Research article

Predictability of viral load dynamics in the early phases of SARS-CoV-2 through a model-based approach

Andrea Bondesan[1]**, Antonio Piralla**[2]**, Elena Ballante**[3]**, Antonino Maria Guglielmo Pitrolo**[2]**, Silvia Figini**[3]**, Fausto Baldanti**[2,4] **and Mattia Zanella**[5,*]

[1] Department of Mathematical, Physical and Computer Sciences, University of Parma, Parma, Italy

[2] Microbiology and Virology Department, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

[3] Department of Political and Social Sciences, University of Pavia, Pavia, Italy

[4] Department of Clinical, Surgical, Diagnostic and Pediatric Sciences, University of Pavia, Pavia, Italy

[5] Department of Mathematics "F. Casorati", University of Pavia, Pavia, Italy

* Correspondence: Email: mattia.zanella@unipv.it.

Abstract: A pipeline to evaluate the evolution of viral dynamics based on a new model-driven approach has been developed in the present study. The proposed methods exploit real data and the multiscale structure of the infection dynamics to provide robust predictions of the epidemic dynamics. We focused on viral load kinetics whose dynamical features are typically available in the symptomatic stage of the infection. Hence, the epidemiological evolution was obtained by relying on a compartmental approach characterized by a varying infection rate to estimate early-stage viral load dynamics, of which few data are available. We tested the proposed approach with real data of SARS-CoV-2 viral load kinetics collected from patients living in an Italian province. The considered database refers to early-phase infections, whose viral load kinetics have not been affected by the mass vaccination policies in Italy.

Keywords: mathematical epidemiology; SARS-CoV-2 viral load; SIR with age of infection; uncertainty quantification

1. Introduction

Infection dynamics are typically shaped by heterogeneous and interconnected factors, including exogenous individual-based aspects such as the individual response to threat, the collective compliance to non-pharmaceutical interventions, and variable endogenous characteristics of a pathogen [1–10]. Indeed, the infectivity of individuals may vary strongly from the onset of the infection to its end. In a purely data-oriented approach the day of first contact with an infected individual is typically unknown

or unreported and the infectivity levels may vary in time. Therefore, the trend of an epidemic naturally involves a plurality of timescales: one that is linked to the evolution of infected cases among the population and a second based on the change of infectivity related to in-host viral dynamics [11–13].

In this direction, viral load (VL) dynamics represent a significant aspect influencing the disease progression and transmission. The evolution of VL constitutes a quantitative marker for assessing viral kinetics [14–16]. Nevertheless, the reconstruction of these trajectories is a challenging problem since data are incomplete and biased toward the peak viral load, which is typically related to the onset of symptoms in infected patients. We mention in this direction the empirical studies, see, e.g., [17]. Hence, VL kinetics are strongly affected by uncertainties stemming from data assimilation processes. Existing methods are not well-suited to consider possible mutations of the virus, resulting in a modification of the infectiousness curve associated with each individual illness. Therefore, a purely data-oriented approach should be complemented by a multiscale model-oriented approach to provide robust long-term predictions of epidemiological dynamics [18]. In the last few years, several mathematical approaches have been designed to assess the impact of infectivity dynamics at the population level [19, 20]. These works are based on kinetic-type equations and on simplified dynamics for which observable quantities are analytically computable. The developed methods are capable of linking agent-based dynamics to available data, providing effective transition rates in macroscopic compartmental models, see also [10, 21–26].

In this study, we focus on the definition of a new pipeline capable of extracting information starting from VL kinetics through a model-driven approach that couples the multiple timescales of infection dynamics. In detail, we exploit VL kinetics obtained in a series of SARS-CoV-2 infections that are not affected by exogenous influencing factors in the dynamics, e.g., the implementation of mass vaccination policies. We interface data-oriented VL dynamics with a simple compartmental model involving the age of infection [27, 28] to describe the evolution of infected cases from SARS-CoV-2 in Italy. A new approach to evaluate the dynamics of the viral load has been developed and compared with original data obtained during the initial waves of the SARS-CoV-2 pandemic.

Comprehending the factors influencing the in-host viral dynamics represents a fundamental tool to provide robust public health strategies. This pilot study could be implemented in further investigations involving other respiratory viruses to better clarify the process of viral dynamics as a preparatory action for future pandemics.

2. Data and modeling assumptions

2.1. Samples and SARS-CoV-2 kinetics

Since February 2020, Fondazione IRCCS Policlinico San Matteo (Pavia, Italy) has been identified as the reference regional center for the diagnosis of SARS-CoV-2 infections through the analysis of respiratory samples. In these analyses, RNA was extracted by using the MGISP-960 automated workstation and the MGI Easy Magnetic Beads Virus DNA/RNA Extraction Kit (MGI Technologies, Shenzhen, China). SARS-CoV-2 RNA was detected using the SARS-CoV-2 variants ELITe MGB kit (ELITechGroup, Puteaux, France; cat. no. RTS170ING) targeting ORF8 and RdRp genes. Reactions were carried out on the CFX96 Touch Real-time PCR detection system (BioRad, Mississauga, ON, Canada).

Individuals that resulted positive from SARS-CoV-2 RNA may leave isolation only when two

consecutive negative respiratory samples occurred. Thus, a series of follow-up samples were analyzed to evaluate the structure of SARS-CoV-2 viral loads and used to monitor the evolution of SARS-CoV-2 VL kinetics. For this reason, the VL kinetics of each considered patient are described by at least two samples, one accounting for the time of testing and the other for the measured viral load. In more detail, the experimental data for each patient consist of two n-tuples, $n \in \mathbb{N}$, where n is the number of successive molecular tests taken by the individual during the monitoring period: the first n-tuple contains the times (in days) when respiratory samples were collected and analyzed (where 0 always represents the day of the first test) and the second contains the kinetics Ct detected by each of these tests (where 40 Ct means complete negativity to the virus and thus recovery). So, for example, a patient that took three successive molecular tests, the first being strongly positive (~ 20 Ct), the second barely positive (~ 36 Ct) after 9 days, and finally the third negative after 7 more days, would be characterized by the two triples $(0, 9, 16)$ for the times and $(20, 36, 40)$ for the kinetics. The values of the corresponding viral loads in terms of RNA copies/ml of respiratory samples may then be deduced from the kinetics of cycle threshold (Ct) using the conversion formula for the quantification of the copy number in clinical samples proposed in [29, 30], which is based on the following proportion:

$$1 \frac{\text{copies}}{\text{ml}} = 16.6 \times 10^{-0.2814 \times 1\text{Ct} + 11.232}, \tag{2.1}$$

where 16.6 is the dilution factor needed to report VL as RNA copies/ml of respiratory samples. Laboratory data were anonymized and retrospectively analyzed. The study was conducted in accordance with the Declaration of Helsinki and no clinical information were available except for the age and gender.

2.2. Statistical analysis

Numerical variables are described as mean ± standard deviation and categorical variables as count and percentages. The comparison of the two time periods (November–December 2020 and January–May 2021) considered is performed in terms of the t-test for quantitative variables and Fisher exact test for quantitative variables. Considering the longitudinal structure of the data, they are presented as daily averaged values. The comparison between the two periods and the investigation of the influence of age and gender on the kinetics are performed in terms of a linear mixed-effects model. The temporal evolution of VL kinetics is obtained as a Gamma distribution using the MATLAB built-in `fit` function, with the LAR method to ensure that the extreme values have a lesser influence on the result. The significance threshold is set to $p < 0.05$.

2.3. Data description

The study analyzed a total of 233 VL kinetics, corresponding to approximately 700 respiratory samples, collected from patients across two distinct periods. The first period goes from November to December 2020, before the circulation of any variants of concern (VOCs), and included 71 patients. The second period, going from January to May 2021 during the circulation of VOCs, involved 162 patients. The subjects considered had a mean age of 42.3 ± 14.7 years, with 55% being females (120) and 45% being males (113). The demographic distribution, in terms of age and gender, does not statistically differ between the two periods. Using mixed-effects models we found that age and gender did not significantly influence VL kinetics. Time to negativization has a significant effect, showing a decreasing pattern

of VL kinetics (estimate = −482,193.7, p-value < 0.001). Additionally, a significant difference in VL kinetics was observed between the two periods, with lower values recorded in 2021 compared to 2020 (estimate = −2,122,031.9, p-value = 0.015).

