Modeling long COVID dynamics: Impact of underlying health conditions

Jie Bai [a,*], Jin Wang [b]

[a] School of Mathematics and Statistics, Liaoning University, Shenyang 110036, China
[b] Department of Mathematics, University of Tennessee at Chattanooga, Chattanooga TN 37403, USA

ARTICLE INFO

ABSTRACT

We propose a new mathematical model to investigate the population dynamics of long COVID, with a focus on the impact of chronic health conditions. Our model connects long COVID with the transmission of COVID-19 so as to accurately predict the prevalence of long COVID from the progression of the infection in the host population. The model additionally incorporates the effects of COVID-19 vaccination. We implement the model with data from both the US and the UK to demonstrate the real-world applications of this modeling framework.

1. Introduction

Long COVID, also known as post-COVID-19 syndrome or post-acute sequelae of COVID-19, is a long-lasting disorder that arises following infection with SARS-CoV-2 (Alwan and Johnson, 2021; Davis et al., 2021; Anon, 2023a). Long COVID can be caused by many different mechanisms, including consequences from acute SARS-CoV-2 injury to one or more organs, autoimmunity due to molecular mimicry between the pathogen and host proteins, re-activation of neurotrophic pathogens under conditions of COVID-19 immune dysregulation, SARS-CoV-2 interactions with host microbiome communities, and dysfunctional brainstem/vagus nerve signaling (Proal and VanElzakker, 2021). In addition, the virus may linger in the body and trigger persistent immune response that drives the symptoms (Marshall, 2021). The situation is further complicated by the fact that individuals with mild symptoms could become COVID long-haulers (Nalbandian et al., 2021). Moreover, recent publications reported long-term COVID-19 symptoms among fully vaccinated people who developed breakthrough infections (Bergwerk et al., 2021; Massey et al., 2021).

Underlying health conditions have been strongly linked to both the severity of COVID-19 infection and the risk of long COVID. It was estimated that more than 1.7 billion people worldwide had at least one underlying medical conditions that put them at increased risk for severe illness caused by COVID-19 (Clark et al., 2020). A study conducted by CDC found that among individuals with reported underlying conditions, hospitalizations were 6 times higher and deaths were 12 times higher compared to those without a chronic condition (Stokes et al., 2020). Meanwhile, it was found that people with asthma, lung disease and several other chronic conditions were more likely to develop long COVID (Sudre et al., 2021). Furthermore, a recent review article (Nalbandian et al., 2021) concluded that COVID-19 patients with

underlying conditions such as obesity, type-2 diabetes, hypertension, chronic pulmonary disease, and compromised immunity were at high risk for long COVID. It is clear that these underlying health conditions play an important role in shaping the prevalence of long COVID.

A systematic review of 33 long COVID cohort studies showed that the post-COVID-19 symptoms and their prevalence rates, ranging between 10% and 67%, are highly heterogeneous across different populations (Fernández-de Las-Peñas et al., 2021). Another review article (Nalbandian et al., 2021) provided an overview of studies based on various COVID-19 cohorts (hospitalized and non-hospitalized) around the world and found estimates for the prevalence of long COVID ranging from 12% to 55%. Such a heterogeneous pattern for long COVID was also highlighted in other systematic reviews and meta-analyses (Cares-Marambio et al., 2021; López-León et al., 2021; Willi et al., 2021). At the population level, it is critical to predict and evaluate the burden of long COVID to assess its impact on the healthcare system and to design appropriate strategies for resource distribution. However, the heterogeneity and variation for the COVID syndrome indicate that there is generally no ready-to-use formula to quantify long COVID in a given population, and the prediction of long COVID prevalence has to take into account the epidemic/pandemic progression and the specific population characteristics, especially the prevalence of underlying health conditions.

Mathematical and computational modeling could help us to address such a pressing need and gain a deeper understanding of the population dynamics associated with long COVID. A large number of mathematical, statistical and computational models have already been developed for COVID-19 dynamics (see reviews in Wang (2020), Afzal et al. (2022), Napolitano et al. (2022) and Padmanabhan et al. (2021)). To our knowledge, however, no mechanistic models have been published

for long COVID thus far. Some predictive work for potential patients with long COVID has been conducted using simple statistical and learning models (Pfaff et al., 2022; Cervia et al., 2022). Particularly, a machine learning study based on the random forest algorithm was published to identify individuals at risk for long COVID (Sudre et al., 2021), and a statistical classification algorithm was applied to identify phenotypes that indicate long COVID (Estiri et al., 2021). In addition, a regression model based on a single exponential decay function was proposed to estimate the lost quality-adjusted life years (QALYs) from long COVID in the UK (Martin et al., 2021). These existing models, however, are not helpful in assessing the population dynamics of long COVID.

As a pilot study, we propose a mechanistic model based on differential equations in this paper to investigate long COVID dynamics and population-level prevalence, with a focus on the impact of chronic health conditions. Our model will divide the host population into two groups depending on whether they have underlying conditions, and each group is associated with different infection characteristics and different risk of developing long COVID. We will connect the population dynamics of long COVID with the transmission dynamics of COVID-19, so as to accurately predict the prevalence of long COVID from the progression of the acute infection. We will additionally incorporate the effects of COVID-19 vaccination in our model.

The remainder of this paper proceeds as follows. The mathematical formulation of our model is presented in Section 2. Main mathematical results concerned with this model are summarized in Section 3, with details provided in Appendix. Numerical simulation is conducted in Section 4 to illustrate the applications of our model. Finally, conclusions are drawn and some discussion is made in Section 5.

2. Model formulation

We utilize mathematical modeling with differential equations to investigate the population dynamics of long COVID with the impact of underlying health conditions. We divide the host population into two groups: Group 1 for individuals with underlying conditions who are more likely to develop long COVID, and Group 2 for individuals without underlying conditions who are less likely to develop long COVID.

The following equations describe the transmission dynamics of COVID-19 in Group 1:

$$
\begin{aligned}
\frac{dS_1}{dt} &= \Lambda_1 - \beta S_1(I_1 + I_2) - (\phi + \mu)S_1, \\
\frac{dV_1}{dt} &= \phi S_1 - \theta \beta V_1(I_1 + I_2) - \mu V_1, \\
\frac{dI_1}{dt} &= \beta(I_1 + I_2)(S_1 + \theta V_1) - (\gamma_1 + w_1 + \mu)I_1.
\end{aligned}
\tag{2.1}
$$

The following equations describe the transmission dynamics of COVID-19 in Group 2:

$$
\begin{aligned}
\frac{dS_2}{dt} &= \Lambda_2 - \beta S_2(I_1 + I_2) - (\phi + \mu)S_2, \\
\frac{dV_2}{dt} &= \phi S_2 - \theta \beta V_2(I_1 + I_2) - \mu V_2, \\
\frac{dI_2}{dt} &= \beta(I_1 + I_2)(S_2 + \theta V_2) - (\gamma_2 + w_2 + \mu)I_2.
\end{aligned}
\tag{2.2}
$$

For each group $j = 1, 2$, the variables S_j, V_j and I_j represent the susceptible, vaccinated, and (short-term) infected individuals, and the parameters Λ_j, γ_j and w_j denote the population influx rate, exit rate from the acute infection, and disease-induced death rate, respectively. Since infected individuals with underlying medical conditions usually exhibit more severe illness than those without such conditions, we assume that $\gamma_1 < \gamma_2$ and $w_1 > w_2$. Meanwhile, the parameter β is the disease transmission rate which may be time-dependent, μ is the natural death rate for the human hosts, and ϕ is the vaccination rate, which take the same values for both groups. In addition, we assume that the vaccine has a degree of protection, or efficacy rate, $1 - \theta$; that

is, a portion of θ in vaccinated individuals are at risk for breakthrough infections (Kates et al., 2021).

Let I_L denote the number of individuals with long COVID, and R denote the number of recovered individuals. The time evolution of the long COVID compartment is described by the following equation:

$$
\frac{dI_L}{dt} = \gamma_1 p_1 I_1 + \gamma_2 p_2 I_2 - (\gamma_L + w_L + \mu)I_L.
\tag{2.3}
$$

Meanwhile, the recovered compartment satisfies the equation below:

$$
\frac{dR}{dt} = \gamma_1(1 - p_1)I_1 + \gamma_2(1 - p_2)I_2 + \gamma_L I_L - \mu R.
\tag{2.4}
$$

Among the individuals who exit from the acute infection period in group j, we assume that a portion p_j of them will develop long COVID and enter the I_L compartment, whereas the other portion $1 - p_j$ will truly recover from the disease and enter the R compartment, for $j = 1, 2$. We have $p_1 > p_2$, since individuals in group 1 are more likely to develop long COVID. The parameter γ_L denotes the recovery rate, and w_L denotes the disease-induced death rates, for individuals in the long COVID state. We note that COVID long-haulers exhibit higher risk of death compared to those without long COVID (Al-Aly et al., 2021). We make an additional assumption that individuals with long CVOID are not contagious, based on a systematic review of 79 clinical studies on SARS-CoV-2 (Cevik et al., 2021) which found that the duration of infectiousness was limited to a relatively short period (up to 9 days).

3. Mathematical results

The main dynamical properties of our mathematical model can be summarized by the following theorems.

Theorem 3.1. *The system* (2.1)–(2.4) *has a unique disease-free equilibrium (DFE)*

$$
\begin{aligned}
x_0 &= (S_{10}, V_{10}, 0, S_{20}, V_{20}, 0, 0, 0) \\
&= \left(\frac{\Lambda_1}{\phi + \mu}, \frac{\Lambda_1 \phi}{(\phi + \mu)\mu}, 0, \frac{\Lambda_2}{\phi + \mu}, \frac{\Lambda_2 \phi}{(\phi + \mu)\mu}, 0, 0, 0 \right)
\end{aligned}
\tag{3.1}
$$

in a positively invariant domain

$$
\begin{aligned}
\Omega = \Big\{ &(S_1, V_1, I_1, S_2, V_2, I_2, I_L, R) \,\Big|\, S_1, V_1, I_1, S_2, V_2, I_2, I_L, R \geq 0, \\
&S_1 \leq S_{10}, \; S_2 \leq S_{20}, \; V_1 \leq V_{10}, \\
&V_2 \leq V_{20}, \; S_1 + V_1 + I_1 + S_2 + V_2 + I_2 + I_L + R \leq \frac{\Lambda_1 + \Lambda_2}{\mu} \Big\}.
\end{aligned}
\tag{3.2}
$$

The basic reproduction number of the model is given by

$$
\mathcal{R}_0 = \mathcal{R}_{01} + \mathcal{R}_{02} = \frac{\beta(S_{10} + \theta V_{10})}{\gamma_1 + \omega_1 + \mu} + \frac{\beta(S_{20} + \theta V_{20})}{\gamma_2 + \omega_2 + \mu},
\tag{3.3}
$$

where $\mathcal{R}_{01}$ and $\mathcal{R}_{02}$ are the individual reproduction numbers of Groups 1 and 2, respectively.

