

A structured Markov chain model to investigate the effects of pre-exposure vaccines in tuberculosis control

R. Fernández-Peralta^a, A. Gómez-Corral^{b,*}

^a*Department of Mathematics and Computer Science, University of Balearic Islands,
Ctra. Valldemossa, Km. 7.5, 07122-Palma, Spain*

^b*Department of Statistics and Operations Research, School of Mathematical Sciences,
Plaza de Ciencias 3, Complutense University of Madrid, 28040-Madrid, Spain*

Abstract

In this paper, the interest is in a structured Markov chain model to describe the transmission dynamics of tuberculosis (TB) in the setting of small communities of hosts sharing confined spaces, and to explore the potential impact of new pre-exposure vaccines on reducing the number of new TB cases during an outbreak of the disease. The model under consideration incorporates endogenous reactivation of latent tubercle bacilli, exogenous reinfection of latently infected TB hosts, loss of effectiveness of the vaccine protection, and death of hosts due to tubercle bacilli and from causes beyond TB. Various probabilistic measures are defined and analytically studied to describe extreme values and the number of vaccinations during an outbreak, and a random version of the basic reproduction number is used to measure the transmission potential during the initial phase of the epidemic. Our numerical experiments allow us to compare different pre-exposure vaccines versus the level of coverage in terms of these probabilistic measures.

*Corresponding author

Email addresses: `r.fernandez@uib.es` (R. Fernández-Peralta), `agcorral@ucm.es` (A. Gómez-Corral)

URL: `personal.uib.eu/r.fernandez` (R. Fernández-Peralta),
`blogs.mat.ucm.es/agomez-corral` (A. Gómez-Corral)

Keywords: Epidemiological modeling; *Mycobacterium tuberculosis*;

Pre-exposure vaccination; Quasi-birth-death processes

2000 MSC: 60J28 Applications of continuous-time Markov processes on discrete state spaces; 92B05 General biology and biomathematics

1. Introduction

Tuberculosis (TB), which is a communicable disease caused by the bacterial species *Mycobacterium tuberculosis*, is one of the top ten world's major causes of death and physical and economic deprivation in human and animal populations and the leading cause of mortality from a single infectious agent. Based on seven million new cases which were notified to the World Health Organization [55], it is estimated that ten million people fell ill with TB in 2018, but the burden of disease is observed to vary notably among countries. Most new cases of TB were reported from regions in South-East Asia (44%), with smaller shares in the Americas (3%) and Europe (3%). Eight countries –more concretely, India (27%), China (9%), Indonesia (8%), the Philippines (6%), Pakistan (6%), Nigeria (4%), Bangladesh (4%) and South Africa (3%)– accounted for two thirds of the global total; see e.g. [55, Chapter 3].

The tubercle bacilli live typically in the alveoli of the lungs, but they can also affect other sites, and are spread when a host that has an *active* TB infection expels bacilli –for example, by speaking, coughing or sneezing–, and a susceptible host inhales bacilli from the air. A higher risk of becoming infected with bacilli is therefore linked to susceptible hosts sharing the same environment with hosts with active TB disease. Upon infection, most infected TB hosts do not develop symptoms, which amounts to a *latent*

TB infection. Hosts with latent TB disease are not clinically ill nor capable of transmitting TB, but they can develop then active TB disease as a consequence of either exogenous reinfection or an endogenous reactivation of latent bacilli. Although only a small proportion (estimated at 5–10%) of latently infected TB hosts will develop active TB disease during the lifetime, the probability of progressing towards active TB disease becomes higher among those hosts affected by health-related risk factors, such as smoking, diabetes, HIV, harmful use of alcohol and undernourishment.

A prompt and accurate diagnosis followed by provision of effective anti-tuberculosis chemotherapy (van den Boogaard et al. [50]) prevents deaths among actively infected TB hosts, and limits further transmission of bacilli to others; see [55, Chapter 4]. However, the use of antibiotics has side effects and takes a long time, varying from 6 to 24 months, which can lead to non-adherence to treatment (Awofeso [9]; Mekonnen and Azagew [41]) and the development of antibiotic resistant TB (Lange et al. [34]; Nachega and Chaisson [44]). Prevention of new infections of tubercle bacilli through prophylactic health care interventions, and their progression to active TB disease through therapeutic vaccination strategies (Gomes et al. [27]), is crucial for reduction in active TB cases and deaths, and to achieve the End TB Strategy targets (Dye et al. [20]; World Health Organization [55, Chapter 5]) of an 80% reduction in incidence from the 2015 level by 2030, and of a 90% reduction by 2035. An important milestone for this formidable challenge (Abu-Raddad et al. [1]; Ziv et al. [58]) is the development of more effective vaccines than the only licensed BCG (*Bacillus Calmette-Guérin*) vaccine. The BCG vaccine was first introduced in 1921 and has been widely used as a neonatal vaccine having consistently high efficacy against childhood tuberculous meningitis and miliary TB (World Health Organization [54]).

However, it has demonstrated limited effectiveness in protecting against pulmonary TB in adolescents and adults –especially, in immuno-compromised subjects– and BCG revaccination in adolescents deserves further investigation (Dye [19]). We refer the reader to the paper by Méndez-Sampeiro [42] for more details about the main TB vaccine candidates currently in clinical phase; see also Britton and Palendira [13], and Harris et al. [31]. For a related work, see Knight et al. [32], where the impact and cost-effectiveness of new TB vaccines in low- and middle-income countries is analyzed.

There is an extensive literature of models describing the transmission dynamics of TB, and prophylactic and therapeutic interventions, including the use of vaccines. For instances, different models have investigated the effects of socio-demographic factors (Guzzeta et al. [29]), age-structured populations (Abu-Raddad et al. [1]; Castillo-Chavez and Feng [15]; Renardy and Kirschner [47]), treatment regimes and diagnosis (Abu-Raddad et al. [1]; Gerberry [25]), multi-level contact structures (Song et al. [49]), antibiotic resistance (Agusto et al. [2]; Castillo-Chavez and Feng [14]; Fofana et al. [23]; Mishra and Srivastava [43]; Yu et al. [56]), and pre- and post-exposure vaccines (Abu-Raddad et al. [1]; Bhunu et al. [11]; Gerberry [25, 26]; Gomes et al. [27]; Lietman and Blower [37]; Liu and Zhang [38]; Ziv et al. [58]), among other aspects. A review of models in the study of TB dynamics is the paper by Castillo-Chavez and Song [16], and a survey on mathematical models for epidemiological impact of TB vaccines is provided in Harris et al. [31]. Newly notified TB cases, properly combined with migration statistics, permit (Anzai et al. [7]) to evaluate the possibility of introducing pre-entry TB screening requirements for foreign-born citizens entering a country in order to estimate the prevalence of TB infection by country of birth.

Despite the significant literature in this field, the above-mentioned mod-

els evaluate TB dynamics and vaccination strategies in terms of systems of differential equations, instead of using stochastic modelling. Both deterministic and stochastic perspectives are important for practical use, but stochastic models are usually preferable (Artalejo and López-Herrero [8]; Britton [12]) in the setting of small communities sharing confined spaces, such as intensive care units, households, schools, work places, prisons and elderly residences (Almaraz and Gómez-Corral [4]; Amador et al. [5]; Economou et al. [21]; Gómez-Corral and López-García [28]). In the setting of stochastic vaccination models, recent work (Chen and Kang [17]; Kribs-Zaleta and Martchena [33]; Liu and Jiang [39]; Shim [48]; Witbooi et al. [53]; Zhang et al. [57]) incorporates stochastic fluctuations into deterministic models by means of a suitably defined system of stochastic differential equations. Another way to describe stochastic dynamics in epidemic models with vaccination is related to the use of a continuous-time Markov chain (CTMC) describing the dynamics of the underlying compartments at an arbitrary time. Assuming that the vaccine is imperfect, Gamboa and López-Herrero [24] present a Markov chain model with infection reintroduction and define alternative measures to the basic reproduction number \mathcal{R}_0 in order to study the random effects of the coverage and vaccine efficacy on the transmission of bacilli.

Here we analyze a Markov chain model to explore the effects of a pre-exposure vaccine on a finite population of hosts exposed to tubercle bacilli during an outbreak of the TB disease. Specifically, our distributional assumptions are based on the main features of the deterministic pre-exposure vaccine model of Lietman and Blower [37] (see also Castillo-Chavez and Song [16, Section 8.1]), who show that moderately effective vaccines can have a significant effect on reducing TB when they are suitably coupled with mod-

erate to high treatment rates. More particularly, three different mechanisms of vaccine failure –termed *take*, *degree* and *duration*– are exhaustively investigated by Lietman and Blower [37] in order to identify which pre-exposure vaccines can be found as equivalent in terms of their potential epidemic-control effects. Unlike the deterministic model of Lietman and Blower [37], which is best suited for a large population, the Markov chain model we propose here is related to small communities, such as child protection centers, hospital centers for long-stay patients with specific pathologies, and households or other small transmission units in a mini-community design to assert the indirect effect of vaccination; see e.g. Allen [3], Britton [12] and Halloran [30]. In these real-world settings, it is natural to assume that the community is homogeneous and that hosts mix uniformly with each other, while at the same time it seems reasonable to assume some randomness, for example, in the duration of the epidemic.

This paper is organized as follows. In Section 2, we make specific distributional assumptions leading to a pre-exposure vaccine model with re-activation of a pre-existing dormant TB infection, exogenous reinfection of latently infected TB hosts, the loss of effectiveness of the vaccine protection, and the death of hosts either due to the TB infection or from causes beyond TB. The resulting Markov chain model can be seen as a finite level-dependent quasi-birth-death (LD-QBD; Latouche and Ramaswami [35]) process \mathcal{X} , from which the random length of an outbreak is characterized as a phase-type (PH; Neuts [46]) random variable. In Section 3, the focus is on various probabilistic measures describing the transmission of tubercle bacilli during an outbreak, including the peak of TB infection during the outbreak and the final status of the population (Section 3.1), the cumulative number of vaccinations (Section 3.2), and a random version of the basic reproduction

number (Section 3.3). In Section 4, the algorithmic solution in Section 3 is used to compare three different vaccines versus three levels of coverage by considering key parameters. Finally, some concluding remarks are given in Section 5.

2. The pre-exposure vaccine model and a related LD-QBD process

We consider a Markov chain model for the transmission of TB among a homogeneously well-mixed and finite population of hosts, with constant population size N . A vaccine is administrated before evidence of TB infection, whence the vaccination strategy is assumed to be pre-exposure, also called prophylactic or pre-infection. The host population is decomposed into the following epidemiological groups or *compartments*: unvaccinated susceptible (S_U) and vaccinated susceptible (S_V) hosts; unvaccinated latently infected (L_U) and vaccinated latently infected (L_V) hosts; and actively infected (I) hosts.

A susceptible host becomes infected due to the contact with an actively infected TB host at rates α_U or α_V *per year* if the susceptible host is either unvaccinated or vaccinated, respectively, in such a way that it becomes either actively infected with respective probabilities p_U and p_V , or latently infected with respective probabilities $1 - p_U$ and $1 - p_V$. Under the assumption that the population contains i infected hosts with active TB, a latently infected TB host becomes infected with active TB due to either the endogenous reactivation of latent bacilli or the exogenous reinfection of such a host – acquiring a new infection with active TB from another actively infected TB host – at rates $r_U(i)$ and $r_V(i)$ *per year* if it is unvaccinated and vaccinated, respectively; see Feng et al. [22], and Nematollahi et al. [45].

[Insert here Table 1]

Table 1: Parameters in the stochastic model with pre-exposure vaccination strategy.

Every host, regardless of its current status, can die from causes beyond TB at rate δ_D *per year*, and an actively infected TB host can also die due to the TB infection at rate δ_{TB} *per year*. When a host dies, it is assumed that a new susceptible host replaces it, being the new susceptible host vaccinated with probability $\eta(q, C)$, where q denotes the take (or fraction of vaccinated hosts in whom the vaccine induces some level of protective immunologic response) and C is the coverage level (or proportion of population to be vaccinated); see Castillo-Chavez and Song [16], and Lietman and Blower [37].

When an infected host with active TB recovers, which occurs at rate δ_R *per year*, it is assumed that it becomes either infected with latent TB or susceptible with probabilities ξ and $1 - \xi$, respectively; more particularly, it becomes either unvaccinated latently infected with probability $\xi\theta_U$ or vaccinated latently infected with probability $\xi(1 - \theta_U)$, and it becomes either unvaccinated susceptible with probability $(1 - \xi)\tilde{\theta}_U$ or vaccinated susceptible with probability $(1 - \xi)(1 - \tilde{\theta}_U)$. The expected duration of vaccine-induced immunity in a susceptible host is given by γ^{-1} *years*.

Finally, infectious contacts, and latently infectious and actively infectious periods are assumed to be mutually independent and exponentially distributed, and are also independent of the assignment procedure of probabilities (i.e., p_U , p_V , $\eta(q, C)$, ξ , θ_U and $\tilde{\theta}_U$) governing transitions among compartments. The parameters of the model are summarized in Table 1.

The dynamics of the pre-exposure vaccine model can be described by the

[Insert here Table 2]

Table 2: Transitions between states and underlying rates of \mathcal{X} , and labels used in Figure 1.

[Insert here Figure 1]

Figure 1: Diagram of transitions among compartments and underlying events (Table 2) in the pre-exposure vaccine model.

process $\mathcal{X} = \{X(t) : t \geq 0\}$ with $X(t) = (I(t), L_U(t), L_V(t), S_U(t), S_V(t))$ and $I(t) + L_U(t) + L_V(t) + S_U(t) + S_V(t) = N$, where $I(t)$ is the number of actively infected TB hosts, $L_U(t)$ and $L_V(t)$ are the numbers of unvaccinated and vaccinated latently infected TB hosts, and $S_U(t)$ and $S_V(t)$ denote the numbers of unvaccinated and vaccinated susceptible hosts, respectively, at time t . Clearly, the process \mathcal{X} is a time-homogenous CTMC taking values in the state space

$$\mathcal{S} = \{(i, l_U, l_V, s_U, s_V) : i, l_U, l_V, s_U, s_V \in \mathbb{N}_0, i + l_U + l_V + s_U + s_V = N\}.$$

The non-null transition rates of \mathcal{X} are briefly summarized in Table 2; see also Figure 1 for a diagram of transitions among compartments and related events.

For later use, we define the i th level $l(i)$ consisting of those states with i infected hosts with active TB (i.e., states (i, l_U, l_V, s_U, s_V) verifying $l_U + l_V + s_U \in \{0, \dots, N - i\}$ and $s_V = N - i - l_U - l_V - s_U$ with $l_U, l_V, s_U, s_V \in \mathbb{N}_0$, for integers $i \in \{0, \dots, N\}$), and we decompose \mathcal{S} into subsets \mathcal{S}_A and \mathcal{S}_T , where the class of absorbing states $\mathcal{S}_A = l(0)$ is related to the ultimate extinction of the disease, and $\mathcal{S}_T = \cup_{i=1}^N l(i)$ contains transient states. As a result, the i th level $l(i)$ consists of $J'(i) = \binom{N-i+3}{3}$ states, for integers $i \in \{0, \dots, N\}$,

and the subsets \mathcal{S}_A and \mathcal{S}_T contain $J_A = \binom{N+3}{3}$ and $J_T = \binom{N+4}{4} - \binom{N+3}{3}$ states, respectively.

A suitable labeling of states (Appendix A) results in the following block-structured form for the infinitesimal generator \mathbf{Q} of \mathcal{X} :

$$\mathbf{Q} = \begin{pmatrix} \mathbf{0}_{J_A \times J_A} & \mathbf{0}_{J_A \times J_T} \\ \mathbf{Q}_{T,A} & \mathbf{Q}_{T,T} \end{pmatrix},$$

where the sub-matrices $\mathbf{Q}_{T,A}$ and $\mathbf{Q}_{T,T}$ are related to jumps of \mathcal{X} from states in \mathcal{S}_T to states in \mathcal{S}_A and \mathcal{S}_T , respectively. More concretely, these sub-matrices have the form

$$\mathbf{Q}_{T,T} = \begin{pmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & & & \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \mathbf{A}_{2,3} & & \\ & \ddots & \ddots & \ddots & \\ & & \mathbf{A}_{N-1,N-2} & \mathbf{A}_{N-1,N-1} & \mathbf{A}_{N-1,N} \\ & & & \mathbf{A}_{N,N-1} & \mathbf{A}_{N,N} \end{pmatrix}, \quad (1)$$

$$\mathbf{Q}_{T,A} = \begin{pmatrix} \mathbf{A}_{1,0} \\ \mathbf{0}_{(J_T - \binom{N+2}{3}) \times J_A} \end{pmatrix}, \quad (2)$$

where $\mathbf{0}_{a \times b}$ denotes the null matrix of dimension $a \times b$ and sub-matrices $\mathbf{A}_{i,i'}$ record infinitesimal rates for jumps of \mathcal{X} from states in level $l(i)$ to states in $l(i')$, for $i' \in \{i-1, i, i+1\}$. Specifically, the non-null entries of $\mathbf{A}_{i,i-1}$, for integers $i \in \{1, \dots, N\}$, (Appendix A.1) are linked to those jumps for which the number of actively infected TB hosts decreases by one unit (transitions H-M in Figure 1 and Table 2); off-diagonal entries of $\mathbf{A}_{i,i}$, for $i \in \{1, \dots, N\}$, (Appendix A.2) are related to the loss of efficacy of the vaccine (transition A), the development of latent TB infection (transitions B-C) and the replacement of a susceptible or latently infected TB host by a new susceptible host due to the death of the former (transitions N-U); finally,

sub-matrices $\mathbf{A}_{i,i+1}$, for $i \in \{1, \dots, N-1\}$, (Appendix A.3) are associated with jumps which imply the development of active TB (transitions D-G).

