

The impact of the distributions of the latent and infectious periods on epidemic dynamics

Supplementary Material

Paulo C. Ventura¹, Allisandra G. Kummer¹, Yong-Dam Jeong^{2,3}, Shreeya Mhade¹, Shingo Iwami³, Stefano Merler⁴, Hongjie Yu^{5,6}, Keisuke Ejima⁷, Maria Litvinova⁸, Marco Ajelli^{1,*}

¹ Laboratory for Computational Epidemiology and Public Health, Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN, USA

² Department of Applied Mathematics, Pukyong National University, Busan, Republic of Korea

³ interdisciplinary Biology Laboratory (iBLab), Division of Biological Science, Graduate School of Science, Nagoya University, Nagoya, Japan

⁴ Center for Health Emergencies, Bruno Kessler Foundation, Trento, Italy

⁵ Shanghai Institute of Infectious Disease and Biosecurity, Department of Epidemiology, School of Public Health, Fudan University, Shanghai, China

⁶ Department of Epidemiology, School of Public Health, Fudan University, Key Laboratory of Public Health Safety, Ministry of Education, Shanghai, China

⁷ Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

⁸ Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN, USA

*Corresponding Author. marco.ajelli@gmail.com

Methods

Viral load driven parameterization (VLP)

In the VIBES model, a pathogen spreads between individuals through contacts in the described synthetic population network. At each time step during a simulation, each individual is assigned to one of the following epidemiologic statuses: susceptible, latent, infectious pre-symptomatic, infectious symptomatic, infectious asymptomatic and removed. Susceptible individuals can acquire the infection from infectious individuals that share a cluster in the population. Once infected, we determine whether the individual will develop symptoms as a random event with an age-dependent probability as provided in Poletti et al¹ for SARS-CoV-2 infections. For individuals assigned to develop symptoms, we sample an element from the ensemble of viral load trajectories and incubation period of symptomatic individuals, generated by a viral dynamics model. For individuals assigned to not develop symptoms, we sample from the viral load trajectories of asymptomatic individuals. Details on the viral dynamics model can be found in our previous publication².

The assigned viral load trajectory determines the progression of an individual through epidemiologic statuses during an infection. The individual is assigned to the latent status at the time step immediately following their infection. Over the next simulation steps, when the viral load exceeds the infectiousness threshold (see the table of parameters, Table S1), the individual is assigned to the infectious status and becomes able to transmit forward the pathogen. Once the viral load goes below the same infectiousness threshold, the individual is assigned to the removed status and remains so until the end of the simulation. If the individual was assigned to develop symptoms, they start the infectious period as infectious pre-symptomatic, then move to the infectious symptomatic status once the incubation period is completed.

In the VLP, the viral load of an infectious individual determines the infectiousness profile over time. The probability $p_{i \rightarrow j}(t)$ that an infectious individual i infects a susceptible individual j that is in contact with i in setting $l \in \{\text{household, workplace, school, community}\}$ at time t is defined as:

$$p_{i \rightarrow j}(t) = \alpha \cdot \frac{\beta_l}{N_{i,l}} \cdot g_i \cdot \psi_i \cdot u_i(t - \tau_i) \cdot \Delta t, \quad (1)$$

Where:

- α is the overall transmission rate;
- β_l is the layer-specific relative transmission risk;
- $N_{i,l}$ is the number of individuals in the given cluster of layer l ;
- $u_i(t - \tau_i)$ is the infectiousness profile function of individual i at step $t - \tau_i$, where τ_i represents the time of infection of individual i ;
- g_i is a gamma-distributed random variable determining the overall infectiousness of individual i , with shape 0.235 and scale 4.26 to obtain the distribution of the number of secondary infections reported in Sun et al¹³.
- ψ_i is the infectiousness of symptomatic individuals relative to asymptomatic individuals; namely, $\psi_i = 1$ if i is asymptomatic and $\psi_i = \psi$ if i is symptomatic.
- $\Delta t = 0.25$ days is the duration of a single time step in the simulation.

The geometric parameterization (GP)

The geometric parameterization was implemented on the same modeling framework of VIBES, with adaptations that allows it to run without a controlled infectiousness profile. The epidemiological statuses, synthetic population and symptom-development probabilities are the same as those used in the VLP. However, instead of using a viral dynamics model to determine viral load curves, we used a simpler algorithm that assumes memory-free (Markovian) transitions.