Theorem 3.2. *When $\mathcal{R}_0 < 1$, the DFE x_0 is the only equilibrium of the system* (2.1)–(2.4) *in Ω. When $\mathcal{R}_0 > 1$, in addition to the DFE, the system* (2.1)–(2.4) *has a unique positive endemic equilibrium*

$$
x_* = (S_{1*}, V_{1*}, I_{1*}, S_{2*}, V_{2*}, I_{2*}, I_{L*}, R_*)
\tag{3.4}
$$

in the domain Ω.

Theorem 3.3. *When $\mathcal{R}_0 < 1$, the disease-free equilibrium x_0 is globally asymptotically stable in the domain Ω.*

Theorem 3.4. *When $\mathcal{R}_0 > 1$, the endemic equilibrium x_* is globally asymptotically stable in the domain Ω.*

These results are standard, showing a threshold at $\mathcal{R}_0 = 1$ that separates two distinct types of dynamical behaviors: disease eradication when $\mathcal{R}_0 < 1$ and disease persistence when $\mathcal{R}_0 > 1$. Our special

Table 1
Parameter values in Application 1.

Parameter	Description	Value	Unit	Source
N	Population size	367 804	person	Anon (2023b)
μ	Natural death rate	2.74×10^{-5}	per day	Anon (2023b)
Λ_1	Population influx rate in Group 1	4.031	person/day	Razzaghi et al. (2020)
Λ_2	Population influx rate in Group 2	6.047	person/day	Razzaghi et al. (2020)
θ	Probability for breakthrough infection	5%	–	Kates et al. (2021)
w_1	Disease-induced death rate in Group 1	0.0144	per day	Stokes et al. (2020)
w_2	Disease-induced death rate in Group 2	0.0012	per day	Yang and Wang (2021)
γ_1	Recovery rate in Group 1	0.08	per day	Anon (2023f)
γ_2	Recovery rate in Group 2	0.12	per day	Anon (2023f)
β	Disease transmission rate	–	/person/day	Fitted
ϕ	Vaccination rate	–	per day	Fitted
w_L	Long COVID related death rate	0.00012	per day	Assumed
γ_L	Recovery rate for Long COVID	1/90	per day	Assumed
p_1	The portion of infected individuals in Group 1 developing long COVID	–	per day	Varied
p_2	The portion of infected individuals in Group 2 developing long COVID	–	per day	Varied

interest here, however, is that these analytical results confirm the impact of the underlying health conditions, as can be seen from the expression of $\mathcal{R}_0$ in Eq. (3.3), and the connection between long COVID (represented by I_L in the model) and the acute infection (represented by I_1 and I_2 in the model). Since long COVID stems from the acute infection, the transmission dynamics of COVID-19 play an important role in shaping the evolution of the long COVID prevalence. Consequently, Theorems 3.3 indicates that if the COVID-19 infection is totally eliminated, long COVID will also be eradicated eventually, whereas Theorem 3.4 indicates that if the COVID-19 infection remains endemic, long COVID will also persist in the host population.

Proofs of Theorems 3.1–3.4 can be established through standard mathematical analysis, and details are provided in Appendix. In what follows, we focus on the application of our model.

4. Simulation results

We have incorporated published data to demonstrate the real-world applications of our model. At present, time series data for long COVID at the population level are very rare. One exception is the UK Office for National Statistics (ONS) that published monthly data for long COVID prevalence in the UK population (Anon, 2023e). In our study, we conducted two modeling applications: one for a place in the US which has high prevalence of chronic conditions and which has surveillance data for COVID-19 but not for long COVID, and the other for the UK where data for both COVID-19 infection and long COVID are available.

Application 1. We first applied our model to Hamilton County, one of the most populous counties in the US state of Tennessee. The prevalence of underlying health conditions, particularly chronic heart disease, chronic obstructive pulmonary disease, diabetes, and obesity, in Hamilton County is as high as 40% (Razzaghi et al., 2020). Thus, a significant portion of the population in Hamilton County is considered highly vulnerable to COVID-19 infection and subsequent long COVID.

There are currently no population-level long COVID data available in the US. Thus our model fitting is based on COVID-19 data, using Eqs. (2.1) and (2.2) only. Hamilton County has a total population of $N = 367,804$ (Anon, 2023b). The population sizes of the two groups in our model were set as $N_1 = 0.4\,N$ and $N_2 = 0.6\,N$, based on an estimate from CDC that about 40% of the residents in the county have at least one underlying health conditions (Razzaghi et al., 2020). We then computed the influx rate of the susceptible individuals in each group by $\Lambda_i = \mu_i N_i$ ($i = 1, 2$), where we set $\mu_1 = \mu_2 = \mu$ as the natural birth and death rate in the region. The recovery period from acute COVID-19 infection varies among different patients (Anon, 2023f), depending on their age, severity of illness, and overall health conditions. We assumed that individuals in Group 1 (with chronic conditions) and Group 2

Table 2
Fitting results in Application 1.

Parameter/Quantity	Period 1	Period 2	Period 3
β	1.094×10^{-6}	9.58×10^{-7}	7.20×10^{-7}
ϕ	0.0285	0.0033	0.0031
NMSE	0.0131	0.00051	0.0059

(without chronic conditions) would recover, on average, in 12.5 days and 8.5 days, respectively, which led to $\gamma_1 = 0.08$ per day and $\gamma_2 = 0.12$ per day. We used the disease-induced mortality rate from Yang and Wang (2021) for Group 2, $w_2 = 1.2 \times 10^{-3}$ per day. Meanwhile, base on the CDC study that COVID-19 related deaths were 12 times higher among those with reported underlying conditions than those without such conditions (Stokes et al., 2020), we took $w_1 = 1.44 \times 10^{-2}$ per day. We assumed that the breakthrough infection rate was approximately $\theta = 5\%$ (Kates et al., 2021). The definitions and values of these parameters are listed in Table 1. Additionally, the values for the disease transmission rate β and the vaccination rate ϕ were obtained from data fitting.

We considered a time frame of 11 months from January 1, 2022 to November 30, 2022. We collected the COVID-19 surveillance data from the Tennessee Department of Health (Anon, 2023d) that included the numbers of daily new infections, active infections, and cumulative cases, as well as vaccination coverage. We then fitted Eqs. (2.1) and (2.2) to the reported data for the number of the daily reported active cases, using the least squares method. Specifically, the error function to be minimized is give by $\sum_{j=1}^{m} \left[I_1(t_j) + I_2(t_j) - Y_j \right]^2$, where m is the number of days in the period of consideration, $I_1(t_j)$ and $I_2(t_j)$ are the model predictions for the active infections on the jth day for Groups 1 and 2, respectively, and Y_j is the reported number of active infections on the jth day, for $j = 1, 2, \ldots, m$.

Our preliminary fitting based on the entire time frame of 11 months yielded unsatisfactory results, mainly due to the different patterns exhibited by the data at different times. To overcome this challenge, we divided the time domain into 3 periods:

Period 1: January 1, 2022 to April 7, 2022;
Period 2: April 8, 2022 to June 30, 2022;
Period 3: July 1, 2022 to November 30, 2022.

We then fitted our model separately to these 3 periods for the two unknown parameters β and ϕ. The reported data provided initial conditions for each period.

The fitted values for β and ϕ are presented in Table 2. We can see that the disease transmission rate β is relatively stable in the entire time frame, though its value is slightly decreased from Period 1 to
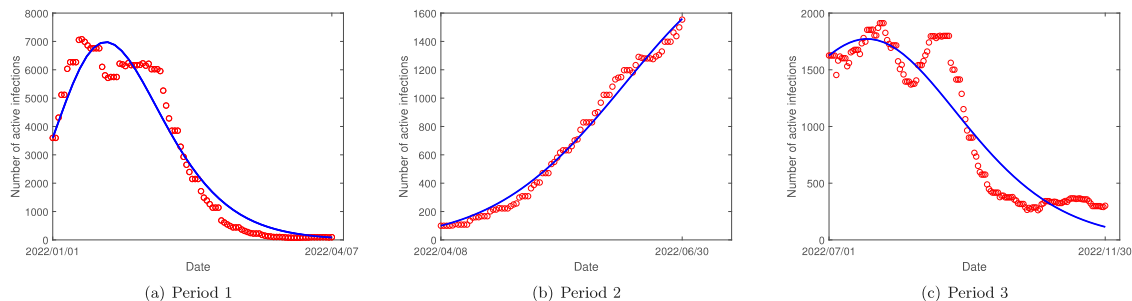
Fig. 1. Number of active COVID-19 cases in Application 1. Red circles represent the reported data and blue solid lines represent the fitting results.

Period 2 and then to Period 3. Meanwhile, the vaccination rate ϕ has a significant reduction from Period 1 to Period 2, and then a small decrease from Period 2 to Period 3. This pattern of change for the values for ϕ reflects the fact that by April 2022, the majority of the population in Hamilton County had received at least two doses of vaccines (primary series plus boosters) (Anon, 2023c) and fewer people would need vaccination afterwards. In addition, we calculated the normalized mean square error (NMSE) for the fitting in each period and listed the results in the last row of Table 2.

The fitting curves for the number of active cases in all the three periods are presented in Fig. 1, in comparison with the reported data. We clearly observe three distinct shapes for these curves: an approximate parabola followed by a gradual decrease in Period 1, a monotonically increasing curve in Period 2, and a dominantly decreasing curve in Period 3. These curves match the reported data well, as evidenced by the computed NMSE (shown in Table 2). This provides a quantitative justification for our piecewise fitting.