3. Probabilistic descriptors

Let us assume that, for a predetermined state $(i_0, l_U, l_V, s_U, s_V) \in \mathcal{S}_T$, an outbreak begins at time $t = 0$ under the assumption of $I(0) = i_0$ actively infected TB hosts, $L_U(0) = l_U$ and $L_V(0) = l_V$ unvaccinated and vaccinated hosts with latent bacilli, and $S_U(0) = s_U$ and $S_V(0) = s_V$ unvaccinated and vaccinated susceptible hosts. It is clear that the random length τ of the outbreak can be seen as a first-passage time of \mathcal{X} since $\tau = \inf\{t > 0 : X(t) \in l(0)\}$, which leads us to assert that τ is a continuous PH random variable of order J_T and representation $(\pi_T, \mathbf{Q}_{T,T})$, where the row vector π_T records initial probabilities of \mathcal{X} on transient states; i.e., π_T is a null vector of order J_T with a single 1 at the entry associated with the initial state $(i_0, l_U, l_V, s_U, s_V)$.

It is then immediately found that

$$\begin{aligned} P(\tau \leq t | X(0) = (i_0, l_U, l_V, s_U, s_V)) &= 1 - \pi_T e^{\mathbf{Q}_{T,T} t} \mathbf{1}_{J_T}, \quad t > 0, \\ E \left[\tau^k | X(0) = (i_0, l_U, l_V, s_U, s_V) \right] &= k! \pi_T \left(-\mathbf{Q}_{T,T}^{-k} \right) \mathbf{1}_{J_T}, \quad k \in \mathbb{N}, \end{aligned}$$

where the column vector $\mathbf{1}_a$ denotes the unit vector of order a ; see e.g. Latouche and Ramaswami [35, Chapter 2]. It is also found that, based on an analytical treatment of the transition function of \mathcal{X} , the time-dependent vector $\mathbf{f}(t; \mathcal{S}_A)$ with entries $P(\tau \leq t, X(\tau) = (0, l'_U, l'_V, s'_U, s'_V) | X(0) = (i_0, l_U, l_V, s_U, s_V))$, for states $(0, l'_U, l'_V, s'_U, s'_V) \in l(0)$ and $(i_0, l_U, l_V, s_U, s_V) \in \mathcal{S}_T$, can be expressed in terms of

$$\mathbf{f}(t; \mathcal{S}_A) = \pi_T \left(\mathbf{I}_{J_T} - e^{\mathbf{Q}_{T,T} t} \right) \left(-\mathbf{Q}_{T,T}^{-1} \right) \mathbf{Q}_{T,A}, \quad t > 0,$$

from which it follows that

$$\mathbf{f}(\infty; \mathcal{S}_A) = \pi_T \left(-\mathbf{Q}_{T,T}^{-1} \right) \mathbf{Q}_{T,A} \quad (3)$$

contains hitting probabilities of \mathcal{X} , which permit to describe the population of susceptible and latently infected TB hosts at the end of the outbreak.

The discussion which follows in Sections 3.1 and 3.2 is based on a detailed analysis of the first-passage time τ to the level $l(0)$, starting from an initial state $(i_0, l_U, l_V, s_U, s_V) \in \mathcal{S}_T$, whereas in Section 3.3 the focus is on an invasion time and the initial phase of the epidemic.

3.1. Extreme values during an outbreak

The peak of TB infection is commonly analyzed by using the maximum number $I_{\max} = \max\{I(t) : t \in [0, \tau)\}$ of hosts who are simultaneously infected with active TB at any time of an outbreak. A value of I_{\max} is not specific to an outbreak, and large (respectively, small) values of I_{\max} are not necessarily linked to large (respectively, small) values of the expected length $E[\tau | X(0) = (i_0, l_U, l_V, s_U, s_V)]$ of the outbreak. Thus, from the public health perspective, reducing peak incidence, regardless of the random length of the outbreak, is important if planning capacity to respond to a potential TB epidemic through intervention programs.

Under the assumption of $I(0) = i_0$ actively infected TB hosts, $L_U(0) = l_U$ and $L_V(0) = l_V$ unvaccinated and vaccinated infected hosts with latent bacilli, and $S_U(0) = s_U$ and $S_V(0) = s_V$ unvaccinated and vaccinated susceptible hosts, for $(i_0, l_U, l_V, s_U, s_V) \in \mathcal{S}_T$, the probability law of I_{\max} can be characterized by the conditional probabilities

$$P(I_{\max} \leq i | X(0) = (i_0, l_U, l_V, s_U, s_V)), \quad i \in \{0, \dots, N\},$$

which are equal to 0 if $i \in \{0, \dots, i_0 - 1\}$; in the case of an integer $i \in \{i_0, \dots, N\}$, $P(I_{\max} \leq i | X(0) = (i_0, l_U, l_V, s_U, s_V))$ is equivalent to the probability that, starting from the transient state $(i_0, l_U, l_V, s_U, s_V)$, the process \mathcal{X} reaches states in level $l(0)$, but avoiding states in the subclass $\cup_{i'=i+1}^N l(i')$ of transient states.

In the framework of LD-QBD processes (Amador and Gómez-Corral [6]), the analytical derivation of $P(I_{\max} \leq i | X(0) = (i_0, l_U, l_V, s_U, s_V))$, for integers $i \in \{i_0, \dots, N\}$, amounts to the dynamics of an auxiliary absorbing LD-QBD process $\mathcal{X}(i)$ taking values in the state space

$$\mathcal{S}(i) = \{0\} \cup \bigcup_{i'=1}^i l(i') \cup \{i+1\},$$

where 0 and $i+1$ are absorbing states and states in $\cup_{i'=1}^i l(i')$ are assumed to be transient; more particularly, states 0 and $i+1$ are obtained, respectively, by lumping together those states in the class \mathcal{S}_A of absorbing states and the subclass $\cup_{i'=i+1}^N l(i')$ of transient states, which are all accessible from states in $\cup_{i'=1}^i l(i')$. Specifically, the infinitesimal generator $\mathbf{Q}(i)$ of $\mathcal{X}(i)$ has the block-structured form

$$\mathbf{Q}(i) = \begin{pmatrix} 0 & \mathbf{0}_{J(i)}^T & 0 \\ \mathbf{t}_0(i) & \mathbf{T}(i) & \mathbf{t}_{i+1}(i) \\ 0 & \mathbf{0}_{J(i)}^T & 0 \end{pmatrix},$$

where $J(i) = \sum_{i'=1}^i J'(i')$ denotes the cardinality of $\cup_{i'=1}^i l(i')$, $\mathbf{T}(i)$ is obtained from the sub-matrix $\mathbf{Q}_{T,T}$ by removing rows and columns associated with states in $\cup_{i'=i+1}^N l(i')$, and the column vectors $\mathbf{t}_0(i)$ and $\mathbf{t}_{i+1}(i)$ are given by

$$\mathbf{t}_0(i) = \begin{pmatrix} \mathbf{A}_{1,0} \mathbf{1}_{J'(0)} \\ \mathbf{0}_{J(i)-J'(1)} \end{pmatrix},$$

$$\mathbf{t}_{i+1}(i) = \begin{pmatrix} \mathbf{0}_{J(i-1)} \\ \mathbf{A}_{i,i+1} \mathbf{1}_{J'(i+1)} \end{pmatrix}.$$

In a similar manner to Amador et al. [5], Theorem 1 may be used to design an algorithm for the calculation of the probability law of I_{\max} . Its proof is mainly based on the fact that eigenvalues of $\mathbf{T}(i)$ have strictly negative real part and the equality $\int_0^\infty e^{\mathbf{T}(i)u} \mathbf{t}_0(i) du = -\mathbf{T}^{-1}(i) \mathbf{t}_0(i)$, since the sub-matrix $\mathbf{T}(i)$ is stable.

Theorem 1 *For a predetermined integer $i \in \{i_0, \dots, N\}$, the vector $\mathbf{p}(i)$ with entries $P(I_{\max} \leq i | X(0) = (i_0, l_U, l_V, s_U, s_V))$, for states $(i_0, l_U, l_V, s_U, s_V) \in \mathcal{S}_T$, has the form $\mathbf{p}(i) = -\mathbf{T}^{-1}(i) \mathbf{t}_0(i)$ and, starting from $-\mathbf{T}^{-1}(1) = -\mathbf{A}_{1,1}^{-1}$, it can be evaluated from the recursive expression*

$$\mathbf{p}(i) = \begin{pmatrix} (\mathbf{I}_{J(i)} + \mathbf{V}_{1,2}(i) \mathbf{U}_{2,1}(i)) \mathbf{p}(i-1) \\ \mathbf{V}_{2,2}(i) \mathbf{U}_{2,1}(i) \mathbf{p}(i-1) \end{pmatrix},$$

where $\mathbf{U}_{2,1}(i) = (\mathbf{0}_{J'(i) \times J(i-2)}, \mathbf{A}_{i,i-1})$. We further have that sub-matrices $\mathbf{V}_{1,2}(i)$ and $\mathbf{V}_{2,2}(i)$ are given by

$$\begin{aligned} \mathbf{V}_{1,2}(i) &= -\mathbf{T}^{-1}(i-1) \mathbf{U}_{1,2}(i) \mathbf{V}_{2,2}(i), \\ \mathbf{V}_{2,2}(i) &= (-\mathbf{A}_{i,i} - \mathbf{U}_{2,1}(i) (-\mathbf{T}^{-1}(i-1)) \mathbf{U}_{1,2}(i))^{-1}, \end{aligned}$$

where $\mathbf{U}_{1,2}(i) = \begin{pmatrix} \mathbf{0}_{J(i-2) \times J'(i)} \\ \mathbf{A}_{i-1,i} \end{pmatrix}$, and verify

$$-\mathbf{T}^{-1}(i) = \begin{pmatrix} \mathbf{V}_{1,1}(i) & \mathbf{V}_{1,2}(i) \\ \mathbf{V}_{2,1}(i) & \mathbf{V}_{2,2}(i) \end{pmatrix}, \quad (4)$$

with $\mathbf{V}_{1,1}(i) = -\mathbf{T}^{-1}(i-1)(\mathbf{I}_{J(i-1)} + \mathbf{U}_{1,2}(i) \mathbf{V}_{2,1}(i))$ and $\mathbf{V}_{2,1}(i) = \mathbf{V}_{2,2}(i) \mathbf{U}_{2,1}(i) (-\mathbf{T}^{-1}(i-1))$.

The detailed algorithmic solution (Algorithm 1) is specified in Appendix B; note that, since $\mathbf{T}(N) = \mathbf{Q}_{T,T}$, the expected length of an outbreak and hitting probabilities of \mathcal{X} in Step 2 of Algorithm 1 are routinely derived from Eq. (4).

3.2. Cumulative number of vaccinations

Our objective here is to characterize the probability law of the cumulative number V of vaccinations during an outbreak by deriving recurrence equations for the conditional probabilities

$$P_{(i_0, l_U, l_V, s_U, s_V)}(v) = P(V = v | X(0) = (i_0, l_U, l_V, s_U, s_V)), \quad v \in \mathbb{N}_0,$$

for a fixed state $(i_0, l_U, l_V, s_U, s_V) \in \mathcal{S}_T$.

To that end, we first distinguish between those events involving a new vaccination –which are related to transitions M, O, Q, S and U (Figure 1 and Table 1)– and those which do not. Based on the fact that the transition M is linked to jumps of \mathcal{X} from states in level $l(i)$ to states in $l(i-1)$, and transitions O, Q, S and U are linked to jumps from states in $l(i)$ to states in $l(i)$, we split sub-matrices $\mathbf{A}_{i,i-1}$ and $\mathbf{A}_{i,i}$ as

$$\begin{aligned} \mathbf{A}_{i,i-1} &= \mathbf{A}_{i,i-1}^V + \mathbf{A}_{i,i-1}^U, \\ \mathbf{A}_{i,i} &= \mathbf{A}_{i,i}^* + \mathbf{A}_{i,i}^V + \mathbf{A}_{i,i}^U, \end{aligned}$$

for integers $i \in \{1, \dots, N\}$, where $\mathbf{A}_{i,i-1}^V$ and $\mathbf{A}_{i,i-1}^U$ record infinitesimal rates associated with the transition M and other transitions, respectively, from states in $l(i)$ to states in $l(i-1)$; $\mathbf{A}_{i,i}^V$ and $\mathbf{A}_{i,i}^U$ consist of infinitesimal rates associated with transitions O, Q, S and U, and other transitions, respectively, between states of $l(i)$; and $\mathbf{A}_{i,i}^*$ is the resulting diagonal matrix.

Then, by conditioning on the state visited by the process \mathcal{X} in its first transition, it is seen that the column vectors $\mathbf{p}(v; i)$ with entries $P(V =$

$v|X(0) = (i, l_U, l_V, s_U, s_V)$, for states $(i, l_U, l_V, s_U, s_V) \in l(i)$ with $i \in \{1, \dots, N\}$, satisfy the following equations:

$$- (\mathbf{A}_{1,1}^* + \mathbf{A}_{1,1}^U) \mathbf{p}(v; 1) = (\delta_{0,v} \mathbf{A}_{1,0}^U + \delta_{1,v} \mathbf{A}_{1,0}^V) \mathbf{1}_{\binom{N+3}{3}} + (1 - \delta_{0,v}) \mathbf{A}_{1,1}^V \mathbf{p}(v-1; 1) + \mathbf{A}_{1,2} \mathbf{p}(v; 2), \quad (5)$$

$$- (\mathbf{A}_{i,i}^* + \mathbf{A}_{i,i}^U) \mathbf{p}(v; i) = (1 - \delta_{0,v}) \mathbf{A}_{i,i-1}^V \mathbf{p}(v-1; i-1) + \mathbf{A}_{i,i-1}^U \mathbf{p}(v; i-1) + (1 - \delta_{0,v}) \mathbf{A}_{i,i}^V \mathbf{p}(v-1; i) + \mathbf{A}_{i,i+1} \mathbf{p}(v; i+1), \quad i \in \{2, \dots, N-1\}, \quad (6)$$

$$- (\mathbf{A}_{N,N}^* + \mathbf{A}_{N,N}^U) \mathbf{p}(v; N) = (1 - \delta_{0,v}) \mathbf{A}_{N,N-1}^V \mathbf{p}(v-1; N-1) + \mathbf{A}_{N,N-1}^U \mathbf{p}(v; N-1) + (1 - \delta_{0,v}) \mathbf{A}_{N,N}^V \mathbf{p}(v-1; N), \quad (7)$$

for integers $v \in \mathbb{N}_0$, where $\delta_{a,b}$ denotes the Kronecker delta.

In solving Eqs. (5)-(7), a simple and efficient iterative procedure can be derived by using block-Gaussian elimination, which allows us to write vectors $\mathbf{p}(v; i)$, for integers $i \in \{1, \dots, N\}$, in terms of $\mathbf{p}(v-1; i)$. This is summarized below; see also Algorithm 2 (Appendix C).

Theorem 2: *For integers $v \in \mathbb{N}_0$, the vectors $\{\mathbf{p}(v; i) : i \in \{1, \dots, N\}\}$ satisfy the recurrence equations*

$$\mathbf{p}(v; i) = \mathbf{a}(v; i) - \mathbf{W}_i^{-1} \mathbf{A}_{i,i+1} \mathbf{p}(v; i+1), \quad i \in \{1, \dots, N-1\}, \quad (8)$$

$$\mathbf{p}(v; N) = \mathbf{a}(v; N), \quad (9)$$

with $\mathbf{W}_i = \mathbf{A}_{i,i}^* + \mathbf{A}_{i,i}^U + (1 - \delta_{1,i}) \mathbf{A}_{i,i-1}^U (-\mathbf{W}_{i-1}^{-1}) \mathbf{A}_{i-1,i}$ and

$$\begin{aligned} \mathbf{a}(v; i) = & -\mathbf{W}_i^{-1} \left(\delta_{1,i} (\delta_{0,v} \mathbf{A}_{i,i-1}^U + \delta_{1,v} \mathbf{A}_{i,i-1}^V) \mathbf{1}_{\binom{N+3}{3}} \right. \\ & + (1 - \delta_{1,i}) (\mathbf{A}_{i,i-1}^U \mathbf{a}(v; i-1) + (1 - \delta_{0,v}) \mathbf{A}_{i,i-1}^V \mathbf{p}(v-1; i-1)) \\ & \left. + (1 - \delta_{0,v}) \mathbf{A}_{i,i}^V \mathbf{p}(v-1; i) \right). \end{aligned}$$

In a similar manner to (8)-(9), recurrence equations for the column vectors $\mathbf{v}^{(k)}(i)$ with entries

$$E[V(V-1)\cdots(V-k+1)|X(0) = (i, l_U, l_V, s_U, s_V)],$$

for integers $k \in \mathbb{N}$, states $(i, l_U, l_V, s_U, s_V) \in l(i)$ and $i \in \{1, \dots, N\}$, can be derived from Eqs. (5)-(7) by introducing generating functions and taking derivatives on the resulting equations; see Algorithm 3 (Appendix C).

