Efficient coupling of within-and between-host infectious disease dynamics

Cameron A. Smith [a,d] , Ben Ashby [b,c,d,*]

[a] *Department of Biology University of Oxford Oxford UK*
[b] *Department of Mathematics, Simon Fraser University Burnaby BC Canada*
[c] *Pacific Institute on Pathogens, Pandemics and Society Burnaby BC Canada*
[d] *Department of Mathematical Sciences, University of Bath Bath UK*

A B S T R A C T

Mathematical models of infectious disease transmission typically neglect within-host dynamics. Yet within-host dynamics – including pathogen replication, host immune responses, and interactions with microbiota – are crucial not only for determining the progression of disease at the individual level, but also for driving within-host evolution and onwards transmission, and therefore shape dynamics at the population level. Various approaches have been proposed to model both within- and between-host dynamics, but these typically require considerable simplifying assumptions to couple processes at contrasting scales (e.g., the within-host dynamics quickly reach a steady state) or are computationally intensive. Here we propose a novel, readily adaptable and broadly applicable method for modelling both within- and between-host processes which can fully couple dynamics across scales and is both realistic and computationally efficient. By individually tracking the deterministic within-host dynamics of infected individuals, and stochastically coupling these to continuous host state variables at the population-level, we take advantage of fast numerical methods at both scales while still capturing individual transient within-host dynamics and stochasticity in transmission between hosts. Our approach closely agrees with full stochastic individual-based simulations and is especially useful when the within-host dynamics do not rapidly reach a steady state or over longer timescales to track pathogen evolution. By applying our method to different pathogen growth scenarios we show how common simplifying assumptions fundamentally change epidemiological and evolutionary dynamics.

1. Introduction

Since their inception nearly a century ago, compartmental models of infectious diseases have proved to be incredibly powerful for modelling epidemiological dynamics while requiring relatively few parameters and assumptions (Anderson and May 1991; Kermack and McKendrick, 1927). Compartmental models are not only widely used for understanding and predicting disease transmission, but also provide crucial insights into fundamental evolutionary processes in pathogens, including selection for virulence (Cressler et al., 2016), immune evasion (Saad-Roy, 2021), and anti-microbial resistance (Blanquart, 2019). Such models also shed light on pathogen coevolution with their host populations, including humans (Penman et al., 2013), animals (May and Anderson, 1983), plants (Frank, 1992), fungi (Sandhu et al., 2021), and bacteria (Koonin and Wolf, 2015). The development of realistic and efficient modelling frameworks has therefore been identified as a key goal for the modelling community, especially those that link processes across scales (Gog, 2015).

For simplicity, compartmental models typically only focus on epidemiological dynamics at the between-host level (e.g., transmission) and neglect within-host processes (e.g., pathogen growth and immune response dynamics). This greatly simplifies the difficulty of modelling nonlinear dynamics in large populations as one can classify hosts into distinct groups according to their current disease status (e.g., "susceptible", "infected", or "recovered") without needing to track individual pathogen dynamics within each infected host. Movement between disease-status compartments occurs due to processes such as infection and recovery that do not depend on individual within-host properties such as pathogen load or T-cell count. When fitting to real-world data, this means that key disease parameters, such as the basic reproduction number or serial interval, may be inferred from population-level data without knowledge of the dynamics that occur within hosts.

Although between-host compartmental models give vital insights into pathogen epidemiology and evolution, within-host dynamics

---

including pathogen replication and growth, the host immune response, pharmacokinetics and pharmacodynamics, and interactions with the microbiome and other pathogens are also known to play a critical role (Mideo et al., 2008). For example, transmission and virulence are both likely to depend on pathogen load. Moreover, theoretical studies have shown that within-host dynamics can play a crucial role in pathogen evolution, leading to immune escape (Smith and Ashby, 2023) or more varied predictions for virulence evolution (Alizon and Lion, 2011). Dynamics at the within-host level can clearly have important implications at the population-level, which can be captured using nested or hybrid models that couple dynamics at the intra- and inter-host scales (Mideo et al., 2008).

While multiple approaches to nested models have been developed (Mideo et al., 2008), these methods are not widely used because tracking the within-host dynamics of all infected individuals in a population is computationally intensive and analytically intractable. For example, in a standard *SIR* model lacking within-host dynamics, only three equations are required to track the population dynamics, whereas a full nested model requires at least the same number of equations as infected individuals. Moreover, if the transmission rate depends on pathogen load, which varies through time at the individual level, then each infected host has a unique time-dependent transmission rate. Many nested models therefore assume that within-host processes rapidly reach a steady state as this eliminates the need to track transient within-host dynamics. However, this separation of timescales does not capture the impact of transient within-host dynamics on transmission and virulence and fails when the within-host dynamics are slower or do not reach a steady state (e.g., acute infections). Stochastic individual-based models (IBMs) (Althouse and Hanley, 2015; McKenzie and Bossert, 2005; Kennedy and Dwyer, 2018; Lindberg et al., 2018) or integro-differential equations (IDEs) (Handel et al., 2013; Gilchrist and Sasaki, 2002; Coombs et al., 2007; Azevedo et al., 2018) have been used to fully capture transient dynamics, but IBMs are computationally intensive, and IDEs can only be used when the within-host dynamics are relatively simple. More recently, there has been a larger push towards fully stochastic methods to capture the two scales. For example, Yin *et al.* proposed a fully stochastic method which is able to capture individual within-host dynamics (Yin et al., 2024). The method uses an event driven approach, which has the ability to capture significant detail, but at the cost of simulation speed. Another recent study, by Zhang *et al.*, proposed a multiscale method for modelling within-host evolution and infection on networks (Zhang, 2022). This method uses a Markov-chain to determine mutations at the within-host scale, with mixtures of different mutants determining fitness at both the within- and between-host level. Transmission then occurs on a network. This method has many applications for determining when mutations can cause epidemics and pandemics, but doesn't include mechanistic within-host dynamics. Generally, methods that fully capture transient and mechanistic within-host dynamics and are computationally efficient are currently lacking.

Here, we propose a novel hybrid method for coupling transient within- and between-host dynamics that is both computationally efficient and realistic. Our method leverages different modelling paradigms across scales to track individual-level within-host dynamics for a discrete number of infected hosts, which are then stochastically coupled to continuous population-level processes. The main novelty of this framework is the ability to integrate multiple methods to leverage the advantages of fast numerical solvers for ODEs for the within-host dynamics of each infected individual and for demographic processes at the population level, while still capturing stochasticity through the coupling process between scales. This allows us to more efficiently capture transient within-host dynamics for each infected individual than previous methods. Further, we demonstrate the ability for the method to track genetic evolution and construct phylogenetic trees. We illustrate our method with contrasting examples of within-host pathogen dynamics, showing the critical impact on epidemiological and evolutionary dynamics at the population level.