To generate each infectiousness profile with the geometric parameterization, we sampled the duration of the latent period from a geometric distribution with success probability η , resulting in an average latent period of $\Delta t + 1/\eta$, where a simulation step length Δt is added due to the model constraint of preventing the individual to become infectious on the same time step of infection. The infectious period was sampled from a geometric distribution with success probability γ , resulting in an average duration of $\Delta t + 1/\gamma$. We used the same latent and infectious period parameters for the profiles of symptomatic and asymptomatic individuals. Specifically for symptomatic infectiousness profiles, the onset of symptoms was determined to occur T_{OS} days after the start of the infectious period. Once the durations of the latent and infectious periods are determined, the i -th infectiousness profile $u_i(t)$ is defined as:

$$u_i(T) = \begin{cases} 1, & T \in [T_{i,1}, T_{i,2}], \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

Where T is the time since infection, $T_{i,1}$ is the time between infection of individual i and the start of the infectious period, while $T_{i,2}$ is the time between infection and the end of the infectious period.

The disease progression and assignment of a symptom development status follow the same rules as used for the VLP. The infection probability $p_{i \rightarrow j}(t)$ that an infectious individual i infects a susceptible individual j at time t is the same as that in equation 1, except that $g_i = 1$ for all individuals so that individual infectiousness is determined only by the length of the infectious period. All parameter values are reported in the Table of Parameters.

Model parameterization and initialization

In both GP and VLP, the layer-specific transmission risks β_l were adjusted along with the overall transmission rate α such that:

- a) The proportion of infections in each of the four layer match the target values (see Table of Parameters) for a reproduction number $R = 1.3$ (calculation explained in section “Calculation of metrics – Reproduction number” of this supplementary text).
- b) At $\alpha = 1$, the reproduction number is $R = 2.5$.

Each simulation was initialized with 30 randomly selected individuals in the latent status, with all other individuals initialized as susceptibles.

Grid of transmission rate values

To study key epidemiological metrics as a function of the reproduction number, we ran simulations on a grid of values of the overall transmission rate α , with all other parameters unchanged. For both GP and VLP, the grid of α values was [0.35, 0.50, 0.67, 0.85, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0]. The layer-specific transmission risk coefficients β_l were kept constant and set such that $R = 2.5$ when $\alpha = 1$.

Calculation of metrics

Reproduction number

We calculated the reproduction number during the initial growth phase of the epidemic based on start and stop thresholds, using the same algorithm for both GP and VLP. For each simulation, we determined the start of the calculation window as the first time $t = t_{start}$ in which the cumulative number of infections is equal to or greater than the calculation start threshold $c_{start} = 1000$. We determined the end of the calculation window as the first time $t = t_{end}$ in which the cumulative number of infections is equal to or greater than $c_{end} = 7000$, the calculation end threshold. If in a particular simulation the cumulative number of infections did not reach c_{end} , we considered that an outbreak did not occur and did not calculate the reproduction number for that simulation. Thus, our criterion for the occurrence of an outbreak was that cumulative cases were greater than or equals to c_{end} .

The reproduction number of a simulation is the number average number of secondary infections from primary infections created during the calculation window. Specifically, defining $W = \{t \mid t \geq t_{start}, t < t_{end}\}$ as the set of time steps within the calculation window (if an outbreak occurred), the reproduction number was calculated as:

$$R = \frac{1}{M_W} \cdot \sum_{i \mid t_{i,0} \in W} r_i$$

where:

- $M_W = \sum_{i \mid t_{i,0} \in W} 1$ is the number of primary infections that occurred during the calculation window W ;
- r_i is the number of secondary infections caused by an infected individual i during their entire infectious period.
- $t_{i,0}$ is the time of infection of individual i ;

For each set of simulations with the same overall transmission rate α , the associated reproduction number was an average R over all simulations for which an outbreak occurred (i.e., in which the calculation window W was defined).

