

Accepted Manuscript

Stochastic dynamics of an epidemic with recurrent spillovers from an endemic reservoir

Marina Voinson, Alexandra Alvergne, Sylvain Billiard, Charline Smadi

PII: S0022-5193(18)30396-5
DOI: <https://doi.org/10.1016/j.jtbi.2018.08.017>
Reference: YJTBI 9582



To appear in: *Journal of Theoretical Biology*

Received date: 2 November 2017
Revised date: 12 July 2018
Accepted date: 12 August 2018

Please cite this article as: Marina Voinson, Alexandra Alvergne, Sylvain Billiard, Charline Smadi, Stochastic dynamics of an epidemic with recurrent spillovers from an endemic reservoir, *Journal of Theoretical Biology* (2018), doi: <https://doi.org/10.1016/j.jtbi.2018.08.017>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Highlights

- Stochastic SIR with reservoir to describe the dynamics of pathogens in a host
- Branching process approximations are provided
- Recurrent spillover cause multiple outbreaks even for a pathogen barely contagious
- Spillover and direct transmission have similar importance for pathogens dynamics

Stochastic dynamics of an epidemic with recurrent spillovers from an endemic reservoir

Marina Voinson^{a,*}, Alexandra Alvergne^b, Sylvain Billiard^a, Charline Smadi^{c,d}

^aUniv. Lille, CNRS, UMR 8198 - Evo-Eco-Paleo, F-59000 Lille, France

^bSchool of Anthropology and Museum Ethnography, University of Oxford, OX2 6PE Oxford, UK

^cIRSTEA UR LISC, Laboratoire d'ingénierie pour les Systèmes Complexes, 9 avenue Blaise-Pascal CS 20085, 63178 Aubière, France

^dComplex Systems Institute of Paris Ile-de-France, 113 rue Nationale, Paris, France

Abstract

Most emerging human infectious diseases have an animal origin. While zoonotic diseases originate from a reservoir, most theoretical studies have principally focused on single-host processes, either exclusively humans or exclusively animals, without considering the importance of animal to human transmission (i.e. spillover transmission) for understanding the dynamics of emerging infectious diseases. Here we aim to investigate the importance of spillover transmission for explaining the number and the size of outbreaks. We propose a simple continuous time stochastic Susceptible-Infected-Recovered model with a recurrent infection of an incidental host from a reservoir (e.g. humans by a zoonotic species), considering two modes of transmission, (1) animal-to-human and (2) human-to-human. The model assumes that (i) epidemiological processes are faster than other processes such as demographics or pathogen evolution and that (ii) an epidemic occurs until there are no susceptible individuals left. The results show that during an epidemic, even when the pathogens are barely contagious, multiple outbreaks are observed due to spillover transmission. Overall, the findings demonstrate that the only consideration of direct transmission between individuals is not sufficient to explain the dynamics of zoonotic pathogens

*Corresponding author

Email address: marina.voinson.etu@univ-lille.fr (Marina Voinson)

in an incidental host.

Keywords: Branching processes, multi-host processes, animal-human interface, emerging infectious diseases, spillover transmission, R_0 , zoonotic pathogens

1. Introduction

Recent decades have seen a surge of emerging infectious diseases (EIDs), with up to forty new diseases recorded since the 1970s [1]. Sixty percent of emerging human infectious diseases are zoonotic, i.e. are caused by pathogens that have an animal origin [1, 2]. The World Health Organization defines zoonotic pathogens as “pathogens that are naturally transmitted to humans via vertebrate animals”. The epidemics caused by EIDs impact the societal and economical equilibria of countries by causing unexpected deaths and illnesses thereby increasing the need for health care infrastructures and by interfering with travel [3]. Moreover, the risk of EIDs being transmitted to humans from wildlife is increasing because of the recent growth and geographic expansion of human populations, climate change and deforestation, which all increase the number of contacts between humans and potential new pathogens [1, 4, 5]. Given most EIDs have an animal origin, it is crucially important to understand how infections spread from animal to human populations, i.e. by spillover transmission.

There is numerous empirical evidence that the epidemiological dynamics of infectious diseases is highly dependent on the transmission from the reservoir (the reservoir will be defined following Ashford’s definition (1997), i.e. a pathogen is persistent in the environment of the incidental host, see Table 1 for details). The start of an outbreak is promoted by a primary contact between the reservoir and the incidental host (i.e. host that becomes infected but is not part of the reservoir) leading to the potential transmission of the infection to the host population. Moreover, multiple outbreaks are commonly observed during an epidemic of zoonotic pathogens in human populations, for instance in the case of the epidemic of the Nipah Virus between 2001 and 2007 [6]. With

regards to the Ebola virus, some twenty outbreaks have been recorded since the discovery of the virus in 1976 [7]. This number of outbreaks undoubtedly underestimates the total number of emergences because not all emergences necessarily lead to the spread of the infection from an animal reservoir to the host population [8]. While the reservoir has an important role for causing the emergence of outbreaks, the role of spillover transmission on the incidental epidemiological dynamics is rarely discussed.

Empirically, it is generally difficult to distinguish between direct transmission and transmission from the reservoir. Only in the case of non-communicable diseases it is easily possible to measure the importance of the recurrent transmission from the reservoir, since all infected individuals originate from a contact with the reservoir. For instance, the H7N9 virus, for which most human cases are due to a contact with an infected poultry, approximately 132 spillovers have been listed during the epidemic of 2013 [9]. For pathogens that are able to propagate from one individual to another, the origin of the infection can be established according to patterns of contacts during the incubation period [6, 10]. Most often, if an infected individual has been in contact with another infected individual in his recent past, direct transmission is considered as the likeliest origin of the infection. However, both individuals might have shared the same environment and thus might have been independently infected by the reservoir. This leads to overestimating the proportion of cases that result from person-to-person transmission. Moreover, when the pathogen infects an individual and the latter does not produce secondary cases then the detection of emergence is unlikely.

Pathogen spillover is often neglected in epidemiological theoretical models. It is generally assumed that the epidemiological dynamics of outbreaks is driven by the ability of the pathogen to propagate within hosts. For instance, a classification scheme for pathogens has been proposed by Wolfe et al. (2007), including five evolutionary stages in which the pathogen may evolve ranging from an exclusive animal infection (Stage I) to an exclusive human infection (Stage V) (Figure 1) [15]. The intermediate stages are those found for the

Definitions of a reservoir	Authors	Refs
“any animal, person, plant, soil, substance or combination of any of these in which the infectious agent normally lives”	CDC	[11]
“all hosts, incidental or not, that contribute to the transmission to the target host (i.e. the population of interest) in which the pathogen can be permanently maintained”	Haydon et al.	[12]
“an ecologic system in which an infectious agent survives indefinitely”	Ashford	[13, 14]

Table 1: Definitions of a reservoir from the literature. The reservoir is mostly used as defined by the Centre for Disease Control and prevention (CDC). Two other definitions have been proposed to clarify and complete the notion of reservoir in the case of zoonotic pathogens. On the one hand, Haydon et al. (2002) define the reservoir from a practical point of view in order to take into account all hosts epidemiologically connected to the host of interest (i.e. target host), to implement better control strategies. On the other hand, Ashford (1997) establishes a more generalizable definition: for a given pathogen there is a single reservoir.

zoonotic pathogens (Stages II-IV). Lloyd-Smith et al. (2009), proposed to enhance the classification scheme by differentiating the Stages II-IV by the ability of pathogens to propagate between individuals in the incidental host (i.e. as
60 a function of the basic reproductive ratio R_0): the non-contagious pathogens ($R_0 = 0$, Stage II), pathogens barely contagious inducing stuttering chains of transmission ($0 < R_0 < 1$, Stage III) and contagious pathogens inducing large outbreaks ($R_0 > 1$, Stage IV) [16]. However, the role of the reservoir is not clearly defined, and spillover effects on the epidemiological dynamics are not
65 discussed.

Only a few models have investigated the dynamics of EIDs by taking into account explicitly the transmission from the reservoir to the incidental host. Lloyd-Smith et al. (2009) have analysed 442 modelling studies of zoonotic pathogens and concluded that models incorporating spillover transmission are
70 dismayingly rare [16]. More recent models aimed at investigating the dynamics of EIDs by taking into account the spread of the pathogen using multi-host processes but disregarding the persistence of the pathogen in the reservoir [17], or by focusing on the dynamics and conditions of persistence of the pathogen between two populations [18]. Models that have considered an endemic reservoir
75 are disease-specific and do not generate generalizable dynamics [10, 19]. More

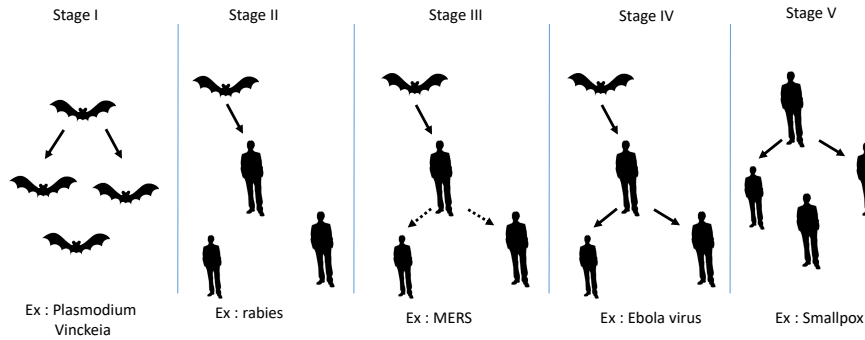


Figure 1: Representation of the classification scheme of pathogens proposed by Wolfe et al. (2007). A pathogen may evolve from infecting only animals (Stage I) to infecting only humans (Stage V). Each stage corresponds to a specific epidemiological dynamics in the incidental host. Stage II corresponds to few spillovers from animals (e.g. bats) to humans with no possible transmission between humans. Stage III corresponds to few stuttering chains of transmission between humans that go extinct (no outbreaks). Stage IV corresponds to large outbreaks in human population but the pathogen cannot be maintained without the reservoir.

recently, Singh and Myers (2014) developed a Susceptible-Infected-Recovered (SIR) stochastic model coupled with a constant force of infection. The authors are mostly interested in the effect of population size and its impact on the size of an outbreak [20]. However, this approach does not allow teasing apart the contribution of the incidental host transmission from that of the transmission from the reservoir in modulating the dynamics of zoonotic pathogens.