2.4. Viral load kinetics

In this section we propose a reconstruction based on raw data of VL kinetics, to understand the main features of their evolution. Based on existing works in the medical literature about the shape of the SARS-CoV-2 infectiousness [9, 31] we expect the viral load to be characterized by a rapid growth in the few days after the first contact with an infective individual, until the infection reaches its maximum, followed by a slower decay up to the complete negativization of the subject after several days. We recall that a patient is considered negative if Ct ≥ 40 is detected, based on the conversion formula (2.1). In order to compare VL kinetics from different patients, we determine the first day of negativization for each subject and we save all their preceding, patient-specific, VL measurements. Since there is no a priori information on the day when the peak of infectivity is reached, we align the different VL kinetics of all subjects based on the day of negativization, being this value determined within the whole dataset. In particular, the considered alignment does not depend on the observed peak of individual viral loads, which in turn is often the first measurement. Therefore, an alignment by negativization makes it possible to observe an aggregate trend in the dataset. In Figure 1 we present the raw data of VL kinetics for the whole cohort of patients, aligned by the day of first negativization, over the two distinct time periods during which their infectiousness was monitored.
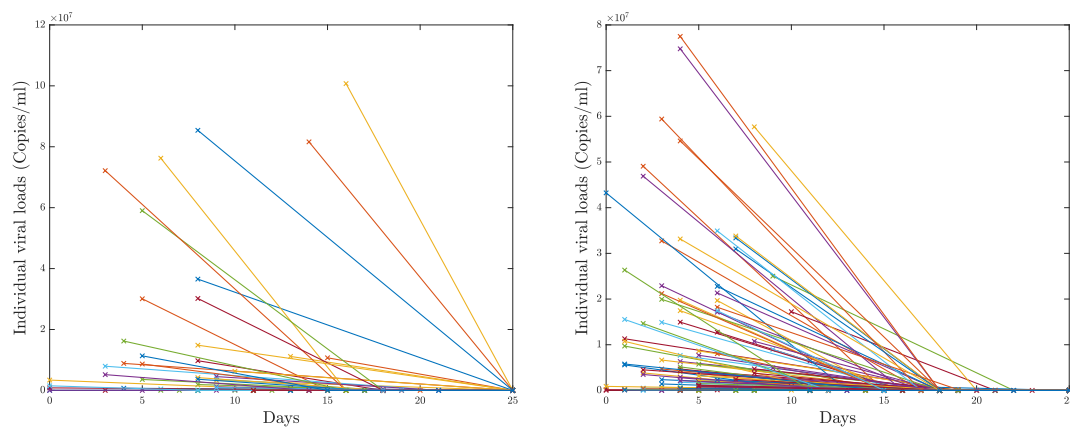

Figure 1. Evolution of individuals' viral load kinetics over time with an ordering based on the first negative test taken by the infective patients, where a patient is considered negative if Ct ≥ 40 is detected, based on the conversion formula (2.1). We use a simple linear interpolation to connect values that are relative to the same subject. We distinguish between the two observation periods of November–December 2020, accounting for 71 patients (left), and January–May 2021, accounting for 162 patients (right).

Based on this alignment process, we consider the daily average of the viral load with respect to the number of patients that got tested on the corresponding day. We then infer the evolution of the viral load as the best fit of these averages, weighted by the daily patients' count. This means that the more patients tested on one specific day, the more importance the corresponding average VL kinetics will have in the fitting. The outcomes of this analysis are shown in Figure 2. The infection persists for at most 25

days and its peak is reached 3–4 days after the first contact with an infected individual (4 days for the viral loads considered from the period November–December 2020 and 3 days for those belonging to the window January–May 2021). The red curves represent the best weighted fits for the daily averaged viral loads and have the shape of a Gamma distribution

$$\hat{\beta}(\tau) = \alpha_0(\tau + \alpha_1)^{\alpha_2} \exp\{-\alpha_3(\tau + \alpha_1)\}, \tag{2.2}$$

where $\alpha_0 = 10^7$ is a scale factor used to capture the magnitude of individuals' viral load and the coefficients $\alpha_1 \in \mathbb{R}$, $\alpha_2, \alpha_3 \geq 0$ are reported in Table 1. In particular, α_1 represents the time of first contact (measured in days), while α_2 and α_3 characterize the shape of the viral load, hence the probability of viral transmission between individuals.
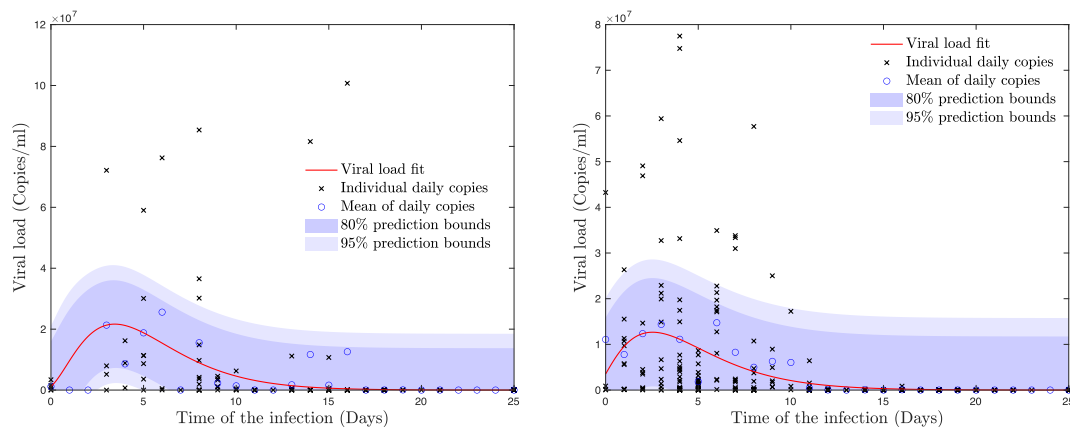
Figure 2. Best fit of the patients' viral loads depending on the time of the infection, for the two monitoring periods of November–December 2020 (left) and January–May 2021 (right). The black crosses denote all the individuals' infection cycles ordered based on their first negativization. The blue dots represent the averages of the daily viral loads. The red curves are the best fit for these averages, weighted by the number of patients counted on each day. The shaded region gives the 80% and 95% prediction bounds for the whole dataset.

Table 1. Optimised coefficients for viral load kinetics, for the two observation periods of November–December 2020 (left) and January–May 2021 (right). The parameter α_1 denotes the time of first contact (measured in days) while the shape parameters α_2 and α_3 characterize the probability of transmitting the infection.

	November–December 2020's sample			January–May 2021's sample	
	Estimated value	95% CI		Estimated value	95% CI
α_1	0.4189 d	[-0.4865, 1.3245]	α_1	0.6262 d	[0.1588, 1.0936]
α_2	2.2015	[1.6124, 2.7906]	α_2	1.5687	[1.0462, 2.0912]
α_3	0.5701	[0.3928, 0.7475]	α_3	0.4964	[0.3348, 0.6580]

2.5. Epidemic model with uncertain quantities

Compartmental models are classically defined to describe mathematically the spread of epidemics in a population [32]. The main idea of such models is to divide the population into different subgroups,

each identified by a specific epidemiologically relevant status. Suitable transition rates are introduced to describe the switches between compartments. Generalizations of the mentioned compartmental models include additional factors characterizing the transmission dynamics such as external influence, age structure, variable contact dynamics, and mobility of agents, see, e.g., [33–36]. From a model perspective, the choice of the compartmentalization is crucial to provide reliable predictions and, therefore, to effectively understand the evolution of the infection. In the following, we will adopt a SIR-type model with the age of infection, where the compartment propagating the infection is the one identified by the population of infected agents. In particular, assuming the infectivity to be modulated by an additional variable identifying the age of infection, we can keep track of the disease's variability inside each individual, as it evolves in parallel with the main timescale of the epidemic. In other words, we assume that the viral load heterogeneity is capable of characterizing the population propagating the infection [11]. Experimental evidence suggests that in the growth phase of VL dynamics no symptoms are typically detected, see [37]. This allows us to introduce a local incidence rate function of the age of infection, modeling the average viral load of the infected individuals and whose shape can be inferred by the information coming from real data on molecular tests. To take into account the structural uncertainties stemming from data assimilation processes of VL kinetics, we extend the model approach to consider the presence of uncertain quantities as a structural feature of the epidemic dynamics. To date, few results are available regarding the development of uncertainty quantification (UQ) methods in epidemic systems, see [33, 38–40, 51]. The development of UQ methods are based on an increased dimensionality of the problem, which is made dependent on a random variable $\mathbf{z} \in \mathbb{R}^{d_z}$, of which the distribution $p(\mathbf{z})$ is known [41–43]. The extrapolation of statistics is classically obtained by looking at quantities of interest, with respect to $\mathbf{z}$, representing the expected solution of the problem.