3.3. Exact reproduction number

In this section, the interest is in an invasion time; i.e., $X(0) = (1, 0, 0, N-1, 0)$. The aim is to characterize the probability law of the exact number $\mathcal{R}_{exact,0}$ of secondary cases generated by the initially infected host with active TB before it recovers or dies, by evaluating the conditional probabilities

$$Q_{(1,0,0,N-1,0)}(r) = P(\mathcal{R}_{exact,0} = r | X(0) = (1, 0, 0, N-1, 0)), \quad r \in \mathbb{N}_0.$$

Clearly, $\mathcal{R}_{exact,0}$ is related to a random interval $(0, \tau')$, where τ' is an exponentially distributed random variable with mean $(\delta_R + \delta_{TB} + \delta_D)^{-1}$, whence $\mathcal{R}_{exact,0}$ allows us to analyze trends in the early phase of the disease since $\tau' < \tau$. More generally, we define conditional probabilities $Q_{(i,l_U,l_V,s_U,s_V)}(r)$ that, provided that (i, l_U, l_V, s_U, s_V) is the *current* state of \mathcal{X} and the infectious period of the initially infected host with active TB is in process at time $t' < \tau'$, the initially infected host with active TB generates r secondary cases during the *residual* interval (t', τ') .

It is important to stress that, in our analysis, secondary cases contributing to $\mathcal{R}_{exact,0}$ are not associated with endogenous reactivation of latent bacilli –not even when the latent infection of a host had been caused by

the initially infected host with active TB—nor exogenous reinfection caused by contacts with other actively infected TB hosts. This implies that, under the assumption that the infectious period of the initially infected host with active TB is still in process, the reactivation/reinfection rates $r_U(i)$ and $r_V(i)$ of unvaccinated and vaccinated latently infected TB hosts can be appropriately decomposed into two contributions as follows:

$$\begin{aligned} r_U(i) &= \bar{r}_U(i) + \tilde{r}_U(i), \\ r_V(i) &= \bar{r}_V(i) + \tilde{r}_V(i), \end{aligned}$$

where the terms $\bar{r}_U(i) = b_U$ and $\bar{r}_V(i) = b_V$ are related to the exogenous reinfection caused by contacts between unvaccinated and vaccinated latently infected TB hosts and the initially infected host with active TB, and $\tilde{r}_U(i) = a_U + (i-1)b_U$ and $\tilde{r}_V(i) = a_V + (i-1)b_V$ correspond to either the endogenous reactivation of latent tubercle bacilli in unvaccinated and vaccinated latently infected hosts, or the exogenous reinfection caused by contacts between these hosts and other actively infected TB hosts, provided that i is the number of infected hosts with active TB.

The above comments also motivate us to distinguish between jumps of \mathcal{X} from states in level $l(i)$ to states in $l(i+1)$ which contribute to $\mathcal{R}_{exact,0}$ and those jumps which do not. This leads us to write down

$$\mathbf{A}_{i,i+1} = \bar{\mathbf{A}}_{i,i+1} + \tilde{\mathbf{A}}_{i,i+1},$$

for integers $i \in \{1, \dots, N-1\}$, where $\bar{\mathbf{A}}_{i,i+1}$ records infinitesimal rates associated with transitions D-G (Figure 1 and Table 1) caused by the initially infected host with active TB, and $\tilde{\mathbf{A}}_{i,i+1}$ consists of infinitesimal rates associated with transitions D-G caused by other actively infected TB hosts. In a similar manner, transitions H-M (Figure 1 and Table 1) correspond to

deaths and recoveries of actively infected TB hosts, from which sub-matrix $\mathbf{A}_{i,i-1}$ is specified by

$$\mathbf{A}_{i,i-1} = \overline{\mathbf{A}}_{i,i-1} + \tilde{\mathbf{A}}_{i,i-1},$$

for integers $i \in \{1, \dots, N\}$, where $\overline{\mathbf{A}}_{i,i-1}$ is associated with the recovery and death of the initially infected host with active TB, and $\tilde{\mathbf{A}}_{i,i-1}$ records infinitesimal rates of other jumps of \mathcal{X} from states in level $l(i)$ to states in $l(i-1)$.

By conditioning on the first passage from states in level $l(i)$ to states in levels $l(i')$, for $i' \in \{i-1, i, i+1\}$, under the assumption that the infectious period of the initially infected host with active TB is in process, it is found that the column vectors $\mathbf{q}(r; i)$ with entries $Q_{(i, l_U, l_V, s_U, s_V)}(r)$, for states $(i, l_U, l_V, s_U, s_V) \in l(i)$, $i \in \{1, \dots, N\}$ and $r \in \mathbb{N}_0$, satisfy

$$-\mathbf{A}_{1,1}\mathbf{q}(r; 1) = \delta_{0,r}\overline{\mathbf{A}}_{1,0}\mathbf{1}_{\binom{N+3}{3}} + (1 - \delta_{0,r})\overline{\mathbf{A}}_{1,2}\mathbf{q}(r-1; 2) + \tilde{\mathbf{A}}_{1,2}\mathbf{q}(r; 2), \quad (10)$$

$$\begin{aligned} -\mathbf{A}_{i,i}\mathbf{q}(r; i) &= \delta_{0,r}\overline{\mathbf{A}}_{i,i-1}\mathbf{1}_{\binom{N-i+4}{3}} + \tilde{\mathbf{A}}_{i,i-1}\mathbf{q}(r; i-1) \\ &\quad + (1 - \delta_{0,r})\overline{\mathbf{A}}_{i,i+1}\mathbf{q}(r-1; i+1) + \tilde{\mathbf{A}}_{i,i+1}\mathbf{q}(r; i+1), \\ &\quad i \in \{2, \dots, N-1\}, \end{aligned} \quad (11)$$

$$-\mathbf{A}_{N,N}\mathbf{q}(r; N) = \delta_{0,r}\overline{\mathbf{A}}_{N,N-1}\mathbf{1}_{\binom{4}{3}} + \tilde{\mathbf{A}}_{N,N-1}\mathbf{q}(r; N-1), \quad (12)$$

for integers $r \in \mathbb{N}_0$. We may then proceed as in Section 3.2 and apply block-Gaussian elimination to the above equations. This yields the following result and Algorithm 4 (Appendix D).

Theorem 3 *For integers $r \in \mathbb{N}_0$, the vectors $\{\mathbf{q}(r; i) : i \in \{1, \dots, N\}\}$ are given by*

$$\mathbf{q}(r; i) = \mathbf{c}(r; i) - \widehat{\mathbf{W}}_i^{-1}\tilde{\mathbf{A}}_{i,i+1}\mathbf{q}(r; i+1), \quad i \in \{1, \dots, N-1\},$$

$$\mathbf{q}(r; N) = \mathbf{c}(r; N),$$

where $\widehat{\mathbf{W}}_i = \mathbf{A}_{i,i} + (1 - \delta_{1,i})\widetilde{\mathbf{A}}_{i,i-1}(-\widehat{\mathbf{W}}_{i-1}^{-1})\widetilde{\mathbf{A}}_{i-1,i}$, and

$$\begin{aligned} \mathbf{c}(r; i) = & -\widehat{\mathbf{W}}_i^{-1} \left(\delta_{0,r} \overline{\mathbf{A}}_{i,i-1} \mathbf{1}_{(N-i+4)} + (1 - \delta_{1,i}) \widetilde{\mathbf{A}}_{i,i-1} \mathbf{c}(r; i-1) \right. \\ & \left. + (1 - \delta_{0,r})(1 - \delta_{N,i}) \overline{\mathbf{A}}_{i,i+1} \mathbf{q}(r-1; i+1) \right). \end{aligned}$$

In a similar manner to Algorithm 3, it is a simple matter, using Eqs. (10)-(12), to derive recurrence equations for the k th factorial moment $\overline{\mathcal{R}}_{exact,0}^{(k)} = E[\mathcal{R}_{exact,0}(\mathcal{R}_{exact,0} - 1) \cdots (\mathcal{R}_{exact,0} - k + 1) | X(0) = (1, 0, 0, N - 1, 0)]$, for integers $k \in \mathbb{N}$; see Algorithm 5 (Appendix D).

4. Numerical experiments

We carry out numerical examples to exemplify analytical results in Section 3 as well as to show how, in the setting of a mini-community design, the process \mathcal{X} may serve as a baseline for evaluating the prospective impact of pre-exposure vaccines. In a hypothetical case study (Table 3), twenty 30-35 aged individuals are recruited into the study under the assumption that they have not been previously vaccinated against TB and are initially free of tubercle bacilli, with the exception of one of them who is actively infected. Every individual, independently of his/her status, is assumed to die due to causes beyond TB or to abandon the study at rate $\delta_D = 35^{-1}$ *per year*; in the case of an actively infected TB individual, the rate $\delta_{TB} = 5^{-1}$ *per year* is related to the death due to TB infection (Renardy and Kirschner [47]). Individuals who die or abandon the study are replaced instantly by new individuals who are not colonized with tubercle bacilli.

[Insert here Table 3]

Table 3: Parameters, values and related references.

The prophylactic use of a vaccine is related to the fact that only newly admitted individuals can be vaccinated. More particularly, a newly admitted individual is labelled as either *vaccinated* or *unvaccinated* susceptible with probabilities qC and $1 - qC$, respectively; i.e., we let $\eta(q, C)$ be defined by qC . The objective is to analyze the impact of three vaccines, which are assumed to be 100 percent effective (Knight et al. [32]) and uniquely characterized by the value of take q varying among 0.4 (Scenario 1), 0.6 (Scenario 2) and 0.8 (Scenario 3), and their protection may wane over time after a mean duration of $\gamma^{-1} = 10$ years. In our examples, we consider coverage levels $C \in \{0.2, 0.5, 0.8\}$ to predict the joint impact of pair (q, C) on the spread of bacilli.

Susceptible individuals become infected with tubercle bacilli due to infectious respiratory contacts with actively infected TB individuals, in such a way that one effective respiratory contact per week is assumed to occur and, in accordance to the World Health Organization [54], only in 10% of these respiratory contacts bacilli are transmitted in sufficient numbers to infect a susceptible individual who is in close contact with an actively infected TB individual. To be consistent with new vaccines against TB (Méndez-Samperio [42]), vaccines prevent active TB, but not latent infection and, regardless of the status of susceptible individuals involved in infectious respiratory contacts, bacilli settle into the organism at rates $\alpha_U = \alpha_V = 5.2$ *per year*. The probability of a fast progression to active TB is specified by $p_U = 0.15$ on unvaccinated susceptible individuals; note that, by the perfect

[Insert here Table 4]

Table 4: Expectations of $\mathcal{R}_{exact,0}$, τ , I_{\max} and V at an invasion time versus $C \in \{0.2, 0.5, 0.8\}$, for values $q = 0.4$ (Scenario 1), 0.6 (Scenario 2) and 0.8 (Scenario 3).

protection of vaccines against TB, we consider the value $p_V = 0.0$.

Latent bacilli on an unvaccinated infected individual become active due to either endogenous reactivation at rate $a_U = 1.5 \times 10^{-4}$ *per year* (Knight et al. [32]) or exogenous reinfection at rate b_U *per year*. Latent infection is assumed (Knight et al. [32]) to provide some degree of protection in further infectious respiratory contacts. Specifically, latent infection protects an individual in 65% of infectious respiratory contacts and, consequently, we have that $r_U(i) = a_U + b_U i$ with $b_U = 0.35a_U$, provided that the compartment of actively infected TB hosts consists of i individuals. Since vaccines are assumed to provide perfect protection, vaccinated infected individuals with latent TB cannot suffer from neither endogenous reactivation nor exogenous reinfection and, consequently, we have that $a_V = b_V = 0.0$; as a result, $r_V(i) = 0$.

According to the average duration of TB treatment (World Health Organization [54]), actively infected TB individuals recover after $\delta_R^{-1} = 9$ *months*—on average—and become latently infected almost surely; i.e., $\xi = 1.0$. The immunity structure in subsequent infections suffered by recovered individuals is not well documented (Cliff et al. [18]), whence the selection $\theta_U = 0.5$ in our examples is merely illustrative; since no recovered individuals become susceptible, there is no need for any specification of $\tilde{\theta}_U$.

4.1. The exact reproduction number and the length of an outbreak

In Table 4, we list values $E[\mathcal{R}_{exact,0}|X(0) = (1, 0, 0, 19, 0)]$, $E[\tau|X(0) = (1, 0, 0, 19, 0)]$, $E[I_{\max}|X(0) = (1, 0, 0, 19, 0)]$ and $E[V|X(0) = (1, 0, 0, 19, 0)]$, for coverage levels $C \in \{0.2, 0.5, 0.8\}$ in Scenarios 1-3. It is observed that, since $E[\mathcal{R}_{exact,0}|X(0) = (1, 0, 0, 19, 0)] \in (3.2, 3.5)$ in our examples, the TB disease will start spreading in the community, apparently, with similar propagation *speed*. However, vaccines in Scenarios 1-3 will not lead to similar outbreaks. In particular, it is seen that the expected length of the outbreak depends strongly on the selection of pair (q, C) . As intuition tells us, more efficient vaccines (i.e., increasing values of q) and/or increasing coverage levels C lead to decreasing values of the expected length of the outbreak, with the *longest* and the *shortest* duration of the outbreak for selections $(q, C) = (0.4, 0.2)$ and $(q, C) = (0.8, 0.8)$, respectively, in our examples.

A particularly appealing feature of our Markov chain model for TB is related to the certain extinction of the disease in a finite time, regardless of the initial state and the underlying parameters. This property contrasts with the fact that, for our selection of parameters and initial condition, the analogous deterministic model converges to an endemic equilibrium point $(i^*, l_U^*, l_V^*, s_U^*, s_V^*)$; i.e., with $i^* > 0$. Therefore, the evaluation of our descriptors in Subsections 4.2-4.4 is not plausible in the case of the ODE model, since in the numerical study the TB disease is not eradicated and remains in the population for the analogous ODE model.

4.2. The peak of infection

It is seen that, in a similar manner to the exact reproduction number, values of $E[I_{\max}|X(0) = (1, 0, 0, 19, 0)]$ in Table 4 do not change notably as a function of pair (q, C) . Thus, expected values of I_{\max} do not appear to

[Insert here Figure 2]

Figure 2: Mass function of I_{\max} , provided that $X(0) = (1, 0, 0, 19, 0)$, versus $C \in \{0.2, 0.5, 0.8\}$, for values $q = 0.4$ (Scenario 1), 0.6 (Scenario 2) and 0.8 (Scenario 3).

play a significant role to prevent a high degree of TB colonization, but it is observed that they decrease when the vaccine effectiveness increases and/or with increasing values of the coverage level C . In Figure 2, the probability law of I_{\max} is seen to be a bimodal distribution with two peaks (or modes) with different heights. The first peak amounts to the smallest value of I_{\max} (i.e., the event $\{I_{\max} = 1\}$), regardless of q and C , which does essentially correspond to the recovery of the initially infected individual with active TB before spreading the disease to other individuals; in this case, the effectiveness of a vaccine on the dynamics of the outbreak is not relevant and, consequently, the height of the first peak in Figure 2 remains almost constant. The second peak is related to a larger degree $i \in \{15, \dots, 18\}$ of TB colonization, and decreases when the vaccine effectiveness increases and/or with increasing coverage levels C . This means that, when the initially infected individual with active TB spreads the disease to other individuals, a significant proportion of individuals will be simultaneously colonized with tubercle bacilli during the outbreak. The height of the second peak is notably higher than the height of the first one, but differences become smaller when the vaccine effectiveness increases and/or with increasing values of C . Indeed, as the take q and the coverage level C increase, the probability law of I_{\max} shows a more significant dispersion in terms of higher values for the number of actively infected TB individuals; for instance, probabilities associated with relevant numbers I_{\max} of simultaneously infected individuals

[Insert here Figure 3]

Figure 3: A graphical representation of the threshold η_{\max} , the resulting regions R_1 and $R_2 = R'_2 \cup R''_2$, and the pairs (q, C) for $q = 0.4$ (Scenario 1), 0.6 (Scenario 2) and 0.8 (Scenario 3), and $C \in \{0.2, 0.5, 0.8\}$.

with active TB are more significant in an outbreak with longer duration than their counterparts in an outbreak with shorter duration. Based on Figure 2 and numerical work not reported here, the graphical differences in the bimodal distribution of I_{\max} become more apparent with either smaller values of the take q and the coverage level C , or higher values of q and C .

4.3. The cumulative number of vaccinations

It is worth noting that, in Table 4, $E[V|X(0) = (1, 0, 0, 19, 0)]$ does not behave necessarily as a monotone function of q depending on the coverage level C . To be concrete, the expected value of V in the case $C = 0.2$ starts growing until it reaches its maximum value and then decreases when the effectiveness of the vaccine increases, while it always decreases when the vaccine effectiveness increases for selections $C \in \{0.5, 0.8\}$. This behaviour shows the existence of a certain threshold C' for the coverage level such that values $C > C'$ permit to reduce the number V of vaccinations by using a more effective vaccine, whereas coverage levels $C < C'$ do not allow to predict the number of vaccinations in terms of the take q of a vaccine.

In order to determine the threshold C' for the level of coverage C and an analogous threshold q' for the take q of a vaccine, it should be pointed out that those pairs $(q, C) \in (0, 1] \times (0, 1]$ satisfying $qC = \eta_0$, for a pre-defined probability $\eta_0 \in (0, 1]$, yield the same value for the expected number $E[V|X(0) = (1, 0, 0, 19, 0)]$ of vaccinations. This allows to decompose

the unit square $(0, 1] \times (0, 1]$ in the plane (q, C) into the regions $R_1 = \{(q, C) \in (0, 1] \times (0, 1] : \min\{q, C\} \leq \eta_{\max}\}$ and $R_2 = \{(q, C) \in (0, 1] \times (0, 1] : \min\{q, C\} > \eta_{\max}\}$, where the value $\eta_{\max} = 0.11457$ amounts to the probability $\eta(q, C)$ of vaccination that leads to the maximum value of $E[V|X(0) = (1, 0, 0, 19, 0)]$ within the unit square; see Figure 3. More concretely, the probability η_{\max} is valuated as $\eta(q_{\max}(C), C)$, where the take $q_{\max}(C) = 0.57287$ yields the maximum value of $E[V|X(0) = (1, 0, 0, 19, 0)]$ for the selected level of coverage $C = 0.2$, which is given by 93.26586. This implies that

$$\max\{E[V|X(0) = (1, 0, 0, 19, 0)] : (q, C) \in (0, 1] \times (0, 1]\} = 93.26586,$$

since every pair (q, C) satisfying $\eta(q, C) = \eta_{\max}$ yields this maximum value for the expected number of vaccinations during the outbreak.