Following the data fitting, we conducted a series of simulations for the possible prevalence levels of long COVID in Hamilton County, using Eq. (2.3). We assumed that $\omega_L = 0.1\,\omega_2$ and $\gamma_L = 1/(90 \text{ days})$. We then picked three different values for p_1, the portion of infected individuals from Group 1 who would go on to develop long COVID: $p_1 = 12\%$, 24%, and 36%. Based on the CDC estimates (Stokes et al., 2020) that individuals with underlying health conditions were 6 times more likely to be hospitalized than those without such conditions, we assumed that the risk of developing long COVID in Group 1 is 6 times that in Group 2. This led to $p_2 = 2\%$, 4%, and 6%, respectively. We then ran the simulation for Eq. (2.3), together with Eqs. (2.1) and (2.2), for each of the three periods.

Fig. 2 displays the simulation curves for the number of active long COVID cases, I_L, and for S, V and I in each group, for all the three periods. In particular, the panels in the top row show that the peak of I_L ranges from about 1500 (when $p_1 = 12\%$) to nearly 4500 (when $p_1 = 36\%$). Even with the minimal estimate of $p_1 = 12\%$, the lowest point on the simulation curves in all the three periods is $I_L = 400$, indicating a significant public health burden caused by long COVID. We were not able to fit and predict the evolution of long COVID in this case, due to the unavailability of long COVID data. Nevertheless, our simulation results provided possible ranges for the prevalence of long COVID in the given population, which could inform public health agencies in developing their policies and scaling their efforts.

Additionally, the panels in the bottom row of Fig. 2 shows that the number of infections is disproportionally distributed between the two groups. Although the population size in Group 2 is about 60% of the total; i.e., about 1.5 times of the population size in Group 1, the value of I_2 is only slightly larger than, or even overlapping with, that of I_1 for most of the time. These results confirm the higher risk of infection for individuals in Group 1 (with underlying health conditions).

Application 2. In this study, we applied our model to the UK population. The Office for National Statistics (ONS) published survey data for the prevalence of long COVID in the UK (Anon, 2023e). The period covered by the ONS data started from February 6, 2021. The

Table 3
Number of active long COVID cases in the UK (Anon, 2023e).

Date (year/month/day)	Cases	Date (year/month/day)	Cases
2021/02/06	1 094 000	2021/04/04	1 021 000
2021/05/09	962 000	2021/06/06	945 000
2021/07/04	970 000	2021/08/08	1 086 000
2021/09/04	1 202 000	2021/10/03	1 209 000
2021/11/09	1 266 000	2021/12/06	1 332 000
2022/01/04	1 528 000	2022/02/06	1 724 000
2022/03/07	1 796 000	2022/04/04	1 988 000
2022/05/08	1 950 000	2022/06/05	1 790 000
2022/07/04	1 985 000	2022/08/07	2 290 000
2022/09/04	2 134 000	2022/10/10	2 180 000
2022/11/07	2 131 000	2022/12/06	1 977 000

time series data were published approximately once a month. Based on the ONS data, we list the number of active long COVID cases in the UK from February 6, 2021 to December 6, 2022 in Table 3, which were used for our model fitting and testing.

Instead of fitting again the regular COVID-19 data in the UK (as what we did for Hamilton County in the US), we took advantage of the available long COVID data and focused our attention on the fitting of active long COVID cases (see Table 3) using Eq. (2.3) and the least squares method. We chose the period from February 6, 2021 to July 4, 2022 (about 17 months) for model fitting, and the period from July 5, 2022 to December 6, 2022 (about 5 months) for model testing/prediction.

We took the same values as before for the parameters γ_1, γ_2, and w_L. The natural death rate μ was calculated using the demographic information of the UK population. Data for the daily number of active COVID-19 infections reported in the UK were used to determine I_1 and I_2 in Eq. (2.3), where we incorporated the finding from Walker et al. (2021) that about 25% of the UK population had at least one underlying health conditions that put them at higher risk for COVID-19. We then fitted Eq. (2.3) to the monthly reported long COVID data for the three key parameters p_1, p_2, and γ_L associated with the long COVID prevalence over the 17-month fitting period. Next, we conducted numerical simulation to generate a prediction for the 5-month testing period, using the parameter values estimated from the fitting period.

Table 4 lists the fitted parameter values. Fig. 3 displays the simulation curves for the number of active long COVID cases in the UK, together with the published ONS data. The vertical dashed line separates the fitting and prediction periods. We found that NMSE ≈ 0.019 for the fitting, and NMSE ≈ 0.013 for the testing.

Based on our numerical results, about 35% of the infected individuals in Group 1 and about 28% of the infected individuals in Group 2 went on to develop long COVID in the UK, confirming the higher risk of long COVID for individuals with underlying conditions. Meanwhile, the average recovery rate for long COVID was $\gamma_L \approx 0.012$ per day. This means that long COVID, on average, would last about $1/\gamma_L \approx 83$ days among the UK population. These numbers would provide useful
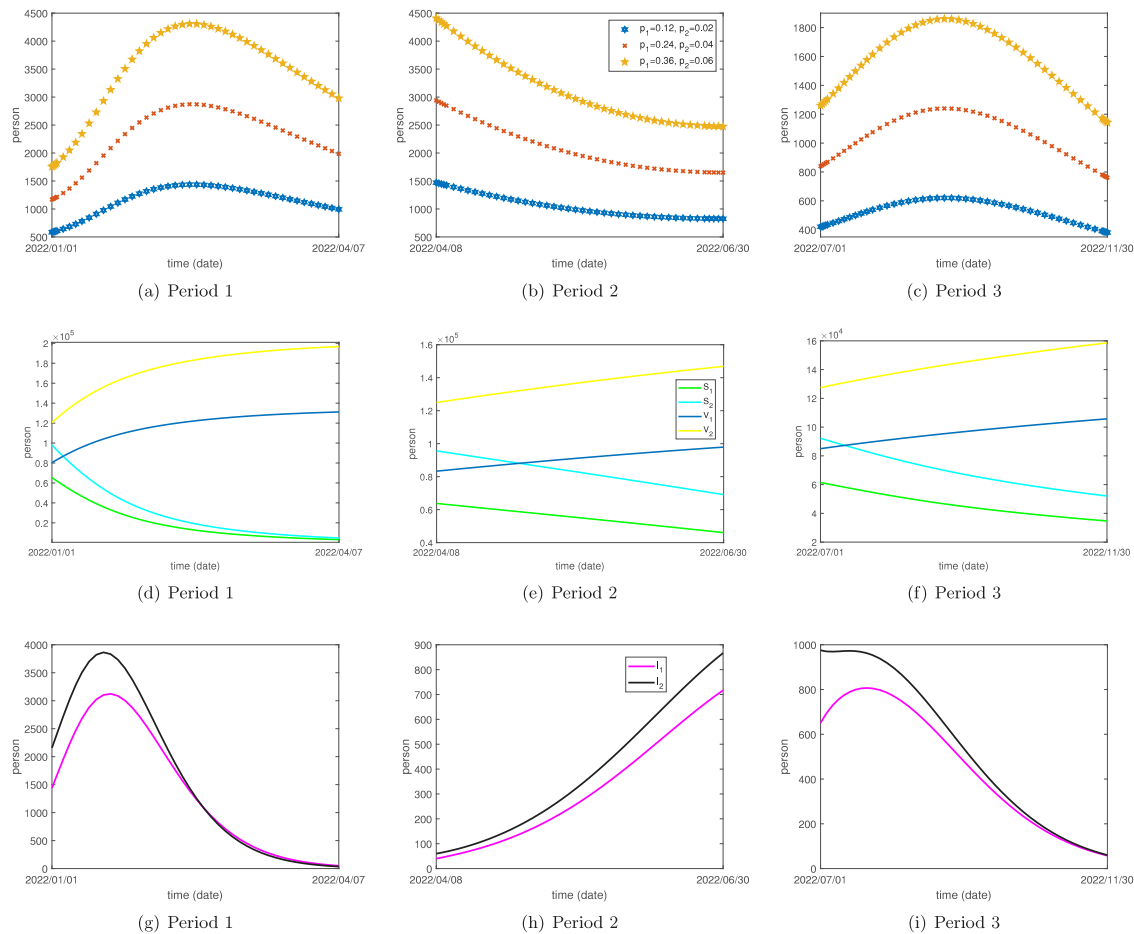
Fig. 2. Simulation curves for Application 1. (a)–(c): Long COVID prevalence with three choices of p_1 and p_2; (d)–(f): Number of susceptible and vaccinated individuals in Groups 1 and 2; (g)–(i): Number of active infections in Groups 1 and 2.

Table 4
Parameter values in Application 2.

Parameter	Description	Value	Unit	Source
μ	Natural death rate	3.91×10^{-5}	per day	Rahman and Kuddus (2021)
γ_1	Recovery rate in Group 1	0.08	per day	Anon (2023f)
γ_2	Recovery rate in Group 2	0.12	per day	Anon (2023f)
w_L	Long COVID related death rate	0.00012	per day	Assumed
γ_L	Recovery rate for Long COVID	0.01198	per day	Fitted
p_1	The portion of infected individuals in Group 1 developing long COVID	0.345	per day	Fitted
p_2	The portion of infected individuals in Group 2 developing long COVID	0.280	per day	Fitted

information to assess the burden of long COVID and to guide policy development and resource allocation.

We make two additional remarks for our numerical results in the UK application. First, we have used the same values for the disease recovery rates (γ_1 and γ_2) in both the US and UK populations, as shown in Table 1 and Table 4. To examine whether this is reasonable, we conduct another numerical test by adding γ_1 and γ_2 into the list of parameters to be fitted; i.e., we use the long COVID data in the UK to fit the parameters γ_1 and γ_2, together with p_1, p_2 and γ_L. We obtain $\gamma_1 \approx 0.097$ and $\gamma_2 \approx 0.118$ in this case, which are very close to the values given in Table 4. Meanwhile, the fitting and prediction curves for the UK long COVID cases are presented in Fig. 4 for this setting. Compared to Fig. 3, we see that the results are almost identical.

Second, we note that Fig. 3 provides a baseline scenario for our model outcomes based on the simplest fitting technique, with all parameters fixed as constants during the entire period. There are various ways to improve the fitting and prediction. For example, Fig. 5 shows a scenario where we divide the fitting period into sub-period 1 (from

February 6, 2021 to April 3, 2022) and sub-period 2 (from April 4, 2022 to July 4, 2022) and where we consider the 7-day moving average of the UK data. We then conduct piecewise fitting separately for these two sub-periods, where the fitting result from sub-period 1 is used to provide the initial condition for sub-period 2, and the fitting result from sub-period 2 is used to generate the prediction in the prediction period. Compared to Fig. 3, we observe better performance for both the fitting and prediction in Fig. 5.