Since the first day of infectivity of an infectious individual is unknown, we consider a compartmental modeling approach with uncertainty. We consider two temporal scales, one for the time $t \geq 0$ of evolution for the epidemic and another for the time $\tau \geq 0$ of evolution for the infection. At any given time $t \geq 0$, we then divide the population into susceptible individuals $S = S(t, \mathbf{z})$, recovered individuals $R = R(t, \mathbf{z})$ and infected individuals $i(t, \tau, \mathbf{z})$, with variable infection over time $\tau \geq 0$. We then consider the triple $(S(t, \mathbf{z}), i(t, \tau, \mathbf{z}), R(t, \mathbf{z}))$ solution to

$$
\begin{cases}
\dfrac{\mathrm{d}S(t, \mathbf{z})}{\mathrm{d}t} = -\bar{\beta} S(t, \mathbf{z}) \displaystyle\int_{\mathbb{R}_+} \beta(\tau, \mathbf{z}) i(t, \tau, \mathbf{z}) \mathrm{d}\tau, \\[2mm]
\partial_t i(t, \tau, \mathbf{z}) + \partial_\tau i(t, \tau, \mathbf{z}) = -\bar{\gamma} \gamma(\tau, \mathbf{z}) i(t, \tau, \mathbf{z}), \\[2mm]
\dfrac{\mathrm{d}R(t, \mathbf{z})}{\mathrm{d}t} = \bar{\gamma} \displaystyle\int_{\mathbb{R}_+} \gamma(\tau, \mathbf{z}) i(t, \tau, \mathbf{z}) \mathrm{d}\tau,
\end{cases}
\tag{2.3}
$$

where $\beta(\tau, \mathbf{z}) \geq 0$ and $\gamma(\tau, \mathbf{z}) \geq 0$ represent, respectively, the uncertain infection rate and the uncertain recovery rate of the infected individuals, depending on the age of infection $\tau \geq 0$, while the constants $\bar{\beta} > 0$ and $\bar{\gamma} > 0$ enclose, respectively, the baseline contact rate and the baseline recovery rate. In (2.3) we indicated with ∂_t the partial derivative with respect to the epidemic timescale $t \geq 0$ and with ∂_τ the partial derivative computed with respect to the infection timescale $\tau \geq 0$, see [27]. In particular, we stress once more that since the observable number of infected individuals also varies based on the age of infection $\tau \geq 0$, the quantity $i(t, \tau, \mathbf{z}) \geq 0$ depends on both timescales (epidemic and infection course)

and is related to the total current number of infected individuals $I = I(t, \mathbf{z})$ through the relation

$$I(t, \mathbf{z}) = \int_{\mathbb{R}_+} i(t, \tau, \mathbf{z}) \mathrm{d}\tau.$$

Therefore, this model couples the timescale of the epidemic with the one characterizing the evolution of the virus in terms of its average infectiousness. For each $\mathbf{z}$, the system is finally completed with suitable initial and boundary conditions for the variables $S(t, \mathbf{z})$, $i(t, \tau, \mathbf{z})$, and $R(t, \mathbf{z})$. A natural choice is

$$S(0) = S_0, \quad i(0, \tau, \mathbf{z}) = i_0(\tau, \mathbf{z}), \quad R(0) = R_0, \tag{2.4}$$

$$i(t, 0, \mathbf{z}) = \bar{\beta} S(t, \mathbf{z}) \int_{\mathbb{R}_+} \beta(\tau, \mathbf{z}) i(t, \tau, \mathbf{z}) \mathrm{d}\tau, \tag{2.5}$$

where $S_0, i_0(\tau, \mathbf{z}), R_0 \in \mathbb{R}_+$ are such that

$$S_0 + \int_{\mathbb{R}^{d_z}} \int_{\mathbb{R}_+} i_0(\tau, \mathbf{z}) \mathrm{d}\tau \mathrm{d}p(\mathbf{z}) + R_0 = 1,$$

and $p(\mathbf{z})$ is the distribution of the random variable $\mathbf{z}$.

The introduced simplified infection dynamics may be further modified to take into account a more complex compartmentalization of the society, the influence of external interventions, and infection delays, among others, see [4, 7, 44] and the references therein. In the following, we will always assume a constant $\gamma(\tau, \mathbf{z}) \equiv 1$, so that the recovery rate is simply given by $\bar{\gamma}$. We point the interested reader to [10] for a more precise estimation of such parameters.

3. Inference of viral load dynamics

In this section we define a pipeline to effectively estimate VL dynamics $\beta(\tau, \mathbf{z})$ by using available epidemiological data and introducing uncertainty in the day of first contact, expressed by the coefficient α_1 in (2.2). To this end, we consider an uncertain parameter $\alpha_1(\mathbf{z}) \in \mathbb{R}$ to make the local incidence rate random. Moreover, we take into account data from the COVID-19 pandemic in Italy, at both the local and national level, by first considering the evolution of infected individuals in the province of Pavia and by then extending our analysis to the whole territory of Italy, where the central government implemented several non-pharmaceutical interventions (NPIs) during the pandemic. In Table 2 we report the timeline of the main NPIs taken before the mass vaccination campaign started in mid-2021. It is however worth mentioning that local additional measures, like isolation of small portions of a territory, had often been implemented as well.

Before advancing to the estimation of VL dynamics, it is crucial to address some key points. We emphasize the inherent complexity in calibrating epidemiological models, which is particularly demanding from a numerical perspective. Indeed, the process of aligning these models with real-world scenarios requires a careful approach to parameter estimation. Furthermore, available data on VL kinetics tend to hide the initial evolution of their dynamics since agents are typically tested after the manifestation of the symptoms. This bias is primarily attributed to limited testing capacity, resulting in a dataset that often reflects a conservative estimate. Consequently, any attempts to determine the VL

dynamics should be approached with a mindful consideration of these intrinsic limitations, coupled with a commitment to refining the model's accuracy through continuous reassessment and adaptation.

Table 2. Timeline of the main NPIs employed in Italy before the mass vaccination campaign started in mid-2021. The average recovery time from the infection was 14 days.

Implemented NPI	Date	Average recovery time
First lockdown	March 2020	
Curfew measures	October 2020	14 d
Strengthening of measures	March 2021	

We shall focus on three different periods of time. We start by analyzing the first COVID-19 epidemic wave in Italy, which began in February–March 2020. We then proceed with the study of the second wave of October 2020 and conclude with the third major wave of February–March 2021. We stress that the data on VL kinetics are limited to the sole province of Pavia and have been collected during the two periods of November–December 2020 (characterized by the absence of VOCs) and January–May 2021 (involving the circulation of VOCs). Moreover, we point out that the epidemiological data at our disposal for the province of Pavia were only acquired up to January 18, 2021, and we will therefore narrow down our analysis of the third epidemic wave solely to the case of Italy, for which we shall instead take advantage of the open access data collected by Johns Hopkins University (github.com/CSSEGISandData/COVID-19), reporting in detail the daily number of infected, recovered, and deceased individuals in Italy up to August 4, 2021. Let $C_z = \{\mathbf{z}_1, \ldots, \mathbf{z}_M\}$ be a sample of the random variable $\mathbf{z}$ having probability density function $p(\mathbf{z})$. We adopt a multi-level approach to estimate the parameters $(\alpha_i)_{i=1,2,3}$ characterizing $\beta(\tau, \mathbf{z}_j) \geq 0$, which consists of two steps.