Therefore, the expected number $E[V|X(0) = (1, 0, 0, 19, 0)]$ of vaccinations behaves as a non-monotone function of the probability $\eta(q, C)$ of vaccination. In particular, it is seen that $E[V|X(0) = (1, 0, 0, 19, 0)]$ grows with increasing values of $\eta(q, C)$ until it reaches its maximum value at point $\eta(q, C) = \min\{\eta_{\max}, 1\}$ (i.e., $(q, C) \in R_1 \cup R'_2$ with $R'_2 = \{(q, C) \in R_2 : \eta(q, C) \leq \eta_{\max}\}$), and it then decreases with increasing values of $\eta(q, C)$ satisfying $\eta_{\max} < \eta(q, C) \leq 1$ (i.e., $(q, C) \in R''_2$ with $R''_2 = \{(q, C) \in R_2 : \eta(q, C) > \eta_{\max}\}$). In terms of the take q and the level of coverage C , this can be translated into the following two characteristic behaviors:

- (i) For a fixed value C with $0 < C \leq \eta_{\max}$ (respectively, q with $0 < q \leq \eta_{\max}$), the expected number of vaccinations increases when the effectiveness of the vaccine (respectively, the level of coverage) increases, since the resulting pairs (q, C) belong to the region R_1 in Figure 3. Thus, more effective vaccination strategies –based on either a level of

coverage C with $0 < C \leq \eta_{\max}$ and/or a vaccine with take q satisfying $0 < q \leq \eta_{\max}^-$ will result in decreasing lengths of the outbreak and larger values for the expected number of vaccinations.

- (ii) For a fixed value C with $\eta_{\max} < C \leq 1$ (respectively, q with $\eta_{\max} < q \leq 1$), the expected number of vaccinations increases when, under the assumption that the take q (respectively, the level of coverage C) is not greater than η_{\max} , the effectiveness of the vaccine (respectively, the level of coverage) increases, whereas it decreases when $\eta_{\max} < q \leq 1$ (respectively, $\eta_{\max} < C \leq 1$); note that the former is related to pairs $(q, C) \in R'_2$ in Figure 3, and the latter is associated with pairs $(q, C) \in R''_2$. A smaller number of vaccinations and a smaller length of the outbreak will be then observed by implementing more effective vaccination strategies –i.e., by increasing the proportion C of individuals to be vaccinated (respectively, the take q) with $C > C'$ and $C' = q^{-1}\eta_{\max}$ (respectively, $q > q'$ and $q' = C^{-1}\eta_{\max}$)–, provided that the take of the vaccine and the level of coverage are both greater than η_{\max} .

These comments show how the dynamics of the pathogen may influence on unvaccinated individuals when a pre-exposure strategy is implemented and the take of the vaccine or the proportion of individuals to be vaccinated are not sufficiently large. For the sake of completeness, we mention here that, in our numerical examples, the thresholds are given by $C' = 0.28646$ in Scenario 1 ($q = 0.4$), 0.19095 in Scenario 2 ($q = 0.6$), and 0.14323 in Scenario 3 ($q = 0.8$), and $q' = 0.57287$ if $C = 0.2$, 0.22913 if $C = 0.5$, and 0.14323 if $C = 0.8$.

Figure 4 shows that the mass function of V corresponds to a bimodal

[Insert here Figure 4]

Figure 4: Mass function of V , provided that $X(0) = (1, 0, 0, 19, 0)$, versus $C \in \{0.2, 0.5, 0.8\}$, for values $q = 0.4$ (Scenario 1), 0.6 (Scenario 2) and 0.8 (Scenario 3). For convenience, the mass function of V has been plotted on a common set of integers $v \in \{0, \dots, v'\}$, with $v' \geq V_{q'}$, for all selection of the pair (q, C) and $q' = 0.99$.

[Insert here Table 5]

Table 5: Conditional expectations of the random numbers $L_U(\tau)$, $L_V(\tau)$, $S_U(\tau)$ and $S_V(\tau)$, provided that $X(0) = (1, 0, 0, 19, 0)$, versus $C \in \{0.2, 0.5, 0.8\}$, for values $q = 0.4$ (Scenario 1), 0.6 (Scenario 2) and 0.8 (Scenario 3).

distribution with a first peak within its smallest values, and a second peak in moderate/higher numbers, regardless of the pair (q, C) . These peaks are inherently linked to the existence of two different patterns: a fast recovery/death/abandonment of the initially infected individual with active TB, which leads to a small number of vaccinations; and a high degree of TB colonization, which results in a higher number of vaccinations as a consequence of an increasing number of newly admitted individuals. The height of the first peak decreases slightly when the vaccine effectiveness increases, and more apparently with increasing coverage levels C , whereas the height of the second peak behaves as an increasing function of q and/or C .

4.4. The final status of the disease

The status of the population at the end of an outbreak is illustrated in Table 5 in terms of the expectations of $L_U(\tau)$, $L_V(\tau)$, $S_U(\tau)$ and $S_V(\tau)$, provided that $X(0) = (1, 0, 0, 19, 0)$. In a similar manner to the expectation of I_{\max} (Table 4), it is noted that expected values of $L_U(\tau)$, $L_V(\tau)$, $S_U(\tau)$ and $S_V(\tau)$ are not strongly influenced by the effectiveness of the vaccine

[Insert here Figure 5]

Figure 5: Mass functions of $L_U(\tau)$ (dark green), $L_V(\tau)$ (light green), $S_U(\tau)$ (dark blue) and $S_V(\tau)$ (light blue), provided that $X(0) = (1, 0, 0, 19, 0)$, versus $C \in \{0.2, 0.5, 0.8\}$, for values $q = 0.4$ (Scenario 1), 0.6 (Scenario 2) and 0.8 (Scenario 3). For convenience, a polygonal line is used instead of the corresponding histogram.

and the coverage level. These expected values behave as a monotone function of q and C ; more particularly, the final numbers of latently infected TB and susceptible individuals who are vaccinated increase, whereas their unvaccinated counterparts decrease when the vaccine effectiveness increases and/or with increasing coverage levels C . This behaviour is intuitive since higher values of C and q are related to a more frequent use of more effective vaccines, being the subpopulation of vaccinated infected individuals with latent TB the most representative compartment at the end of the outbreak. This is closely related to the assumption that vaccines under study prevent endogenous reactivation and exogenous reinfection on latently infected TB individuals. Moreover, since the vaccine does not provide protection towards latent infection, the number of susceptible individuals at the end of the outbreak is, on average, significantly less than the number of infected individuals with latent TB.

The probability laws of $L_U(\tau)$, $L_V(\tau)$, $S_U(\tau)$ and $S_V(\tau)$ are shown in Figure 5 to have different shape. To be concrete, the mass function of $L_V(\tau)$ is a bimodal distribution with a first peak in its smallest value (i.e., the event $\{L_V(\tau) = 0\}$) and a second peak in an intermediate value $l_V \in \{13, \dots, 17\}$, which moves towards higher values l_V with increasing values of q and/or C . The probability law of $S_U(\tau)$ is a bimodal distribution with a first peak $s_U \in \{0, 1, 2\}$, which moves from higher to smaller values with increasing values

of q and/or C , and a second peak in higher values $s_U \in \{18, 19, 20\}$. The height of the first peak is significantly more representative than the height of the second one, which is related to a fast recovery/death/abandonment of the initially infected individual with active TB. The probability laws of $L_U(\tau)$ and $S_V(\tau)$ are seen to be unimodal distributions, whence $L_U(\tau)$ and $S_V(\tau)$ are not seriously affected by the occurrence of a fast end of the outbreak nor a high degree of TB colonization. It is remarkable to note that the mass function of $S_V(\tau)$ has a significant peak at value $s_V = 0$ in the case of an outbreak with longer duration, whereas this peak becomes less representative with increasing values of q and C —which yield an outbreak with shorter duration—, even changing its structural shape.

5. Conclusions

In this paper, we have proposed and analyzed a Markov chain model to study the dynamics of TB in the context of small communities of hosts sharing confined spaces, as well as to evaluate the prospective use of pre-exposure vaccines. Child protection centers, long-stay hospital units, and households or other small transmission units in a mini-community design are examples of small communities where random effects shall be accounted for under the assumption of a homogeneously well-mixed population of hosts. In constructing the model, the susceptible and latently infected TB subpopulations have been divided into unvaccinated and vaccinated subclasses, resulting in five compartments by adding the subpopulation of infected hosts with active TB. Distributional assumptions were based on the deterministic model of Lietman and Blower [37] (Castillo-Chavez and Song [16, Section 8.1]), allowing us to incorporate exogenous reinfection and endogenous reactivation

in disease progression from latent to active TB, among other features.

The Markov chain model has been formulated as an absorbing level-dependent quasi-birth-death process, whence the random length τ of an outbreak has been thought of as a phase-type random variable (Section 3). Extreme value properties (Section 3.1), the cumulative number V of vaccinations (Section 3.2) and a random version $\mathcal{R}_{exact,0}$ of the basic reproduction number (Section 3.3) have been analytically derived by using absorption probabilities and first-passage times in suitably defined quasi-birth-death processes, and Algorithms 1-5 have been inspired from Gaussian elimination (Latouche and Ramaswami [35, Chapter 10]), which is applicable in all descriptors of Sections 3.1-3.3 since underlying rate matrices are band matrices. The analytical approach is inherently linked to the use of block-structured Markov processes (Baumann and Sandmann [10]; Lefèvre and Simon [36]) and phase-type random variables (Neuts [46]) in epidemic modelling (Almaraz and Gómez-Corral [4]; Amador et al. [5]; Gómez-Corral and López-García [28]).

In the setting of a mini-community design, we have exemplified the analytical solution for a hypothetical case study. Parameters used in our examples have been selected from the existing literature and available information on TB vaccines in development; see e.g. Knight et al. [32], Méndez-Samperio [42], Renardy and Kirschner [47], and World Health Organization [54, 55]. It has been seen that the exact reproduction number $\mathcal{R}_{exact,0}$ is not very informative when comparing vaccines at the initial phase of the epidemic. This is closely related to the fact that, at an invasion time, the probability that a new vaccination occurs during the *generalized* infectious period of the initially infected individual with active TB is not significant, since the underlying event –which should occur during a random interval of expected

length $(\delta_R + \delta_{TB} + \delta_D)^{-1}$ involves the death or abandonment of an individual and the replacement by a new individual who will be vaccinated with probability qC , but who may also not be vaccinated with probability $1 - qC$. However, the use of vaccines (not necessarily the most effective one), appropriately combined with increasing levels of coverage, has been found to have a significant impact on the duration of the outbreak. It has been noted that the probability law of certain descriptors, such as I_{\max} and V , has a bimodal shape. This property of the underlying mass function constitutes the most notable difference between these stochastic descriptors and their deterministic counterparts, and is closely related to the recovery of the initially infected individual with active TB before the disease spreads to other susceptible individuals. In the opposite case, the peak of infection measured in terms of I_{\max} will lead to a high degree of colonization with tubercle bacilli with a significant probability.

It should be pointed out that the disease begins to disappear gradually, towards its extinction, from the time τ_{\max} when the maximum number I_{\max} of simultaneously infected hosts with active TB is reached. At time τ_{\max} , the population can be said to have achieved herd immunity to the disease, since the proportion of vaccinated (susceptible or latently infected TB) hosts at time τ_{\max} will guarantee, almost surely, the extinction of the disease over time. Figure 2 could thus be used to have a preliminary relationship between herd immunity and the maximum number I_{\max} of infected hosts with active TB, but there is clearly future work to be done on the joint distribution of the random vector (I_{\max}, τ_{\max}) and the probability law of the proportion $(L_V(\tau_{\max}), S_V(\tau_{\max}))$ of vaccinated hosts yielding herd immunity to the tubercle bacilli.

The most interesting result in our examples is related to the behavior

of V depending on the efficacy of a vaccine and the level of coverage. It has been found that the take q of a vaccine greater than η_{\max} implies the existence of the threshold $C' = q^{-1}\eta_{\max}$ for the level of coverage such that a higher proportion C of individuals to be vaccinated with $C > C'$ will result in a smaller expected number of vaccinations during the outbreak. A similar threshold $q' = C^{-1}\eta_{\max}$ for the take q is obtained when the level of coverage C is greater than η_{\max} , which allows us to assert that the expected number of vaccinations decreases when the effectiveness of a vaccine increases with the take q satisfying $q > q'$. On the contrary, the use of a vaccine with the take $q \leq \eta_{\max}$ (respectively, a level of coverage $C \leq \eta_{\max}$) does not permit to reduce the expected number of vaccinations by increasing the level of coverage (respectively, by using a more effective vaccine). It is likely that in this way pre-exposure vaccination strategies could be improved in the particular setting of small host communities.

It is important to clarify that the aforementioned behavior –regarding the number of vaccinations– is inherently related to the assumption that only newly admitted hosts can be vaccinated when other hosts die in our model. Therefore, the reader is alerted to the fact that there is no evidence that our conclusions can be directly extended to a more general context, and future work must be addressed, for example, in the case of births and neonatal use of the vaccine.

Our Markov chain model for TB is, in some aspects, unrealistic. For example, age of infection as well as chronological age are critical factors in disease progression and, however, the process \mathcal{X} did not take into account. Models that incorporate age structure and, more generally, population structure play a significant role in the design of new vaccines against TB but may not be numerically tractable by using an extension of \mathcal{X} . Some of our earlier

studies address these issues in the context of SIS-like epidemic models on networks without vaccination (Economou et al. [21]), showing that the analytical treatment of *exact* Markov chain models will have a more theoretical rather than practical motivation in the case of population sizes $N \geq 30$; for a related work, see also the paper by López-García [40] on SIR-like epidemic networks. In a similar manner to network dynamics, the state space \mathcal{S} of \mathcal{X} grows dramatically with the population size N , so can be extremely large for even fairly small populations. Another aspect of interest is how the size N of the population influences the computational efficiency of our algorithmic solutions. More concretely, Table 6 shows the effect that the population size N , the number $J'(i)$ of states in the i th level $l(i)$, and the cardinality $J(i)$ of the subset $\cup_{i'=1}^i l(i')$ of states, for integers $i \in \{1, \dots, N\}$, has on the computational complexity of Algorithms 1-5. It is observed that Algorithms 2 and 4 (respectively, Algorithms 3 and 5) have complexities similar to each other, because the most demanding effort resides in the inversion of matrices with dimension $J'(i)$ and the product between matrices with dimension $J'(i) \times J'(i')$ and column vectors of order $J'(i')$, with $i' \in \{i-1, i, i+1\}$, for $i, i' \in \{1, \dots, N\}$. It is observed that Algorithm 1 has computational complexity similar to the computational complexity of the procedure for computing the inverse $-\mathbf{T}^{-1}(N)$. Note that the matrix $\mathbf{T}(N)$ is a square matrix of order $J_T = 8855$ in our numerical experiments (i.e., with $N = 20$) and, for example, this order becomes $J_T = 4421275$ in the case $N = 100$. This makes the memory requirements of Algorithm 1, but also of Algorithms 2-5 –especially for storing auxiliary matrices– very demanding even for moderate values of N . Therefore, one aspect that deserves further exploration is how to reduce, for moderate/large population sizes N , the state-space size through graph-automorphism lumping (Ward and Evans [51]) by exploiting

[Insert here Table 6]

Table 6: Computational complexities of Algorithms 1-5.

symmetries of the resulting contact network (Ward and López-García [52]), although it is not clear what features of the population-structured vaccination model may generally ensure this.

Acknowledgements. – The authors thank the editor and two anonymous referees for their constructive comments, which have improved the presentation of the paper. This research was supported by the Government of Spain (Ministry of Science and Innovation), project PGC2018-097704-B-I00 (A. Gómez-Corral), and was carried out while the first author (R. Fernández-Peralta) was visiting the Institute of Mathematical Sciences with a grant “Excelencia Severo Ochoa-CSIC” for master students.

Appendix A. Sub-matrices $\mathbf{A}_{i,i'}$, for $i' \in \{i-1, i, i+1\}$

We construct sub-matrices $\mathbf{A}_{i,i'}$, for integers $i' \in \{i-1, i, i+1\}$, in Eqs. (1)-(2) by using the following ordering for states $(i, l_U, l_V, s_U, s_V), (i', l'_U, l'_V, s'_U, s'_V) \in \mathcal{S}$ with $(i', l'_U, l'_V, s'_U, s'_V) \neq (i, l_U, l_V, s_U, s_V)$:

$$\begin{aligned}
 (i, l_U, l_V, s_U, s_V) \prec (i', l'_U, l'_V, s'_U, s'_V) \quad & \text{if and only if} \quad (i < i') \\
 & \vee ((i = i') \wedge (l_U + l_V < l'_U + l'_V)) \\
 & \vee ((i = i') \wedge (l_U + l_V = l'_U + l'_V) \wedge (l_U < l'_U)) \\
 & \vee ((i = i') \wedge (l_U + l_V = l'_U + l'_V) \wedge (l_U = l'_U) \wedge (s_U < s'_U)).
 \end{aligned}$$

Note that, by rewriting states $(i, l_U, l_V, s_U, s_V) \in \mathcal{S}$ in terms of $(i, l_U + l_V, l_U, s_U)$, the above labelling of states can be thought of as the lexico-

graphical ordering for the resulting modified set of states.