2. Methods

We introduce a novel method for coupling infectious disease dynamics within individual hosts with population-level processes. For simplicity, we focus on a single directly transmitted pathogen in a homogeneous, randomly mixing host population, but our approach may be readily generalised for multiple co-circulating pathogens or symbionts, alternative modes of transmission (e.g., vector-borne or sexually transmitted), or for heterogeneous or structured populations. The population consists of $N(t)$ individuals at time t. Each host has a given disease state (e.g., susceptible, infected, or recovered), with the density of each state tracked at the population level (*see Host demographic* dynamics; HDD), and a within-host state (e.g., pathogen load, T-cell count), where applicable (see *Within-host dynamics*; WHD*)*. To couple the different scales efficiently, we convert between a discrete number of hosts at the WHD level and a continuous density of hosts at the population-level (see *Coupling* and schematic in Fig. 1). Below, we first describe general frameworks for modelling the within-host dynamics of an infectious disease and the demographic dynamics of the host population, before introducing our novel hybrid method for coupling dynamics across these two scales.

2.1. Within-host dynamics

The within-host dynamics (WHD) of an infectious disease may include replication and growth of the pathogen population from an initial infectious dose, the host immune response and its interactions with the pathogen, positive or negative interactions with other pathogens or constituents of the host microbiome, and within-host evolution of the pathogen through selection or drift. To model the WHD, we assign a unique set of ordinary differential equations (ODEs) to every individual in the population (although in practice one might limit the within-host dynamics only to currently infected or recently recovered individuals).

Consider an individual indexed by i who was infected T_i time-units ago. At time T_i, their within-host state is given by the function $\boldsymbol{W}_i(\theta_W, T_i)$, where θ_W contains all parameters that are required for the evaluation of the within-host state. These parameters could include, but are not limited to, rates for pathogen replication and the upregulation of an immune response. The within-host dynamics are then given by:

$$\frac{d\boldsymbol{W}_i}{dT_i} = \boldsymbol{f}_W(\boldsymbol{W}_i, \theta_W, T_i), \boldsymbol{W}_i(\theta_W, 0) = \widetilde{\boldsymbol{W}}(\theta_w). \tag{1}$$

where $\boldsymbol{f}_W(\boldsymbol{W}_i, \theta_W, T_i)$ describes how the within-host states change subject to the initial condition $\widetilde{\boldsymbol{W}}(\theta_W)$, which could be a function of the infecting individual (see *Coupling*).

2.2. Host demographic dynamics

At the population level, we distinguish host demographic dynamics (HDD), such as births and background mortality, from disease-related processes such as transmission, recovery and disease-associated mortality (virulence), which are captured through the coupling of the two scales (see *Coupling*). In principle, one could include other processes such as migration or mating dynamics (for non-sexually transmitted pathogens in animal populations) at this scale.

As with the WHD, one can model the HDD stochastically, or deterministically using the following general ODEs:

$$\frac{d\boldsymbol{D}}{dt} = \boldsymbol{f}_D(\boldsymbol{D}, \theta_D, t), \boldsymbol{D}(\theta_D, 0) = \widetilde{\boldsymbol{D}}(\theta_D), \tag{2}$$

Where the host demographic state at time t is given by $\boldsymbol{D}(\theta_D, t)$ with parameter set θ_D and initial condition $\widetilde{\boldsymbol{D}}(\theta_D)$, and the function $\boldsymbol{f}_D(\boldsymbol{D}, \theta_D, t)$ describes the change in the density of each host state due to non-disease related processes (e.g., births and background mortality). Parameters at
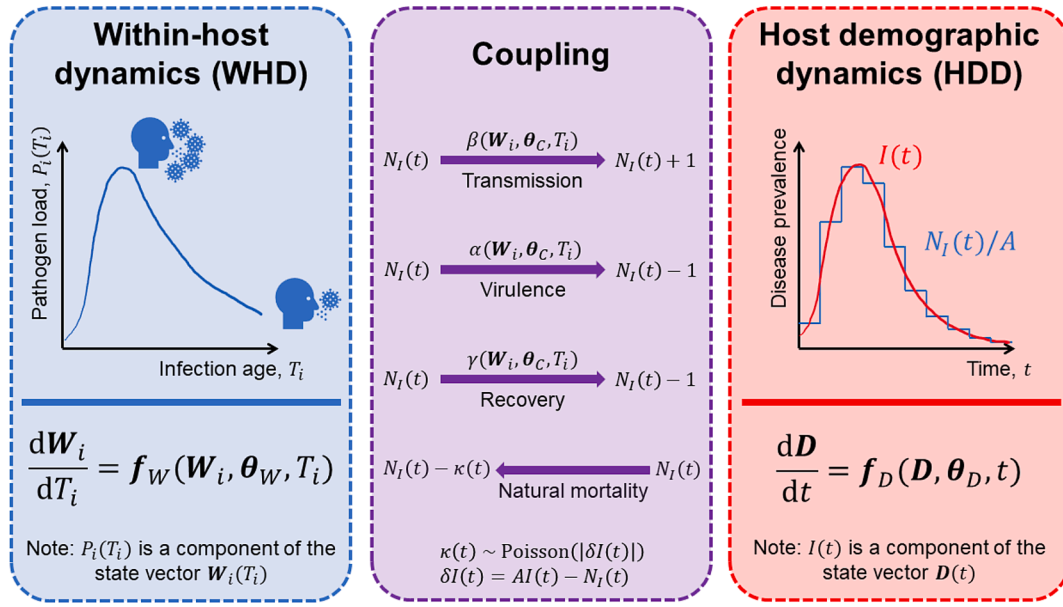
Fig. 1. Model schematic for the hybrid method. The within-host dynamics (WHD) track pathogen replication and immune dynamics for each infected host, i at infection time T_i, in state vector $W_i(T_i)$ for a discrete number of infected hosts. The WHD have instantaneous rate of change $f_W(W_i, \theta_W, T_i)$, where θ_W is a vector of within-host parameters. The host-demographic dynamics (HDD) track continuous densities of population-level disease states and include births, natural mortality and migration in state vector $D(t)$, which changes according to instantaneous rate of change $f_D(D, \theta_D, t)$, where θ_D is a vector of host demographic parameters and t is the population-level time. The coupling functions link these two scales together stochastically through epidemiological processes (transmission, mortality virulence, and recovery) to update the host disease states at the HDD level, and by removing infected individuals at the WHD level due to natural mortality.

this scale would include, for example, rates for host birth and death, or migration.

