

Epidemic dynamics and host immune response: a nested approach

Alberto Gandolfi¹, Andrea Pugliese², Carmela Sinisgalli¹

¹Istituto di Analisi dei Sistemi ed Informatica “A. Ruberti” - CNR
Viale Manzoni 30, 00185 Roma, Italy
alberto.gandolfi@iasi.cnr.it, carmela.sinisgalli@iasi.cnr.it

²Department of Mathematics, University of Trento
Via Sommarive 14, 38050 Povo (Trento), Italy
pugliese@science.unitn.it

Abstract

This paper proposes an approach for building epidemiological models that incorporate the intra-host pathogen-immunity dynamics. The infected population is structured in terms of pathogen load and level of immunity, and the initial infection load may depend on the load of the individual from which the infection is acquired. In particular, we focus on the case in which the initial inoculum is taken proportional to the load of the infectant. Possible reinfections are disregarded. Such an approach is applied to formulate an epidemic model with isolation in a closed population by introducing a specific intra-host dynamics. A numerical scheme for the solution of model equations is developed, and some numerical results illustrating the role of the initial inoculum, of the isolation threshold and of the pathogen dynamics on the epidemic evolution are presented. From the simulations the distributions of latency, infectivity, and isolation times can be also derived; however the predictions of the present models differ qualitatively from those of traditional SEIHR models with distributed latency, infectivity and isolation periods.

Keywords: Epidemic modelling, Pathogen dynamics, Immune response, Structured population models.

MSC: 92D30, 92D25, 35L04.

1 Introduction

In recent years there have been several attempts to connect epidemiological aspects of infections (transmission rate, incubation and infectious periods)

to the within-host dynamics of pathogens. There are several reasons why this connection may be relevant.

On the one hand, classical epidemic models classify individuals in one of few discrete states (e.g., susceptible or infected or immune), while reality is more complex and this may be relevant, especially when considering waning immunity (Breda *et al.*, 2012). Explicit consideration of immune response seems particularly relevant for infections (such as dengue or varicella/zoster) where **the severity of the** disease depends on immune response.

Moreover, Dushoff (1996) suggested that immunological considerations predict that there may be a connection between the disease level in the population and the distribution of exposure intensities that individuals face (see Paunio *et al.* (1998) for some empirical support to the hypothesis) and that the intensity of initial exposure may have an effect on immunological outcome. In this sense, within-host dynamics cannot be decoupled from infection distribution in the population, and vice versa (see Hellriegel (2001) for a general review on the topic). However Dushoff himself resorts to a compartmental model in order to approximate such effects.

Including host immune response is also a natural way to account for host heterogeneity (Pugliese, 2011), while integrating host immune response in population models seems relevant in studies of the dynamics of macroparasites (Woolhouse, 1992), where immunity grows gradually with exposure.

We conclude this list of reason, that does not intend to be exhaustive, for integrating within-host response in epidemic models, by recalling several recent theoretical studies on the evolution of virulence (Gilchrist and Sasaki, 2002; Gilchrist and Coombs, 2006; André and Gandon, 2006; Boldin and Diekman, 2008) where virulence-transmissibility ‘trade-offs’ emerge from a mechanistic within-host models. In the so-called ‘nested’ approach (Gilchrist and Sasaki, 2002; Gilchrist and Coombs, 2006; Mideo *et al.*, 2008) an explicit (simplified) model of hosts’ immune response (such as the one by Perelson *et al.* (1993) for HIV, or one proposed in Nowak and May (2000) or Moshashemi and Levins (2001)) is used, over which epidemic dynamics is built.

Here we follow this approach, generalising the structure relatively to previous authors, but limiting ourselves to analyse the main dynamical features of the model for a single epidemic. We allow however for rather general models for within-host dynamics and infection transmission, and study how model features (especially how initial exposure is modelled) affect the overall epidemic dynamics. This approach is quite similar to that used in a recent paper by Angulo *et al.* (to appear). Other recent papers considering a model integrating within-host dynamics and epidemic transmission are due to Kostova (2007) and Martcheva and Pilyugin (2006).

The content of the present paper can be outlined as follows. In Section 2 we present a general epidemic model including within-host dynamics where the initial infection load may depend on the load of the individual from which the infection is acquired. In our model the infected population

is structured in terms of both pathogen load and level of immunity. In the subsequent two sections we consider the epidemic model in two special cases. In Section 3 the initial inoculum is independent of the pathogen load of the infectant: this case may be reduced to an epidemic model structured through the infection age and thus classical threshold results apply. In Section 4 the initial infection load is taken proportional to the load of the infectant, giving rise to a model that poses novel mathematical problems, and whose qualitative behaviour is unexplored. In Section 5 the above modelling approach is applied introducing a specific intra-host dynamics to build an epidemic model that includes isolation in a closed population. A numerical scheme for the solution of model equations is developed, and some numerical results illustrating the role of the initial inoculum, of the isolation threshold and of the pathogen dynamics on the epidemic evolution are presented. From the simulations the distributions of latency, infectivity, and isolation times are also derived; however the predictions of our model differ qualitatively from those of traditional SEIHR models with distributed latency, infectivity and isolation periods.

2 Introducing pathogen dynamics in epidemic modelling

2.1 Infected structured by pathogen load and immunity level

The main point of this manuscript is structuring the compartment of infected individuals by their pathogen load and by their level of specific immunity. Let us denote these quantities, perhaps after a suitable scaling, by x and y , respectively. The *state* of an infected individual will be identified by the pair (x, y) . The number of individuals at time t that have never been infected, and hence are free of pathogens and capable of infection (susceptible individuals), will be denoted by $S(t)$. The population of individuals that have been infected will be described by a density $i(t, x, y)$, $x > 0$, $y > 0$, such that $i(t, x, y) dx dy$ gives the number of individuals at time t having pathogen content between x and $x + dx$, and immunity level between y and $y + dy$.

In principle, one could allow the individuals that have been infected to revert to a pathogen-free state (and perhaps also with basal specific immunity) in a finite time. However, this cannot occur in the framework we are considering; hence, we restrain from further discussing it.

The within-host pathogen-immunity dynamics, assumed to be the same for each individual in the population, will be expressed by a deterministic time-invariant system that we generically describe as:

$$x' = F(x, y), \tag{1}$$

$$y' = G(x, y). \tag{2}$$

A specific example will be given below.

We require that for Eqs. (1), (2) any Cauchy problem in $\{x > 0, y > 0\}$ has a unique solution. We assume that before any infection the immunity attains the basal level $y = \bar{y} > 0$. This implies that $(0, \bar{y})$ must be an equilibrium point for the above system. Moreover, $G(x, y)$ is such that $G(x, \bar{y}) > 0$ for $x > 0$, **i.e. the immunity level can only increase upon infection**. The last property actually restricts the meaningful domain of the density i to the set $\{x > 0, y \geq \bar{y}\}$.

Beyond the equations describing within-host dynamics, there are two modelling choices that need **to be made explicit**. One can allow only for a single infection event per individual, or allow for multiple infections; the second choice is whether the initial pathogen load of an infected individual is independent of the level of infection in the infectant, or does depend on that. In both the latter cases, assuming a single infection event per individual leads to a much simpler modelling, and here we hold to this assumption; possible extensions are discussed in the Concluding Remarks.

We then assume that, once an individual is infected at some time, its internal dynamics is given by equations (1), (2), disregarding any possible reinfection. One has only to specify the mortality rate, that we write as $\mu_0 + \mu_i(x, y)$; μ_0 is the (constant) death rate of uninfected individuals, while $\mu_i \geq 0$ specifies the additional death rate caused by the infection, according to pathogen load and immunity level (indeed mortality could depend directly on the action of the pathogen, or could be mediated by the stress induced in mounting the immune response, as assumed in Gilchrist and Sasaki (2002)).

Using the standard technique for going from individual to population dynamics (Metz and Diekmann, 1986), one obtains the equation

$$\frac{\partial i}{\partial t} + \frac{\partial(F(x, y)i)}{\partial x} + \frac{\partial(G(x, y)i)}{\partial y} = -(\mu_0 + \mu_i(x, y))i(t, x, y), \quad x > 0, y > \bar{y}, \quad (3)$$

where the left hand side explicitly writes the term

$$\frac{\partial i}{\partial t} + \operatorname{div} \left(\begin{pmatrix} F \\ G \end{pmatrix} i \right).$$

One has to add to (3) the boundary condition (at $y = \bar{y}$) specifying the infection rate.

We assume that the infectiousness (which includes the contact rate, and the probability of infection upon contact) of infected individuals with state (x, y) , is given by a function $\beta(x, y)$ that reflects both biological and behavioural factors. In fact, individuals with high pathogen load are more likely to transmit the infection at each contact; on the other hand, individuals with **very high pathogen load are likely so sick as to have few or no contacts**. Moreover, when the pathogen load is very small, it is likely that no actual infection occurs, since the transmission mechanism

may be impaired and/or the first immune barrier of susceptible may block a very small inoculum (such defense might not be included in the intra-host dynamical model). Summarising, we assume that two positive values x_m and x_M exist such that:

$$\beta(x, y) = 0, \quad \text{if } 0 \leq x \leq x_m \text{ or } x \geq x_M. \quad (4)$$

The definition of the support of β , in the strip $x_m < x < x_M$, may also involve the variable y , and an example of such a case is given in Section 5, **where isolation of infected is explicitly introduced.**

Furthermore, when an infection occurs, we allow for the inoculum size to depend, in general, on the pathogen load of the infectant. Precisely, we assume that there exists a function $p(x, x_0)$ such that, if a susceptible is infected by an individual with pathogen load x , $p(x, x_0) dx_0$ is the probability that he/she will start as an infected with pathogen load in $(x_0, x_0 + dx_0)$. Generically, we need only to assume

$$\int_0^\infty p(x, x_0) dx_0 = 1,$$

although some other properties, such as $p(x, x_0) = 0$ for $x \leq x_0$, are natural. **In words, initial pathogen level in a newly infected individual cannot exceed the pathogen load of the individual from whom the infection is acquired.**

From these assumptions, assuming random contacts among all individuals, the rate of new infections at time t is given by

$$S(t)\lambda(t)$$

with

$$\lambda(t) = \int_0^\infty \int_{\bar{y}}^\infty \beta(x, y) i(t, x, y) dy dx. \quad (5)$$

Note that this formulation follows the mass-action law, while it is often considered more realistic assuming the standard incidence, i.e. dividing (5) by the total population size. This does not pose any particular difficulty and, since we are actually interested in understanding the effect of including within-host dynamics on epidemic models rather than in being realistic for any specific infection, we stick to the formulation (5).

The quantity $S(t)\lambda(t)$ gives the total rate of entrance into the infected class. In order to write the proper boundary condition of (3), we need to consider also the pathogen level at entry in the class through the function p . Remember that we assumed that the initial value of y is the constant \bar{y} , so that the boundary will be the half-line $(0, \infty) \times \bar{y}$ (or, more realistically,

a segment of it). Letting \vec{n} be the inward normal, i.e. $(0, 1)$, we can then write on the boundary:

$$\left\langle \begin{pmatrix} F \\ G \end{pmatrix} i, \vec{n} \right\rangle = G(x, \bar{y})i(t, x, \bar{y}) = S(t) \int_0^\infty \int_{\bar{y}}^\infty \beta(x', y)p(x', x)i(t, x', y) dy dx'. \quad (6)$$

Equations (3) and (6) must be complemented (beyond the initial conditions) also by an equation for the susceptibles $S(t)$. Since the aim of this research is to assess the relevance of the detailed description of within-host dynamics, we keep other aspects as simple as possible. Thus, we assume a constant inflow Λ into the susceptible population, yielding the equation

$$S'(t) = \Lambda - (\mu_0 + \lambda(t))S(t) \quad (7)$$

where $\lambda(t)$ is defined in (5).

2.2 Infected structured by age of infection and initial pathogen load

When reinfections are not allowed (as in the present paper), it appears convenient to change the variables of the internal state of infected individuals from (x, y) to ϑ , the time since infection, and x_0 , the pathogen load at infection. **This change facilitates the numerical solution of the model, also permitting the use of known analytical tools (as we will see in the next subsection).**

Indeed, since we assume that the initial immune level is fixed at \bar{y} , the pathogen load and the immune level of an individual infected time ϑ ago with an initial load x_0 can be written as $\xi(\vartheta, x_0)$ and $\eta(\vartheta, x_0)$, where ξ and η are the solutions of the Cauchy problem

$$\begin{cases} \frac{d\xi}{d\vartheta} &= F(\xi, \eta) \\ \frac{d\eta}{d\vartheta} &= G(\xi, \eta) \\ \xi(0, x_0) &= x_0 \\ \eta(0, x_0) &= \bar{y}. \end{cases} \quad (8)$$

This transformation is **one-to-one** between $\{(\vartheta, x_0), \vartheta \geq 0, x_0 > 0\}$ and a subset Ω of the positive quadrant (x, y) formed by the points that can be reached by trajectories of (1)–(2) starting from (x_0, \bar{y}) for some $x_0 > 0$. Depending on the specific system (1)–(2), Ω may be smaller than \mathbb{R}_+^2 .

