medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

An exact method for quantifying the reliability of end-of-epidemic declarations in real time

Kris V Parag[†, 1], Christl A Donnelly[1, 2], Rahul Jha[3], and Robin N Thompson[4]

[1]MRC Centre for Global Infectious Disease Analysis, Imperial College London, London, W2 1PG, UK
[2]Department of Statistics, University of Oxford, Oxford, OX1 3LB, UK
[3]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK
[4]Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK
[†]Email: k.parag@imperial.ac.uk

Abstract

We derive and validate a novel and analytic method for estimating the probability that an epidemic has been eliminated (i.e. that no future local cases will emerge) in real time. When this probability crosses 0.95 an outbreak can be declared over with 95% confidence. Our method is easy to compute, only requires knowledge of the incidence curve and the serial interval distribution, and evaluates the statistical lifetime of the outbreak of interest. Using this approach, we rigorously show how the time-varying under-reporting of infected cases will artificially inflate the inferred probability of elimination and hence lead to early (false-positive) end-of-epidemic declarations. Contrastingly, we prove that incorrectly identifying imported cases as local will deceptively decrease this probability, resulting in late (false-negative) declarations. Failing to sustain intensive surveillance during the later phases of an epidemic can therefore substantially mislead policymakers on when it is safe to remove travel bans or relax quarantine and social distancing advisories. World Health Organisation guidelines recommend fixed (though disease-specific) waiting times for end-of-epidemic declarations that cannot accommodate these variations. Consequently, there is an unequivocal need for more active and specialised metrics for reliably identifying the conclusion of an epidemic.

Key-words: epidemic elimination; renewal models; effective reproduction number; epidemic curves; Bayesian statistics; infectious disease.

I. Introduction

The timing of an end-of-epidemic declaration can have significant economic and public health consequences. Early declarations can negate the benefits of prior control measures (e.g. quarantines or lockdown), leaving a population at an elevated risk to the resurgence of the infectious disease. The Ebola virus epidemic in Liberia (2014-2016), for example, featured several declarations that were followed by additional waves of infections [1]. Late declarations, however, can unnecessarily stifle commercial sectors such as agriculture, trade and tourism, leading to notable financial and livelihood losses. One of the first studies advocating the need for improved end-of-epidemic metrics suggested that the MERS-CoV epidemic in South Korea was declared over at least one week later than was necessary [2]. Balancing the health risk of a second wave of infections against the benefits of reopening the economy earlier is a non-trivial problem currently being faced by many countries as the COVID-19 pandemic enters a more controlled phase.

Current World Health Organisation (WHO) guidelines adopt a time-triggered (i.e. decisions are enacted after a fixed, deterministic time) approach to end-of-epidemic declarations, recommending that officials wait for some prescribed period after the last observed infected case before adjudging the outbreak over. The most common waiting time, which applies to Ebola virus and MERS-CoV among others, involves twice the maximum incubation period of the disease [3]. While having a fixed decision time is simple and actionable, it neglects the stochastic variation that is inherently possible at the tail of an outbreak. Recent studies have started to question this time-triggered heuristic and investigate the factors that could limit its practical reliability.

Specifically, [2] made initial advances in this direction and derived mathematical formulae for assessing the end of an epidemic in a data-driven manner. This method uses the time-series of new cases (incidence) across an epidemic together with estimates of its serial interval distribution, which describes the random inter-event times between infections, and the basic reproduction number

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

(the average number of secondary infections per primary infection at the start of an epidemic) to compute the probability that the outbreak is over at any moment. This leads to an epidemiologically informed statistical measure of confidence in an end-of-outbreak declaration.

This approach is important but not perfect. It assumes that infected cases are reported without any error and it depends on parameters that relate to the initial growth phase of the epidemic. Moreover, to maintain simplicity, it adopts a mathematically conservative description of transmission, making its end-of-epidemic declaration time estimates likely to be late [2]. More recent studies [4, 5] have applied forward simulation to explore the tail dynamics of an outbreak. These have revealed the impact of the constant under-reporting of cases [4] and demonstrated the sensitivity of declarations to the effective reproduction number [5], a parameter that remains relevant across all phases of the epidemic. The effect of different routes of transmission on declarations has also been examined in [1] using the framework of [2].

However, there is still much we do not know about the dynamics of an outbreak as it approaches its end. Specifically, analytic and general insight into the sensitivity of end-of-epidemic declarations to practical surveillance imperfections is needed. Real incidence data is corrupted by time-varying trends in under-reporting, delays in case notification and influenced by the interaction of imported and local cases [6, 7, 8]. Previous works have either assumed perfect reporting [2] or treated constant under-reporting within some simulated scenarios [4, 5]. Here we attempt to expose the implications of more realistic types of data corruption, particularly time-varying case under-reporting and importation, by developing an exact framework that provides broad and provable insights.

We build on the renewal or branching process transmission model from [9, 10], to derive and test a novel and exact real-time method for estimating the probability of elimination; defined as the probability that no future local cases will emerge conditioned on the past epidemic incidence. We explain this model in Fig. 1. Using this probability, we define an event-triggered [11, 12] declaration metric that guarantees confidence in that declaration provided the assumptions of the model hold. The trigger is the first time that this probability crosses a threshold e.g. we are 95% confident in our declaration if the threshold is 0.95. Event-triggered decision-making was essentially proposed by [2], has proven effective in other fields [13, 14, 15] and belies the time-triggered WHO approach, which fixes the time (elapsed since the last case) but not the confidence in declaration.

We benchmark our estimate against the true probability of elimination i.e. the probability if the statistics and effective reproduction number of the epidemic were known precisely and show consistency under the perfect conditions in [2] but with the caveat that we estimate effective reproduction numbers from the incidence curve in real time. We find that even the true elimination probabilities strongly depend on the specific stochastic incidence curve observed, confirming that time-triggered decision heuristics are unwarranted. Using our exact framework we prove two key results about imperfect surveillance. First, any type of time-varying under-reporting will lead to early or false-positive event-triggers and hence declarations, unless explicit knowledge of the under-reporting scheme is available. Second, a failure to identify and account for the differences between local and imported cases will result in late or false-negative event-triggers, regardless of the dynamics of case importation.

Many infectious disease epidemics, including the ongoing COVID-19 pandemic, are known to feature extensive time-varying under-reporting and repeated importations from different regions [16, 17]. As this pandemic progresses into the controlled phase in several countries, public health authorities will need to decide when to relax existing intervention measures such as lockdowns, social distancing policies or travel bans. Our work suggests that intensive surveillance, both of cases and their origin, must be sustained to make informed, reliable and adaptive decisions about the threat posed by the virus in the late stages of the outbreak, even if reported case numbers remain at zero for consecutive days. We hope that our method will aid understanding and assessment of the tail kinetics of infectious epidemics.

II. Methods

A. *Infectious disease transmission models*

We can mathematically describe the transmission of an infection within a population over time with a branching or reproductive process based on the fundamental Euler-Lotka equation from ecology and demography [18]. This process models communicable pathogen spread from a primary (infected) case to secondary ones at some time s using two key variables: the effective reproduction number, R_s, and the generation time distribution with probabilities $\{w_u\}$ for all times u. Here R_s defines the number of secondary cases at time $s+1$ one primary case at s infects on average, while w_u is the probability that the time for a primary to secondary transmission is u units [18]. We make the common assumption that the serial interval and generation time distributions are the same, known, and do not vary with time [10].