In this paper, we aim to provide general insights into the dynamics of a zoonotic pathogen (i.e. pathogens classified in stages II-IV) emerging from a reservoir and its ability to propagate in an incidental host. To do so, we developed a continuous time stochastic model that can dissociate the effect of between-host (i.e. direct) transmission from the effect of spillover (i.e. reservoir) transmission. A multi-host process with a reservoir and an incidental host is considered. The epidemiological processes are stochastic, which is particularly relevant in the case of transmission from the reservoir and more realistic because

only a small number of individuals are expected to be infected at the beginning of an outbreak. The model makes a number of assumptions. First, the epidemiological processes are much faster than the demographic processes. Second, the pathogen in the reservoir is considered as endemic and might contaminate recurrently the incidental host. Third, an individual cannot become susceptible after having been infected. As a consequence, the total number of susceptible individuals in the incidental host decreases during the epidemic. This is what is expected for an epidemic spreading locally during a short period of time (at the scale of a few thousands individuals during weeks or months, depending on the disease and populations considered). We then harness the model to predict the effects of both spillover transmission and direct transmission on the number and the size of outbreaks. Outbreaks occur when the number of cases of disease increases above the epidemiological threshold. In the case of non emerging infectious diseases, an epidemiological threshold is used to gauge the start of outbreaks. For instance for the seasonal influenza the epidemiological threshold is calculated depending on the incidence of the disease during the previous years [21]. In the case of emerging infectious diseases, no incidence is normally expected in the population so from a small number of infected individuals, the outbreak can be considered to spread. We show that, regarding the epidemiological dynamics, the recurrent emergence of the pathogen from the reservoir in the incidental host is as important as the transmission between individuals of the incidental host. We conclude by discussing the implications of these results for the classification of pathogens proposed by Lloyd-Smith et al. (2009).

2. Model

A continuous time stochastic Susceptible-Infected-Recovered (SIR) compartmental transmission model [22] with recurrent introduction of the infection into an incidental host by a reservoir is considered (Figure 2). Our goal here is not to study a disease in particular but to provide general insights of the reservoir effect on the epidemiological dynamics of the incidental host. The infection is assumed

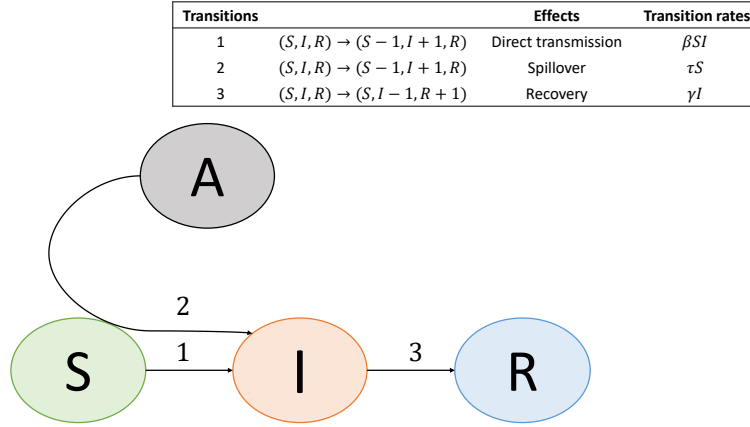


Figure 2: Representation of the stochastic model with transitions. A reservoir (A) has been added to a classical SIR model where the pathogen is persistent. Individuals are characterized by their epidemiological status in the incidental host (S : susceptible; I : infected; R : recovered). S , I and R are measured in numbers of individuals. All stochastic transitions with associated rates are listed in the table. A susceptible individual becomes infected through the transmission by contact at rate βSI or through the reservoir at rate τS . An infected individual recovers at rate γ .

to propagate quickly relatively to other processes such as pathogen evolution
 120 and demographic processes. The reservoir is defined as a compartment where
 the pathogen is persistently maintained, this pathogen is then considered as
 endemic. The population is fully mixed. An individual can be infected through
 two types of transmission, from the reservoir by the spillover transmission and
 by direct contact between individuals. We neglect the possibility for reverse
 125 infection from the incidental host to the reservoir.

The incidental host is composed of N individuals. The infection can spillover
 by contact between the reservoir and the incidental host at rate τS where S is
 the number of susceptible individuals and τ is the rate at which an individ-
 ual becomes infected from the reservoir. In the incidental host, the infection
 130 can propagate by direct contact at rate βSI where I is the number of infected
 individuals and β is the individual rate of infection transmission. An infected in-
 dividual can recover at rate γ . The propensity of the pathogen to be transmitted
 between individuals within host is expressed in terms of the basic reproductive

ratio of the pathogen, R_0 , which is widely used in epidemiology. R_0 corresponds to the average number of secondary infections produced by an infected individual in an otherwise susceptible population. In a deterministic model, for a pathogen to invade the population, R_0 must be larger than 1 in the absence of reservoir. In a stochastic model, the higher the R_0 the higher the probability for the pathogen to invade the population. In a SIR model, the basic reproductive ratio R_0 equals to $\beta N/\gamma$. Individuals in the recovered compartment do not contribute anymore to the transmission process. Since we assume that demographic processes are slower than epidemic processes, the number of susceptible individuals decreases during the epidemic due to the consumption of susceptibles by the infection until the extinction of the susceptible population. In other words, in our model, R_0 will decrease because of the successive spillovers from the reservoir. We expect this to occur especially at short space and time scales (a local population during the course of weeks or months).

To analyse the dynamics in the incidental host, three statistics will be studied (i) the mean number of outbreaks, (ii) the mean size of the recurrent outbreaks during an epidemic and (iii) the mean size of the largest outbreak occurring during an epidemic. We consider the appearance of an outbreak when the incidence of the infection exceeds the threshold c and define the maximum size of an outbreak as the largest number of infected individuals during the largest outbreak.

2.1. Analysis of the model

Stochastic simulations. The epidemiological dynamics described previously can be simulated with the following algorithm (simulations were run in C++). The population state is assumed to be known at time t . A total event rate (Ω), only depending of the state of the population at time t , is calculated for each iteration.

- A) The total event rate Ω of the continuous time stochastic *SIR* model is given by:

$$\Omega = \beta SI + \tau S + \gamma I.$$

B) The next event time is $t' = t + \delta$ where δ is exponentially distributed with parameter Ω .

C) The next event to occur is randomly chosen: direct transmission, spillover transmission or recover with respective probabilities $\beta SI/\Omega$, $\tau S/\Omega$ and $\gamma I/\Omega$.

165

We performed stochastic individual-based simulations of the epidemics with spillover transmission, using rates as presented in Figure 2. The incidental host is initially ($t = 0$) composed of 1000 susceptible individuals ($N = S = 1000$). The infection is considered as endemic in the reservoir. Simulations are stopped when there are no susceptible individuals anymore. An outbreak begins when the number of infected individuals reaches the epidemiological threshold c ($c = 5$ infected individuals in the simulations) and ends when there is no infected individuals anymore ($I = 0$). Stochastic simulations were run for values of the basic reproductive ratio (R_0) ranging from 0 to 10 and of the spillover transmission (τ) ranging from 10^{-10} to 10^{-1} , 10 000 simulations are performed for each parameter set. All other parameter values are detailed in Table 2.

170

175

Parameters	Values
R_0	variable
γ	0.1 UT^{-1}
τ	variable (UT^{-1})
c	5 infected individuals
N	1000 individuals

Table 2: Parameters used and their values. UT denotes the unit of time which can be expressed in days or weeks.

Approximation by a branching process. The epidemiological model with recurrent introduction of the infection into an incidental host by a reservoir can be approximated by a branching process with immigration from the reservoir to the incidental host at the beginning of the infectious process (thus assuming that individual “birth” and “death” rates of infected individuals are constant

180

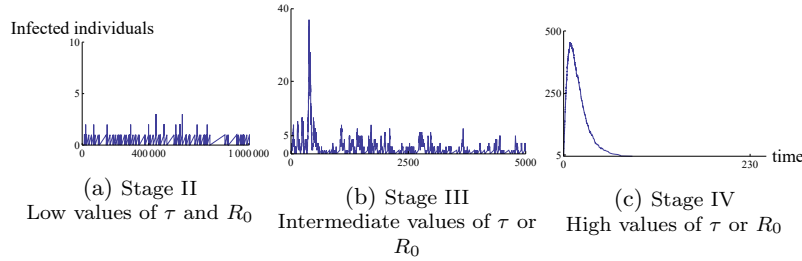


Figure 3: Examples of stochastic simulations. Three examples of epidemiological dynamics corresponding to the three Stages II, III and IV (see Figure 1) respectively with low values of both direct and spillover transmissions ($R_0 = 0.2$ and $\tau = 10^{-7}$), intermediate values of direct or spillover transmission ($R_0 = 1.5$ and $\tau = 10^{-4}$), high value of direct or spillover transmission ($R_0 = 2$ and $\tau = 10^{-1}$).

during the starting phase of an outbreak). The individual birth and death rates are respectively βN , the transmission rate and γ , the recovery rate. The immigration rate corresponds to the spillover rate, τN , at the beginning of the infection. In other words, we assume that the number of susceptibles is N to study the beginning of the infection, which is a good approximation as long as few individuals have been infected. We distinguish between two regimes in the incidental host, the subcritical regime when $R_0 < 1$ and the supercritical regime when $R_0 > 1$. We suppose that at time $t = 0$ a single individual is infected by the spillover transmission.

3. Results

3.1. The epidemiological dynamics in the incidental host

As illustrated in Figure 3, three patterns are observed (i) a stuttering chain of transmission that goes extinct, i.e. infection spreads inefficiently, (corresponding to Stage II in Wolfe's classification, see Figure 1), (ii) a large outbreak and few stuttering chains of transmission (corresponding to Stage III, see Figure 1) and (iii) a single large outbreak consuming a large number of susceptible individuals (corresponding to Stage IV see Figure 1).

Figure 4 shows the roles of the direct transmission, R_0 , and the spillover transmission, τ , in the occurrence of the three patterns depicted in Figure 3, in

the case of a threshold $c = 5$ and $c = 10$. Stuttering chains of transmission occur when the pathogen is barely contagious between individuals (small R_0) and when the recurrent emergence of the pathogen (τ) is low. At the opposite, when the pathogen is highly contagious (large R_0) or when the spillover transmission is high (τ), only one large outbreak is observed. Finally, both dynamics (a large outbreak and stuttering chains of transmission) are observed for intermediate value of both direct and spillover transmission. The dynamics observed in the three stages depend both of the value on the direct transmission (R_0) and the effect of the reservoir (τ).

The three stages are observed when spillover transmission (τ) or the direct transmission (R_0) is low, two stages appear when direct or spillover transmission is intermediate and only one stage is observed when at least one of the two transmission is high. Hence, both types of transmission are important in the emergence of epidemiological dynamics.

Figure 4 shows that the epidemiological threshold c little affects the observed dynamics. For both values $c = 5$ and $c = 10$, the three patterns are observed. However, the values of both transmissions (τ and R_0) on the frontiers between stages are different. The parameter range for which the patterns of the stages II and IV are observed are wider for a higher threshold value (compare Fig. 4a and 4b), whereas stage III is narrower. When the direct and the spillover transmissions are low, it is more difficult for the infection to reach a higher threshold. Thus, there are more stuttering chains of transmission. In the same way, when the direct or the spillover transmission is high, a large outbreak is observed then some stuttering chains of transmission occur but do not reach the epidemiological threshold.