Then, by setting

$$u(t, \vartheta, x_0) = i(t, \xi(\vartheta, x_0), \eta(\vartheta, x_0)) \begin{vmatrix} \xi_\vartheta & \xi_{x_0} \\ \eta_\vartheta & \eta_{x_0} \end{vmatrix}, \quad (9)$$

one can check (the determinant in (9) is needed because of the rule for the change of variables in integrals) that u satisfies the standard equation for epidemics with infection age (see, e.g. Iannelli (1994)) with the boundary condition at $\vartheta = 0$ being an **intuitively understandable** transformation of (6). Precisely, one obtains

$$\frac{\partial u(t, \vartheta, x_0)}{\partial t} + \frac{\partial u(t, \vartheta, x_0)}{\partial \vartheta} = -\tilde{\mu}(\vartheta, x_0)u(t, \vartheta, x_0) \quad (10)$$

$$u(t, 0, x_0) = S(t) \int_0^\infty \int_0^\infty \tilde{\beta}(\vartheta, x'_0) \tilde{p}(\vartheta, x'_0, x_0) u(t, \vartheta, x'_0) dx'_0 d\vartheta \quad (11)$$

$$S'(t) = \Lambda - S(t) \int_0^\infty \int_0^\infty \tilde{\beta}(\vartheta, x) u(t, \vartheta, x) dx d\vartheta - \mu_0 S(t) \quad (12)$$

where

$$\begin{aligned} \tilde{\mu}(\vartheta, x_0) &= \mu_0 + \mu_i(\xi(\vartheta, x_0), \eta(\vartheta, x_0)) \\ \tilde{\beta}(\vartheta, x_0) &= \beta(\xi(\vartheta, x_0), \eta(\vartheta, x_0)) \\ \tilde{p}(\vartheta, x'_0, x_0) &= p(\xi(\vartheta, x'_0), x_0). \end{aligned}$$

The initial conditions will be given by

$$S(0) = S_0 \geq 0, \quad u(0, \vartheta, x_0) = u_0(\vartheta, x_0) \in L^1_+((0, \infty) \times (0, \infty)).$$

2.3 A renewal equation formulation

Following the usual steps in the analysis of age-structured problems, the equations for u can be transformed into a renewal equation, albeit an infinite-dimensional one. Precisely, integrating (10) along the characteristics one obtains

$$u(t, \vartheta, x_0) = \begin{cases} u(t - \vartheta, 0, x_0) \tilde{\pi}(\vartheta, x_0) & t > \vartheta \\ u_0(\vartheta - t, x_0) \frac{\tilde{\pi}(\vartheta, x_0)}{\tilde{\pi}(\vartheta - t, x_0)} & t < \vartheta \end{cases}. \quad (13)$$

$$\text{where } \tilde{\pi}(\vartheta, x_0) = \exp\left\{-\int_0^\vartheta \tilde{\mu}(s, x_0) ds\right\}.$$

Setting $B(t, x_0) = u(t, 0, x_0)$ and substituting (13) into (11), one arrives at

$$\begin{aligned} B(t, x_0) &= S(t) \int_0^t \int_0^\infty \tilde{p}(\vartheta, x'_0, x_0) \tilde{\beta}(\vartheta, x'_0) \tilde{\pi}(\vartheta, x'_0) B(t - \vartheta, x'_0) dx'_0 d\vartheta \\ &+ S(t) \int_t^\infty \int_0^\infty \tilde{p}(\vartheta, x'_0, x_0) \tilde{\beta}(\vartheta, x'_0) \frac{\tilde{\pi}(\vartheta, x'_0)}{\tilde{\pi}(\vartheta - t, x'_0)} u_0(\vartheta - t, x'_0) dx'_0 d\vartheta. \end{aligned} \quad (14)$$

Letting

$$\begin{aligned} A(\vartheta) &: L^1(\mathbb{R}_+) \rightarrow L^1(\mathbb{R}_+) \\ (A(\vartheta)\varphi)(x_0) &= \int_0^\infty \tilde{p}(\vartheta, x'_0, x_0) \tilde{\beta}(\vartheta, x'_0) \tilde{\pi}(\vartheta, x'_0) \varphi(x'_0) dx'_0 \end{aligned}$$

(14) can be seen as the renewal equation in $L^1(\mathbb{R}_+)$:

$$B(t) = S(t) \left(\int_0^t A(\vartheta) B(t - \vartheta) d\vartheta + P(t) \right) \quad (15)$$

where $B(t)$ is seen as an element of $L^1(\mathbb{R}_+)$, $S(t)$ is a scalar obtained as the solution of (12) and $P(t)$ is the element of $L^1(\mathbb{R}_+)$ given by the second term in (14), the one involving the initial conditions at $t = 0$.

We can rewrite (12) in terms of B as

$$\begin{aligned} S'(t) &= \Lambda - S(t) \left[\int_0^t \int_0^\infty \tilde{\beta}(\vartheta, x) \tilde{\pi}(\vartheta, x) B(t - \vartheta)(x) dx d\vartheta \right. \\ &\quad \left. + \int_t^\infty \int_0^\infty \tilde{\beta}(\vartheta, x) \frac{\tilde{\pi}(\vartheta, x)}{\tilde{\pi}(\vartheta - t, x)} u_0(\vartheta - t, x) dx d\vartheta \right] - \mu_0 S(t). \end{aligned} \quad (16)$$

We do not delve here into conditions that guarantee existence and uniqueness of solutions of (15)–(16), since the thrust of this manuscript is on the modelling part. However, the use of fixed point techniques should yield existence and uniqueness results.

We end this part by noting that this setting yields a natural definition of the basic reproduction number R_0 , following the framework by Diekmann and Heesterbeek (2000). Precisely, if we consider the operator

$$\begin{aligned} K : L^1(\mathbb{R}_+) &\rightarrow L^1(\mathbb{R}_+) \\ (K\varphi)(x_0) &= \frac{\Lambda}{\mu_0} \int_0^\infty (A(\vartheta)\varphi)(x_0) d\vartheta \end{aligned} \quad (17)$$

R_0 can be obtained as the spectral radius of K .

In the following, we will consider in greater detail two cases of this general framework, in order of increasing complexity. In the first, the pathogen load at infection is chosen according to a fixed distribution (possibly concentrated in a single value \bar{x}_0). In the second, the initial pathogen load is a fixed fraction of the pathogen load of the infecting individual.

3 Infections with independent initial pathogen load

A simple case of system (10)–(12) occurs if the initial pathogen load follows a fixed probability density $q(\cdot)$, with a **bounded** support $[q_1, q_2]$, independent of the pathogen load of the infectant. To guarantee an initial inoculum never greater than the pathogen load of the infectant, we must assume $q_2 < x_m$. Then

$$\tilde{p}(\vartheta, x'_0, x_0) = p(\xi(\vartheta, x'_0), x_0) = q(x_0),$$

and we can take this term out of the integral in (11) to get

$$u(t, 0, x_0) = q(x_0) S(t) \lambda(t) \quad (18)$$

$$\text{with } \lambda(t) = \int_0^\infty \int_0^\infty \tilde{\beta}(\vartheta, x') u(t, \vartheta, x') dx' d\vartheta. \quad (19)$$

The equation for the susceptibles is still given by (7).

Finally, using (13) with (18) in the definition (19), we obtain

$$\lambda(t) = \int_0^t S(t-\vartheta)\lambda(t-\vartheta)k(\vartheta) d\vartheta + \lambda_0(t) \quad (20)$$

where

$$k(\vartheta) = \int_0^\infty q(x_0)\tilde{\beta}(\vartheta, x_0)\tilde{\pi}(\vartheta, x_0) dx_0 \quad (21)$$

and

$$\lambda_0(t) = \int_t^\infty \int_0^\infty \tilde{\beta}(\vartheta, x_0)u_0(\vartheta-t, x_0)\frac{\tilde{\pi}(\vartheta, x_0)}{\tilde{\pi}(\vartheta-t, x_0)} dx_0 d\vartheta. \quad (22)$$

One has then reduced the problem to the system of an integral equation, (20), and a differential equation, (7), for the unknown $S(t)$ and $\lambda(t)$ that needs only the initial condition $S(0) = S_0$. This system has been studied by Thieme and Castillo-Chavez (1993) in a slightly different form and letting the contact rate depend on the total population size.

In Thieme and Castillo-Chavez (1993) the well-posedness of the system (under very mild properties on the kernel $k(\cdot)$) has been proved, as well as several properties concerning its asymptotic behaviour. In our notation, Thieme and Castillo-Chavez defined R_0 as

$$R_0 = \frac{\Lambda}{\mu_0} \int_0^\infty k(\vartheta) d\vartheta \quad (23)$$

which corresponds to the general definition given in the previous Section. They showed that, if $R_0 < 1$, the infection-free equilibrium ($S \equiv \Lambda/\mu_0, \lambda \equiv 0$) is globally asymptotically stable, while, if $R_0 > 1$, $(\Lambda/\mu_0, 0)$ is unstable and there exists a unique positive equilibrium (S^*, λ^*) of the limiting system

$$\lambda(t) = \int_0^\infty S(t-\vartheta)\lambda(t-\vartheta)k(\vartheta) d\vartheta. \quad (24)$$

Furthermore, they found several sufficient conditions for the asymptotic stability of (S^*, λ^*) ; for instance, the equilibrium is always stable in case of a contact rate linear in the population size (as implicitly assumed in (20),(7)). With other choices of contact rate, they also found cases where the endemic equilibrium is unstable and attracting periodic solutions exist (see also Milner and Pugliese (1999)).

It would be interesting to know whether this is possible also when the kernel k is given by (21), and biologically reasonable assumptions are made on the infectiousness $\beta(x, y)$ and on the disease-induced mortality $\mu_i(x, y)$.

Note that system (20),(7) is the one implicitly used by Gilchrist and Sasaki (2002) with the following choice of the functions F and G

$$F(x, y) = rx - cxy, \quad G(x, y) = bxy,$$

and assuming the initial pathogen load x_0 concentrated at the single value $x_0 = \bar{x}_0$. In this case, however, we cannot really speak of a density function $i(t, x, y)$ or $u(t, \vartheta, x_0)$, but of a Dirac mass concentrated on a single trajectory of (1)–(2) when using variables (x, y) , or on the line $\{x_0 = \bar{x}_0, \vartheta \geq 0\}$ in the variables (x_0, ϑ) .

4 Infections with proportional initial pathogen load

The difference from the previous case lies in the assumption made here that the initial pathogen load of an infected individual is *proportional* to the pathogen load of the infectant. This choice is mainly motivated with reference to pathogens with aerial transmission, in which it is reasonable to suppose that the amount of pathogen particles released in the environment per unit time by an infectious individual is proportional to his pathogen burden. During the contact, a susceptible individual will “capture” a portion of this release. **Although the initial inoculum should actually depend on the contact duration, for simplicity it may be assumed that contact durations are constant and equal to an average value. Different contact durations could be taken into account in the definition of the probability density $p(x, x_0)$.** The scheme can be generalized to other mechanisms of transmission, for instance to all the cases in which infection occurs through the contact with organic fluids containing the pathogen particles.

In the mechanisms of transmission above described, we assume that the amount of pathogens of the infectant is not altered by the infective encounter. This is strictly true when the infection exploits an intrinsic pathogen loss independent of the individual-to-individual interaction.

Precisely, we assume that the inoculum received, x_0 , is equal to ρx , where x is the pathogen burden of the infectant and $\rho \ll 1$ is a fixed constant. Equations (3), (6) will still hold, except that $p(x, x_0)$ will not be a function, but a Dirac mass concentrated at ρx .

Instead of setting the problem for measure-valued functions, we arrive at a proper formulation of this problem, in terms of density functions, via a limiting procedure. First of all, we assume that

$$p(x, x_0) = \begin{cases} \frac{1}{\Delta x} & \rho x \leq x_0 < \rho x + \Delta x, \\ 0 & \text{otherwise,} \end{cases} \quad (25)$$

and eventually we will let Δx go to 0.

In this case, equation (6) can be explicitly written as

$$G(x, \bar{y})i(t, x, \bar{y}) = \frac{S(t)}{\Delta x} \int_{\frac{x-\Delta x}{\rho}}^{\frac{x}{\rho}} \int_{\bar{y}}^{\infty} \beta(x', y)i(t, x', y) dy dx'. \quad (26)$$

Letting now Δx go to 0, the condition becomes

$$G(x, \bar{y})i(t, x, \bar{y}) = \frac{S(t)}{\rho} \int_{\bar{y}}^{\infty} \beta\left(\frac{x}{\rho}, y\right) i\left(t, \frac{x}{\rho}, y\right) dy. \quad (27)$$

The system to be solved will then consist of (3), (27) and (7) plus the initial conditions.

We wish now to write the system in the coordinates ϑ and x_0 as in (10), (11); the computations are more involved, but the final result appears more suitable for the analysis and simulation.