2.3. Coupling

We now describe coupling of the within-host (WHD) and host demographic (HDD) dynamics, together with the change in host state due to disease-related processes (infection, recovery, death) (Fig. 1). Recall that hosts are tracked individually for the WHD, but the HDD use population densities. We focus on infected individuals, but the same approach could be applied to any host state. The density of infected hosts at time t, $I(t) \in \mathbb{R}_+$, is given by $I(t) \approx N_I(t)/A$, where $N_I(t) \in \mathbb{N}$ is the number of infected individuals at time t and A is a positive constant representing the area over which individuals interact. Note that continuous compartmental models of infectious diseases track population densities rather than population sizes (Anderson and May 1991). We use stochastic updates to determine the number of times each event takes place over a small time-interval $[t, t + \Delta]$, using the tau-leap method (Gillespie, 2001). Throughout, we will set $\Delta = 0.5$ days. The HDD state variables and the discrete individuals are then updated simultaneously. We focus on infection, recovery, and virulence as disease-related processes, but the method could be readily extended to other processes (e.g., transitions to or from latency, or between asymptomatic and symptomatic disease).

The general transmission function, $\beta(W, \theta_C, T)$, describes the rate at which a host with infection age T and within-host state W transmits the pathogen to any individual currently susceptible to the infection, where θ_C denotes any coupling parameters (e.g., defining the relationship between pathogen load and transmission). We use the function $\beta(W, \theta_C, T)$ to stochastically calculate the number of new infections arising from each infected individual in a short period of time, $[T, T + \Delta]$, using the τ-leap method (Gillespie, 2001). The number of new infections produced by infected individual i during this time interval is drawn from the probability distribution

$$n_i(t) \sim Poisson(\beta(W_i, \theta_C, T_i)S(t)\Delta). \tag{3}$$

We then update the number of infected individuals, $N_I(t)$, and the population-level host states, $D(t)$, accordingly by a mass equivalent to one individual $(1/A)$ from the susceptible to the infected state:

$$S(t + \Delta) = S(t) - \frac{1}{A}\sum_{i=1}^{N_I(t)} n_i(t), \tag{4}$$

$$I(t + \Delta) = I(t) + \frac{1}{A}\sum_{i=1}^{N_I(t)} n_i(t). \tag{5}$$

The rate at which an individual with infection age T recovers is given by $\gamma(W, \theta_C, T)$. For example, the recovery rate may depend on the current pathogen load and T-cell count, or for simplicity may be treated as a constant. To determine if an infected individual i has recovered in this time interval, we draw a uniform random number between 0 and 1, $u_{r,i}$, and if $u_{r,i} < \gamma(W_i, \theta_C, T_i)\Delta$ then the individual has recovered and we update the population-level and within-host states accordingly and no longer track their within-host dynamics. We use a similar process to determine if an individual has died from disease during this time interval, where the disease-associated mortality rate is $\alpha(W, \theta_C, T)$. We update the number of infected individuals $N_I(t + \Delta)$ accordingly, and the population-level states by subtracting $1/A$ (the density of a single individual) from the infected state, $I(t)$, for each recovery or mortality event, and adding $1/A$ per recovery event to the recovered or susceptible states depending on the assumptions regarding recovery.

The final stage of the coupling reconciles changes in the HDD with the number of individuals tracked for the WHD. For example, infected individuals may die due to other causes, or the number of infected individuals may change due to migration. We therefore need to ensure that the number of individuals at the WHD level is approximately equivalent to the appropriate HDD state variable. For simplicity, we assume that only infected individuals are tracked for the WHD and that the HDD may result in non-disease-associated mortality, but an analogous approach would apply to other scenarios.

Recall that the relationship between the continuous density of infected individuals in the HDD, $I(t)$, and the discrete number of infected individuals in the WHD, $N_I(t)$, is given by $I(t) \approx N_I(t)/A$. If $\delta I(t) = AI(t) - N_I(t) > 0$ then the number of tracked individuals in the

WHD is fewer than the number of individuals according to the HDD. Conversely, if $\delta I(t) < 0$, then the number of tracked individuals is greater than in the HDD. This discrepancy is corrected probabilistically. We draw a Poisson random variable $\kappa(t) \sim Poisson(|\delta I(t)|)$ which determines the number of tracked individuals to either add ($\delta I(t) > 0$) or remove ($\delta I(t) < 0$) at random. In our simulations, we cannot have $\delta I(t) > 0$ due to the absence of migration. When $\delta I(t) < 0$, we choose $\kappa(t)$ infected individuals at random to be removed due to non-disease-associated mortality.

2.4. Applications

We illustrate possible applications of our novel coupling method for two pathogen growth scenarios. In the first case, we assume that following infection the pathogen load grows logistically to a steady state and is not cleared by an immune response, whereas in the second case we assume that the pathogen load grows exponentially for a short period before decaying to a lower level due to a build-up of immunity within the host. For simplicity of illustration (but not a limitation of the framework), we do not explicitly model the immune response, but the dynamics of T cells, for example, could readily be incorporated into the WHD. The pathogen load dynamics presented below are qualitatively similar to models with explicit immunity. We compare the full transient dynamics for pathogen load with a steady state approximation, which is commonly used in hybrid models of within- and between-host dynamics (Mideo et al., 2008; Gilchrist and Sasaki, 2002; Gilchrist and Coombs, 2006; Almocera and Hernandez-Vargas, 2019; Garira, 2018). We also demonstrate how our modelling framework can be used to study pathogen evolution and construct phylogenetic trees. Neither of these examples are meant to be representations of all diseases, nor should their quantitative results be perceived as such. The purpose of these examples is to demonstrate the method with different qualitative within-host dynamics which are designed to be proxies for a wide range of possible infections experienced by hosts across the tree of life.