For both values of epidemiological threshold (Figure 4a and 4b), a “bulb” is observed in the Stage III where the direct transmission is high. After the occurrence of a large outbreak, the susceptible population becomes small. Hence the next excursion is very unlikely to reach the epidemiological threshold. However, a high enough spillover transmission rate is able to counterbalance the small effective R_0 and to produce other outbreaks after the large one. The

“bulb” is less pronounced in the case of a higher threshold (Figure 4b) because the susceptible population consumed during the large outbreak is important leading to the failure for the next excursion to reach a high epidemiological threshold.

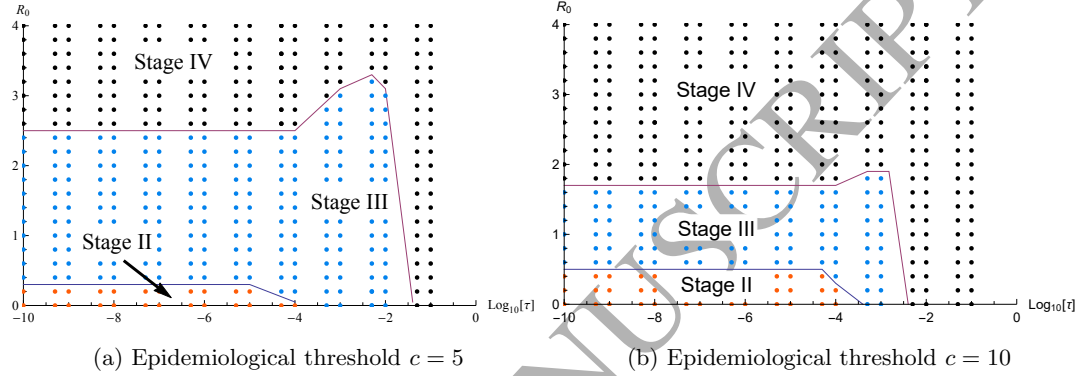


Figure 4: Classification scheme of the epidemiological dynamics observed in simulations. The general epidemiological dynamics is depicted as a function of the direct transmission (R_0) and the spillover transmission (τ). The epidemiological dynamics of stochastic simulations are classified following the stages described by Wolfe et al. (2007), Stage II: stuttering chains of transmission (i.e. less than one outbreak), Stage III: one large outbreak and stuttering chains of transmission (i.e. more than one outbreak) and Stage IV: a single large outbreak.

3.2. Number of outbreaks when the effect of the reservoir is low

3.2.1. Case of a barely contagious pathogen ($R_0 < 1$):

We aim at approximating the mean number of outbreaks in the case where the spillover transmission rate τ and the reproductive number R_0 are small (subcritical case corresponding to $R_0 < 1$). The method of approximation is the following: let us denote by S_i the number of susceptible individuals at the beginning of the i -th excursion. During the i -th excursion, we set this number of susceptibles to its initial value S_i , and consider that the rate of new infections is $\beta I S_i$. We thus obtain a branching process with individual birth (infection) rate βS_i and individual death (recovery) rate γ . When there is no more infected individuals, we compute the mean number of recovered individuals produced by this branching process excursion, denoted by $\mathbb{E}[K(S_i, \beta, \gamma)]$, and make the

approximation that

$$S_{i+1} = S_i - \mathbb{E}[K(S_i, \beta, \gamma)], \quad (1)$$

where $\mathbb{E}[K(S_i, \beta, \gamma)]$ can be computed and equals (see Appendix A):

$$\mathbb{E}[K(S_i, \beta, \gamma)] = \frac{\gamma}{\gamma + \beta S_i} \sum_{k=0}^{\infty} \frac{(2k)!}{(k!)^2} \left(\frac{\gamma \beta S_i}{(\gamma + \beta S_i)^2} \right)^k. \quad (2)$$

In other words, the initial number of susceptible individuals for the $i + 1$ -th excursion is the initial number of susceptible individuals for the i -th excursion minus the mean number of recovered individuals produced during the i -th excursion under our branching process approximation. We repeat the procedure for the $i + 1$ -th excursion, and so on, until k satisfies $S_k > 0$ and $S_{k+1} \leq 0$ (no susceptible anymore). In order to be considered as an outbreak, an excursion has to exceed c individuals, where we recall that c is the epidemiological threshold. Under our branching process approximation, the probability for the i -th excursion to reach the epidemiological threshold (see Appendix A) is:

$$P(S_i, \beta, \gamma) = \frac{(\gamma/\beta S_i) - 1}{(\gamma/\beta S_i)^c - 1}. \quad (3)$$

As a consequence, our approximation of the mean number of outbreaks ($\mathbb{E}[O(N, \beta, \gamma)]$) reads:

$$\mathbb{E}[O(N, \beta, \gamma)] = \sum_{i=0}^k P(S_i, \beta, \gamma), \quad (4)$$

where $S_1 = N$, and the S_i 's are computed as described in (1).

The mean number of outbreaks computed with the branching process is a good approximation compared to numerical simulations for a small spillover transmission ($10^{-10} \leq \tau \leq 10^{-6}$) (Figure 5). The spillover transmission added in our model introduces the infection recurrently and allows the infection to spread even for a pathogen barely contagious ($R_0 < 1$). According to Figure 5 when $R_0 < 1$ the number of outbreaks increases when the direct transmission between individuals increases. Indeed, the higher the direct transmission, the higher the probability for the excursions to reach the epidemiological threshold

Average number of Outbreaks

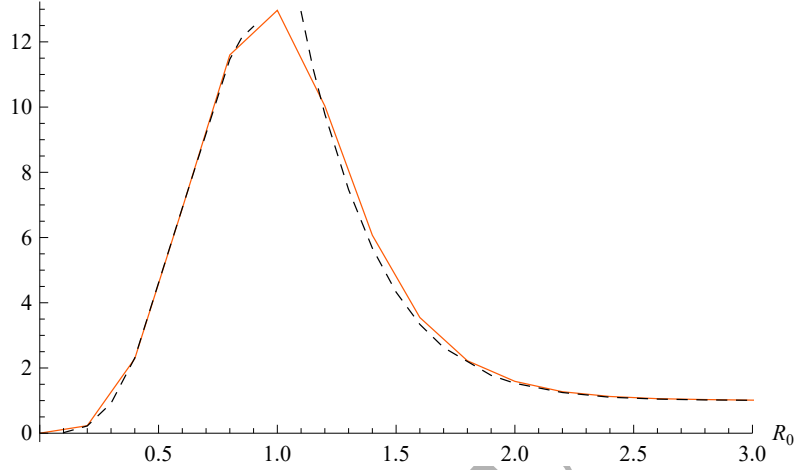


Figure 5: Average number of outbreaks for a low spillover transmission. The orange and black dotted curves represent the results, from numerical simulations and branching process approximation, respectively. The average number of outbreaks approximated is evaluated when the spillover transmission τ is small. For the numerical simulations, $\tau = 10^{-10}$ has been chosen. There is a break in the dotted curve (branching process) because our approximation is not valid in the critical regime (when R_0 is close to 1).

(c). The number of outbreaks can be high because when the direct transmission is smaller than 1, the infection spreads inefficiently and does not consume a large number of susceptibles allowing the next excursion to exceed the epidemiological threshold.

Figure 6 shows that the average number of outbreaks is a non-monotonic function of the direct transmission (R_0) and the spillover transmission (τ). More precisely, figure 6b shows that for intermediate and low values of spillover transmission ($10^{-6} \leq \tau \leq 10^{-4}$), the average number of outbreaks increases until $R_0 \sim 1$ then decreases until it reaches one outbreak when the direct transmission is high ($R_0 > 2.5$). Moreover, we observed an increasing number of outbreaks with τ when the pathogen is barely contagious. By contrast, in figure 6a, we show that the average number of outbreaks decreases when τ becomes large ($\tau \geq 5 \cdot 10^{-4}$).

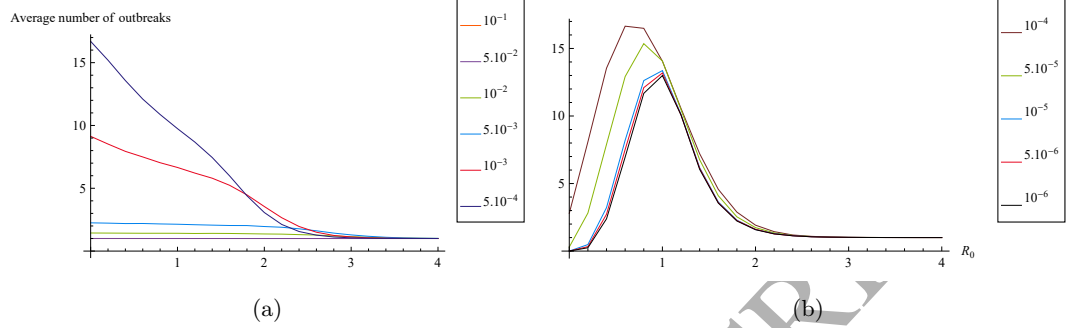


Figure 6: Average number of outbreaks obtained from stochastic simulations. The average number of outbreaks is depicted as a function of the direct transmission $0 < R_0 < 4$ and the spillover transmission (a) $5.10^{-4} \leq \tau \leq 10^{-1}$ and (b) $10^{-6} \leq \tau \leq 10^{-4}$ and for an epidemiological threshold $c = 5$.

3.2.2. Case of a contagious pathogen $R_0 > 1$:

The supercritical case ($R_0 > 1$) is now considered and the spillover transmission rate (τ) is still supposed small.

In this case, two different types of excursions occur in the incidental host: (i) a large outbreak which consumes, with a probability close to one, a large proportion of susceptible individuals and (ii) multiple excursions before and after a large outbreak which each consumes few susceptible individuals. We let $O_{before}(N, \beta, \gamma)$ and $O_{after}(N, \beta, \gamma)$ denote the number of outbreaks occurring respectively before and after the large outbreak. Because $R_0 > 1$, the probability to have one large outbreak is close to one. Hence we make the approximation that one large outbreak occurs during the epidemic, and the total number of outbreaks ($O_{total}(N, \beta, \gamma)$) can be approximated by:

$$O_{total}(N, \beta, \gamma) = O_{before}(N, \beta, \gamma) + 1 + O_{after}(N, \beta, \gamma). \quad (5)$$

To be part of outbreaks occurring before the large one, an excursion has to satisfy two conditions (i) to have a size higher than the epidemiological threshold c , and (ii) to be of a size not too large otherwise it would correspond to the large outbreak. To be more precise, this condition will correspond to the fact that the supercritical branching process used to approximate this excursion does not

go to infinity. As a consequence, $O_{before}(N, \beta, \gamma)$ can be approximated by (See Appendix B):

$$O_{before}(N, \beta, \gamma) = \frac{1}{(\beta N / \gamma)^c - 1}. \quad (6)$$

To approximate the number of outbreaks after the large outbreak ($O_{after}(N, \beta, \gamma)$), we need to know how many susceptible individuals remain in the population.

265 The number of susceptibles consumed before the large outbreak is negligible with respect to the number of susceptibles consumed during the large outbreak. Hence we can consider that the initial state of the large outbreak is N susceptibles, one infected individual and no recovered individual. The number of susceptibles remaining after the large outbreak can be approximated with the
270 deterministic SIR model.