If I_s counts the newly observed infected cases at s and a Poisson (Poiss) model is used to represent

medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

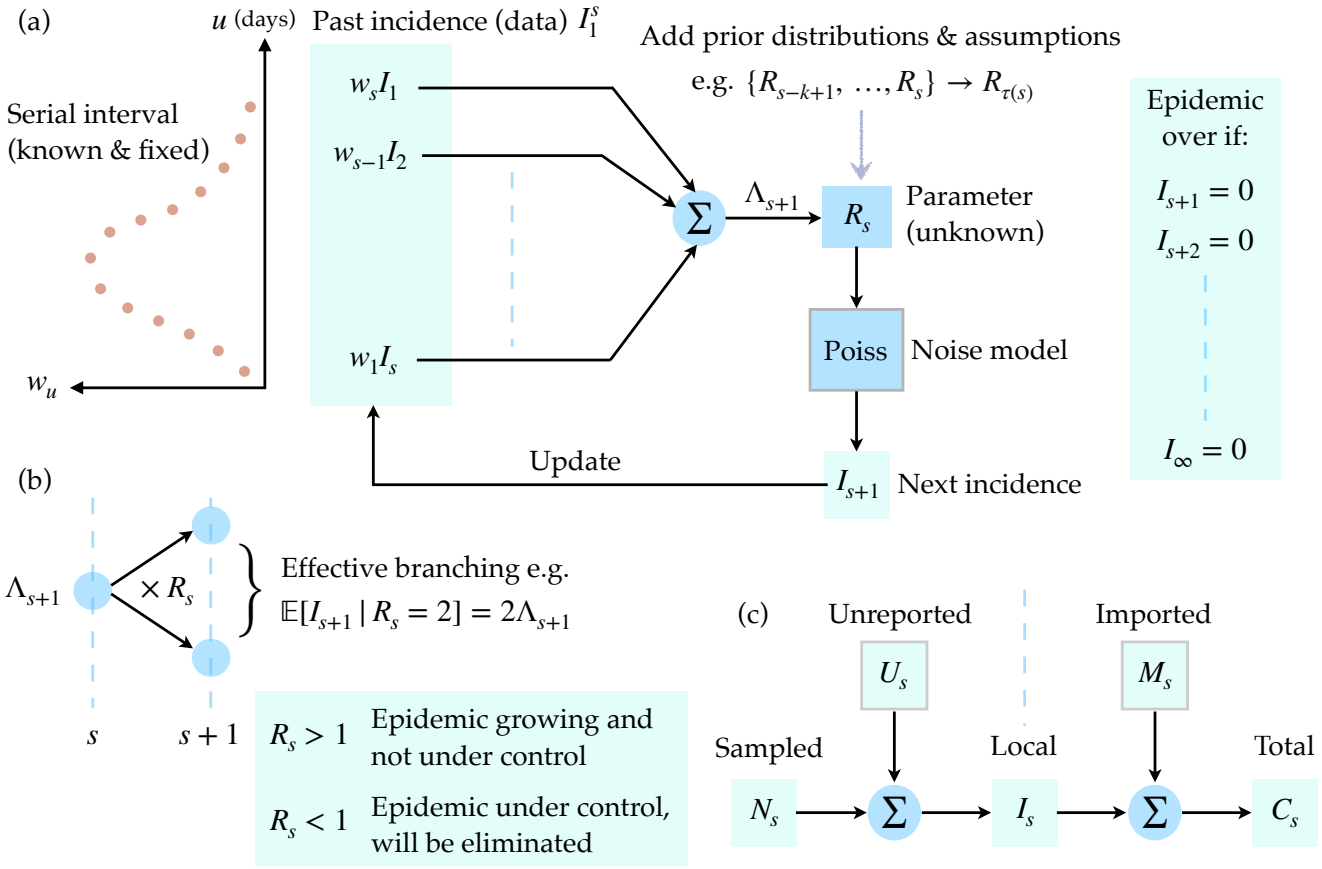

Fig. 1: **Transmission dynamics of an infectious disease.** The branching process or renewal approach to infection propagation is outlined under a Poisson noise model in panel (a). Past, observed infected cases I_1^s, which form an incidence curve, seed new infections with probabilities proportional to w_u defined by the serial interval distribution of the disease. The total infectiousness Λ_{s+1} sums these contributions. The effective reproduction number R_s determines how many effective infections are passed on to the next time unit $s+1$. It is common to group R_s values over a window $\tau(s)$ to improve estimation reliability. When all future incidence values are zero we conclude that the epidemic is over or eliminated. Panel (b) shows how R_s acts as a branching parameter, controlling whether the epidemic grows or dies out. This parameter is therefore be essential to predicting the dynamics of an epidemic. Panel (c) provides a breakdown of more realistic observation assumptions, where we might not be able to directly measure the local and complete incidence I_s due to unreported U_s or imported (migrating) M_s cases. If we can only observe sampled cases, N_s, or the total number of cases, C_s, then our epidemic predictions will be biased.

the noise in these observations then the renewal model captures the branching dynamics of infectious disease transmission with $I_s \sim \text{Poiss}(R_{s-1}\Lambda_s)$ [19]. Here $\Lambda_s := \sum_{u=1}^{s-1} I_{s-u} w_u$ is the total infectiousness of the disease up to time $s-1$ and summarises how previous cases contribute to upcoming cases on day s. We use $I_1^s := \{I_1, I_2, \ldots, I_s\}$ to represent the incidence curve from time 1 to s. A schematic of this reproductive approach to epidemic transmission is given in Fig. 1. Usually we are interested in estimating the R_s numbers in real time from the progressing I_1^s [10, 20, 21].

This effective reproduction number is important for forecasting the kinetics of the epidemic. If $R_s > 1$ then we can expect the number of infections to increase monotonically with time. However, if $R_s < 1$ is sustained then we can be confident that the epidemic is being controlled and will, eventually, be eliminated [22]. In order to enhance the reliability of these estimates we usually assume that the epidemic transmission properties are stable over a look-back window of size k defined at time s as $\tau(s) := \{s, s-1, \ldots, s-k+1\}$ [10, 23]. We let the reproduction number over this window be $R_{\tau(s)}$ and apply a conjugate gamma (Gam) prior distribution assumption: $R_{\tau(s)} \sim \text{Gam}(a, 1/c)$ with a and c as shape-

medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

scale hyperparameters. This formulation, together with the use of gamma prior distributions, is standard in current renewal model frameworks [10, 20].

The posterior distribution of $R_{\tau(s)}$ given the relevant window of the past incidence curve of data i.e. $I_{\tau(s)} := I^s_{s-k+1}$ is also gamma distributed as [21]

$$R_{\tau(s)} \,|\, I_{\tau(s)} \sim \text{Gam}\left(a + i_{\tau(s)},\, \tfrac{1}{c+\lambda_{\tau(s)}}\right), \qquad (1)$$

with grouped sums $i_{\tau(s)} := \sum_{u\in\tau(s)} I_u$ and $\lambda_{\tau(s)} := \sum_{u\in\tau(s)} \Lambda_u$. If some variable $y \sim \text{Gam}(\alpha, \beta)$ then $\mathbb{P}(y) = {y^{\alpha-1}e^{-y/\beta}}/{\beta^\alpha\Gamma(\alpha)}$ and $\mathbb{E}[y] = \alpha\beta$. As a result, Eq. (1) yields the posterior mean estimate, $\hat{R}_{\tau(s)} = \alpha_{\tau(s)}\beta_{\tau(s)}$ with $\alpha_{\tau(s)} := a + i_{\tau(s)}$, $\beta_{\tau(s)} := {1}/{c+\lambda_{\tau(s)}}$. Eq. (1) allows us to infer the grouped or averaged effective reproduction number over the window $\tau(s)$.

We can derive the posterior predictive distribution of the next incidence value (at time $s+1$) by marginalising over the domain of $R_{\tau(s)}$ as in [21]. If the space of possible predictions at $s+1$ is $x \,|\, I_{\tau(s)}$ and NB indicates a negative binomial distribution then

$$x \,|\, I_{\tau(s)} \sim \text{NB}\left(\alpha_{\tau(s)},\, p_{\tau(s)} := \tfrac{\Lambda_{s+1}\beta_{\tau(s)}}{1+\Lambda_{s+1}\beta_{\tau(s)}}\right). \qquad (2)$$

Eq. (2) completely describes the uncertainty surrounding one-step-ahead incidence predictions and is causal because all of its terms (including Λ_{s+1}) only depend on the past observed incidence curve I^s_1 [21].