Step 1 – Determination of the epidemiological parameters. We initially assess the (deterministic) value of α_1 dictated by the epidemic during the two weeks preceding any of the three main NPIs presented in Table 2 and we estimate at the same time the relevant epidemiological parameter $\bar{\beta} > 0$ which defines the baseline contact rate, assuming no restrictions on social contacts were in place. Notice that we have to discard the variable $\mathbf{z}$ at this stage, because we first need to ensure that the day of first contact α_1 is compatible with the epidemiological data, in order to quantify the shift between the infection and the epidemic timescales. For this, we look for the optimal α_1 and $\bar{\beta}$ minimizing the relative weighted L^2 norm between the current number of infected and recovered individuals (in Pavia or in Italy, depending on the local or national setting under consideration), which we denote as $\hat{I}(t)$ and $\hat{R}(t)$, and the theoretical evolution of these quantities, $I(t) = \int_{\mathbb{R}_+} i(\tau) d\tau$ and $R(t)$, given by the model (2.3) in absence of uncertainty. We notice in particular that, since (2.3) only involves the recovered individuals inside the population, we take $\hat{R}(t)$ to be the sum of the reported numbers of recovered and deceased individuals. We thus solve the minimization problem

$$\min_{\alpha_1, \bar{\beta}} (1 - \eta) \frac{\|I(t) - \hat{I}(t)\|_{L^2([t_0, t_f])}}{\|\hat{I}(t)\|_{L^2([t_0, t_f])}} + \eta \frac{\|I(t) + R(t) - \hat{I}(t) - \hat{R}(t)\|_{L^2([t_0, t_f])}}{\|\hat{I}(t) + \hat{R}(t)\|_{L^2([t_0, t_f])}} \tag{3.1}$$

over the time horizon $[t_0, t_f]$ representing any of the three periods that can be identified with the initial spreading phase of the three epidemic waves under study: February 24–March 9, 2020, October 7–October 22, 2020 and February 24–March 10, 2021. In particular, we choose here the infection rate $\beta(\tau)$ to be the Gamma function (2.2) obtained in the previous section from the viral load measurements,

rescaled by the factor $\alpha_0 = 10^7$. We use the fitting from November–December 2020 as the initial guess for the optimization (3.1) during the first two epidemic waves, while we opt for the January–May 2021's fitting to study the third wave. Notice that at this stage the parameters α_2 and α_3 are fixed from these two fits, since we only need to ensure a proper initial calibration of the time shift α_1 to correctly align our model with the epidemiological data. Finally, we always consider a constant recovery rate $\bar{\gamma} = 1/14$ to lower the complexity of the minimization. This value is consistent with the existing literature on the COVID-19 infection [5, 10, 45], for which the time to viral clearance has been reported to span approximately between 10 and 21 days on average over the different variants. In Table 3 we collect all the parameters obtained from this initial analysis, for each of the three waves.

Table 3. Epidemiological parameters and coefficients of the infectiousness function for the three periods of time under consideration, both at a local (province of Pavia) and a national (Italy) level.

| | Province of Pavia | | | Italy | | |
	First wave	Second wave		First wave	Second wave	Third wave
α_1	1.8110 d	-0.6706 d	α_1	6.6849 d	-0.6032 d	-0.8251 d
α_2	2.2015	2.2015	α_2	2.2015	2.2015	1.5687
α_3	0.5701	0.5701	α_3	0.5701	0.5701	0.4964
$\bar{\beta}$	0.1486 d^{-1}	0.1603 d^{-1}	$\bar{\beta}$	0.3992 d^{-1}	0.1225 d^{-1}	0.1323 d^{-1}
$\bar{\gamma}$	1/14 d^{-1}	1/14 d^{-1}	$\bar{\gamma}$	1/14 d^{-1}	1/14 d^{-1}	1/14 d^{-1}

Step 2 – Inclusion of the uncertainty. We now go further in the analysis by including the presence of uncertainty and by looking at the full model (2.3). The aim is to perform an uncertainty quantification on the shape of $\beta(\tau, \mathbf{z}_j)$ with respect to the randomness $(\mathbf{z}_j)_{j=1}^{M}$ that may appear when estimating $\alpha_1(\mathbf{z}_j)$, the day of first contact. For this, we perturb the optimal value α_1 computed in the previous step by a centered uniform random distribution $\mathcal{U}([-2, 2])$, i.e., $\alpha_1(\mathbf{z}_j) = \alpha_1 + \mathbf{z}_j$ with $\mathbf{z}_j \sim \mathcal{U}([-2, 2])$, and we assess the impact of this addition on the parameters $\alpha_2(\mathbf{z}_j)$ and $\alpha_3(\mathbf{z}_j)$ that define the shape of $\beta(\tau, \mathbf{z}_j)$. The (randomized) optimal shape is thus determined by fitting the model (2.3) to reproduce the data on infected and recovered cases, through a minimization problem similar to (3.1), but with parameters $\alpha_1(\mathbf{z}_j)$ and $\alpha_2(\mathbf{z}_j)$ instead of α_1 and $\bar{\beta}$, while of course the latter and $\bar{\gamma}$ are those obtained from the previous step. This procedure is repeated for the whole sample C_z and we can finally infer the optimal shape of the infectiousness function with UQ to be the expected value $\mathbb{E}_z(\beta(\tau, \mathbf{z}))$ with respect to $\mathbf{z}$, based on the probability density $p(\mathbf{z}) = \mathcal{U}([-2, 2])(\mathbf{z})$.

The outcomes of our two-step procedure are reported in Figure 3 for the case of Pavia and in Figure 4 for the case of Italy. In the left figures we plot in red the curves providing the mean best shape of the infectiousness function over the three epidemic waves that we have considered. The shaded regions show the 80% and 95% percentiles of the different shapes, based on the sample C_z. We then perform numerical simulations to assess the viability of model (2.3) to predict the evolution of infected individuals during these subsequent phases of the pandemic. We take the infectiousness function to be the computed mean best shapes over the three periods and we choose $\gamma(\tau, \mathbf{z})$ constant and equal to 1 for simplicity. In particular, the duration of the infection is solely determined by $\bar{\gamma}$ and the contact rate is given by the optimal $\bar{\beta}$ determined from Step 1 and reported in Table 3. We show in Figures 3 and 4, on the right, the comparisons between the reported number of infected individuals in Pavia and in Italy