We specialize to the case where the solutions of (1), (2) have properties that we consider **biologically significant**: we assume that there exist two threshold quantities x_1 and x_2 with $0 < x_1 < x_2 < +\infty$ such that:

- if $x_1 < x_0 < x_2$, then $\xi(\vartheta, x_0)$ monotonically increases to a maximum value, and thereafter monotonically decreases towards 0 (normal acute infection). Let $M(x_0) > x_0$ be that maximum, and $T(x_0)$ the age of infection at which $x = M(x_0)$ is attained starting from $x = x_0$;
- if $x_0 > x_2$, $\xi(\vartheta, x_0)$ grows monotonically to $+\infty$ (uncontrolled infection);
- if $x_0 < x_1$, $\xi(\vartheta, x_0)$ decreases monotonically to 0 (infection controlled by innate immunity).

Then, we can define the time τ_1 to reach a certain level x while the pathogen load is increasing, starting from x_0 :

- for $x_1 < x_0 < x_2$

$$\tau_1(x, x_0) = \inf\{\vartheta \geq 0 : \xi(\vartheta, x_0) = x\}, \quad x_0 \leq x \leq M(x_0),$$

- for $x_0 > x_2$

$$\tau_1(x, x_0) = \inf\{\vartheta \geq 0 : \xi(\vartheta, x_0) = x\}, \quad x \geq x_0,$$

and the time τ_2 to reach x while pathogens are decreasing:

- for $0 < x_0 < x_1$

$$\tau_2(x, x_0) = \inf\{\vartheta \geq 0 : \xi(\vartheta, x_0) = x\}, \quad 0 < x \leq x_0,$$

- for $x_1 < x_0 < x_2$

$$\tau_2(x, x_0) = \inf\{\vartheta > T(x_0) : \xi(\vartheta, x_0) = x\}, \quad 0 < x < M(x_0).$$

The functions $\tau_1(x, x_0)$ and $\tau_2(x, x_0)$ are continuous with respect to x and differentiable in their definition domains. It will be convenient to conventionally set $\tau_2(x, x_0) = +\infty$ for $x_0 > x_2$, $0 < x < +\infty$.

For further use, we observe that for the function τ_1 and τ_2 we have

$$\frac{\partial \tau_i}{\partial x}(x, x_0) = \frac{1}{F(x, \eta(\tau_i(x, x_0), x_0))}, \quad i = 1, 2, \quad (28)$$

as it can be seen through the implicit function theorem. Moreover, we have $\frac{\partial \tau_1}{\partial x} > 0$ and $\frac{\partial \tau_2}{\partial x} < 0$.

Since the solutions of (1), (2) are unique, and then the trajectories $(\xi(\vartheta, x_0), \eta(\vartheta, x_0))$ cannot intersect, we can note that $M(x_0)$ is necessarily increasing with $x_0 \in (x_1, x_2)$. So, in addition, we can define $h(x)$ as the inverse function of $M(\cdot)$, i.e.:

$$h(x) = M^{-1}(x), \quad M(x_1^+) < x < +\infty. \quad (29)$$

$M(x_1^+) = \lim_{x_0 \rightarrow x_1^+} M(x_0)$ may be larger than x_1 ; this would be the case when the presence of an interior saddle point divides the trajectories starting with $x_0 \leq x_1$ from those with $x_0 > x_1$ (see Pugliese and Gandolfi (2008) and Fig. 2). **Being $h(x)$ the initial inoculum such that $M(h(x)) = x$, from the very definition of the function M , one may also note that $h(x) < x$.**

Table 1 summarizes the meaning of some characteristic quantities here introduced.

Table 1: Summary of some useful quantities

Name	Meaning
x_m	If $x \leq x_m$, $\beta(x, y) = 0$
x_M	If $x \geq x_M$, $\beta(x, y) = 0$
x_1	If $x_0 < x_1$, the pathogen load decays monotonically to zero
x_2	If $x_0 > x_2$, the pathogen load grows monotonically to $+\infty$
$\tau_1(x, x_0)$	Age of infection at which the pathogen load reaches increasing the value x , when the initial inoculum is x_0
$\tau_2(x, x_0)$	Age of infection at which the pathogen load reaches decreasing the value x , when the initial inoculum is x_0
$M(x_0)$	Maximum value of the pathogen load, when the initial inoculum is x_0 (for $x_1 < x_0 < x_2$)
$T(x_0)$	Age of infection at which $x = M(x_0)$ is attained
$h(x)$	Initial inoculum such that $M(h(x)) = x$

In view of assumption (4), it is sufficient to define the boundary condition (11) for $x_0 > \rho x_m$. We show here the case where $x_m < x_1$, which is more involved. If $x_m \geq x_1$ (probably, a more realistic case) the boundary condition can be obtained by simple changes as shown below.

Using (25) in (11), and passing to the limit for Δx tending to zero, we get the following (see Appendix A):

i) $x_m < \frac{x_0}{\rho} < x_1$

$$u(t, 0, x_0) = \frac{S(t)}{\rho} \left[- \int_{\frac{x_0}{\rho}}^{x_2} K_2(t, x_0, x'_0) dx'_0 \right], \quad (30)$$

ii) $x_1 < \frac{x_0}{\rho} < x_2$

$$u(t, 0, x_0) = \frac{S(t)}{\rho} \left[\int_{h(\frac{x_0}{\rho})}^{\frac{x_0}{\rho}} K_1(t, x_0, x'_0) dx'_0 - \int_{h(\frac{x_0}{\rho})}^{x_2} K_2(t, x_0, x'_0) dx'_0 \right], \quad (31)$$

iii) $\frac{x_0}{\rho} > x_2$

$$u(t, 0, x_0) = \frac{S(t)}{\rho} \left[\int_{h(\frac{x_0}{\rho})}^{\frac{x_0}{\rho}} K_1(t, x_0, x'_0) dx'_0 - \int_{h(\frac{x_0}{\rho})}^{x_2} K_2(t, x_0, x'_0) dx'_0 \right]. \quad (32)$$

where

$$K_i(t, x_0, x'_0) = \frac{\tilde{\beta}(\tau_i(\frac{x_0}{\rho}, x'_0), x'_0) u(t, \tau_i(\frac{x_0}{\rho}, x'_0), x'_0)}{F(\frac{x_0}{\rho}, \eta(\tau_i(\frac{x_0}{\rho}, x'_0), x'_0))}, \quad i = 1, 2. \quad (33)$$

The condition given by (30)-(32) substitutes the more general boundary condition specified by (11), yielding the system to be solved.

If $x_m \geq x_1$, then the boundary conditions consist of ii) for $\rho x_m < x_0 < x_2$, and of iii).

One may **restrict** (15) to this special case, but there does not seem to be a particular insight to be gained doing so. Instead, in the next section a numerical method for this case is developed and applied to a model example where β is either 0 or a positive constant.

In this special case, R_0 , defined as the spectral radius of K given by (17), may be simpler to compute as K becomes, using the definitions of \tilde{p} and $\tilde{\beta}$,

$$\begin{aligned} K &: L^1(\mathbb{R}_+) \rightarrow L^1(\mathbb{R}_+) \\ (K\varphi)(x_0) &= \int_{h(\frac{x_0}{\rho})}^{\frac{x_0}{\rho}} \beta\left(\frac{x_0}{\rho}, \eta\left(\tau_1\left(\frac{x_0}{\rho}, x'_0\right), x'_0\right)\right) \tilde{\pi}\left(\tau_1\left(\frac{x_0}{\rho}, x'_0\right), x'_0\right) \varphi(x'_0) dx'_0 \\ &+ \int_{\frac{x_0}{\rho}}^{x_2} \beta\left(\frac{x_0}{\rho}, \eta\left(\tau_2\left(\frac{x_0}{\rho}, x'_0\right), x'_0\right)\right) \tilde{\pi}\left(\tau_2\left(\frac{x_0}{\rho}, x'_0\right), x'_0\right) \varphi(x'_0) dx'_0. \end{aligned} \quad (34)$$

In (34) we take the second integral to be 0 if $x_0 > \rho x_2$, and the first integral to be 0 if $x_0 \leq \rho x_1$ so that $h(x_0/\rho)$ is not defined. Finally if $x_1 < x_0/\rho \leq M(x_1^+)$ (in which case too h was not defined through (29)), we take $h(x_0/\rho) = x_1$.

Note that (34) can be further simplified if, as it may be reasonable, β does not depend on its second argument (host's immune response).

Beyond the simpler expression, one can gather some properties of K from (34). First of all, remembering that it was assumed $\beta(x, y) = 0$ if $x \geq x_M$, the support of $K\varphi$ is contained in $[0, \rho x_M]$ for each φ .

Second, assume that the internal dynamics were such that $\rho M(x_0) < x_0$ for each $x_0 < \rho x_M$; then if the maximum of the support of φ were equal to L , a simple computation shows that the maximum of the support of $K\varphi$ would be $\rho M(L) < L$. It follows that K has no positive eigenvectors, hence $R_0 = 0$, independently of the transmission rate β . A natural assumption for a reasonable model is then $M(x_0) > x_0/\rho$, at least for a range of initial values x_0 .

5 An epidemic model with isolation

5.1 The model and a numerical scheme

In this section, we illustrate an application of the approach outlined in the previous Section, giving an example of the epidemic model built on the within-host pathogen dynamics described in Pugliese and Gandolfi (2008).

Precisely, equations (1)–(2) become

$$\begin{cases} \dot{x} &= \alpha x - \frac{xy}{1 + \beta_s x} - \frac{mx}{1 + \beta_u x} \\ \dot{y} &= \frac{xy}{1 + \gamma x} - y + \eta \end{cases} \quad (35)$$

where x and y represent (after appropriate scaling) pathogen load and specific immune response, respectively, while time is measured in terms of loss of immune response (i.e., it declines at rate 1 in these units).

In particular, we assume parameter values such that $\alpha < \eta + m$, $\alpha > 1/\gamma - 1$, $\beta_s > 0$, and such that two positive equilibria are present, a saddle point and an unstable focus. In this case, the pathogen-free state $(0, \bar{y})$, with $\bar{y} = \eta$, will be locally asymptotically stable, and unbounded trajectories will exist (see Pugliese and Gandolfi (2008) and d'Onofrio (2010)).

Thus the behaviour of the within-host model is exactly of the type assumed in the previous Section, although x_1 and x_2 cannot be computed explicitly.

Figure 1 shows the time course of the host pathogen load $x(t)$ for two selected sets of parameter values and for different values of the initial condition $x(0) = x_0$, keeping $y(0) = \bar{y}$,

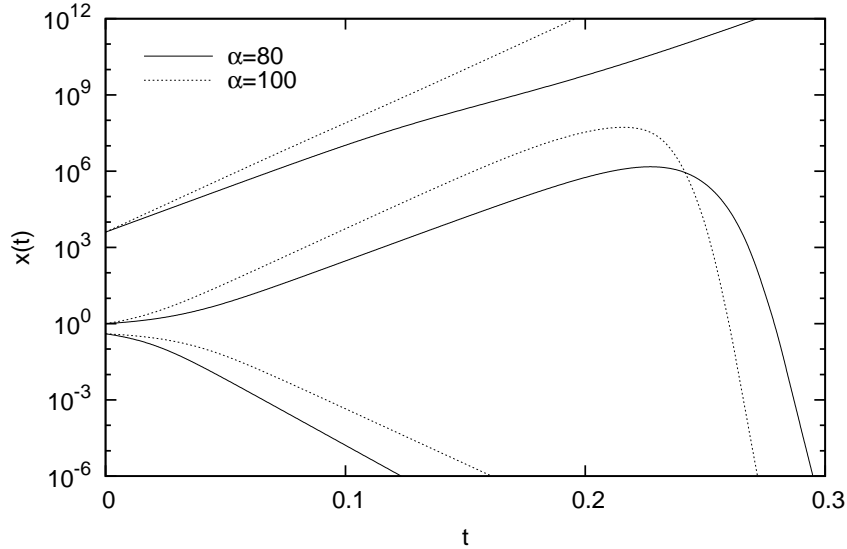


Figure 1: Time course of the pathogen load $x(t)$ for different values of x_0 . $\alpha = 80$ (solid lines), $\alpha = 100$ (dotted lines). Other parameters: $\eta = 0.05$, $\gamma = 0.02$, $m = 200$, $\beta_s = 10^{-8}$, $\beta_u = 2$.