If a random variable $y \sim \text{NB}(\alpha, p)$ then $\mathbb{P}(y) := \binom{\alpha+y-1}{y}(1-p)^\alpha p^y$ and $\mathbb{E}[y] = {p\alpha}/{1-p}$. Hence our posterior mean prediction is $\hat{I}_{s+1} = \mathbb{E}[x \,|\, I_{\tau(s)}] = \Lambda_{s+1}\hat{R}_{\tau(s)}$. The current estimate of $R_{\tau(s)}$ influences our ability to predict upcoming incidence points. Thus, we expect that reliable estimation of the effective reproduction number is necessary for projecting the future behaviour of an infectious disease epidemic. In Results we rigorously extend and apply this insight to derive an exact method for computing the probability that an epidemic is reliably over i.e. that no future infections will occur.

Under-reported and imported cases

The above formulation assumes perfect case reporting and that all cases, I^s_1, are local to the region being monitored. We now relax these assumptions. First, we consider more realistic scenarios where only some fraction of the local cases are reported or observed at any time. We use N_s and U_s for the number of sampled and unreported cases at time s. We consider a general time-varying binomial (Bin) sampling model with $0 \le \rho_s \le 1$ as the probability that a true case is sampled at time s (hence $1-\rho_s$ is the under-reporting probability). Then $N_s \sim \text{Bin}(I_s, \rho_s)$. The smaller ρ_s is, the less representative the sampled curve N^s_1 is of the true I^s_1.

This is a standard model for under-reporting [6, 24] and implies the following statistical relationship

$$I_s = N_s + U_s, \qquad N_s \sim \text{Poiss}(\rho_s R_{s-1}\Lambda_s). \qquad (3)$$

Raikov's theorem [25] states that if the sum of two independent variables is Poisson then each variable is also Poisson. Consequently, U_s is Poisson with mean $(1-\rho_s)R_{s-1}\Lambda_s$. Most studies investigating this model make the simplifying assumption that $\rho_s = \rho$ for all s i.e. that constant under-reporting occurs. The persistence of the Poisson relationship in Eq. (3) means that we can directly apply the forecasting and estimation results of the previous section to N_s. Practically, if we observe only N^s_1 then unless we have independent knowledge of ρ_s (this can often be difficult to ascertain reliably [16, 24]) we can only construct an approximation to $\rho_s\Lambda_s$ as $\tilde{\Lambda}_s := \sum_{u=1}^{s-1} w_u N_{s-u}$ with $\mathbb{E}[\tilde{\Lambda}(s)] = \rho_s\Lambda_s$.

Second, we investigate when imported or migrating cases from other regions, denoted by count M_s at time s, are introduced, resulting in the total number of observed cases being C_s. Within this framework we ignore the under-reporting of cases and assume that I_s is observed to avoid confounding factors. We follow the approach of [7] and describe M_s as a Poisson number with some mean at time s of ϵ_s. Using Raikov's theorem we obtain

$$C_s = I_s + M_s, \qquad C_s \sim \text{Poiss}(R_{s-1}\Lambda_s + \epsilon_s). \qquad (4)$$

Eq. (4) models how imported cases combine with existing local ones to propagate future local infections.

While our work does not require assumptions on ϵ_s, for ease of comparison later on we adopt the convention that the sum of imports and local cases drive the epidemic forward with the same reproduction number and serial interval [26]. Consequently, $I_s \sim \text{Poiss}(R_{s-1}\bar{\Lambda}_s)$ with $\bar{\Lambda}_s := \sum_{u=1}^{s-1} w_u C_{s-u}$. Practically, if surveillance is poor and one assumes that all observed cases are local then the approximate model $C_s \sim \text{Poiss}(R_{s-1}\bar{\Lambda}_s)$ results. The forecasting and estimation results of the previous section therefore apply here as well.

In Results we examine the impact of imperfect (our null hypothesis $\mathcal{H}_0$) and ideal (the alternative $\mathcal{H}_1$) surveillance within the context of under-reporting and importation in turn. Ideal surveillance represents the ability to know either U_s or M_s and hence account for their contributions. Imperfect surveillance refers to only having knowledge of N_s or C_s and basing inferences on these curves under the strong assumption that they approximate the true incidence. This assumption is often made in the literature [2, 10, 20] for the purposes of tractability and means Eq. (1) and Eq. (2) are valid. Fig. 1 summarises the relationships from Eq. (3) and Eq. (4).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

RESULTS

An exact method for declaring an outbreak over

We define an epidemic to be eliminated or over [22] at time s if no future, local or indigenous infected cases are observed i.e. $I_{s+1} = I_{s+2} = \cdots = I_\infty = 0$. We can define the estimated probability of elimination, z_s, as

$$z_s := \mathbb{P}\left(\wedge_{j=s}^{\infty} I_{j+1} = 0 \,|\, I_1^s\right), \tag{5}$$

with I_1^s as the incidence curve (data), observed until time s. We refer to z_s as an estimated probability because we do not have perfect knowledge of the epidemic statistics e.g. we cannot know R_s precisely. The importance of this distinction will become clear in the subsequent section (see Eq. (10)). However, we observe that if we could have this idealised knowledge then Eq. (5) would exactly define the probability of no future cases given I_1^s.

Declaring the end of an epidemic with confidence $\mu\%$ translates into solving the optimal stopping time problem

$$t_\mu = \arg\min_s z_s \geq \tfrac{\mu}{100}, \tag{6}$$

with t_{95}, for example, signifying the first time that we are at least 95% sure that the epidemic has ended. Note that z_s is a function of I_1^s and practically characterises our uncertainty in the outcome of the epidemic (i.e. if it is over or not). This uncertainty derives from the fact that a range of possible epidemics with distinct future incidences I_{s+1}^∞ can possess the same I_1^s and R_1^s values. Some uncertainty exists even if R_1^s is known perfectly.

Eq. (6) presents an event-triggered approach to declaring the end of an epidemic with the μ threshold serving as an informative trigger. Event-triggered formulations have the advantage of being robust to changes in the observed data. For example, various incidence curves, I_1^s, can be observed under the same reproduction number time-series R_1^s. Defining t_μ as in Eq. (6) ensures that we guarantee our confidence in the end-of-epidemic declaration irrespective of the specific trajectory I_1^s takes. While Eq. (6) is written in absolute time, we may also measure it relative to the time of the last observed case, t_0. Our waiting time until declaration is then $t_\mu - t_0$. Time-triggered approaches set some fixed waiting time from t_0 of d so that declarations occur at $d+t_0$. However, since z_{d+t_0} can vary considerably among realisations of incidence curves from the same disease, these approaches can offer no confidence guarantees.

Previous works on end-of-epidemic declarations have either approximated z_s with a simpler, more conservative probability [2] or used simulations to estimate a quantity similar to z_s that is averaged over those simulations [4] [5]. Further, no study has yet included real-time estimates of R_s, within its assessment of epidemic elimination, despite the importance of this parameter in preventing and describing continued transmission [22]. By taking the renewal process approach to epidemic propagation, as shown in Fig. 1, we explicitly embed uncertainty about R_s estimates and obtain an analytic and insightful expression for the probability that the outbreak is over given the observed cases (Eq. (5)).

We derive this by inferring R_s within a sequential Bayesian framework from I_1^s, by using a moving window of length k time units. We denote this estimate $R_{\tau(s)}$ with window $\tau(s)$ spanning I_{s-k+1}^s [10, 21]. Our main result is summarised as a theorem below (see Methods for further details). Fig. 2 illustrates how our computed z_s probability varies across the lifetime of an example incidence curve, thus providing a real-time, causal and dynamically updating view of our confidence in its end.

