



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Sub- or supercritical transmissibilities in a finite disease outbreak: Symmetry in outbreak properties of a disease conditioned on extinction

David Waxman^a, Pierre Nouvellet^{b,*}

^a Centre for Computational Systems Biology ISTBI, Fudan University, 220 Handan Road, Shanghai 200433, People's Republic of China

^b School of Life Sciences, University of Sussex, Brighton BN1 9QG, UK



ARTICLE INFO

Article history:

Received 5 February 2018

Revised 24 January 2019

Accepted 30 January 2019

Available online 31 January 2019

Keywords:

Branching process

Stochastic extinction

Disease outbreak

Emerging diseases

ABSTRACT

This work is concerned with the transmissibility of a disease, on observation of an outbreak of limited size. When such an outbreak occurs, an accurate estimate of the transmissibility of the responsible pathogen is essential for an appropriate response to future outbreaks. Transmissibility is usually characterised in terms of the reproduction number, R , which is the mean number of new cases of infection produced by a single infectious individual. A subcritical reproduction number ($R < 1$) guarantees that an outbreak will eventually die out of its own accord. By contrast, a supercritical reproduction number ($R > 1$) does *not* guarantee spread of the disease, since even with appreciable transmissibility, an outbreak may become extinct due to stochastic effects associated with a small number of infected individuals. Once the number of infectious individuals is conditioned on extinction, we show that an exact symmetry of the underlying theory ensures two distinct values of R , one larger than unity, the other smaller than unity, for which all outbreak properties are identical, with no signature of difference. Therefore a disease with a subcritical R , or its supercritical counterpart, when conditioned on extinction, have, for a given outbreak, identical individual likelihoods. In the full likelihood, this symmetry is lost, since the individual likelihood for a subcritical R is weighted by an extinction probability of unity, but the individual likelihood of a supercritical R is weighted by a sub-unity extinction probability. However, the inference can still benefit from the underlying symmetry, since it yields a mapping of all supercritical reproduction numbers onto the subcritical domain ($R < 1$), thereby speeding up evaluation of the likelihood profile. The symmetry holds in the standard situation, where the distribution of secondary cases is Poisson, as well as where this distribution has a negative binomial form and super-spreading can occur.

© 2019 The Authors. Published by Elsevier Ltd.

This is an open access article under the CC BY license. (<http://creativecommons.org/licenses/by/4.0/>)

1. Introduction

Over the past two decades, outbreaks of SARS (Riley et al., 2003), pandemic flu (Cauchemez et al., 2013), MERS-CoV (Cauchemez et al., 2016), Ebola (WHO Ebola Response Team, 2014) and more recently Zika (Ferguson et al., 2016), have illustrated the importance of early characterisation of the transmissibility of emerging or re-emerging pathogens. Typically, characterisation is achieved by estimating either the basic reproduction number R_0 , which applies in a large population that is fully susceptible, or the instantaneous reproduction number R_t , which applies at time t after start of an outbreak, and which accounts for past exposure

and any interventions which have mitigated the spread of the pathogen. Such estimates of transmissibility have been used to assess the level of intervention required to control an outbreak. For instance, if $R_0 > 1$ and vaccination is possible, then treating a proportion of the population that exceeds $1-1/R_0$ will bring an outbreak under control (Fine et al., 2011). Furthermore, the reproduction number (either R_t or R_0) has multiple uses including: (i) forecasting future incidence of a pathogen (Nouvellet et al., 2018), (ii) monitoring the impact of newly implemented control strategies (International Ebola Response Team, 2016), and (iii) predicting current and future attack rates (Ferguson et al., 2016).

Future preparedness, however, ultimately relies on the early assessment of future threats, ideally before a full blown outbreak or epidemic occurs (Karesh et al., 2012). This requires characterising the transmissibility of pathogens that are responsible for short lived outbreaks which have become extinct without any preven-

* Corresponding author at: School of Life Sciences, University of Sussex, Brighton BN1 9QG, UK.

E-mail address: Pierre.nouvellet@sussex.ac.uk (P. Nouvellet).

tative interventions (Farrington et al., 2003; Farrington and Grant, 1999). Intuitively, the limited nature of such an outbreak could reflect a low transmissibility of the pathogen or it could reflect stochastic behaviour of the early transmission process.

Here, we provide properties and a characterisation of outbreaks that have become extinct in the absence of any intervention. A key parameter of such outbreaks is the *reproduction number* R , namely the mean number of new infections produced by a single infectious individual (Cori et al., 2013; Wallinga and Teunis, 2004). We shall assume that R remains constant throughout an outbreak that goes extinct.

Due to individual or environmental variation, there is variation in the number of new infections produced by different infected individuals. The number of new infections (secondary cases) thus follows a probability distribution which is usually referred to as the *offspring distribution* (Farrington et al., 2003; Farrington and Grant, 1999). By definition, the mean of this distribution is the reproduction number R .

2. Choice of offspring distribution

Typically, the offspring distribution is modelled by either a Poisson distribution, a geometric distribution, or negative binomial distribution (Garske and Rhodes, 2008; Lloyd-Smith, 2007; Lloyd-Smith et al., 2005). In this work we shall adopt a negative binomial form for the offspring distribution. This is a somewhat flexible choice since this distribution depends on two parameters and has two standard distributions as limiting cases (see below).