Appendix A.1. Sub-matrices $\mathbf{A}_{i,i-1}$, for integers $i \in \{1, \dots, N\}$

Based on the above labelling of states, it is readily seen that sub-matrices $\mathbf{A}_{i,i-1}$, for $i \in \{1, \dots, N\}$, have dimension $\binom{N-i+3}{3} \times \binom{N-i+4}{3}$ and are specified by the block-structured form

$$\begin{pmatrix} \mathbf{C}_{0,0}(i) & \mathbf{C}_{0,1}(i) & & & \\ & \mathbf{C}_{1,1}(i) & \mathbf{C}_{1,2}(i) & & \\ & & \ddots & \ddots & \\ & & & \mathbf{C}_{N-i,N-i}(i) & \mathbf{C}_{N-i,N-i+1}(i) \end{pmatrix},$$

where $\mathbf{C}_{j,j'}(i)$ is related to jumps of \mathcal{X} from states in $l(i|j)$ to states in $l(i-1|j')$ with *sub-levels* $l(i|j) = \{(i, l_U, l_V, s_U, s_V) \in l(i) : l_U + l_V = j\}$, for $j \in \{0, \dots, N-i\}$. Then, an appropriate decomposition of sub-levels $l(i|j)$ into subsets $l(i|j|k) = \{(i, l_U, l_V, s_U, s_V) \in l(i|j) : l_U = k\}$, for $k \in \{0, \dots, j\}$, allows us to express $\mathbf{C}_{j,j}(i)$ and $\mathbf{C}_{j,j+1}(i)$ as follows:

$$\begin{aligned} \mathbf{C}_{j,j}(i) &= \begin{pmatrix} \mathbf{J}_{0,0}(i|j) & & & \\ & \mathbf{J}_{1,1}(i|j) & & \\ & & \ddots & \\ & & & \mathbf{J}_{j,j}(i|j) \end{pmatrix}, \\ \mathbf{C}_{j,j+1}(i) &= \begin{pmatrix} \mathbf{H}_{0,0}(i|j) & \mathbf{H}_{0,1}(i|j) & & \\ & \mathbf{H}_{1,1}(i|j) & \mathbf{H}_{1,2}(i|j) & \\ & & \ddots & \ddots \\ & & & \mathbf{H}_{j,j}(i|j) & \mathbf{H}_{j,j+1}(i|j) \end{pmatrix}, \end{aligned}$$

where sub-matrices $\mathbf{J}_{k,k}(i|j)$ have dimension $(N-i-j+1) \times (N-i-j+2)$, and $\mathbf{H}_{k,k'}(i|j)$ are square matrices of order $N-i-j+1$, for $k \in \{0, \dots, j\}$ and $k' \in \{k, k+1\}$. They are given by $\mathbf{H}_{k,k}(i|j) = \delta_R \xi (1 - \theta_U) i \mathbf{I}_{N-i-j+1}$

and $\mathbf{H}_{k,k+1}(i|j) = \delta_R \xi \theta_U^i \mathbf{I}_{N-i-j+1}$, where \mathbf{I}_a is the identity matrix of order a , and

$$\mathbf{J}_{k,k}(i|j) = \begin{pmatrix} a(i) & b(i) & & & \\ & a(i) & b(i) & & \\ & & \ddots & \ddots & \\ & & & a(i) & b(i) \end{pmatrix},$$

with $a(i) = (\delta_R(1 - \xi)(1 - \tilde{\theta}_U) + (\delta_{TB} + \delta_D)\eta(q, C))i$ and $b(i) = (\delta_R(1 - \xi)\tilde{\theta}_U + (\delta_{TB} + \delta_D)(1 - \eta(q, C)))i$.

Appendix A.2. Sub-matrices $\mathbf{A}_{i,i}$, for integers $i \in \{1, \dots, N\}$

Sub-matrices $\mathbf{A}_{i,i}$, for $i \in \{1, \dots, N\}$, are seen to be square matrices of order $\binom{N-i+3}{3}$ and have the block-structured form

$$\begin{pmatrix} \mathbf{B}_{0,0}(i) & \mathbf{B}_{0,1}(i) & & & \\ \mathbf{B}_{1,0}(i) & \mathbf{B}_{1,1}(i) & \mathbf{B}_{1,2}(i) & & \\ & \ddots & \ddots & \ddots & \\ & & \mathbf{B}_{N-i-1,N-i-2}(i) & \mathbf{B}_{N-i-1,N-i-1}(i) & \mathbf{B}_{N-i-1,N-i}(i) \\ & & & \mathbf{B}_{N-i,N-i-1}(i) & \mathbf{B}_{N-i,N-i}(i) \end{pmatrix},$$

where sub-matrices $\mathbf{B}_{j,j'}(i)$ are associated with jumps of \mathcal{X} from states in $l(i|j)$ to states in $l(i|j')$ with $j' \in \{j-1, j, j+1\}$. By using subsets $l(i|j|k)$, they can be written as follows:

$$\mathbf{B}_{j,j-1}(i) = \begin{pmatrix} \mathbf{G}_{0,0}(i|j) & & & & \\ \mathbf{G}_{1,0}(i|j) & \mathbf{G}_{1,1}(i|j) & & & \\ & \ddots & \ddots & & \\ & & \mathbf{G}_{j-1,j-2}(i|j) & \mathbf{G}_{j-1,j-1}(i|j) & \\ & & & \mathbf{G}_{j,j-1}(i|j) \end{pmatrix},$$

$$\mathbf{B}_{j,j}(i) = \text{diag}(\mathbf{F}_{0,0}(i|j), \dots, \mathbf{F}_{j,j}(i|j)),$$

$$\mathbf{B}_{j,j+1}(i) = \begin{pmatrix} \mathbf{E}_{0,0}(i|j) & \mathbf{E}_{0,1}(i|j) & & & \\ & \mathbf{E}_{1,1}(i|j) & \mathbf{E}_{1,2}(i|j) & & \\ & & \ddots & \ddots & \\ & & & \mathbf{E}_{j,j}(i|j) & \mathbf{E}_{j,j+1}(i|j) \end{pmatrix},$$

where sub-matrices $\mathbf{G}_{k,k-1}(i|j)$, for $k \in \{1, \dots, j\}$, and $\mathbf{G}_{k,k}(i|j)$, for $k \in \{0, \dots, j-1\}$, have dimension $(N-i-j+1) \times (N-i-j+2)$, $\mathbf{F}_{k,k}(i|j)$ are square matrices of order $N-i-j+1$, for $k \in \{0, \dots, j\}$, and sub-matrices $\mathbf{E}_{k,k'}(i|j)$, for $k' \in \{k, k+1\}$, have dimension $(N-i-j+1) \times (N-i-j)$, for $k \in \{0, \dots, j\}$. They can be written as

$$\mathbf{G}_{k,k-1}(i|j) = \begin{pmatrix} \delta_D \eta(q, C)k & \delta_D(1 - \eta(q, C))k & & & \\ & \ddots & \ddots & & \\ & & \delta_D \eta(q, C)k & \delta_D(1 - \eta(q, C))k \end{pmatrix},$$

$$\mathbf{G}_{k,k}(i|j) = \mathbf{G}_{j-k,j-k-1}(i|j),$$

$$\mathbf{F}_{k,k}(i|j) = \text{diag}(q_{(i,k,j-k,0,N-i-j)}, \dots, q_{(i,k,j-k,N-i-j,0)})$$

$$+ \begin{pmatrix} 0 & c(N-i-j) & & & \\ d(1) & 0 & c(N-i-j-1) & & \\ & \ddots & \ddots & \ddots & \\ & & d(N-i-j-1) & 0 & c(1) \\ & & & d(N-i-j) & 0 \end{pmatrix},$$

$$\mathbf{E}_{k,k}(i|j) = \begin{pmatrix} \alpha_V(1-p_V)(N-i-j)i & & & & \\ & \alpha_V(1-p_V)(N-i-j-1)i & & & \\ & & \ddots & & \\ & & & \alpha_V(1-p_V)i & \\ & & & & 0 \end{pmatrix},$$

$$\mathbf{E}_{k,k+1}(i|j) = \begin{pmatrix} 0 & & & \\ \alpha_U(1-p_U)i & & & \\ & \ddots & & \\ & & \alpha_U(1-p_U)(N-i-j-1)i & \\ & & & \alpha_U(1-p_U)(N-i-j)i \end{pmatrix},$$

with $c(i') = (\gamma + \delta_D(1 - \eta(q, C))i')$ and $d(i') = \delta_D\eta(q, C)i'$, for $i' \in \{1, \dots, N - i - j\}$, and $q_{(i,k,j-k,s,N-i-j-s)} = -((\delta_D + \delta_{TB} + \delta_R)i + \delta_D j + r_V(i)(j - k) + r_U(i)k + (\alpha_U i + \delta_D\eta(q, C))s + (\gamma + \alpha_V i + \delta_D(1 - \eta(q, C)))(N - i - j - s))$, for $s \in \{0, \dots, N - i - j\}$.

Appendix A.3. Sub-matrices $\mathbf{A}_{i,i+1}$, for integers $i \in \{1, \dots, N - 1\}$

Sub-matrices $\mathbf{A}_{i,i+1}$, for $i \in \{1, \dots, N - 1\}$, have dimension $\binom{N-i+3}{3} \times \binom{N-i+2}{3}$ and their block-structured form is given by

$$\begin{pmatrix} \mathbf{D}_{0,0}(i) & & & \\ \mathbf{D}_{1,0}(i) & \mathbf{D}_{1,1}(i) & & \\ & \ddots & \ddots & \\ & & \mathbf{D}_{N-i-1,N-i-2}(i) & \mathbf{D}_{N-i-1,N-i-1}(i) \\ & & & \mathbf{D}_{N-i,N-i-1}(i) \end{pmatrix},$$

where $\mathbf{D}_{j,j'}(i)$ is related to jumps of \mathcal{X} from states in $l(i|j)$ to states in $l(i + 1|j')$, with $j' \in \{j-1, j\}$, which results in $\mathbf{D}_{j,j}(i) = \text{diag}(\mathbf{L}_{0,0}(i|j), \dots, \mathbf{L}_{j,j}(i|j))$ and

$$\mathbf{D}_{j,j-1}(i) = \begin{pmatrix} \mathbf{M}_{0,0}(i|j) & & & \\ \mathbf{M}_{1,0}(i|j) & \mathbf{M}_{1,1}(i|j) & & \\ & \ddots & \ddots & \\ & & \mathbf{M}_{j-1,j-2}(i|j) & \mathbf{M}_{j-1,j-1}(i|j) \\ & & & \mathbf{M}_{j,j-1}(i|j) \end{pmatrix}.$$

In these expressions, $\mathbf{M}_{k,k}(i|j)$, for $k \in \{0, \dots, j-1\}$, and $\mathbf{M}_{k,k-1}(i|j)$, for $k \in \{1, \dots, j\}$, are square matrices of order $N - i - j + 1$, and sub-matrices $\mathbf{L}_{k,k}(i|j)$ have dimension $(N - i - j + 1) \times (N - i - j)$, for $k \in \{0, \dots, j\}$. More particularly, it is easily seen that $\mathbf{M}_{k,k-1}(i|j) = r_U(i)k\mathbf{I}_{N-i-j+1}$ and $\mathbf{M}_{k,k}(i|j) = r_V(i)(j-k)\mathbf{I}_{N-i-j+1}$, and $\mathbf{L}_{k,k}(i|j)$ is specified by

$$\begin{pmatrix} \alpha_V p_V(N-i-j)i & & & & & \\ & \alpha_U p_U i & & & & \\ & & \alpha_V p_V(N-i-j-1)i & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & \alpha_U p_U(N-i-j-1)i & \alpha_V p_V i \\ & & & & & & \alpha_U p_U(N-i-j)i \end{pmatrix}.$$

**Appendix B. Probabilities $P(I_{\max} \leq i | X(0) = (i_0, l_U, l_V, s_U, s_V))$,
for initial states $(i_0, l_U, l_V, s_U, s_V) \in \mathcal{S}_T$**

In Algorithm 1, the rows of the matrix $\hat{\mathbf{f}}(\infty; \mathcal{S}_A)$ record probabilities of absorption into states in $l(0)$ and the entries of the column vector \mathbf{g} specify expected values $E[\tau | X(0) = (i_0, l_U, l_V, s_U, s_V)]$, for every initial state $(i_0, l_U, l_V, s_U, s_V) \in \mathcal{S}_T$. This means that $P(I_{\max} \leq i | X(0) = (i_0, l_U, l_V, s_U, s_V)) = \pi_T(i)\mathbf{p}(i)$, $\mathbf{f}(\infty; \mathcal{S}_A) = \pi_T \hat{\mathbf{f}}(\infty; \mathcal{S}_A)$ and $E[\tau^k | X(0) = (i_0, l_U, l_V, s_U, s_V)] = \pi_T \mathbf{g}$, where π_T is the initial probability vector of \mathcal{X} —on the states of \mathcal{S}_T —in Section 3, provided that $X(0) = (i_0, l_U, l_V, s_U, s_V)$, and $\pi_T(i)$ is obtained from π_T by removing entries associated with states in $\cup_{i'=i+1}^N l(i')$.

Algorithm 1: *Computation of the conditional probabilities $P(I_{\max} \leq i | X(0) = (i_0, l_U, l_V, s_U, s_V))$, for integers $i \in \{i_0, \dots, N\}$, the hitting probabilities of \mathcal{X} and the expected value $E[\tau^k | X(0) = (i_0, l_U, l_V, s_U, s_V)]$, for $k \in \mathbb{N}$ and an initial state $(i_0, l_U, l_V, s_U, s_V) \in \mathcal{S}_T$.*

Step 0: $i := 1$;

$$-\mathbf{T}^{-1}(i) := -\mathbf{A}_{i,i}^{-1};$$

$$\mathbf{t}_0(i) := \mathbf{A}_{i,i-1} \mathbf{1}_{J_A};$$

$$\mathbf{p}(i) := (-\mathbf{T}^{-1}(i)) \mathbf{t}_0(i).$$

Step 1: While $i < N$, do

$$i := i + 1;$$

$$\mathbf{U}_{2,1}(i) := (\mathbf{0}_{J'(i) \times J(i-2)}, \mathbf{A}_{i,i-1});$$

$$\mathbf{U}_{1,2}(i) := \begin{pmatrix} \mathbf{0}_{J(i-2) \times J'(i)} \\ \mathbf{A}_{i-1,i} \end{pmatrix};$$

$$\mathbf{V}_{2,2}(i) := (-\mathbf{A}_{i,i} - \mathbf{U}_{2,1}(i)(-\mathbf{T}^{-1}(i-1))\mathbf{U}_{1,2}(i))^{-1};$$

$$\mathbf{V}_{1,2}(i) := -\mathbf{T}^{-1}(i-1)\mathbf{U}_{1,2}(i)\mathbf{V}_{2,2}(i);$$

$$\mathbf{V}_{2,1}(i) := \mathbf{V}_{2,2}(i)\mathbf{U}_{2,1}(i)(-\mathbf{T}^{-1}(i-1));$$

$$\mathbf{V}_{1,1}(i) := -\mathbf{T}^{-1}(i-1)(\mathbf{I}_{J(i-1)} + \mathbf{U}_{1,2}(i)\mathbf{V}_{2,1}(i));$$

$$-\mathbf{T}^{-1}(i) := \begin{pmatrix} \mathbf{V}_{1,1}(i) & \mathbf{V}_{1,2}(i) \\ \mathbf{V}_{2,1}(i) & \mathbf{V}_{2,2}(i) \end{pmatrix};$$

$$\mathbf{p}(i) := \begin{pmatrix} (\mathbf{I}_{J(i-1)} + \mathbf{V}_{1,2}(i)\mathbf{U}_{2,1}(i)) \mathbf{p}(i-1) \\ \mathbf{V}_{2,2}(i)\mathbf{U}_{2,1}(i)\mathbf{p}(i-1) \end{pmatrix};$$

end while;

$$\widehat{\mathbf{f}}(\infty; \mathcal{S}_A) := (-\mathbf{T}^{-1}(i)) \mathbf{Q}_{T,A};$$

$$\mathbf{g} := k!(-\mathbf{T}^{-k}(i)) \mathbf{1}_{J(i)}.$$

Step 2: $\mathbf{f}(\infty; \mathcal{S}_A) := \pi_T \widehat{\mathbf{f}}(\infty; \mathcal{S}_A)$;

$$E[\tau^k | X(0) = (i_0, l_U, l_V, s_U, s_V)] := \pi_T \mathbf{g};$$

$$P(I_{\max} \leq i | X(0) = (i_0, l_U, l_V, s_U, s_V)) := 1;$$

while $i > i_0$, do

$$i := i - 1;$$

$$P(I_{\max} \leq i | X(0) = (i_0, l_U, l_V, s_U, s_V)) := \pi_T(i) \mathbf{p}(i);$$

end while.