Realized latent period

The latent period of an individual is the time between their infection and the start of their infectious period. We calculate the *realized latent period* of an epidemic outbreak as the average latent period of the infectors over all infector-infectee pairs. This means, for a simulation that produces an outbreak (i.e. for which the reproduction number calculation window has finite bounds t_{start} and t_{end}), the realized latent period was calculated as:

$$T_{Lr} = \frac{1}{N_P} \sum_{(A,B) \in P} (t_{A,1} - t_{A,0})$$

where:

- P is the set of all infector-infectee pairs (A, B) within the simulation, where A is the infector and B is the infectee;
- N_P is the number of elements in P ;
- $t_{i,0}$ is the time of infection of individual i (in this case, $i = A$ is the infector).
- $t_{i,1}$ is the time when individual i became infectious.

Notice that in the above expression the latent period of each infector i is counted r_i times, where r_i is the number of secondary infections generated by the individual. This makes the realized latent period to be influenced by the number of secondary infections, with individuals that created more infectious contributing more to the average value.

Realized infectious period

Following a similar logic as for the realized latent period, we define the *realized infectious period* of an infector-infectee pair as the time between the infector becoming infectious and the secondary infection. For a simulation that produces an outbreak, the realized infectious period was calculated as:

$$T_{Ir} = \frac{1}{N_p} \sum_{(A,B) \in P} (t_{B,0} - t_{A,1})$$

Where:

- P is the set of all infector-infectee pairs (A, B) within the simulation, where A is the infector and B is the infectee;
- N_p is the number of elements in P ;
- $t_{i,0}$ is the time of infection of individual i ;
- $t_{i,1}$ is the time when individual i becomes infectious.

Since the realized infectious period considers the time of the secondary infection rather than the end of the infectious period itself, it accounts the effective time until infections occur. For example: If an individual is infectious for 9 days, but creates a secondary infection 2 days after becoming infectious, the realized infectious period of this pair is only 2 days, reflecting the fact that most of the time during which the infector was infectious was not “utilized”. This difference in definition from the *intrinsic* infectious period is intentional, as it reveals the effects of the transmission chain in time-to-event metrics, making the realized infectious period ideal for comparing the VLP and GP formulations.

Realized generation time

The generation time of an infector-infectee pair is the time between the primary and secondary infections – i.e. between the infection of the infector and the infection of the infectee. The *realized generation time* of a simulation that produces an outbreak was calculated as:

$$T_{Gr} = \frac{1}{N_p} \sum_{(A,B) \in P} (t_{B,0} - t_{A,0})$$

where:

- P is the set of all infector-infectee pairs (A, B) within the simulation, where A is the infector and B is the infectee;
- N_p is the number of elements in P ;
- $t_{i,0}$ is the time of infection of individual i .

Once again, the realized generation time is averaged over infector-infectee pairs rather than just infectors, meaning that individuals with more secondary infections will have more significant contributions to the average.

Infection attack rate

The infection attack rate is the fraction of the entire population that was infected any time during the entire outbreak. For any simulation, the infection attack rate is calculated as:

$$IAR = \frac{N_I}{N}$$

where N_I is the number of individuals who were infected at some point, either index cases or during simulations, and N is the number of individuals in the population.

Infection incidence peak time and size

The daily infection incidence I_d , where d is a given day during the simulation, is defined as the number of infections that occurred during that day. For a simulation that produces an outbreak, the infection incidence peak day is calculated as d such that I_d is greater than or equals to any other day during the simulation. If more than one day satisfies this condition, the latest day was chosen. The peak incidence of daily infections is defined as I_d at the peak day d .

Average numbers of secondary infections by latent and effective infectious periods

In Figures 2.B and 3.B of the main text, we showed the average number of secondary infections as a function of the latent period (2.B) and effective infectious period (3.B). These were calculated as binned averages over the set of infector-infectee pairs, where bins were defined by the selected time-to-event metric, and different stochastic executions were aggregated together into a single average.

For example, select the latent period as the time-to-event metric. We first aggregated the latent periods of all infector-infectee pairs into daily bins, constructing a subset of the infector-infectee pairs $P_d = \{ (A, B) \mid T_L(B|A) \in [d - 1, d] \}$ for each period $d = 1, 2, 3, \dots$ days, where $T_L(B|A)$ is the latent period of the infector A of the infector-infectee pair (A, B) . In other words, P_d is the set of all infector-infectee pairs with a latent period of d days rounded up. For a simulation that produced outbreak, the average number of secondary infections binned by a latent period of d days was calculated as:

$$R_d = \frac{1}{M_{P_d}} \cdot \sum_{(A,B) \in P_d} r_A$$

where:

- r_i is the number of secondary infections generated by individual i ;
- M_{P_d} is the number of elements in set P_d .