Figure 2 depicts in the plane $(\ln x, \ln y)$ a set of trajectories starting from different initial points (x_0, \bar{y}) , namely the evolution of the infection in individuals with initial immunity at the basal level and infected with different pathogen amounts. We can note the presence of the critical values x_1 , such that for $x_0 < x_1$ the infection does not develop and the pathogens decrease to zero, and x_2 , such that for $x_0 \geq x_2$ the infection cannot be controlled by the immune system and the pathogen load goes to infinity. Note that because of the existence of the unstable focus, as x_0 tends to x_1^+ the trajectories accumulate and we have $M(x_1^+) > x_1$. Since for $x_0 < x_1$ the pathogen decrease to zero without an appreciable increase of the immunity, we disregard infections with initial load $x_0 < x_1$ by setting $x_m = x_1/\rho$ so that $\beta(x, y) = 0$ in the set

$$\Omega_1 = \{(x, y) : 0 \leq x \leq x_1/\rho, y \geq \bar{y}\}.$$

In the present model, we explicitly consider that the infected individuals can be isolated from further contacts during the time-course of the disease, because of home care or hospitalization (without, however, any specific anti-pathogen therapy). We assume that an individual enters the *isolation* state when its pathogen load reaches the value x^* (or is infected with an initial pathogen load greater than x^*), and is released from such a state when its pathogen load is lowered to the value x^{**} , $x^{**} < x^*$. Thus we define Ω_2 as the image of the trajectories between the times $\tau_1(x^*, x_0)$ and $\tau_2(x^{**}, x_0)$. The isolated subjects will be then the individuals with (x, y) in the set Ω_2 ;

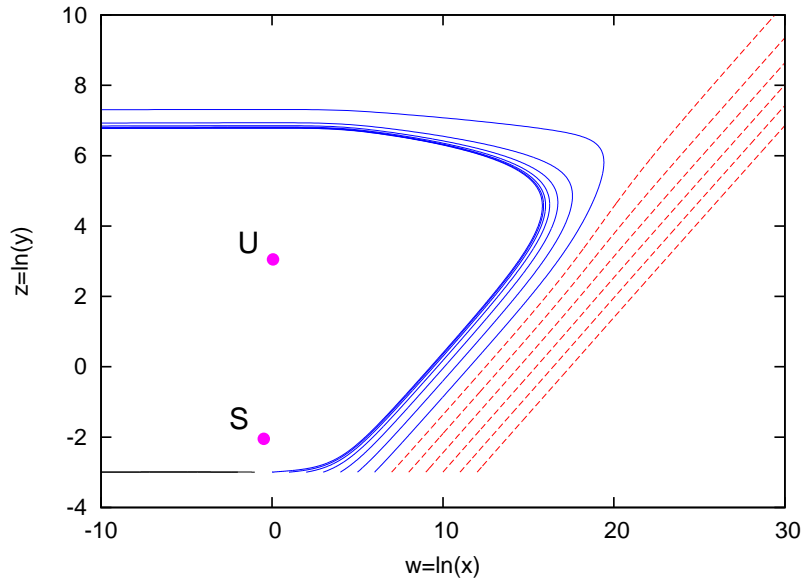


Figure 2: Trajectories in the plane $(\ln x, \ln y)$ all starting from basal-level immunity and different initial pathogen load. Parameters as in Fig. 1 with $\alpha = 90$. The positive equilibrium points S (saddle point) and U (unstable equilibrium) are also reported. Critical values: $\ln x_1 \approx -0.48$, $\ln x_2 \approx 6.32$. Dashed trajectories are diverging, whereas along the solid trajectories the pathogen load eventually decays to zero (monotonically or not).

consequently we take $\beta = 0$ in Ω_2 . Summarising we assume the expression:

$$\beta(x, y) = \begin{cases} 0 & \text{in } \Omega_1 \cup \Omega_2 \\ \beta_0 & \text{otherwise,} \end{cases} \quad (36)$$

with $\beta_0 > 0$ constant for simplicity. We assume $\rho x^* > x_1$, otherwise β would be identically zero and the epidemics could not develop.

As mentioned above, the use of the age since infection, ϑ , appears convenient. For numerical purposes, since the variables x and y typically span several orders of magnitudes over the course of an infection, it is also convenient to introduce the variables $w = \ln x$, $z = \ln y$, and

$$w_0 = \ln x_0, \quad (37)$$

and consider as new state variables the pair (ϑ, w_0) . Given (w_0, ϑ) , the corresponding pair (x, y) will be given by $x = \xi(\vartheta, e^{w_0})$, $y = \eta(\vartheta, e^{w_0})$. Conversely, given (x, y) in Ω , there will be a unique trajectory passing for such a point and this trajectory will identify (ϑ, w_0) .

We then define the density $\hat{u}(t, \vartheta, w_0) = u(t, \vartheta, e^{w_0})e^{w_0}$: hence the number of individuals whose initial (logarithmic) pathogen load was between w_0

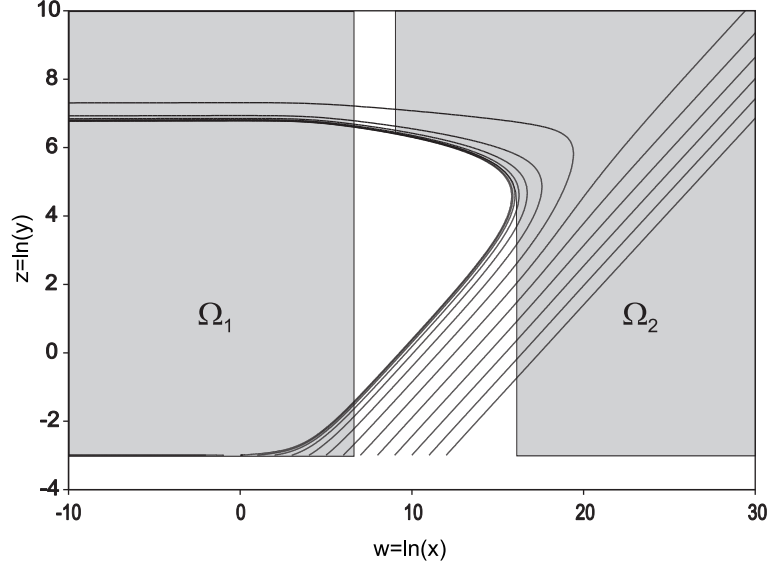


Figure 3: Regions Ω_1 and Ω_2 in the plane $(\ln x, \ln y)$ where the infection rate β is zero. $\ln x_m = \ln x_1 + \ln 10^3 = -0.48 + \ln 10^3$, $\ln x^* = 16$ and $\ln x^{**} = 9$.

and $w_0 + dw_0$, and that have been infected between ϑ and $\vartheta + d\vartheta$ time before is given by $\hat{u}(t, \vartheta, w_0)dw_0d\vartheta$.

From (10) one immediately obtains

$$\frac{\partial \hat{u}}{\partial t} + \frac{\partial \hat{u}}{\partial \vartheta} = -\hat{\mu}(\vartheta, w_0)\hat{u}(t, \vartheta, w_0), \quad w_0 \in \mathbb{R}, \vartheta > 0, \quad (38)$$

where:

$$\hat{\mu}(\vartheta, w_0) = \tilde{\mu}(\vartheta, e^{w_0}).$$

Since from (36) the support of β lies in the strip $x_m < x < x^*$, with $x_m = x_1/\rho$, the boundary condition of \hat{u} may be different from zero only for $w_m - \Delta < w_0 < w^* - \Delta$, where $\Delta = -\ln(\rho)$, $w_m = \ln x_m$, and $w^* = \ln x^*$ (without loss of generality we take $x^* > x_2$).

From the expressions (31)-(32) we can derive:

- for $w_m < w_0 + \Delta < w_2$

$$\begin{aligned} \hat{u}(t, 0, w_0) &= S(t)e^{w_0+\Delta} \left[\int_{\hat{h}(w_0+\Delta)}^{w_0+\Delta} \hat{K}_1(t, w_0, w'_0) dw'_0 \right. \\ &\quad \left. - \int_{\hat{h}(w_0+\Delta)}^{w_2} \hat{K}_2(t, w_0, w'_0) dw'_0 \right], \end{aligned} \quad (39)$$

- for $w_2 < w_0 + \Delta < w^*$

$$\begin{aligned}\hat{u}(t, 0, w_0) &= S(t)e^{w_0+\Delta} \left[\int_{h(w_0+\Delta)}^{w_0+\Delta} \hat{K}_1(t, w_0, w'_0) dw'_0 \right. \\ &\quad \left. - \int_{h(w_0+\Delta)}^{w_2} \hat{K}_2(t, w_0, w'_0) dw'_0 \right],\end{aligned}\quad (40)$$

where $w_2 = \ln x_2$, $\hat{h}(w) = \ln(h(e^w))$, and, for $i = 1, 2$,

$$\hat{K}_i(t, w_0, w'_0) = \frac{\tilde{\beta}(\hat{\tau}_i(w_0 + \Delta, w'_0), e^{w'_0}) \hat{u}(t, \hat{\tau}_i(w_0 + \Delta, w'_0), w'_0)}{F(e^{w_0+\Delta}, \eta(\hat{\tau}_i(w_0 + \Delta, w'_0), e^{w'_0}))}, \quad (41)$$

$$\hat{\tau}_i(w, w_0) = \tau_i(e^w, e^{w_0}). \quad (42)$$

The number of susceptibles $S(t)$ is given by Eq. (7) with:

$$\lambda(t) = \int_{-\infty}^{+\infty} \int_0^{+\infty} \tilde{\beta}(\vartheta, e^{w_0}) \hat{u}(t, \vartheta, w_0) d\vartheta dw_0. \quad (43)$$

In order to proceed with the numerical approximation of system (38)–(40), we choose a grid of points $w_m = w_0^0 < w_0^1 < \dots < w_0^M$ and $0 = \vartheta^0 < \vartheta^1 < \dots < \vartheta^N$, and aim at approximating

$$\int_{w_0^{i-1}}^{w_0^i} \hat{u}(t, 0, w_0) dw_0.$$

Through several intermediate steps, shown in Appendix B, one arrives at

Proposition 1

$$\int_{w_0^{i-1}}^{w_0^i} \hat{u}(t, 0, w_0) dw_0 = \int_0^\infty \int_{\mathcal{A}_i^\vartheta} \tilde{\beta}(\vartheta, e^{w'_0}) \hat{u}(t, \vartheta, w'_0) dw'_0 d\vartheta \quad (44)$$

with

$$\mathcal{A}_i^\vartheta = \left\{ w'_0 : w_0^{i-1} \leq \log(\xi(\vartheta, e^{w'_0})) - \Delta \leq w_0^i \right\}.$$

Note that equation (44) is very intuitive, much more than the expressions (30)–(32) from which it has been derived. It seems likely that setting the equations in terms of measures (instead of densities) could lead to an equally intuitive equation without special assumptions on the system (1)–(2) at the price, however, of introducing more advanced mathematical concepts (see e.g. Gwiazda *et al.* (2010)).

The numerical method, briefly described in Appendix C, is developed on the basis of (44). The method can be considered an adaptation of the so-called escalator boxcar train (De Roos, 1988; Brännström *et al.*, 2013) introduced for physiologically structured population models.

5.2 Numerical results

With the aim of highlighting the role of the initial inoculum, of the isolation threshold, and of the pathogen dynamics on the epidemic evolution, we have simplified the dynamics of susceptibles by taking $\Lambda = 0$ and $\mu_0 = 0$. Moreover, we assumed $\hat{\mu}(\vartheta, w_0) = 0$. We thus consider a closed population without vital dynamics, and no deaths due to the infection. This is an appropriate model for a single epidemic wave.

The density $\hat{u}(t, \vartheta, w_0)$ was approximated on the domain $D = [w_{in}, w_{fin}] \times [0, \vartheta_{max}]$ by means of the numerical scheme described in Appendix C. In the following, unless otherwise stated, we set: $w_{in} = -1$, $w_{fin} = 12$, $\vartheta_{max} = 5$, $w_m = -0.48 + \Delta$, $w^* = 16$, $w^{**} = 9$. Note that, according to the chosen shape of β , infections with initial pathogen load smaller than -0.48 and larger than $w^* - \Delta$ are not possible, so that the chosen values of w_{in} , w_{fin} guarantee the correct computation of the evolution for a reasonable range of Δ (we always take $\Delta > 10$ assuming that the initial inoculum will be very small compared to the pathogen load of an acute infection). We assumed $S(0) = 1000$, whereas to simulate the initial presence of a concentrated sub-population of moderately infected subjects, the initial condition for \hat{u} was taken constant in the rectangle $[0.8, 1.2] \times [0.005, 0.045]$ with integral equal to 1, and zero otherwise.