The large outbreak stops when there is no infected individual anymore in the incidental host. Using that $\frac{\dot{S}}{S} = -\beta I = \frac{-\beta}{\gamma} \dot{R}$, we get:

$$(N - \frac{\gamma}{\beta} \log N) - (S - \frac{\gamma}{\beta} \log S) = 0 \quad (7)$$

which has one trivial solution ($S = N$) and a non-trivial solution with no explicit expression denoted $N_{after}(N, \beta, \gamma)$. After the large outbreak, the reproductive ratio for the next excursions, denoted $R_{0_{after}}$, is subcritical ($R_{0_{after}} < 1$) (see Appendix B) and the number of outbreaks after the large one, denoted O_{after} ,
275 can be approximated using Equations (2) to (4).

The branching process approximations of the mean number of outbreaks in the supercritical regime, depicted in Figure 5, are close to the mean number of outbreaks computed by numerical simulations when the recurrent infection from the reservoir is small. The number of outbreaks decreases when the pathogen
280 becomes highly contagious to reach one outbreak when $R_0 > 2.5$. When the infection is introduced in the incidental host by the spillover transmission, the probability to reach the epidemiological threshold depends on the direct transmission between individuals. When the direct transmission increases the infection spreads more efficiently consuming a large number of susceptible individuals

allowing few or no other excursion to reach the epidemiological threshold and producing only one outbreak when $R_0 > 2.5$.

3.3. What is the effect of the reservoir on the number of outbreaks?

We now focus on the effect of the spillover transmission with a pathogen barely contagious ($R_0 < 1$) on the number of outbreaks. We exclude for the sake of simplicity the cases very close to the critical case, that is to say, $1 - R_0$ is not too close to 0. Because we consider the subcritical case ($R_0 < 1$), the excursions are small and at the beginning of the epidemiological dynamics, we make the approximation that the spillover transmission rate is constant equal to τN , and the direct transmission rate is equal to βNI . We thus consider a birth and death process with constant immigration rate τN , individual birth rate βN and individual death rate γ . We are interested in the effect of the parameter τ on the mean number of outbreaks. In particular we aim at estimating the value of τ maximising the mean number of outbreaks, denoted τ_{opt} .

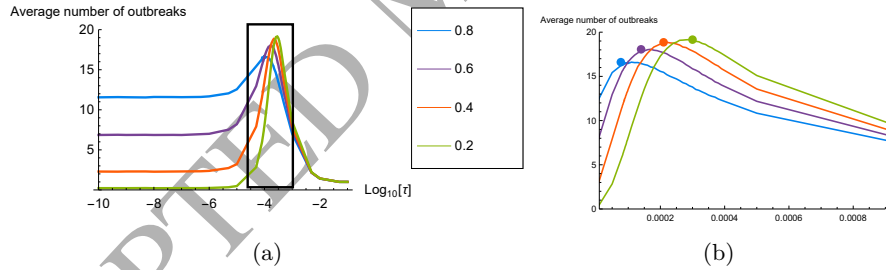


Figure 7: Average number of outbreaks as a function of the spillover transmission. The average number of outbreaks is represented as a function of the basic reproductive ratio ($0.2 < R_0 < 0.8$) and the spillover transmission rate of a pathogen ($\text{Log}_{10}[10^{-10}] \leq \tau \leq \text{Log}_{10}[10^{-1}]$). The rectangle in figure (a) represents the results enlarged in figure (b). In figure (b) the dots represent the estimation of the value of the spillover transmission (τ_{opt}) for each basic reproductive ratio where the number of outbreaks is maximal.

A first quantity which will help giving us an idea of the order of magnitude of the values of τ to be considered is the mean number of infected individuals at large times. This quantity, denoted m_I , equals (see for instance Equation

(8.74) in [23]):

$$m_I = \frac{\tau N}{\gamma - \beta N} = \frac{\tau N/\gamma}{1 - R_0}.$$

In particular, when

$$m_I \gg c \iff \frac{\tau N}{\gamma} \gg c(1 - R_0), \quad (8)$$

the mean number of infected individuals is much larger than c , and when on the contrary

$$m_I \ll c \iff \frac{\tau N}{\gamma} \ll c(1 - R_0), \quad (9)$$

the mean number of infected individuals is negligible with respect to c .

Let us first consider the first case (equation (8)), and choose $\alpha > 1$ such that

$$\frac{\tau N}{\gamma} \geq \alpha(c - 1)(1 - R_0).$$

Then we can show (see Appendix C.1), that the probability p_c that a first infection by the reservoir gives rise to an outbreak (that is to say the number of infected individuals reaches c before 0) is larger than:

$$p_c \geq \frac{(\alpha - 1)(1 - R_0)}{1 + \alpha(1 - R_0)} \xrightarrow{\alpha \rightarrow \infty} 1. \quad (10)$$

Moreover, we can show that if an excursion reaches the level c , it has a probability close to one to lead to a large outbreak consuming a large number of susceptible individuals. Thus only few stuttering chains of transmission will emerge. In Figure 7a, when τ is large ($\tau > 10^{-2}$ thus $\tau N/(c\gamma(1 - R_0)) \geq 25$), only one outbreak is observed because the large number of spillovers prevents the outbreak from dying out.

Let us now consider the second case (equation (9)). Recall that in the case of emerging infectious diseases, the threshold c can be considered as small. Hence

we may consider without loss of generality that (9) implies:

$$\frac{\tau N}{\gamma} \leq \frac{1 - R_0}{2}. \quad (11)$$

In this case, we can prove (see Appendix C.2) that the probability that the number of infected individuals is higher than the threshold c is :

$$\mathbb{P}(I > c) \leq \frac{\tau N}{c\gamma(1 - R_0)} \left(\frac{1 + R_0}{1 - R_0} \right) \left(\frac{1 + R_0}{2} \right)^{1/\ln(2/(1+R_0))-1}.$$

Thus the probability for the number of infected individuals to reach the epidemiological threshold c is small under condition (9). As a consequence, few outbreaks will occur. Indeed, the successive spillovers by the reservoir will produce outbreaks with a small probability, but will nevertheless consume susceptible individuals, until there is no more susceptible in the population. According to Figure 7a, when a small effect of spillover transmission ($\tau < 10^{-6}$) and a small reproductive ratio ($R_0 \leq 0.6$) are considered ($\tau N/(c\gamma(1 - R_0)) < 2.10^{-2}$) then the number of outbreaks is small. In the case of a slightly higher direct transmission rate ($R_0 = 0.8$), each spillover has a non negligible probability to become an outbreak (more precisely 0.12 when $c = 5$) and the number of outbreaks is higher.

We thus predict that the number of outbreaks will tend to be large when the average size of an excursion is close to the epidemiological threshold ($m_I \simeq c$). These observations allow us to give a first rough upper bound of the optimal value τ_{opt} . Indeed, if the mean number of infected individuals ($E[I]$) is equal to c , the ratio $Var(I)/\mathbb{E}^2[I]$ equals $1/(c(1 - R_0))$ (where $Var[I]$ represents the variance of the number of infected individuals, see Appendix C.3) and whose value belongs to $[0.25, 1]$ when $c = 5$ for the values of R_0 considered, which is large. Moreover in this case the distribution of I is skewed to the right (see Figure C.1). This implies that the number of infected individuals will be larger than c a large fraction of the time producing outbreaks which do not go extinct

before a new infection by the reservoir, and thus producing few outbreaks. Hence
 330 we may conclude that τ_{opt} is smaller than the τ leading to a mean number of
 infected individuals c . For instance for the parameters considered in Figure 7a,
 this gives that τ_{opt} is smaller than:

- 10^{-4} when $R_0 = 0.8$,
- 2.10^{-4} when $R_0 = 0.6$,
- 335 • 3.10^{-4} when $R_0 = 0.4$,
- 4.10^{-4} when $R_0 = 0.2$,

Let us now be more precise on the estimation of τ_{opt} . To this aim, we
 will apply two results of the theory of branching processes with immigration.
 The first one, which can be found in [23] (Equation (8.74)) describes the total
 340 infectious lines over the course of the infection, denoted by m :

$$m = (1 - R_0)^{-\frac{\tau}{\beta}} = (1 - R_0)^{-\frac{\tau N}{\gamma R_0}}. \quad (12)$$

Notice that m is necessarily larger than 1 as an infection from the reservoir
 is needed to generate the first infectious line.

The second result is the mean number of infectious lines expected to be
 present at any time, that is to say in the theory of branching processes the
 number of distinct immigrants which have descendants alive at a given moment.
 For large times, this mean number (M_I) is equal to:

$$M_I = \tau N \mathbb{E}[T_0]$$

(see Chapter 8.7 in [24]), where $\mathbb{E}[T_0]$ denotes the mean lifetime of a branching
 process without immigration, with individual birth rate βN , individual death
 rate γ , and initial state 1. This expression can be computed explicitly (see
 Appendix C.4) and equals:

$$\mathbb{E}[T_0] = \frac{1}{\beta N} \log \left(\frac{1}{1 - R_0} \right),$$

thus leading to the expression:

$$M_I = \frac{\tau N}{\gamma R_0} \log \left(\frac{1}{1 - R_0} \right). \quad (13)$$

We will divide an excursion into m/M_I blocks of M_I simultaneous infectious lines (thus without immigration). The idea for such an estimation is the following: it is known that if a Poisson process has k jumps during a time interval, the jumps are uniformly distributed during this time interval. As the infections by the reservoir follow approximately a Poisson process with parameter τN , and we know that in expectation m infections by the reservoir occur before all infected individuals are removed, we divide the epidemic in homogeneous blocks. We choose these blocks to have an initial number of M_I infected individuals to allow the use of results on branching processes with immigration. The initial number of infected individuals in each block is thus M_I , and as a consequence the infection has a probability

$$\frac{R_0^{-M_I} - 1}{R_0^{-c} - 1}$$

to reach the threshold c (see Appendix C.6). Hence the probability for the whole excursion to reach the threshold c can be approximated by

$$1 - \left(1 - \frac{R_0^{-M_I} - 1}{R_0^{-c} - 1} \right)^{m/M_I}.$$

We want this probability to be not too close to 0, otherwise most susceptible individuals would be consumed without giving rise to an outbreak. We also want this probability to be not too close to 1. Indeed, as we have shown in the beginning of this section, this would correspond to a case where $\tau N/\gamma$ is much larger than $c(1 - R_0)$ and once the infected number of individuals has reached the value c it would be very likely to reach a large value and consume a large number of susceptible individuals. As a consequence, we would have at the limit only one large outbreak. We thus choose to equalize this probability

to one half to get an estimation of τ_{opt} . Notice that this choice is arbitrary but has only a small effect on the final results. For instance, a choice of 0.3 or 0.7 would give very close results. The most important is to stay away from 0 and 1. As a conclusion, τ_{opt} is estimated as the unique solution to:

$$\left(1 - \frac{R_0^{-M_I} - 1}{R_0^{-c} - 1}\right)^{\frac{m}{M_I}} = \left(1 - \frac{R_0^{\frac{\tau_{opt} N \log(1-R_0)}{\gamma R_0}} - 1}{R_0^{-c} - 1}\right)^{-\frac{\gamma R_0(1-R_0)^{-\tau_{opt} N/(\gamma R_0)}}{\tau_{opt} N \log(1-R_0)}} = \frac{1}{2}. \quad (14)$$

The unicity of the solution is proved in Appendix C.6.