For our first example, the within-host pathogen load grows logistically with growth rate r and carrying capacity K. This may be a reasonable model for infections that are not cleared by the immune system, such as certain sexually transmitted pathogens such as *Chlamydia trachomatis* and opportunistic pathogens in immunocompromised hosts (e.g., *Pseudomonas aeruginosa* infections in cystic fibrosis patients), and for non-pathogenic gut microbiota. We assume that the HDD feature

a density-dependent birth term and that all individuals have the same background mortality rate, regardless of their disease status. We include two epidemiological coupling functions: one for transmission (which we assume to be proportional to pathogen load); and one for virulence (which we assume is proportional to the square of the pathogen load). We also assume that there is no recovery from infection. A full description of the model is given in the *Supplementary material,* Section S1. Unsurprisingly, both the epidemic growth rate and the size of the epidemic is typically much lower when the transient WHD are taken into account compared to the steady state approximation, especially when the within-host growth rate (r) is slow relative to the average infectious period (Fig. 2). This is because the transmission rate increases with pathogen load and the steady state approximation therefore overestimates transmission, whereas the transient WHD accounts for a period of relatively low transmission early in the infection.

In our second example, the pathogen load initially grows exponentially before peaking and decaying due to an immune response. This is a reasonable approximation for many acute infections, including COVID-19, influenza, and measles, and for the early stages of many chronic infections, such as Syphilis and HIV, where pathogen load peaks soon after infection then falls to lower levels for a long period of time before eventual resurgence. The WHD for pathogen load are controlled by three parameters: r, which scales the overall pathogen load; η, which controls the strength of pathogen decay due to an immune response; and P^*, which is the pathogen load steady state. For comparison with the steady state approximation, we assume that the steady state is non-zero. The transmission and virulence functions are the same as in the previous example, but now we also include a recovery rate, which increases exponentially as pathogen load falls. If an individual recovers, they gain full immunity and cannot be reinfected, and their pathogen load is set to zero. A full description of the model can be found in *Supplementary Material* Section S2.

As in the previous example, modelling the full WHD using our coupling method leads to stark differences in epidemic dynamics compared to a steady state approximation (Fig. 3). Since the steady state is relatively low compared to peak pathogen load, this approximation neglects a significantly higher transmission rate during the early stages of infection, and therefore underestimates both the size and growth rate of the epidemic. Naturally, this effect will be stronger the greater the difference between the peak pathogen load and the steady state, and
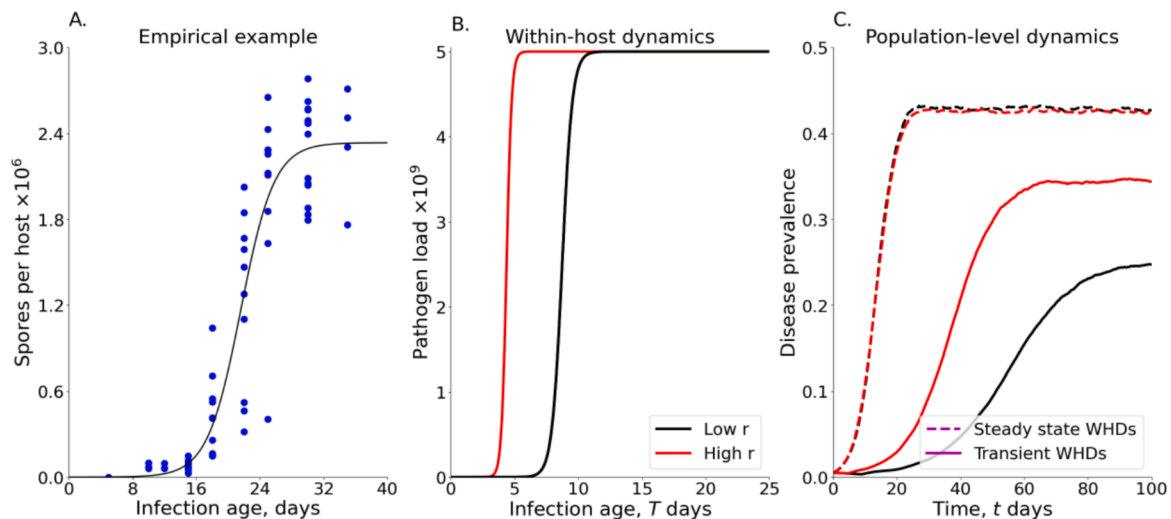

Fig. 2. Epidemiological dynamics with logistic pathogen growth. A. Logistic curve fitted to within-host growth of the bacterial pathogen Pasteuria ramosa in water fleas (Daphnia dentifera) from (Auld et al., 2014). B. WHD of pathogen load with relatively slow (r = 2.5; black) and fast (r = 5; red) pathogen load growth rates. C. Disease prevalence in the host population for relatively slow (r = 2.5; black) and fast (r = 5; red) pathogen load growth rates, with a steady-state approximation (dashed) or using the full transient dynamics with our novel coupling method (solid). Other parameters as detailed in the Supplementary Material (section S1.5). Units for all parameters are dimensionless and are chosen simply to illustrate the links between within and between-host dynamics for a generic pathogen.
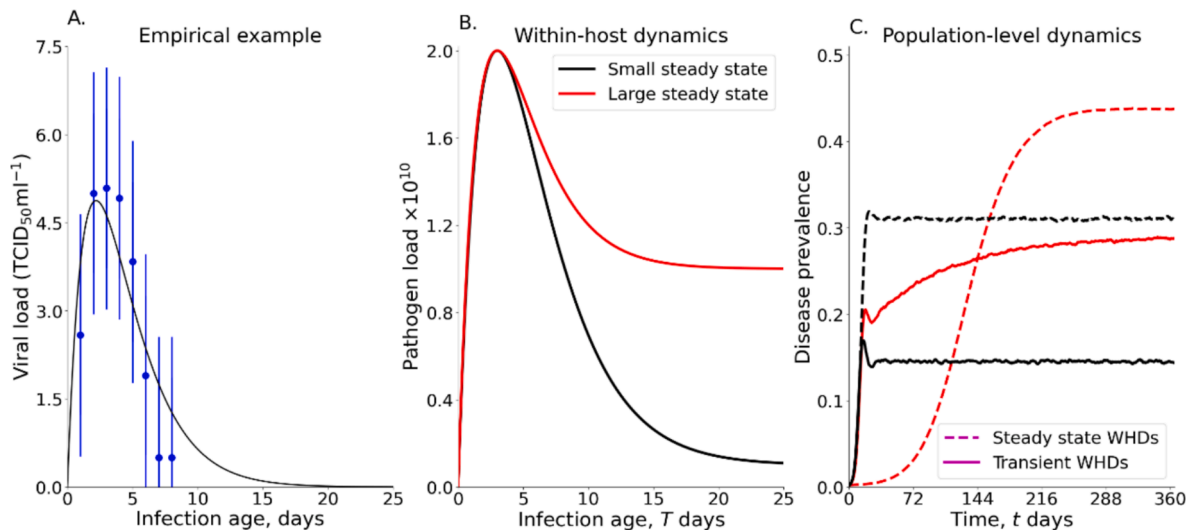
Fig. 3. Epidemiological dynamics when pathogen load peaks and decays. A. Pathogen load (equation (S4); black) fitted to the within-host viral load of averaged influenza A cases (blue) from (Baccam et al., 2006). B. WHD of pathogen load with relatively low ($P^* = 10^9$; black) and high ($P^* = 10^{10}$; red) long-term pathogen loads. C. Disease prevalence in the host population for small ($P^* = 10^9$; black) and large ($P^* = 10^{10}$; red) long-term pathogen loads at the within-host level, with a steady-state approximation (dashed) or using the full transient dynamics with our novel coupling method (solid). Other parameters as detailed in the Supplementary Material (section S2.5). Units for all parameters are dimensionless and are chosen simply to illustrate the links between within and between-host dynamics for a generic pathogen.