We write the negative binomial distribution as $\text{NB}(\delta, p)$ where δ and p are parameters of the distribution, and lie in the ranges $\delta > 0$ and $0 < p < 1$. In terms of these parameters, the form of the negative binomial distribution we adopt is such that the mean and variance are given by $p\delta/(1-p)$ and $p\delta/(1-p)^2$, respectively.

In what follows, we shall take $p = R/(R + \delta)$, in which case the mean and the variance of the offspring distribution are R and $R + R^2/\delta$, respectively. The mean, R , is thus the reproduction number and the quantity δ is conventionally referred to, in the literature, as the *dispersion*. Generally, large values of δ correspond to low levels of heterogeneity (i.e., a low variance) in realised values of the infectiousness, while small values of δ correspond to high levels of heterogeneity. For more details about heterogeneity, see the work of Garske and Rhodes (2008), Lloyd-Smith et al. (2005). In the particular case of no over-dispersion ($\delta = \infty$), the negative binomial distribution reduces to a Poisson distribution, while when there is unit dispersion ($\delta = 1$) the negative binomial distribution reduces to a geometric distribution.

3. Modelling the dynamics

In modelling the dynamics of infection we shall, for the sake of simplicity, work with a discrete time model. We define the mean time between a primary case becoming infected and the resulting production of a secondary case as the *generation time*. The model then describes the number of new infections at discrete multiples of the generation time—i.e., in different generations.

We proceed by first defining the disease incidence. This is the random number of infected individuals at a given time. For generation $t = 0, 1, 2, \dots$, we denote the disease incidence by X_t , and model the dynamics of disease incidence as a branching process (Lloyd-Smith et al., 2005), based on an offspring distribution of negative binomial form. We then have

$$X_{t+1} = \sum_{j=1}^{X_t} \xi_j \quad (1)$$

where ξ_j is the number of newly infected individuals (number of ‘offspring’) arising from the j th infected individual. The ξ_j are independent and identically distributed (iid) random variables that are drawn from the negative binomial distribution described above, hence $\xi_j \sim \text{NB}(\delta, \frac{R}{R+\delta})$. This form of offspring distribution leads to the distribution of X_{t+1} , conditional on the value of X_t , also having a negative binomial distribution, which is given by

$$X_{t+1}|X_t \sim \text{NB}\left(X_t \delta, \frac{R}{R+\delta}\right) \quad (2)$$

(see the Supplementary Information). Eq. (2) is a convenient representation since even in an appreciable population, only a single random number needs to be generated to produce all ‘offspring’ of a generation.

3.1. Two possible reproduction numbers

An outbreak that has become extinct can always be described by a ‘small’ reproduction number R_s that lies below unity (thus R_s is subcritical). However, as we show in the Supplementary Information, all dynamical properties of such an outbreak are completely *indistinguishable* from an outbreak characterised by a ‘large’ reproduction number R_l that lies *above* unity (thus R_l is supercritical), but where the incidence of the disease is conditioned so that it eventually achieves extinction. The two reproduction numbers satisfy the inequality $R_l > 1 > R_s > 0$.

Generally, all dynamical results calculated for R_l , when conditioned on eventually achieving extinction, are identical to the results calculated for R_s . Another way of saying this is that given the eventual occurrence of extinction, and given that there were X_t cases in generation t , the probability of observing X_{t+s} cases in generation $t+s$ is the same for outbreaks with either R_s or R_l . This is ultimately a property of Eq. (1) and follows from mathematical results which apply for branching processes (Athreya and Ney, 1972; Waugh, 1958). We shall refer to the identity of results for R_s and R_l as the *symmetry* of the problem.

To the best of our knowledge, this symmetry was first implied by Kendall (1956) in the context of stochastic dynamics of epidemics. The symmetry was then further explored in the context of Markov processes (Athreya and Ney, 1972; Daly, 1979; Waugh, 1958) and more recently in the context of infectious disease outbreaks (Guttorp and Perlman, 2015).

In the Supplementary Information we use elementary approaches to provide intuition about the origin of this symmetry and establish some relevant results. Our demonstration of the symmetry, unlike previous ones does not rely on probability generating functions, but rather is derived directly from the properties of the transition matrix.

In what follows, when we discuss the large reproduction number R_l , we shall often leave implicit the fact that the population has been conditioned upon the occurrence of eventual extinction.

To establish some useful results and relations, we have found it useful to introduce the quantity ε , which is the probability of eventual extinction of a disease outbreak when: (i) the outbreak starts with a single infected individual, and (ii) the reproduction number is *larger than unity*. We note that condition (i) readily generalises: when there are n infected individuals, the probability of eventual extinction is ε^n , (see the Supplementary Information). Furthermore, because of condition (ii) eventual extinction is not inevitable and so $\varepsilon < 1$.

It is convenient to proceed using ε and δ as the independent parameters in the problem, and then determining other quantities in terms of these. We find that for $\varepsilon < 1$, independent of the actual initial size, n , of the outbreak, the two reproduction numbers are

given by

$$R_s = g(\delta, \varepsilon) \quad (3a)$$

and

$$R_l = g\left(\delta, \frac{1}{\varepsilon}\right) \quad (3b)$$

where the function $g(d, x)$ is given by

$$g(d, x) = d \frac{x^{1/d} - 1}{1 - x^{-1}} \quad (3c)$$

(see the Supplementary Information).