To trace the evolution of the infection in the population, we computed the following integrals:

$$\begin{aligned} I(t) &= \iint_D \hat{u}(t, \vartheta, w_0) d\vartheta dw_0, \\ I_{\bar{w}}(t) &= \iint_{D_{\bar{w}}} \hat{u}(t, \vartheta, w_0) d\vartheta dw_0, \\ H(t) &= \iint_{\widehat{\Omega}_2} \hat{u}(t, \vartheta, w_0) d\vartheta dw_0, \\ R(t) &= \iint_R \hat{u}(t, \vartheta, w_0) d\vartheta dw_0, \end{aligned}$$

where

$$\begin{aligned} D_{\bar{w}} &= \{(w_0, \vartheta) \in D : \ln \xi(\vartheta, e^{w_0}) \geq \bar{w}\}, \\ \widehat{\Omega}_2 &= \{(w_0, \vartheta) \in D : (\xi(\vartheta, e^{w_0}), \eta(\vartheta, e^{w_0})) \in \Omega_2\}, \\ R &= \{(w_0, \vartheta) \in D : \xi(\vartheta, e^{w_0}) \leq x_1\}. \end{aligned}$$

The quantity $I(t)$ gives the total number of individuals that have been infected within time t , $I_{\bar{w}}(t)$ the number of infected with a pathogen burden above the threshold $e^{\bar{w}}$, here set to $\bar{w} = 12$, $H(t)$ the number of isolated subjects, and $R(t)$ the number of recovered. The number of recovered *and* immune individuals might be computed by integrating \hat{u} over the (w_0, ϑ) domain such that $\xi(\vartheta, e^{w_0}) \leq x_1$ and $\eta(\vartheta, e^{w_0}) \geq y^*$, with y^* some suitable

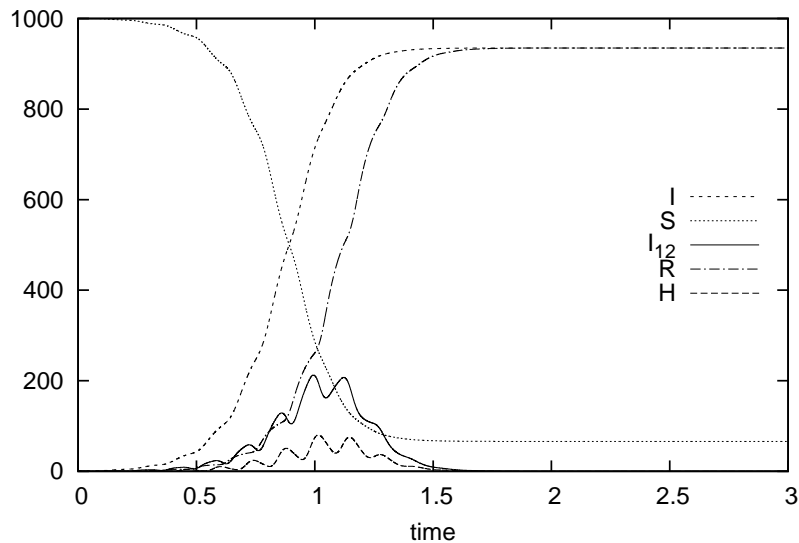


Figure 4: Time course of the epidemic for $\Delta = 11.51$, $\beta_0 = 0.035$, $w^* = 16$, $w^{**} = 9$. Other parameters as in Fig. 1, $\alpha = 90$.

threshold. Note that, according to our modelling, the subpopulations $I_{\bar{w}}$, H , R at time t are *all* included in the population $I(t)$.

Figure 4 shows the simulated evolution of the epidemic when $\Delta = 11.51$ (i.e. $\rho = 10^{-5}$) and $\beta_0 = 0.035$. In this case, the number of susceptibles tends to about 5% of the initial value, and both the numbers of subjects infected with $w > 12$ and of isolated individuals reach a peak and then decline with some oscillations. All the infected become eventually recovered. Since $\Lambda = \mu_0 = \mu_i = 0$, we have $S(t) + I(t) = \text{const.}$

A peculiar feature of this model is to account separately for the quantity of pathogens transmitted at the act of infection, and for the rate of infective encounters. Figure 5 depicts the epidemics evolution for $\Delta = 13.82$, keeping $\beta_0 = 0.035$. The comparison to Fig. 4 shows that decreasing the fraction of transferred pathogens (i.e. increasing Δ), the epidemics develops more slowly, the peak value of I_{12} decreases, the isolated population almost vanishes, while the susceptibles tend asymptotically to a greater value. Moreover, we can observe that the typical oscillations in the subpopulation evolutions tend to be more marked. All these aspects may be explained taking into account that in the (w, z) plane the window of infective individuals narrows increasing Δ (see Fig. 3), and that the initial condition for \hat{u} is here assumed to have a small support. Indeed, when the window gets narrow, the new infections are “synchronized” and rise in waves, with a period approximately given by the time necessary to reach the region $\beta > 0$ along the trajectories starting from x_0 close to x_1 (and $y = \bar{y}$).

The asymptotic fraction of susceptibles in the population will increase,

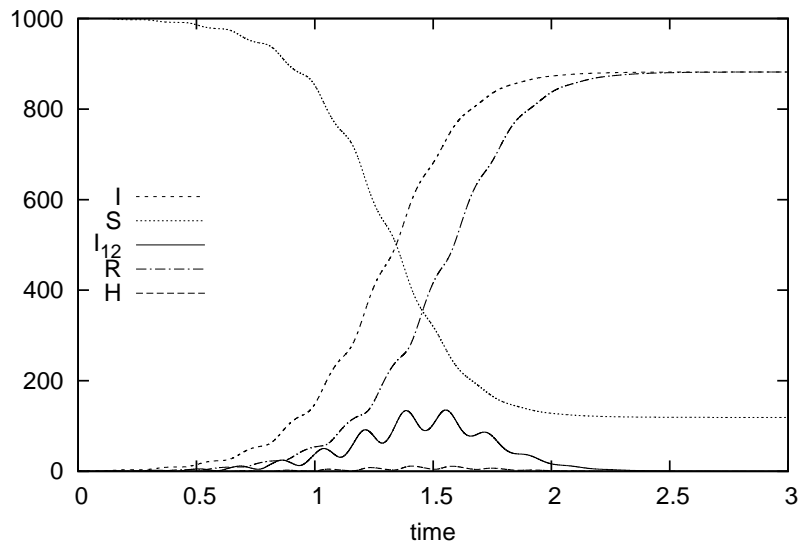


Figure 5: Time course of the epidemic for $\Delta = 13.82$. Other parameters as in Fig. 4.

as expected, even lowering the value of β_0 keeping Δ constant (data not shown). In the present model, such a fraction is indeed modulated by both the parameters Δ and β_0 (as well as by the parameter w^* characterizing the isolation). Figure 6 shows the effect of changing the isolation threshold, w^* . Including also $w^* = 16$ in the comparison, we observe that as w^* decreases (earlier isolation of infected individuals) the infection not only reduces markedly, but its progression slows down. For $w^* = 15$ or 16 no individual was affected by an infection whose pathogen load diverges. However this case can be achieved if $w^* - \Delta > \ln x_2$. The lower panel of Fig. 6 when w^* is increased to 20 shows an example of such an evolution where not all the infected eventually recover.

It can be of interest to examine the influence of stronger pathogen virulence on the epidemic evolution. In Figure 7 all the parameters are as in Fig. 4 except the pathogen replication rate $\alpha = 100$. Note that the isolated subpopulation becomes larger than in Fig. 4, and a substantial portion of the isolated subjects does not ever recover. Moreover, the fraction that remains susceptible is higher than in Fig. 4, possibly because of the larger fraction of infected that remain isolated.

The present model allows also the computation of the distributions of times of epidemiological interest, such as the latency time, the infectivity time and the time of isolation. For each individual receiving the initial inoculum $w_0 > \ln x_1$, the ages at which the individual enters and leaves the region of infectivity ($\beta > 0$), and the ages to enter and leave the region of isolation, can be examined. Of course, some of these ages may be not

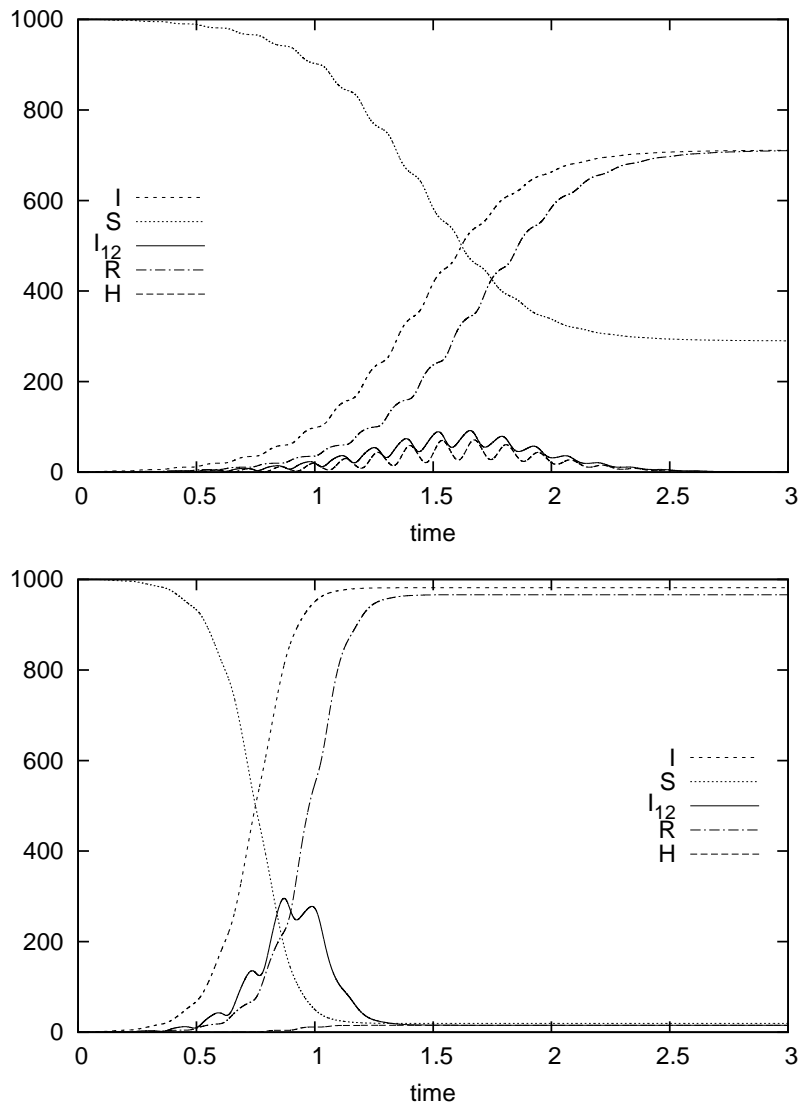


Figure 6: Effect of changing w^* . Upper panel $w^* = 15$, lower panel $w^* = 20$. Other parameters as in Fig. 4.

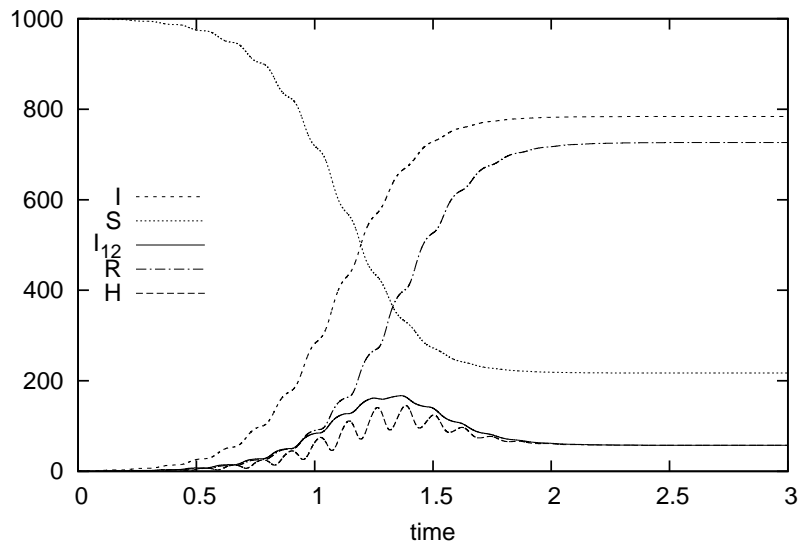


Figure 7: Time course of the epidemic for $\Delta = 11.51$ and $\alpha = 100$. Other parameters as in Fig. 4.

defined or be equal to $+\infty$ for particular w_0 values. To each w_0 we can associate (when it exists) the time needed to enter the infectivity region, i.e. the time required to reach $x = x_m$ (latency time), the time interval spent in the infectivity region (infectivity time), and the time interval during which the individual is isolated (isolation time). The distribution p_{w_0} of the initial inoculum in $[0, T]$ can be computed as

$$p_{w_0}(w_0) = \frac{Q(w_0)}{\int_{w_{in}}^{w_{fin}} Q(w_0) dw_0},$$

where:

$$Q(w_0) = \int_0^T \hat{u}(t, 0, w_0) dt.$$

By means of the density p_{w_0} , the distributions of the latency time, infectivity time and isolation time can be obtained. Figure 8, upper panel, reports such distributions for the reference parameters of Fig. 4. It can be noted that the infectivity time histogram has bimodal shape, whereas the latency time has a broader distribution skewed to the right. Concerning the isolation time, the peak at $t = 0$ means that a substantial fraction ($\simeq 40\%$) of the new infections never attains the isolation threshold. This fraction corresponds to trajectories that stay in the infectivity region for a relatively long time, bringing about the second peak of the infectivity time histogram. When the isolation threshold is lowered to $w^* = 15$, all the new infections give rise to trajectories entering isolation, and the histogram of the infectivity time is unimodal (see Figure 8, lower panel).

