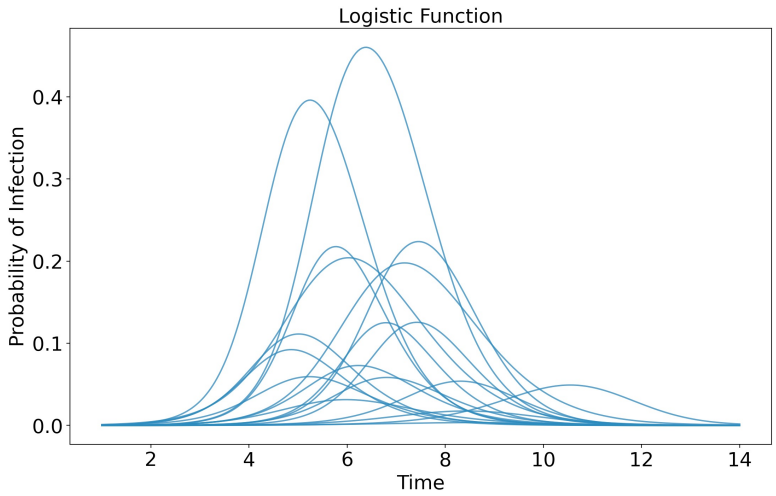

Figure 11: Viral load dependent probability of infection driven by exponential function 15 for 17 patients.

world noise, or perhaps due to deposition of particle sizes or infection site) and will be considered further on TRACK project.

4 Discussion

Within-host models play an important role in explaining data and making predictions for future impacts. In this work, we borrowed the previous work in Xu et al. (2022) and developed the multi-compartment model to consider the viral dynamics in different organs. In Section 2, we assumed different scenarios. Scenarios 1 and 2 only have two compartments, the nose and throat, both of which can interpret the nose and throat data. Scenario 3 considers human disease characteristics such as coughing and sneezing using the model in scenario 2, however, it should be noticed that we only assumed the rate of coughing and sneezing and the percentage of viral shed each time, which requires further study or data to verify. Since the 14-day nose and throat data from the human challenge study is the only data set we could access, it is difficult to consider other organs without causing parameter unidentifiability issue. Based on the assumption in Scenario 2, we attempted to use a Feller process to simulate the viral trajectory of the lung without introducing a new parameter. Besides, we assumed two-way migration between each compartment and the current model cannot explain the impact

of viral migration as the viral migration leaving one compartment can be cancelled by the viral migration from another compartment. More information of data about the magnitude of viral migration will be beneficial in this case. Due to the limitation of evidence base, we only develop a conceptual model in Section 3. This lack of empirical data has also stalled work on infection site (eye/nose/mouth). The conceptual model contains too many unidentifiable parameters and changing one or a few of them will significantly change the simulation result. We will leave it for future study (for example planned collaboration with the University of Leicester).

Alongside developing multi-compartment models, we also adopted the dose-response model from Xu et al. (2022) and linked the viral load and infectiousness in a quantitative way. We observed that a larger viral load leads to higher infectiousness. This is consistent with the previous studies; Ke et al. (2021) notices that the probability of infectiousness becomes high when the viral load exceeds between 10^6 and 10^7 RNA copies/ML. The exponential dose-response model and logistic model indicate that only a few patients showed high infectiousness during the infection while the approximate beta-Poisson dose-response model illustrated that most of the patients will show similar infectiousness and larger infectious window. Another purpose of the dose-response section is to investigate the infectious window. Compared with the conclusion of Goyal et al. (2021), we noticed that overall, each patient demonstrated a broader infectious window. However, the infectious window ultimately relies on how we define infectiousness Carruthers et al. (2022). The work of Ke et al. (2021) defines the 10% of the maximum probability of infection as the threshold and observed the mean infectious window is 5.5 days. Under the same assumption, we noticed a similar mean infectious window from the exponential dose-response model and logistic dose-response model. However, the approximate beta-Poisson model illustrates a much longer infectious window, up to 9.00 days but the approximate-Beta Poisson model require careful justification, as it is by definition an approximation. Continued development of dose response models should consider early inhibition of viral growth inside the host due to *Allee* type effects this was started in PROTECT phase 3 and will be continued on the TRACK project.

References

Abuin, P., Anderson, A., Ferramosca, A., Hernandez-Vargas, E. A. & Gonzalez, A. H. (2020), 'Characterization of sars-cov-2 dynamics in the host', *Annual reviews in control* **50**, 457–468.

Afonyushkin, V. N., Akberdin, I. R., Kozlova, Y. N., Schukin, I. A., Mironova, T. E., Bobikova, A. S., Cherepushkina, V. S., Donchenko, N. A., Poletaeva, Y. E. & Kolpakov, F. A. (2022), 'Multicompartmental mathematical model of sars-cov-2 distribution in human organs and their treatment', *Mathematics* **10**(11), 1925.