Theorem 1. If the posterior distribution of the grouped effective reproduction number, $R_{\tau(s)}$, given the incidence curve I_1^s has form $\text{Gam}\left(\alpha_{\tau(s)}, \beta_{\tau(s)}\right)$ then the estimated probability that this epidemic has been eliminated at time s is $z_s = \prod_{j=s}^{\infty}\left(1 + \frac{\hat{I}_{j+1}}{\alpha_{\tau(j)}}\right)^{-\alpha_{\tau(j)}}$ with $\hat{I}_{j+1} = \Lambda_{j+1}\hat{R}_{\tau(j)}$ and $\hat{R}_{\tau(j)} = \alpha_{\tau(j)}\beta_{\tau(j)}$ as the mean posterior incidence prediction and effective reproduction number estimate at time j.

We outline the development of this theorem. First, we decompose Eq. (5) into sequentially predictive terms as:

$$z_s = \mathbb{P}\left(I_{s+1} = 0 \,|\, I_1^s\right) \prod_{j=s+1}^{\infty} \mathbb{P}\left(I_{j+1} = 0 \,|\, I_1^j\right). \tag{7}$$

For simplicity, we rewrite Eq. (7) as $z_s = q_0 \prod_{j=1}^{\infty} q_j$. The factor q_j conditions on I_1^{s+j}, which includes all the epidemic data, I_1^s and the sequence of assumed zeros beyond that i.e. $I_{s+1}^{s+j} = 0$ for $j \geq 1$. This sequence is treated as pseudo-data. Note that q_0 is just a one-step-ahead prediction of 0 from the available incidence curve.

We solve Eq. (7) by making use of known renewal model results derived in [10, 21, 23] and outlined in Methods. The renewal transmission model allows us to estimate the effective reproduction number R_s and hence compute z_s in real time (see Fig. 1). This estimate at time s, $R_{\tau(s)}$, uses the look-back window $\tau(s)$ of k time units (e.g. days). The posterior over $R_{\tau(s)}$ is shape-scale gamma distributed as $\text{Gam}\left(\alpha_{\tau(s)}, \beta_{\tau(s)}\right)$ with $\alpha_{\tau(s)} := a + i_{\tau(s)}$ and $\beta_{\tau(s)} := \frac{1}{c+\lambda_{\tau(s)}}$ (see Eq. (1)). Here (a, c) are hyperparameters of a gamma prior distribution placed on $R_{\tau(s)}$ and $i_{\tau(s)}$ and $\lambda_{\tau(s)}$ are grouped sums of the incidence I_u and total infectiousness Λ_u for $u \in \tau(s)$. The total infectiousness describes the cumulative impact of past cases and is defined in Methods.

Under this formulation, the posterior predictive dis-

medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

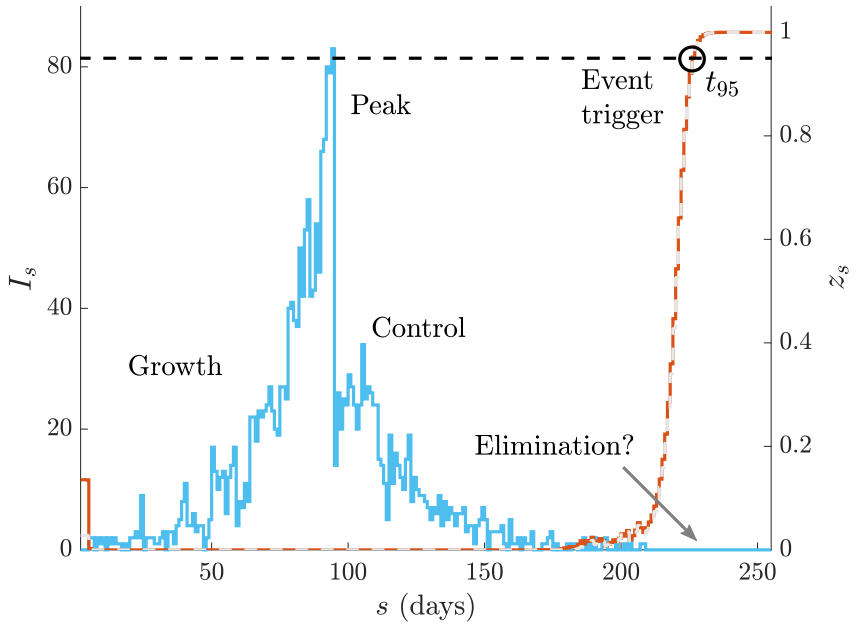

Fig. 2: **Elimination probabilities across the lifetime of an epidemic.** We simulate a single epidemic curve, I_s (blue, case counts on left y-axis), under a gamma distributed serial interval distribution (similar to that used in [27] for Ebola virus) and a true R_s profile that step changes from 2 to 0.5 at $s = 100$ days. We compute the true and estimated elimination probabilities, z_s^* and z_s, conditional on all cases observed up to time s in grey and red respectively (right y-axis). The circle (black) indicates when the outbreak can be declared over with 95% confidence. Observe how z_s and z_s^* respond to the low I_s at the beginning of the epidemic before remaining 0 until we get to the tail of the outbreak. The central question in this study is how few cases need to be observed in the recent past before we can be confident that the epidemic has been eliminated.

tribution of the incidence at $s+1$ is negative binomially distributed (NB) (see Eq. (2)). The probability of I_{s+1} being zero from this distribution gives $q_0 = (1+\Lambda_{s+1}\beta_{\tau(s)})^{-\alpha_{\tau(s)}}$ by substitution. The next term, q_1, can be computed similarly because we condition on $I_{s+1} = 0$ as pseudo-data (i.e. the sequential terms in Eq. (7)) and update Λ_{s+2}, $\beta_{\tau(s+1)}$ and $\alpha_{\tau(s+1)}$ with this zero. Iterating for all terms yields

$$z_s = \prod_{j=s}^{\infty}(1+\Lambda_{j+1}\beta_{\tau(j)})^{-\alpha_{\tau(j)}}, \qquad (8)$$

which is an exact expression for z_s. As zero incidence values accumulate with time $\Lambda_{j+1} \to 0$ and hence $q_j \to 1$. As a result, only a finite number of terms in Eq. (8) need to be computed and the initial ones are the most important for evaluating z_s.

The posterior mean estimate of $R_{\tau(s)}$ is $\hat{R}_{\tau(s)} = \mathbb{E}[R_{\tau(s)} \,|\, I_1^s] = \mathbb{E}[R_{\tau(s)} \,|\, I_{\tau(s)}] = \alpha_{\tau(s)}\beta_{\tau(s)}$ with $I_{\tau(s)}$ as the incidence values in the $\tau(s)$ window (the remaining I_1^{s-k} are assumed uninformative [10]). This follows from the Gam distribution and implies a posterior mean incidence prediction $\hat{I}_{s+1} = \mathbb{E}[I_{s+1} \,|\, I_{\tau(s)}] = \Lambda_{s+1}\hat{R}_{\tau(s)}$ from the NB posterior predictive distribution [21]. Substituting these into Eq. (8) gives:

$$z_s = \prod_{j=s}^{\infty}\left(1+\frac{\hat{I}_{j+1}=\Lambda_{j+1}\hat{R}_{\tau(j)}}{\alpha_{\tau(j)}}\right)^{-\alpha_{\tau(j)}}. \qquad (9)$$

This completes the derivation. Theorem 1, when combined with Eq. (6), provides a new, analytic and event-triggered approach to adjudging when an outbreak has ended. Eq. (9) provides direct and quantifiable insight into what controls the elimination of an epidemic and can be easily computed and updated in real time.