345 Figure 7b presents the values of τ maximising the number of outbreaks and their estimations (dots) obtained by the branching process approximations. The estimates derived under the branching process approximation give good results, with error ranging from 3 to 33 % regardless the value of the epidemiological threshold (Table 3).

R_0	c	τ_{opt} esti	τ_{opt} simu	error
0.8	5	$7.7 \cdot 10^{-5}$	$1.1 \cdot 10^{-4}$	30
0.8	10	$1.4 \cdot 10^{-4}$	$2.1 \cdot 10^{-4}$	33
0.6	5	$1.4 \cdot 10^{-4}$	$1.7 \cdot 10^{-4}$	18
0.6	10	$2.7 \cdot 10^{-4}$	$2.8 \cdot 10^{-4}$	4
0.4	5	$2.1 \cdot 10^{-4}$	$2.2 \cdot 10^{-4}$	5
0.4	10	$4.3 \cdot 10^{-4}$	$4.1 \cdot 10^{-4}$	5
0.2	5	$3.0 \cdot 10^{-4}$	$2.9 \cdot 10^{-4}$	3
0.2	10	$6 \cdot 10^{-4}$	$4.9 \cdot 10^{-4}$	22

Table 3: Values of the optimal spillover transmission from numerical simulations and estimations. The optimal spillover transmission is calculated for R_0 being equal to 0.2, 0.4, 0.6, and 0.8. We present the values of the optimal spillover for two values of epidemiological threshold $c = 5$ (blue lines) and $c = 10$ (white lines). The errors are ranging from 3 to 33 % with a mean error of 15 %.

350 To get the estimation of τ_{opt} we have made several approximations. First we have considered that the spillover rate by the reservoir is constant equal to τN , whereas it is decreasing and equals to τS , and that the rate of direct trans-

mission due to infected individuals in the population is βNI and not βSI . We believe that these approximations are reasonable because the probability for an excursion to reach the threshold decreases with the consumption of susceptible individuals, and as a consequence, most of the outbreaks will occur at the beginning of the process. However, the real τ_{opt} should be a little bit higher than the one we estimate, to counterbalance the fact that the real infection rates (by the reservoir and the infected individuals) are smaller than the one we use in our calculations. This may explain why in most of the cases we underestimate the real τ_{opt} (see Table 3).

The following approximation we made is the decomposition of the excursions into blocks with an initial number of individuals M_I . In the real process there are no simultaneous infections by the reservoir. However this approximation allows to take into account previous infections by the reservoir whose infectious lines are still present.

When the ratio, $Var[I]/\mathbb{E}^2[I] = \tau N/\gamma$, is small (see Appendix C.3), the value of I stays close to its expectation and few outbreaks occur, as the number of infected individuals rarely reaches 0. For instance, for $R_0 = 0.2$ and $\tau = 4.9 \cdot 10^{-4}$, $\tau N/\gamma \sim 0.2$. As decreasing τ increases this ratio, this could explain why we overestimate τ_{opt} for small values of R_0 (because smaller values of R_0 necessitate higher values of τ to get the same probability to reach the threshold c).

3.4. *What is the effect of the reservoir on the expected size of the largest outbreak?*

During the epidemic, a large outbreak can occur depending on the value of the direct transmission (R_0) and the spillover transmission (τ) and corresponds to the largest number of infected individuals. To analyse the effect of the recurrent emergence of the pathogen on the size of the largest outbreak, we model the largest outbreak by a SIR deterministic model with a spillover transmission:

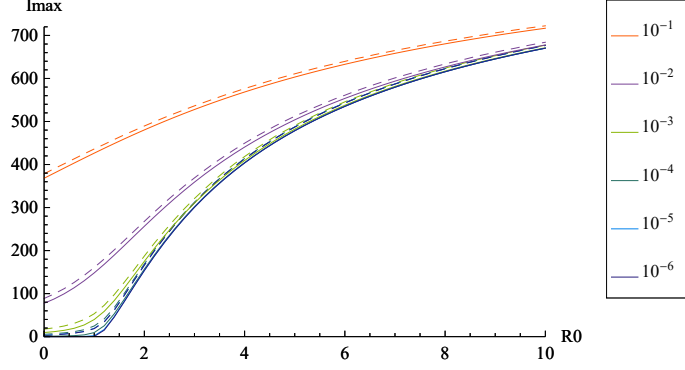


Figure 8: The average maximum number of infected individuals during the largest outbreak. The number of infected individuals during the largest outbreak (I_{max}) is numerically found with the deterministic model (line) and the numerical simulations of the continuous time stochastic model (dashed line). The situation is indicated for a reproductive ratio (R_0) varying from 0 to 10 and with the recurrent emergence rate of the pathogen (τ) varying from 10^{-6} to 10^{-1} .

$$\begin{cases} \dot{S} = -\beta SI - \tau S \\ \dot{I} = \beta SI + \tau S - \gamma I \\ \dot{R} = \gamma I \end{cases} \quad (15)$$

Since no explicit expression of the size of the outbreak can be obtained with the deterministic model, we estimated it using numerical analyses.

Figure 8 shows that the maximal number of infected individuals during the largest outbreak increases with the direct transmission (R_0) and the spillover transmission (τ). When the direct transmission (R_0) is small, the size of the largest outbreak can differ by orders of magnitude with varying spillover transmission (τ). Furthermore, a large outbreak can be observed for a pathogen barely contagious ($R_0 < 1$) when the recurrent emergence of the pathogen is high ($\tau \geq 10^{-3}$).

4. Discussion

Zoonotic pathogens constitute one of the most pressing concerns with regards to future emerging diseases, but studies investigating the importance of the role

of animal reservoirs for the epidemiological dynamics of infectious diseases are lacking. Indeed, most theoretical works only consider pathogen transmission between conspecifics for predicting disease epidemiology. Here, we build a continuous time stochastic SIR model to consider the statistical process underlying a spillover transmission, i.e transmission from an animal reservoir to a host. We analyse the model to predict the number and the size of outbreaks as a function of both the spillover transmission and within host. The model shows that spillover transmission influences the epidemiological dynamics as much as the transmission by direct contact between individuals. Three different dynamics are observed, ranging from the absence of outbreaks to a single large outbreak. The findings have implications for (1) modelling the dynamics of EIDs, (2) understanding the occurrence of outbreaks in the case of pathogens that are barely contagious and (3) control strategies.

In our results, the appearance of outbreaks depends on both the transmission from the reservoir and the direct transmission between individuals. Generally, the occurrence of epidemics in humans is attributed to the ability of the pathogen to propagate between individuals. In the case of a single-host process, the notion of the basic reproductive ratio R_0 seems sufficient to evaluate the spread of the pathogen in a population entirely composed of susceptible individuals. In EIDs, R_0 is also used to gauge the risk of pandemics. In this way, Lloyd-Smith et al. (2009) delineate the three stages identified for a zoonotic pathogen [15] by using the ability of the pathogen to spread between individuals. Each stage corresponds to a specific epidemiological dynamics ranging from a non-contagious pathogen making an outbreak impossible (Stage II, $R_0 = 0$) to a barely contagious pathogen with few outbreaks and stuttering chains of transmission (Stage III, $R_0 < 1$) to a contagious pathogen making a large outbreak possible (Stage IV, $R_0 > 1$). The aim of the Wolfe's classification is to establish each stage at which a zoonotic pathogen may evolve to be adapted to human transmission only, in order to identify pathogens at potential risk of pandemics. However, in our model, by taking into account the recurrent emergence of the pathogen from the reservoir, the three dynamics that define the three

stages will depend on both the spillover transmission and the direct transmission of the pathogen between individuals. The results suggest that in the case of pathogen spilling recurrently over an incidental host, the direct transmission should not be the only parameter to consider.

The presence of a reservoir and its associated recurrent spillovers dramatically impact the epidemiological dynamics of infectious diseases in the incidental host. Without transmission from the reservoir, the probability to have an outbreak when the pathogen is barely contagious only depends on the direct transmission between individuals, and the outbreak rapidly goes extinct. By contrast, the results show that the recurrent emergence of the pathogen from a reservoir increases the probability to observe an outbreak. Spillover transmission enhances the probability to both observe longer chains of transmission and reach the epidemiological threshold (i.e. threshold from which an outbreak is considered) even for a pathogen barely contagious. Moreover, this coupling model (reservoir-human transmission) allows the appearance of multiple outbreaks depending on both the ability of the pathogen to propagate in the population and the transmission from the reservoir. Zoonotic pathogens such as MERS, Ebola or Nipah are poorly transmitted between individuals (R_0 estimated to be less than 1) [25, 10, 26, 6] yet outbreaks of dozens/hundreds/thousands of infected individuals are observed. We argue that, as suggested by our model, the human epidemic caused by EIDs could be due to a recurrent spillover from an animal reservoir.

In the case of zoonotic pathogens, it is of primary importance to distinguish between primary cases (i.e. individuals infected from the reservoir) and secondary cases (i.e. individuals infected from another infected individual) to specify which control strategies to implement and how to optimize public health resources. According to the stochastic SIR model coupled with a reservoir analysed here, the same dynamics can be observed depending on the relative contribution of the transmission from the reservoir and the direct transmission by contact with an infected individual (see Figure 4). For example, a large outbreak is observed either for a high spillover transmission or for a high direct

transmission. The proposed stochastic model makes it possible to understand the effects of the infection from the reservoir or from direct transmission on the epidemiological dynamics in an incidental host when empirically this distinction is difficult. Empirically, the origin of the infection is established by determining the contact patterns of infected individuals during the incubation period. Thereafter, the role of control programs implemented could be evaluated in order to determine better strategies.

We have considered that the reservoir is a unique population in which the pathogen can persist, which is a simplifying assumption. The pathogen is then endemic in the reservoir and the reservoir has a constant force of infection on the incidental host. The reservoir can be seen as an ecological system comprising several species or populations in order to maintain the pathogen indefinitely [12]. For example, bat and dromedary camel (*Camelus Dromedarius*) populations are involved in the persistence of MERS-CoV and in the transmission to human populations [27]. In these cases, the assumptions of a constant force of infection can be valid because the pathogen is endemic. However, zoonotic pathogens can spill over multiple incidental hosts and they can infect each other. In the case of the Ebola virus, which infects multiple incidental hosts such as apes, gorillas and monkeys [28], the principal mode of contamination of the human population is the transmission from non-human primate populations. Moreover, the contact patterns between animals and humans is one of the most important risk factors in the emergence of avian influenza outbreaks [29]. These different epidemiological dynamics with transmission either from the reservoir or from other incidental hosts can largely impact the dynamics of infection observed in the human population, and the investigations of those effects can enhance our understanding of zoonotic pathogens dynamics.