over the three epidemic waves and the corresponding approximations provided by model (2.3), with the aforementioned choices of the parameters. We notice that with the progression of the pandemic, specifically over the third wave of February–March 2021, the model loses some accuracy since it is less appropriate to describe slower growths of the epidemic, which stems from the numerous containment measures introduced by the Italian government in the course of 2021 to reduce contacts among the population. This suggests in particular that, in order to capture such intricate effects and minimize the error made by the model, one should consider a variable contact rate $\bar{\beta}(t)$ that continuously evolves over the epidemic time t.
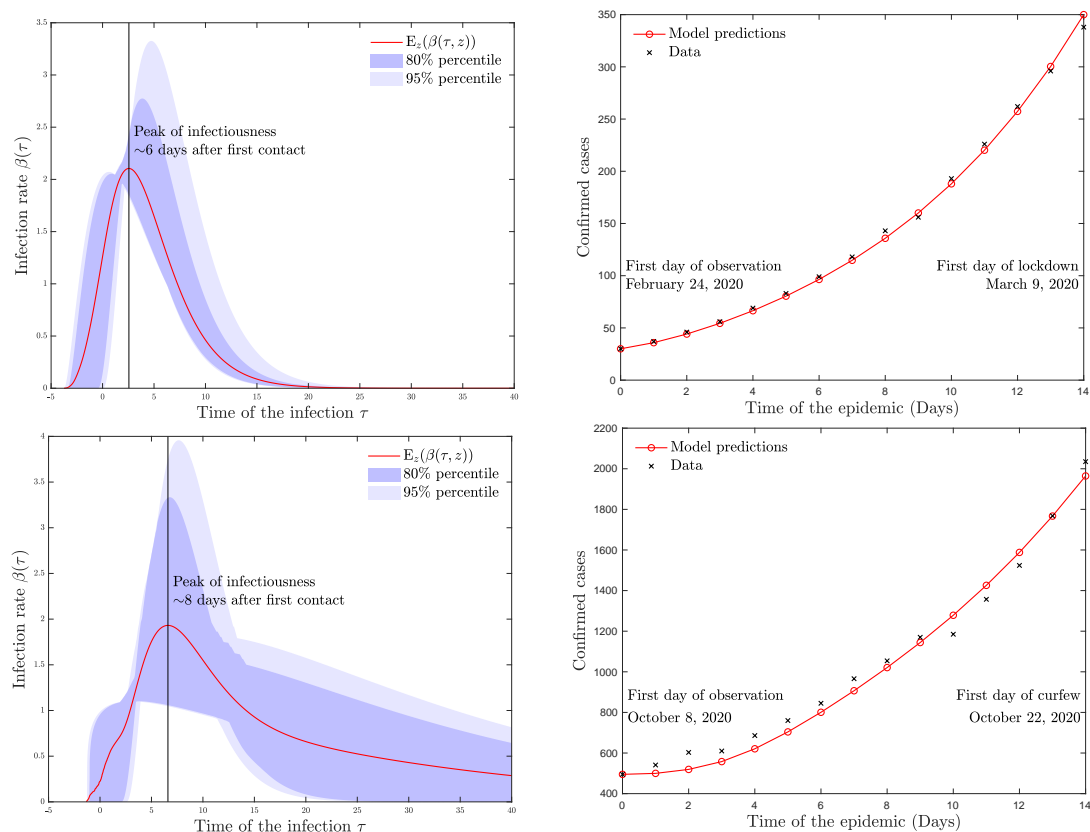
Figure 3. Data- and model-driven inference of the optimal shape of the infection rate in the province of Pavia (Italy) over two different epidemic waves: one occurring at the beginning of the pandemic in 2020, with a focus on the two weeks (February 24–March 9) preceding the first lockdown (top), and the other occurring in the fall of the same year, with a focus on the two weeks (October 7–22) preceding the imposition of a national curfew (bottom). The figures on the left provide the optimal shapes $\mathbb{E}_z(\beta(\tau, \mathbf{z}))$ of the infectiousness function (red curves) determined for these two periods, starting from the data collected by Fondazione IRCCS Policlinico San Matteo to infer an initial guess of the infection rate through the function (2.2) rescaled by the factor $\alpha_0 = 10^7$. The shaded areas show how the infectiousness function is altered when a uniform noise $\mathcal{U}([-2, 2])$ modifies the day of first contact. The figures on the right present the evolution of infected individuals from SARS-CoV-2 in the province of Pavia over the considered windows of time. We compare the computed (in red) and reported (in black) number of cases.
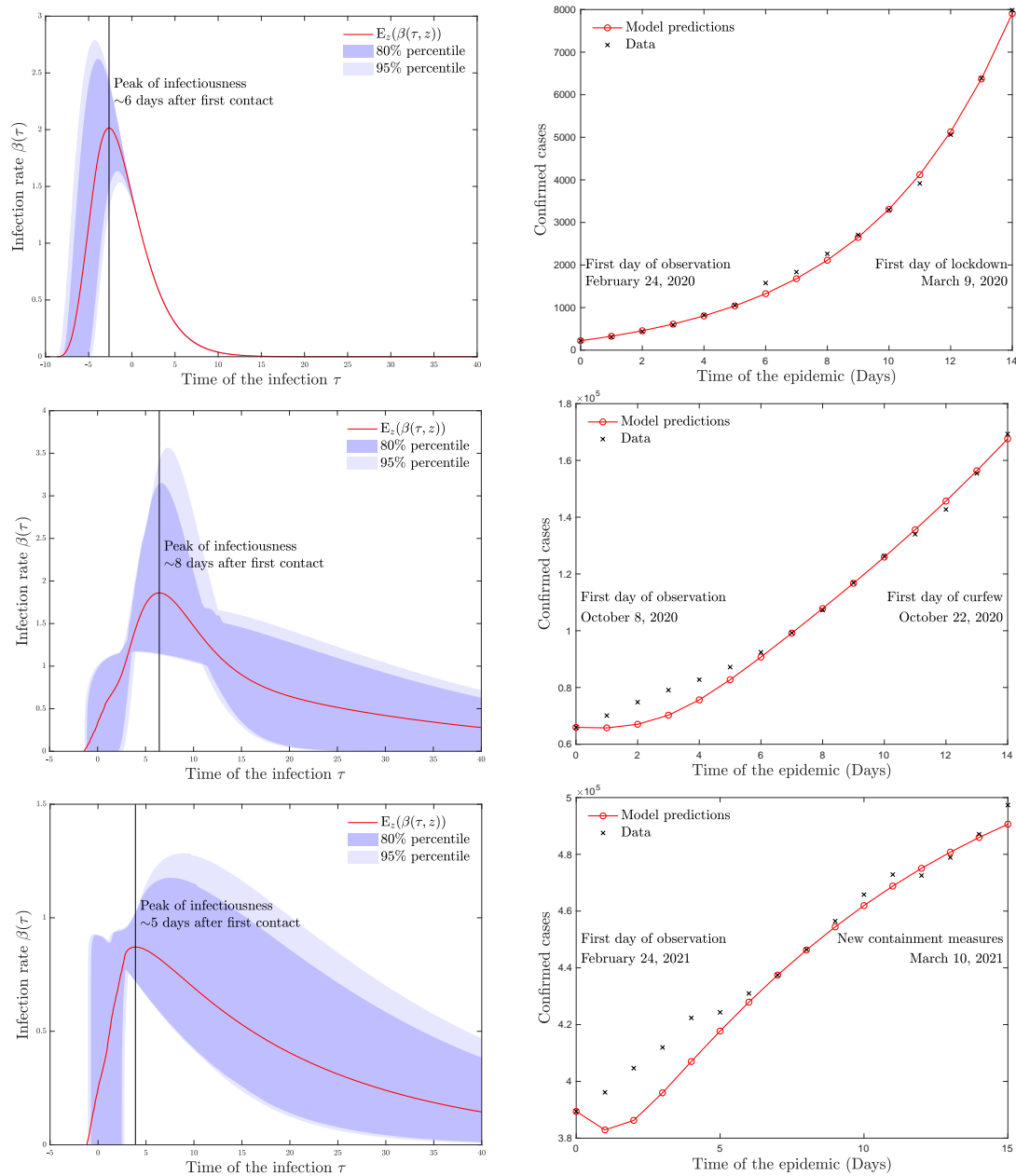
Figure 4. Investigations on the SARS-CoV-2 pandemic in Italy during three distinct epidemic waves, starting from the data collected by Fondazione IRCCS Policlinico San Matteo to infer an initial guess of the infection rate through the function (2.2). Top row: Analysis of the epidemic wave that started in February–March 2020 showing the optimal shape $\mathbb{E}_z(\beta(\tau, \mathbf{z}))$ of the infectiousness function (left) and model approximation (right) for the cases of infection registered during the period February 24–March 9, 2020. Middle row: Analysis of the new epidemic outburst that took place in October 2020 showing the optimal shape of the infection rate (left) and model approximation (right) for the cases of infection registered during the period October 7–22, 2020. Bottom row: Analysis of the epidemic wave that occurred from February to March 2021.