Understanding the probability of elimination

We dissect and verify the implications of Theorem 1, which presents an exact and novel method for estimating the probability that any infectious disease epidemic has been eliminated. Eq. (8) formalises the expectation that any decrease in case incidence increases z_s. This results because $\partial q_j/\partial\alpha_{\tau(j)} > 0$ for all $\alpha_{\tau(j)}$, meaning that q_j is monotonically increasing in $\alpha_{\tau(j)}$ and hence $i_{\tau(j)}$. As z_s is a product of q_j and every q_j is positive then z_s is also monotonically increasing in all incidence window sums. Consequently, any process that reduces incidence surely increases the probability of elimination.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The main variable controlling z_s is the average predicted incidence $\hat{I}_{j+1}$. Reducing either Λ_{j+1} or $\hat{R}_{\tau(j)}$ will increase our confidence in a declaration made after a fixed time (the time-triggered approach) or, decrease the time of declaration for a fixed confidence (the event-triggered approach). Since Λ_{j+1} depends on the serial interval distribution, which is characteristic of the infectious disease of interest, some epidemics will be intrinsically harder to control and hence eliminate [28]. The only factor we can manipulate is $\hat{R}_{\tau(j)}$, which is reduced by initiating interventions e.g. vaccination, social-distancing or quarantine. As a result, sustained control efforts will, as expected, increase z_s [22].

Interestingly, z_s is insensitive to the uncertainty in our estimates or predictions, despite its derivation from the posterior distributions of Eq. (1) and Eq. (2). This is a consequence of the inherent data shortage at the tail of an epidemic (there are necessarily many zero incidence points), which likely precludes the inference of anything more complex than mean statistics [23]. Moreover, when the incidence is small stochastic fluctuations can dominate epidemic dynamics. Consequently, to maximise the reliability of our z_s estimates we recommend using long windows (large k) for $\hat{R}_{\tau(j)}$. Short windows are more sensitive to recent fluctuations and are more prone to yielding uninformative estimates when many zero incidence points occur [23].

Last, we validate the correctness of our estimated z_s by considering a hypothetical setting in which the true reproduction number, $\{R_s : s \geq 0\}$, is known without error. This allows us to derive the true (but generally unknowable) probability of elimination z_s^* at time s, given complete information of the epidemic statistics. Under the renewal model $\mathbb{P}\left(I_{s+1} = 0 \,|\, I_1^s\right) = e^{-R_s \Lambda_{s+1}}$. Repeating this process for future zero infected cases (akin to describing its likelihood) gives:

$$z_s^* = \prod_{j=s}^{\infty} e^{-\Lambda_{j+1} R_j} = e^{-\sum_{j=s}^{\infty} \Lambda_{j+1} R_j}. \qquad (10)$$

Observe that z_s^* depends on the serial interval distribution and the level of implemented control, which modulates R_s. These are the two main factors underlying the transmission of the infectious disease.

The true declaration time with confidence $\mu\%$ is then $t_\mu^* = \arg\min_s z_s^* \geq \frac{\mu}{100}$ (see Eq. (6)). We can verify our approach to end-of-epidemic declarations if we can prove that t_μ sensibly converges to t_μ^*. At the limit of $\alpha_{\tau(j)} \to i_{\tau(j)} \to \infty$, the estimated $\hat{R}_{\tau(j)}$ tends to the true R_j because under those conditions the posterior mean estimate coincides with the grouped maximum likelihood estimate of R_j, which is unbiased. Applying this limit to q_j in Eq. (9) we find that as $\hat{R}_{\tau(j)} \to R_j$:

$$\lim_{i_{\tau(j)} \to \infty} \left(1 + \frac{\Lambda_{j+1} \hat{R}_{\tau(j)}}{i_{\tau(j)}}\right)^{-i_{\tau(j)}} = e^{-\Lambda_{j+1} R_j}, \qquad (11)$$

implying that $z_s \to z_s^*$ and hence $t_\mu \to t_\mu^*$.

This asymptotic consistency suggests that z_s and t_μ indeed approximate the true but unknowable probability of elimination z_s^* and declaration time t_μ^*. Other end-of-epidemic metrics in the literature have not shown such theoretical justification. We illustrate z_s and z_s^* across a simulated and representative incidence curve in Fig. 2. There we find a close correspondence between these probabilities and observe a clear sensitivity to changes in incidence at the beginning and end of this outbreak. Note that z_s and z_s^* (and hence declaration times derived from them) are deterministic functions of I_1^s and are more precisely written: $z_s \,|\, I_1^s$ and $z_s^* \,|\, I_1^s$.

Given this dependence, it is often more meaningful to characterise the relative declaration time of the epidemic $\Delta t_\mu = t_\mu - t_0$ with t_0 as the time of the last observed case. This allows us to sensibly compare z_s values from various realisations of I_1^s and to compute confidence intervals on Δt_μ from either simulated or empirical data. In both cases we first generate M conditionally independent $I_1^s[u]$ trajectories (e.g. by bootstrapping over the original I_1^s time-series). Here u counts the 1 to M trajectories. Every $z_s \,|\, I_1^s[u]$ provides a sample of Δt_μ. We can then obtain confidence intervals by inverting frequentist probabilities e.g. $[a, b]$ forms a 95% interval if $0.95 = \mathbb{P}(a \leq \Delta t_\mu \leq b) = \frac{1}{M} \sum_{u=1}^{M} 1(\Delta t_\mu[u] < b \wedge \Delta t_\mu[u] > a)$ with $1(.)$ as an indicator function.

Practical comparisons and verification

We have only validated our approach at an asymptotic limit that is not realistic for elimination i.e. the proof that z_s and t_μ converge to their true counterparts requires infinite incidence. While this proof suggests our formulation is mathematically correct, it does not indicate its performance on actual elimination problems. We now verify out method more practically. We first use simulated data to show that Δt_μ and Δt_μ^* correspond well over several end-of-epidemic problems, where we are far from this limit. These simulations also demonstrate why time-triggered approaches can be misleading; depending on the specific instance of I_1^s observed a fixed time can lead to either early or late declarations. We then provide a direct comparison with the approach of [2] on empirical data. We find that our method performs well even when tested on bootstrapped incidence curves resulting from fitting the empirical data to the model of [2], which assumes different transmission dynamics.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

