

RESEARCH ARTICLE

Open Access



Modelling upper respiratory viral load dynamics of SARS-CoV-2

Joseph D. Challenger^{1*} , Cher Y. Foo², Yue Wu³, Ada W. C. Yan⁴, Mahdi Moradi Marjaneh⁵, Felicity Liew⁶, Ryan S. Thwaites⁶, Lucy C. Okell¹ and Aubrey J. Cunnington^{5,7}

Abstract

Relationships between viral load, severity of illness, and transmissibility of virus are fundamental to understanding pathogenesis and devising better therapeutic and prevention strategies for COVID-19. Here we present within-host modelling of viral load dynamics observed in the upper respiratory tract (URT), drawing upon 2172 serial measurements from 605 subjects, collected from 17 different studies. We developed a mechanistic model to describe viral load dynamics and host response and contrast this with simpler mixed-effects regression analysis of peak viral load and its subsequent decline. We observed wide variation in URT viral load between individuals, over 5 orders of magnitude, at any given point in time since symptom onset. This variation was not explained by age, sex, or severity of illness, and these variables were not associated with the modelled early or late phases of immune-mediated control of viral load. We explored the application of the mechanistic model to identify measured immune responses associated with the control of the viral load. Neutralising antibodies correlated strongly with modelled immune-mediated control of viral load amongst subjects who produced neutralising antibodies. Our models can be used to identify host and viral factors which control URT viral load dynamics, informing future treatment and transmission blocking interventions.

Introduction

COVID-19 exhibits a wide range of severity, from asymptomatic infection to severe disease leading to hospitalisation and death. Age and sex have emerged as important risk factors for poor outcome [1, 2]. Viral load in the respiratory tract has been reported as an additional determinant of severity of illness [3, 4] and also a determinant of likelihood of transmission [5]. However, viral load varies over the course of illness due to dynamic interaction with the host immune response, and measurements at single points in time provide limited insight into this dynamic process. Within-host models of viral load can help to distinguish the sequence of events by tracking both viral dynamics and host response over

time, accounting for the effect of multiple factors simultaneously [4, 6, 7].

Studies measuring viral load over time in COVID-19 are beginning to establish viral dynamics and explore correlates of protection in the host response, although findings to date are somewhat contradictory and these relationships are still not well characterised [8]. Viral load in the upper respiratory tract (URT) peaks early in infection, usually before or within a few days of symptom onset [8–12]. Some studies suggest that viral load in the lower respiratory tract (LRT) may peak later, in the second week after symptoms [9], but this is much harder to measure in a serial manner. Viral load at a given time after diagnosis or detection tends to be similar between asymptomatic and symptomatic cases [13], but evidence tends towards a longer duration of viral shedding in more severe cases and older individuals [14, 15]. Despite the detection of viral material in URT samples from some individuals several weeks after onset of symptoms,

* Correspondence: j.challenger@imperial.ac.uk

¹Medical Research Council Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, Imperial College London, London, UK

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

infectious virus is not usually present beyond 8–14 days [16–19]. Viral shedding can also be detected in faeces, often persisting longer than URT viral detection [20], but there is little evidence that this contributes significantly to transmission [21].

The host response during COVID-19 has many features typical of an anti-viral immune response, including the generation of antibodies and T cell populations [22]. Antibodies are chiefly considered to contribute to viral clearance through pathogen neutralisation, though other effector functions may be of significance to COVID-19 [23]. A diverse array of T cell populations are active during COVID-19, including cytotoxic and helper populations. Whilst these cells contribute to viral clearance, emerging data indicates that T cells may also contribute to immunopathology in severe cases of COVID-19 [24]. Additionally, COVID-19 is associated with lymphopenia in peripheral blood [25], possibly reflecting the migration of cells to the site(s) of viral infection. In addition to these adaptive immune processes, innate immune processes are considered to play a major part in the pro-inflammatory state that scales with COVID-19 severity [26].

Within-host models are being developed to characterise viral kinetics in relation to host responses and disease outcomes and to guide therapeutic development. For example, Néant et al. [27] found an association between higher viral load late in infection and mortality. Goyal et al. [28] inferred different stages of host response from observing three stages of viral decline: a rapid drop following peak viral load, a period of slower decline, then rapid elimination of the virus. Benefield et al. estimated that viral load peaks prior to symptoms, suggesting substantial pre-symptomatic transmission [29]. Other within-host models have been used to explore the potential effects of antivirals, immunotherapies, and prophylactic treatment [30, 31]. More detailed models have simulated viral load in different tissues and detailed components of the innate and cellular immune response [32, 33]. However, many modelling studies to date have been calibrated to limited longitudinal data on viral load and host responses, which reduces parameter identifiability and the ability to infer pathways of pathogenesis and protection.

In this study, we aimed to improve understanding of how disease severity and immune response relate to URT viral load by developing a model framework which incorporates the dynamics of URT viral load over time. First, we sought to combine longitudinal data from published studies to characterise URT viral load dynamics (including peak viral load and rate of decline in viral load) in a large number of individuals, and to use linear regression models to assess whether variation in URT viral load was associated with age, sex or severity of

illness. Subsequently, we sought to fit a mechanistic model to this data, relating viral load to viral replication and its control by the immune response, and to investigate whether this model could facilitate identification of mechanisms which control URT viral load. We pragmatically represented the complex host response to the virus in two phases: an early phase which restricts the initial rate of viral replication, and a later phase which acts to clear the virus. We explored relationships between these components of the model and measured T cell and neutralising antibody measurements. This model represents the first step towards a well-validated, flexible, and open-source framework which can be utilised to better interpret immune responses in COVID-19 in the context of viral load and to understand how different treatments given at different stages of illness might influence viral load, transmission potential, host response, and outcome.

Results

Our literature search revealed 53 studies reporting longitudinal viral load measurements. We successfully obtained individual-level viral load measurements from 19 studies (either from the supplementary materials of the publication or preprint or by emailing the corresponding authors). The analysis presented here utilises data from 17 of these studies (summarised in Table 1). Studies that were identified but did not provide data are shown in Supplementary Table 1. We excluded two studies that contained little or no longitudinal data for URT samples [49, 50]. In all studies, a description of disease severity for each patient was available, although the level of detail varied between studies. In six studies viral loads were fully quantified from concurrent standard curves, whilst in the remaining 11 studies, cycle-threshold (Ct) values were reported. To combine data from different studies we generated an “average” standard curve, using data from 7 previously reported standard curves to convert Ct values to viral load per ml (Supplementary Fig. 1). Details on antiviral or immunomodulatory treatments were available for some studies (Supplementary Table 2).

In total, 2172 URT samples from 605 patients were used for analysis. Subject-level data on age and sex were available for 576 subjects and missing for 29 subjects. The majority (492 out of 576; 85%) of subjects were under 60, compared to 84 (15%) aged 60 or over; 321 (55%) patients were female. Disease severity was categorised using the WHO scale [51], where 501 (83%) patients (contributing 1698 (78%) samples) experienced mild COVID-19 illness, 65 (11%) patients (contributing 359 (17%) samples) had moderate severity illness, and 39 (6%) patients (contributing 115 (5%) samples) had severe illness. Supplementary Fig. 2 shows the viral load data

Table 1 Summary of participant characteristics and sample collection protocols for each study, from which data was collected to model viral load dynamics

Study Reference	No. of patients	Country	Healthcare setting	Recruitment period	Age, years Median (IQR)	Non-Caucasian ethnicity, n (%)	Male, n (%)	Remdesivir therapy, n (%)	WHO Severity score, Range	Mortality, n (%)	No. of samples, n	Method of sample collection	No. of samples used for modelling	Unitst
1 Kim JY et al., <i>J Korean Med Sci</i> , 2020 [34]	2	Korea	Hospital	January 2020	35 (10–50)	2(100)	1 (50)	0	3–4	0	26	Clinical	13	3
2 Lui G et al., <i>J Infect</i> , 2020 [35]	11	Hong Kong	Multicentre, Hospital	February 2020	58 (42–70)	NK	7(64)	0	3–6/7	0	73	Research protocol	34	3
3 Scott S et al., <i>Clin Infect Dis</i> , 2020 [36]	1	USA	Community	January 2020	26	NK	1(100)	0	2	0	12	Research protocol	6	1
4 Kim SE, <i>Int J Infect Dis</i> , 2020 [37]	3	Korea	Tertiary Hospital	February–April 2020	30 (25–30)	NK	2 (67)	0	3	0	21	Research protocol	21	1
5 Gautret P et al., <i>Int J Antimicrob Agents</i> , 2020 [38]	19	France	Tertiary Hospital	March 2020	49 (36.5–57)	NK	9 (47)	0	3–6/7	0	126	Research protocol	125	1
6 Young B et al., <i>JAMA</i> , 2020 [39]	18	Singapore	Multicentre, Tertiary Hospital	January–February 2020	47 (31–73*)	16 (89)	9 (50)	0	3–6/7	0	216	Research protocol	160	1
7 The COVID–19 Investigation team, <i>Nat Med</i> , 2020 [40]	12	USA	Multicentre, Community and Hospital	January 2020	53 (21–68*)	NK	8 (67)	3 (25)	2–5	0	121	Clinical	69	1
8 Wölfel R et al., <i>Nature</i> , 2020 [41]	9	Germany	Hospital	January 2020	41 (33–49)	NK	9 (100)	NK	3–4	0	96	Research protocol	79	3
9 Vetter P et al., <i>mSphere</i> , 2020 [42]	5	Switzerland	Hospital	February 2020	28 (24–55)	NK	5 (100)	0	2–4**	0	39	Research protocol	37	3
10 Lavezzo E et al., <i>Nature</i> , 2020 [43]	37	Italy	Community and Hospital	February–March 2020	65 (55–75)	NK	22 (59)	NK	2–10	3 (8)	55	Research protocol	6	3
11 Xu Y et al., <i>Nat Med</i> , 2020 [44]	6	China	Paediatric cohort, Tertiary Hospital	January–February 2020	6 (4–11)	NK	3 (50)	0	3	0	37	Research protocol	34	1
12 Shrestha N et al., <i>Clin Infect Dis</i> , 2020 [45]	230	USA	Healthcare worker cohort, non-hospitalised	March–April 2020	36 (29–51)	70 (30)	59 (26)	0	2	0	528	Clinical	46	2
13 Fajnzylber J et al., <i>Nat Commun</i> , 2020 [4]	64	USA	Multicentre, Tertiary Hospital	NK	56 (42–69)	76(43)	40 (63)	16 (25)	2–10	8 (13)	93	Research Protocol	0	3

Table 1 Summary of participant characteristics and sample collection protocols for each study, from which data was collected to model viral load dynamics (Continued)