Baccam, P., Beauchemin, C., Macken, C. A., Hayden, F. G. & Perelson, A. S. (2006), 'Kinetics of influenza a virus infection in humans', *Journal of virology* **80**(15), 7590–7599.

Beauchemin, C. A., McSharry, J. J., Drusano, G. L., Nguyen, J. T., Went, G. T., Ribeiro, R. M. & Perelson, A. S. (2008), 'Modeling amantadine treatment of influenza a virus in vitro', *Journal of theoretical biology* **254**(2), 439–451.

Carruthers, J., Xu, J., Finnie, T. J. R. & Hall, I. (2022), 'A within-host model of sars-cov-2 infection', *medRxiv* .

Challenger, J. D., Foo, C. Y., Wu, Y., Yan, A. W., Marjaneh, M. M., Liew, F., Thwaites, R. S., Okell, L. C. & Cunnington, A. J. (2022), 'Modelling upper respiratory viral load dynamics of sars-cov-2', *BMC medicine* **20**(1), 1–20.

Dawood, A. A. (2021), 'Transmission of sars cov-2 virus through the ocular mucosa worth taking precautions', *Vacunas (English Edition)* **22**(1), 56.

Dogra, P., Ruiz-Ramírez, J., Sinha, K., Butner, J. D., Peláez, M. J., Rawat, M., Yellepeddi, V. K., Pasqualini, R., Arap, W., Sostman, H. D. et al. (2020), 'Innate immunity plays a key role in controlling viral load in covid-19: mechanistic insights from a whole-body infection dynamics model', *ACS pharmacology & translational science* **4**(1), 248–265.

Filippi, S., Barnes, C., Cornebise, J. & MPH, S. (2013), 'On optimality of kernels for approximate bayesian computation using sequential monte carlo', *Statistical applications in genetics and molecular biology* **12**(1), 87–107.

Gonçalves, A., Bertrand, J., Ke, R., Comets, E., De Lamballerie, X., Malvy, D. & Pizzorno, A. e. a. (2020), 'Timing of antiviral treatment initiation is critical to reduce sars-cov-2 viral load', *CPT: Pharmacometrics & Systems Pharmacology* **9**(9), 509–514.

Gonçalves, A., Maisonnasse, P., Donati, F., Albert, M., Behillil, S. & Contreras, V. e. a. (2021), 'Sars-cov-2 viral dynamics in non-human primates', *PLoS Computational Biology* **17**(3), e1008785.

Goyal, A., Cardozo-Ojeda, E. F. & Schiffer, J. T. (2020), 'Potency and timing of antiviral therapy as determinants of duration of sars-cov-2 shedding and intensity of inflammatory response', *Science advances* **6**(47), eabc7112.

Goyal, A., Reeves, D. B., Cardozo-Ojeda, E. F., Schiffer, J. T. & Mayer, B. T. (2021), 'Viral load and contact heterogeneity predict sars-cov-2 transmission and super-spreading events', *Elife* .

Haas, C. N., Rose, J. B. & Gerba, C. P. (2014), *Quantitative Microbial Risk Assessment*, John Wiley & Sons.

He, X., Lau, E. H., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y. C., Wong, J. Y., Guan, Y., Tan, X. et al. (2020), 'Temporal dynamics in viral shedding and transmissibility of covid-19', *Nature medicine* **26**(5), 672–675.

Heitzman-Breen, N. & Ciupe, S. M. (2022), 'Modeling within-host and aerosol dynamics of sars-cov-2: The relationship with infectiousness', *PLoS computational biology* **18**(8), e1009997.

Hernandez-Vargas, E. A. & Velasco-Hernandez, J. X. (2020), 'In-host mathematical modelling of covid-19 in humans', *Annual reviews in control* **50**, 448–456.

Holder, B. P., Simon, P., Liao, L. E., Abed, Y., Bouhy, X., Beauchemin, C. A. & Boivin, G. (2011), 'Assessing the in vitro fitness of an oseltamivir-resistant seasonal a/h1n1 influenza strain using a mathematical model', *PloS one* **6**(3), e14767.

Ke, R., Zitzmann, C., Ho, D. D., Ribeiro, R. M. & Perelson, A. S. (2021), 'In vivo kinetics of sars-cov-2 infection and its relationship with a person's infectiousness', *Proceedings of the National Academy of Sciences* **118**(49), e2111477118.

Ke, R., Zitzmann, C., Ribeiro, R. M. & Perelson, A. S. (2020), 'Kinetics of sars-cov-2 infection in the human upper and lower respiratory tracts and their relationship with infectiousness', *MedRXiv* .