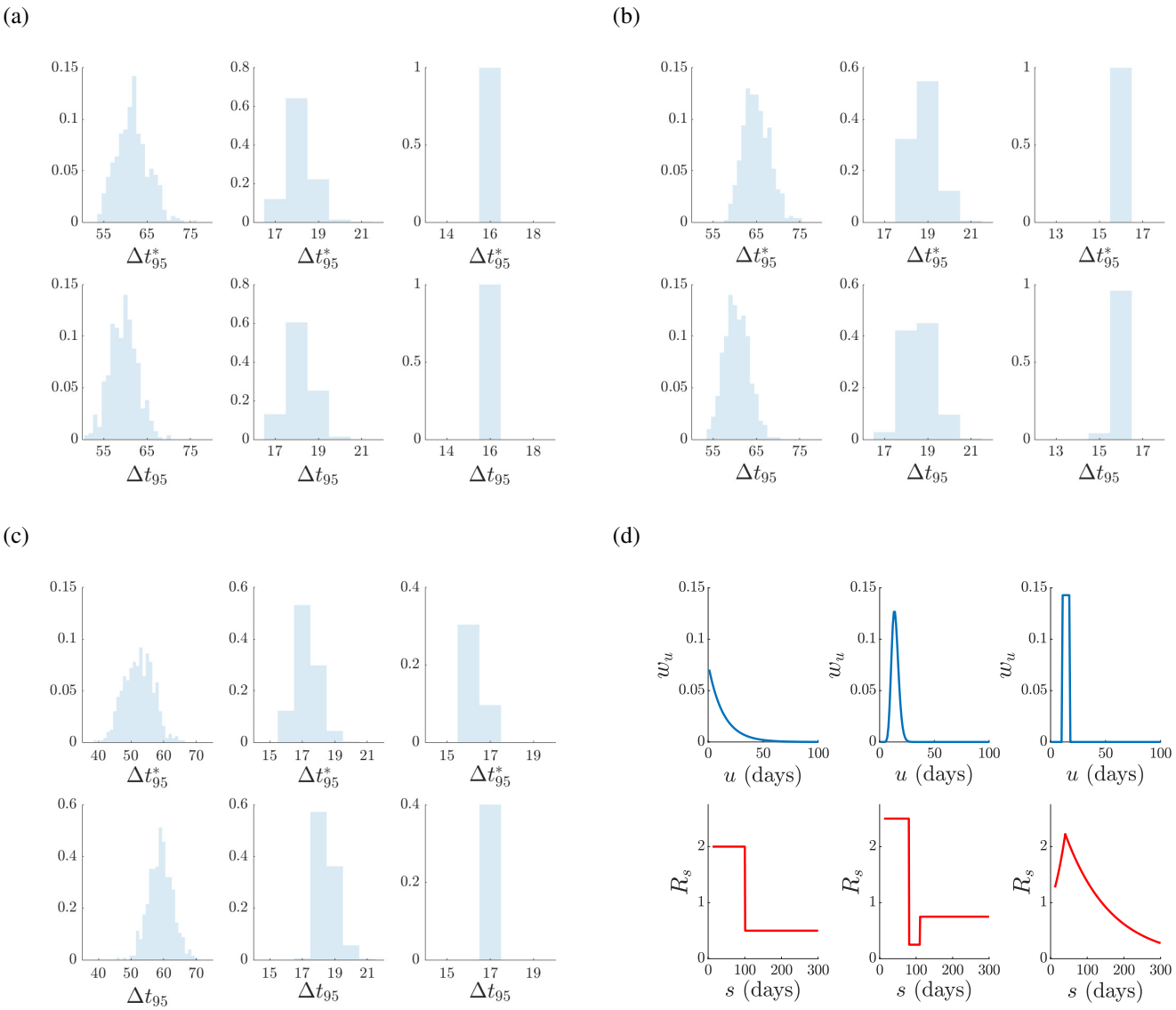

Fig. 3: **True and estimated declaration times.** We simulate $M = 500$ independent incidence curves under renewal models with R_s profiles indicating (a) rapidly controlled, (b) recovering and (c) rising and then decaying transmission and provide normalised histograms of the true (Δt^*_{95}) and estimated (Δt_{95}) 95% relative declaration times. The bottom row of (d) plots the true R_s curves (red) in absolute time underlying (a)-(c) in order. The columns of (a)-(c) correspond to exponential, gamma and approximately delta serial interval distributions with the same mean. The top row of (d) (blue) plots these distributions over absolute time. Generally we find that $\Delta t_{95} \approx \Delta t^*_{95}$. This approximation is at its worst when the serial interval is exponential and hence maximally variable for a given mean.

We start by investigating true R_s profiles describing epidemics with (a) rapidly controlled, (b) partially recovering and (c) exponentially rising and falling transmission. For each profile we simulate $M = 500$ conditionally independent I^s_1 curves and compute $z_s \,|\, I^s_1$ and $z^*_s \,|\, I^s_1$ using Eq. (9) and Eq. (10), We then obtain relative declaration times Δt_{95} and Δt^*_{95} for each curve using Eq. (6) and the time of the last case, t_0, of that curve. This yields the normalised histograms of Fig. 3, with panels (a)-(c) identifying respective R_s profiles, which are plotted in the bottom row of (d) in order. Columns of (a)-(c) correspond to exponential, gamma (with parameters taken from [27] to match Ebola virus epidemics) and approximately delta distributed serial interval distributions, shown in the top row of (d). All distributions in (d) have approximately the same mean.

These simulated examples cover many practical models (and epidemic growth patterns) as described in [18]

medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

and provide some key insight into end-of-epidemic declarations. Specifically, we find that the variability of the tail of the serial interval distribution (for a given mean) controls the variance and mean of Δt_μ and how well it approximates Δt_μ^*. This is especially obvious in the simulations featuring the exponential distribution, which is maximally variable for a given mean, where relative declaration times can vary of the order of months making time-triggered approaches likely to be strongly biased and unreliable. While in the gamma distributed case this bias falls to the order of one week, time-triggered approaches can only be justified in the limit of very tight serial intervals as in the delta distributed case.

Generally, we find that $\Delta t_\mu^* \approx \Delta t_\mu$ and note that the biggest discrepancy, which is under the exponential distribution is still reasonable given the variation in the individual relative declaration times. This also represents the worst case performance. While most realistic serial intervals are likely to be represented by a gamma-type distribution, which has an analogue to the infectious and latent periods of an SEIR compartmental model, the exponential and delta distribution also have meaning as reflecting the dynamics of SIR models and classical branching processes respectively [18, 29]. When the serial interval is tightly specified, it appears that the specific dynamics of R_s are not as important in determining the declaration times provided it remains notably below 1. Last, we comment (not shown) that the variability of the relative declaration times also increases as μ decreases.

At present, we have only verified our method for under ideal reporting conditions. Practical surveillance will be investigated in subsequent sections. We now compare our method to that of [2], which assumes ideal surveillance. This approach describes epidemic transmission using a NB branching process that is strictly only valid at the beginning of the outbreak and which differs noticeably from our renewal model. We compare both methods on MERS-CoV data from South Korea, first investigated in [2]. Note that the elimination probabilities derived in [2] are a mathematically conservative approximation of our z_s. We use the same set of bootstrapped incidence curves generated from fitting the model of [2] to the MERS-CoV data to obtain confidence intervals over the probability of elimination from each method.

Fig. 4 presents our main results with time relative to the last observed case in each bootstrap (Δs). While the median 95% relative declaration times (black circles) are reasonably close, the approach of [2] leads to a late declaration. This effect is reduced if we use the lower bound of the z_s curves instead of their median. When z_s is small (which is not practical for defining end-of-epidemic declarations) we find that the methods are less consistent. The WHO declaration time for this epidemic is at least one week later than the time proposed by both methods [2]. While our method shows wider uncertainty, the similarity of these intervals suggests that our formulation is robust to moderate model mismatch.

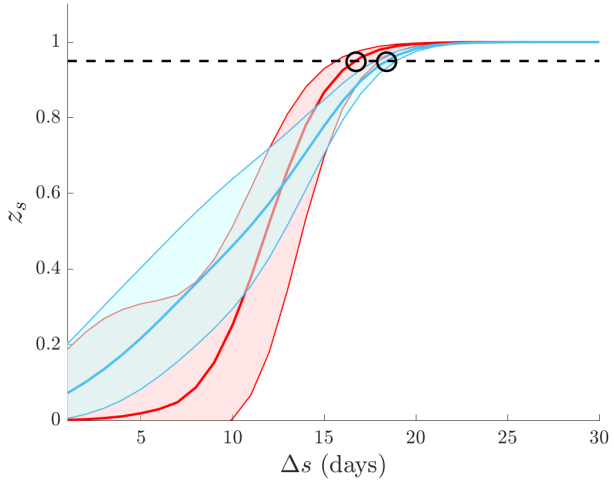

Fig. 4: **Empirical method comparison.** We compare 95% confidence intervals on the elimination probability from [2] (blue) and z_s from Eq. (9) (red) from bootstrapped epidemics based on the MERS-CoV data from South Korea used in [2]. Black circles define the median time when each method deems the epidemic to be over with 95% confidence (the event trigger). Time is relative to the last observed case in each epidemic bootstrap.

Under-reporting leads to premature declarations

Having verified z_s and hence t_μ as reliable and sensible means of assessing the conclusion of an epidemic, we now investigate the effect of model mismatch due to imperfect surveillance. We start with case under-reporting, which affects all infectious disease outbreaks to some degree. While previous works have drawn attention to how constant under-reporting can bias end-of-epidemic declarations [4] [5], no analytic results are available. Moreover, the impact of time-varying under-reporting, which models a wider range of more realistic surveillance scenarios [6, 30], remains unstudied. We provide some mathematical background for under-reporting in the renewal process framework in the Methods.