Study Reference	No. of patients	Country	Healthcare setting	Recruitment period	Age, years Median (IQR)	Non-Caucasian ethnicity, n (%)	Male, n (%)	Remdesivir therapy, n (%)	WHO Severity score, Range	Mortality, n (%)	No. of samples, n	Method of sample collection	No. of samples used for modelling	Unitst
14 Yilmaz A et al., <i>J Infec Dis</i> , 2020 [46]	54	Sweden	Tertiary Hospital	February–April 2020	50.5 (36–58)	NK	31(57)	NK	2–10	3 (5)	349	Research Protocol	137	2
15 Alsharrah et al., <i>J Med Virol</i> , 2020 [47]	29	Kuwait	Paediatric cohort, Tertiary Hospital	February–April 2020	8.8 (4.7–12.4)	NK	16 (55)	0	3	0	75	Clinical	9	1
16 Tan A et al., <i>Cell Reports</i> , 2020 [7]	12	Singapore	Tertiary Hospital	NK	52.5 (32–65)	10 (83)	6(50)	NK	3–10	1 (8)	82	Research Protocol	53	1
17 Salvatore PP et al., <i>Clin Infec Dis</i> , 2020 [48]	93	USA	Community	March–May 2020	37 (21–53)	24 (22)	43 (46)	0	2	0	223	Research Protocol	41	1
Total	605	-	-	-	42 (28–56)	-	271 (45)	-	2-10	15 (2)	2172	-	870	-

Only cases of symptomatic COVID-19 were included
 Hospital = non-tertiary hospital or hospital for which the tertiary status has not been reported by the authors
 NK not known

*Value expressed as median (range)

**All cases hospitalised due to public health policy at the time of the study

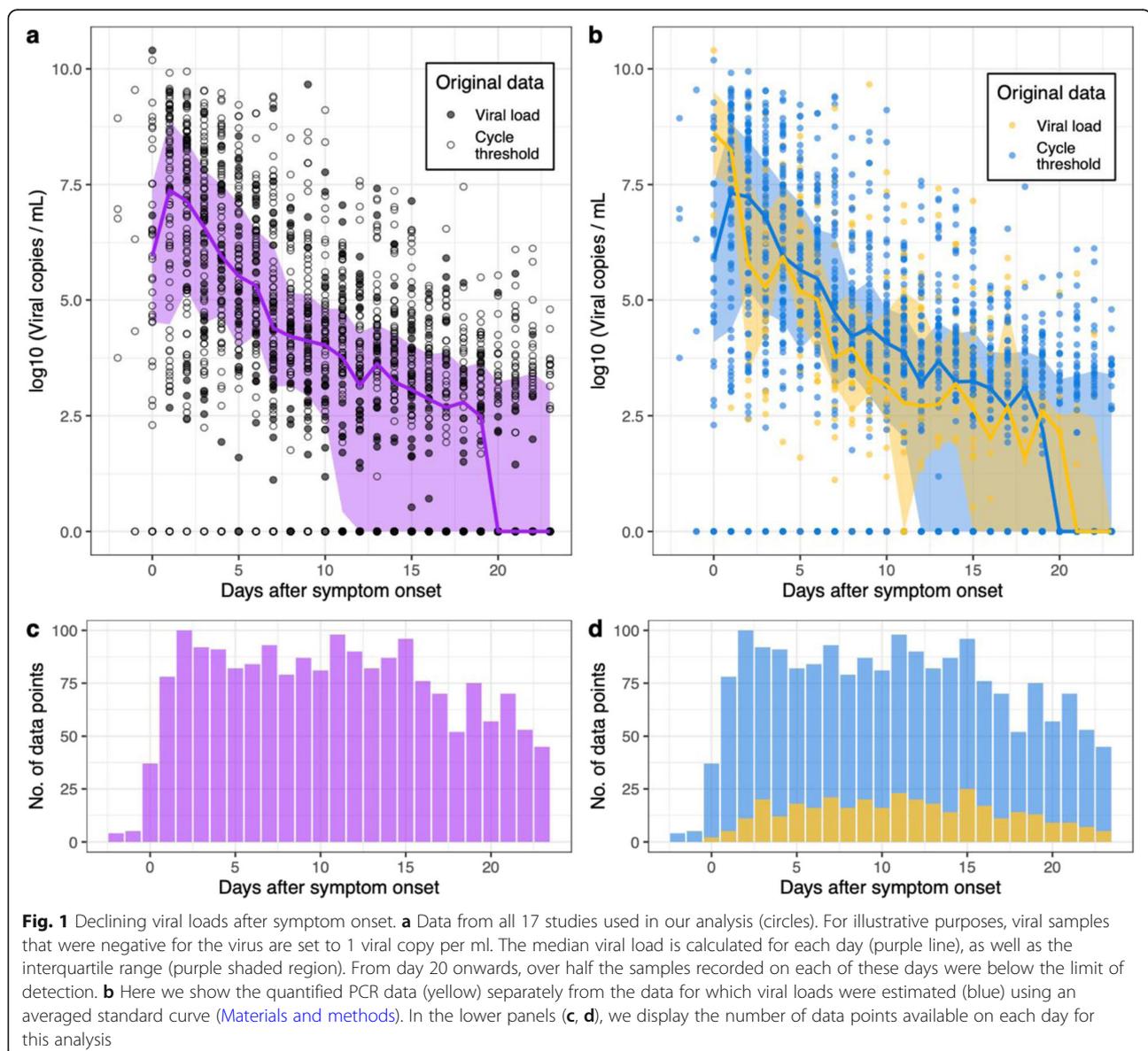
#1: Cycle threshold values; 2: Viral load calculated either by using the study's own conversion formula for Ct value to viral load or a standard curve calibrated to other samples (e.g. viral load in plasma); 3: Viral load calculated from standard curve run with each sample

⊗ Models fitted to data from subjects for whom at least 3 samples were taken in the first 15 days after symptom onset

for all patients, presented separately for each study. The vast majority (2163, or 99.6%) of samples were taken after the onset of symptoms, although samples were collected from a minority (8 patients across all 17 studies) prior to symptom onset e.g. if they were identified as a close contact of another patient. The median timing of swabs was 12 days after symptom onset (interquartile range: 6–19; range: 2 days before symptom onset to 54 days afterwards). Pooling the data from all studies, we found that the median viral load peaked one day after symptom onset (Fig. 1a), although less data was available on the day of symptom onset compared to subsequent days (Fig. 1c). 421 patients had more than one URT swab recorded: for 60.1% of these patients, the first sample had the highest recorded viral load. Viral load

estimates at corresponding times after the onset of symptoms did not differ systematically between studies in which viral load was calculated by the authors of the original study or inferred from our averaged standard curve (Fig. 1b), although less data was contributed by studies which used concurrent standard curve quantification (Fig. 1d).

The timing of the first sample obtained relative to the onset of symptoms varied with the severity of illness. Subjects with moderate or severe disease had first samples collected later in their illness than those with mild disease (Supplementary Fig. 3a). Accordingly, the first viral load and maximum viral load measurements for subjects with moderate or severe disease in these studies



were lower than in those with mild disease (Supplementary Figs. 3b and 3c).

Fitting a regression model to the viral load data

We fitted two types of models to the viral load data from the first 15 days from onset of symptoms (see the “Materials and methods” section), using only subjects with at least three samples during this time period (models fitted to 870 samples from 155 patients from 16 studies). The first was a linear regression model, fitted to log-transformed viral loads. This model included patient- and study-specific random effects, which captured variation from the average behaviour observed across 16 datasets. There was variation in the peak and slope of the viral loads across different studies (Fig. 2, Supplementary Table 3). The variation in the (log-transformed) peak viral load varied over several orders of magnitude. Some of this variation was explained by study-specific differences: the standard deviation of the between-study variation in the peak viral load was 0.84, i.e. nearly one order of magnitude (Supplementary Table 3). This could be due to a number of factors, such as the method of sample collection, quantification method or characteristics of included patients. Furthermore, viral loads in several studies were estimated using an averaged standard curve, which introduces some uncertainty into the magnitude of the viral load. However, the inclusion of study-specific random effects allows such data to be appraised alongside data from other studies.

Within this regression framework, we incorporated information on age, disease severity and sex (Materials and methods) to see if the goodness of fit could be improved. We added fixed effects for these three variables, both separately and in combination. We separately examined age as a continuous variable or a dichotomous variable (stipulating whether patients were 60 years of age or over). The goodness of fit to the data did not vary appreciably: here we report results for age as a dichotomous variable. As we had relatively few (29) samples from patients with severe disease in the subset of the data considered here, we pooled patients with moderate disease and severe disease together. The inclusion of the fixed effects for age, sex, or severity did not improve the model fit (Table 2, Supplementary Table 4). In other words, the wide variation between individuals observed in the viral load dynamics could not be explained by the inclusion of these variables (Fig. 3).

We used simulation-based methods (Materials and methods) to estimate the power of our analysis to detect different effect sizes for severity, age and sex on peak viral load (Supplementary Fig. 4). For severity and sex, our analysis has 80% power to detect a difference of around 1.1 in the \log_{10} viral load (about a 12-fold difference in viral load), whereas power was lower for

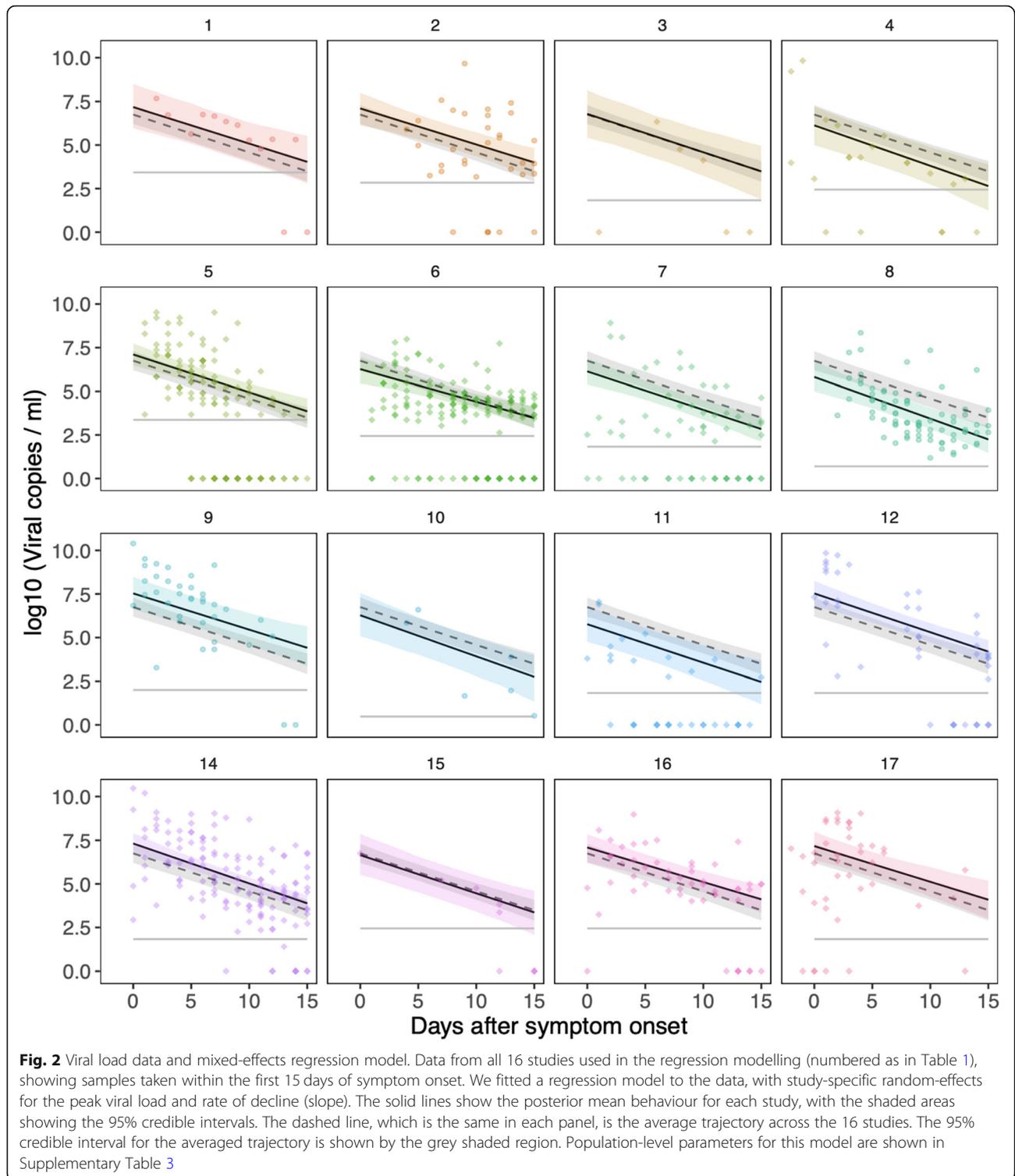
detecting the effect of age. Importantly the detectable differences are considerably smaller than the inter-individual variation in viral load at any given time point (Supplementary Table 3), indicating that these are not the major determinants of viral load.