Keeling, M. & Rohani, P. (2007), *Modelling Infectious Diseases in Humans and Animal*, Princeton University Press.

Kelsall, A., Decalmer, S., McGuinness, K., Woodcock, A. & Smith, J. A. (2009), 'Sex differences and predictors of objective cough frequency in chronic cough', *Thorax* **64**(5), 393–398.

Killingley, B., Mann, A. J., Kalinova, M., Boyers, A., Goonawardane, N., Zhou, J., Lindsell, K., Hare, S. S., Brown, J., Frise, R. et al. (2022), 'Safety, tolerability and viral kinetics during sars-cov-2 human challenge in young adults', *Nature Medicine* **28**(5), 1031–1041.

Lee, K. K., Matos, S., Evans, D. H., White, P., Pavord, I. D. & Birring, S. S. (2013), 'A longitudinal assessment of acute cough', *American journal of respiratory and critical care medicine* **187**(9), 991–997.

Li, C., Xu, J., Liu, J. & Zhou, Y. (2020), 'The within-host viral kinetics of sars-cov-2', *bioRxiv* .

Li, J., Wu, J., Zhang, J., Tang, L., Mei, H., Hu, Y. & Li, F. (2022), 'A multicompartment mathematical model based on host immunity for dissecting covid-19 heterogeneity', *Heliyon* p. e09488.

Madelain, V., Baize, S., Jacquot, F., Reynard, S., Fizet, A., Barron, S., Solas, C., Lacarelle, B., Carbonnelle, C., Mentré, F. et al. (2018), 'Ebola viral dynamics in nonhuman primates provides insights into virus immuno-pathogenesis and antiviral strategies', *Nature communications* **9**(1), 1–11.

Minter, A. & Retkute, R. (2019), 'Approximate bayesian computation for infectious disease modelling', *Epidemics* **29**, 100368.

Nath, B. J., Dehingia, K., Mishra, V. N., Chu, Y.-M. & Sarmah, H. K. (2021), 'Mathematical analysis of a within-host model of sars-cov-2', *Advances in Difference Equations* **2021**(1), 1–11.

Néant, N., Lingas, G., Le Hingrat, Q., Ghosn, J., Engelmann, I., Lepiller, Q., Gaymard, A., Ferré, V., Hartard, C., Plantier, J.-C. et al. (2021), 'Modeling sars-cov-2 viral kinetics and association with mortality in hospitalized patients from the french covid cohort', *Proceedings of the National Academy of Sciences* **118**(8), e2017962118.

of Canada, P. H. A. (2020), 'Coronavirus disease 2019 (covid-19): Epidemiology update'.

Pratt, A., Gilliard, J., Leach, S. & Ian, H. (2020), 'Dose-response modelling: extrapolating from experimental data to exposure in heterogeneous populations', *Risk Analysis* .

Smith, A. M. & Perelson, A. S. (2011), 'Influenza a virus infection kinetics: quantitative data and models', *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* **3**(4), 429–445.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A. & Stumpf, M. (2009), 'Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems', *Journal of the Royal Society Interface* **6**(31), 187–202.

Wang, S., Pan, Y., Wang, Q., Miao, H., Brown, A. N. & Rong, L. (2020), 'Modeling the viral dynamics of sars-cov-2 infection', *Mathematical biosciences* **328**, 108438.

Williamson, B. N., Feldmann, F., Schwarz, B., Meade-White, K., Porter, D. P., Schulz, J., Van Doremalen, N., Leighton, I., Yinda, C. K., Pérez-Pérez, L. et al. (2020), 'Clinical benefit of remdesivir in rhesus macaques infected with sars-cov-2', *Nature* **585**(7824), 273–276.

Wölfel, R., Corman, V. M., Guggemos, W., Seilmaier, M., Zange, S., Müller, M. A., Niemeyer, D., Jones, T. C., Vollmar, P., Rothe, C. et al. (2020), 'Virological assessment of hospitalized patients with covid-2019', *Nature* **581**(7809), 465–469.

Xu, J., Carruthers, J., Finnie, T. J. R. & Hall, I. (2022), 'Simplified within host and dose-response models of sars-cov-2', *medRxiv* pp. 2022–09.

The PROTECT COVID-19 National Core Study on transmission and environment is a UK-wide research programme improving our understanding of how SARS-CoV-2 (the virus that causes COVID-19) is transmitted from person to person, and how this varies in different settings and environments. This improved understanding is enabling more effective measures to reduce transmission – saving lives and getting society back towards 'normal'.

Published by the PROTECT COVID-19 National Core Study 01/2023