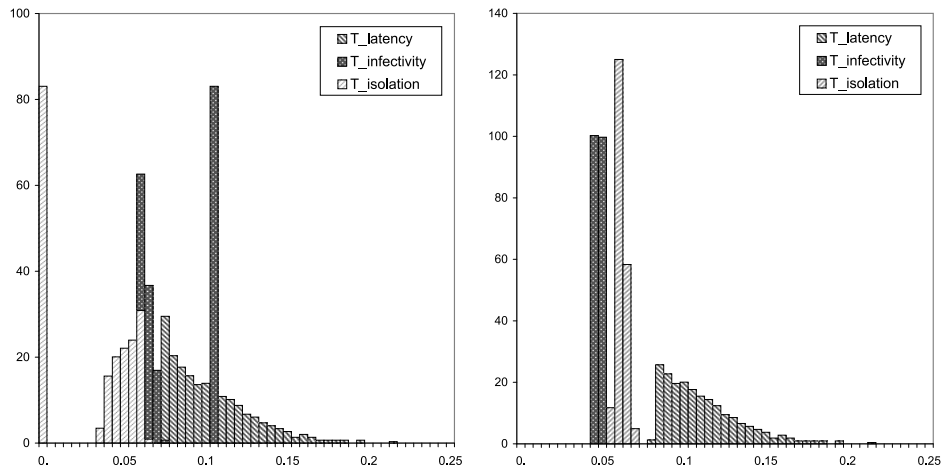


Figure 8: Histograms of latency, infectivity and isolation times. Parameter values as in Fig. 4, left panel. Parameter values as in Fig. 4 except $w^* = 15$, right panel. $T = 3$.

Note that the dynamical behaviour of our model is quite different from the behaviour of a standard infection-age model in which the same time distributions are assumed. As evidence for this fact, we show a simulation of an SEIHR model with age of infection where distributions of the latency, infectivity and isolation periods are given by the left panel of Fig. 8, while infectiousness is constant during the infectious period. The result, shown in Fig. 9, shows a different pattern with a single epidemic peak, lacking the repeated waves, corresponding roughly to the generations of infected individuals, displayed by the present model.

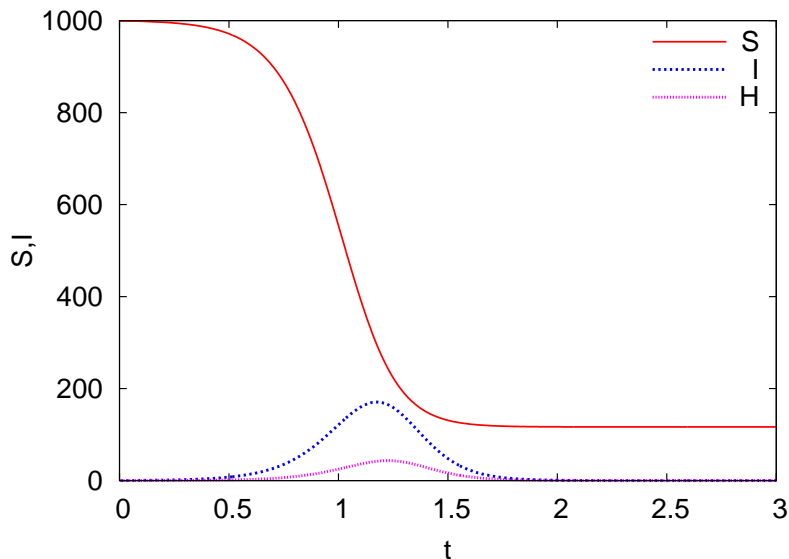


Figure 9: Numerical solution of an SEIHR model with distributions of the stage lengths given by the left panel of Fig. 8, and a fixed contact rate adjusted to yield a value of R_0 similar to that in Fig. 4.

6 Concluding remarks

In this paper we have developed a modelling approach to couple within-host and between-host dynamics. The approach follows the methods used for physiologically structured populations, starting from a PDE formulation and arriving at an abstract renewal equation.

While the steps performed in this paper show a possible way to prove the well-posedness of the resulting problem, a rigorous proof of existence, uniqueness and regularity of solutions is beyond the scope of the present paper, and will be the subject of future work.

When the initial size of the infecting inoculum is independent of the state of the infectant, the resulting system can be written as an epidemic model structured in terms of age since infection, and thus fits in a well-known class of models. Still, the within-host model can be used to derive how infectiousness and survival probability depend on age since infection on the basis of fundamental mechanisms acting in the interactions between pathogens and host immune response. For instance, one can see the effect of pathogen replication rate on the overall epidemic dynamics, and this has been used extensively when studying virulence evolution (Gilchrist and Sasaki, 2002; Gilchrist and Coombs, 2006; André and Gandon, 2006).

On the other hand, when the initial inoculum (and thus the future course of the infection) depends on the current state of the infectant, as in the model

with proportional inoculum size, it is not possible to decouple within-host from population dynamics, and it is necessary to proceed with both at the same time. The resulting model often yields an epidemic curve superimposed with several smaller waves (see Figs. 4–7) in contrast with classical epidemic models (Fig. 9) in which incidence curve is generally unimodal. Smaller waves in incidence curves can emerge in infection-age epidemic models if the infectiousness period is very narrow and not much variable among infected individuals; this is the usual explanation for the profiles sometimes observed in emerging epidemics, where subsequent generations of infected can be recognized in data. In the present model, subsequent generations of infected are recognizable, but the generation time (hence, the interval between waves) changes during the epidemic course, giving rise to an overall distribution of latent and infectious periods which is not so narrow. Indeed, we have no evidence that such a mechanism may operate in reality, but may be worth investigating.

The model of Section 5 includes isolation of infected individuals. This feature has been introduced either to show the flexibility of our modelling approach, and to limit the possible size of the initial inoculum, allowing an accurate numerical simulation. Note that under the condition $\rho M(x_0) > x_0$ and in the absence of isolation (and of the limitation of the support of β expressed by (4)), the initial inoculum size could unboundedly increase. The effect of reducing the value of the isolation threshold appears remarkable: the infection substantially reduces and slows down. On the contrary, high and increasing threshold values produce marginally different behaviours, suggesting that the asymptotic fraction of susceptibles tends to zero in absence of isolation.

The model can include only some of complexities tied to the interactions between immunological and epidemic processes. A fundamental missing feature is host heterogeneity; the model assumes identical immune systems for all individuals, while we know that individuals differ deeply. In principle, it would not be difficult allowing for the parameter values of system (1)–(2) to be distributed in the population, though the system would certainly be more complicated. A relatively simple example, demonstrating the relevance of host heterogeneity for the evolution of virulence, is provided by Pugliese (2011).

The within-host dynamics considered are also relatively simple and far from being realistic, by neglecting most of complexities of immune system, and not distinguishing between antibodies, T-cells, memory cells, etc. It would be certainly possible using a more detailed (and higher-dimensional) model for the within-host system; it is not clear to us whether this increase in realism would justify the added complexity, but this is certainly an area worth exploring.

The transmission mechanism of specific pathogens could be in-

investigated in detail to give realistic estimates of the probability density $p(x, x_0)$ or of the transmission coefficient ρ , though we did not face this task in the present paper. We just mention that in recent years significant studies and measurements have been produced about the airborne transmission of influenza A virus (see e.g. Tellier, 2009 and Yang *et al.*, 2011). Such results could help to relate the virus emission to the pathogen burden of the infectant, and to assess the amount of virions taken up during different contact modalities.

Finally, we did not allow, as already discussed, for reinfections of an already infected individual. From the mathematical point of view, it would not be difficult considering them in system (3); one just needs to add two terms, the decrease of i due to reinfections at a value (x, y) , and the consequent increase at $(x + k, y)$ where k is the (fixed or variable) increase in pathogen level due to a reinfection.

However system (3) does not appear easily amenable to analysis or numerical simulation, even without reinfections. We found system (10)–(12), based on age since infection, much more convenient. Clearly, the concept of age since infection loses sense if reinfections would be allowed. One may try to distinguish then between individuals infected once, those infected twice, and so forth, and structure each of these classes by the ages after each infection. The model would become a nightmare in complexity, and it is not even clear whether such a fine description is worthwhile.

It is generally believed that reinfections are relevant, either shortly after the initial infection, when they may help in overcoming the defences of the innate immune system; or long after the first infection, after a decline in the specific immune response. On the other hand, during the phase of acute infection or shortly afterwards, possible reinfections do not seem to play a role. Multiple exposures in a short period are a mechanism to realise the idea by Dushoff (1996) that population prevalence may influence epidemic outcome, here mediated instead by initial dose.

Hence, instead of allowing for a general model with reinfections, it appears more promising aiming at simplified versions targeted to a specific aspect (for instance, reinfections shortly after the initial infection, or after decline of specific immune response). Developing and analysing such types of simplified models will be an important aim in our future research activities.

References

- André JB, Gandon S (2006) Vaccination, within-host dynamics, and virulence evolution. *Evolution* 60:13–23
- Angulo O, Milner F, Segal L, A SIR epidemic model structured by immunological variables. *J Biol Systems*. To appear

- Boldin B, Diekmann O (2008) Superinfections can induce evolutionarily stable coexistence of pathogens. *J Math Biol* 56:635–672
- Brännström A, Carlsson L, Simpson D (2013) On the convergence of the Escalator Boxcar Train. Preprint
- Breda D, Diekmann O, de Graaf WF, Pugliese A, Vermiglio AR (2012) On the formulation of epidemic models (an appraisal of Kermack and McKendrick). *J Biol Dynamics* 6:103–117
- de Roos AM (1988) Numerical methods for structured population models: the escalator boxcar train. *Numer Methods PDE* 4:173–195
- Diekmann O, Heesterbeek JAP (2000) *Mathematical Epidemiology of Infectious Diseases*. Wiley, New York
- d’Onofrio A (2010) On the interaction between the immune system and an exponentially replicating pathogen. *Math Biosci Eng* 7:579–602
- Dushoff J (1996) Incorporating immunological ideas in epidemiological models. *J Theor Biol* 180:181–187
- Gilchrist M, Sasaki A (2002) Modeling host-parasite coevolution. *J Theor Biol* 218:289–308
- Gilchrist M, Coombs D (2006) Evolution of virulence: interdependence, constraints and selection using nested models. *Theor Pop Biol* 69:145–153
- Gwiazda P, Lorenz T, Marciniak-Czochra A (2010) A nonlinear structured population model: Lipschitz continuity of measure-valued solutions with respect to model ingredients. *J Differential Equations* 248:2703–2735
- Hellriegel B (2001) Immunoepidemiology – bridging the gap between immunology and epidemiology. *Trends Parasitol* 17:102–106
- Iannelli M (1994) *Mathematical theory of age-structured population dynamics*. Giardini, Pisa
- Kostova T (2007) Persistence of viral infections on the population level explained by an immunoepidemiological model. *Math Biosci* 206:309–319
- Martcheva M, Pilyugin S (2006) An epidemic model structured by host immunity. *J Biol Systems* 14:185–203
- Metz JAJ, Diekmann O (1986) *The dynamics of physiologically structured populations*. Lecture Notes in Biomathematics Vol 68, Springer, Berlin
- 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

- Milner FA, Pugliese A (1999) Periodic solutions: A robust numerical method for an $S - I - R$ model of epidemics. *J Math Biol* 39:471–492
- Mohtashemi M, Levins R (2001) Transient dynamics and early diagnosis in infectious disease. *J Math Biol* 43:446–470
- Nowak MA, May R (2000) *Virus dynamics: Mathematical principles of immunology and virology*. Oxford University Press, Oxford
- Paunio M, Peltola H, Heinonen OP (1998) Explosive school-based measles outbreak: Intense exposure may have resulted in high risk, even among revaccinees. *American J Epidem* 148:1103–1110
- Perelson AS, Kirschner DE, De Boer R (1993) Dynamics of HIV infection of CD4+T cells. *Math Biosci* 114:81–125
- Pugliese A, Gandolfi A (2008) A simple model of pathogen-immune dynamics including specific and non-specific immunity. *Math Biosci* 214:73–80
- Pugliese A (2011) The role of host population heterogeneity in the evolution of virulence. *J Biol Dynamics* 5:104–119
- Tellier R (2009) Aerosol transmission of influenza A virus: a review of new studies. *J R Soc Interface* 6:S783–S790
- Thieme HR, Castillo-Chavez C (1993) How may infection-age-dependent infectivity affect the dynamics of HIV/AIDS? *SIAM J Appl Math* 53:1447–1479
- Woolhouse M (1992) A theoretical framework for the immunoepidemiology of helminth infection. *Parasite Immunol* 14:563–578
- Yang W, Elankumaran S, Marr LC (2011) Concentrations and size distributions of airborne influenza A viruses measured indoors at a health centre, a day-care centre and on aeroplanes. *J R Soc Interface* 8: 1176–1184

7 Appendix A: derivation of the boundary condition for $u(t, \vartheta, x_0)$, single infection with proportional pathogen load

Suppose $x_m < x_1$. Then, x_0/ρ lies in one of the following intervals: i) $x_m < x_0/\rho < x_1$, ii) $x_1 < x_0/\rho < x_2$, iii) $x_0/\rho > x_2$. Since we will make Δx in (25) tend to zero, we can take Δx sufficiently small so that $\frac{x_0 - \Delta x}{\rho}$ belongs to the same interval containing $\frac{x_0}{\rho}$. Using (25) in (11) we obtain the boundary condition for $\Delta x > 0$ in each interval of interest (see Figure 10):