Reconstructing the experimental data over the whole timeframe. We now perform a calibration of our model to cover the whole timespan of observation, connecting the epidemic waves between them and providing a full reconstruction of the experimental data on the evolution of infected individuals. We initially observe that, as the epidemiological parameter $\bar{\beta}$ models the contact rate between individuals, its value is highly influenced by the restrictions imposed by a government. This fact needs to be taken into account when extending our model's predictions from one epidemic wave to the other, since various containment measures (we mention in particular the lockdown imposed between March and May 2020) have been put in place by the Italian government over time to slow down the epidemic spreading. Since we do not need to evaluate the shape of the infection rate in these connecting timeframes (the previous two-step analysis is relevant only when the virus is able to spread free among the population, allowing for a proper calibration of the epidemiological parameters), in order to connect the observations we can rely on a simpler model that allows us to speed up the computational time. Specifically, we make use of system (2.3) with the choices $\beta(\tau, \mathbf{z}), \gamma(\tau, \mathbf{z}) \equiv 1$. From the above considerations, we account for the variable contact restrictions by assuming that the contact rate varies over the epidemic time as $\bar{\beta} = \bar{\beta}(t)$. Then, we determine the value $\bar{\beta}(t)$ by solving an optimization problem for a sequence of time steps t_i (days) over a moving time window of three days (namely, averaging the fitting over three days), varying in any timespan $t \in [t_0 + 1, t_f]$ where t_0 and t_f correspond, respectively, to the end of an epidemic wave and the beginning of the next one. This strategy is needed in order to correctly assess the evolution of the contact rate depending on the containment measures that were in place. More precisely, we solve the following constrained least-square problem:

$$\min_{\bar{\beta}(t_i)} (1 - \eta) \frac{\|I(t) - \hat{I}(t)\|_{L^2([t_i-1, t_i+2])}}{\|\hat{I}(t)\|_{L^2([t_i-1, t_i+2])}} + \eta \frac{\|I(t) + R(t) - \hat{I}(t) - \hat{R}(t)\|_{L^2([t_i-1, t_i+2])}}{\|\hat{I}(t) + \hat{R}(t)\|_{L^2([t_i-1, t_i+2])}}, \tag{3.2}$$

where the recovery rate $\bar{\gamma}$ is still kept fixed to the value $1/14$. We complete our study by using the model to reconstruct the experimental data on the epidemic up to the last observation times, being January 18, 2021, for the case of the province of Pavia and August 4, 2021, for the case of Italy. By means of the above procedure we connect the observations by calibrating $\bar{\beta}(t)$ with the minimization strategy (3.2). Using all these optimized parameters we can finally run our model over the whole timeframe of observation. We depict the resulting approximations for the evolution of infected individuals in the province of Pavia and in Italy in Figures 5 and 6, respectively, on the right. The evolution of the confirmed cases is plotted in black while the model approximations are shown in red.

At last, in Figures 5 and 6 on the left, we provide a direct comparison between the optimal infection rates from different epidemic waves, obtained via the previous two-step analysis, together with their 80% prediction bounds. These are aligned by their respective peak of infectiousness (which correspond to day 0 on the graph). One can observe that the infection's peak progressively decreases, while the tails of the infection rate tend to grow. This suggests that the virus becomes weaker but more effective over a longer period of time and individuals stay infective for more days on average. Moreover, the decrease in the curve's growth in the beginning of the infection could suggest a slower progression of the symptoms and explain the greater spreading of the epidemic with each new wave.
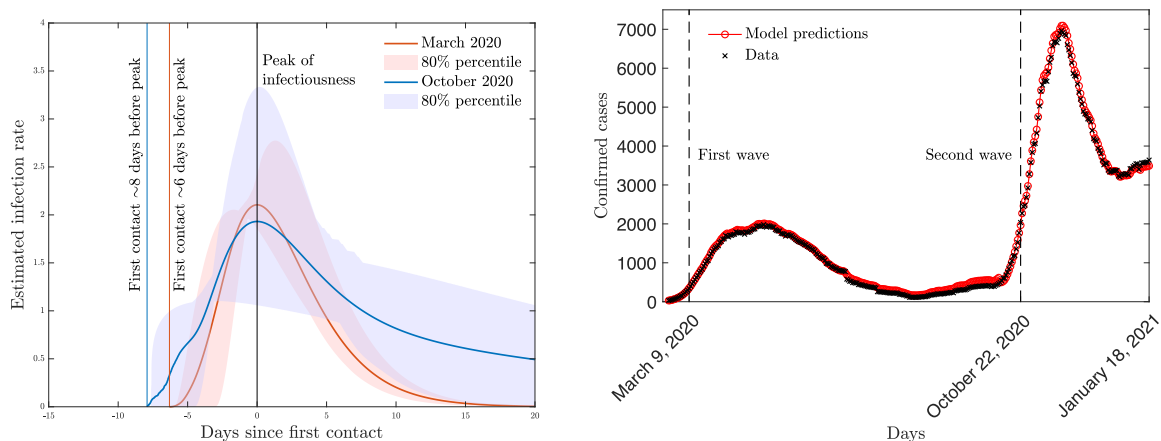
Figure 5. Left: Comparison between infection rates from the first (in orange, March 2020) and the second (in blue, October 2020) epidemic waves in the province of Pavia. In order for the model (2.3) to best fit the data on confirmed cases, the peak of the infection should be reached about 6 to 8 days after the first contact with an infected subject. Right: Evolution of infected individuals from SARS-CoV-2 in Pavia from February 24, 2020, to January 18, 2021. Comparison between computed (in red) and reported (in black) number of cases.
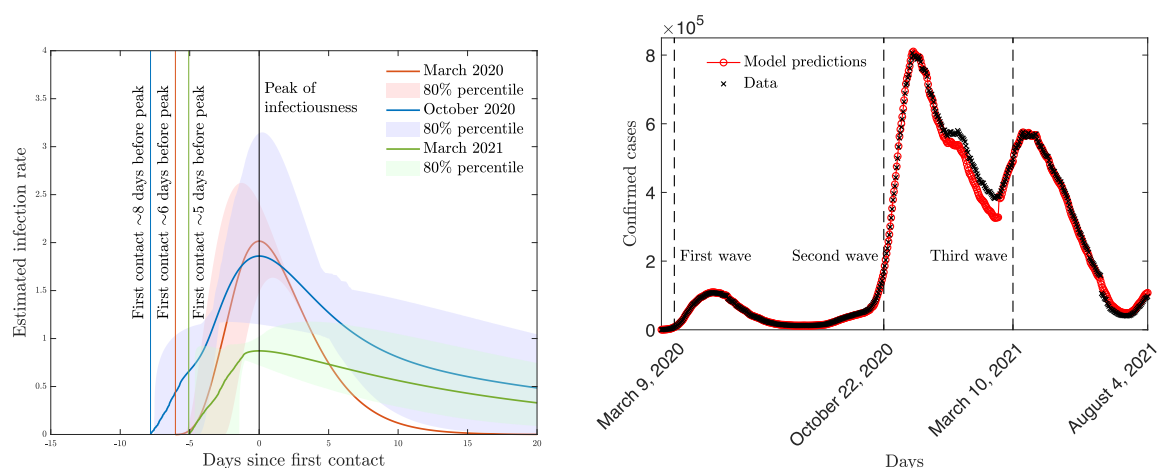
Figure 6. Left: Comparison between infection rates from three distinct epidemic waves in Italy which happened, respectively, in March 2020 (orange), in October 2020 (blue), and in March 2021 (green). In order for the model (2.3) to best fit the data on confirmed cases, the peak of the infection should be reached about 5 to 8 days after the first contact with an infected subject. Right: Evolution of infected from SARS-CoV-2 in Italy from February 24, 2020, to August 4, 2021. Comparison between computed (in red) and reported (in black) number of cases.

4. Conclusions

Understanding the structure and evolution of VL dynamics is crucial in the context of public health when setting up quarantine and isolation measures of positive subjects. In particular, one of the critical issues that has emerged since the onset of the recent SARS-CoV-2 pandemic has been the need of

criteria to assess the infectivity of patients and decide when to release them from prolonged quarantine, safely admitting them back to work duties and social activities. It was demonstrated that high Ct values (low VL), greater than 30–35 Ct, were associated with a reduced infectivity of the patients [46–48]. Nevertheless, existing studies indicate that the peak viral load is closely linked with the onset of symptoms, making the available datasets partial and of difficult interpretation. In this work, we have established a pipeline to assess the evolution of in-host VL from respiratory samples of patients. To this end, we combined the use of real data and model-based methods. Furthermore, the proposed approach is capable all the whole VL dynamics. In particular, our model helps to provide a data-oriented understanding of the non-infectious status of people that were under quarantine with high Ct.