Fitting a mechanistic model to the viral load data

In addition to modelling viral load decline using regression models, we also developed a mechanistic model which we fitted to the dataset. We elected to keep the model relatively simple due to a lack of identifiability between more complex model structures. We represented the multi-faceted immune response to the infection via two components (Fig. 4). First, the exponential growth of viral load in the initial phase of infection is brought under control by an *early immune response*. Subsequent to this, the infection is gradually cleared by a *late immune response*. The early immune response is stimulated by a high load of infected cells and starts to block viral replication and the invasion of susceptible cells. The late response requires a maturation phase before it becomes effective and is, therefore, more representative of the adaptive immune response. However, we do not attempt to fully distinguish between innate and adaptive responses in this model, as the interplay between them is complex.

To guide the model fitting (see the “Materials and methods” section), we made the assumption that both the peak viral load and the activation of the early immune response should roughly be concurrent with the onset of symptoms. The vast majority of subjects in the dataset were only under observation after the onset of symptoms, which means we are unable to infer the rate of exponential growth during the initial phase of the infection. For each subject, we fitted two parameters in the mechanistic model, holding all other parameters fixed (see the “Materials and methods” section). One of these free parameters governs the density of infected cells required to activate the early response, whilst the other determines the rate at which the late response clears infected cells and, therefore, the rate at which the viral load declines. As we did for the regression modelling, we used a nested random effect structure, with study- and patient-specific random offsets for both of these parameters (Materials and methods). In addition, we incorporated data points that fell below the limit of detection in each study by accounting for censoring in the likelihood (Materials and methods).

This approach allowed us to characterise the average time-course of an infection at the population-level, i.e. after removing study- and patient-level offsets (Fig. 5, Supplementary Table 5), as well as showing the dynamics for the study-specific fits to the data. There was no significant relationship between the subject-specific



random effects for either of the model parameters and maximum disease severity (Supplementary Fig. 5), consistent with the outcome of the preceding regression modelling (Table 2). Similarly, there was no significant relationship between the subject-specific random effects

and subject age or sex (not shown). Supplementary Fig. 6 shows the 95% credible intervals for the study-specific random effects, showing the between-study variation in both the peak viral load (panel a) and the rate at which the infection is cleared in the URT (panel b).

Table 2 Assessing the goodness of fit of the regression models

Fixed effects included	WAIC	Δ WAIC	dSE	Model weight in ensemble
Data from all studies (n=16)				
None	2222.5	-	-	0.62
Severity	2223.5	1.0	1.1	0.38
Data from studies (n=14) with demographic information				
None	1796.6	-	-	0.33
Age	1797.4	0.8	1.2	0.22
Severity	1798.8	2.2	1.0	0.11
Sex	1798.9	2.2	0.9	0.11
Severity + sex	1799.7	3.0	1.4	0.07
Age + sex	1799.7	3.1	1.1	0.07
Age + sex + severity	1800.7	4.0	1.9	0.04
Age + severity	1800.7	4.0	1.7	0.04

Three fixed effects (severity, age, sex) were added to linear regression models of viral load over time, separately or in combinations (see the “Materials and methods” section), to determine the extent to which they explained variation present in the data. All studies included in this analysis described the severity of disease for each patient. However, patient-level demographic information was only provided in 15 of the 17 studies. Hence, a separate analysis was carried out for these studies. In each analysis, the best-fitting model is the one with the lowest WAIC value, as indicated in bold. The value of Δ WAIC indicates the difference in goodness of fit between each model and the one that provided the best fit to the data, whilst dSE denotes the standard error of the difference in WAIC between each model and the best-fitting one. In each analysis, we provide the Akaike weight of each model in the ensemble of candidate models that were analysed (Materials and methods). As a sensitivity analysis, we also assessed the goodness of fit by leave-one-out cross-validation [52], which resulted in very similar conclusions (see Supplementary Table 4)

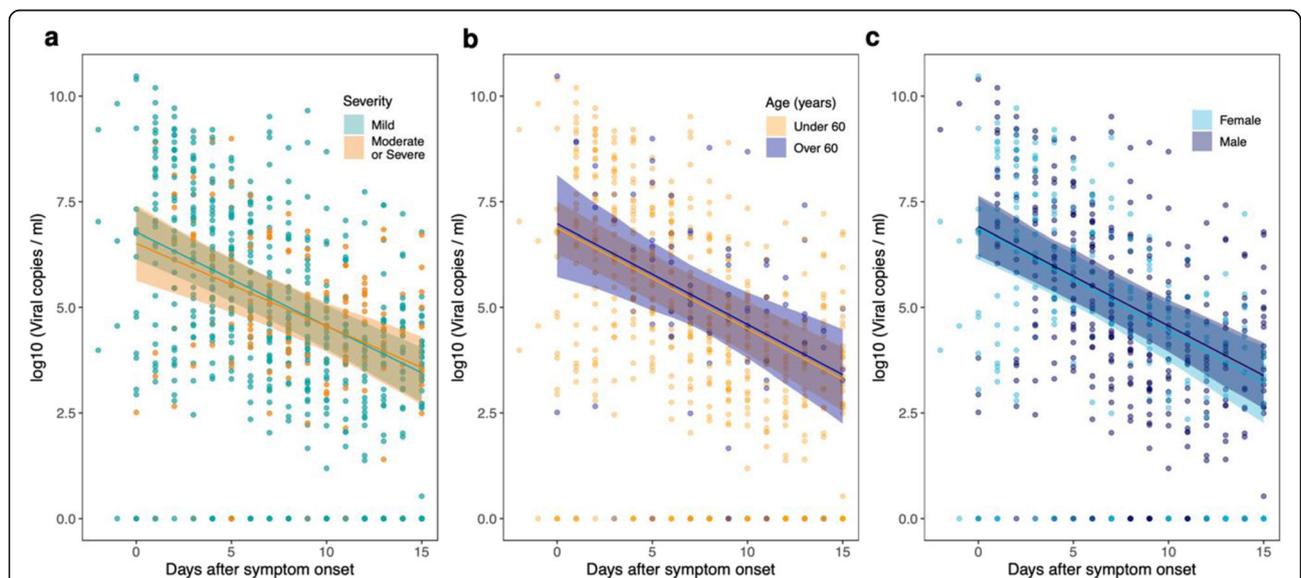
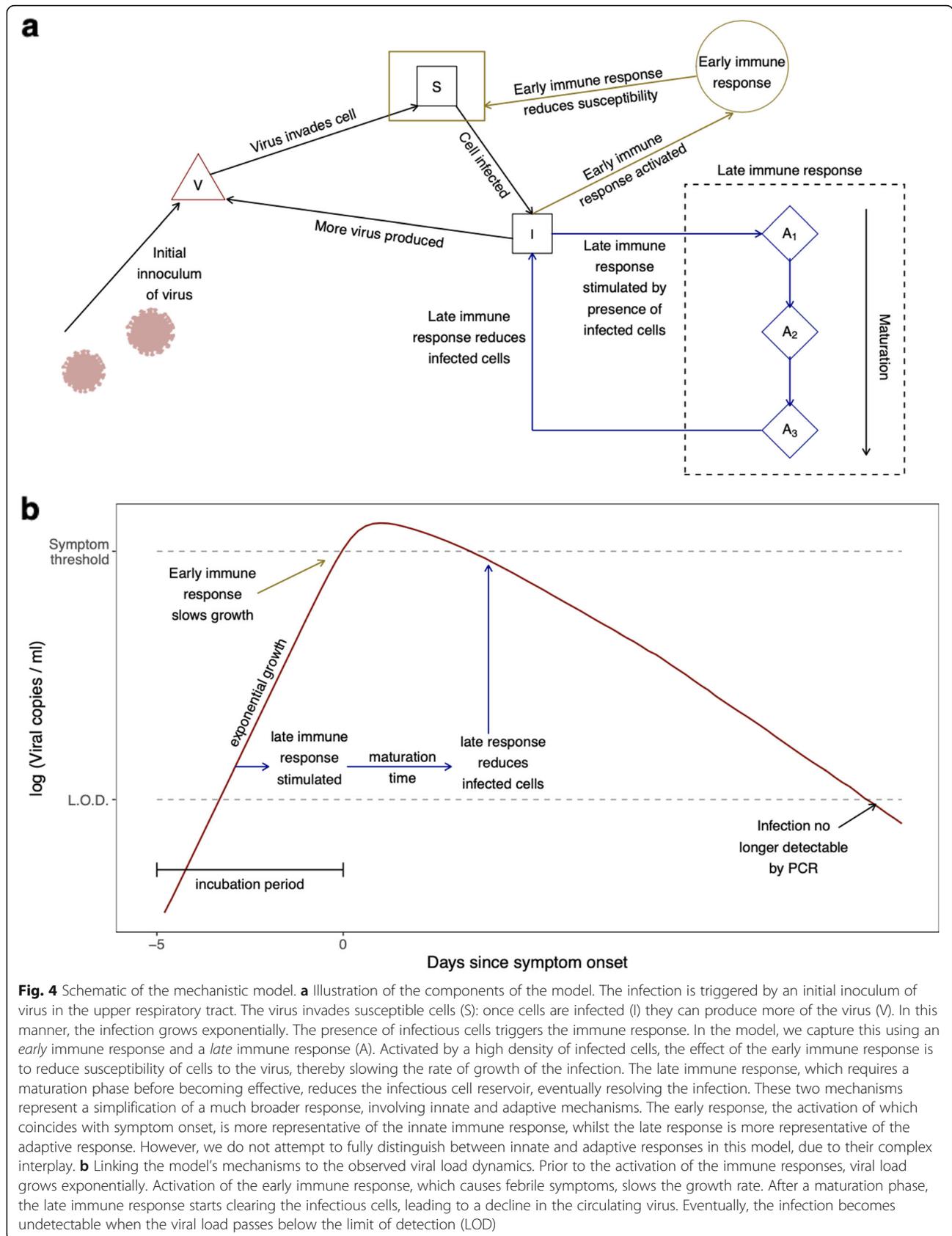