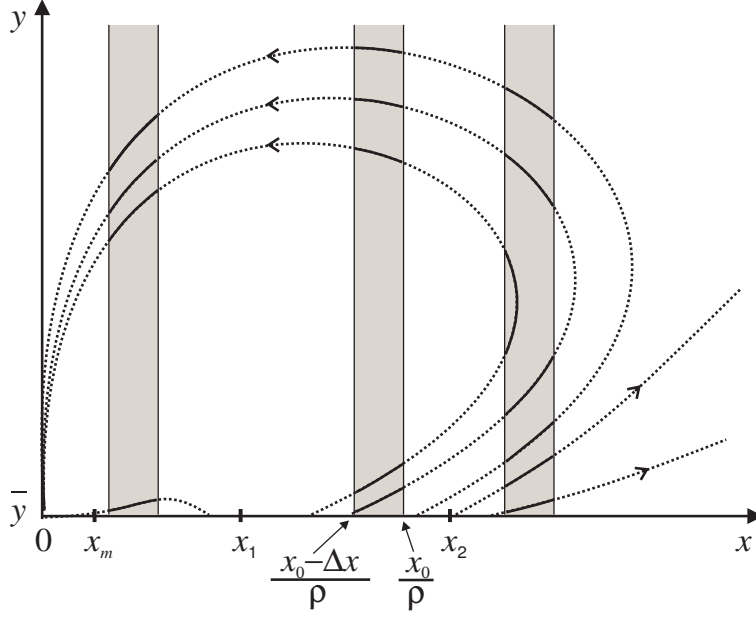


Figure 10: Sketch illustrating the computation in Eqs. (A.1)-(A.3). Examples of trajectories converging to $(0, \eta)$ or diverging are shown as dashed lines. The interval $[\frac{x_0 - \Delta x}{\rho}, \frac{x_0}{\rho}]$ is drawn in three different positions accordingly to cases i), ii) and iii). Integrations in Eqs. (A.1)-(A.3) with respect to ϑ correspond to integrating along the trajectory segments (solid lines) singled out by the shaded areas.

i) $x_m < \frac{x_0 - \Delta x}{\rho} < \frac{x_0}{\rho} < x_1$

$$\begin{aligned}
u(t, 0, x_0) = & \frac{S(t)}{\Delta x} \left[\int_{\frac{x_0 - \Delta x}{\rho}}^{\frac{x_0}{\rho}} \int_0^{\tau_2(\frac{x_0 - \Delta x}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta dx'_0 \right. \\
& + \int_{\frac{x_0}{\rho}}^{x_1} \int_{\tau_2(\frac{x_0}{\rho}, x'_0)}^{\tau_2(\frac{x_0 - \Delta x}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta dx'_0 \\
& \left. + \int_{x_1}^{x_2} \int_{\tau_2(\frac{x_0}{\rho}, x'_0)}^{\tau_2(\frac{x_0 - \Delta x}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta dx'_0 \right], \quad (\text{A.1})
\end{aligned}$$

$$\text{ii) } x_1 < \frac{x_0 - \Delta x}{\rho} < \frac{x_0}{\rho} < x_2$$

$$\begin{aligned}
u(t, 0, x_0) &= \frac{S(t)}{\Delta x} \left[\int_{h(\frac{x_0 - \Delta x}{\rho})}^{h(\frac{x_0}{\rho})} \int_{\tau_1(\frac{x_0 - \Delta x}{\rho}, x'_0)}^{\tau_2(\frac{x_0 - \Delta x}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta dx'_0 \right. \\
&+ \int_{h(\frac{x_0}{\rho})}^{\frac{x_0 - \Delta x}{\rho}} \left(\int_{\tau_1(\frac{x_0 - \Delta x}{\rho}, x'_0)}^{\tau_1(\frac{x_0}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta \right. \\
&+ \left. \int_{\tau_2(\frac{x_0}{\rho}, x'_0)}^{\tau_2(\frac{x_0 - \Delta x}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta \right) dx'_0 \\
&+ \int_{\frac{x_0 - \Delta x}{\rho}}^{\frac{x_0}{\rho}} \left(\int_0^{\tau_1(\frac{x_0}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta \right. \\
&+ \left. \int_{\tau_2(\frac{x_0}{\rho}, x'_0)}^{\tau_2(\frac{x_0 - \Delta x}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta \right) dx'_0 \\
&\left. + \int_{\frac{x_0}{\rho}}^{x_2} \int_{\tau_2(\frac{x_0}{\rho}, x'_0)}^{\tau_2(\frac{x_0 - \Delta x}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta dx'_0 \right], \quad (\text{A.2})
\end{aligned}$$

$$\text{iii) } x_2 < \frac{x_0 - \Delta x}{\rho} < \frac{x_0}{\rho}$$

$$\begin{aligned}
u(t, 0, x_0) &= \frac{S(t)}{\Delta x} \left[\int_{h(\frac{x_0 - \Delta x}{\rho})}^{h(\frac{x_0}{\rho})} \int_{\tau_1(\frac{x_0 - \Delta x}{\rho}, x'_0)}^{\tau_2(\frac{x_0 - \Delta x}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta dx'_0 \right. \\
&+ \int_{h(\frac{x_0}{\rho})}^{x_2} \left(\int_{\tau_1(\frac{x_0 - \Delta x}{\rho}, x'_0)}^{\tau_1(\frac{x_0}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta \right. \\
&+ \left. \int_{\tau_2(\frac{x_0}{\rho}, x'_0)}^{\tau_2(\frac{x_0 - \Delta x}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta \right) dx'_0 \\
&+ \int_{x_2}^{\frac{x_0 - \Delta x}{\rho}} \int_{\tau_1(\frac{x_0 - \Delta x}{\rho}, x'_0)}^{\tau_1(\frac{x_0}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta dx'_0 \\
&+ \left. \int_{\frac{x_0 - \Delta x}{\rho}}^{\frac{x_0}{\rho}} \int_0^{\tau_1(\frac{x_0}{\rho}, x'_0)} \tilde{\beta}(\vartheta, x'_0) u(t, \vartheta, x'_0) d\vartheta dx'_0 \right]. \quad (\text{A.3})
\end{aligned}$$

Let us consider now the above expressions in the limit $\Delta x \rightarrow 0$. We get:

$$\text{i) } x_m < \frac{x_0}{\rho} < x_1$$

$$\begin{aligned}
u(t, 0, x_0) &= \\
&\frac{S(t)}{\rho} \left[- \int_{\frac{x_0}{\rho}}^{x_1} \tilde{\beta}(\tau_2(\frac{x_0}{\rho}, x'_0), x'_0) u(t, \tau_2(\frac{x_0}{\rho}, x'_0), x'_0) \frac{\partial \tau_2}{\partial x}(\frac{x_0}{\rho}, x'_0) dx'_0 \right. \\
&\left. - \int_{x_1}^{x_2} \tilde{\beta}(\tau_2(\frac{x_0}{\rho}, x'_0), x'_0) u(t, \tau_2(\frac{x_0}{\rho}, x'_0), x'_0) \frac{\partial \tau_2}{\partial x}(\frac{x_0}{\rho}, x'_0) dx'_0 \right], \quad (\text{A.4})
\end{aligned}$$

ii) $x_1 < \frac{x_0}{\rho} < x_2$

$$\begin{aligned}
u(t, 0, x_0) = & \\
& \frac{S(t)}{\rho} \left[\int_{h(\frac{x_0}{\rho})}^{\frac{x_0}{\rho}} \left(\tilde{\beta}(\tau_1(\frac{x_0}{\rho}, x'_0), x'_0) u(t, \tau_1(\frac{x_0}{\rho}, x'_0), x'_0) \frac{\partial \tau_1}{\partial x}(\frac{x_0}{\rho}, x'_0) \right. \right. \\
& \left. \left. - \tilde{\beta}(\tau_2(\frac{x_0}{\rho}, x'_0), x'_0) u(t, \tau_2(\frac{x_0}{\rho}, x'_0), x'_0) \frac{\partial \tau_2}{\partial x}(\frac{x_0}{\rho}, x'_0)) \right) dx'_0 \right. \\
& \left. - \int_{\frac{x_0}{\rho}}^{x_2} \tilde{\beta}(\tau_2(\frac{x_0}{\rho}, x'_0), x'_0) u(t, \tau_2(\frac{x_0}{\rho}, x'_0), x'_0) \frac{\partial \tau_2}{\partial x}(\frac{x_0}{\rho}, x'_0) dx'_0 \right], \quad (\text{A.5})
\end{aligned}$$

iii) $\frac{x_0}{\rho} > x_2$

$$\begin{aligned}
u(t, 0, x_0) = & \\
& \frac{S(t)}{\rho} \left[\int_{h(\frac{x_0}{\rho})}^{x_2} \left(\tilde{\beta}(\tau_1(\frac{x_0}{\rho}, x'_0), x'_0) u(t, \tau_1(\frac{x_0}{\rho}, x'_0), x'_0) \frac{\partial \tau_1}{\partial x}(\frac{x_0}{\rho}, x'_0) \right. \right. \\
& \left. \left. - \tilde{\beta}(\tau_2(\frac{x_0}{\rho}, x'_0), x'_0) u(t, \tau_2(\frac{x_0}{\rho}, x'_0), x'_0) \frac{\partial \tau_2}{\partial x}(\frac{x_0}{\rho}, x'_0)) \right) dx'_0 \right. \\
& \left. + \int_{x_2}^{\frac{x_0}{\rho}} \tilde{\beta}(\tau_1(\frac{x_0}{\rho}, x'_0), x'_0) u(t, \tau_1(\frac{x_0}{\rho}, x'_0), x'_0) \frac{\partial \tau_1}{\partial x}(\frac{x_0}{\rho}, x'_0) dx'_0 \right]. \quad (\text{A.6})
\end{aligned}$$

Note that the first double integrals in (A.2) and (A.3) yield the contributions of trajectories such that $\frac{x_0 - \Delta x}{\rho} < M(x'_0) \leq \frac{x_0}{\rho}$. Such contributions, in the limit $\Delta x \rightarrow 0$, vanish since the domain of integration over ϑ vanishes.

Taking into account (28), the final expressions (30)-(32) are readily obtained. In case of $x_m \geq x_1$, the boundary condition can be derived through obvious modifications. For instance, if $x_1 \leq x_m < x_2$, the boundary condition will be given for $x_m < \frac{x_0}{\rho} < x_2$ by (A.5), and for $\frac{x_0}{\rho} > x_2$ by (A.6).

8 Appendix B: derivation of (44)

Integrating (39) with respect to w_0 , with $w_0^i < w_2$, the first addend will be given by

$$S(t) \int_{w_0^{i-1}}^{w_0^i} \int_{\hat{h}(w_0 + \Delta)}^{w_0 + \Delta} e^{w_0 + \Delta} \hat{K}_1(t, w_0, w'_0) dw'_0 dw_0.$$

Exchanging the order of integration, this can be rewritten (assuming $\hat{h}(w_0^i + \Delta) < w_0^{i-1} + \Delta$ which will always hold as long as the step size is small enough)

as

$$S(t) \left[\begin{aligned} & \int_{\hat{h}(w_0^{i-1}+\Delta)}^{\hat{h}(w_0^i+\Delta)} \int_{w_0^{i-1}}^{\hat{M}(w_0')-\Delta} e^{w_0+\Delta} \hat{K}_1(t, w_0, w_0') dw_0 dw_0' \\ & + \int_{\hat{h}(w_0^i+\Delta)}^{w_0^{i-1}+\Delta} \int_{w_0^{i-1}}^{w_0^i} e^{w_0+\Delta} \hat{K}_1(t, w_0, w_0') dw_0 dw_0' \\ & + \int_{w_0^{i-1}+\Delta}^{w_0^i+\Delta} \int_{w_0'-\Delta}^{w_0^i} e^{w_0+\Delta} \hat{K}_1(t, w_0, w_0') dw_0 dw_0' \end{aligned} \right],$$

where $\hat{M}(w) = \ln M(e^w)$ is the inverse function of $\hat{h}(w)$.