5. Conclusions

Despite many theoretical discoveries and clinical advances (including the development of efficacious vaccines) for COVID-19, our current knowledge and intervention strategies for long COVID remain very limited. Quantitative and predictive studies are urgently needed to determine at-risk population groups for long COVID, to engage in science-based policy development and resource allocation, and to target early intervention strategies and clinical services. This paper represents
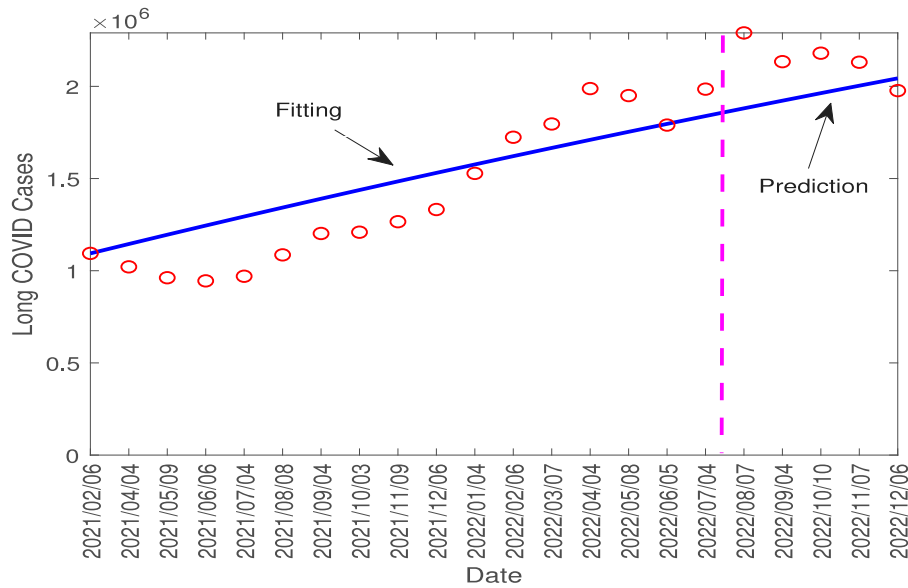
Fig. 3. Fitting and prediction results for the long COVID cases in Application 2. The red circles represent the reported data and the blue solid line represents the numerical results.
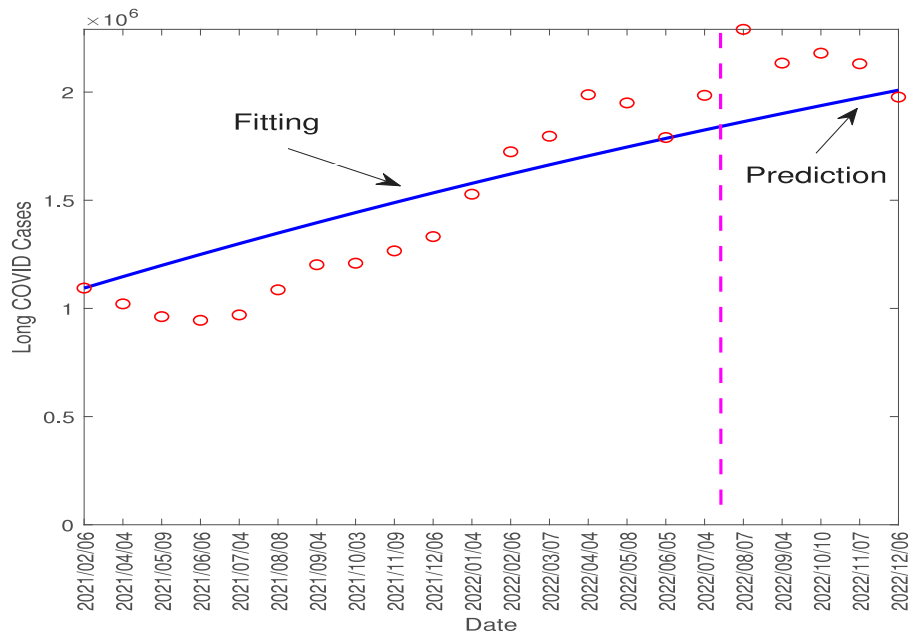

Fig. 4. Fitting and prediction results for the long COVID cases in Application 2, with γ_1 and γ_2 added into the set of parameters to be fitted. The red circles represent the reported data and the blue solid line represents the numerical results.

a proof of concept for the investigation of long COVID through quantitative and computational means. We have developed a novel mechanistic model based on differential equations to study the population dynamics of long COVID with the impact of underlying health conditions. Our model connects long COVID prevalence with the population-level transmission of COVID-19, emphasizing the interplay between the COVID-19 infection, vaccination, chronic conditions, and long COVID dynamics. We have carried out a detailed mathematical analysis for the model. We have also conducted numerical simulation to validate this modeling framework using real data from the US and the UK.

Our first model application is based on Hamilton County in the US, where data for COVID-19 are available but those for long COVID are not, representing a typical scenario at present in terms of data availability. The place also has high prevalence of underlying health conditions. We fitted our model to the COVID-19 data, based on which we simulated the progression of long COVID using several prescribed values for

the parameters p_1 and p_2 that characterize the risk of developing long COVID. The simulation results could provide useful information for the range of the long COVID prevalence at the population level.

Our second model application targets the UK which, as an exceptional case, has published data for both COVID-19 and long COVID at the population level. We fitted our model to the monthly long COVID data over a period of 17 months, and then tested the results by running the model for another 5 months immediately following the fitting period. We found that the percentage of infected individuals with underlying medical conditions to develop long COVID (35%) was significantly higher than that of infected individuals without underlying conditions (28%). The findings confirm that underlying health conditions contribute to increased risk of long COVID.

These two application studies demonstrate the utility and validity of our modeling framework. They have shown that our model can not only analyze the population dynamics of long COVID, but also predict
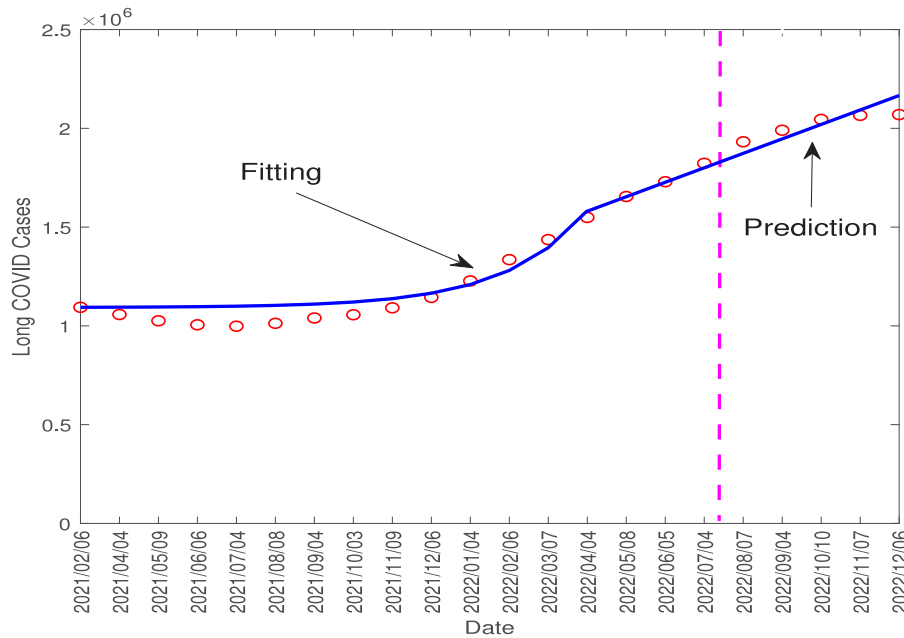
Fig. 5. Fitting and prediction results for the long COVID cases in Application 2, with the fitting period divided into two sub-periods based on piecewise fitting. The red circles represent the reported data and the blue solid line represents the numerical results.

the progression of long COVID in the future (with the availability of historical data). Our work is a pilot effort toward a better understanding of the population dynamics of long COVID. The findings in this study contribute to quantitative knowledge associated with the population-level prevalence and burden of long COVID, and their interplay with the transmission of COVID-19, the impact of underlying health conditions, and the progression from acute infection to long-lasting disorder.

It is anticipated that more long COVID data at the population level will be generated and published in the near future. For example, the US Department of Health and Human Services initiated a large-scale research action plan for long COVID data collection (Anon, 2022) in August 2022 and this project is currently ongoing. The availability of high-quality long COVID data will facilitate broader applications of computational models and will motivate extensions and improvements of the current modeling framework.

The model proposed in this paper can be extended in several directions. For example, we may conduct an optimal control study, utilizing both mathematical analysis and numerical simulation, to explore an effective intervention strategy for long COVID that could balance the effects of the control measures and the related costs in an optimal way. Meanwhile, the model can be naturally extended to study the impact of long COVID on the economy and society. Long COVID has been recognized as a disability under the Americans with Disabilities Act since 2021. We may use model predictions to estimate the total years lived with disability associated with long COVID, and to measure the burden of long COVID on quality of life and on society. Furthermore, we may introduce additional variables such as the employment rate and the economic development level into the modeling framework to study the interplay between long COVID and the economy, particularly the negatively impacted labor force and the lost productivity due to long COVID.

Declaration of competing interest

We declare that there is no conflict of interest in this work. We also confirm that this paper has not been published elsewhere and that it is not under consideration for publication by another journal.