Fig. 3 Inclusion of fixed effects in the regression models. All regression models included study- and patient-specific random effects for the peak and slope (i.e. rate of decline) of the viral load. We then added fixed effects, both separately and in combination, to see if the model fit could be improved (Table 2). These fixed effects were: age, sex, and severity of disease. Here we show results for the three models containing one fixed effect (left: severity; middle: age; right: sex). The inclusion of severity, age, or sex did not improve the goodness of fit (Table 2). In the modelling here, age was included as a binary variable (under or over 60 years of age)



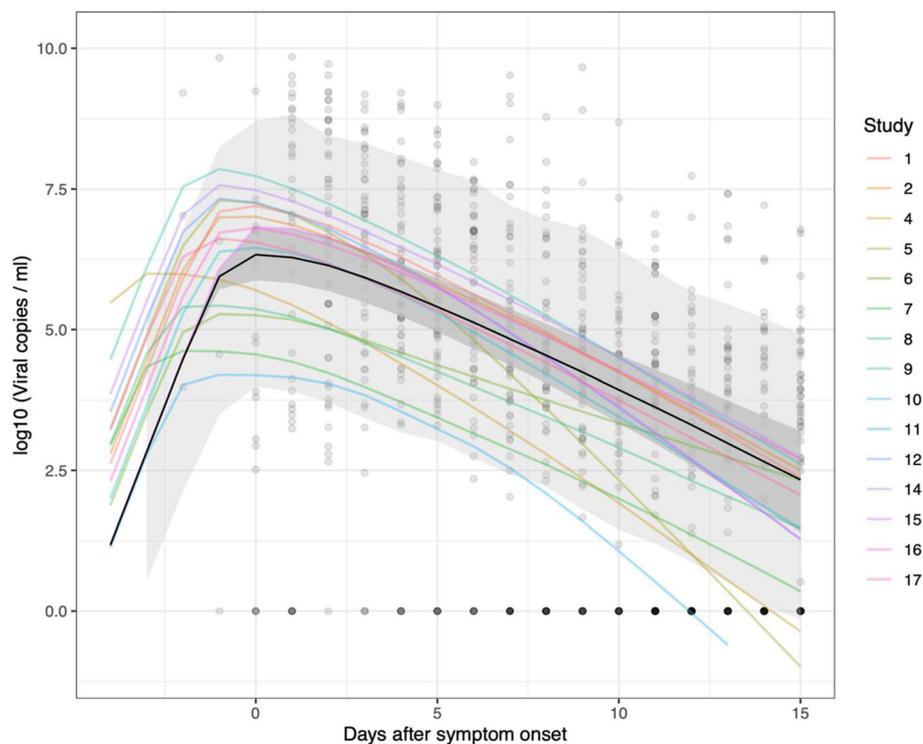


Fig. 5 Averaged trajectories obtained from the fitted mechanistic model. In this plot, we show the average trajectory predicted for each study (coloured lines), generated using the median value used for the initial viral load at $t=0$ in each case. We also show the average trajectory across all the studies, indicated by the black line. The dark grey shaded area indicates the 95% credible interval for this average trajectory. The light grey area accounts for the variation observed around the average trajectory (generated using samples from σ , as defined in in Eq. 8, and calculating the 95% prediction interval for the population-level dynamics). The fit to data from Study 3 is not shown, as this study only contained one patient, which means one cannot distinguish between study- and patient-specific random effects. The opaque black circles are the data points from the 16 studies used to fit the model. For illustrative purposes, viral samples that were negative for the virus are set to 1 viral copy per ml (i.e. 0 on the log-scale). The results from the mechanistic model presented here were obtained using 1500 samples from the posterior distribution, with the median trajectories plotted

Supplementary Fig. 8 shows modelled viral loads for each of the 155 subjects used to fit the mechanistic model.

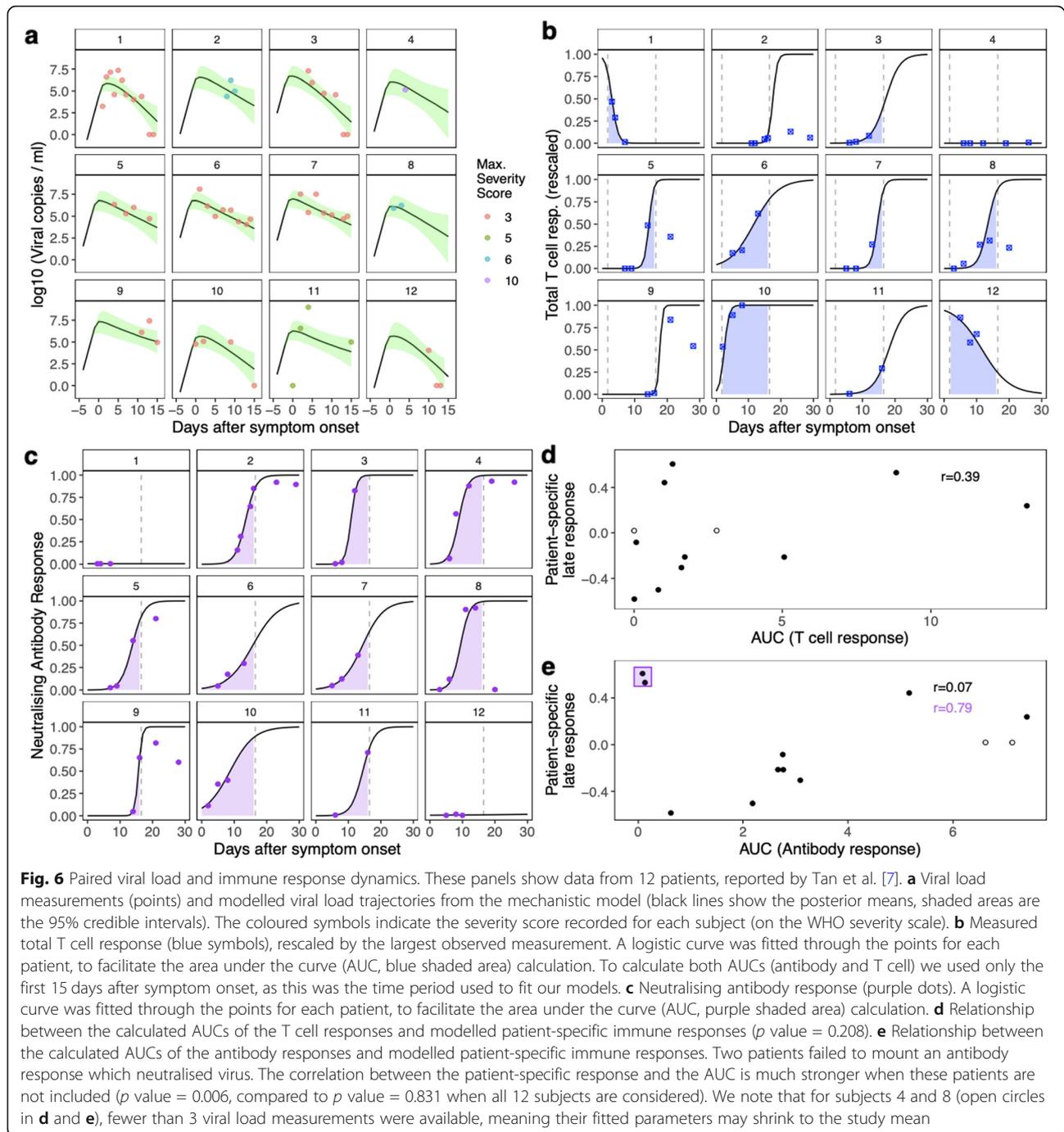
Relating viral load dynamics to SARS-CoV-2-specific immune responses

To demonstrate the potential to link viral load dynamics to SARS-CoV-2-specific adaptive immune responses in individual subjects, we made use of detailed data on neutralising antibody and T cell responses from the study by Tan et al. [7]. We compared this immunological data with the patient-specific late immune response parameter fitted in our mechanistic model in these 12 patients, to attempt to gain insight into within-host mechanisms that might control the viral load (Fig. 6). Specifically, we assessed the total interferon- γ (IFN- γ) T cell response to SARS-CoV-2 peptides measured in an ELISPOT assay and results from a surrogate virus antibody neutralisation assay. Simple logistic curves were fitted to the data from each subject's measured immune responses, and the area under each curve (AUC) was used as a measure

of the magnitude of each subject's immune responses (Fig. 6b, c). Overall, no significant correlation was observed between the modelled late immune response and the total interferon- γ (IFN- γ) T cell response ($r = 0.39$, p value = 0.208, see Fig. 6d) or the neutralising antibody response ($r = 0.07$, p value = 0.831, see Fig. 6e). However, we noted two subjects (subjects 1 and 12) displayed quite distinct immune responses to the others, characterised by absent neutralising antibody, and initially high but subsequently declining total interferon- γ (IFN- γ) T cell responses, possibly indicating a qualitatively different response to SARS-CoV-2. When these two subjects were removed from the analysis, a much stronger correlation was observed between the modelled late immune response and surrogate virus antibody neutralisation ($r = 0.79$, p value = 0.006).

Discussion

Understanding the causes and consequences of variation in pathogen load is fundamental to infectious disease research [53]. Increasing pathogen load can drive both



pathogenesis and transmission of infection [54]. High viral load in COVID-19 has been associated with severity of illness in some studies [3, 4], but not others (see e.g. Ref. [55]), and has been associated with risk of transmission [5]. Pathogen load is dynamic, it varies over time, and is often considered to be the stimulus for the host response, as well as a target of the host response. Therefore, any attempt to establish the determinants of pathogen load and relationships between pathogen load,

severity, and transmissibility should account for these dynamics.

We collated longitudinal URT viral load data from 2172 samples taken from 605 subjects with SARS-CoV-2 infection in 17 studies, to investigate the association of viral load dynamics with age, sex, and severity of illness. We used the WHO clinical progression scale to standardise varied descriptions of disease severity reported in different studies. We used two distinct modelling

approaches to characterise viral load dynamics, accounting for systematic differences in viral load estimation between studies. We found no evidence to support the hypotheses that URT viral load dynamics are substantially influenced by age or sex, or that URT viral load dynamics influence the severity of illness. We also found no association between severity and the latent variables describing early and late immune responses to the URT virus in our mechanistic model. Nevertheless, we identified considerable inter-individual variation in URT viral load dynamics. Understanding the biological basis for this variation could help to identify new approaches to reduce transmission of SARS-CoV-2.

The lack of association between URT viral load dynamics and severity of illness is particularly interesting. On one hand, this is not necessarily surprising since severe COVID-19 is predominantly a consequence of LRT and systemic disease processes, and some studies have indicated that viral load in LRT samples is more strongly associated with severity of illness [9]. On the other hand, this would imply that distinct processes govern URT and LRT viral load dynamics, or that the extension of infection from URT to LRT and other systemic locations is controlled by different mechanisms to those controlling local viral load. Although it is difficult to obtain serial LRT viral load measurements, evaluating distinct mechanisms controlling local and spatial viral dynamics could be important to understand the pathogenesis of COVID-19 and other respiratory infections.