In order to analyse the efficiency of Algorithm 1 in terms of the number $J'(i)$ of states in the i th level $l(i)$, and the number $J(i)$ of states in the subset $\cup_{i'=1}^i l(i')$, it must be observed that, in evaluating the column vectors $\mathbf{p}(i)$ for integers $i \in \{1, \dots, N\}$, the most intensive computational effort lies in the computation of the matrices $\mathbf{V}_{1,1}(i)$ in Step 1 and the corresponding complexity is $O\left(\sum_{i=1}^{N-1} (J(i))^3\right)$. Then, by noting that the complexities $O(J_A J'(1) J(N))$ and $O(\delta_{1,k}(J(N))^2 + (1 - \delta_{1,k})(J(N))^3)$ are related to the computation of the matrix $\hat{\mathbf{f}}(\infty; \mathcal{S}_A)$ and the vector \mathbf{g} , respectively, in Step 1, we may propose that the computational complexity of Algorithm 1 is $O\left(\sum_{i=1}^N (J(i))^3\right)$, regardless of the integer $k \in \mathbb{N}$, since $O\left(\max\{\sum_{i=1}^{N-1} (J(i))^3, J_A J'(1) J(N), \delta_{1,k}(J(N))^2 + (1 - \delta_{1,k})(J(N))^3\}\right) \subset O\left(\sum_{i=1}^N (J(i))^3\right)$.

Appendix C. Probability mass function of V and factorial moments

For convenience (Section 4), we select $(i_0, l_U, l_V, s_U, s_V) = (1, 0, 0, N - 1, 0)$ and let Algorithm 2 compute probabilities $P_{(i_0, l_U, l_V, s_U, s_V)}(v)$, for integers $v \in \{0, \dots, V_{q'}\}$, where $V_{q'}$ is the $(100q')$ -th percentile of the cumulative number V of vaccinations, for a predetermined probability $q' \in (0, 1)$; i.e., $V_{q'}$ satisfies

$$F_{(i_0, l_U, l_V, s_U, s_V)}(V_{q'} - 1) \leq q' < F_{(i_0, l_U, l_V, s_U, s_V)}(V_{q'}),$$

with $F_{(i_0, l_U, l_V, s_U, s_V)}(v) = \sum_{v'=0}^v P_{(i_0, l_U, l_V, s_U, s_V)}(v')$. Note that, for initial states $(i_0, l_U, l_V, s_U, s_V) \in \mathcal{S}_T \setminus \{(1, 0, 0, N - 1, 0)\}$, Algorithm 2 can be readily adapted by replacing $\mathbf{e}_{J(1)}(N)\mathbf{p}(v; 1)$ in Steps 1 and 2 by $\bar{\mathbf{e}}_{J(i_0)}\mathbf{p}(v; i_0)$,

where $\bar{\mathbf{e}}_{J(i_0)}$ denotes a null vector of order $J(i_0)$ with a single 1 at the entry corresponding to state $(i_0, l_U, l_V, s_U, s_V)$.

Algorithm 2: *Computation of the probabilities $P_{(i_0, l_U, l_V, s_U, s_V)}(v)$, for integers $v \in \{0, \dots, V_{q'}\}$, state $(i_0, l_U, l_V, s_U, s_V) = (1, 0, 0, N-1, 0)$ and $q' \in (0, 1)$.*

Step 0: $i := 1$;

$v := 0$;

$\mathbf{W}_i := \mathbf{A}_{i,i}^* + \mathbf{A}_{i,i}^U$;

$\mathbf{a}(v; i) := -\mathbf{W}_i^{-1} \mathbf{A}_{i,i-1}^U \mathbf{1}_{\binom{N+3}{3}}$;

while $i < N$, do

$i := i + 1$;

$\mathbf{W}_i := \mathbf{A}_{i,i}^* + \mathbf{A}_{i,i}^U + \mathbf{A}_{i,i-1}^U (-\mathbf{W}_{i-1}^{-1}) \mathbf{A}_{i-1,i}$;

$\mathbf{a}(v; i) := -\mathbf{W}_i^{-1} \mathbf{A}_{i,i-1}^U \mathbf{a}(v; i-1)$;

end while.

Step 1: $\mathbf{p}(v; i) := \mathbf{a}(v; i)$;

while $i > 1$, do

$i := i - 1$;

$\mathbf{p}(v; i) := \mathbf{a}(v; i) - \mathbf{W}_i^{-1} \mathbf{A}_{i,i+1} \mathbf{p}(v; i+1)$;

end while;

$P_{(1,0,0,N-1,0)}(v) := \mathbf{e}_{J(1)}(N) \mathbf{p}(v; 1)$;

$F_{(1,0,0,N-1,0)}(v) := P_{(1,0,0,N-1,0)}(v)$.

Step 2: While $F_{(1,0,0,N-1,0)}(v) < q'$, do

$v := v + 1$;

$\mathbf{a}(v; i) := -\mathbf{W}_i^{-1} (\delta_{1,v} \mathbf{A}_{i,i-1}^V \mathbf{1}_{\binom{N+3}{3}} + \mathbf{A}_{i,i}^V \mathbf{p}(v-1; i))$;

while $i < N$, do

```

 $i := i + 1;$ 
 $\mathbf{a}(v; i) := -\mathbf{W}_i^{-1}(\mathbf{A}_{i,i-1}^U \mathbf{a}(v; i-1) + \mathbf{A}_{i,i-1}^V \mathbf{p}(v-1; i-1)$ 
 $\quad + \mathbf{A}_{i,i}^V \mathbf{p}(v-1; i));$ 
end while;
 $\mathbf{p}(v; i) := \mathbf{a}(v; i);$ 
while  $i > 1$ , do
 $i := i - 1;$ 
 $\mathbf{p}(v; i) := \mathbf{a}(v; i) - \mathbf{W}_i^{-1} \mathbf{A}_{i,i+1} \mathbf{p}(v; i+1);$ 
end while;
 $P_{(1,0,0,N-1,0)}(v) := \mathbf{e}_{J(1)}(N) \mathbf{p}(v; 1);$ 
 $F_{(1,0,0,N-1,0)}(v) := F_{(1,0,0,N-1,0)}(v-1) + P_{(1,0,0,N-1,0)}(v);$ 
end while.

```

Algorithm 3: *Computation of sub-vectors $\mathbf{v}^{(k)}(i)$, for $i \in \{1, \dots, N\}$ and a predetermined integer $k \in \mathbb{N}$.*

```

Step 0:  $i := 1;$ 
 $k' := 0;$ 
for  $i' = 1, \dots, N$ , do
 $\mathbf{v}^{(k')}(i') := \mathbf{1}_{\binom{N-i'+3}{3}};$ 
end for;
 $k' = k' + 1;$ 
 $\mathbf{W}_i := \mathbf{A}_{i,i};$ 
 $\mathbf{b}(i) := -\mathbf{W}_i^{-1}(\mathbf{A}_{i,i-1}^V \mathbf{1}_{\binom{N+3}{3}} + \mathbf{A}_{i,i}^V \mathbf{v}^{(k'-1)}(i));$ 
while  $i < N$ , do
 $i := i + 1;$ 
 $\mathbf{W}_i := \mathbf{A}_{i,i} + \mathbf{A}_{i,i-1}(-\mathbf{W}_{i-1}^{-1})\mathbf{A}_{i-1,i};$ 

```

$$\mathbf{b}(i) := -\mathbf{W}_i^{-1}(\mathbf{A}_{i,i-1}\mathbf{b}(i-1) + k'(\mathbf{A}_{i,i-1}^V \mathbf{v}^{(k'-1)}(i-1) + \mathbf{A}_{i,i}^V \mathbf{v}^{(k'-1)}(i)));$$

end while.

Step 1: $\mathbf{v}^{(k')}(i) := \mathbf{b}(i);$

while $i > 1$, do

$i := i - 1;$

$\mathbf{v}^{(k')}(i) := \mathbf{b}(i) - \mathbf{W}_i^{-1}\mathbf{A}_{i,i+1}\mathbf{v}^{(k')}(i+1);$

end while.

Step 2: While $k' < k$, do

$k' := k' + 1;$

$\mathbf{b}(i) := -k'\mathbf{W}_i^{-1}\mathbf{A}_{i,i}^V \mathbf{v}^{(k'-1)}(i);$

while $i < N$, do

$i := i + 1;$

$$\mathbf{b}(i) := -\mathbf{W}_i^{-1}(\mathbf{A}_{i,i-1}\mathbf{b}(i-1) + k'(\mathbf{A}_{i,i-1}^V \mathbf{v}^{(k'-1)}(i-1) + \mathbf{A}_{i,i}^V \mathbf{v}^{(k'-1)}(i)));$$

end while;

$\mathbf{v}^{(k')}(i) := \mathbf{b}(i);$

while $i > 1$, do

$i := i - 1;$

$\mathbf{v}^{(k')}(i) := \mathbf{b}(i) - \mathbf{W}_i^{-1}\mathbf{A}_{i,i+1}\mathbf{v}^{(k')}(i+1);$

end while;

end while.

In evaluating the conditional probabilities $P_{(i_0, l_U, l_V, s_U, s_V)}(v)$, for integers $v \in \{0, \dots, V_{q'}\}$, from Algorithm 2, the crucial steps are related to the inversion of the matrices \mathbf{W}_i , for integers $i \in \{1, \dots, N\}$, in Step 0 and the product

of matrices/vectors in Steps 1-2, whose complexities are $O\left(\sum_{i=1}^N (J'(i))^3\right)$ and $O\left(\sum_{i=1}^N (J'(i))^2\right)$, respectively. Then, by noting that the $(100q')$ -th percentile $V_{q'}$ of the cumulative number V of vaccinations should be seen as a function of the population size N , it is found that the complexity of Algorithm 2 is $O\left(\sum_{i=1}^N (J'(i))^3 + (1 + V_{q'}) \sum_{i=1}^N (J'(i))^2\right)$. In a similar manner, the complexities of Step 0 and Steps 1-2 in Algorithm 3 are $O\left(\sum_{i=1}^N (J'(i))^3\right)$ and $O\left(\sum_{i=1}^N (J'(i))^2\right)$, respectively, from which it follows that the complexity of Algorithm 3 is $O\left(\sum_{i=1}^N (J'(i))^3\right)$ for a predefined integer $k \in \mathbb{N}$, since $O\left(\sum_{i=1}^N (J'(i))^3 + k \sum_{i=1}^N (J'(i))^2\right) \subset O\left(\sum_{i=1}^N (J'(i))^3\right)$.

Appendix D. Probability mass function of $\mathcal{R}_{exact,0}$ and factorial moments

In Algorithm 4, we let $R_{q'}$ denote the $(100q')$ -th percentile of $\mathcal{R}_{exact,0}$ and it is thus verified that $G_{(1,0,0,N-1,0)}(R_{q'} - 1) \leq q' < G_{(1,0,0,N-1,0)}(R_{q'})$, with $G_{(1,0,0,N-1,0)}(r) = \sum_{r'=0}^r Q_{(1,0,0,N-1,0)}(r')$, for $r \in \mathbb{N}_0$ and $q' \in (0, 1)$.

Algorithm 4: *Computation of the probabilities $Q_{(1,0,0,N-1,0)}(r)$, for integers $r \in \{0, \dots, R_{q'}\}$ and $q' \in (0, 1)$.*

Step 0: $i := 1$;

$r := 0$;

$\widehat{\mathbf{W}}_i := \mathbf{A}_{i,i}$;

$\mathbf{c}(r; i) := -\widehat{\mathbf{W}}_i^{-1} \overline{\mathbf{A}}_{i,i-1} \mathbf{1}_{\binom{N+3}{3}}$;

while $i < N$, do

$i := i + 1$;

$\widehat{\mathbf{W}}_i := \mathbf{A}_{i,i} + \widetilde{\mathbf{A}}_{i,i-1} (-\widehat{\mathbf{W}}_{i-1}^{-1}) \widetilde{\mathbf{A}}_{i-1,i}$;

$$\mathbf{c}(r; i) := -\widehat{\mathbf{W}}_i^{-1}(\overline{\mathbf{A}}_{i,i-1}\mathbf{1}_{\binom{N-i+4}{3}} + \widetilde{\mathbf{A}}_{i,i-1}\mathbf{c}(r; i-1));$$

end while.

Step 1: $\mathbf{q}(r; i) := \mathbf{c}(r; i);$

while $i > 1$, do

$$i := i - 1;$$

$$\mathbf{q}(r; i) := \mathbf{c}(r; i) - \widehat{\mathbf{W}}_i^{-1}\widetilde{\mathbf{A}}_{i,i+1}\mathbf{q}(r; i+1);$$

end while;

$$Q_{(1,0,0,N-1,0)}(r) := \mathbf{e}_{J(1)}(N)\mathbf{q}(r; 1);$$

$$G_{(1,0,0,N-1,0)}(r) := Q_{(1,0,0,N-1,0)}(r).$$

Step 2: While $G_{(1,0,0,N-1,0)}(r) < q'$, do

$$r := r + 1;$$

$$\mathbf{c}(r; i) := -\widehat{\mathbf{W}}_i^{-1}\overline{\mathbf{A}}_{i,i+1}\mathbf{p}(r-1; i+1);$$

while $i < N$, do

$$i := i + 1;$$

$$\begin{aligned} \mathbf{c}(r; i) := & -\widehat{\mathbf{W}}_i^{-1}(\widetilde{\mathbf{A}}_{i,i-1}\mathbf{c}(r; i-1) \\ & + (1 - \delta_{N,i})\overline{\mathbf{A}}_{i,i+1}^V\mathbf{q}(r-1; i+1)); \end{aligned}$$

end while;

$$\mathbf{q}(r; i) := \mathbf{c}(r; i);$$

while $i > 1$, do

$$i := i - 1;$$

$$\mathbf{q}(r; i) := \mathbf{c}(r; i) - \widehat{\mathbf{W}}_i^{-1}\widetilde{\mathbf{A}}_{i,i+1}\mathbf{q}(r; i+1);$$

end while;

$$Q_{(1,0,0,N-1,0)}(r) := \mathbf{e}_{J(1)}(N)\mathbf{q}(r; 1);$$

$$G_{(1,0,0,N-1,0)}(r) := G_{(1,0,0,N-1,0)}(r-1) + Q_{(1,0,0,N-1,0)}(r);$$

end while.

Algorithm 5: *Computation of the k th factorial moment $\overline{\mathcal{R}}_{exact,0}^{(k)}$, for a predetermined integer $k \in \mathbb{N}$.*

Step 0: $i := 1$;

$k' := 0$;

for $i' = 1, \dots, N$, do

$$\mathbf{r}^{(k')}(i') := \mathbf{1}_{\binom{N-i'+3}{3}};$$

end for;

$k' = k' + 1$;

$\widehat{\mathbf{W}}_i := \mathbf{A}_{i,i}$;

$\mathbf{d}(i) := -\widehat{\mathbf{W}}_i^{-1} \overline{\mathbf{A}}_{i,i+1} \mathbf{r}^{(k'-1)}(i+1)$;

while $i < N$, do

$i := i + 1$;

$\widehat{\mathbf{W}}_i := \mathbf{A}_{i,i} + \widetilde{\mathbf{A}}_{i,i-1} (-\widehat{\mathbf{W}}_{i-1}^{-1}) \mathbf{A}_{i-1,i}$;

$\mathbf{d}(i) := -\widehat{\mathbf{W}}_i^{-1} (\widetilde{\mathbf{A}}_{i,i-1} \mathbf{d}(i-1) + k'(1 - \delta_{N,i}) \overline{\mathbf{A}}_{i,i+1} \mathbf{r}^{(k'-1)}(i+1))$;

end while.

Step 1: $\mathbf{r}^{(k')}(i) := \mathbf{d}(i)$;

while $i > 1$, do

$i := i - 1$;

$\mathbf{r}^{(k')}(i) := \mathbf{d}(i) - \widehat{\mathbf{W}}_i^{-1} \mathbf{A}_{i,i+1} \mathbf{r}^{(k')}(i+1)$;

end while.

Step 2: While $k' < k$, do

$k' := k' + 1$;

$\mathbf{d}(i) := -k' \widehat{\mathbf{W}}_i^{-1} \overline{\mathbf{A}}_{i,i+1} \mathbf{r}^{(k'-1)}(i+1)$;

while $i < N$, do

$i := i + 1$;

$$\mathbf{d}(i) := -\widehat{\mathbf{W}}_i^{-1}(\widetilde{\mathbf{A}}_{i,i-1}\mathbf{d}(i-1) + (1-\delta_{N,i})k'\overline{\mathbf{A}}_{i,i+1}\mathbf{r}^{(k')}(i+1));$$

end while;

$$\mathbf{r}^{(k')}(i) := \mathbf{d}(i);$$

while $i > 1$, do

$$i := i - 1;$$

$$\mathbf{r}^{(k')}(i) := \mathbf{d}(i) - \widehat{\mathbf{W}}_i^{-1}\mathbf{A}_{i,i+1}\mathbf{r}^{(k')}(i+1);$$

end while;

end while;

$$\overline{\mathcal{R}}_{exact,0}^{(k)} := \mathbf{e}_{J(i)}\mathbf{r}^{(k')}(i).$$

In a similar manner to Appendix C, it is found that the complexities of Algorithms 4 and 5 are $O\left(\sum_{i=1}^N(J'(i))^3 + (1 + R_{q'})\sum_{i=1}^N(J'(i))^2\right)$ and $O\left(\sum_{i=1}^N(J'(i))^3\right)$, respectively.

References

- [1] Abu-Raddad LJ, Sabatelli L, Achterberg JT, Sugimoto JD, Longini IM Jr, Dye C, Halloran ME (2009) Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. *Proceedings of the National Academic of Sciences* 106: 13980–13985.
- [2] Agosto FB, Cook J, Shelton PD, Wickers MG (2015) Mathematical model of MDR-TB and XDR-TB with isolation and lost to follow-up. *Abstract and Applied Analysis* 2015: 828461.
- [3] Allen LJS (2017) A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis. *Infectious Disease Modelling* 2: 128–142.