Fig. 1 illustrates how under-reporting results in only a portion, N_s, of the total local cases, I_s being sampled or observed. We use $U_s = I_s - N_s \geq 0$ to denote the unreported cases. We investigate two hypotheses or models about the incidence curve, a null one, $\mathcal{H}_0$, where we assume that the observed cases N_1^s represent all

medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

the infected individuals and an alternative hypothesis $\mathcal{H}_1$, in which the unreported cases U_1^s (and hence I_1^s) are known and distinguished. The estimated elimination probabilities under both surveillance models are:

$$\begin{aligned}\mathcal{H}_0 &: z_s \,|\, N_1^s = \mathbb{P}\left(\wedge_{j=s}^{\infty} N_{j+1} = 0 \,|\, N_1^s\right) \text{ and} \\ \mathcal{H}_1 &: z_s \,|\, I_1^s = \mathbb{P}\left(\wedge_{j=s}^{\infty} I_{j+1} = 0 \,|\, N_1^s \wedge U_1^s\right).\end{aligned} \tag{12}$$

Here $\mathcal{H}_0$ portrays a naive interpretation of the observed (N_s) incidence, while $\mathcal{H}_1$ indicates ideal surveillance. Intensive and targeted population testing should interpolate between $\mathcal{H}_0$ and $\mathcal{H}_1$. We compute $z_s \,|\, N_1^s$ by constructing the sampled total infectiousness $\tilde{\Lambda}_s := \sum_{u=1}^{s-1} w_u N_{s-u}$ and then applying Theorem 1. This follows because N_s can also be described by a Poisson renewal model (see Methods for details). We therefore find that $z_s \,|\, N_1^s = \prod_{j=s}^{\infty}(1+\tilde{\Lambda}_{j+1}\tilde{\beta}_{\tau(j)})^{-a-n_{\tau(j)}}$ with $n_{\tau(j)}$ and $\tilde{\lambda}_{\tau(j)}$ as the sums of N_u and $\tilde{\Lambda}_u$ within the $\tau(j)$ window and $\tilde{\beta}_{\tau(j)} = 1/c+\tilde{\lambda}_{\tau(j)}$. We get $z_s \,|\, I_1^s$ directly from Eq. (8) since this is the perfect surveillance case.

Since $N_s \leq I_s$ for all s then $\lambda_{\tau(j)} \geq \tilde{\lambda}_{\tau(j)}$ for all j implying that $\beta_{\tau(j)} \leq \tilde{\beta}_{\tau(j)}$. This means that $z_s \,|\, N_1^s \geq \prod_{j=s}^{\infty}(1+\tilde{\Lambda}_{j+1}\beta_{\tau(j)})^{-a-n_{\tau(j)}} := c$. From Eq. (8) we can rewrite $z_s \,|\, I_1^s = \prod_{j=s}^{\infty}(1+\Lambda_{j+1}\beta_{\tau(j)})^{-n_{\tau(j)}-u_{\tau(j)}-a}$ with $u_{\tau(j)} = i_{\tau(j)} - n_{\tau(j)}$ as the total number of unreported cases in the window $\tau(j)$. We examine the ratio of $z_s \,|\, N_1^s$ to $z_s \,|\, I_1^s$, which is at least as large as $c/(z_s \,|\, I_1^s)$. If this ratio is above 1 then the elimination probability is being inflated by imperfect surveillance. We find that $c/(z_s \,|\, I_1^s) = \prod_{j=s}^{\infty}(1+\Lambda_{j+1}\beta_{\tau(j)})^{u_{\tau(j)}}\left(\frac{1+\Lambda_{j+1}\beta_{\tau(j)}}{1+\tilde{\Lambda}_{j+1}\beta_{\tau(j)}}\right)$. Since $\Lambda_j \geq \tilde{\Lambda}_j$ at every j and the remaining term is always ≥ 1 we do find this inflation and consequently

$$z_s \,|\, N_1^s \geq z_s \,|\, I_1^s \implies t_\mu \,|\, \mathcal{H}_0 \leq t_\mu \,|\, \mathcal{H}_1. \tag{13}$$

At no point have we assumed any form for the under-reporting fraction, denoted ρ_s at time s (see Methods).

Thus any under-reporting, whether constant (i.e. all ρ_s are the same) or time-varying will engender early or false-positive end-of-epidemic declarations provided N_s is randomly sampled from I_s (so Theorem 1 holds; see Eq. (3)). We highlight this principle by examining a random sampling scheme using empirical SARS 2003 data from Hong Kong [10]. We binomially sample the SARS incidence with random probability $\rho_s \sim \text{Beta}(a, b)$. We set $b = 40$ and compute a so that the mean sampling fraction $\mathbb{E}[\rho_s] = f_\rho$ takes some desired (fixed) value. We investigate various f_ρ and show that early declarations are guaranteed in (a) and (b) of Fig. 5. The impact of ρ_s is especially large when under-reporting leads to early but false sequences of 0 cases. We present results in absolute time to showcase this effect.

Importation results in late declarations

The influence of imported cases on end-of-epidemic declarations has not been investigated in the literature. Repeated importations or migrations of infected cases are a common means of seeding and re-seeding local infectious epidemics. We assume that I_s is the total count of local cases in our region of interest but that at time s there are also M_s imported cases that have migrated from neighbouring regions. The total number of infected cases observed is $C_s = I_s + M_s$ as displayed in Fig. 1. We provide mathematical background on how importations are included within the renewal framework in Methods. We consider two hypotheses about our observed incidence data that reflect real epidemic scenarios.

Under the null hypothesis, $\mathcal{H}_0$, we assume that all cases are local and so we cannot disaggregate the components of C_s. The alternative, $\mathcal{H}_1$, assumes perfect surveillance. Imported cases are distinguished from local ones under $\mathcal{H}_1$ and their differing impact considered. The relevant elimination probabilities for each model are

$$\begin{aligned}\mathcal{H}_0 &: z_s \,|\, C_1^s = \mathbb{P}\left(\wedge_{j=s}^{\infty} C_{j+1} = 0 \,|\, C_1^s\right) \text{ and} \\ \mathcal{H}_1 &: z_s \,|\, I_1^s = \mathbb{P}\left(\wedge_{j=s}^{\infty} I_{j+1} = 0 \,|\, I_1^s \wedge M_1^s\right).\end{aligned} \tag{14}$$

Since $\mathcal{H}_0$ deems all cases local, it models C_s as a renewal process with total infectiousness $\bar{\Lambda}_s := \sum_{u=1}^{s-1} C_{s-u} w_s$. Thus we use Theorem 1 to obtain the j^{th} factor of $z_s \,|\, C_1^s$ as $q_j \,|\, C_1^s = (1+\bar{\Lambda}_{j+1}\bar{\beta}_{\tau(j)})^{-a-c_{\tau(j)}}$ with $\bar{\beta}_{\tau(j)} = 1/c+\bar{\lambda}_{\tau(j)}$. Here $c_{\tau(j)}$ and $\bar{\lambda}_{\tau(j)}$ are sums of C_u and $\bar{\Lambda}_u$ over window $\tau(j)$.

Under $\mathcal{H}_1$ the imported cases are distinguished but all cases still contribute to ongoing local transmission [7, 26]. Consequently, I_s still adheres to a renewal transmission process and Theorem 1 yields the j^{th} factor of $z_s \,|\, I_1^s$ as $q_j \,|\, I_1^s = (1+\bar{\Lambda}_{j+1}\bar{\beta}_{\tau(j)})^{-a-i_{\tau(j)}}$. We compare $q_j \,|\, I_1^s$ with $q_j \,|\, C_1^s$ directly to easily prove that

$$z_s \,|\, C_1^s \leq z_s \,|\, I_1^s \implies t_\mu \,|\, \mathcal{H}_0 \geq t_\mu \,|\, \mathcal{H}_1. \tag{15}$$

Not accounting for migrations shrinks the elimination probability leading to false-negative or unnecessarily late declarations. This result makes no assumption on the dynamics for importation other than it possesses Poisson noise (so Theorem 1 is valid for C_s) and so holds quite generally (see Methods for further details).