Appendix. Mathematical analysis

Proof of Theorem 3.1. It is obvious that x_0 in Eq. (3.1) is a disease-free equilibrium (DFE) of the system, and the DFE is unique. The positive invariance of the domain Ω can also be easily verified. In particular, adding up all the equations in (2.1)–(2.4), we obtain

$$\frac{dN}{dt} = \Lambda_1 + \Lambda_2 - w_1 I_1 - w_2 I_2 - w_L I_L - \mu N, \tag{A.1}$$

where $N = S_1 + V_1 + I_1 + S_2 + V_2 + I_2 + I_L + R$. Eq. (A.1) yields

$$N \leq \frac{\Lambda_1 + \Lambda_2}{\mu}. \tag{A.2}$$

The basic reproduction number $\mathcal{R}_0$ can be derived from the classical next-generation matrix technique (van den Driessche and Watmough, 2002), with the new infection matrix F and the transition matrix G given by

$$F = \begin{pmatrix} \beta S_{10} + \beta\theta V_{10} & \beta S_{10} + \beta\theta V_{10} & 0 \\ \beta S_{20} + \beta\theta V_{20} & \beta S_{20} + \beta\theta V_{20} & 0 \\ \gamma_1 p_1 & \gamma_2 p_2 & 0 \end{pmatrix}, \tag{A.3}$$

and

$$G = \begin{pmatrix} \gamma_1 + \omega_1 + \mu & 0 & 0 \\ 0 & \gamma_2 + \omega_2 + \mu & 0 \\ 0 & 0 & \gamma_L + \omega_L + \mu \end{pmatrix}. \tag{A.4}$$

The basic reproduction number is the spectral radius of the next-generation matrix FG^{-1}; i.e.,

$$\mathcal{R}_0 = \rho(FG^{-1}) = \mathcal{R}_{01} + \mathcal{R}_{02} = \frac{\beta S_{10} + \beta\theta V_{10}}{\gamma_1 + \omega_1 + \mu} + \frac{\beta S_{20} + \beta\theta V_{20}}{\gamma_2 + \omega_2 + \mu}. \tag{A.5}$$

Each of the two individual reproduction numbers $\mathcal{R}_{01}$ and $\mathcal{R}_{02}$ is a combination of the disease risk from the primary infection and that from the breakthrough infection.

Proof of Theorem 3.2. At an equilibrium of the system, we have

$$\Lambda_1 - \beta S_1(I_1 + I_2) - (\phi + \mu)S_1 = 0, \tag{A.6}$$

$$\phi S_1 - \theta \beta V_1(I_1 + I_2) - \mu V_1 = 0, \tag{A.7}$$

$$\beta(I_1 + I_2)(S_1 + \theta V_1) - (\gamma_1 + w_1 + \mu)I_1 = 0, \tag{A.8}$$

$$\Lambda_2 - \beta S_2(I_1 + I_2) - (\phi + \mu)S_2 = 0, \tag{A.9}$$

$$\phi S_2 - \theta \beta V_2(I_1 + I_2) - \mu V_2 = 0, \tag{A.10}$$

$$\beta(I_1 + I_2)(S_2 + \theta V_2) - (\gamma_2 + w_2 + \mu)I_2 = 0, \tag{A.11}$$

$$\gamma_1 p_1 I_1 + \gamma_2 p_2 I_2 - (\gamma_L + w_L + \mu)I_L = 0, \tag{A.12}$$

$$\gamma_1(1 - p_1)I_1 + \gamma_2(1 - p_2)I_2 + \gamma_L I_L - \mu R = 0. \tag{A.13}$$

According to Eq. (A.6), we obtain

$$I_2 = \frac{\Lambda_1 - (\phi + \mu)S_1 - \beta S_1 I_1}{\beta S_1}. \tag{A.14}$$

Multiply Eq. (A.6) by S_2 and Eq. (A.9) by S_1, and subtract one from the other. This yields

$$S_1 = \frac{\Lambda_1 S_2}{\Lambda_2}. \tag{A.15}$$

Similarly, multiplying Eq. (A.7) by V_2 and Eq. (A.10) by V_1, and subtracting one from the other, we obtain

$$V_1 = \frac{S_1 V_2}{S_2} = \frac{\Lambda_1 V_2}{\Lambda_2}, \tag{A.16}$$

where the second equality follows Eq. (A.15). Substituting Eqs. (A.15) and (A.16) into (A.8), we have

$$\beta(I_1 + I_2)\left(\frac{\Lambda_1 S_2}{\Lambda_2} + \theta \frac{\Lambda_1 V_2}{\Lambda_2}\right) = (\gamma_1 + w_1 + \mu)I_1. \tag{A.17}$$

By Eq. (A.11), we observe

$$\beta(I_1 + I_2)(S_2 + \theta V_2) = (\gamma_2 + w_2 + \mu)I_2. \tag{A.18}$$

Dividing (A.17) from (A.18), we obtain

$$\frac{\Lambda_1}{\Lambda_2} = \frac{(\gamma_1 + w_1 + \mu)I_1}{(\gamma_2 + w_2 + \mu)I_2}, \tag{A.19}$$

and

$$I_2 = \frac{\Lambda_2(\gamma_1 + w_1 + \mu)I_1}{\Lambda_1(\gamma_2 + w_2 + \mu)}. \tag{A.20}$$

Equating (A.14) and (A.20), we obtain

$$I_1 = \frac{\frac{\Lambda_1}{S_1} - (\phi + \mu)}{\beta(1 + \eta)}, \quad \text{with} \quad \eta = \frac{\Lambda_2(\gamma_1 + \omega_1 + \mu)}{\Lambda_1(\gamma_2 + \omega_2 + \mu)}. \tag{A.21}$$

According to (A.6), we have

$$\beta(I_1 + I_2) = \frac{\Lambda_1 - (\phi + \mu)S_1}{S_1}. \tag{A.22}$$

Substituting Eq. (A.22) separately to (A.7) and (A.8), we obtain

$$V_1 = \frac{\phi S_1}{\theta[\frac{\Lambda_1}{S_1} - (\phi + \mu)] + \mu}, \tag{A.23}$$

and

$$I_1 = \frac{[\frac{\Lambda_1}{S_1} - (\phi + \mu)](S_1 + \theta V_1)}{\gamma_1 + \omega_1 + \mu} = \frac{[\frac{\Lambda_1}{S_1} - (\phi + \mu)](S_1 + \theta \frac{\phi S_1}{\theta[\frac{\Lambda_1}{S_1} - (\phi + \mu)] + \mu})}{\gamma_1 + \omega_1 + \mu}. \tag{A.24}$$

Equating (A.21) and (A.24), we obtain the equation below

$$\left[\frac{\Lambda_1}{S_1} - (\phi + \mu)\right] \cdot \left\{S_1 + \theta \frac{\phi S_1}{\theta[\frac{\Lambda_1}{S_1} - (\phi + \mu)] + \mu} - \frac{\gamma_1 + \omega_1 + \mu}{\beta(1 + \eta)}\right\} = 0. \tag{A.25}$$

The first factor of Eq. (A.25), $\frac{\Lambda_1}{S_1} - (\phi + \mu) = 0$, obviously leads to $S_1 = S_{10}$ which subsequently determines the DFE x_0.

The second factor of Eq. (A.25) can be written as the following equation

$$f(S_1) = \frac{\gamma_1 + \omega_1 + \mu}{\beta(1 + \eta)}, \tag{A.26}$$

where the function f is defined as

$$f(S_1) = S_1 + \theta \frac{\phi S_1}{\theta[\frac{\Lambda_1}{S_1} - (\phi + \mu)] + \mu}, \quad 0 \le S_1 \le \frac{\Lambda_1}{\phi + \mu}. \tag{A.27}$$

Clearly, $f(0) = 0$, and $f(S_1)$ is strictly increasing for $0 \le S_1 \le \frac{\Lambda_1}{\phi + \mu} = S_{10}$. Hence, Eq. (A.26) has a unique positive root $S_1 = S_{1*}$ that satisfies $0 < S_{1*} < S_{10}$, if and only if

$$f(S_{10}) > \frac{\gamma_1 + \omega_1 + \mu}{\beta(1 + \eta)}. \tag{A.28}$$

Through direct algebraic manipulation, it can be easily verified that the condition (A.28) is equivalent to

$$\mathcal{R}_0 > 1, \tag{A.29}$$

where $\mathcal{R}_0$ is defined in Eq. (A.5).

Hence, when $\mathcal{R}_0 > 1$, a positive root $S_1 = S_{1*}$ satisfying $0 < S_{1*} < S_{10}$ can be uniquely determined. Consequently, $V_1, I_1, S_2, V_2, I_2, I_L$, and R at the equilibrium can all be uniquely determined. From the equations derived above, it is straightforward to observe that these components are all positive and that the unique positive equilibrium $x_* = (S_{1*}, V_{1*}, I_{1*}, S_{2*}, V_{2*}, I_{2*}, I_{L*}, R_*) \in \Omega$. In contrast, when $\mathcal{R}_0 < 1$, the system has only one equilibrium in Ω which is the DFE x_0.

Proof of Theorem 3.3. From the system (2.1)–(2.4), we obtain

$$\frac{dI_1}{dt} \le \beta(I_1 + I_2)(S_{10} + \theta V_{10}) - (\gamma_1 + w_1 + \mu)I_1,$$

$$\frac{dI_2}{dt} \le \beta(I_1 + I_2)(S_{20} + \theta V_{20}) - (\gamma_2 + w_2 + \mu)I_2,$$

$$\frac{dI_L}{dt} \le \gamma_1 p_1 I_1 + \gamma_2 p_2 I_2 - (\gamma_L + w_L + \mu)I_L.$$

Let $Y = (I_1, I_2, I_L)^T$. Then we have

$$\frac{dY}{dt} \le (F - G)Y, \tag{A.30}$$

where the matrices F and G are defined in Eqs. (A.3) and (A.4). By the Perron–Frobenius theorem, there exists a non-negative left eigenvector u of the non-negative matrix $G^{-1}F$ with respect to the eigenvalue $\mathcal{R}_0 = \rho(FG^{-1}) = \rho(G^{-1}F)$. We define the Lyapunov function:

$$L = u^T G^{-1} Y. \tag{A.31}$$

Differentiating L along the solutions of the system yields

$$L' = u^T G^{-1} \frac{dY}{dt} \le u^T G^{-1}(F - G)Y = (\mathcal{R}_0 - 1)u^T Y. \tag{A.32}$$

If $\mathcal{R}_0 < 1$, then $L' \le 0$. The equality $L' = 0$ leads to $u^T Y = 0$. Therefore, at least one of the three equations $I_1 = 0$, $I_2 = 0$ and $I_L = 0$ must hold. By one of the three equations, we can obtain that the other two equations also hold. Consequently, the largest invariant set where $L' = 0$ is the DFE $x_0 = (S_{10}, V_{10}, 0, S_{20}, V_{20}, 0, 0, 0)$. By LaSalle's Invariance Principle, the DFE is globally asymptotically stable in Ω when $\mathcal{R}_0 < 1$.