when most transmission occurs during the early stages of an infection.

Not only is our method more realistic than a steady-state approximation, if is also much more computationally efficient than a full stochastic simulation algorithm (SSA). For example, the epidemic simulations in Fig. 2 were approximately 3.5 times faster using our modelling framework than a classical SSA (see *Supplementary Material* Section S4), and 1.7 times faster for the simulations in Fig. 3.

Our modelling framework can also be applied to understand the effects of WHD on pathogen evolution and patterns that emerge at the population level. For example, suppose we consider the evolution of the pathogen growth rate, *r*, in our first example (for a full description of the following simulations, see *Supplementary Material*, Section S1.4). In the steady state approximation, the growth rate has no effect on pathogen fitness, which only depends on the carrying capacity, *K*. The pathogen therefore only experiences drift rather than selection, resulting in a random walk through trait space (Fig. 4A). In contrast, when the transient WHD are modelled using our framework, the pathogen growth rate impacts on transmission and virulence, and therefore fitness, leading to selection towards an evolutionarily stable relationship between pathogen load, transmission, and virulence (Fig. 4B).

As our method tracks infected individuals, along with who infects whom, it can also be readily used to construct true phylogenetic or transmission trees from epidemic simulations with perfect data on infection histories (Fig. 5; *Supplementary Material section S2.4*). This not only has potential uses for efficiently simulating realistic epidemics for testing novel transmission tree reconstruction methods (Didelot et al., 2017), but could also provide new insights into how evolution driven by within-host dynamics impacts the relationship between transmission trees and phylogenetic trees at the population level.

3. Discussion

Understanding how within-host processes impact on the epidemiological and evolutionary dynamics of infectious diseases is a key objective for the infectious disease community. Yet fully capturing transient within-host dynamics in nested models is a major challenge due to the computational burden of tracking individual infection processes, and so many studies assume a separation of timescales so that the within-host dynamics rapidly reach steady state. We have developed a

novel hybrid method for efficiently and realistically modelling infectious disease dynamics across scales, capturing both population-level processes (e.g., births and deaths) and individual disease-related processes (e.g., pathogen growth and immune responses). By leveraging fast deterministic models at the population level and for each infected host, and stochastically coupling these scales to enable changes in host states (e.g., susceptible to infected), we make significant speed improvements in even the simplest of models while still fully capturing transient within-host dynamics. We illustrated the power of our framework with two examples of within-host pathogen growth, showing how it can be readily used to model both epidemiological and evolutionary dynamics, often leading to contrasting outcomes compared to widely used steady-state approximations.

A critical feature of our framework is that hosts are tracked as discrete entities for the within-host dynamics, but host states are tracked continuously for the population level dynamics. At the within-host level, each infected individual has their own unique set of ODEs to describe changes in internal states such as pathogen load and T-cell count, but at the population-level host states are tracked continuously using a density-based approximation with a single set of ODEs. This approach allows us to leverage fast ODE-solvers at both scales, which are then coupled using a stochastic tau-leaping method to keep the discrete and continuous levels closely aligned. However, the modular nature of our method means that it can be readily adapted to incorporate different types of models. For example, each of these different scales could be modelled using partial or stochastic differential equations, or with stochastic simulation algorithms. The only requirement is that the coupling correctly aligns the number of individuals tracked for the within-host dynamics with the density of the corresponding host states at the population level. Thus, when an individual is infected, for example, the number of individuals tracked at the within-host level is increased by one, and one unit of host "mass" is moved from the susceptible to infected state at the population-level. Naturally, this means that there will usually be slight discrepancies between the two scales (as illustrated in Fig. 1C), but the density of infected hosts never differs by more than one unit of host mass.