The metric above was used to produce the numbers for Figure 2.B. The same logic was applied to calculate the average number of secondary infections as a function of the effective infectious period (Figure 3.B).

Distributions: Generation time

In Figure 4.B of the main text, we showed probability distributions of the generation time for different reproduction numbers. For each simulation, we calculated this by counting the number of infector-infectee pairs based on their associated generation time, then normalizing it as a probability density function. The probability density F_T of a given value T of the generation time was calculated as:

$$F_T = \frac{N_{G_T}}{dt \cdot N_p}$$

where:

- N_{G_T} is the number of infector-infectee pairs during the simulation for which the generation time was exactly T ;
- N_p is the number of infector-infectee pairs during the entire simulation;
- $dt = 0.25$ days is the length of the step size of the simulation.

Metrics interpolation as a function of the reproduction number

Figures 2.A, 3.A and 4.A show different time-to-event metrics as a function of the reproduction number during the growth phase. Since we only simulated VIBES for a discrete grid of transmission rate values (α), we have interpolated these metrics to obtain smoother plots and metric estimates between the discretely obtained values. Specifically, for each curve of a metric Y as a function of the reproduction number R , we first obtained the points $\{(R_s, Y_s)\}$ of mean reproduction number R_s and mean time-to-event metric Y_s for each set s of simulations with the same transmission rate α , then interpolated these points with an exact second-order spline. This interpolation was then used to obtain estimates of Y for any reproduction number R within the minimum and maximum values of R achieved by the simulations. Besides the mean estimates, the same interpolation was used to calculate the 95% interquartile ranges and the standard deviations.

Table S1. List of parameters used for the VIBES simulations reported in the main analysis.

Parameter description	Parameter value
Synthetic population	
Study population	Indiana, USA
Number of individuals (population size)	502,697
Number of households	200,000
Transmission model parameters	
Overall transmission rate	$\alpha \in \{0.35, 0.50, 0.67, 0.85, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0\}$
Setting-specific transmission risks (fraction of transmission by setting) [#] :	
Households	1.0 {35%}
Schools	0.418 {15%}
Workplaces	0.307 {15%}
Community	0.304 {35%}
Individual infectiousness (VLP only)	$g_i \sim \text{Gamma}$ shape= 0.235

	scale= 4.26
GP: Average latent period duration	$\Delta t + \frac{1}{\eta} = 4.00$ days
GP: Average infectious period duration	$\Delta t + \frac{1}{\gamma} = 3.75$ days
GP: Symptom onset delay after the start of the infectious period	$T_{OS} = 2.25$ days
VIBES: Infectiousness threshold	10^5 copies/ml (Zou et al ³)
Infectiousness of symptomatic individuals relative to asymptomatic individuals	$\psi = 1$
Probability of developing symptoms by age group ^{\$} :	
0-19 years old	18.1%
20-39 years old	22.4%
40-59 years old	30.5%
60-79 years old	35.5%
80+ years old	54.6%
Simulation parameters	
Length of a time step	$\Delta t = 0.25$ days
Number of simulations for each value of α	50
Number of index cases (infections at the start of each simulation)	30
Reproduction number calculation	
Start threshold for calculating the reproduction number	$c_0 = 1000$
Stop threshold for calculating the reproduction number	$c_{end} = 7000$

Transmission risks and fraction of transmission by setting refer to the scenario $R=1.3$, as estimated in Hu et al.⁴ and Liu et al.⁵

\$ Reference: Poletti et al. ¹

Additional results

Temporal dynamics of each parameterization

In the main text, we showed how different model parameterizations impact outbreak features such as the realized generation time, latent period and infectious period. These features are directly linked to the time evolution of an outbreak, therefore different parameterizations should display different dynamics over time. Since the GP was parameterized to match the generation time of VLP at $R = 2.5$, simulations from both parameterizations with the same reproduction number of $R = 2.5$ produce similar infection incidence profiles over time (Fig. S1A). However, by changing the reproduction number to either lower or higher values, a mismatch between the temporal

profiles of the two parameterizations emerges (Fig. S1B and S1C). This happens because the time-to-event metrics change at different rates with the reproduction number between the two parameterizations, as discussed in the main text. More specifically, the realized generation time decreases as R increases in both parameterizations, but this happens at a lower rate for the VLP compared with the GP. Since the generation time dictates the temporal evolution of infection chains, simulations with the GP predict a shorter outbreak than those with the VLP if $R > 2.5$, even though both simulations have the same reproduction number. A similar pattern can be observed in the temporal evolution of the instantaneous reproduction number (Fig. S1D, E and F).