3.2. Relationships involving the reproduction numbers and parameters

We note that Eq. (3b) relates the large reproduction number, R_l , to the dispersion, δ , and to the probability of extinction given a single infected individual, ε . Thus given a reproduction number that is greater than unity, and given the value of the dispersion, we can use Eq. (3b) to compute the probability of extinction. This result allows us to illustrate that increasing the dispersion (i.e., decreasing δ) reduces the probability of extinction (Blumberg and Lloyd-Smith, 2013a,b; Lloyd-Smith et al., 2005), see Fig. 2A.

We additionally note that using Eqs. (3a-c) allows us to determine a very simple relation between R_s and R_l , namely

$$R_s = \varepsilon^{(1+\frac{1}{\delta})} R_l. \quad (4)$$

This can be interpreted as saying that the small reproduction number R_s is a discounted version of the large reproduction number R_l , with the discount a power of the extinction probability, ε , associated with R_l .

Additionally, writing $\varepsilon = e^{-\alpha}$, (3b) and (3c) yield the exact result

$$\alpha = \begin{cases} R_l + W(-R_l e^{-R_l}) & \text{for a Poisson distribution of offspring (} W(x) \text{ is Lambert's W function)} \\ \ln(R_l) & \text{for a geometric distribution of offspring} \\ -\ln\left[\left(\frac{1-p}{1-pe^{-\alpha}}\right)^\delta\right] & \text{for a negative binomial distribution of offspring (} p = \frac{R_l}{R_l + \delta} \text{)} \end{cases}$$

(see the Supplementary Information).

Eq. (4) provides a simple way to map reproduction numbers that are above unity (supercritical R values) to dynamically equivalent reproduction numbers of a finite outbreak that all lie below unity (subcritical R values). The result of Eq. (4) also allows us to show how, for a given value of R_l , the corresponding value of R_s is higher as over-dispersion increases (i.e., δ decreases), see Fig. 2B.

4. Statistics

4.1. Outbreak size

Finite outbreaks are often quantified in terms of the *outbreak size* (Farrington et al., 2003) (also called *final size*), which is a random variable we denote by Z and which is defined by $Z = \sum_{t=0}^{\infty} X_t$. This random variable counts the total number of individuals who suffered infection during the entire course of the outbreak. In the Supplementary Information, we provide an analytical treatment of the outbreak size distribution, when the offspring distribution is a negative binomial. This treatment can be used to estimate the reproduction number given observation(s) of outbreak size(s). We also present a numerical method to evaluate the distribution of outbreak sizes associated with outbreaks that eventually become extinct.

We verify to high accuracy, with simulations, that the reproduction numbers R_s and R_l of Eqs. (3a) and (3b) lead to the same distribution of outbreak sizes (see Fig. 1A).

4.2. Outbreak duration

The outbreak duration (Farrington and Grant, 1999) is a random time we denote by T_{ext} , that corresponds to the number of generations where the incidence of the disease is different from zero. This also has identical dynamics for the reproduction numbers R_s and R_l of Eqs. (3a) and (3b), see Fig. 1B. An analytical treatment of the outbreak duration distribution can be found in the Supplementary Information.

4.3. Mean and standard deviation of final size and duration

For a finite outbreak, rather than characterising the complete distribution of the final-size and the complete distribution of the duration (as in Fig. 1), we note that knowledge of the mean and standard deviation of those distributions can be sufficient to characterise broad dynamical patterns (Fine et al., 2011). In the Supplementary Information, we provide a simple formulation that allows precise numerical computation of the mean and variance of both the final size distribution and the duration distribution, see Fig. 2C and D. Again, we show that the means and standard deviations for both of these distributions are identical when the reproduction number is R_s and when it is R_l and eventual extinction is conditioned upon.

4.4. Estimate of reproduction number

From a practical point of view, on observing an extinct outbreak of final size z (with $z = 1, 2, 3 \dots$), the key issue is to estimate the reproduction number characterising the disease. Estimating such reproduction number involves computing the probability

of observing a final outbreak size, given a value of the reproduction number, i.e. the likelihood of the parameter. Conditional on eventual extinction, given the symmetry in outbreak size distributions for either R_l or R_s , the individual likelihoods of observing the final size, given a reproduction number of either R_l or R_s are also identical (Fig. 3A).

However,¹ the full likelihood, i.e. unconditional on extinction, must be weighted by the probability of extinction having occurred. Given this weight is unity for R_s and ε^n for R_l , the symmetry in the likelihood is lost after accounting for extinction, as illustrated in Fig. 3B (see the Supplementary information).

Conveniently, when estimating the reproduction number, the mapping of R_s onto R_l , means that the likelihood can be evaluated only on the R_s domain (i.e., 0–1) and the likelihood of R_l follows by simply weighting the results on R_s by the probability of extinction.