Our study provides one of the most comprehensive assessments of URT viral load dynamics, but despite collating data from a large number of studies we had a relatively low proportion of patients with very severe illness and very few from those with fatal infection, potentially reducing our ability to distinguish different viral dynamics in these groups. We also had very little data on URT viral loads before the onset of symptoms, which limits our ability to model variation in the rate of increase in viral load early in infection or consider the incubation period as a possible covariate. We lacked data on the interval from exposure to symptoms, forcing us to fix this parameter in our mechanistic model. This means we were unable to properly explore the role that the initial viral dose plays in the dynamics. Furthermore, we did not have sufficient data on ethnicity, or other host characteristics besides age and sex, for which it may have been instructive to examine associations with viral load dynamics. We extracted data on antiviral treatment in different studies where possible (Supplementary Table 2), but there were too few subjects receiving each treatment to allow meaningful analysis.

We used two modelling approaches to analyse the data. Mixed-effects regression modelling sought to determine if any of the wide variation observed in patients'

URT viral loads could be explained by age, sex, or disease severity. The inclusion of these variables, separately and in combination, did not improve the model fit. However, our power analysis suggests that we cannot confidently exclude smaller differences in viral load (e.g. an increase or decrease in peak viral load of less than tenfold) due to these variables. We believe our power analysis provides useful insight into the capacity of such models to detect differences in viral load dynamics in different subpopulations. Another limitation is that we did not consider correlations between covariates independent of the viral load. One could imagine a more complex model where age or sex (or both) could be associated with severity both through changes in viral load and through changes in severity at a given viral load. However, given that each covariate separately did not have a significant effect on viral load, considering correlations between these variables is unlikely to affect our results.

In order to utilise as much data as possible, we have included semi-quantitative data (presented as Ct values) alongside fully quantitative viral load data. We have converted the former using an averaged standard curve. Modelling the data with study-specific random effects provides a way to analyse all the data collected, without requiring study-specific standard curve data. Adding data from more studies in future will allow us to examine whether the conclusions of our analyses are influenced by this approach.

It is useful to compare the two modelling approaches—fitting a regression model to capture the peak viral load and the rate of its decline versus using a mechanistic model. With the former approach, we avoid making assumptions about the mechanisms controlling viral load or overfitting the data, given that only a few viral load data points are available for each individual and immune data were not available for most individuals. On the other hand, no insight is obtained about the mechanisms underlying the differences between peak and slope. The mechanistic model provided more insight in this regard, but would benefit from longitudinal data for components other than the viral load.

It is interesting to compare our mechanistic model to others that have been fitted to viral load data. Goyal et al. [28] developed a more complex model, which was able to capture a changing rate of decline in viral load observed in the patient data available to them. However, one should note that in our model we are fitting to a narrower timeframe (no more than 15 days after symptom onset, due to lack of viable infective virus after this time), so it may be that the changing rate at which viral load declines is not so noticeable during this phase of the infection. This narrower time frame was one factor that led us to choose a relatively simple model structure.

Another difference between the two approaches is that we use the time of symptom onset to centre the dynamics (i.e. set time equal to zero), whereas Goyal et al. used the time of the first swab. When considering data from many different subjects, using symptom onset for temporal alignment helps adjust for the wide variation in the time of sample collection, particularly as we observed a relationship between disease severity and the time of the first sample collection in the studies considered here (Supplementary Fig. 3). Amongst the within-host models already published, one finds a wide variation in model complexity. Some models, such as the one presented here, have erred on the side of parsimony, whilst others have sought to capture very intricate within-host processes (see e.g. Ref. [33]). Our parsimonious approach was influenced by the lack of immunological data for the vast majority of subjects considered here, which hampered our ability to elucidate specific within-host mechanisms in more detail. An interesting avenue for further work would be to use goodness-of-fit criteria, such as the one used here for the regression modelling, to explore the extent to which more complicated models explain more of the variation present in the data.

We have demonstrated the potential to relate modelled viral load dynamics, and the immunological determinants of the model, to measured immunological data for individual subjects. It is not surprising that circulating SARS-CoV-2-specific T cell responses are poor correlates of the late immune response controlling viral load, because these cells would need to migrate from the circulation to other locations like the respiratory tract to control virus. It is reassuring that when we considered individuals who did mount a neutralising antibody response, we saw that antibody neutralisation did correlate well with the late immune response parameter of our model, consistent with the evolving evidence that a neutralising antibody does indeed play an important role. However, it is now well established that some individuals do not mount a detectable serum antibody response to SARS-CoV-2; nevertheless, they have protective immunity against re-infection [56], and applying this modelling approach to much larger numbers of subjects might help to identify alternative or additional protective mechanisms. Due to the dynamic interplay between viral load and the immune response, more biological insight can be gained from mechanistic modelling, compared to using summary statistics or regression modelling.

In conclusion, our study clearly illustrates that the relationships between URT viral load and COVID-19 severity or immune responses need to be assessed in the context of the dynamic changes in URT viral load over time. The lack of association between URT viral load dynamics and severity, age or sex in our analysis, indicates

that measuring the effects of anti-viral or adjunctive treatments on URT viral load may not be very useful for predicting their effects on severe COVID-19 illness. However, the mechanistic model we have developed provides a framework for the identification of immunological mechanisms which may control the transmission of SARS-CoV-2 through their action on viral load. We believe our mechanistic model has the potential to accelerate the discovery of mechanisms and therapeutics which suppress URT viral load and therefore reduce transmission. We present this tool for other researchers to use with different patient populations, combining sequential viral load measurements and paired measurements of immune response parameters, to identify mediators associated with early or late phase immune responses controlling URT viral load. This would be particularly well-suited to prioritisation of candidate mechanistic correlates in high dimensional omics datasets and may provide leads for desirable vaccine-induced responses or novel therapeutic approaches to reduce transmission of SARS-CoV-2.

Materials and methods

Data

To collect data on viral load dynamics, we searched PubMed and medRxiv for studies that recorded longitudinal viral load data from individuals with symptomatic COVID-19 infections. Searches were carried out between May 20, 2020, and February 11, 2021. In particular, we searched for studies that reported the timing of symptom onset for each patient, which we used to temporally align samples from different patients. We identified 53 studies, data from 5 of which could be extracted from the publication or preprint. We contacted the authors of 48 studies by email to request access to patient-level data, and we received data from 14 of these studies. Two of the 19 studies were dropped from this analysis, as they contained little or no longitudinal data for URT samples. Studies for which data was obtained are summarised in Table 1: all remaining studies identified by our literature search are summarised in Supplementary Table 1. We extracted as much information as possible on the severity of illness experienced by the patients and extracted demographic information on the patients where available. Two authors independently studied the severity information and matched the descriptions to the WHO clinical progression scale [51]. Some articles contained detailed information on the course of disease for each patient, but this was not available in all the studies. We considered scores of 1–3 to represent mild disease, 4–6 to represent moderate disease and scores above 6 to represent severe disease. This is slightly different to the patient state descriptors of the WHO scale, because the majority of studies did not provide sufficient detail about

the method of oxygen delivery to allow us to distinguish between scores of 5 and 6. In some studies, the severity of symptoms was recorded at multiple timepoints, over the course of the infection. Here, we use the term ‘severity’ to describe the maximum severity of disease experienced by each patient.

Although some studies with a long follow-up period demonstrate that some patients can remain PCR-positive for the virus for well over a month [57, 58], several studies [16, 17] have demonstrated that very few people are infectious, based on URT swabs, beyond 10 or 14 days after symptom onset. Data collected after this period is likely to reflect RNA debris that remains in the URT. Therefore, our models of viral load are fitted to data points within the first 15 days of symptoms. This means that we have fewer data points to fit to (Table 1), but we believe that the meaningful dynamics, in terms of how the host controls the infection, are found in this time period.

Viral load quantification

In all studies, real-time polymerase chain reaction (PCR) assays were used to quantify URT virus either as Ct values, or as viral copies per ml (*V*). As discussed in Ref. [59], calculation of viral copies per ml from Ct values requires the use of a ‘standard curve’, which is calibrated to the experimental set-up in a particular laboratory using reference samples. In general, these curves have the form:

$$V = 10^{a-bCt} \tag{1}$$

Here *a* and *b* are positive numbers that fully specify the viral load for a given Ct value. This relationship indicates that there is a linear relationship between log-transformed viral loads and the raw Ct values, with higher Ct values representing lower viral loads. The units of *V* vary between studies (e.g. viral copies per ml, per swab, or per 1000 cells), all studies collected here that have quantified viral load have used viral copies per unit of volume. We collected as many different standard curves as possible, from the studies included in this analysis [34, 43] and from other papers in the COVID-19 literature [50, 60–62], to understand the variation observed. From these, we determined an ‘averaged’ standard curve (Supplementary Fig. 1), using the mean observed values for parameters *a* and *b*, which we used to estimate the viral loads for studies for which only Ct values were available. This enabled us to pool data from all 17 studies.

Models

We sought to explain variation in viral load (either its peak value or its rate of decline over time) amongst

patients, due to (e.g.) age, sex, and severity of the disease. We did this using two types of models, a linear mixed effects regression approach and a mechanistic model, which took the form of a system of first-order differential equations. Both models were fitted to viral load data from within 15 days of symptom onset, using only subjects with at least three samples taken during this time period (models fitted to 870 samples from 155 patients from 16 studies). Therefore, as indicated in Table 1, no data from study 13 was used to fit the models.

Linear regression models

Bayesian regression models were fitted using RStan [63], with some of the analysis carried out using the rethinking package [64]. Linear models were fitted to log-transformed viral loads, with random effects for each patient and study applied to the parameters determining the peak viral load, which we assumed coincides with the onset of symptoms, and rate of its decline over time. Samples for which no virus was detected were treated as being below limit of detection (LOD), rather than truly virus negative. We show the general form of the regression model here, where *L* is the likelihood of each data point, to illustrate the random-effect structure and the censoring of the data:

$$\begin{aligned} V > 0 : L &= N(\log_{10} V | \mu, \sigma) \\ V = 0 : L &= N_{CDF}(LOD | \mu, \sigma) \\ \mu &= a_0 + a_1[\text{study}] + a_2[\text{patient}] + \\ &\quad (b_0 + b_1[\text{study}] + b_2[\text{patient}])\text{Day} \\ a_0 &\sim N(5, 5) \\ b_0 &\sim N(0, 1) \\ (a_1[\text{study}], b_1[\text{study}]) &\sim N_2\left((0, 0), \rho_{\text{study}}, \sigma_{\text{study}}\right) \\ (a_2[\text{patient}], b_2[\text{patient}]) &\sim N_2\left((0, 0), \rho_{\text{patient}}, \sigma_{\text{patient}}\right) \\ \sigma_{\text{study}} &\sim \text{Cauchy}(0, 2) \\ \sigma_{\text{patient}} &\sim \text{Cauchy}(0, 2) \\ \rho_{\text{study}} &\sim \text{LKJ}(2) \\ \rho_{\text{patient}} &\sim \text{LKJ}(2) \\ \sigma &\sim \text{Cauchy}(0, 2) \end{aligned} \tag{2}$$

This model is relatively simple: the log-transformed viral loads are captured by a linear model, with the intercept describing the viral load at the time of symptom onset, and the slope capturing the rate at which the viral load subsequently declines. The log-transformed viral load for a given patient is normally distributed around a modelled trajectory, which is described by μ , with a standard deviation given by σ . For data points where no virus was detected, the likelihood takes a different form. We calculate the likelihood as the probability of the viral load being below the LOD, writing N_{CDF} as the

cumulative distribution function of the normal distribution. We allow for the fact that the LOD is study-specific (indicated in Fig. 2). All remaining terms in Eq. 2 define the prior distributions for the parameters. Both the slope and the intercept have a parameter that captures the average behaviour across all patients in all studies (a_0 and b_0 respectively). Random effects capture patient- and study-specific deviations from the average behaviour. Both the patient- and study-specific random effects are normally distributed with zero mean: here we write N_2 to indicate the bivariate normal distribution. We allow correlation between the random effects on the intercept and slope at each level, as indicated by the presence of parameters ρ_{study} and ρ_{patient} . For these parameters, we use the LKJ distribution [65] as a prior.