Our method was inspired by hybrid methods for modelling particle diffusion and interaction across a spatial domain with large disparities in particle number. Such spatial hybrid methods (Smith and Yates, 2018)
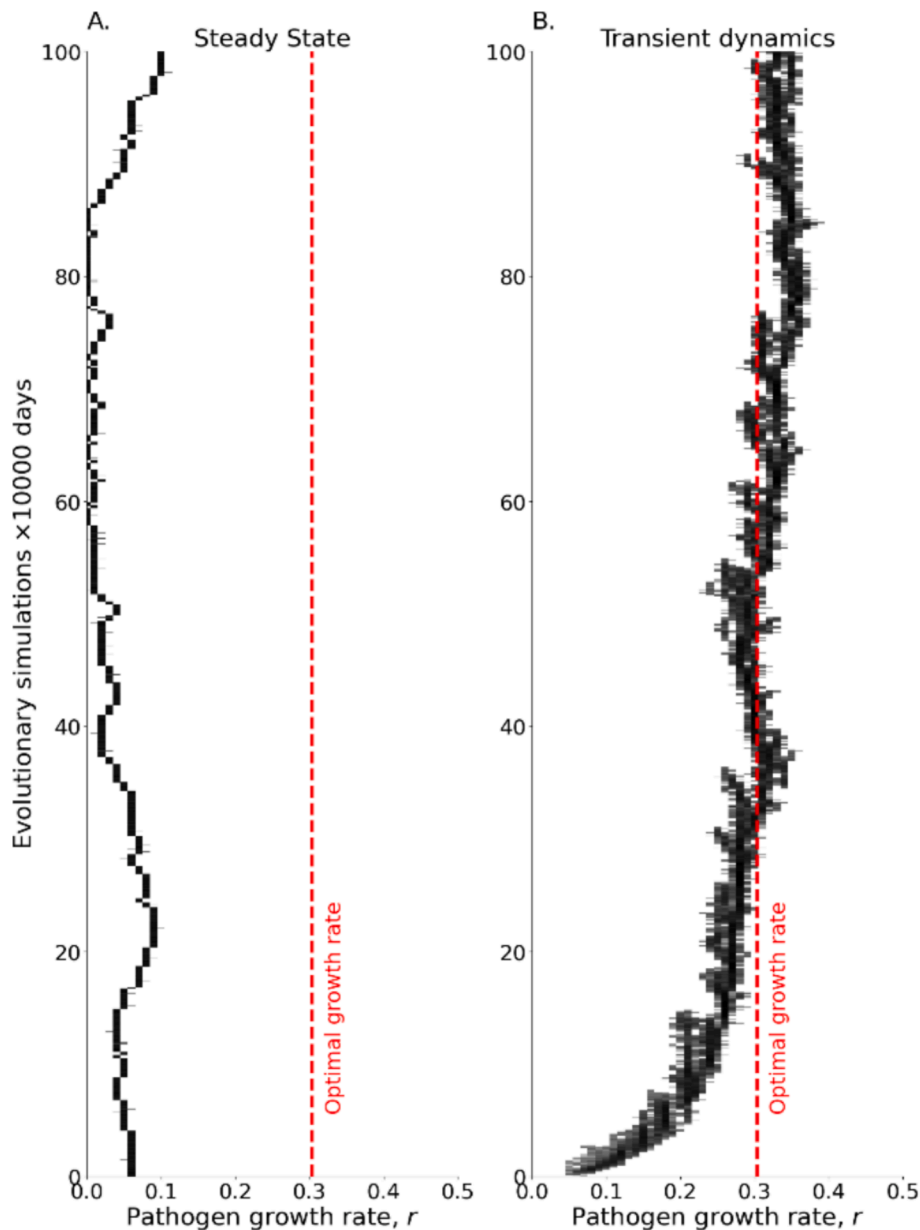
Fig. 4. Evolutionary simulations with logistic pathogen growth. A. Evolutionary simulation using the steady-state approximation for the WHD. B. Evolutionary simulation using transient WHD. The red line denotes the evolutionarily stable level of virulence. All simulations conducted using 500 individuals, with density-dependent term $q = 9.9 \times 10^{-4}$. Other parameters as detailed in the Supplementary Material (section S1.5).

employ different modelling paradigms across the spatial domain, with regions containing large numbers of particles using coarser continuum approaches (e.g., PDEs), while low copy number regions use stochastic methods (e.g., agent-based) where continuum limits are invalid. While the two approaches are related in terms of how they handle contrasting scales, our hybrid epidemiological framework uses distinct modelling paradigms for different processes, as opposed to modelling the same phenomena using disparate modelling paradigms.

We demonstrated our method for two contrasting scenarios for within-host pathogen dynamics – logistic growth and exponential growth followed by decay. The former is a reasonable approximation for certain chronic infections or opportunistic pathogens in immunocompromised hosts where the pathogen is not cleared (e.g., *P. aeruginosa* infections in cystic fibrosis patients), whereas the latter is a good approximation for many acute infections that are cleared due to the host's immune response (e.g., influenza, measles), or for chronic infections where the pathogen load peaks before falling to lower levels (e.

g., HIV, syphilis). In both cases, we show that the transient within-host dynamics lead to substantial shifts in epidemic dynamics at the population-level compared to steady-state approximations. In the case of logistic growth, the steady-state approximation for within-host pathogen load always overestimates the transmission rate, which leads to a much faster growth rate and higher endemic prevalence. With exponential growth and decay in pathogen load, the steady-state approximation can lead to a slower growth rate and lower endemic prevalence. The latter scenario especially highlights issues with using a steady-state approximation, as the steady state for acute infections is a pathogen load of zero, which would clearly prevent any transmission. To facilitate comparisons with the steady state approximation we assumed a non-zero pathogen load, but this is not a limitation of our model (one could set the steady state pathogen load to zero but maintain the pathogen in the population due to the transient dynamics). Our framework therefore allows for acute infections with full pathogen clearance whereas the steady-state approximation does not. A steady state
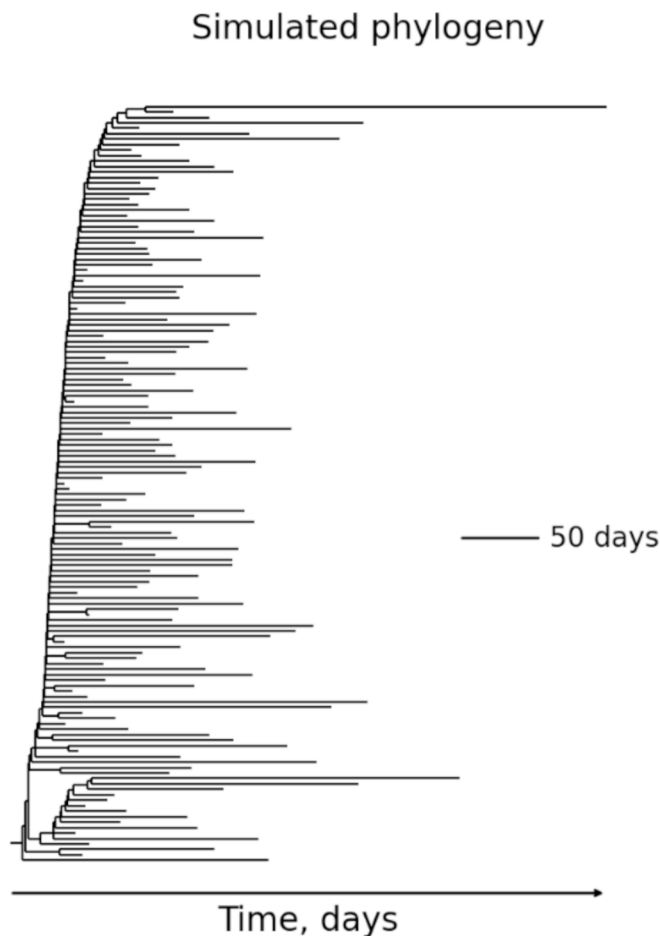
Simulated phylogeny