We note that the increased level of randomness/uncertainty associated with a small outbreak leads, perhaps counterintuitively, to a greater chance of inferring a larger reproduction number than when a large outbreak is observed. For example, Fig. 3C shows that there is a larger probability of R_l being greater than 2, when the final outbreak size is small, than when the final outbreak size is large (e.g., $\Pr(R_l > 2|Z = 1) > \Pr(R_l > 2|Z = 10)$).

¹ An equivalent relationship applies for the probabilities of outbreak duration (Fig. S1.2).

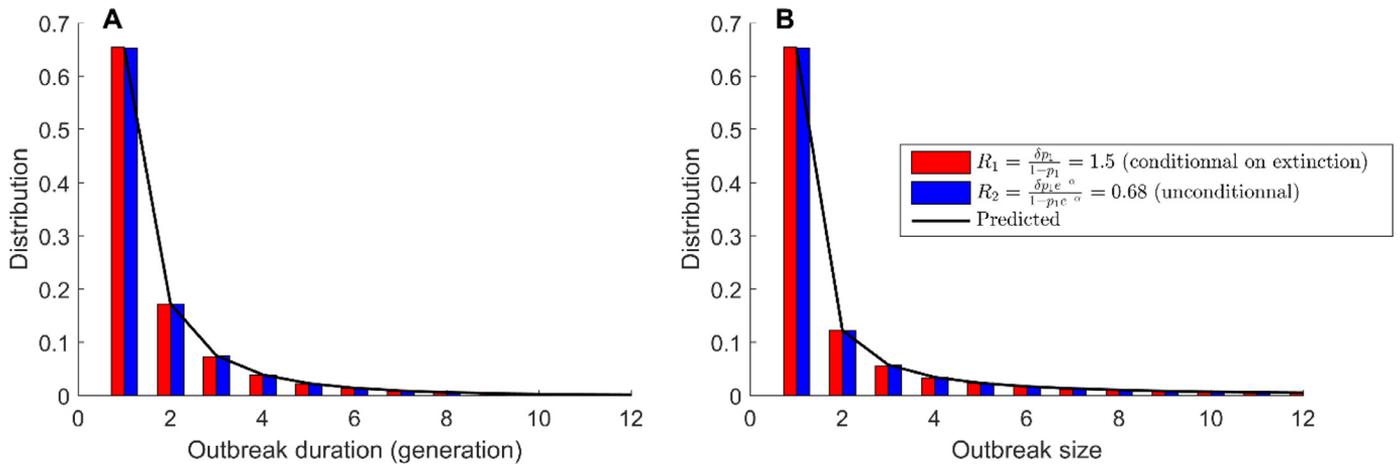


Fig. 1. Distribution of outbreak duration (A) and final-size (B) given extinction ultimately occurs. The distributions were simulated for two reproduction numbers given by Eqs. (3) and (4), where one was greater than unity (red) and one smaller than unity (blue), using a negative binomial offspring distribution with $\delta = 0.5$. The solid lines show the predicted distributions (see Supplementary Information). (An equivalent figure, using a Poisson offspring distribution, can be found in the Supplementary information.)