In our model, we make a second simplifying assumption by considering that the infection propagates quickly relatively to other processes such as pathogen evolution and demographic processes. This assumption can be not valid in the case of low emergence of the pathogen from the reservoir. Indeed, the time between two spillovers can be long and makes the evolution of the pathogen

possible inside the reservoir. Moreover, during the time between two spillovers, the demography in the incidental host can vary and impact the propagation of the pathogen. In the case of low spillover transmission in the incidental host, the effect of both pathogen evolution and demographic processes can be a topic for
485 future research on the epidemiological dynamics of emerging infectious diseases.

In this paper, we have argued that the conventional way for modelling the epidemiological dynamics of endemic pathogens in an incidental host should be enhanced to account for spillover transmission in addition to conspecifics transmission. We have shown that our continuous time stochastic SIR model with
490 a reservoir produces similar dynamics to those found empirically (see the classification scheme for pathogens from [15]). This model can be used to better understand the ways in which EIDs transmission routes impact disease dynamics.

Appendices

495 In this appendix, we derive results on the branching process approximation stated in Section 3. The main idea of this approximation is the following: when the epidemiological process is subcritical ($R_0 < 1$), an excursion will modify the state of a small number of individuals with respect to the total population size. During the i -th excursions, the direct transmission rate βSI will stay close
500 to $\beta S_i I$ where S_i denotes the number of susceptibles at the beginning of the i -th excursion. Hence, if we are interested in the infected population, the rate $\beta S_i I$ can be seen as a constant individual birth rate βS_i . Similarly, γI , which is the rate at which an individual in the population recovers, can be interpreted as a constant individual death rate γ in the population of infected individuals.

505 A. Number of outbreaks in the subcritical case ($R_0 < 1$)

In this section, we focus on the number of outbreaks when $R_0 < 1$ and when the rate of introduction of the infection by the reservoir is small ($\tau \ll 1$). That is to say, we consider that each introduction of the infection by the reservoir occurs

after the end of the previous excursion created by the previous introduction of the infection by the reservoir. According to Equation (12), this approximation is valid as long as the ratio τ/β is small. We first approximate the mean number of susceptible individuals consumed by an excursion. Let us consider a subcritical branching process with individual birth rate βN and individual death rate γ . As this process is subcritical, we know that the excursion will die out in a finite time and produce a finite number of individuals. Then from [30] or [31], if we denote by $K[N, \beta, \gamma]$ the total number of individuals born during the lifetime of this branching process (counting the initial individual), we know that:

$$\mathbb{P}(K(N, \beta, \gamma) = k) = \frac{(2k-2)!}{k!(k-1)!} \left(\frac{\beta N}{\gamma + \beta N} \right)^k \left(\frac{\gamma}{\gamma + \beta N} \right)^{k-1},$$

where \mathbb{P} denotes a probability, and hence

$$\mathbb{E}[K(N, \beta, \gamma)] = \frac{\beta N}{\gamma + \beta N} \sum_{k=0}^{\infty} \frac{(2k)!}{(k!)^2} \left(\frac{\gamma \beta N}{(\gamma + \beta N)^2} \right)^k,$$

where \mathbb{E} is the expectation.

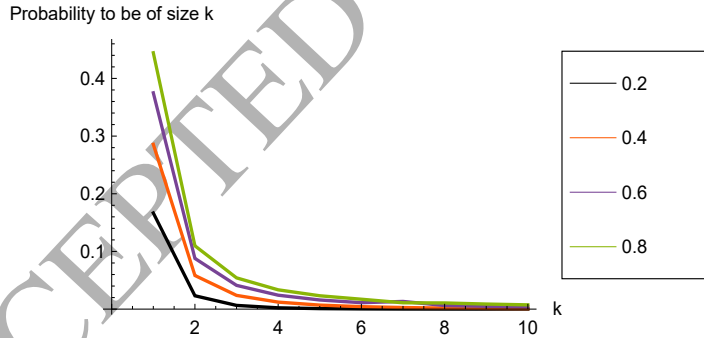


Figure A.1: Probability for an excursion to be of size k . The situation is indicated for a reproductive ratio (R_0) varying from 0.2 to 0.8.

By definition, an excursion is considered as an outbreak only if the maximal number of individuals infected at the same time during this excursion is larger than an epidemiological threshold that we have denoted by c . Hence in order to

approximate the number of outbreaks we still have to compute the probability for an excursion to be an outbreak. This is a classical result in branching process theory which can be found in [32] for instance.

$$P(N, \beta, \gamma) := \mathbb{P}(\text{more than } c \text{ individuals infected at a given time}) = \frac{\gamma/\beta N - 1}{(\gamma/\beta N)^c - 1}. \quad (\text{A.1})$$

With these results in hands, the method to approximate the mean number of outbreaks is the following: the probability that the first excursion is an outbreak is

$$\frac{\gamma/\beta N - 1}{(\gamma/\beta N)^c - 1}.$$

The number of susceptibles at the beginning of the second excursion is approximated by

$$S_2 = N - \mathbb{E}[K(N, \beta, \gamma)].$$

The second excursion has a probability

$$\frac{\gamma/\beta S_2 - 1}{(\gamma/\beta S_2)^c - 1}$$

to be an outbreak. The number of susceptibles at the beginning of the third excursion is approximated by

$$S_3 = S_2 - \mathbb{E}[K(S_2, \beta, \gamma)],$$

and the third excursion has a probability

$$\frac{\gamma/\beta S_3 - 1}{(\gamma/\beta S_3)^c - 1}$$

to be an outbreak. The procedure is iterated as long as there is still a positive number of susceptible individuals. This gives 4.

B. Number of outbreaks in the supercritical case ($R_0 > 1$)

We now focus on the case $R_0 = \beta N / \gamma > 1$. In this case the approximating branching process is supercritical and go to infinity with a positive probability. In the case where the epidemic process describes small excursions, the branching process approximation is still valid, but in the case where it describes a large excursion, then a large fraction of susceptible individuals is consumed and the branching approximation is not valid anymore. However, as all the quantities (susceptible, infected and recovered individuals) are large, a mean field approximation is a good approximation of the process. Here the mean field approximation is the deterministic SIR process, whose dynamics is given by:

$$\begin{cases} \dot{S} = -\beta SI \\ \dot{I} = \beta SI - \gamma I \\ \dot{R} = \gamma I. \end{cases} \quad (\text{B.1})$$

510 Let us first focus on the small excursions occurring before the large one.

As they are small, they can be approximated by a branching process. Here, unlike in the previous section, the approximating branching process Z is supercritical, as $\beta N > \gamma$. We compute its probability to drift to infinity:

$$p_G := \mathbb{P}(Z_\infty = \infty) = \frac{\beta N - \gamma}{\beta N}.$$

As we will see, a supercritical branching process with individual birth rate βN and individual death rate γ conditioned to go extinct has the same law as a subcritical branching process with individual birth rate γ and individual death rate βN . Indeed, if we denote by Z_n the successive values of this branching

process, we get for every couple of natural numbers (n, k) :

$$\begin{aligned} \mathbb{P}(Z_{n+1} = k+1 | Z_n = k, Z_\infty = 0) &= \frac{\mathbb{P}(Z_{n+1} = k+1, Z_n = k, Z_\infty = 0)}{\mathbb{P}(Z_n = k, Z_\infty = 0)} \\ &= \frac{\mathbb{P}(Z_\infty = 0 | Z_{n+1} = k+1, Z_n = k) \mathbb{P}(Z_{n+1} = k+1 | Z_n = k)}{\mathbb{P}(Z_\infty = 0 | Z_n = k)} \\ &= \frac{\mathbb{P}(Z_\infty = 0 | Z_{n+1} = k+1)}{\mathbb{P}(Z_\infty = 0 | Z_n = k)} \frac{\beta N}{\beta N + \gamma} = \left(\frac{\gamma}{\beta N}\right)^{k+1} \left(\frac{\gamma}{\beta N}\right)^{-k} \frac{\beta N}{\beta N + \gamma} = \frac{\gamma}{\beta N + \gamma}. \end{aligned}$$

where $\mathbb{P}(A|B)$ denotes the probability of the event A when B is realised. We used again in this series of equalities classical results on branching processes that can be found in [32]. As a consequence, if we denote by $G[N, \beta, \gamma]$ the number of susceptible individuals consumed by the excursion of a supercritical branching process with individual birth rate βN and individual death rate γ conditioned to go extinct, we get:

$$\mathbb{E}[G(N, \beta, \gamma)] = \frac{\gamma}{\gamma + \beta N} \sum_{k=0}^{\infty} \frac{(2k)!}{(k!)^2} \left(\frac{\gamma \beta N}{(\gamma + \beta N)^2}\right)^k,$$

and the probability for this excursion to have a size bigger than the epidemiological threshold c is

$$\frac{\beta N / \gamma - 1}{(\beta N / \gamma)^c - 1}.$$

As the number of susceptible individuals stays large until the large excursion occurs, we may keep N as the initial number of susceptibles at the beginning of the excursions instead of replacing it by their mean value, as we have done in the previous section.

The different quantities we have just computed allow us to approximate the number of small excursions before the large excursion: in expectation, we have

$$\sum_{k=1}^{\infty} (k-1) p_G (1 - p_G)^{k-1} = \frac{1 - p_G}{p_G} = \frac{\gamma}{\beta N - \gamma}$$

small excursions, which consume

$$\frac{\gamma^2}{(\beta N)^2 - \gamma^2} \sum_{k=0}^{\infty} \frac{(2k)!}{(k!)^2} \left(\frac{\gamma \beta N}{(\gamma + \beta N)^2} \right)^k$$

susceptibles and produce

$$\frac{\gamma}{\beta N - \gamma} \frac{\beta N / \gamma - 1}{(\beta N / \gamma)^c - 1} = \frac{1}{(\beta N / \gamma)^c - 1}$$

515 outbreaks.

Now we focus on the large excursion. We use Equation (B.1) to approximate its dynamics. This equation is well known, and it is easy to obtain the equation satisfied by the final number of susceptible individuals: from (B.1)

$$\frac{\dot{S}}{S} = -\beta I = -\frac{\beta}{\gamma} \dot{R}.$$

Hence

$$\ln(S(t)) - \ln(S(0)) = -\frac{\beta}{\gamma} (R(t) - R(0)) = -\frac{\beta}{\gamma} (N - I(t) - S(t)).$$

In particular, if we are interested in the time T_f when there is no more infected individual and we suppose that at time 0 there is only one infected individual we get

$$\ln(S(T_f)) - \ln(S(0)) = \frac{\beta}{\gamma} (S(T_f) - S(0)).$$

That is to say, $S(T_f)$ and $S(0)$ are related by the equation

$$S(T_f) - \frac{\gamma}{\beta} \ln(S(T_f)) = S(0) - \frac{\gamma}{\beta} \ln(S(0)).$$

Rigorously, the value of $S(0)$ depends on the number of susceptible individuals consumed by the small excursions before the large excursion, but we have seen that this number is small compared to the population size, N . Hence the number of susceptible individuals remaining after the large excursion can be

approximated by the smallest solution of:

$$S(T_f) - \frac{\gamma}{\beta} \ln(S(T_f)) = N - \frac{\gamma}{\beta} \ln(N). \quad (\text{B.2})$$

(as the largest solution is $S(T_f) = N$).