Within this framework, fixed effects could then be added to the model, to assess whether (e.g.) severity of disease, or the sex of patient affected the typical viral load trajectory observed. To do this, we simply add terms to μ in Eq. 1. Here we illustrate this for sex, making female the reference category:

$$\begin{aligned} \mu &= a_0 + a_1[\text{study}] + a_2[\text{patient}] + a_M\text{Male} + \\ &(b_0 + b_1[\text{study}] + b_2[\text{patient}] + b_M\text{Male})\text{Day} \\ a_M &\sim N(0, 1) \\ b_M &\sim N(0, 1). \end{aligned} \tag{3}$$

Here variable ‘Male’ is equal to 1 if the patient is male (0 otherwise), and we have also added zero-mean priors for the new parameters. As mentioned above, severity was expressed on the WHO scale, which takes values from 1 to 10. As 1 is asymptomatic, severity in our dataset is limited from 2 (mild symptoms, not hospitalised) to 10 (dead). As the dataset contains relatively few samples from patients with severe disease (score 7 to 10 on the WHO scale), we chose to make two groups from these categories: mild (2 or 3) and moderate or severe (4–10). When severity is included in the model, μ becomes:

$$\begin{aligned} \mu &= a_0 + a_1[\text{study}] + a_2[\text{patient}] + a_{SM}\text{ModSev} + \\ &(b_0 + b_1[\text{study}] + b_2[\text{patient}] + b_{SM}\text{ModSev})\text{Day} \\ a_{SM} &\sim N(0, 1) \\ b_{SM} &\sim N(0, 1). \end{aligned} \tag{4}$$

Here variable ‘ModSev’ is equal to 1 for patients with moderate or severe disease (0 otherwise). In most studies, patient age is given in years, as an integer value. Some studies expressed the age of patients in decades, e.g. 10–20 years of age. For these studies, we used the midpoint of the age range given for each patient. We grouped age into two groups: < 60 years of age and ≥ 60

years of age. In the regression modelling, we added a fixed effect for age, making the youngest age group the default. Here, μ takes the form:

$$\begin{aligned} \mu &= a_0 + a_1[\text{study}] + a_2[\text{patient}] + a_A\text{Age} + \\ &(b_0 + b_1[\text{study}] + b_2[\text{patient}] + b_A\text{Age})\text{Day} \\ a_A &\sim N(0, 1) \\ b_A &\sim N(0, 1). \end{aligned} \tag{5}$$

Here variable ‘Age’ is equal to 1 for patients aged 60 or over (equal to 0 otherwise). We also looked at including subject age as a continuous variable, but the goodness of fit did not change appreciably. We added the 3 terms (age, severity, sex) to the regression models both separately and in combination. We assessed the goodness of fit using the Watanabe Akaike Information Criterion (WAIC) [65], with the best-fitting model having the lowest WAIC value. As the WAIC for each candidate model is estimated from a finite sample, its standard error was calculated using the rethinking package [64] to appreciate the uncertainty in its value. These standard errors are useful when appraising the differences in WAIC between candidate models. The relative goodness of fit of a given model can also be appraised by calculating its Akaike weight amongst the set of all considered models. This can be interpreted as the probability that this model, out of the set of models considered, would provide the best fit to new (i.e. out of sample) data [65].

We use simulation-based estimation of statistical power to assess our capacity to detect a difference in viral load dynamics due to one of the three factors (sex, age, severity of disease) assessed here. We generated synthetic datasets of the same size as the one considered here, with study- and patient-specific random effects of the same magnitude (Supplementary Table 3). For simplicity, we generated datasets with an equal number of samples per subject, and the sampling times were randomly generated from a uniform distribution. When generating the data, we assumed that the peak viral load was influenced by one of the three aforementioned factors. We then ran the regression analysis, to see if the modelled effect could be detected. To reduce computation time, we here used frequentist regression via the lme4 package in R [66], with a p value < 0.05 for the included fixed effect indicating a significant finding. Generating 1000 synthetic datasets, the statistical power can be estimated as the percentage of datasets for which the regression analysis found a significant effect. Supplementary Fig. 4 shows how the statistical power varies with the magnitude of the modelled effect, which is here expressed as a fold difference in the peak viral load. The results of the power analysis suggest that we were not

well powered to measure relatively small differences in peak viral load (under 10-fold difference). The power to measure sex-specific differences in peak viral load was slightly higher than the power to measure severity- or age-specific differences, as we have a better balance between samples from male and female samples than we do between samples in under 60 s and over 60 s, or samples from subjects from mild disease versus samples from subjects with moderate or severe disease.

Mechanistic model

We also developed a mechanistic model to describe the viral dynamics. We started modelling the infection 5 days before the onset of symptoms. At this time, an initial dose of virus V_0 is introduced. These virions start invading susceptible cells at rate β . Infected cells then produce more virus at rate p , which can then invade subsequent susceptible cells. Meanwhile, free virions are cleared at rate γ . In this way, the viral population grows exponentially. In the model, this growth is brought under control by an early phase of the immune response, and then the infection is gradually cleared by a later phase of the immune response. We do not seek to specify the immunological mechanisms contributing to these phases of the immune response. In each infection, the exponential growth slows as the population of infected cells approaches a certain value (I_{max}). This reflects a combination of two within-host mechanisms: the depletion of susceptible cells in the URT and the effect of an early phase immune response which is triggered at a high level of infection. Since these mechanisms may be linked, i.e. the immune response may modify the susceptibility of target cells, we do not attempt to distinguish between them, or to explicitly model the population of susceptible cells.

After the exponential growth is brought under control, the infection is cleared by the late immune response. In this model, this is triggered by a certain density of infected cells but requires a maturation stage before it becomes effective at clearing infected cells. We write the system of equations as:

$$\begin{aligned}
 \frac{dI}{dt} &= \beta V \left[1 - \left(\frac{I}{I_{max}} \right) \right] - k_a I \frac{A_3^2}{A_3^2 + A_{50}^2} \\
 \frac{dV}{dt} &= pI - \gamma V \\
 \frac{dA_1}{dt} &= \frac{I}{I + I_{50}} - k_c A_1 \\
 \frac{dA_2}{dt} &= k_c A_1 - k_c A_2 \\
 \frac{dA_3}{dt} &= k_c A_2.
 \end{aligned}
 \tag{6}$$

The late immune response, represented here by variables (A_1, A_2, A_3) , is stimulated by the presence of

infected cells. A Hill function is used to make the adaptive response dimensionless, and to rescale its value (daily input into compartment A_1 scales between 0 and 1). Parameter I_{50} (here fixed to a value of 1000 cells) determines the magnitude of the density of infected cells required to stimulate this response. It requires time to mature (governed by rate parameter k_c , which we fix at 0.33 day^{-1}), meaning that only stage A_3 is able to clear infected cells. Once mature, the late response does not wane in this model, as we are only interested in its ability to clear an infection, not how long it persists after the infection has been resolved. In this way, the late response recapitulates features of adaptive cell-mediated and humoral immunity, without needing to specify their relative contributions. Parameter k_a represents the maximum rate at which the late response can clear the infection, and A_{50} governs the magnitude of the adaptive response required to effectively clear infected cells. We note that the magnitude of the late immune response has been set to be of order 1 for convenience, and this determines the magnitude of A_{50} .

For the vast majority of patients for whom we have data, only samples taken after the onset of symptoms are recorded. This means that we are unable to estimate the duration of the incubation period, the initial dose of virus that causes the infection, or the rate at which the virus reproduces before being acted upon by the immune response. We elect to model a five-day incubation period and fix the rate of exponential growth to be the same for all patients, setting $\beta = 0.8 \text{ ml day}^{-1} \text{ virion}^{-1}$, $\gamma = 13 \text{ day}^{-1}$, $p = 80 \text{ day}^{-1}$, with these values informed by parameter values chosen in published models [27, 28, 30]. This is a simplification: it is unclear whether, in reality, the time from infection to symptom onset varies by age, sex or disease severity. For each patient, the infection starts at day -5 (in the dataset, day 0 is the day of symptom onset, not the day of infection onset) with initial conditions given by $(I = 0, V = V_0, A_1 = 0, A_2 = 0, A_3 = 0)$ i.e. the infection starts with an initial viral dose, no infected cells, and no prior exposure to the virus (late immune response completely inactive).

The initial viral dose, V_0^i , for each patient is chosen with the idea in mind that the viral load should peak at or close to the time of symptom onset. Ideally, V_0^i , should be fitted simultaneously with the free parameters for each patient, but this has proved challenging to date, especially given that peak viral load should coincide approximately with symptom onset but may also not be observed. Briefly, V_0^i is chosen by considering the linear subsystem of equations that govern the growth of the virus prior to the activation of the immune response to the infection. As this system of equations is linear, it can be solved exactly (we use a matrix exponential). We

require that V_0^i be chosen so that, just prior to the onset of symptoms, an uncontrolled infection would reach the maximum viral load observed for patient i in the dataset. This ensures that the peak viral load modelled for each patient is controlled by the early-stage immune response, as it is too soon for the adaptive response to be effective at clearing infected cells.

When fitting to the data, we allowed the values of k_a and I_{max} to vary and we fixed all the other parameters. As we did for the regression modelling, we included random effects in the model, to account for different behaviour in different patients and across different studies. In the case where we have P patients taken from S studies, we write:

$$\begin{aligned}
 k_a &= k_a^0 + k_a^p[i] + k_a^s[j] \\
 k_a^p[i] &\sim N(0, \sigma_{kp}) \quad i = (1, 2, \dots, P) \\
 k_a^s[j] &\sim N(0, \sigma_{ks}) \quad i = (1, 2, \dots, S) \\
 I_{max} &= I_{max}^0 + I_{max}^p[i] + I_{max}^s[j] \\
 I_{max}^p[i] &\sim N(0, \sigma_{Ip}) \quad i = (1, 2, \dots, P) \\
 I_{max}^s[j] &\sim N(0, \sigma_{Is}) \quad j = (1, 2, \dots, S)
 \end{aligned} \tag{7}$$

This means that k_a^0 and I_{max}^0 represents the population-level average value of each parameter. The system of equations was solved numerically in R, using the dopri function from the dde package [67]. We show how this numerical solution can be obtained in the code repository that accompanies this article (see the section ‘Availability of Code’). We write the modelled viral load trajectory at time t for patient i in study j as $V_{ij}(t)$. Samples from the posterior distribution of the fitted parameter were obtained using Markov Chain Monte Carlo methods, via the R package drjacob [68]. As with the regression modelling, the form of the likelihood for each data point depends on whether the sample is above the study-specific LOD or is recorded as being virus-negative. We write D_{ijk} to indicate the k th viral load sample, taken t_k after the onset of symptoms, from patient i in study j . The likelihood for each data point has the form:

$$\begin{aligned}
 V > 0 : L &= N(\log_{10}(D_{ijk}) \mid \log_{10} V_{ij}(t_k), \sigma) \\
 V = 0 : L &= N_{CDF}(\log_{10} LOD \mid \log_{10} V_{ij}(t_k), \sigma)
 \end{aligned} \tag{8}$$

The global likelihood was then obtained by multiplying together the likelihoods for each data point. We provided weakly informative prior distributions for the parameters to be fitted. These were chosen to guide the fitting process, rather than being reflective of information gleaned from other studies. These distributions were:

$$\begin{aligned}
 k_a^0 &\sim N(6.8, 1.5) \\
 I_{max}^0 &\sim N(18, 2.0) \\
 \sigma_{kp} &\sim LN(0.2, 0.7) \\
 \sigma_{ks} &\sim LN(0.2, 0.7) \\
 \sigma_{Ip} &\sim LN(0.2, 0.7) \\
 \sigma_{Is} &\sim LN(0.2, 0.7) \\
 \sigma &\sim \text{Exp}(0.8).
 \end{aligned} \tag{9}$$

Supplementary Fig. 7 shows the model output obtained for both the prior and posterior distributions.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12916-021-02220-0>.