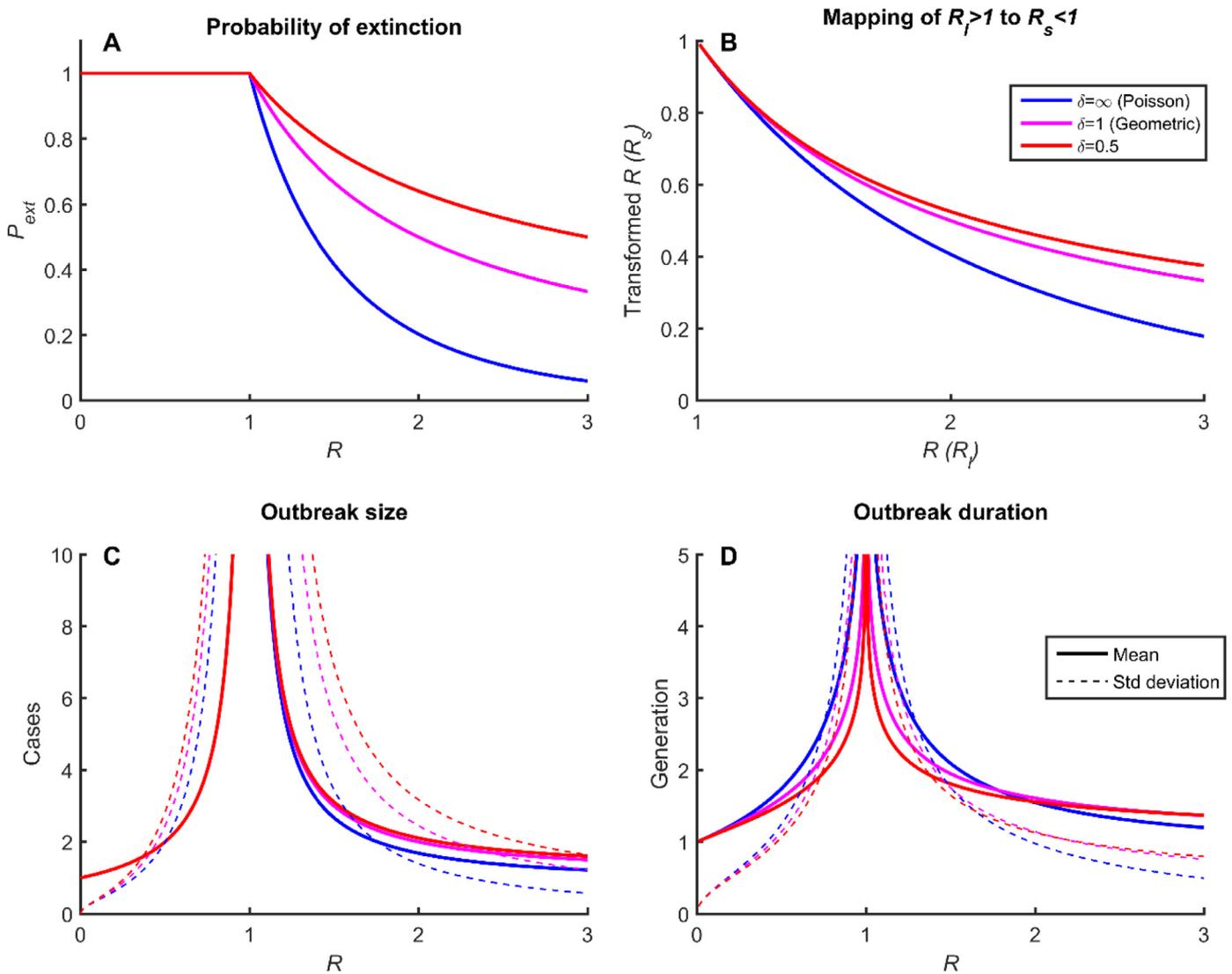


Fig. 2. Characterising extinct outbreaks and their cryptic symmetry. (A) Probability of extinction (P_{ext}) for outbreaks characterised by different reproduction numbers (R), and different offspring distributions. (B) Mapping of R_1 onto R_2 : extinct outbreaks originating from pathogens with mean transmissibility R_1 or R_2 have identical properties. Finally, given that extinction occurs, the symmetry demonstrated in this work allows the mean and standard deviation of outbreak size (C) and duration (D), for R 's greater than unity, to be easily obtained from the analytical formulation of those quantities that applies for R 's below unity.