Notice that it is easy to have an idea of the error caused by a small variation of the initial state. Indeed, if we denote by S_f the smallest solution of (B.2) and by $S_f - l(k)$ the solution when $S(0) = N - k$ for a k small with respect to N , we get:

$$(S_f - \frac{\gamma}{\beta} \ln(S_f)) - (S_f - l(k) - \frac{\gamma}{\beta} \ln(S_f - l(k))) = (N - \frac{\gamma}{\beta} \ln(N)) - (N - k - \frac{\gamma}{\beta} \ln(N - k)),$$

or in other terms

$$l(k) + \frac{\gamma}{\beta} \ln(1 - l(k)/S_f) = k + \frac{\gamma}{\beta} \ln(1 - k/N).$$

As k and $l(k)$ are small with respect to N , this can be approximated by

$$l(k) \sim k \left(1 - \frac{\gamma}{\beta N}\right) / \left(1 - \frac{\gamma}{\beta S_f}\right).$$

Finally, notice that in (B.1), S is a decreasing quantity, and I is a non-negative quantity, which varies continuously. Hence $\dot{I} = I(\beta S - \gamma)$ has to be negative before I hits 0. As a consequence,

$$\frac{\beta S_f}{\gamma} < 1.$$

This ensures that the epidemic is subcritical after the large outbreak.

C. Effect of the reservoir on the number of outbreaks

In this section, we focus on the effect of the reservoir transmission rate (parameter τ) on the number of outbreaks when the infection is subcritical ($R_0 < 1$). The idea is the following: first, as excursions of subcritical branching processes are small, we can make the approximation that, at the beginning, the

infection rate by the reservoir is constant equal to τN , and that the direct transmission rate is equal to βNI . Making this approximation allows us to handle
 525 the two processes of infection (by contact and by the reservoir) independently, and to use known results on branching processes with immigration.

C.1. Large τ

Recall that I denotes the number of infected individuals. We first assume (8) and prove inequality (10). Let us choose $\alpha > 1$ such that

$$\frac{\tau N}{\gamma} \geq \alpha(c-1)(1-R_0).$$

Then for any $1 \leq k \leq c-1$, the jump rates of the process I are:

$$\tau(k \rightarrow k+1) = \tau N + \beta N k = \gamma \left(\frac{\tau N}{\gamma} + R_0 k \right) \geq \gamma k((1-R_0)\alpha + R_0)$$

$$\tau(k \rightarrow k-1) = \gamma k.$$

Hence

$$\frac{\tau(k \rightarrow k+1)}{\tau(k \rightarrow k-1)} \geq 1 + (\alpha-1)(1-R_0).$$

This implies that once one individual is infected, the probability for the number of simultaneously infected individuals to reach c before the recovery of all infected individuals is larger than the probability that a birth and death process with initial state 1 and birth (b) and death (d) rates satisfying

$$\frac{b}{d} = 1 + (\alpha-1)(1-R_0)$$

reaches the state c . Applying (A.1), we deduce that this probability p_c
 satisfies:

$$\begin{aligned} p_c &= \frac{1 - (1 + (1-R_0)\alpha + R_0)^{-1}}{1 - (1 + (1-R_0)\alpha + R_0)^{-c}} \geq 1 - (1 + (1-R_0)\alpha + R_0)^{-1} \\ &= \frac{(\alpha-1)(1-R_0)}{1 + (\alpha-1)(1-R_0)}. \end{aligned}$$

This proves (10).

Moreover, as α has been taken large, the infectious process stays supercritical (in the sense that the next event is more likely to be an infection than a recovery) until a size k satisfying

$$\gamma k = \tau N + \beta N k \geq \gamma(\alpha(c-1)(1-R_0) + R_0),$$

and thus if the number of infected individuals reaches c it is likely to reach a large value and consume a large number of susceptible individuals.

C.2. Small τ

As we approximate the infection process by a branching process with constant immigration, the law of I under this approximation converges to a well known law, provided by Equation (8.75) in [23]:

$$\mathbb{E}[x^I] = \left(\frac{1-R_0x}{1-R_0} \right)^{-\frac{\tau N}{\gamma R_0}}, \quad 0 \leq x \leq 1. \quad (\text{C.1})$$

From this law, we can deduce the probability for I to be equal to any integer k :

$$\mathbb{P}(I = k) = (1-R_0)^{\frac{\tau N}{\gamma R_0}} \frac{1}{k!} \prod_{i=0}^{k-1} \left(\frac{\tau N}{\gamma} + i R_0 \right),$$

and thus, the probability for I to be larger than c is

$$\mathbb{P}(I > c) = (1-R_0)^{\frac{\tau N}{\gamma R_0}} \sum_{k=c+1}^{\infty} \frac{1}{k!} \prod_{i=0}^{k-1} \left(\frac{\tau N}{\gamma} + i R_0 \right).$$

Recall that we assumed that

$$\frac{\tau N}{\gamma} \leq \frac{1-R_0}{2}.$$

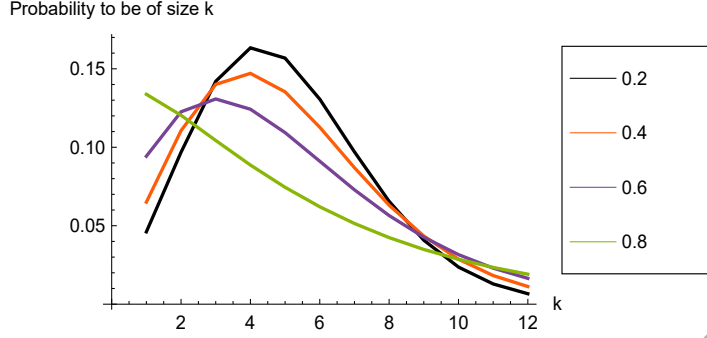


Figure C.1: Distribution of the number of infected individuals in a branching process with immigration for large times. The legend corresponds to the values of R_0 considered.

We thus get the following inequality

$$\mathbb{P}(I > c) \leq \frac{\tau N}{\gamma(1-R_0)} \sum_{k=c+1}^{\infty} \frac{1}{k!} \prod_{i=1}^{k-1} \left(\frac{1-R_0}{2} + iR_0 \right).$$

Now let us notice that for any $i \geq 1$:

$$\frac{1-R_0}{2} + (i-1)R_0 < i \frac{1+R_0}{2}.$$

We thus get

$$\begin{aligned} \mathbb{P}(I > c) &\leq \frac{\tau N}{\gamma(1-R_0)} \sum_{k=c+1}^{\infty} \prod_{i=2}^k \frac{(1-R_0)/2 + (i-1)R_0}{i} \\ &\leq \frac{\tau N}{\gamma(1-R_0)} \sum_{k=c+1}^{\infty} \left(\frac{1+R_0}{2} \right)^{k-1} \\ &= \frac{\tau N}{\gamma(1-R_0)} \left(\frac{1+R_0}{2} \right)^{c-1} \frac{2}{1-R_0} \\ &\leq \frac{\tau N}{c\gamma(1-R_0)} \left(\frac{1+R_0}{1-R_0} \right) \left(\frac{1+R_0}{2} \right)^{1/\ln(2/(1+R_0))-1}, \end{aligned}$$

where to get the last inequality we computed the maximum of the function $x \mapsto xa^{x-1}$ for $a = (1+R_0)/2$. As a conclusion, for a fixed $R_0 < 1$ and a small enough τ , the probability for the number of infected individuals at a given time

535 to be higher than c is bounded by a function of R_0 time $\tau N/(c\gamma(1-R_0))$. This probability is thus small when the last term is small.

C.3. Variance of the process

Recall Equation (C.1). It allows us to compute the variance of I , as follows:

$$\mathbb{E}[I] = \partial_x \mathbb{E}[x^I]_{|x=1} = \frac{\tau N}{\gamma R_0} \frac{R_0}{1-R_0} = \frac{\tau N}{\gamma(1-R_0)},$$

and

$$\mathbb{E}[I(I-1)] = \partial_{xx} \mathbb{E}[x^I]_{|x=1} = \frac{\tau N}{\gamma R_0} \left(\frac{\tau N}{\gamma R_0} + 1 \right) \left(\frac{R_0}{1-R_0} \right)^2 = \frac{\tau N}{\gamma(1-R_0)^2} \left(\frac{\tau N}{\gamma} + R_0 \right).$$

Thus

$$\begin{aligned} \text{Var}[I] &= \mathbb{E}[I(I-1)] + \mathbb{E}[I] - \mathbb{E}^2[I] \\ &= \frac{\tau N}{\gamma(1-R_0)^2} \left(\frac{\tau N}{\gamma} + R_0 + 1 - R_0 - \frac{\tau N}{\gamma} \right) = \frac{\tau N}{\gamma(1-R_0)^2}. \end{aligned}$$

In particular,

$$\frac{\text{Var}[I]}{\mathbb{E}^2[I]} = \frac{\tau N}{\gamma}.$$

C.4. Duration of the excursion of a subcritical branching process

In this section we provide an expression for the term $\mathbb{E}[T_0]$ which appears in the definition of M_I (see Equation (13)). This expression derives from the following equality, which can be found in [32]. Let $t \geq 0$ and T_0 denotes the extinction time of the excursion of a branching process, with one individual at time 0. Then we have

$$\mathbb{P}(T_0 > t) = \frac{\gamma - \beta N}{\gamma e^{(\gamma - \beta N)t} - \beta N}.$$

This allows one to compute the expectation of T_0 as follows:

$$\mathbb{E}_1[T_0] = \int_0^\infty t \mathbb{P}(t < T_0 \leq t + dt) = - \int_0^\infty t \partial_t \mathbb{P}(T_0 > t) dt = \int_0^\infty \mathbb{P}(T_0 > t) dt,$$

where we made an integration by parts. Thus

$$\begin{aligned}\mathbb{E}_1[T_0] &= \int_0^\infty \frac{\gamma - \beta N}{\gamma e^{(\gamma - \beta N)t} - \beta N} \\ &= \lim_{t \rightarrow \infty} \frac{1}{\beta N} \log \left(\frac{\gamma - \beta N e^{-(\gamma - \beta N)t}}{\gamma - \beta N} \right) = \frac{1}{\beta N} \log \left(\frac{\gamma}{\gamma - \beta N} \right).\end{aligned}$$

C.5. Approximating blocks

Recall that if we consider a branching process with individual birth rate βN , individual death rate γ , and initial state $k \leq c$, the probability for this process to reach the size c is

$$\frac{R_0^{-k} - 1}{R_0^{-c} - 1}$$

(see [32] for instance). We use this result to approximate the probability for a block of the excursion to reach the threshold c by

$$\frac{R_0^{-M_I} - 1}{R_0^{-c} - 1},$$

540 where we recall that M_I is the mean number of simultaneous excursions generated by different infections from the reservoir. Notice that this is an approximation, as M_I is not necessarily an integer.