Proof of Theorem 3.4. By Theorem 3.2, there exists a unique positive equilibrium x_* when $\mathcal{R}_0 > 1$. Define the Lyapunov function

$$L = \int_{S_{1*}}^{S_1} \frac{u - S_{1*}}{u} du + \int_{V_{1*}}^{V_1} \frac{u - V_{1*}}{u} du + \int_{I_{1*}}^{I_1} \frac{u - I_{1*}}{u} du$$

$$+ \frac{S_{1*} I_{2*}}{S_{2*} I_{1*}} \int_{S_{2*}}^{S_2} \frac{u - S_{2*}}{u} du + \frac{V_{1*} I_{2*}}{V_{2*} I_{1*}} \int_{V_{2*}}^{V_2} \frac{u - V_{2*}}{u} du + \frac{V_{1*} I_{2*}}{V_{2*} I_{1*}} \int_{I_{2*}}^{I_2} \frac{u - I_{2*}}{u} du.$$

According to (A.16), we have

$$\frac{S_{1*}I_{2*}}{S_{2*}I_{1*}} = \frac{V_{1*}I_{2*}}{V_{2*}I_{1*}}.$$

Using Eqs. (A.6)–(A.11), we calculate the derivative of L along the solutions of Eqs. (2.1) and (2.2):

$$
\begin{aligned}
L' =\ & (1 - \frac{S_{1*}}{S_1})\{\Lambda_1 - \beta S_1 I_1 - \beta S_1 I_2 - (\phi + \mu)S_1 \\
& - [\Lambda_1 - \beta S_{1*}I_{1*} - \beta S_{1*}I_{2*} - (\phi + \mu)S_{1*}]\} \\
& \qquad \text{(denoted as } k_1) \\
& + (1 - \frac{V_{1*}}{V_1})[\phi S_1 - \theta \beta V_1 I_1 - \theta \beta V_1 I_2 - \mu V_1 \\
& - (\phi S_{1*} - \theta \beta V_{1*}I_{1*} - \theta \beta V_{1*}I_{2*} - \mu V_{1*})] \\
& \qquad \text{(denoted as } k_2) \\
& + (1 - \frac{I_{1*}}{I_1})\{(\beta I_1 S_1 + \beta I_2 S_1 + \theta \beta I_1 V_1 + \theta \beta I_2 V_1) - (\gamma_1 + w_1 + \mu)I_1 \\
& - [(\beta I_{1*}S_{1*} + \beta I_{2*}S_{1*} + \theta \beta I_{1*}V_{1*} + \theta \beta I_{2*}V_{1*}) - (\gamma_1 + w_1 + \mu)I_{1*}]\} \\
& \qquad \text{(denoted as } k_3) \\
& + \frac{S_{1*}I_{2*}}{S_{2*}I_{1*}}(1 - \frac{S_{2*}}{S_2})\{\Lambda_2 - \beta S_2 I_1 - \beta S_2 I_2 - (\phi + \mu)S_2 \\
& - [\Lambda_2 - \beta S_{2*}I_{1*} - \beta S_{2*}I_{2*} - (\phi + \mu)S_{2*}]\} \\
& \qquad \text{(denoted as } k_4) \\
& + \frac{V_{1*}I_{2*}}{V_{2*}I_{1*}}(1 - \frac{V_{2*}}{V_2})[\phi S_2 - \theta \beta V_2 I_1 - \theta \beta V_2 I_2 - \mu V_2 \\
& - (\phi S_{2*} - \theta \beta V_{2*}I_{1*} - \theta \beta V_{2*}I_{2*} - \mu V_{2*})] \\
& \qquad \text{(denoted as } k_5) \\
& + \frac{V_{1*}I_{2*}}{V_{2*}I_{1*}}(1 - \frac{I_{2*}}{I_2})\{(\beta I_1 S_2 + \beta I_2 S_2 + \theta \beta I_1 V_2 + \theta \beta I_2 V_2) - (\gamma_2 + w_2 + \mu)I_2 \\
& - [(\beta I_{1*}S_{2*} + \beta I_{2*}S_{2*} + \theta \beta I_{1*}V_{2*} + \theta \beta I_{2*}V_{2*}) - (\gamma_2 + w_2 + \mu)I_{2*}]\} \\
& \qquad \text{(denoted as } k_6).
\end{aligned}
$$

To facilitate the algebraic manipulation, we re-arrange k_1, k_2, k_3, k_4, k_5, k_6 as follows, with each k_i represented as a summation of K_{ij} ($j = 0, 1, 2, \dots$):

$$
\begin{aligned}
k_1 =\ & \phi S_{1*}(2 - \frac{S_1}{S_{1*}} - \frac{S_{1*}}{S_1}) \quad K_{10} + \mu S_{1*}(2 - \frac{S_1}{S_{1*}} - \frac{S_{1*}}{S_1}) \quad K_{11} \\
& + \beta(S_{1*}I_{1*} + S_{1*}I_{2*} - S_1 I_1 - S_1 I_2) \quad K_{12} \\
& - \beta(\frac{S_1^2 I_{1*}}{S_1} + \frac{S_1^2 I_{2*}}{S_1} - S_{1*}I_1 - S_{1*}I_2) \quad K_{13},
\end{aligned}
$$

$$
\begin{aligned}
k_2 =\ & \phi(S_1 - S_{1*}) - \phi(\frac{V_{1*}S_1}{V_1} - \frac{V_{1*}S_{1*}}{V_1}) \quad K_{20} \\
& + \theta \beta(V_{1*}I_{1*} + V_{1*}I_{2*} - V_1 I_1 - V_1 I_2) \quad K_{21} \\
& - \theta \beta(\frac{V_{1*}^2 I_{1*}}{V_1} + \frac{V_{1*}^2 I_{2*}}{V_1} - V_{1*}I_1 - V_{1*}I_2) \quad K_{22} \\
& + \mu V_{1*}(2 - \frac{V_1}{V_{1*}} - \frac{V_{1*}}{V_1}) \quad K_{23},
\end{aligned}
$$

$$
\begin{aligned}
k_3 =\ & \beta(I_1 S_1 + I_2 S_1 + \theta I_1 V_1 + \theta I_2 V_1 - I_{1*}S_{1*} \\
& \quad - I_{2*}S_{1*} - \theta I_{1*}V_{1*} - \theta I_{2*}V_{1*}) \quad K_{30} \\
& - \beta(I_{1*}S_1 + \frac{I_{1*}I_2 S_1}{I_1} + \theta I_{1*}V_1 + \frac{\theta I_{1*}I_2 V_1}{I_1} - \frac{I_{1*}^2 S_{1*}}{I_1} \\
& \quad - \frac{I_{1*}I_{2*}S_{1*}}{I_1} - \frac{\theta I_{1*}^2 V_{1*}}{I_1} - \frac{\theta I_{1*}I_{2*}V_{1*}}{I_1}) \quad K_{31} \\
& + (\gamma_1 + \omega_1 + \mu)I_{1*}(2 - \frac{I_1}{I_{1*}} - \frac{I_{1*}}{I_1}) \quad K_{32},
\end{aligned}
$$

$$
\begin{aligned}
k_4 =\ & \frac{S_{1*}I_{2*}}{S_{2*}I_{1*}}\Big[\phi S_{2*}(2 - \frac{S_2}{S_{2*}} - \frac{S_{2*}}{S_2}) \quad K_{40} + \mu S_{2*}(2 - \frac{S_2}{S_{2*}} - \frac{S_{2*}}{S_2}) \quad K_{41} \\
& + \beta(S_{2*}I_{1*} + S_{2*}I_{2*} - S_2 I_1 - S_2 I_2) \quad K_{42}
\end{aligned}
$$

$$
\begin{aligned}
& - \beta(\frac{S_{2*}^2 I_{1*}}{S_2} + \frac{S_{2*}^2 I_{2*}}{S_2} - S_{2*}I_1 - S_{2*}I_2) \quad K_{43} \Big],
\end{aligned}
$$

$$
\begin{aligned}
k_5 =\ & \frac{V_{1*}I_{2*}}{V_{2*}I_{1*}}\Big[\phi(S_2 - S_{2*}) - \phi(\frac{V_{2*}S_2}{V_2} - \frac{V_{2*}S_{2*}}{V_2}) \quad K_{50} \\
& + \theta \beta(V_{2*}I_{1*} + V_{2*}I_{2*} - V_2 I_1 - V_2 I_2) \quad K_{51} \\
& - \theta \beta(\frac{V_{2*}^2 I_{1*}}{V_2} + \frac{V_{2*}^2 I_{2*}}{V_2} - V_{2*}I_1 - V_{2*}I_2) \quad K_{52} \\
& + \mu V_{2*}(2 - \frac{V_2}{V_{2*}} - \frac{V_{2*}}{V_2}) \quad K_{53} \Big],
\end{aligned}
$$

$$
\begin{aligned}
k_6 =\ & \frac{V_{1*}I_{2*}}{V_{2*}I_{1*}}\Big[\beta(I_1 S_2 + I_2 S_2 + \theta I_1 V_2 + \theta I_2 V_2 - I_{1*}S_{2*} \\
& \quad - I_{2*}S_{2*} - \theta I_{1*}V_{2*} - \theta I_{2*}V_{2*}) \quad K_{60} \\
& - \beta(I_{2*}S_2 + \frac{I_{2*}I_1 S_2}{I_2} + \theta I_{2*}V_2 + \frac{\theta I_{2*}I_1 V_2}{I_2} - \frac{I_{2*}^2 S_{2*}}{I_2} - \frac{I_{1*}I_{2*}S_{2*}}{I_2} \\
& \quad - \frac{\theta I_{2*}^2 V_{2*}}{I_2} - \frac{\theta I_{1*}I_{2*}V_{2*}}{I_2}) \quad K_{61} \\
& + (\gamma_2 + \omega_2 + \mu)I_{2*}(2 - \frac{I_2}{I_{2*}} - \frac{I_{2*}}{I_2}) \quad K_{62} \Big].
\end{aligned}
$$