By setting

$$\vartheta = \hat{\tau}_1(w_0 + \Delta, w_0'),$$

and recognizing that

$$d\vartheta = \frac{e^{w_0+\Delta}}{F(e^{w_0+\Delta}, \eta(\hat{\tau}_1(w_0 + \Delta, w_0'), e^{w_0'}))} dw_0,$$

we obtain

$$S(t) \left[\begin{aligned} & \int_{\hat{h}(w_0^{i-1}+\Delta)}^{\hat{h}(w_0^i+\Delta)} \int_{\hat{\tau}_1(w_0^{i-1}+\Delta, w_0')}^{T(e^{w_0'})} \tilde{\beta}(\vartheta, e^{w_0'}) \hat{u}(t, \vartheta, w_0') d\vartheta dw_0' \\ & + \int_{\hat{h}(w_0^i+\Delta)}^{w_0^{i-1}+\Delta} \int_{\hat{\tau}_1(w_0^{i-1}+\Delta, w_0')}^{\hat{\tau}_1(w_0^i+\Delta, w_0')} \tilde{\beta}(\vartheta, e^{w_0'}) \hat{u}(t, \vartheta, w_0') d\vartheta dw_0' \\ & + \int_{w_0^{i-1}+\Delta}^{w_0^i+\Delta} \int_0^{\hat{\tau}_1(w_0^i+\Delta, w_0')} \tilde{\beta}(\vartheta, e^{w_0'}) \hat{u}(t, \vartheta, w_0') d\vartheta dw_0' \end{aligned} \right]$$

We remind the reader that $T(x_0)$ is the time needed for a solution starting with $x = x_0$ to reach the maximum value in x , $M(x_0)$, of the solution. The next step will be slightly different according to whether T is increasing or decreasing (or non-monotone) in $(\hat{h}(w_0^{i-1} + \Delta), \hat{h}(w_0^i + \Delta))$; to be definite, we assume that it is decreasing. For convenience we also define $\hat{T}(w) = T(h(e^w))$, i.e. the time needed for a solution having maximum value e^w to reach it from its initial point.

Now we exchange again the integration order; we need to consider the inverse of the function $\hat{\tau}_1$ (it can be indeed shown that it is monotone in the second argument for $x_0 \in (x_1, +\infty)$): let then $r_1(\vartheta, w_0)$ defined as the value such that

$$\hat{\tau}_1(w_0, r_1(\vartheta, w_0)) = \vartheta.$$

With this definition, exchanging the order of integration yields:

$$S(t) \left[\begin{aligned} & \int_0^{\hat{T}(w_0^i+\Delta)} \int_{r_1(\vartheta, w_0^{i-1}+\Delta)}^{r_1(\vartheta, w_0^i+\Delta)} \tilde{\beta}(\vartheta, e^{w_0'}) \hat{u}(t, \vartheta, w_0') dw_0' d\vartheta \\ & + \int_{\hat{T}(w_0^i+\Delta)}^{\hat{T}(w_0^{i-1}+\Delta)} \int_{r_1(\vartheta, w_0^{i-1}+\Delta)}^{\log(T^{-1}(\vartheta))} \tilde{\beta}(\vartheta, e^{w_0'}) \hat{u}(t, \vartheta, w_0') dw_0' d\vartheta \end{aligned} \right]$$

Note that if w'_0 belongs to $[r_1(\vartheta, w_0^{i-1} + \Delta), r_1(\vartheta, w_0^i + \Delta)]$, by definition

$$e^{w_0^{i-1} + \Delta} \leq \xi(\vartheta, e^{w'_0}) \leq e^{w_0^i + \Delta}.$$

They are exactly those values w'_0 such that the solution starting from $e^{w'_0}$ is at time ϑ for the first time between $e^{w_0^{i-1} + \Delta}$ and $e^{w_0^i + \Delta}$. The same is true when $\hat{T}(w_0^i + \Delta) < \vartheta \leq \hat{T}(w_0^{i-1} + \Delta)$, except that, at such values of ϑ the maximum x that can be reached will not be equal to $e^{w_0^i + \Delta}$ but something less.

We perform now the same steps with the second term in (39). Integrating (39) with respect to w_0 , with $w_0^i < w_2$, the second addend will be given by

$$S(t) \int_{w_0^{i-1}}^{w_0^i} \int_{\hat{h}(w_0 + \Delta)}^{w_2} e^{w_0 + \Delta} (-\hat{K}_2(t, w_0, w'_0)) dw'_0 dw_0.$$

Exchanging the order of integration, the above expression can be rewritten as

$$S(t) \left[\int_{\hat{h}(w_0^{i-1} + \Delta)}^{\hat{h}(w_0^i + \Delta)} \int_{w_0^{i-1}}^{\hat{M}(w'_0) - \Delta} e^{w_0 + \Delta} (-\hat{K}_2(t, w_0, w'_0)) dw_0 dw'_0 + \int_{\hat{h}(w_0^i + \Delta)}^{w_2} \int_{w_0^{i-1}}^{w_0^i} e^{w_0 + \Delta} (-\hat{K}_2(t, w_0, w'_0)) dw_0 dw'_0 \right].$$

By setting

$$\vartheta = \hat{\tau}_2(w_0 + \Delta, w'_0),$$

and recognizing that

$$d\vartheta = \frac{e^{w_0 + \Delta}}{F(e^{w_0 + \Delta}, \eta(\hat{\tau}_2(w_0 + \Delta, w'_0), e^{w'_0}))} dw_0,$$

we obtain

$$S(t) \left[\int_{\hat{h}(w_0^{i-1} + \Delta)}^{\hat{h}(w_0^i + \Delta)} \int_{\hat{\tau}_2(w_0^{i-1} + \Delta, w'_0)}^{T(e^{w'_0})} (-\tilde{\beta}(\vartheta, e^{w'_0}) \hat{u}(t, \vartheta, w'_0)) d\vartheta dw'_0 + \int_{\hat{h}(w_0^i + \Delta)}^{w_2} \int_{\hat{\tau}_2(w_0^{i-1} + \Delta, w'_0)}^{\hat{\tau}_2(w_0^i + \Delta, w'_0)} (-\tilde{\beta}(\vartheta, e^{w'_0}) \hat{u}(t, \vartheta, w'_0)) d\vartheta dw'_0 \right],$$

and then

$$S(t) \left[\int_{\hat{h}(w_0^{i-1} + \Delta)}^{\hat{h}(w_0^i + \Delta)} \int_{T(e^{w'_0})}^{\hat{\tau}_2(w_0^{i-1} + \Delta, w'_0)} \tilde{\beta}(\vartheta, e^{w'_0}) \hat{u}(t, \vartheta, w'_0) d\vartheta dw'_0 + \int_{\hat{h}(w_0^i + \Delta)}^{w_2} \int_{\hat{\tau}_2(w_0^i + \Delta, w'_0)}^{\hat{\tau}_2(w_0^{i-1} + \Delta, w'_0)} \tilde{\beta}(\vartheta, e^{w'_0}) \hat{u}(t, \vartheta, w'_0) d\vartheta dw'_0 \right].$$

To exchange again the integration order, we introduce the inverse of the function $\hat{\tau}_2$ with respect to the second argument, i.e. the function $r_2(\vartheta, w_0)$ defined as the value such that

$$\hat{\tau}_2(w_0, r_2(\vartheta, w_0)) = \vartheta.$$

Since $\lim_{w'_0 \rightarrow w_2^-} \hat{\tau}_2(w, w'_0) = +\infty$ for any w , we get:

$$S(t) \left[\int_{\hat{T}(w_0^i + \Delta)}^{\hat{T}(w_0^{i-1} + \Delta)} \int_{\log(T^{-1}(\vartheta))}^{r_2(\vartheta, w_0^i + \Delta)} \tilde{\beta}(\vartheta, e^{w'_0}) \hat{u}(t, \vartheta, w'_0) dw'_0 d\vartheta + \int_{\hat{T}(w_0^{i-1} + \Delta)}^{\infty} \int_{r_2(\vartheta, w_0^{i-1} + \Delta)}^{r_2(\vartheta, w_0^i + \Delta)} \tilde{\beta}(\vartheta, e^{w'_0}) \hat{u}(t, \vartheta, w'_0) dw'_0 d\vartheta \right].$$

Summing the contribution of the first and of the second addend, the following expression is obtained:

$$S(t) \left[\int_0^{\hat{T}(w_0^i + \Delta)} \int_{r_1(\vartheta, w_0^{i-1} + \Delta)}^{r_1(\vartheta, w_0^i + \Delta)} \tilde{\beta}(\vartheta, e^{w'_0}) \hat{u}(t, \vartheta, w'_0) dw'_0 d\vartheta + \int_{\hat{T}(w_0^i + \Delta)}^{\hat{T}(w_0^{i-1} + \Delta)} \int_{r_1(\vartheta, w_0^{i-1} + \Delta)}^{r_2(\vartheta, w_0^i + \Delta)} \tilde{\beta}(\vartheta, e^{w'_0}) \hat{u}(t, \vartheta, w'_0) dw'_0 d\vartheta + \int_{\hat{T}(w_0^i + \Delta)}^{\infty} \int_{r_2(\vartheta, w_0^{i-1} + \Delta)}^{r_2(\vartheta, w_0^i + \Delta)} \tilde{\beta}(\vartheta, e^{w'_0}) \hat{u}(t, \vartheta, w'_0) dw'_0 d\vartheta \right].$$

Hence, observing that if w'_0 belongs to $[r_1(\vartheta, w_0^{i-1} + \Delta), r_2(\vartheta, w_0^i + \Delta)]$ or to $[r_2(\vartheta, w_0^{i-1} + \Delta), r_2(\vartheta, w_0^i + \Delta)]$ in the appropriate intervals of ϑ , it is

$$e^{w_0^{i-1} + \Delta} \leq \xi(\vartheta, e^{w'_0}) \leq e^{w_0^i + \Delta},$$

one arrives at the expression (44). The same result can be achieved, after obvious changes within the procedure steps, when $T(x_0)$ is increasing or non-monotone in $(\hat{h}(w_0^{i-1} + \Delta), \hat{h}(w_0^i + \Delta))$. Similarly, (44) is obtained also starting from (40).

9 Appendix C: numerical computation of the density $\hat{u}(t, \vartheta, w_0)$

On the domain $[w_{in}, w_{fin}] \times [0, \vartheta_{max}]$ of the plane (w_0, ϑ) , we define the rectangular grid:

$$\begin{aligned} \vartheta_i &= i\delta, & i &= 0, \dots, N, \\ w_0^j &= w_{in} + jh, & j &= 0, \dots, M, \end{aligned}$$

where $\delta = \vartheta_{max}/N$ and $h = (w_{fin} - w_{in})/M$. Denoting by T the maximal integration time so that the integration will be performed over $[0, T]$, we choose $T = L\delta$. We then define the discretized time

$$t_k = k\delta, \quad k = 0, \dots, L.$$

Note that the correct integration of Eq. (38) requires a constraint on T depending on the support of the initial condition $\hat{u}(0, \vartheta, w_0)$. Given ϑ_{max} , if the initial condition is zero for $\vartheta > \vartheta^0$, T cannot exceed $\vartheta_{max} - \vartheta^0$.

Preliminarily, the values

$$W_{ij} = \ln \xi(\vartheta_i, e^{w_0^j}),$$

$$Z_{ij} = \ln \eta(\vartheta_i, e^{w_0^j}),$$

are computed for all the nodes of the grid, solving the system (35) by means of the Cash-Karp 6th-order Runge-Kutta method.

Let us define the array $\hat{U}(i, j, k) = \hat{u}(t_k, \vartheta_i, w_0^j)$. The density $\hat{u}(t, \vartheta, w_0)$ on the grid is then computed along the characteristic lines by the Euler method, according to the following steps:

- Load the initial condition $\hat{U}(i, j, 1)$, $i = 0, \dots, N$, $j = 0, \dots, M$.
- For $k = 0, \dots, L - 1$,
 - i) set $\hat{U}(0, j, k + 1) = 0$, $j = 0, \dots, M$;
 - ii) for $i = 0, \dots, N - 1$, $j = 0, \dots, M$, compute

$$\hat{U}(i + 1, j, k + 1) = \hat{U}(i, j, k) + \Psi(i, j, k)\delta,$$

where $\Psi(i, j, k)$ denotes the r.h.s. of Eq. (38) at $\vartheta = \vartheta_i$, $w_0 = w_0^j$ and $t = t_k$;

- iii) update the boundary condition, computing the new values

$$\hat{U}(0, j, k + 1), j = 0, \dots, M.$$

Concerning the step iii), we compute for $i = 1, \dots, N - 1$ and $j = 0, \dots, M - 1$:

$$B_{ijk} = \tilde{\beta}(\vartheta_i, e^{w_0^j})S(t_{k+1})\hat{U}(i, j, k + 1)h\delta.$$

The quantity $B_{ijk}dt$ approximates the number of new infections caused by infected individuals belonging to the elementary domain $[w_0^j, w_0^j + h] \times [\vartheta_i, \vartheta_i + \delta]$ that occur in the time interval $(t_{k+1}, t_{k+1} + dt)$. As a further

approximation, we assume that all the individuals in the above elementary domain have the same pathogen load W_{ij} . Thus, in view of (44) we have

$$\int_{w_0^j}^{w_0^{j+1}} \hat{u}(t, 0, w_0) dw_0 \simeq \sum_{\{(i,j): W_{i,j} - \Delta \in [w_0^j, w_0^{j+1}]\}} B_{ijk},$$

from which $\hat{U}(0, j, k + 1)$ is derived as:

$$\hat{U}(0, j, k + 1) = \frac{1}{h} \left[\sum_{\{(i,j): W_{i,j} - \Delta \in [w_0^j, w_0^{j+1}]\}} B_{ijk} \right].$$

To compute $S(t_k)$, the integration of (7) is performed by the Euler method. In the calculation of $\lambda(t_k)$ according to (43), the double integral is computed over the domain $[w_{in}, w_{fin}] \times [0, \vartheta_{max}]$.