C.6. Unicity of the solution

We end this appendix with the proof of the unicity of the solution to (14). To simplify the notations, we introduce the function F , which at x associates:

$$F(x) := \left(1 - \frac{R_0^{-x} - 1}{R_0^{-c} - 1} \right)^{e^x/x}.$$

τ_{opt} is thus a solution to

$$F\left(\frac{\tau N}{\gamma R_0} \log \left(\frac{1}{1 - R_0} \right)\right) = \frac{1}{2}. \quad (\text{C.2})$$

First we notice that F is only defined for $x \leq c$. Otherwise the term in brackets would be negative. Second, notice that if $x \leq 1$ and for any $c \geq 2$ and $R_0 < 1$,

$$\frac{R_0^{-x} - 1}{R_0^{-c} - 1} \leq \frac{R_0^{-1} - 1}{R_0^{-c} - 1} \leq \frac{R_0^{-1} - 1}{R_0^{-2} - 1} = \frac{1}{1 + R_0^{-1}} < \frac{1}{2}.$$

This shows that if $x \leq 1$, $F(x) > 1/2$. We now determine the sign of $F'(x)$ for x belonging to the interval $[1, c]$. A direct computation gives:

$$F'(x) = F(x) \left(\frac{e^x R_0^{-x} \log R_0}{(R_0^{-c} - 1)x \left(1 - \frac{R_0^{-x} - 1}{R_0^{-c} - 1}\right)} + \frac{e^x}{x} \left(1 - \frac{1}{x}\right) \log \left(1 - \frac{R_0^{-x} - 1}{R_0^{-c} - 1}\right) \right).$$

As the two logarithms are negative for x belonging to $(1, c)$, we deduce that $F'(x) < 0$ for x belonging to $(1, c)$. As $F(1) > 1/2$ and $F(c) = 0$, we conclude that $F(x) = 1/2$ has a unique solution on the interval $(1, c)$. This is equivalent to the fact that (C.2) has a unique solution τ_{opt} which belongs to

$$\left(\frac{\gamma R_0}{N} \log^{-1} \left(\frac{1}{1 - R_0} \right), c \frac{\gamma R_0}{N} \log^{-1} \left(\frac{1}{1 - R_0} \right) \right).$$

This ends the proof.

References

References

- [1] K. E. Jones, N. G. Patel, M. A. Levy, A. Storeyard, D. Balk, J. L. Gittleman, P. Daszak, Global trends in emerging infectious diseases, *Nature* 451 (2008) 990–994. doi:10.1038/nature06536.
- [2] L. H. Taylor, S. M. Latham, M. E. J. Woolhouse, Risk factors for human disease emergence, *Philosophical Transactions of the Royal Society of London* 356 (2001) 983–989. doi:10.1098/rstb.2001.0888.
- [3] D. M. Morens, A. S. Fauci, Emerging infectious diseases: Threats to human

health and global stability, *PloS Pathogens* 9 (7) (2013) e1003467. doi:
 10.1371/journal.ppat.1003467.

- [4] K. A. Murray, P. Daszak, Human ecology in pathogenic landscapes: two hypotheses on how land use change drives viral emergence, *Current Opinion in Virology* 3 (2013) 79–83. doi:10.1016/j.coviro.2013.01.006.
- [5] F. Keesing, L. K. Belden, P. Daszak, A. Dobson, D. Harvell, R. D. Holt, P. Hudson, A. Jolles, K. E. Jones, C. E. Mitchell, S. S. Myers, B. T. R. S. Ostfeld, Impacts of biodiversity on the emergence and transmission of infectious diseases, *Nature* 468 (647) (2010) 647–652. doi:10.1038/nature09575.
- [6] S. P. Luby, M. J. Hossain, E. S. Gurley, A. Be-N, S. Banu, S. U. Khan, N. Homaira, P. A. Rota, P. E. Rollin, J. A. Comer, E. Kenah, T. G. Ksiazek, M. Rahman, Recurrent zoonotic transmission of nipah virus into humans, bangladesh, 20012007, *Emerging Infectious Diseases* 15 (8) (2009) 1229–1235. doi:10.3201/eid1508.081237.
- [7] M.-A. De La Vega, D. Stein, G. P. Kobinger, Ebolavirus evolution: Past and present, *Plos Pathogens* 11 (11) (2015) 1–10. doi:10.1371/journal.ppat.1005221.
- [8] Z. Jezek, M. Y. Szczeniowski, J. J. Muyembe Tamfum, J. B. McCormick, D. L. Heymann, Ebola between outbreaks: Intensified ebola hemorrhagic fever surveillance in the democratic republic of the congo, 1981 1985, *The Journal of Infectious Diseases* 179 (1999) 60–64. doi:10.1086/514295.
- [9] J. Zhou, D. Wang, R. Gao, B. Zhao, J. Song, X. Qi, Y. Zhang, Y. Shi, L. Yang, W. Zhu, T. Bai, K. Qin, Y. Lan, S. Zou, J. Guo, J. Dong, L. Dong, Y. Zhang, H. Wei, W. Li, J. Lu, L. Liu, X. Zhao, X. Li, W. Huang, Biological features of novel avian influenza a (h7n9) virus, *Nature* 499 (2013) 500–503. doi:10.1038/nature12379.

- [10] G. Chowell, S. Blumberg, L. Simonsen, M. A. Miller, C. Viboud, Synthesizing data and models for the spread of mers-cov, 2013: Key role of index cases and hospital transmission, *Epidemics* 9 (2014) 40–51. doi: 10.1016/j.epidem.2014.09.011.
- 585 [11] Center for disease control and prevention.
URL <https://www.cdc.gov/hantavirus/resources/glossary.html>
- [12] D. T. Haydon, S. Cleaveland, L. H. Taylor, Laurenson, Identifying reservoirs of infection: A conceptual and practical challenge, *Emerging Infectious Diseases* 8 (12) (2002) 1468–1473.
- 590 [13] R. W. Ashford, What it takes to be a reservoir, *The Belgian Journal of Zoology* 127 (1997) 85–90.
- [14] R. W. Ashford, When is a reservoir not a reservoir?, *Emerging Infectious Diseases* 9 (11) (2003) 1495–1496. doi:10.3201/eid0911.030088.
- 595 [15] N. D. Wolfe, C. P. Dunavan, J. Diamond, Origins of major human infectious diseases, *Nature* 447 (2007) 279–283. doi:10.1038/nature05775.
- [16] J. O. Lloyd-Smith, D. George, K. M. Pepin, V. E. Pitzer, J. R. C. Pulliam, A. P. Dobson, P. J. Hudson, B. T. Grenfell, Epidemic dynamics at the human-animal interface, *Science* 326 (2009) 1362–1367. doi: 10.1126/science.1177345.
- 600 [17] S. Singh, D. J. Schneider, C. R. Myers, The structure of infectious disease outbreaks across the animal-human interface, arXiv:1307.4628 (2013) 1–18.
- [18] A. Fenton, A. B. Pedersen, Community epidemiology framework for classifying disease threats, *Emerging Infectious Diseases* 11 (12) (2005) 1815–1821. doi:10.3201/eid1112.050306.
- 605 [19] G. T. Nieddu, L. Billings, J. H. Kaufman, R. Forgoston, S. Bianco, Extinction pathways and outbreak vulnerability in a stochastic ebola model,

- Journal of the Royal Society Interface 14 (2017) 20160847. doi:10.1098/rsif.2016.0847.
- [20] S. Singh, C. R. Myers, Outbreak statistics and scaling laws for externally driven epidemics, *Physical Review* 89 (2014) 042108. doi:10.1103/PhysRevE.89.042108.
- [21] E. L. Tay, K. Grant, M. Kirk, A. Mounts, H. Kelly, Exploring a proposed who method to determine thresholds for seasonal influenza surveillance, *PloS One* 8 (2013) 1–10. doi:10.1371/journal.pone.0077244.
- [22] W. O. Kermack, A. G. McKendrick, A contribution to the mathematical theory of epidemics, *Proceedings of the Royal Society of London* 115 (772) (1927) 700–721. doi:10.1098/rspa.1927.0118.
- [23] N. T. Bailey, *The elements of stochastic processes with applications to the natural sciences*, Vol. 25, John Wiley & Sons, 1990.
- [24] E. Pardoux, *Processus de markov et applications: algorithmes, réseaux, génome et finance: cours et exercices corrigés*.
- [25] A. Zumla, D. S. Hui, S. Perlman, Middle east respiratory syndrome, *Lancet* 386 (2015) 995–1007. doi:10.1016/S0140-6736(15)60454-8.
- [26] C. L. Althaus, Estimating the reproduction number of ebola virus (ebov) during the 2014 outbreak in west africa, *PLOS Currents Outbreaks* 6 (2014) 1–5. doi:10.1371/currents.outbreaks.91afb5e0f279e7f29e7056095255b288.
- [27] J. S. M. Sabir, T. T. Y. Lam, M. M. M. Ahmed, L. Li, Y. Shen, S. E. M. Abo-Aba, M. I. Qureshi, M. Abu-Zeid, Y. Zhang, M. A. Khiyami, N. S. Alharbi, N. H. Hajrah, M. J. Sabir, M. H. Z. Mutwakil, S. A. Kabli, F. A. S. Alsulaimany, A. Y. Obaid, B. Zhou, D. K. Smith, E. C. Holmes, H. Zhu, Y. Guan, Co-circulation of three camel coronavirus species and recombination of mers-covs in saudi arabia, *Science* 351 (6268) (2016) 81–84. doi:10.1126/science.aac8608.

- 635 [28] H. Ghazanfar, F. Orooj, M. A. Abdullah, A. Ghazanfar, Ebola, the killer virus, *Infectious Diseases of Poverty* 4 (15) (2015) 1–6. doi:10.1186/s40249-015-0048-y.
- [29] A. Meyer, T. X. Dinh, T. V. Nhu, L. T. Pham, S. Newman, T. T. T. Nguyen, D. U. Pfeiffer, T. Vergne, Movement and contact patterns of long-
640 distance free-grazing ducks and avian influenza persistence in vietnam, *PloS One* (2017) 1–16doi:10.1371/journal.pone.0178241.
- [30] T. Britton, E. Pardoux, Stochastic epidemics in a homogeneous community, in preparation.
- [31] R. Van Der Hofstad, Random graphs and complex networks (2016).
- 645 [32] K. B. Athreya, P. E. Ney, Branching processes, Springer, Berlin, Heidelberg, 1972.

Acknowledgments

The authors have been supported by the “Chair Modélisation Mathématique et Biodiversité” of Veolia Environnement-Ecole Polytechnique-Museum National
650 d’Histoire Naturelle-Fondation X, France.