It is obvious that $K_{12} + K_{21} + K_{30} = 0$ and $K_{42} + K_{51} + K_{60} = 0$. Because x_* satisfies Eq. (A.8), we have

$$
\begin{aligned}
K_{32} =\ & (\beta I_{1*}S_{1*} + \beta I_{2*}S_{1*} + \theta \beta I_{1*}V_{1*} + \theta \beta I_{2*}V_{1*})(2 - \frac{I_1}{I_{1*}} - \frac{I_{1*}}{I_1}) \\
=\ & \beta I_{1*}S_{1*}(2 - \frac{I_1}{I_{1*}} - \frac{I_{1*}}{I_1}) \quad K_{70} + \beta I_{2*}S_{1*}(2 - \frac{I_1}{I_{1*}} - \frac{I_{1*}}{I_1}) \quad K_{71} \\
& + \theta \beta I_{1*}V_{1*}(2 - \frac{I_1}{I_{1*}} - \frac{I_{1*}}{I_1}) \quad K_{72} + \beta \theta I_{2*}V_{1*}(2 - \frac{I_1}{I_{1*}} - \frac{I_{1*}}{I_1}) \quad K_{73}.
\end{aligned}
$$

Denote the first term of K_{13} by K_{131}, the second term by K_{132}, the third term by K_{133}, etc. With similar notations, we have $K_{131} + K_{311} + K_{701} = \beta I_{1*}S_{1*}(2 - \frac{S_1}{S_{1*}} - \frac{S_{1*}}{S_1}) \le 0$, $K_{133} + K_{702} = 0$, $K_{221} + K_{313} + K_{721} = \theta \beta I_{1*}V_{1*}(2 - \frac{V_1}{V_{1*}} - \frac{V_{1*}}{V_1})$ (denoted as K_8), $K_{223} + K_{722} = 0$, $K_{315} + K_{703} = 0$, $K_{316} + K_{713} = 0$, $K_{317} + K_{723} = 0$, $K_{318} + K_{733} = 0$. Because x_* satisfies Eq. (A.7), we have

$$
\begin{aligned}
K_{23} =\ & (\phi S_{1*} - \theta \beta V_{1*}I_{1*} - \theta \beta V_{1*}I_{2*})(2 - \frac{V_1}{V_{1*}} - \frac{V_{1*}}{V_1}) \\
=\ & \phi S_{1*}(2 - \frac{V_1}{V_{1*}} - \frac{V_{1*}}{V_1}) \quad K_{90} - \theta \beta V_{1*}I_{1*}(2 - \frac{V_1}{V_{1*}} - \frac{V_{1*}}{V_1}) \\
& K_{91} - \theta \beta V_{1*}I_{2*}(2 - \frac{V_1}{V_{1*}} - \frac{V_{1*}}{V_1}) \quad K_{92}.
\end{aligned}
$$

It can be observed that $K_{10} + K_{20} + K_{90} = \phi S_{1*}(3 - \frac{V_{1*}S_1}{V_1 S_{1*}} - \frac{V_1}{V_{1*}} - \frac{S_{1*}}{S_1}) \le 0$, and $K_{91} + K_8 = 0$.

Next, we reorganize the sum of k_1, k_2 and k_3:

$$
\begin{aligned}
& k_1 + k_2 + k_3 \\
=\ & \beta I_{1*}S_{1*}(2 - \frac{S_1}{S_{1*}} - \frac{S_{1*}}{S_1}) + \phi S_{1*}(3 - \frac{V_{1*}S_1}{V_1 S_{1*}} - \frac{V_1}{V_{1*}} - \frac{S_{1*}}{S_1}) \\
& + \mu S_{1*}(2 - \frac{S_1}{S_{1*}} - \frac{S_{1*}}{S_1}) \\
& - \beta \frac{S_1^2 I_{2*}}{S_1} + \beta S_{1*}I_2 - \theta \beta \frac{V_{1*}^2 I_{2*}}{V_1} + \theta \beta V_{1*}I_2 - \theta \beta V_{1*}I_{2*}(2 - \frac{V_1}{V_{1*}} - \frac{V_{1*}}{V_1}) \\
& - \beta \frac{I_{1*}I_2 S_1}{I_1} - \theta \beta \frac{I_{1*}I_2 V_1}{I_1} + \beta I_{2*}S_{1*}(2 - \frac{I_1}{I_{1*}}) + \beta \theta I_{2*}V_{1*}(2 - \frac{I_1}{I_{1*}}) \\
=\ & \beta I_{1*}S_{1*}(2 - \frac{S_1}{S_{1*}} - \frac{S_{1*}}{S_1}) + \phi S_{1*}(3 - \frac{V_{1*}S_1}{V_1 S_{1*}} - \frac{V_1}{V_{1*}} - \frac{S_{1*}}{S_1}) \\
& + \mu S_{1*}(2 - \frac{S_1}{S_{1*}} - \frac{S_{1*}}{S_1}) \\
& + \beta I_{2*}S_{1*}(4 - \frac{I_1}{I_{1*}} - \frac{S_{1*}}{S_1} - \frac{I_{1*}I_2 S_1}{I_1 I_{2*}S_{1*}} - \frac{I_{2*}}{I_2}) - \beta I_{2*}S_{1*}(2 - \frac{I_2}{I_{2*}} - \frac{I_{2*}}{I_2}) \\
& \qquad \text{(denoted as } k_7) \\
& + \beta \theta I_{2*}V_{1*}(4 - \frac{I_1}{I_{1*}} - \frac{V_{1*}}{V_1} - \frac{I_{1*}I_2 V_1}{I_1 I_{2*}V_{1*}} - \frac{I_{2*}}{I_2})
\end{aligned}
$$

$$- \beta \theta I_{2*} V_{1*} (4 - \frac{V_1}{V_{1*}} - \frac{V_{1*}}{V_1} - \frac{I_2}{I_{2*}} - \frac{I_{2*}}{I_2})$$

(denoted as k_8).

With similar manipulations, we obtain

$$k_4 + k_5 + k_6$$

$$= \frac{S_{1*} I_{2*}}{S_{2*} I_{1*}} [\beta I_{2*} S_{2*} (2 - \frac{S_2}{S_{2*}} - \frac{S_{2*}}{S_2}) + \phi S_{2*} (3 - \frac{V_{2*} S_2}{V_2 S_{2*}} - \frac{V_2}{V_{2*}} - \frac{S_{2*}}{S_2})$$

$$+ \mu S_{2*} (2 - \frac{S_2}{S_{2*}} - \frac{S_{2*}}{S_2})]$$

$$+ \frac{S_{1*} I_{2*}}{S_{2*} I_{1*}} [\beta I_{1*} S_{2*} (4 - \frac{I_{1*}}{I_1} - \frac{S_{2*}}{S_2} - \frac{I_{2*} I_1 S_2}{I_2 I_{1*} S_{2*}} - \frac{I_2}{I_{2*}})$$

$$- \beta S_{2*} I_{1*} (2 - \frac{I_1}{I_{1*}} - \frac{I_{1*}}{I_1})]$$

(denoted as k_9)

$$+ \frac{V_{1*} I_{2*}}{V_{2*} I_{1*}} [\beta \theta V_{2*} I_{1*} (4 - \frac{I_{1*}}{I_1} - \frac{V_{2*}}{V_2} - \frac{I_{2*} I_1 V_2}{I_2 I_{1*} V_{2*}} - \frac{I_2}{I_{2*}})$$

$$- \beta \theta V_{2*} I_{1*} (4 - \frac{V_2}{V_{2*}} - \frac{V_{2*}}{V_2} - \frac{I_1}{I_{1*}} - \frac{I_{1*}}{I_1})]$$

(denoted as k_{10}).

$$k_7 + k_9$$

$$= \beta I_{2*} S_{1*} (4 - \frac{I_1}{I_{1*}} - \frac{S_{1*}}{S_1} - \frac{I_{1*} I_2 S_1}{I_1 I_{2*} S_{1*}} - \frac{I_{2*}}{I_2} - 2 + \frac{I_2}{I_{2*}} + \frac{I_{2*}}{I_2}$$

$$+ 4 - \frac{I_{1*}}{I_1} - \frac{S_{2*}}{S_2} - \frac{I_{2*} I_1 S_2}{I_2 I_{1*} S_{2*}} - \frac{I_2}{I_{2*}} - 2 + \frac{I_1}{I_{1*}} + \frac{I_{1*}}{I_1})$$

$$= \beta I_{2*} S_{1*} (4 - \frac{S_{1*}}{S_1} - \frac{I_{1*} I_2 S_1}{I_1 I_{2*} S_{1*}} - \frac{S_{2*}}{S_2} - \frac{I_{2*} I_1 S_2}{I_2 I_{1*} S_{2*}})$$

$$\leq 0.$$

Since

$$\phi S_{1*} - \theta \beta V_{1*} (I_{1*} + I_{2*}) - \mu V_{1*} = 0,$$

$$\phi S_{2*} - \theta \beta V_{2*} (I_{1*} + I_{2*}) - \mu V_{2*} = 0,$$

we have

$$\phi S_{1*} (3 - \frac{V_{1*} S_1}{V_1 S_{1*}} - \frac{V_1}{V_{1*}} - \frac{S_{1*}}{S_1}) \quad \text{(from } k_1 + k_2 + k_3\text{)}$$

$$= (\beta \theta V_{1*} I_{1*} + \mu V_{1*})(3 - \frac{V_{1*} S_1}{V_1 S_{1*}} - \frac{V_1}{V_{1*}} - \frac{S_{1*}}{S_1})$$

$$+ \beta \theta V_{1*} I_{2*} (3 - \frac{V_{1*} S_1}{V_1 S_{1*}} - \frac{V_1}{V_{1*}} - \frac{S_{1*}}{S_1}) \quad \text{(denoted as } k_{11}\text{)},$$

and

$$\frac{S_{1*} I_{2*}}{S_{2*} I_{1*}} \phi S_{2*} (3 - \frac{V_{2*} S_2}{V_2 S_{2*}} - \frac{V_2}{V_{2*}} - \frac{S_{2*}}{S_2}) \quad \text{(from } k_4 + k_5 + k_6\text{)}$$

$$= \frac{V_{1*} I_{2*}}{V_{2*} I_{1*}} \phi S_{2*} (3 - \frac{V_{2*} S_2}{V_2 S_{2*}} - \frac{V_2}{V_{2*}} - \frac{S_{2*}}{S_2})$$

$$= \frac{V_{1*} I_{2*}}{V_{2*} I_{1*}} (\beta \theta V_{2*} I_{2*} + \mu V_{2*})(3 - \frac{V_{2*} S_2}{V_2 S_{2*}} - \frac{V_2}{V_{2*}} - \frac{S_{2*}}{S_2})$$

$$+ \frac{V_{1*} I_{2*}}{V_{2*} I_{1*}} \beta \theta V_{2*} I_{1*} (3 - \frac{V_{2*} S_2}{V_2 S_{2*}} - \frac{V_2}{V_{2*}} - \frac{S_{2*}}{S_2}) \quad \text{(denoted as } k_{12}\text{)}.$$