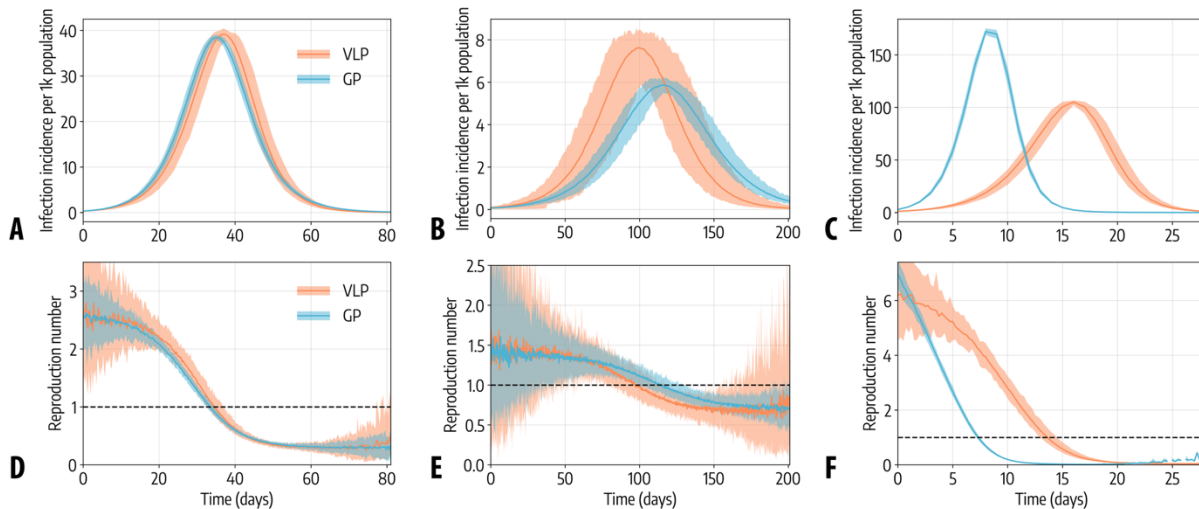


Figure S1. **A.** Incidence of new infections normalized by 1000 individuals as a function of the simulation time in days. Lines represent the mean over the 50 stochastic realizations, while the shaded areas represent the 95% IQR. Red lines and colors represent the VLP, while blue represents the GP. The reproduction number for both parameterizations was set to $R = 2.5$. **B.** Same as A for $R = 1.5$. **C.** Same as A for $R = 6.0$. **D.** Instantaneous reproduction number as a function of the simulation time in days, following the same color code as in A. The initial reproduction number was set to $R = 2.5$. **E.** Same as D for $R = 1.5$. **F.** Same as D for $R = 6.0$.

References

1. Poletti, P. *et al.* Association of Age With Likelihood of Developing Symptoms and Critical Disease Among Close Contacts Exposed to Patients With Confirmed SARS-CoV-2 Infection in Italy. *JAMA Netw. Open* **4**, e211085 (2021).
2. Ventura, P. C. *et al.* VIBES: A multiscale modeling approach integrating within-host and between-hosts dynamics in epidemics. *Proc. Natl. Acad. Sci.* **123**, e2523055123 (2026).
3. Zou, L. *et al.* SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. *N. Engl. J. Med.* **382**, 1177–1179 (2020).

4. Hu, S. *et al.* Infectivity, susceptibility, and risk factors associated with SARS-CoV-2 transmission under intensive contact tracing in Hunan, China. *Nat. Commun.* **12**, 1533 (2021).
5. Liu, Q.-H. *et al.* Model-based evaluation of alternative reactive class closure strategies against COVID-19. *Nat. Commun.* **13**, 322 (2022).