Additional file 1: Supplementary Table S1. Summary of studies which were identified during the literature search but did not provide viral load data when the corresponding authors were contacted. Studies included both male and female patients, unless stated. All ages stated are in years. *For some studies, particularly those carried out early in the pandemic, patients were hospitalised to ensure isolation (i.e. cohorts may include asymptomatic subjects or those with very mild symptoms). † These studies were identified during our literature search, but we were informed by the study authors that quantitative viral load was not available (i.e. only positive or negative result recorded). **Supplementary Table S2.** Summary of antiviral and immunomodulatory treatment in the studies included in analysis. **Supplementary Table S3.** Summary of the population-level (i.e. not study- or patient-specific) parameter values (and 95% credible intervals) obtained for the multi-level regression modelling (as displayed in Fig. 2). Patient- and study-specific random effects were used for both the peak (log-transformed) viral load, and its rate of decline per day. **Supplementary Table S4.** assessing the goodness of fit of the regression models using leave-one-out cross-validation. **Supplementary Table S5.** Summary of the population-level (i.e. not study- or patient-specific) parameter values (and 95% credible intervals) obtained for the mechanistic viral load model (Eqs. 6–8). Samples from $ka0$ and $Imax0$ were used to generate the black line and dark grey shaded area in Fig. 4. **Supplementary Figure S1.** Standard curves relating cycle-threshold (Ct) values to viral load. Seven standard curves, identified from published studies (see Methods) are plotted. **Supplementary Figure S2.** Summary of all the data collected (see Table 1 in the main text). For the studies shown in blue, viral loads have been estimated using an averaged standard curve (see Methods for details). **Supplementary Figure S3** Comparison of timing of first sample and viral load by severity. **Supplementary Figure S4.** Estimations of the statistical power in the regression analyses. **Supplementary Figure S5.** Relationship between patient-specific parameters governing the immune response in the mechanistic model and disease severity. **Supplementary Figure S6.** Posterior means and 95% credible intervals for the study-specific offsets in the mechanistic model. **Supplementary Figure S7.** A comparison of the prior and posterior distributions for the early and late immune responses in the mechanistic model. **Supplementary Figure S8.** (shown over the following 7 pages): Output from the mechanistic model alongside the data, for all 155 patients considered. In the heading of each panel, the first number indicates the study (studies numbered as in Table 1).

Acknowledgements

We would like to express our gratitude to the study authors who shared their data, as well as all those involved in the data collection. In addition, we acknowledge Professor Antonio Bertoletti and colleagues for making subject-level immunology data available. JDC would like to thank Peter Winskill, Robert Verity, and Clare McCormack for useful discussions.

Availability of code

The R scripts used to carry out the analyses presented here and generate the figures are available in a Github repository: https://github.com/JDChallenger/Viral_Load_Covid19.

Authors' contributions

JDC cleaned the data, carried out the analyses and created the figures. JDC and LCO developed the model, with input from all other authors. CYF, YW, and FL extracted the data on disease severity and the antiviral and immunomodulatory treatments administered. All authors contributed to the interpretation of the results. JDC, RST, LCO, and AJC wrote the first draft of the manuscript. AJC designed the study. All authors read and approved the final manuscript.

Funding

This work was supported by a research grant (MR/V027409/1) from UKRI (MRC) and DHSC (NIHR), as well as grant funding from the Imperial College Covid-19 Response Fund. Infrastructure support was provided by the NIHR Imperial Biomedical Research Centre. JDC and LCO acknowledge funding from the MRC Centre for Global Infectious Disease Analysis (reference MR/R015600/1), jointly funded by the UK Medical Research Council (MRC) and the UK Foreign, Commonwealth & Development Office (FCDO), under the MRC/FCDO Concordat agreement and is also part of the EDCTP2 programme supported by the European Union. AWCY was funded by a Wellcome Trust Investigator Award to Becca Asquith (103865Z/14/Z).

Availability of data and materials

Anonymised viral load data is available alongside this article, in the Excel worksheet 'CombinedDataset.xlsx'.

Declarations**Ethics approval and consent to participate**

This study conducted secondary analysis of data from published studies. All studies included were either (i) approved by an institutional ethics review committee and obtained informed consent from all participants or (ii) were exempted from this requirement because they involved identification, control, or prevention of disease in response to an immediate public health threat, and had been determined not to be public health research. Further ethical approval was not required for the present study which used only anonymised non-identifiable information from these studies.

Competing interests

LCO declares grant funding from the Bill and Melinda Gates Foundation.

Author details

¹Medical Research Council Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, Imperial College London, London, UK. ²School of Medicine, Imperial College London, London, UK. ³School of Clinical Medicine, University of Cambridge, Cambridge, UK. ⁴Department of Infectious Disease, Imperial College London, London, UK. ⁵Section of Paediatric Infectious Disease, Department of Infectious Disease, Imperial College London, London, UK. ⁶National Heart and Lung Institute, Imperial College London, London, UK. ⁷Centre for Paediatrics and Child Health, Imperial College London, London, UK.

Received: 18 July 2021 Accepted: 15 December 2021

Published online: 13 January 2022

References

- Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, et al. Risk stratification of patients admitted to hospital with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. *BMJ*. 2020;9.
- Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, et al. Features of 20 133 UK patients in hospital with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ*. 2020;22.
- Pujadas E, Chaudhry F, McBride R, Richter F, Zhao S, Wajnberg A, et al. SARS-CoV-2 viral load predicts COVID-19 mortality. *Lancet Respir*. 2020;5:1.
- Fajnzylber J, Regan J, Coxen K, Corry H, Wong C, Rosenthal A, et al. SARS-CoV-2 viral load is associated with increased disease severity and mortality. *Nat Commun*. 2020;11(1):5493.
- 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. *Lancet Infect Dis*. 2021;1:1–8.
- Lieberman NAP, Peddu V, Xie H, Shrestha L, Huang M-L, Mears MC, et al. In vivo antiviral host transcriptional response to SARS-CoV-2 by viral load, sex, and age. *PLoS Biol*. 2020;8:18(9).
- Tan AT, Linster M, Tan CW, Le Bert N, Chia WN, Kunasegaran K, et al. Early induction of functional SARS-CoV-2-specific T cells associates with rapid viral clearance and mild disease in COVID-19 patients. *Cell Rep*. 2021;34(6):108728.
- 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. *Lancet Microbe*. 2021;2(1):e13–22.
- Walsh KA, Jordan K, Clyne B, Rohde D, Drummond L, Byrne P, et al. SARS-CoV-2 detection, viral load and infectivity over the course of an infection. *J Infect*. 2020;81(3):357–71.
- Zheng S, Fan J, Yu F, Feng B, Lou B, Zou Q, et al. Viral load dynamics and disease severity in patients infected with SARS-CoV-2 in Zhejiang province, China, January–March 2020: retrospective cohort study. *BMJ*. 2020;15:m1443–8.
- Arons MM, Hatfield KM, Reddy SC, Kimball A, James A, Jacobs JR, et al. Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. *N Engl J Med*. 2020;382(22):2081–90. <https://doi.org/10.1056/NEJMoa2008457>.
- Cheng H-Y, Jian S-W, Liu D-P, Ng T-C, Huang W-T, Lin H-H. Contact tracing assessment of COVID-19 transmission dynamics in Taiwan and risk at different exposure periods before and after symptom onset. *JAMA Intern Med*. 2020;180(9):1156–63.
- Lee S, Kim THT, Lee E, Lee C, Kim H, Rhee H, et al. Clinical Course and Molecular Viral Shedding Among Asymptomatic and Symptomatic Patients With SARS-CoV-2 Infection in a Community Treatment Center in the Republic of Korea. *JAMA Intern Med*. 2020;180(11):1–6.
- Liu Y, Yan L-M, Wan L, Xiang T-X, Le A, Liu J-M, et al. Viral dynamics in mild and severe cases of COVID-19. *Lancet Infect Dis*. 2020;20(6):656–7. [https://doi.org/10.1016/S1473-3099\(20\)30232-2](https://doi.org/10.1016/S1473-3099(20)30232-2).
- Zhou C, Zhang T, Ren H, Sun S, Yu X, Sheng J, et al. Impact of age on duration of viral RNA shedding in patients with COVID-19. *Ageing*. 2020;1:12. <https://doi.org/10.18632/aging.104114>.
- Laferl H, Kelani H, Seitz T, Holzer B, Zimpernik I, Steinrigl A, et al. An approach to lifting self-isolation for health care workers with prolonged shedding of SARS-CoV-2 RNA. *Infection*. 2021;49, 95(1):–101.
- Singanayagam A, Patel M, Charlett A, Lopez Bernal J, Saliba V, Ellis J, et al. 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):465. <https://doi.org/10.2807/1560-7917.ES.2020.25.32.2001483>.
- Sohn Y, Jeong SJ, Chung WS, Hyun JH, Baek YJ, Cho Y, et al. Assessing viral shedding and infectivity of asymptomatic or mildly symptomatic patients with COVID-19 in a later phase. *J Clin Med*. 2020;9(9):2924–9. <https://doi.org/10.3390/jcm9092924>.
- Rhee C, Kanjilal S, Baker M, Klompas M. Duration of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infectivity: when is it safe to discontinue isolation? *Clin Infect Dis*. 2020;72(8):1467–74.
- Guo M, Tao W, Flavell RA, Zhu S. Potential intestinal infection and faecal–oral transmission of SARS-CoV-2. *Nat Rev Gastroenterol Hepatol*. 2021;18(4):269–83.
- Pedersen RM, Tomby DS, Bang LL, Madsen LW, Skov MN, Jensen TG, et al. Rectally shed SARS-CoV-2 lacks infectivity: time to rethink faecal–oral transmission? *Nat Rev Gastroenterol Hepatol*. 2021;18(9). <https://doi.org/10.1038/s41575-021-00501-w>.
- Siggins MK, Thwaites RS, Openshaw PJM. Durability of immunity to SARS-CoV-2 and other respiratory viruses. *Trends Microbiol*. 2021;29(7):648–62.
- Krammer F, Simon V. Serology assays to manage COVID-19. *Science* (80-). 2020;368(6495):1060–1.
- Zhao Y, Kilian C, Turner J-E, Bosurgi L, Roedel K, Bartsch P, et al. Clonal expansion and activation of tissue-resident memory-like T_H17 cells expressing GM-CSF in the lungs of patients with severe COVID-19. *Sci Immunol*. 2021;6(56):eabf6692.

25. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020;395(10223):507–13.
26. Thwaites RS, Sanchez Sevilla Uruchurtu A, Siggins MK, Liew F, Russell CD, Moore SC, et al. Inflammatory profiles across the spectrum of disease reveal a distinct role for GM-CSF in severe COVID-19. *Sci Immunol*. 2021;6(57): eabg9873.
27. Néant N, Lingas G, Le Hingrat Q, Ghosn J, Engelmann I, Lepiller Q, et al. Modeling SARS-CoV-2 viral kinetics and association with mortality in hospitalized patients from the French COVID cohort. *Proc Natl Acad Sci*. 2021;118(8). <https://doi.org/10.1073/pnas.2017962118>.
28. Goyal A, Cardozo-Ojeda EF, Schiffer JT. Potency and timing of antiviral therapy as determinants of duration of SARS-CoV-2 shedding and intensity of inflammatory response. *Sci Adv*. 2020;6:eabc7112–24.
29. Benefield AE, Skrip LA, Clement A, Althouse RA, Chang S, Althouse BM. SARS-CoV-2 viral load peaks prior to symptom onset: a systematic review and individual-pooled analysis of coronavirus viral load from 66 studies. *medRxiv*. 2020:1–23.
30. Czuppon P, Débarre F, Gonçalves A, Tenaillon O, Perelson AS, Guedj J, et al. Success of prophylactic antiviral therapy for SARS-CoV-2: Predicted critical efficacies and impact of different drug-specific mechanisms of action. *PLOS Comput Biol*. 2021;17(3). <https://doi.org/10.1371/journal.pcbi.1008752>.
31. Garcia-Cremades M, Solans BP, Hughes E, Ernest JP, Wallender E, Aweeka F, et al. Optimizing hydroxychloroquine dosing for patients with COVID-19: an integrative modeling approach for effective drug repurposing. *Clin Pharmacol Ther*. 2020;12, 108(2):–263. <https://doi.org/10.1002/cpt.1856>.
32. Dogra P, Ruiz-Ramírez J, Sinha K, Butner JD, Peláez MJ, Rawat M, Yellepeddi VK, Pasqualini R, Arap W, Sostman HD, Cristini V, Wang Z. Innate immunity plays a key role in controlling viral load in COVID-19: mechanistic insights from a whole-body infection dynamics model. *ACS Pharmacol Transl Sci*. 2021;4(1), 265. <https://doi.org/10.1021/acspstci.0c00183>.
33. Voutouri C, Nikmaneshi MR, Hardin CC, Patel AB, Verma A, Khandekar MJ, et al. In silico dynamics of COVID-19 phenotypes for optimizing clinical management. *Proc Natl Acad Sci*. 2021;118(3). <https://doi.org/10.1073/pnas.2021642118>.
34. Kim JY, Ko J-H, Kim Y, Kim Y-J, Kim J-M, Chung Y-S, et al. Viral load kinetics of SARS-CoV-2 infection in first two patients in Korea. *J Korean Med Sci*. 2020;35(7):1–7.
35. Lui G, Ling L, Lai CKC, Tso EYK, Fung KSC, Chan V, et al. Viral dynamics of SARS-CoV-2 across a spectrum of disease severity in COVID-19. *J Infect*. 2020;1(2):1–5. <https://doi.org/10.1016/j.jinf.2020.04.014>.
36. Scott SE, Zabel K, Collins J, Hobbs KC, Kretschmer MJ, Lach M, et al. First Mildly Ill, Nonhospitalized Case of Coronavirus Disease 2019 (COVID-19) Without Viral Transmission in the United States—Maricopa County, Arizona, 2020. *Clin Infect Dis*. 2020;382(15):1196–9. <https://doi.org/10.1093/cid/ciaa374>.
37. Kim SE, Jeong HS, Yu Y, Shin SU, Kim S, Oh TH, et al. Viral kinetics of SARS-CoV-2 in asymptomatic carriers and presymptomatic patients. *Int J Infect Dis*. 2020;95:441–3. <https://doi.org/10.1016/j.ijid.2020.04.083>.
38. Gautret P, Lagier J-C, Parola P, Hoang VT, Meddeb L, Mailhe M, et al. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. *Int J Antimicrob Agents*. 2020;56(1):105949. <https://doi.org/10.1016/j.ijantimicag.2020.105949>.
39. Young BE, Ong SWX, Kalimuddin S, Low JG, Tan SY, Loh J, et al. Epidemiologic Features and Clinical Course of Patients Infected With SARS-CoV-2 in Singapore. *JAMA*. 2020;323(15):1415–88. <https://doi.org/10.1001/jama.2020.3204>.
40. The COVID-19 Investigation Team. Clinical and virologic characteristics of the first 12 patients with coronavirus disease 2019 (COVID-19) in the United States. *Nat Med*. 2020;26(6):861–8.
41. Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, et al. Virological assessment of hospitalized patients with COVID-2019. *Nature*. 2020;581(7809):465–9. <https://doi.org/10.1038/s41586-020-2196-x>.
42. Vetter P, Eberhardt CS, Meyer B, Martinez Murillo PA, Torriani G, Pigny F, et al. Daily viral kinetics and innate and adaptive immune response assessment in COVID-19: a case series. *mSphere*. 2020;5(6):e00827–0.
43. Lavezzo E, Franchin E, Ciavarella C, Cuomo-Dannenburg G, Barzon L, Del Vecchio C, et al. Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. *Nature*. 2020;584(7821):425–9. <https://doi.org/10.1038/s41586-020-2488-1>.
44. Xu Y, Li X, Zhu B, Liang H, Fang C, Gong Y, et al. Characteristics of pediatric SARS-CoV-2 infection and potential evidence for persistent fecal viral shedding. *Nat Med*. 2020;26(4):502–5. <https://doi.org/10.1038/s41591-020-0817-4>.
45. Shrestha NK, Marco Canosa F, Nowacki AS, Procop GW, Vogel S, Fraser TG, et al. Distribution of transmission potential during nonsevere COVID-19 illness. *Clin Infect Dis*. 2020;69:476–7.
46. Yilmaz A, Marklund E, Andersson M, Nilsson S, Andersson L-M, Lindh M, et al. Upper Respiratory Tract Levels of Severe Acute Respiratory Syndrome Coronavirus 2 RNA and Duration of Viral RNA Shedding Do Not Differ Between Patients With Mild and Severe/Critical Coronavirus Disease 2019. *J Infect Dis*. 2020;20(1):564–5. <https://doi.org/10.1093/infdis/jiaa632>.
47. Alsharrah D, Alhaddad F, Alyaseen M, Aljamaan S, Almutairi N, Ayed M, et al. Clinical characteristics of pediatric SARS-CoV-2 infection and coronavirus disease 2019 (COVID-19) in Kuwait. *J Med Virol*. 2021;93(5):3246–50. <https://doi.org/10.1002/jmv.26684>.
48. Salvatore PP, Dawson P, Wadhwa A, Rabold EM, Buono S, Dietrich EA, et al. Epidemiological correlates of polymerase chain reaction cycle threshold values in the detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). *Clin Infect Dis*. 2020;81:e45–7.
49. Prebensen C. Severe acute respiratory syndrome coronavirus 2 RNA in plasma is associated with intensive care unit admission and mortality in patients hospitalized with coronavirus disease 2019. *Clin Infect Dis*. 2020;8(3):1–4. <https://doi.org/10.1093/cid/ciaa1338>.
50. Wylie AL, Fournier J, Casanovas-Massana A, Campbell M, Tokuyama M, Vijayakumar P, et al. Saliva or nasopharyngeal swab specimens for detection of SARS-CoV-2. *N Engl J Med*. 2020;28(13):1283–6. <https://doi.org/10.1056/NEJMc2016359>.
51. Marshall JC, Murthy S, Diaz J, Adhikari NK, Angus DC, Arabi YM, et al. A minimal common outcome measure set for COVID-19 clinical research. *Lancet Infect Dis*. 2020;20(8):e192–7.
52. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat Comput*. 2017;27(5):1432. <https://doi.org/10.1007/s11222-016-9696-4>.
53. Cunningham AJ. The Importance of Pathogen Load. *PLoS Pathog*. 2015;11(1). <https://doi.org/10.1371/journal.ppat.1004563>.
54. Georgiadou A, Lee HJ, Walther M, van Beek AE, Friani F, Wouters D, et al. Modelling pathogen load dynamics to elucidate mechanistic determinants of host–Plasmodium falciparum interactions. *Nat Microbiol*. 2019;4(9):1592–602.
55. He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med*. 2020;26(5):672–5. <https://doi.org/10.1038/s41591-020-0869-5>.
56. Breathnach AS, Duncan CJA, El Bouzidi K, Hanrath AT, Payne BAI, Randell PA, et al. Prior COVID-19 protects against reinfection, even in the absence of detectable antibodies. *J Infect*. 2021;83(2):237–79. <https://doi.org/10.1016/j.jinf.2021.05.024>.
57. Zhou B, She J, Wang Y, Ma X. Duration of Viral shedding of discharged patients with severe COVID-19. *Clin Infect Dis*. 2020;17(16):1–3. <https://doi.org/10.1093/cid/ciaa451>.
58. Mallett S, Allen AJ, Graziadio S, Taylor SA, Sakai NS, Green K, et al. At what times during infection is SARS-CoV-2 detectable and no longer detectable using RT-PCR-based tests? A systematic review of individual participant data. *BMC Med*. 2020;18(1):346.
59. Han MS, Byun J-H, Cho Y, Rim JH. RT-PCR for SARS-CoV-2: quantitative versus qualitative. *Lancet Infect Dis*. 2020;19:1.
60. Chan JF-W, Yip CC-Y, To KK-W, Tang TH-C, Wong SC-Y, Leung K-H, et al. Improved molecular diagnosis of COVID-19 by the novel, highly sensitive and specific COVID-19-RdRp/HeI real-time reverse transcription-PCR Assay Validated *In Vitro* and with Clinical Specimens. *J Clin Microbiol*. 2020;4(5): 58(5). <https://doi.org/10.1128/JCM.00310-20>.
61. Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, et al. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *N Engl J Med*. 2020;19:382(12).
62. Yoon JG, Yoon J, Song JY, Yoon S-Y, Lim CS, Seong H, et al. Clinical significance of a high SARS-CoV-2 viral load in the saliva. *J Korean Med Sci*. 2020;35(20):1–6. <https://doi.org/10.3346/jkms.2020.35.e195>.
63. Stan Development Team. RStan: the R interface to Stan. 2020. p. R package version 2.21.1.
64. McElreath R. rethinking: Statistical rethinking book package. 2020. p. R package version 2.01.

65. McElreath R. Statistical rethinking: A bayesian course with examples in R and stan. In: Statistical Rethinking: A Bayesian Course with Examples in R and Stan. 1st ed. Abingdon: Chapman and Hall/CRC; 2016.
66. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw.* 2015;67(1):1–48.
67. FitzJohn R, Hinsley W. dde: Solve Delay Differential Equations. 2021. p. R package version 1.0.3.
68. Verity B, Winskill P. drjacob: Flexible Markov Chain Monte Carlo via Reparameterization. 2021. p. R package version 1.3.0.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