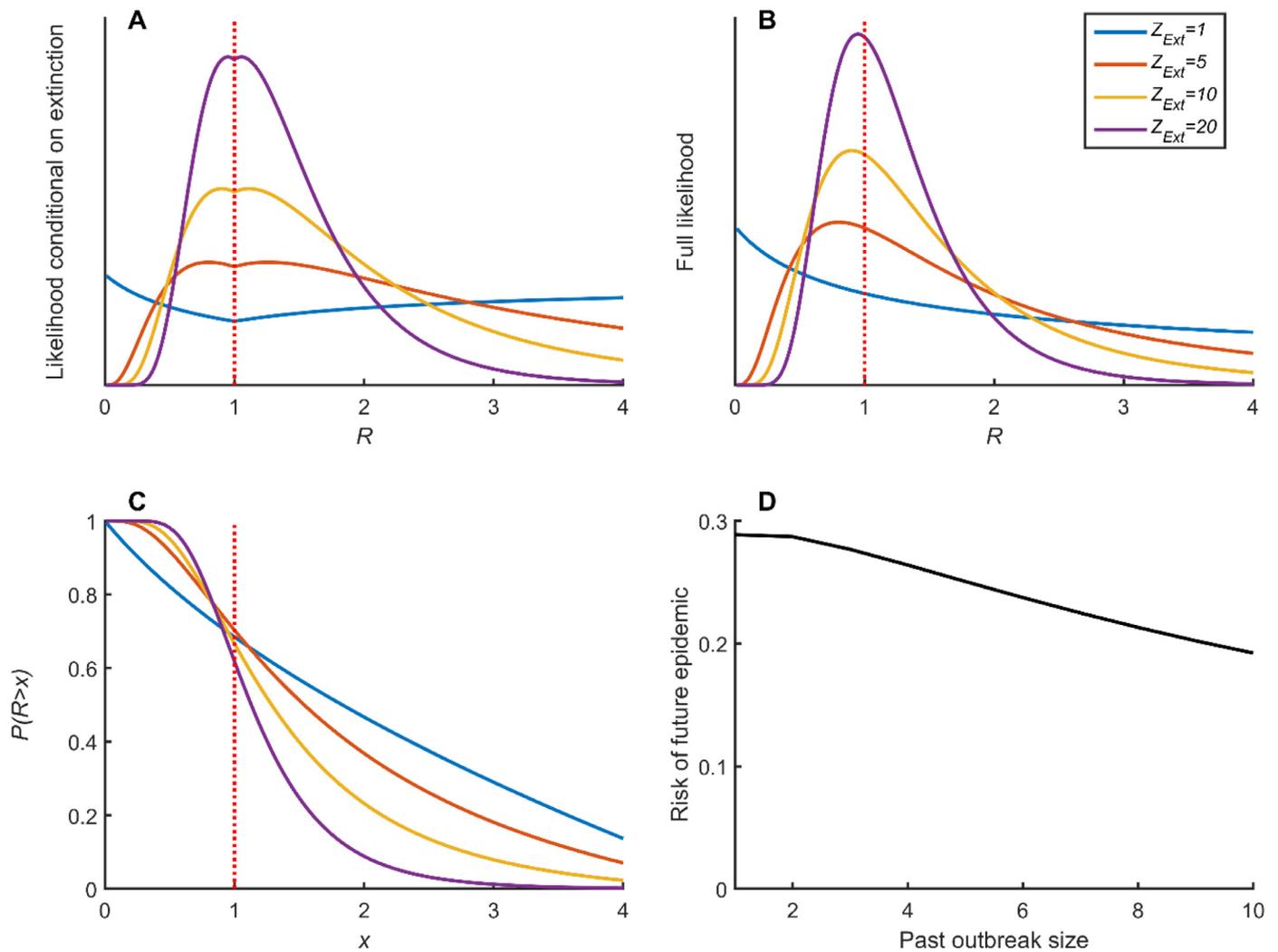


Fig. 3. Inferring the reproduction number and evaluation of future risk from observation of a single outbreak size. (A) Normalised likelihood profile of R , given extinction occurs, and observation of Z cases. The symmetry in the vicinity of $x = 1$ is evident, especially for lower values of Z . (B) The same likelihood profile, without conditioning on extinction. This could be interpreted as a posterior distribution of R given that Z cases were observed. (C) Probability that R is above a threshold, given a past outbreak size, Z . (D) Probability that a future outbreak will not naturally become extinct, given the same pathogen was previously observed to cause an outbreak of size Z : future risk. We assume a negative binomial offspring distribution with $\delta = 0.5$. (Equivalent figures related to time to extinction and using a Poisson offspring distribution can be found in the SI).

Ultimately, given observation of a single outbreak of final size z , one would want to know the risk of a future outbreak of the pathogen running out of control, and leading to an epidemic. The probability of such an epidemic, which we write as $\Pr(\text{Epidemic}|Z=z)$, can be formalised in a branching process as the probability that *extinction will not occur* in a future outbreak, given that a total of z cases were previously observed in an ultimately extinct outbreak. With *FutureExt* the event of extinction given reintroduction of the pathogen, we can write $\Pr(\text{Epidemic}|Z=z) = 1 - \Pr(\text{FutureExt}|Z=z)$ and this is illustrated in Fig. (3D). Interestingly, this result allows us to show that $\Pr(\text{Epidemic}|Z=z)$ decreases as the size of the initial extinct outbreak increases. This illustrates the effect, on prediction, of an increased uncertainty associated with a small outbreak.

5. Discussion

The analysis presented in this work provides a demonstration that the dynamics of a supercritical process conditioned on extinction is in every way identical to the dynamics of a subcritical process. Unlike previous demonstrations (Guttorp and Perlman,

2015; Kendall, 1956; Waugh, 1958), we have relied on the properties of the transition matrix to derive the result. A benefit in this approach is to allow us to derive (i) explicit analytic results for mapping a supercritical reproduction number to its subcritical counterpart, and (ii) explicit analytic results for the expected distributions, means and variances of the final size and duration of the extinct outbreak. The results obtained are applicable when the offspring distribution is characterised by a Poisson, geometric or negative binomial distributions.

While the collapsing of supercritical dynamics conditioned on extinction onto subcritical dynamics has been known (Guttorp and Perlman, 2015; Kendall, 1956; Waugh, 1958), it is clearly underexploited, or has not been appreciated in the literature of infectious disease dynamics. As highlighted by Nishiura et al. (2012), studies of minor outbreaks typically adopt 'the key assumption that a subcritical process (i.e. $R_0 < 1$) resulted in the observation' (Blumberg and Lloyd-Smith, 2013a,b; Cauchemez et al., 2013; Ferguson et al., 2004; Kucharski and Edmunds, 2015). However, we show the likelihood of a supercritical process can also be explored, therefore we see no reason to restrict any such analyses to subcritical process assumptions. We provide a simple way to explore the

full range of reproduction number values (i.e. beyond the subcritical range) using likelihood framework. Such framework is easily applicable and has already been in part implemented in an R package (see <http://github.com/reconhub/branchr>) and used in a recent publication (Cori et al., 2018).

Like many other applications (Blumberg and Lloyd-Smith, 2013a,b; Cauchemez et al., 2013; Ferguson et al., 2004; Kucharski and Edmunds, 2015), our work assumes that the reproduction number remains constant during the duration of the outbreak. While public health and clinical interventions (Cori et al., 2013), as well as behaviour (Funk et al., 2009), can affect the reproduction number during an outbreak, we believe that it is reasonable to assume a constant reproduction number when the outbreaks are of limited durations and sizes. As well as there often being limited available time to intervene, for many diseases (including emerging zoonotic diseases), we have limited knowledge of how to curb transmission, which is further complicated by availability and access to treatment. Additionally, when interventions have been implemented, one could view the estimated reproduction number as an effective reproduction number, see e.g., Blumberg et al. (2015). In practical terms, once again, our conclusions best applies in the context of relatively small outbreaks of short duration, where the conditioning on extinction is realistic. For large outbreaks, it is likely that other factors such local depletion of the susceptible population or the implementation of public health interventions, played a stronger role in the extinction process than stochasticity by reducing the reproduction number.