Fig. 5. Example phylogenetic tree for an evolving pathogen arising from our simulation framework. We have initialised a short length genome (150b in length) and have allowed it to mutate neutrally throughout a population of size 5000. This is the true phylogenetic tree, meaning that all mutation events occurred at the times indicated, rather than a tree reconstruction. Branch length is proportional to time (arbitrary units).

approximation also only allows for unidirectional feedback from within-host to between-host processes (termed "inessential" couplings in [12]) and does not allow for a full ("essential") feedback loop with between-host processes impacting on the within-host dynamics. However, the full feedback loop can readily be captured using our framework, for example by setting the initial inoculum to be a function of the transmission rate. Very few studies incorporate this full feedback loop (Coombs et al., 2007; Alizon and Van Baalen, 2008; Feng et al., 2015), likely due to the predominance of the steady-state approximation, but we expect that our framework will offer new opportunities for exploring the effects of the full within- and between-host feedback loop on epidemic dynamics and pathogen evolution.

We further illustrated how our method can be applied to study pathogen evolution, both phenotypically (e.g., evolution of virulence) and genetically (e.g., for simulating true phylogenetic trees). As with the fundamental shifts observed in epidemiological dynamics, we showed how a steady-state approximation for within-host pathogen dynamics can lead to a random walk through trait space, but there is directional selection towards an evolutionarily stable level of virulence when the transient dynamics are fully captured. This occurs because selection acts on the growth rate of the pathogen, rather than the steady state, which is fixed. Previous theory using IDEs in a few specific scenarios has shown that virulence evolution can be greatly affected by within-host dynamics, which are difficult to capture using conventional methods. These include the order and timing of co-infecting pathogens (Alizon

and Van Baalen, 2008), contrasting selection pressures at the within- and between-host levels (Gilchrist and Coombs, 2006) and the strength of the host immune response (Pugliese, 2011). As our modelling framework does not rely on IDEs, which greatly limit the complexity of within-host dynamics, our method will significantly expand the scope for studying pathogen evolution across scales.

In addition to the phenotypic evolution of key disease traits such as virulence, we have also shown how our framework can be used to efficiently simulate the molecular evolution of pathogens. A major strength of our approach is that infections and transmission dynamics are tracked at the individual level, allowing for the construction of true phylogenetic and transmission trees. This not only offers a novel method to aid in the development of tree reconstruction methods, but would allow, for example, one to explore how different processes at the within-host level impact on phylogenetic patterns that are observed at the population level, or vice versa. While several methods currently exist for simulating phylogenetic and transmission trees using standard epidemiological models (Didelot et al., 2017; Hall et al., 2015), a unique advantage of our approach is that the modular nature of our framework allows one to readily adapt the within-host, population-level, and coupling functions to any feasible scenario. For example, the within-host dynamics could be adapted to account for coinfections of multiple pathogens, pharmaco-dynamics and antimicrobial resistance, interactions with the microbiome, or innate and adaptive responses, and the population-level dynamics could be modified to include migration, vaccination, or non-pharmaceutical interventions. Our framework is therefore likely to provide valuable insights across a wide range of scenarios. Furthermore, while we have focused on host-pathogen interactions our method could readily be applied to model other constituents of the host microbiome, including horizontally or vertically transmitted symbionts (Smith and Ashby, 2023; Ashby and King, 2017) or hyperparasites (Wood and Ashby, 2023).

Our method is highly adaptable and can in principle include any within-host dynamics (e.g. interactions with the immune system or microbiome, complex life-cycles, etc) and any transmission- or population-level features that may be relevant to a particular system (e. g. migration, spatial/contact structure, seasonal effects, etc). *For example, seasonal effects on transmission may be included in the transmission function, so that it becomes a function of the within-host pathogen load, the infection age and the time (as well as any parameters). A simple version of this might be*

$$\beta(P_i, \theta_C, T_i, t) = \widehat{\beta} P_i g(t),$$

where $g(t)$ is a periodic function denoting seasonality. Such modifications can also be made in the right-hand side of any of the ODE functions, so that growth rates are able to be changed seasonally. We efficiently capture full transient dynamics that are often missed by other methods (e.g. steady-state approximations) or which may be more intensive to simulate. However, as our method is simulation-based, other approaches such as integro-differential equations may be able to provide more analytically tractable results (Handel et al., 2013; Gilchrist and Sasaki, 2002; Coombs et al., 2007; Azevedo et al., 2018). Deriving key quantities such as the basic reproductive number or fitness functions analytically is also particularly difficult in multiscale models and requires development of novel methods to address these challenges. While we have demonstrated that our hybrid method is efficient in comparison to the full tau-leap method (see *Supplementary material section S4*), for larger populations a time adaptive method, which will require future development, may be more efficient. Finally, as with any within-host model, parameterisation remains challenging compared to population-level models which are typically simpler and have parameters that are easier to measure. Addressing this challenge will require more detailed longitudinal data of pathogen load and immune responses. However, hybrid models are also useful to understand theoretically how different types of within-host processes (e.g. interactions with the immune system and microbiome)

lead to variation in population-level patterns (e.g. phylodynamics, strength of selection), and hence our framework can also provide support for the development of fundamental theory.Capturing the full complexity of real populations, including all processes at the within-host level, is neither attainable nor desirable. Yet capturing key aspects of within-host pathogen and immune dynamics is crucial for a better understanding of pathogen epidemiology and evolution. While steady-state approximations offer some insights into the impact of within-host processes on population-level dynamics, fully capturing transient interactions between pathogens, host immune responses, and the microbiome is a major challenge that will shed light on pathogen ecology and evolution across scales. Our framework provides a novel, readily adaptable approach for coupling a wide range of within-host and population-level dynamics that is both efficient and realistic.

CRediT authorship contribution statement

Cameron A. Smith: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Ben Ashby:** Writing – review & editing, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jtbi.2025.112061.

References

Anderson, R.M., May, R.M., 1991. *Infectious Diseases of Humans: Dynamics and Control.* Oxford Univ. Press, Oxford.

Kermack, W.O., McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. *Proc. r. Soc. Lond. Ser. Contain. Pap. Math. Phys. Character* 115, 700–721.

Cressler, C.E., McLEOD, D.V., Rozins, C., Van Den Hoogen, J., Day, T., 2016. The adaptive evolution of virulence: a review of theoretical predictions and empirical tests. *Parasitology* 143, 915–930.