We thus obtain

$$k_8 + k_{10} + k_{11} + k_{12}$$

$$= \beta \theta V_{1*} I_{2*} (4 - \frac{I_1}{I_{1*}} - \frac{V_{1*}}{V_1} - \frac{I_{1*} I_2 V_1}{I_1 I_{2*} V_{1*}} - \frac{I_{2*}}{I_2} - 4 + \frac{V_1}{V_{1*}} + \frac{V_{1*}}{V_1} + \frac{I_2}{I_{2*}} + \frac{I_{2*}}{I_2}$$

$$+ 4 - \frac{I_{1*}}{I_1} - \frac{V_{2*}}{V_2} - \frac{I_{2*} I_1 V_2}{I_2 I_{1*} V_{2*}} - \frac{I_2}{I_{2*}} - 4 + \frac{V_2}{V_{2*}} + \frac{V_{2*}}{V_2} + \frac{I_1}{I_{1*}} + \frac{I_{1*}}{I_1}$$

$$+ 3 - \frac{V_{1*} S_1}{V_1 S_{1*}} - \frac{V_1}{V_{1*}} - \frac{S_{1*}}{S_1} + 3 - \frac{V_{2*} S_2}{V_2 S_{2*}} - \frac{V_2}{V_{2*}} - \frac{S_{2*}}{S_2})$$

$$= \beta \theta V_{1*} I_{2*} (6 - \frac{I_{1*} I_2 V_1}{I_1 I_{2*} V_{1*}} - \frac{I_{2*} I_1 V_2}{I_2 I_{1*} V_{2*}} - \frac{V_{1*} S_1}{V_1 S_{1*}} - \frac{S_{1*}}{S_1} - \frac{V_{2*} S_2}{V_2 S_{2*}} - \frac{S_{2*}}{S_2})$$

$$\leq 0.$$

Therefore $L' \leq 0$, and $L' = 0$ if and only if $(S_1, V_1, I_1, S_2, V_2, I_2) = (S_{1*}, V_{1*}, I_{1*}, S_{2*}, V_{2*}, I_{2*})$. Based on LaSalle's Invariance Principle, the endemic equilibrium $(S_{1*}, V_{1*}, I_{1*}, S_{2*}, V_{2*}, I_{2*})$ of the subsystem (2.1) and (2.2) is globally asymptotically stable. Letting $I_1 \to I_{1*}$ and $I_2 \to I_{2*}$ in Eqs. (2.3) and (2.4), we clearly observe that all solutions of I_L approach I_{L*} and that all solutions of R approach R_*. This completes the proof that the endemic equilibrium x_* is globally asymptotically stable in Ω.

References

Afzal, A., Saleel, C.A., Bhattacharyya, S., Satish, N., Samuel, O.D., Badruddin, I.A., 2022. Merits and limitations of mathematical modeling and computational simulations in mitigation of COVID-19 pandemic: A comprehensive review. Arch. Comput. Methods Eng. 29 (2), 1311–1337.

Al-Aly, Z., Xie, Y., Bowe, B., 2021. High-dimensional characterization of post-acute sequelae of COVID-19. Nature 594, 259–264.

Alwan, N.A., Johnson, L., 2021. Defining long COVID: Going back to the start. Med. (N Y) 2 (5), 501–504.

Anon, 2022. Department of Health and Human Services, Office of the Assistant Secretary for Health, National Research Action Plan on Long COVID.

Anon, 2023a. Centers for Disease Control and Prevention: Post-COVID conditions. Available at https://www.cdc.gov/coronavirus/2019-ncov/long-term-effects.html.

Anon, 2023b. Hamilton County Health Department. Available at: https://health.hamiltontn.org.

Anon, 2023c. Hamilton County Health Department: COVID-19 Vaccination Reporting. Available at https://health.hamiltontn.org/en-us/allservices/coronavirus(covid-19).aspx.

Anon, 2023d. Tennessee Department of Health: Coronavirus Disease (COVID-19). Available at https://www.tn.gov/health/cedep/ncov.html.

Anon, 2023e. Office for National Statistics: Prevalence of ongoing symptoms following coronavirus (COVID-19) infection in the UK. Available at https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/.

Anon, 2023f. World Health Organization: Coronavirus disease (COVID-19) situation reports. Available at https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.

Bergwerk, M., Gonen, T., Lustig, Y., et al., 2021. Covid-19 breakthrough infections in vaccinated health care workers. N. Engl. J. Med. 385 (17), 1629–1630.

Cares-Marambio, K., Montenegro-Jiménez, Y., Torres-Castro, R., et al., 2021. Prevalence of potential respiratory symptoms in survivors of hospital admission after coronavirus disease. Chronic Respiratory Dis. 18, 14799731211002240.

Cervia, C., Zurbuchen, Y., Taeschler, P., et al., 2022. Immunoglobulin signature predicts risk of post-acute COVID-19 syndrome. Nature Commun. 13 (1), 446.

Cevik, M., Tate, M., Lloyd, O., Maraolo, A.E., Schafers, J., Ho, A., 2021. 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 2 (1), e13–e22.

Clark, A., Jit, M., Warren-Gash, C., Guthrie, B., Wang, H.H.X., Mercer, S.W., et al., 2020. Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying. Lancet Glob. Health 8, e1003–e1017.

Davis, H.E., Assaf, G.S., McCorkell, L., et al., 2021. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. EClinicalMedicine 38, 101019.

Estiri, H., Strasser, Z.H., Brat, G.A., Semenov, Y.R., 2021. Consortium for characterization of COVID-19 by EHR (4CE), C.J. Patel, and S.N. Murphy, evolving phenotypes of non-hospitalized patients that indicate long COVID. BMC Med. 19, 249.

Kates, J., Dawson, L., Anderson, E., et al., 2021. COVID-19 Vaccine Breakthrough Cases: Data from the States. Kaiser Family Foundation.

Fernández-de Las-Peñas, C., Palacios-Ceña, D., Gómez-Mayordomo, V., Florencio, L.L., Cuadrado, M.L., Plaza-Manzano, G., Navarro-Santana, M., 2021. Prevalence of post-COVID-19 symptoms in hospitalized and non-hospitalized COVID-19 survivors: A systematic review and meta-analysis. Eur. J. Int. Med. 92, 55–70.

López-León, S., Wegman-Ostrosky, T., Perelman, C., Sepulveda, R., Rebolledo, P.A., Cuapio, A., Villapol, S., 2021. More than 50 long-term effects of COVID-19: a systematic review and meta-analysis. Sci. Rep. 11, 16144.

Marshall, M., 2021. The four most urgent questions about long COVID. Nature 594, 168–170.

Martin, C., Luteijn, M., Letton, W., Robertson, J., McDonald, S., 2021. A model framework for projecting the prevalence and impact of long-COVID in the UK. PLoS One 16 (12), e0260843.

Massey, D., Berrent, D., Krumholz, H., 2021. Breakthrough Symptomatic COVID-19 Infections Leading to Long Covid: Report from Long Covid Facebook Group Poll. http://dx.doi.org/10.1101/2021.07.23.21261030, medRxiv.

Nalbandian, A., Sehgal, K., Gupta, A., et al., 2021. Post-acute COVID-19 syndrome. Nat. Med. 27, 601–615.

Napolitano, F., Xu, X., Gao, X., 2022. Impact of computational approaches in the fight against COVID-19: An AI guided review of 17, 000 studies. Brief. Bioinform. 23 (1), bbab456.

Padmanabhan, R., Abed, H.S., Meskin, N., Khattab, T., Shraim, M., Al-Hitmi, M.A., 2021. A review of mathematical model-based scenario analysis and interventions for COVID-19. Comput. Methods Programs Biomed. 209, 106301.

Pfaff, E.R., Girvin, A.T., Bennett, T.D., et al., 2022. Identifying who has long COVID in the USA: a machine learning approach using N3C data. Lancet Digit. Health 4 (7), e532–e541.

Proal, A.D., VanElzakker, M.B., 2021. Long COVID or Post-acute Sequelae of COVID-19 (PASC): An overview of biological factors that may contribute to persistent symptoms. Front. Microbiol. 12, 698169.

Rahman, A., Kuddus, M.A., 2021. Modelling the transmission dynamics of COVID-19 in six high-burden countries. BioMed Res. Int. 2021, 5089184.

Razzaghi, H., Wang, Y., Lu, H., Marshall, K.E., Dowling, N.F., Paz-Bailey, G., et al., 2020. Estimated county-level prevalence of selected underlying medical conditions associated with increased risk. Morb. Mortal. Wkly. Rep. 69, 945–950.

Stokes, E.K., Zambrano, L.D., Anderson, K.N., Marder, E.P., Raz, K.M., Felix, S.E.B., et al., 2020. Coronavirus disease. Morb. Mortal. Wkly. Rep. 69, 759–765.

Sudre, C.H., Murray, B., Varsavsky, T., et al., 2021. Attributes and predictors of long COVID. Nat. Med. 27 (4), 626–631.

van den Driessche, P., Watmough, J., 2002. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Biosci. 180, 29–48.

Walker, J.L., Grint, D.J., Strongman, H., Eggo, R.M., Peppa, M., Minassian, C., et al., 2021. UK prevalence of underlying conditions which increase the risk of severe COVID-19 disease: A point prevalence study using electronic health records. BMC Public Health 21, 484.

Wang, J., 2020. Mathematical models for COVID-19: Applications, limitations, and potentials. J. Public Health Emerg. 4, 9.

Willi, S., Lüthold, R., Hunt, A., et al., 2021. COVID-19 sequelae in adults aged less than 50 years: a systematic review. Travel Med. Infect. Dis. 40, 101995.

Yang, C., Wang, J., 2021. COVID-19 and underlying health conditions: A modeling investigation. Math. Biosci. Eng. 18 (4), 3790–3812.