Our work illustrates the importance of stochastic factors during the early phase of an outbreak (Ferguson et al., 2004; Lloyd-Smith et al., 2005). The symmetrical property of the dynamics emerging for either a small R or a large R conditioned on extinction may appear initially counter-intuitive. However, a simple intuition may come from the following: when R is small, large outbreaks are unlikely to occur, while when R is large, large outbreaks are unlikely to remain finite, therefore after conditioning on extinction, only small outbreaks will be observed. The occurrence of extinction when R is larger than unity must not be under-estimated. For instance, perhaps slightly counterintuitively, when a disease reappears, a full-blown outbreak is more likely to occur following a *small* extinct outbreak rather than following a moderately large one, assuming the same transmissibility applies to the new and extinct outbreaks. However, as the size of an outbreak becomes larger, there is a shift in the impact of stochastic factors toward deterministic behaviour: moderately large (ultimately extinct) outbreaks reflect an R close to unity, but given extinction occurred despite a relatively large outbreak size, stochastic factors have a decreased influence, leading to higher confidence that R lies below 1. Conversely, a small outbreak may reflect a small R , but the large uncertainty means a large R cannot be ruled out.

Overall, we have considered the transmissibility of a disease, on observing a finite outbreak, and believe we have shed light on its nature and provided some useful analytical and numerical methods for its investigation.

Authors' contributions

DW and PN jointly formulated the initial idea, jointly designed and performed all analyses and jointly wrote the manuscript.

Competing interest declaration

We declare no conflict of interest.

Funding

PN acknowledge funding from the [United States Agency for International Development](#) (USAID, AID-OAA-F-16-00115), the [Medical Research Council](#) (MR/R015600/1) and the [National Institute for Health Research](#) (NIHR) Health Protection Research Unit in Modelling Methodology (HPRU-2012-10080). The results of this work does not necessarily reflect the views of USAID.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.jtbi.2019.01.033](https://doi.org/10.1016/j.jtbi.2019.01.033).

References

- Athreya, K.B., Ney, P.E., 1972. *Branching Processes*. Springer Berlin Heidelberg, Berlin, Heidelberg <https://doi.org/10.1007/978-3-642-65371-1>.
- Blumberg, S., Lloyd-Smith, J.O., 2013a. Comparing methods for estimating R_0 from the size distribution of subcritical transmission chains. *Epidemics* 5, 131–145. <https://doi.org/10.1016/j.epidem.2013.05.002>.
- Blumberg, S., Lloyd-Smith, J.O., 2013b. Inference of R_0 and transmission heterogeneity from the size distribution of stuttering chains. *PLoS Comput. Biol.* 9, e1002993. <https://doi.org/10.1371/journal.pcbi.1002993>.
- Blumberg, S., Worden, L., Enanoria, W., Ackley, S., Deiner, M., Liu, F., Gao, D., Lietman, T., Porco, T., 2015. Assessing measles transmission in the United States following a large outbreak in California. *PLoS Curr.* <https://doi.org/10.1371/currents.outbreaks.b497624d7043b1aecfbfd3dfda3e344a>.
- Cauchemez, S., Epperson, S., Biggerstaff, M., Swerdlow, D., Finelli, L., Ferguson, N.M., 2013. Using routine surveillance data to estimate the epidemic potential of emerging zoonoses: application to the emergence of US swine origin influenza A H3N2v virus. *PLoS Med.* <https://doi.org/10.1371/journal.pmed.1001399>.
- Cauchemez, S., Nouvellet, P., Cori, A., Jombart, T., Garske, T., Clapham, H., Moore, S., Mills, H.L., Salje, H., Collins, C., Rodriguez-Barraguer, I., Riley, S., Truelove, S., Al-garni, H., Alhakeem, R., AlHarbi, K., Turkistani, A., Aguas, R.J., Cummings, D.A.T., Van Kerkhove, M.D., Donnelly, C.A., Lessler, J., Fraser, C., Al-Barrak, A., Ferguson, N.M., 2016. Unraveling the drivers of MERS-CoV transmission. *Proc. Natl. Acad. Sci. USA* 113, 9081–9086. <https://doi.org/10.1073/pnas.1519235113>.
- Cori, A., Ferguson, N.M., Fraser, C., Cauchemez, S., 2013. A new framework and software to estimate time-varying reproduction numbers during epidemics. *Am. J. Epidemiol.* 178, 1505–1512. <https://doi.org/10.1093/aje/kwt133>.
- Cori, A., Nouvellet, P., Garske, T., Bourhy, H., Nakoué, E., Jombart, T., 2018 Dec 17. A graph-based evidence synthesis approach to detecting outbreak clusters: an application to dog rabies. *PLoS Comput. Biol.* 14 (12), e1006554. <https://doi.org/10.1371/journal.pcbi.1006554>.
- Daly, F., 1979. Collapsing supercritical branching processes. *J. Appl. Probab* 16, 732–739. <https://doi.org/10.2307/3213140>.
- Farrington, C.P., Grant, A.D., 1999. The distribution of time to extinction in subcritical branching processes: applications to outbreaks of infectious disease. *J. Appl. Probab.* 36, 771–779. <https://doi.org/10.1239/jap/1032374633>.
- Farrington, C.P., Kanaan, M.N., Gay, N.J., 2003. Branching process models for surveillance of infectious diseases controlled by mass vaccination. *Biostatistics* 4, 279–295. <https://doi.org/10.1093/biostatistics/4.2.279>.
- Ferguson, N.M., Cucunubá, Z.M., Dorigatti, I., Nedjati-Gilani, G.L., Donnelly, C.A., Basáñez, M.-G., Nouvellet, P., Lessler, J., 2016. Countering the Zika epidemic in Latin America. *Science* 353, 353–354. <https://doi.org/10.1126/science.aag0219>.
- Ferguson, N.M., Fraser, C., Donnelly, C.A., Ghani, A.C., Anderson, R.M., 2004. Public health risk from the avian H5N1 influenza epidemic. *Science* 304, 968–969. <https://doi.org/10.1126/science.1096898>.
- Fine, P., Eames, K., Heymann, D.L., 2011. "Herd immunity": a Rough Guide. *Clin. Infect. Dis.* 52, 911–916. <https://doi.org/10.1093/cid/cir007>.
- Funk, S., Gilad, E., Watkins, C., Jansen, V.A.A., 2009. The spread of awareness and its impact on epidemic outbreaks. *Proc. Natl. Acad. Sci. USA* 106, 6872–6877. <https://doi.org/10.1073/pnas.0810762106>.
- Garske, T., Rhodes, C.J., 2008. The effect of superspreading on epidemic outbreak size distributions. *J. Theor. Biol.* 253, 228–237. <https://doi.org/10.1016/j.jtbi.2008.02.038>.
- Guttorp, P., Perlman, M.D., 2015. Predicting extinction or explosion in a Galton–Watson branching process with power series offspring distribution. *J. Stat. Plan. Inference* 167, 193–215. <https://doi.org/10.1016/j.jspi.2015.04.009>.
- International Ebola Response Team, 2016. Exposure patterns driving ebola transmission in West Africa: a retrospective observational study. *PLoS Med* 13, e1002170. <https://doi.org/10.1371/journal.pmed.1002170>.
- Karesh, W.B., Dobson, A., Lloyd-Smith, J.O., Lubroth, J., Dixon, M.A., Bennett, M., Aldrich, S., Harrington, T., Formenty, P., Loh, E.H., Machalaba, C.C., Thomas, M.J., Heymann, D.L., 2012. Ecology of zoonoses: natural and unnatural histories. *Lancet* 380, 1936–1945. [https://doi.org/10.1016/S0140-6736\(12\)61678-X](https://doi.org/10.1016/S0140-6736(12)61678-X).
- Kendall, D.G., 1956. *Deterministic and stochastic epidemics in closed populations*. In: *Proc Berkeley Symposium on Math Statist and Probability*, pp. 149–165.
- Kucharski, A.J., Edmunds, W.J., 2015. Characterizing the transmission potential of zoonotic infections from minor outbreaks. *PLoS Comput. Biol.* <https://doi.org/10.1371/journal.pcbi.1004154>.