- [4] Almaraz E, Gómez-Corral A (2019) Number of infections suffered by a focal individual in a two-strain SIS model with partial cross-immunity. *Mathematical Methods in the Applied Sciences* 42: 4318–4330.
- [5] Amador J, Armesto D, Gómez-Corral A (2019) Extreme values in SIR epidemic models with two strains and cross-immunity. *Mathematical Biosciences and Engineering* 16: 1992–2022.
- [6] Amador J, Gómez-Corral A (2020) A stochastic epidemic model with two quarantine states and a limited carrying capacity for quarantine. *Physica A* 544: 121899.
- [7] Anzai A, Kawatsu L, Uchimura K, Nishiura H (2020) Reconstructing the population dynamics of foreign residents in Japan to estimate the prevalence of infection with *Mycobacterium tuberculosis*. *Journal of Theoretical Biology* 489: 110160.
- [8] Artalejo JR, López-Herrero MJ (2013) On the exact measure of disease spread in stochastic epidemic models. *Bulletin of Mathematical Biology* 75: 1031–1050.
- [9] Awofeso N (2008) Anti-tuberculosis medication side-effects constitute major factor for poor adherence to tuberculosis treatment. *Bulletin of the World Health Organization* 86: B–C.
- [10] Baumann H, Sandmann W (2016) Structured modeling and analysis of stochastic epidemics with immigration and demographic effects. *PLoS ONE* 11(3): e0152144.
- [11] Bhunu CP, Garira W, Mukandavire Z, Magombedze G (2008) Modelling

- the effects of pre-exposure and post-exposure vaccines in tuberculosis control. *Journal of Theoretical Biology* 254: 633–649.
- [12] Britton T (2010) Stochastic epidemic models: A survey. *Mathematical Biosciences* 225: 24–35.
 - [13] Britton WJ, Palendira U (2003) Improving vaccines against tuberculosis. *Immunology and Cell Biology* 81: 34–45.
 - [14] Castillo-Chavez C, Feng Z (1997) To treat or not to treat: the case of tuberculosis. *Journal of Mathematical Biology* 35: 629–656.
 - [15] Castillo-Chavez C, Feng Z (1998) Global stability of an age-structure model for TB and its applications to optimal vaccination strategies. *Mathematical Biosciences* 151: 135–154.
 - [16] Castillo-Chavez C, Song B (2004) Dynamical models of tuberculosis and their applications. *Mathematical Biosciences and Engineering* 1: 361–404.
 - [17] Chen C, Kang Y (2016) The asymptotic behavior of a stochastic vaccination model with backward bifurcation. *Applied Mathematical Modelling* 40: 6051–6068.
 - [18] Cliff JM, Kaufmann, McShane H, van Helden P, O’Garra A (2015) The human immune response to tuberculosis and its treatment: a view from the blood. *Immunological Reviews* 264: 88–102.
 - [19] Dye C (2013) Making wider use of the world’s most widely used vaccine: Bacille Calmette–Guérin revaccination reconsidered. *Journal of the Royal Society Interface* 10: 20130365.

- [20] Dye C, Glaziou P, Floyd K, Raviglione M (2013) Prospects for tuberculosis elimination. *Annual Review of Public Health* 34: 271–286.
- [21] Economou A, Gómez-Corral A, López-García M (2015) A stochastic SIS epidemic model with heterogeneous contacts. *Physica A* 421: 78–97.
- [22] Feng Z, Castillo-Chavez C, Capurro AF (2000) A model for tuberculosis with exogenous reinfection. *Theoretical Population Biology* 57: 235–247.
- [23] Fofana MO, Shrestha S, Knight GM, Cohen T, White RG, Cobelens F, Dowdy DW (2017) A multistrain mathematical model to investigate the role of pyrazinamide in the emergence of extensively drug-resistant tuberculosis. *Antimicrobial Agents and Chemotherapy* 61: e00498-16.
- [24] Gamboa M, López-Herrero MJ (2020) Measuring infection transmission in a stochastic SIV model with infection reintroduction and imperfect vaccine. *Acta Biotheoretica*, to appear. doi.org/10.1007/s10441-019-09373-9
- [25] Gerberry DJ (2009) Trade-off between BCG vaccination and the ability to detect and treat latent tuberculosis. *Journal of Theoretical Biology* 261: 548–560.
- [26] Gerberry DJ (2016) Practical aspects of backward bifurcation in a mathematical model for tuberculosis. *Journal of Theoretical Biology* 388: 15–36.
- [27] Gomes MGM, Rodrigues P, Hilker FM, Mantilla-Beniers NB, Muehlen M, Paulo AC, Medley GF (2007) Implications of partial immunity on

- the prospects for tuberculosis control by post-exposure interventions. *Journal of Theoretical Biology* 248: 608–617.
- [28] Gómez-Corral A, López-García M (2018) Perturbation analysis in finite LD-QBD processes and applications to epidemic models. *Numerical Linear Algebra with Applications* 2018: e2160.
 - [29] Guzzetta G, Ajelli M, Yang Z, Merler S, Furlanello C, Kirschner D (2011) Modeling socio-demography to capture tuberculosis transmission dynamics in a low burden setting. *Journal of Theoretical Biology* 289: 197–205.
 - [30] Halloran ME (2012) The minicommunity design to assess indirect effects of vaccination. *Epidemiologic Methods* 1: article 5.
 - [31] Harris RC, Sumner T, Knight GM, White RG (2016) Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines. *Human Vaccines and Immunotherapeutics* 12: 2813–2832.
 - [32] Knight GM, Griffiths UK, Sumner T, Laurence YV, Gheorghe A, Vassall A, Glaziou P, White RG (2014) Impact and cost-effectiveness of new tuberculosis vaccines in low- and middle-income countries. *Proceedings of the National Academic of Sciences* 111: 15520–15525.
 - [33] Kribs-Zaleta CM, Martcheva M (2002) Vaccination strategies and backward bifurcation in an age-since-infection structured model. *Mathematical Biosciences* 177 & 178: 317–332.
 - [34] Lange C, Chesov D, Heyckendorf J, Leung CC, Udawadia Z, Dheda K (2018) Drug-resistant tuberculosis. An update on disease burden, diagnosis and treatment. *Respirology* 23: 656–673.

- [35] Latouche G, Ramaswami V (1999) Introduction to Matrix Analytic Methods in Stochastic Modeling. ASA-SIAM, Philadelphia, PA.
- [36] Lefèvre C, Simon M (2020) SIR-type epidemic models as block-structured Markov processes. *Methodology and Computing in Applied Probability* 22: 433–453.
- [37] Lietman T, Blower SM (2000) Potential impact of tuberculosis vaccines as epidemic control agents. *Clinical Infectious Diseases* 30: S316–322.
- [38] Liu J, Zhang T (2011) Global stability for a tuberculosis model. *Mathematical and Computer Modelling* 54: 836–845.
- [39] Liu Q, Jiang D (2019) The dynamics of a stochastic vaccinated tuberculosis model with treatment. *Physica A* 527: 121274.
- [40] López-García M (2016) Stochastic descriptors in an SIR epidemic model for heterogeneous individuals in small networks. *Mathematical Biosciences* 271: 42–61.
- [41] Mekonnen HS, Azagew AW (2018) Non-adherence to anti-tuberculosis treatment, reasons and associated factors among TB patients attending at Gondar town health centers, Northwest Ethiopia. *BMC Research Notes* 11: 691.
- [42] Méndez-Samperio P (2018) Development of tuberculosis vaccines in clinical trials: Current status. *Scandinavian Journal of Immunology* 88: e12710.
- [43] Mishra BK, Srivastava J (2014) Mathematical model on pulmonary and multidrug-resistant tuberculosis patients with vaccination. *Journal of the Egyptian Mathematical Society* 22: 311–316.

- [44] Nachega JB, Chaisson RE (2003) Tuberculosis drug resistance: A global threat. *Clinical Infectious Diseases* 36: S24-30.
- [45] Nematollahi MH, Vatankhah R, Sharifi M (2020) Nonlinear adaptive control of tuberculosis with consideration of the risk of endogenous reactivation and exogenous reinfection. *Journal of Theoretical Biology* 486: 110081.
- [46] Neuts MF (1981) *Matrix-geometric Solutions in Stochastic Models: An Algorithmic Approach*. The Johns Hopkins University Press, Baltimore.
- [47] Renardy M, Kirschner DE (2019) Evaluating vaccination strategies for tuberculosis in endemic and non-endemic settings. *Journal of Theoretical Biology* 469: 1–11.
- [48] Shim E (2006) A note on epidemic models with infective immigrants and vaccination. *Mathematical Biosciences and Engineering* 3: 557–566.
- [49] Song B, Castillo-Chavez C, Aparicio JP (2002) Tuberculosis models with fast and slow dynamics: the role of close and casual contacts. *Mathematical Biosciences* 180: 187–205.
- [50] Van den Boogaard J, Kibiki GS, Kisanga ER, Boeree MJ, Aarnoutse RE (2009) New drugs against tuberculosis: Problems, progress, and evaluation of agents in clinical development. *Antimicrobial Agents and Chemotherapy* 53: 849-862.
- [51] Ward JA, Evans J (2019) A general model of dynamics on networks with graph automorphism lumping. In: Aiello LM, Cherifi C, Cherifi H, Lambiotte R, Lió P, Rocha LM (eds). *Complex Networks and Their Applications VII*. Springer, Cham, pp. 445-456.

- [52] Ward JA, López-García M (2019) Exact analysis of summary statistics for continuous-time discrete-state Markov processes on networks using graph-automorphism lumping. *Applied Network Science* 4: 108.
- [53] Witbooi PJ, Muller GE, van Schalkwyk GJ (2015) Vaccination control in a stochastic SVIR epidemic model. *Computational and Mathematical Methods in Medicine* 2015: 271654.
- [54] World Health Organization (2018) BCG vaccine: WHO position paper, February 2018 – Recommendations. *Vaccine* 36: 3408–3410.
- [55] World Health Organization (2019) Global Tuberculosis Report 2019. World Health Organization, Geneva.
- [56] Yu Y, Shi Y, Yao W (2018) Dynamic model of tuberculosis considering multi-drug resistance and their applications. *Infectious Disease Modelling* 3: 362–372.
- [57] Zhang X, Jiang D, Hayat T, Ahmad B (2017) Dynamical behavior of a stochastic SVIR epidemic model with vaccination. *Physica A* 483: 94–108.
- [58] Ziv E, Daley CL, Blower S (2004) Potential public health impact of new tuberculosis vaccines. *Emerging Infectious Diseases* 10: 1529–1535.

Parameter	Description
N	Size of the population
α_U	Per capita infectious contact rate (per year) on unvaccinated susceptible hosts
α_V	Per capita infectious contact rate (per year) on vaccinated susceptible hosts
p_U (respectively, $1 - p_U$)	Probability that an unvaccinated susceptible host becomes infected with active (respectively, latent) TB after an infectious contact
p_V (respectively, $1 - p_V$)	Probability that a vaccinated susceptible host becomes infected with active (respectively, latent) TB after an infectious contact
$r_U(i)$	Per capita reactivation/reinfection rate (per year) on unvaccinated infected hosts with latent TB
$r_V(i)$	Per capita reactivation/reinfection rate (per year) on vaccinated infected hosts with latent TB
δ_{TB}	Per capita death rate (per year) due to the TB infection
δ_D	Per capita death rate (per year) from causes beyond TB
δ_R	Per capita effective treatment rate (per year)
q	Take or fraction of vaccinated hosts in who the vaccine induces some degree of protection
C	Coverage level or fraction of hosts to be vaccinated
$\eta(q, C)$ (respectively, $1 - \eta(q, C)$)	Probability that a new susceptible host is vaccinated and receives some degree of protection (respectively, is unvaccinated)
γ	Per capita duration rate (per year) of the vaccine protection
ξ (respectively, $1 - \xi$)	Probability that a recovered host becomes infected with latent TB (respectively, susceptible)
θ_U (respectively, $1 - \theta_U$)	Probability that a recovered host becoming infected with latent TB is unvaccinated (respectively, is vaccinated)
θ_V (respectively, $1 - \theta_V$)	Probability that a recovered host becoming susceptible is unvaccinated (respectively, is vaccinated)

Table 1: Parameters in the stochastic model with pre-exposure vaccination strategy.

Transition	Infinitesimal rate	Label in Figure 1
$(i, l_U, l_V, s_U, s_V) \rightarrow (i, l_U, l_V, s_U + 1, s_V - 1)$	γs_V	A
$(i, l_U, l_V, s_U, s_V) \rightarrow (i, l_U + 1, l_V, s_U - 1, s_V)$	$\alpha_U(1 - p_U)s_U i$	B
$(i, l_U, l_V, s_U, s_V) \rightarrow (i, l_U, l_V + 1, s_U, s_V - 1)$	$\alpha_V(1 - p_V)s_V i$	C
$(i, l_U, l_V, s_U, s_V) \rightarrow (i + 1, l_U - 1, l_V, s_U, s_V)$	$r_U(i)l_U$	D
$(i, l_U, l_V, s_U, s_V) \rightarrow (i + 1, l_U, l_V - 1, s_U, s_V)$	$r_V(i)l_V$	E
$(i, l_U, l_V, s_U, s_V) \rightarrow (i + 1, l_U, l_V, s_U - 1, s_V)$	$\alpha_U p_U s_U i$	F
$(i, l_U, l_V, s_U, s_V) \rightarrow (i + 1, l_U, l_V, s_U, s_V - 1)$	$\alpha_V p_V s_V i$	G
$(i, l_U, l_V, s_U, s_V) \rightarrow (i - 1, l_U, l_V + 1, s_U, s_V)$	$\delta_R \xi(1 - \theta_U)i$	H
$(i, l_U, l_V, s_U, s_V) \rightarrow (i - 1, l_U + 1, l_V, s_U, s_V)$	$\delta_R \xi \theta_U i$	I
$(i, l_U, l_V, s_U, s_V) \rightarrow (i - 1, l_U, l_V, s_U, s_V + 1)$	$\delta_R(1 - \xi)(1 - \tilde{\theta}_U)i$	J
$(i, l_U, l_V, s_U, s_V) \rightarrow (i - 1, l_U, l_V, s_U + 1, s_V)$	$\delta_R(1 - \xi)\tilde{\theta}_U i$	K
$(i, l_U, l_V, s_U, s_V) \rightarrow (i - 1, l_U, l_V, s_U + 1, s_V)$	$(\delta_{TB} + \delta_D)(1 - \eta(q, C))i$	L
$(i, l_U, l_V, s_U, s_V) \rightarrow (i - 1, l_U, l_V, s_U, s_V + 1)$	$(\delta_{TB} + \delta_D)\eta(q, C)i$	M
$(i, l_U, l_V, s_U, s_V) \rightarrow (i, l_U, l_V, s_U, s_V)$	$\delta_D(1 - \eta(q, C))s_U$	N
$(i, l_U, l_V, s_U, s_V) \rightarrow (i, l_U, l_V, s_U - 1, s_V + 1)$	$\delta_D \eta(q, C)s_U$	O
$(i, l_U, l_V, s_U, s_V) \rightarrow (i, l_U, l_V, s_U + 1, s_V - 1)$	$\delta_D(1 - \eta(q, C))s_V$	P
$(i, l_U, l_V, s_U, s_V) \rightarrow (i, l_U, l_V, s_U, s_V)$	$\delta_D \eta(q, C)s_V$	Q
$(i, l_U, l_V, s_U, s_V) \rightarrow (i, l_U - 1, l_V, s_U + 1, s_V)$	$\delta_D(1 - \eta(q, C))l_U$	R
$(i, l_U, l_V, s_U, s_V) \rightarrow (i, l_U - 1, l_V, s_U, s_V + 1)$	$\delta_D \eta(q, C)l_U$	S
$(i, l_U, l_V, s_U, s_V) \rightarrow (i, l_U, l_V - 1, s_U + 1, s_V)$	$\delta_D(1 - \eta(q, C))l_V$	T
$(i, l_U, l_V, s_U, s_V) \rightarrow (i, l_U, l_V - 1, s_U, s_V + 1)$	$\delta_D \eta(q, C)l_V$	U

Table 2: Transitions between states and underlying rates of \mathcal{X} , and labels used in Figure 1.

Parameter	Value	Reference	Description in the case study
N	20	Assumption	Case study on twenty individuals
α_U	5.2 <i>per year</i>	World Health Organization [54]	The 10% of respiratory contacts are infectious and, on average, one respiratory contact is assumed to occur per week
α_V	α_U	Méndez-Samperio [42]	The vaccine does not protect from latent infection, but it prevents a fast progression to active TB
p_U	0.15	Knight et al [32]	Only the 15% of infectious result in a fast progression to active TB
p_V	0.0	Knight et al [32]	Perfect protection of the vaccine against TB
a_U	1.5×10^{-4} <i>per year</i>	Knight et al [32]	Reinfection of unvaccinated infected individuals with latent TB is highly unlikely
b_U	0.35 a_U	Knight et al [32]	Latent infection protects an individual in 65% of infectious respiratory contacts
$r_V(i)$	0.0	Méndez-Samperio [42]	Vaccinated infected individuals with latent TB cannot suffer from neither endogenous reactivation nor exogenous reinfection
δ_{TB}^{-1}	5 <i>years</i>	Renardy and Kirschner [47]	On average, the life expectancy of an individual with untreated TB disease is 5 years
δ_D^{-1}	35 <i>years</i>	Assumption	Remaining lifetime expectancy of 30-35 aged individuals
δ_R^{-1}	9 <i>months</i>	World Health Foundation [54]	On average, the duration of the TB treatment takes 6 – 9 months
q	{0.4, 0.6, 0.8}	Scenarios 1-3	Prospective analysis of three vaccines with small, moderate and high values of the take
C	{0.2, 0.5, 0.8}	Numerical purposes	Small, intermediate and high coverage levels
γ^{-1}	10 <i>years</i>	Knight et al [32]	Expected duration of the vaccine protection
ξ	1.0	Lietman and Blower [37]	A single strain of the pathogen is in progress
θ_U	0.5	Cliffs et al [18]	The immunity structure in subsequent infections suffered by recovered individuals is not well documented

Table 3: Parameters, values and related references.