Saad-Roy, C.M., et al., 2021. Epidemiological and evolutionary considerations of SARS-CoV-2 vaccine dosing regimes. *Science* 372, 363–370.

Blanquart, F., 2019. Evolutionary epidemiology models to predict the dynamics of antibiotic resistance. *Evol. Appl.* 12, 365–383.

Penman, B.S., Ashby, B., Buckee, C.O., Gupta, S., 2013. Pathogen selection drives nonoverlapping associations between HLA loci. *Proc. Natl. Acad. Sci.* 110, 19645–19650.

May, R., Anderson, R., 1983. Epidemiology and genetics in the coevolution of parasites and hosts. *Proc. r. Soc. Lond. B Biol. Sci.* 219, 281–313.

Frank, S., 1992. Models of plant-pathogen coevolution. *Trends Genet.* 8, 213–219.

Sandhu, S.K., Morozov, A.Y., Holt, R.D., Barfield, M., 2021. Revisiting the role of hyperparasitism in the evolution of virulence. *Am. Nat.* 197, 216–235.

Koonin, E.V., Wolf, Y.I., 2015. Evolution of the CRISPR-Cas adaptive immunity systems in prokaryotes: models and observations on virus–host coevolution. *Mol. Biosyst.* 11, 20–27.

Gog, J.R., et al., 2015. Seven challenges in modeling pathogen dynamics within-host and across scales. *Epidemics* 10, 45–48.

Mideo, N., Alizon, S., Day, T., 2008. Linking within- and between-host dynamics in the evolutionary epidemiology of infectious diseases. *Trends Ecol. Evol.* 23, 511–517.

Smith, C.A., Ashby, B., 2023. Antigenic evolution of SARS-CoV-2 in immunocompromised hosts. *Evol. Med. Public Health* 11, 90–100.

Alizon, S., Lion, S., 2011. Within-host parasite cooperation and the evolution of virulence. *Proc. R. Soc. B Biol. Sci.* 278, 3738–3747.

Althouse, B.M., Hanley, K.A., 2015. The tortoise or the hare? Impacts of within-host dynamics on transmission success of arthropod-borne viruses. *Philos. Trans. R. Soc. B Biol. Sci.* 370, 20140299.

McKenzie, F.E., Bossert, W.H., 2005. An integrated model of Plasmodium falciparum dynamics. *J. Theor. Biol.* 232, 411–426.

Kennedy, D.A., Dwyer, G., 2018. Effects of multiple sources of genetic drift on pathogen variation within hosts. *PLOS Biol.* 16, e2004444.

Lindberg, H.M., McKean, K.A., Caraco, T., Wang, I.N., 2018. Within-host dynamics and random duration of pathogen infection: implications for between-host transmission. *J. Theor. Biol.* 446, 137–148.

Handel, A., Brown, J., Stallknecht, D., Rohani, P., 2013. A multi-scale analysis of influenza A virus fitness trade-offs due to temperature-dependent virus persistence. *PLoS Comput. Biol.* 9, e1002989.

Gilchrist, M.A., Sasaki, A., 2002. Modeling host-parasite coevolution: a nested approach based on mechanistic models. *J. Theor. Biol.* 218, 289–308.

Coombs, D., Gilchrist, M.A., Ball, C.L., 2007. Evaluating the importance of within- and between-host selection pressures on the evolution of chronic pathogens. *Theor. Popul. Biol.* 72, 576–591.

Azevedo, F., Amaku, M., Coutinho, F.A.B., Lopez, L.F., Massad, E., 2018. The effect of the infection within the individual host on its propagation in the population. *Infect. Dis. Model.* 3.

Yin, Y., Flegg, J.A., Flegg, M. B. Accurate stochastic simulation algorithm for multiscale models of infectious diseases. Preprint at http://arxiv.org/abs/2406.05058 (2024).

Zhang, X., et al., 2022. Epidemic spreading under mutually independent intra- and inter-host pathogen evolution. *Nat. Commun.* 13, 6218.

Gillespie, D.T., 2001. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* 115, 1716–1733.

Gilchrist, M.A., Coombs, D., 2006. Evolution of virulence: Interdependence, constraints, and selection using nested models. *Theor. Popul. Biol.* 69.

Almocera, A.E.S., Hernandez-Vargas, E.A., 2019. Coupling multiscale within-host dynamics and between-host transmission with recovery (SIR) dynamics. *Math. Biosci.* 309.

Garira, W., 2018. A primer on multiscale modelling of infectious disease systems. *Infect. Dis. Model.* 3, 176–191.

Auld, S.K.J.R., Hall, S.R., Housley Ochs, J., Sebastian, M., Duffy, M.A., 2014. Predators and patterns of within-host growth can mediate both among-host competition and evolution of transmission potential of parasites. *Am. Nat.* 184, S77–S90.

Baccam, P., Beauchemin, C., Macken, C.A., Hayden, F.G., Perelson, A.S., 2006. Kinetics of influenza A virus infection in humans. *J. Virol.* 80, 7590–7599.

Didelot, X., Fraser, C., Gardy, J., Colijn, C., 2017. Genomic infectious disease epidemiology in partially sampled and ongoing outbreaks. *Mol. Biol. Evol.* msw075. https://doi.org/10.1093/molbev/msw275.

Smith, C.A., Yates, C.A., 2018. Spatially extended hybrid methods: a review. *J. Roy. Soc. Interface* 15.

Alizon, S., Van Baalen, M., 2008. Multiple infections, immune dynamics, and the evolution of virulence. *Am. Nat.* 172.

Feng, Z., Cen, X., Zhao, Y., Velasco-Hernandez, J.X., 2015. Coupled within-host and between-host dynamics and evolution of virulence. *Math. Biosci.* 270.

Pugliese, A., 2011. The role of host population heterogeneity in the evolution of virulence. *J. Biol. Dyn.* 5.

Hall, M., Woolhouse, M., Rambaut, A., 2015. Epidemic reconstruction in a phylogenetics framework: transmission trees as partitions of the node set. *PLOS Comput. Biol.* 11, e1004613.

Ashby, B., King, K.C., 2017. Friendly foes: the evolution of host protection by a parasite. *Evol. Lett.* 1, 211–221.

Wood, J. & Ashby, B. *Hyperparasitism and the evolution of parasite virulence.* http://biorxiv.org/lookup/doi/10.1101/2023.01.18.524570 (2023) doi:10.1101/2023.01.18.524570.