- Lloyd-Smith, J.O., 2007. Maximum likelihood estimation of the negative binomial dispersion parameter for highly overdispersed data, with applications to infectious diseases. *PLoS One* 2, e180. <https://doi.org/10.1371/journal.pone.0000180>.
- Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E., Getz, W.M., 2005. Superspreading and the effect of individual variation on disease emergence. *Nature* 438, 355–359. <https://doi.org/10.1038/nature04153>.
- Nishiura, H., Yan, P., Sleeman, C.K., Mode, C.J., 2012. Estimating the transmission potential of supercritical processes based on the final size distribution of minor outbreaks. *J. Theor. Biol.* 294, 48–55. <https://doi.org/10.1016/j.jtbi.2011.10.039>.
- Nouvellet, P., Cori, A., Garske, T., Blake, I.M., Dorigatti, I., Hinsley, W., Jombart, T., Mills, H.L., Nedjati-Gilani, G., Van Kerkhove, M.D., Fraser, C., Donnelly, C.A., Ferguson, N.M., Riley, S., 2018. A simple approach to measure transmissibility and forecast incidence. *Epidemics* 22, 29–35. <https://doi.org/10.1016/j.epidem.2017.02.012>.
- Riley, S., Fraser, C., Donnelly, C.A., Ghani, A.C., Abu-Raddad, L.J., Hedley, A.J., Leung, G.M., Ho, L.-M., Lam, T.-H., Thach, T.Q., Chau, P., Chan, K.-P., Lo, S.-V., Leung, P.-Y., Tsang, T., Ho, W., Lee, K.-H., Lau, E.M.C., Ferguson, N.M., Anderson, R.M., 2003. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science* 300, 1961–1966. <https://doi.org/10.1126/science.1086478>.
- Wallinga, J., Teunis, P., 2004. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am. J. Epidemiol.* 160, 509–516. <https://doi.org/10.1093/aje/kwh255>.
- Waugh, W.A.O., 1958. Conditioned Markov Processes. *Biometrika* 45, 241–249. <https://doi.org/10.1093/biomet/45.1-2.241>.
- WHO Ebola Response Team, 2014. Ebola virus disease in West Africa—the first 9 months of the epidemic and forward projections. *N. Engl. J. Med.* 371, 1481–1495. <https://doi.org/10.1056/NEJMoa1411100>.