Samples included in this study were collected over two periods of time (November–December 2020 and January–May 2021) that include the initial circulation of VOCs and before the extensive introduction of vaccination. We have first determined the shape of the average VL dynamics hinted at by these data (recovering Gamma-type distributions) and we have successively introduced a suitable compartmental model of SIR-type, using a local incidence rate function that varies with the age of infection, to account for the evolution of the epidemic alongside that of the virus. The system of equations belongs to a large class of epidemiological models detailing the transmission dynamics through the inclusion of additional factors such as external influence, age structure, variable contact dynamics, and mobility of agents [33–36, 49]. In this work, we exploited the information coming from real data on molecular tests and on the reported number of infected individuals to deduce the optimal shape of the infectivity function, based on a combined data-driven and model-driven approach. To take into account the extreme variability of VL dynamics, depending on disease severity and patients' features as well [10, 50], we also made use of UQ methods [41–43] adapted for the study of epidemic systems [33, 38–40], to take into account the uncertainties inherent to the data-assimilation processes of VL kinetics. The model allowed us to infer a plausible evolution of the SARS-CoV-2 infectiousness over three distinct periods of the pandemic (February–March 2020, October 2020, and February–March 2021), in agreement with existing works from the medical literature [9, 31]. In particular, we observed that the peak of the infection function decreases over time while its tails tend to grow, suggesting that the virus becomes progressively weaker but more effective, since individuals remain infective for more days on average. This change in shape strongly manifests during the third observation period (February–March 2021) when the Alpha variant was actively circulating in Italy, providing an explanation for the greater spreading of the contagion at the time.

In conclusion, our work aimed to implement strategies of preparedness for a new pandemic, since the model and methods developed here could be easily adapted to investigate real data coming from epidemics that are associated with other respiratory viruses (e.g., influenza and respiratory syncytial virus), in order to assess the evolution of the associated viral infectiousness that may lead to more informed decisions on preventive measures to reduce the spreading of the disease.

Use of AI tools declaration

The authors declare they have not used artificial intelligence (AI) tools in the creation of this article.

Acknowledgments

Authors contributions

Andrea Bondesan: Formal analysis, Investigation, Methodology, Software, Visualization, Writing - original draft. Antonio Piralla: Data curation, Methodology, Validation, Writing - original draft. Elena Ballante: Formal analysis, Methodology. Antonio Maria Guglielmo Pitrolo: Data curation. Silvia Figini: Supervision. Fausto Baldanti: Conceptualization, Resources, Supervision. Mattia Zanella: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Writing - original draft.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. L. Bolzoni, E. Bonacini, C. Soresina, M. Groppi, Time-optimal control strategies in SIR epidemic models, *Math. Biosci.*, **292** (2017), 86–96. https://doi.org/10.1016/j.mbs.2017.07.011

2. F. Brauer, C. Castillo-Chavez, Z. Feng, *Mathematical Models in Epidemiology*, in *Texts in Applied Mathematics*, Springer, New York, 2019.

3. T. Britton, F. Ball, P. Trapman, A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2, *Science*, **369** (2020), 846–849. https://doi.org/10.1126/science.abc6810

4. B. Buonomo, R. Della Marca, Effects of information-induced behavioural changes during the COVID-19 lockdowns: the case of Italy, *R. Soc. Open Sci.*, **7** (2020), 201635. https://doi.org/10.1098/rsos.201635

5. J. Chen, H. Lu, G. Melino, S. Boccia, M. Piacentini, W. Ricciardi, et al., COVID-19 infection: the China and Italy perspectives, *Cell Death Dis.*, **11** (2020), 438. https://doi.org/10.1038/s41419-020-2603-0

6. M. Gatto, E. Bertuzzo, L. Mari, S. Miccoli, L. Carraro, R. Casagrandi, et al., Spread and dynamics of the COVID-19 epidemic in Italy: effect of emergency containment measures, *PNAS*, **117** (2020), 10484–10491. https://doi.org/10.1073/pnas.2004978117

7. G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo, et al., Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy, *Nat. Med.*, **26** (2020), 855–860. https://doi.org/10.1038/s41591-020-0883-7

8. M. J. Keeling, L. Dyson, M. J. Tildesley, E. M. Hill, S. Moore, Comparison of the 2021 COVID-19 roadmap projections against public health data in England, *Nat. Commun.*, **13** (2022). https://doi.org/10.1038/s41467-022-31991-0

9. O. Puhach, B. Meyer, I. Eckerle, SARS-CoV-2 viral load and shedding kinetics, *Nat. Rev. Microbiol.*, **21** (2023), 147–161. https://doi.org/10.1038/s41579-022-00822-w

10. M. Zanella, C. Bardelli, G. Dimarco, S. Deandrea, P. Perotti, M. Azzi, et al., A data-driven epidemic model with social structure for understanding the COVID-19 infection on a heavily affected Italian Province, *Math. Models Methods Appl. Sci.*, **31** (2023), 2533–2570. https://doi.org/10.1142/S021820252150055X

11. T. C. Jones, G. Biele, B. Mühlemann, T. Veith, J. Schneider, J. Beheim-Schwarzbach, et al., Estimating infectiousness throughout SARS-CoV-2 infection course, *Science*, **373** (2021), eabi5273. https://doi.org/10.1126/science.abi5273

12. H. Lu, F. Giannino, D. M. Tartakovsky, Parsimonious models of in-host viral dynamics and immune response, *Appl. Math. Lett.*, **145** (2023), 108781. https://doi.org/10.1016/j.aml.2023.108781

13. G. Bertaglia, L. Pareschi, G. Toscani, Modelling contagious viral dynamics: a kinetic approach based on mutual utility, *Math. Biosci. Eng.*, **21** (2024), 4241–4268. https://doi.org/10.3934/mbe.2024187

14. J. K. Frediani, R. Parsons, K. B. McLendon, A. L. Westbrook, W. Lam, G. Martin, et al., The new normal: delayed peak SARS-CoV-2 viral loads relative to symptom onset and implications for COVID-19 testing programs, *Clin. Infect. Dis.*, **78** (2024), 301–307. https://doi.org/10.1093/cid/ciad582

15. O. Puhach, K. Adea, N. Hulo, P. Sattonnet, C. Genecand, A. Iten, et al., Infectious viral load in unvaccinated and vaccinated individuals infected with ancestral, Delta or Omicron SARS-CoV-2, *Nat. Med.*, **28** (2022), 1491–1500. https://doi.org/10.1038/s41591-022-01816-0

16. R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, et al., Estimates of the severity of coronavirus disease 2019: a model-based analysis, *Lancet Infect. Dis.*, **20** (2020), 669–677. https://doi.org/10.1016/S1473-3099(20)30243-7

17. X. He, E. H. Y. Lau, P. Wu, X. Deng, J. Wang, X. Hao, et al., Temporal dynamics in viral shedding and transmissibility of COVID-19, *Nat. Med.*, **26** (2020), 672–675. https://doi.org/10.1038/s41591-020-0869-5

18. E. Kharazmi, M. Cai, X. Zheng, Z. Zhang, G. Lin, G. E. Karniadakis, Identifiability and predictability of integer- and fractional-order epidemiological models using physics-informed neural networks, *Nature Comput. Sci.*, **1** (2021), 744–753. https://doi.org/10.1038/s43588-021-00158-0

19. R. Della Marca, N. Loy, A. Tosin, An SIR model with viral load-dependent transmission, *J. Math. Biol.*, **86** (2023), 61. https://doi.org/10.1007/s00285-023-01901-z

20. N. Loy, A. Tosin, A viral load-based model for epidemic spread on spatial networks, *Math. Biosci. Eng.*, **18** (2021), 5635–5663. https://doi.org/10.3934/mbe.2021285