C	Descriptor	Scenario 1	Scenario 2	Scenario 3
0.2	$E[\mathcal{R}_{exact,0} X(0) = (1, 0, 0, 19, 0)]$	3.47577	3.46336	3.45073
	$E[\tau X(0) = (1, 0, 0, 19, 0)]$	81.75810	57.42145	41.38551
	$E[I_{\max} X(0) = (1, 0, 0, 19, 0)]$	16.39818	16.28102	16.16770
	$E[V X(0) = (1, 0, 0, 19, 0)]$	88.94354	93.18696	89.08443
0.5	$E[\mathcal{R}_{exact,0} X(0) = (1, 0, 0, 19, 0)]$	3.43788	3.40468	3.36974
	$E[\tau X(0) = (1, 0, 0, 19, 0)]$	30.56972	15.83522	9.31597
	$E[I_{\max} X(0) = (1, 0, 0, 19, 0)]$	16.05697	15.78699	15.52263
	$E[V X(0) = (1, 0, 0, 19, 0)]$	81.85347	62.94112	48.99816
0.8	$E[\mathcal{R}_{exact,0} X(0) = (1, 0, 0, 19, 0)]$	3.39784	3.34033	3.27676
	$E[\tau X(0) = (1, 0, 0, 19, 0)]$	14.10538	6.59369	3.90801
	$E[I_{\max} X(0) = (1, 0, 0, 19, 0)]$	15.73376	15.31383	14.90205
	$E[V X(0) = (1, 0, 0, 19, 0)]$	59.69849	41.44763	32.62727

Table 4: Expectations of $\mathcal{R}_{exact,0}$, τ , I_{\max} and V at an invasion time versus $C \in \{0.2, 0.5, 0.8\}$, for values $q = 0.4$ (Scenario 1), 0.6 (Scenario 2) and 0.8 (Scenario 3).

C	Descriptor	Scenario 1	Scenario 2	Scenario 3
0.2	$E[L_U(\tau) X(0) = (1, 0, 0, 19, 0)]$	3.67901	3.54763	3.41990
	$E[L_V(\tau) X(0) = (1, 0, 0, 19, 0)]$	12.52571	12.66117	12.79243
	$E[S_U(\tau) X(0) = (1, 0, 0, 19, 0)]$	3.63377	3.54960	3.46632
	$E[S_V(\tau) X(0) = (1, 0, 0, 19, 0)]$	0.16151	0.24160	0.32135
0.5	$E[L_U(\tau) X(0) = (1, 0, 0, 19, 0)]$	3.29539	2.99566	2.70842
	$E[L_V(\tau) X(0) = (1, 0, 0, 19, 0)]$	12.91999	13.22537	13.51569
	$E[S_U(\tau) X(0) = (1, 0, 0, 19, 0)]$	3.38384	3.18059	2.98081
	$E[S_V(\tau) X(0) = (1, 0, 0, 19, 0)]$	0.40078	0.59838	0.79508
0.8	$E[L_U(\tau) X(0) = (1, 0, 0, 19, 0)]$	2.93736	2.48484	2.04731
	$E[L_V(\tau) X(0) = (1, 0, 0, 19, 0)]$	13.28448	13.74007	14.17502
	$E[S_U(\tau) X(0) = (1, 0, 0, 19, 0)]$	3.14039	2.82287	2.51056
	$E[S_V(\tau) X(0) = (1, 0, 0, 19, 0)]$	0.63777	0.95222	1.26711

Table 5: Conditional expectations of the random numbers $L_U(\tau)$, $L_V(\tau)$, $S_U(\tau)$ and $S_V(\tau)$, provided that $X(0) = (1, 0, 0, 19, 0)$, versus $C \in \{0.2, 0.5, 0.8\}$, for values $q = 0.4$ (Scenario 1), 0.6 (Scenario 2) and 0.8 (Scenario 3).

Algorithm	Computational complexity
1	$O\left(\sum_{i=1}^N (J(i))^3\right)$
2	$O\left(\sum_{i=1}^N (J'(i))^3 + (1 + V_{q'}) \sum_{i=1}^N (J'(i))^2\right)$
3	$O\left(\sum_{i=1}^N (J'(i))^3\right)$
4	$O\left(\sum_{i=1}^N (J'(i))^3 + (1 + R_{q'}) \sum_{i=1}^N (J'(i))^2\right)$
5	$O\left(\sum_{i=1}^N (J'(i))^3\right)$

Table 6: Computational complexities of Algorithms 1-5.

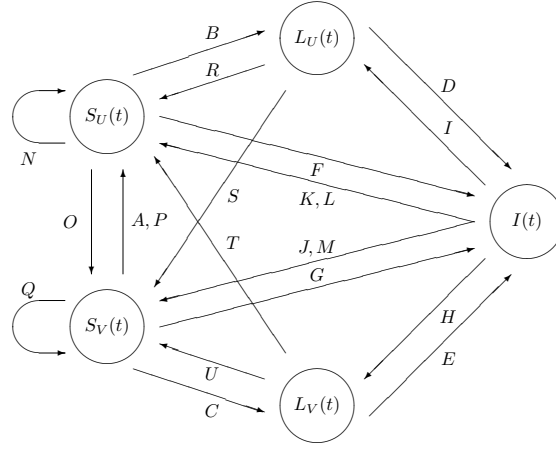


Figure 1: Diagram of transitions among compartments and underlying events (Table 2) in the pre-exposure vaccine model.

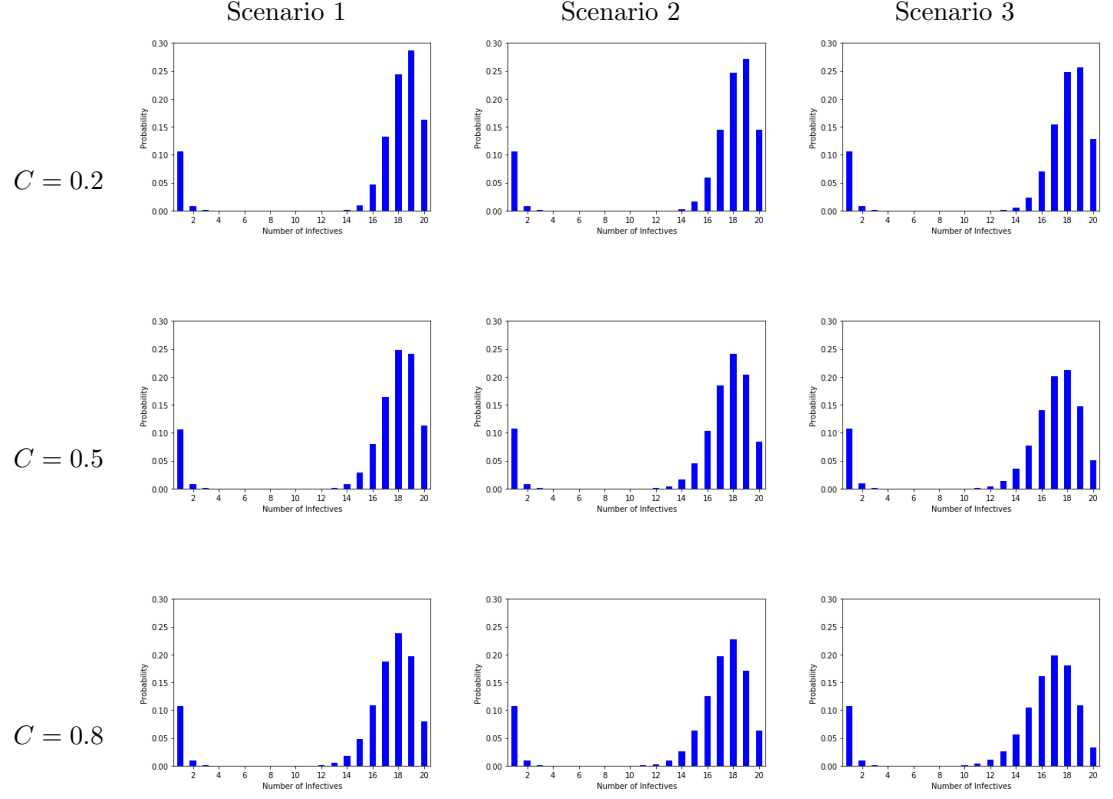


Figure 2: Mass function of I_{\max} , provided that $X(0) = (1, 0, 0, 19, 0)$, versus $C \in \{0.2, 0.5, 0.8\}$, for values $q = 0.4$ (Scenario 1), 0.6 (Scenario 2) and 0.8 (Scenario 3).

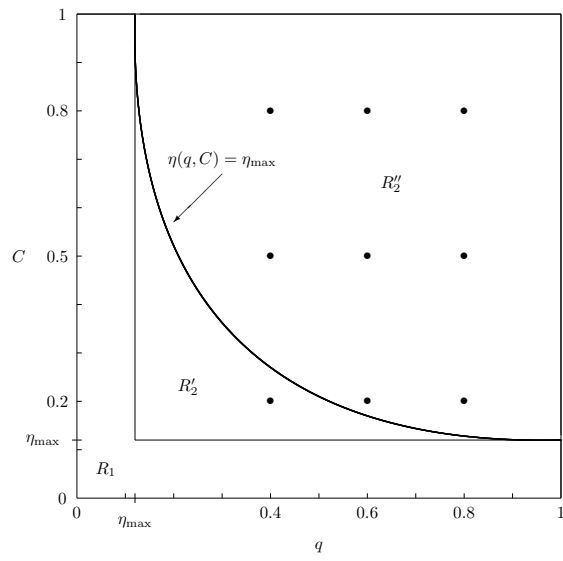


Figure 3: A graphical representation of the threshold η_{\max} , the resulting regions R_1 and $R_2 = R'_2 \cup R''_2$, and the points (q, C) for $q = 0.4$ (Scenario 1), 0.6 (Scenario 2) and 0.8 (Scenario 3), and $C \in \{0.2, 0.5, 0.8\}$.

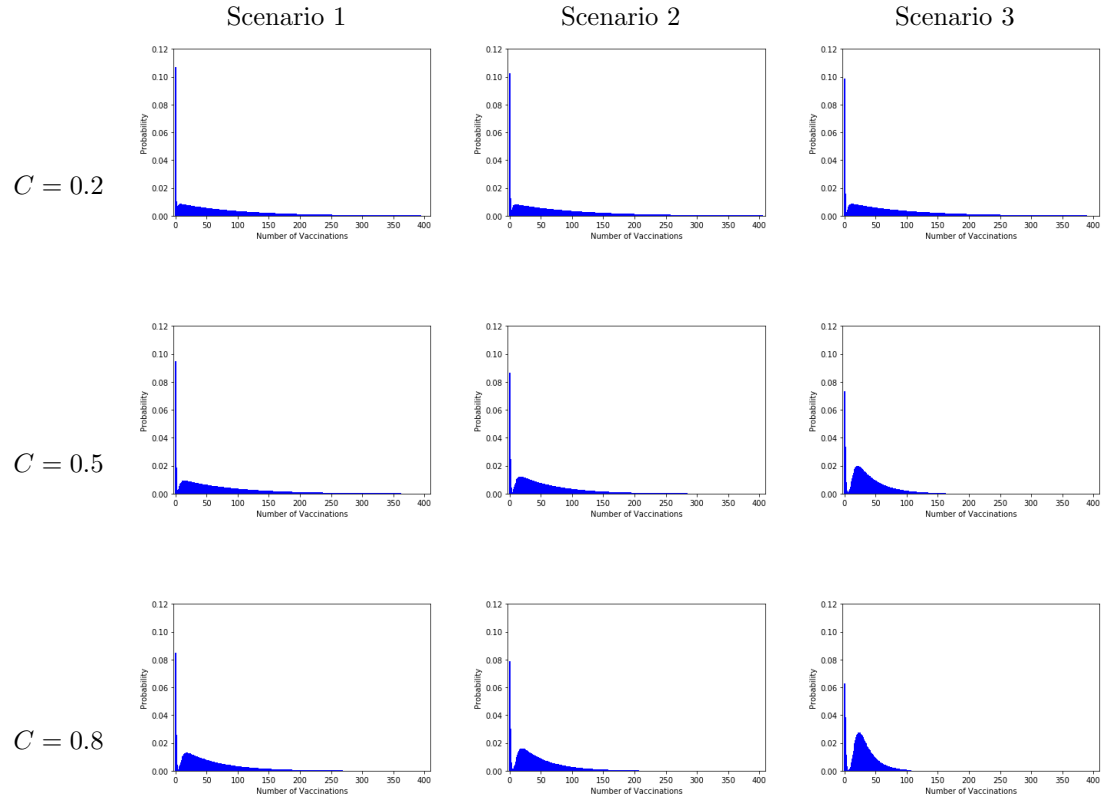


Figure 4: Mass function of V , provided that $X(0) = (1, 0, 0, 19, 0)$, versus $C \in \{0.2, 0.5, 0.8\}$, for values $q = 0.4$ (Scenario 1), 0.6 (Scenario 2) and 0.8 (Scenario 3). For convenience, the mass function of V has been plotted on a common set of integers $v \in \{0, \dots, v'\}$, with $v' \geq V_{q'}$, for all selection of the pair (q, C) and $q' = 0.99$.

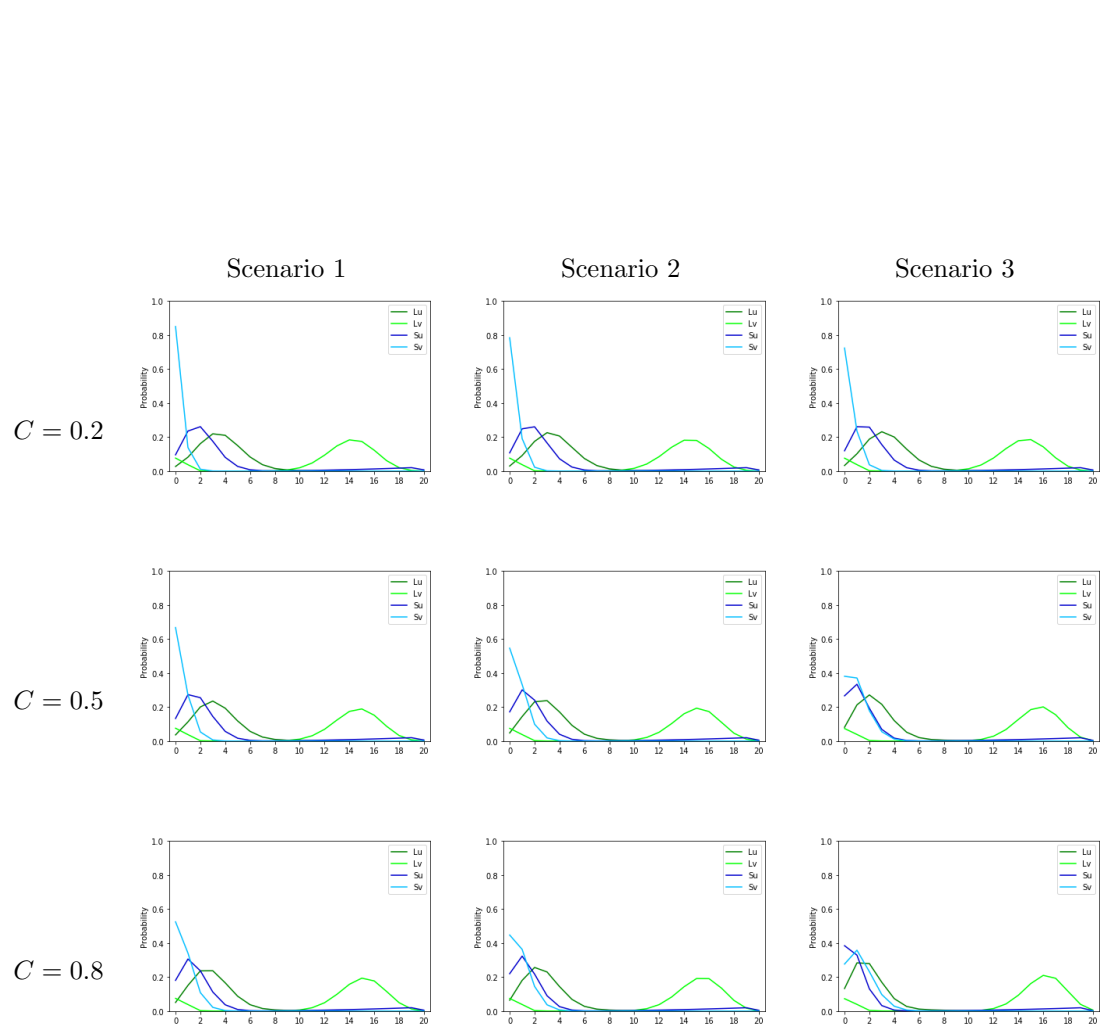


Figure 5: Mass functions of $L_U(\tau)$ (dark green), $L_V(\tau)$ (light green), $S_U(\tau)$ (dark blue) and $S_V(\tau)$ (light blue), provided that $X(0) = (1, 0, 0, 19, 0)$, versus $C \in \{0.2, 0.5, 0.8\}$, for values $q = 0.4$ (Scenario 1), 0.6 (Scenario 2) and 0.8 (Scenario 3). For convenience, a polygonal line is used instead of the corresponding histogram.