21. G. Bertaglia, A. Bondesan, D. Burini, R. Eftimie, L. Pareschi, G. Toscani, New trends on the systems approach to modeling SARS-CoV-2 pandemics in a globally connected planet, *Math. Models Methods Appl. Sci.*, **34** (2024), 1995–2054. https://doi.org/10.1142/S0218202524500301

22. G. Bertaglia, W. Boscheri, G. Dimarco, L. Pareschi, Spatial spread of COVID-19 outbreak in Italy using multiscale kinetic transport equations with uncertainty, *Math. Biosci. Eng.*, **18** (2021), 7028–7059. https://doi.org/10.3934/mbe.2021350

23. A. Bondesan, G. Toscani, M. Zanella, Kinetic compartmental models driven by opinion dynamics: Vaccine hesitancy and social influence, *Math. Models Methods Appl. Sci.*, **34** (2024), 1043–1076. https://doi.org/10.1142/S0218202524400062

24. G. Dimarco, B. Perthame, G. Toscani, M. Zanella, Kinetic models for epidemic dynamics with social heterogeneity, *J. Math. Biol.*, **83** (2021), 4. https://doi.org/10.1007/s00285-021-01630-1

25. J. Franceschi, A. Medaglia, M. Zanella, On the optimal control of kinetic epidemic models with uncertain social features, *Optim. Contr. Appl. Meth.*, **45** (2024), 494–522. https://doi.org/10.1002/oca.3029

26. M. Zanella, Kinetic models for epidemic dynamics in the presence of opinion polarization, *Bull. Math. Biol.*, **85** (2023), 36. https://doi.org/10.1007/s11538-023-01147-2

27. J. Demongeot, Q. Griette, Y. Maday, P. Magal, A Kermack–McKendrick model with age of infection starting from a single or multiple cohorts of infected patients, *Proc. R. Soc. A*, **479** (2023), 20220381. https://doi.org/10.1098/rspa.2022.0381

28. J. Y. Yang, F. Q. Zhang, X. Y. Wang, SIV epidemic models with age of infection, *Int. J. Biomath.*, **2** (2009), 61–67. https://doi.org/10.1142/S1793524509000492

29. A. Piralla, C. Daleno, E. Pariani, P. Conaldi, S. Esposito, A. Zanetti, et al., Virtual quantification of influenza A virus load by real-time RT-PCR, *J. Clin. Virol.*, **56** (2013), 65–68. https://doi.org/10.1016/j.jcv.2012.09.011

30. A. Piralla, A. Girello, M. Premoli, F. Baldanti, A new real-time reverse transcription-PCR assay for detection of human enterovirus 68 in respiratory samples, *J. Clin. Microbiol.*, **53** (2015), 1725–1726. https://doi.org/10.1128/JCM.03691-14

31. M. Cevik, K. Kuppalli, J. Kindrachuk, M. Peiris, Virology, transmission, and pathogenesis of SARS-CoV-2, *BMJ*, **371** (2020), 1–6. https://doi.org/10.1136/bmj.m3862

32. O. Diekmann, J. A. P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*, Wiley, 2000.

33. G. Albi, L. Pareschi, M. Zanella, Control with uncertain data of socially structured compartmental epidemic models, *J. Math. Biol.*, **82** (2021), 63. https://doi.org/10.1007/s00285-021-01617-y

34. B. Bartélemy, A. Barrat, R. Pastor-Satorras, A. Vespignani, Dynamical patterns of epidemic outbreaks in complex heterogeneous networks, *J. Theor. Biol.*, **235** (2005), 275–288. https://doi.org/10.1016/j.jtbi.2005.01.011

35. V. Capasso, G. Serio, A generalization of the Kermack–McKendrick deterministic epidemic model, *Math. Biosci.*, **42** (1978), 43–61. https://doi.org/10.1016/0025-5564(78)90006-8

36. J. Dolbeault, G. Turinici, Heterogeneous social interactions and the COVID-19 lockdown outcome in a multi-group SEIR model, *Math. Model. Net. Pheno.*, **15** (2020), 1–18. https://doi.org/10.1051/mmnp/2020025

37. M. Mozgovoj, M. D. Graham, C. Ferrufino, S. Blanc, A. F. Souto, M. Pilloff, et al., Viral load in symptomatic and asymptomatic patients infected with SARS-CoV-2. What have we learned?, *J. Clin. Virol. Plus*, **3** (2023), 100166.

38. M. A. Capistrán, J. A. Christen, J. X. Velasco-Hernández, Towards uncertainty quantification and inference in the stochastic SIR epidemic model, *Math. Biosci.*, **240** (2012), 250–259. https://doi.org/10.1016/j.mbs.2012.08.005

39. G. Chowell, Fitting dynamic models to epidemic outbreaks with quantified uncertainty: a primer for parameter uncertainty, identifiability, and forecasts, *Infect. Dis. Model.*, **2** (2017), 379–398.

40. M. G. Roberts, Epidemic models with uncertainty in the reproduction number, *J. Math. Biol.*, **66** (2013), 1463–1474. https://doi.org/10.1007/s00285-012-0540-y

41. S. Jin, L. Pareschi, *Uncertainty Quantification for Hyperbolic and Kinetic Equations*, SEMA-SIMAI Springer Series, 2017.

42. L. Pareschi, An introduction to uncertainty quantification for kinetic equations and related problems, in *Trails in Kinetic Theory: Foundational Aspects and Numerical Methods* (eds. G. Albi, S. Merino-Aceituno, A. Nota, M. Zanella), SEMA-SIMAI Springer Series, **25** (2021). https://doi.org/10.1007/978-3-030-67104-4

43. D. Xiu, *Numerical Methods for Stochastic Computations: a Spectral Methods Approach*, Princeton University Press, Princeton, 2010.

44. N. Guglielmi, E. Iacomini, A. Viguerie, Identification of time delays in COVID-19 data, *Epidemiol. Meth.*, **12** (2023), 20220117. https://doi.org/10.1515/em-2022-0117

45. E. Lavezzo, E. Franchin, C. Ciavarella, G. Cuomo-Dannenburg, L. Barzon, C. Del Vecchio, et al., Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo', *Nature*, **584** (2020), 425–429. https://doi.org/10.1038/s41586-020-2488-1

46. B. La Scola, M. Le Bideau, J. Andreani, V. T. Hoang, C. Grimaldier, P. Colson, et al., Viral RNA load as determined by cell culture as a management tool for discharge of SARS-CoV-2 patients from infectious disease wards, *Eur. J. Clin. Microbiol. Infect. Dis.*, **39** (2020), 1059–1061. https://doi.org/10.1007/s10096-020-03913-9

47. A. Singanayagam, M. Patel, A. Charlett, J. L. Bernal, V. Saliba, J. Ellis, et al., Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020, *Eurosurveillance*, **25** (2020), 2001483.

48. A. Piralla, M. Ricchi, M. G. Cusi, P. Prati, N. Vicari, G. Scarsi, et al., Residual SARS-CoV-2 RNA in nasal swabs of convalescent COVID-19 patients: Is prolonged quarantine always justified?, *IJID*, **102** (2021), 299–302. https://doi.org/10.1016/j.ijid.2020.10.072

49. M. J. Keeling, S. Moore, B. S. Penman, E. M. Hill, The impacts of SARS-CoV-2 vaccine dose separation and targeting on the COVID-19 epidemic in England, *Nat. Commun.*, **14** (2023). https://doi.org/10.1038/s41467-023-35943-0

50. Y. Liu, L. M. Yan, L. Wan, T. X. Xiang, A. Le, J. M. Liu, et al., Viral dynamics in mild and severe cases of COVID-19, *Lancet Infect. Dis.*, **20** (2020), 656–657. https://doi.org/10.1016/S1473-3099(20)30232-2

51. A. Medaglia, G. Colelli, L. Farina, A. Bacila, P. Bini, E. Marchioni, et al., Uncertainty quantification and control of kinetic models of tumour growth under clinical uncertainties, *Int. J. NonLin. Mech.*, **141** (2022), 103933. https://doi.org/10.1016/j.ijnonlinmec.2022.103933