We illustrate this phenomenon using empirical MERS-CoV data from Saudi Arabia [31] in (c) and (d) of Fig. 5. Here repeated importations occur as zoonotic camel to human transmissions. We show the increasing effect of importation by adding further (artificial) imports via a Poisson noise variable with mean ϵ (see Eq. (4)). The mean fraction of imported to total cases across the incidence curve is then f_ϵ. In Fig. 5 we see that larger ϵ

medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

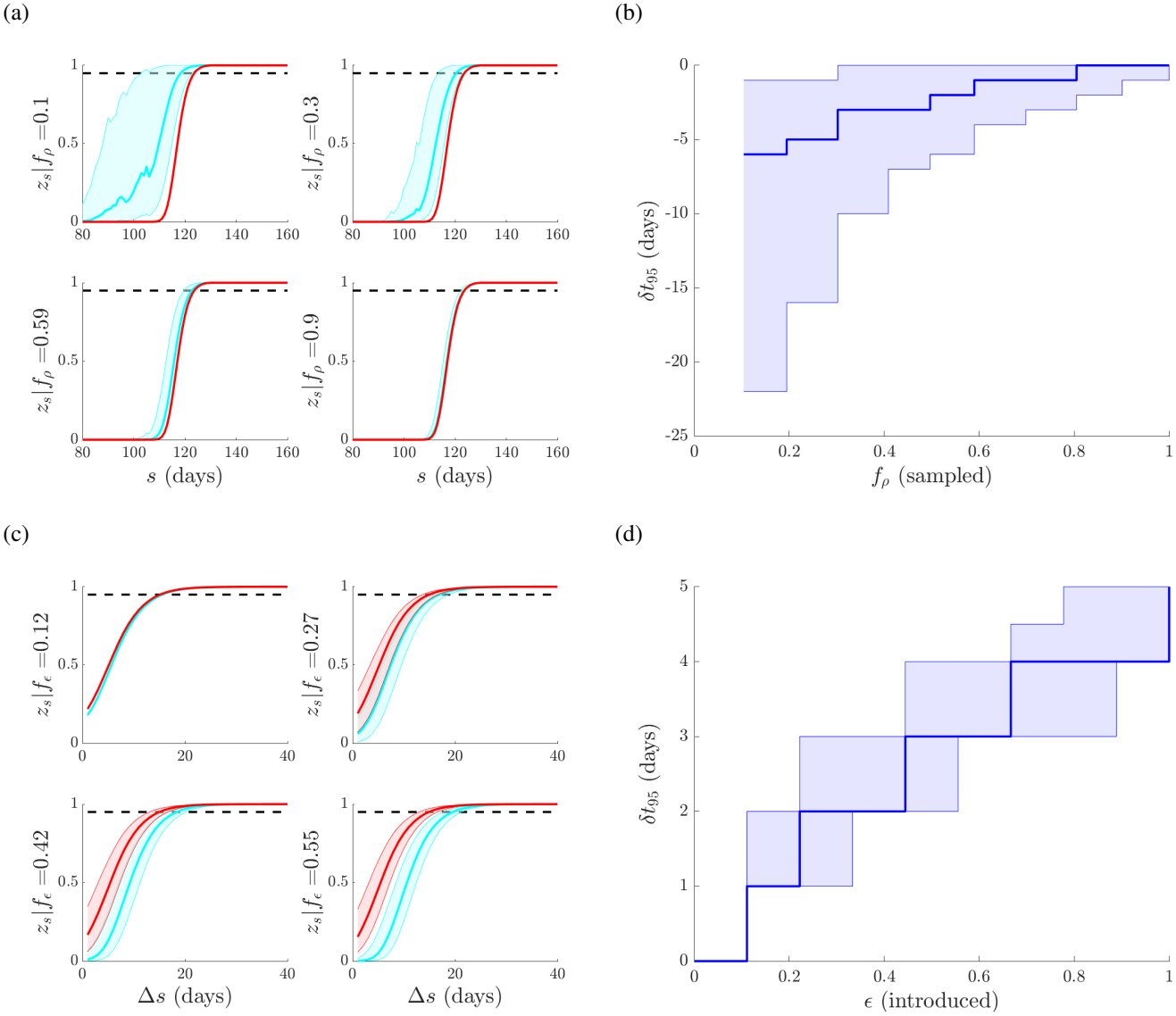

Fig. 5: **Case under-reporting and importation lead to early and late declarations respectively.** In (a) and (b) we binomially sample an empirical SARS 2003 incidence curve from Hong Kong with reporting probabilities drawn from a beta distribution with mean f_ρ. In (a) we plot the elimination probability z_s when surveillance is ideal i.e. there is no underreporting (red) versus when the under-reporting is unknown (blue). The difference in the 95% declaration times, denoted δt_{95}, from these curves is in (b). As f_ρ increases we are more likely to declare early. In (c) and (d) we consider an empirical MERS-CoV 2014-5 incidence curve from Saudi Arabia with local and imported cases. We increase the fraction of imported cases to f_ϵ by adding Poisson imports with mean ϵ and in (c) compute z_s with (red) and without (blue) accounting for the difference between imports and local cases. The change in t_{95} is given in (d). As ϵ and hence f_ϵ increase late declarations become more likely. We repeat our sampling or importation procedure $M = 1000$ times to obtain confidence intervals in (a)–(d). As $f_\epsilon \to 0$ or $f_\rho \to 1$ we attain the ideal of no unreported or imported cases.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

promotes increasingly later declaration times. In Fig. 5 we do not add any noise beyond the time of the last local case. If imports do come after this case it is likely to change the time at which t_0 is assumed and hence will notably worsen the bias from importation.

Discussion

Understanding and predicting the temporal dynamics of infectious disease transmission in real time is crucial to controlling existing epidemics and to thwarting future resurgences of those outbreaks, once controlled [20]. To achieve this understanding it is necessary to characterise and study the infectious disease throughout its lifetime. While many works have focussed on the growth, peak and controlled phases of epidemics (see Fig. 2), relatively less research has examined how the tail of the outbreak shapes the kinetics of its elimination. For example, while much is known about how the basic and effective reproduction numbers influence the growth rate, peak size and controllability of an epidemic [18, 32], the relationship between these numbers and the waiting time to epidemic elimination is still largely unexplored.

However, this relationship has important implications for public health policy. Knowing when to relax non-pharmaceutical interventions, such as social distancing or lockdowns, can be essential to effectively managing and mitigating the financial and social disruption caused by an outbreak as well as to safeguarding populations from the risk of future waves of the disease [1, 2]. The ongoing COVID-19 pandemic for instance, which in many countries is now entering the controlled phase, provides a current and important example where this question might soon become urgent.

Existing WHO guidance on deciding when an outbreak can be safely declared over takes a time-triggered approach. This means a fixed waiting time from the last observed case, usually based on the incubation period of the disease, is adopted [3]. While this approach is easy to follow, it does not change between outbreaks of the same disease, even if the patterns of transmission are very different and cannot provide a measure of the reliability of this suggested declaration time. The few existing studies that have investigated this waiting-time problem [2, 4, 5] have all converged on what is known as an event-triggered solution in control theory [11].

Event-triggered decision-making has been shown to be more effective than acting at deterministic or fixed times for a range of problems including several involving the optimising of waiting or stopping times [12, 13, 14, 15]. Moreover, because it directly couples decision making to observables of interest (in our case the incidence curve), it can better adapt or respond to changes in dynamics. Here we have attempted to build upon these realisations to better characterise the relationship between epidemic transmission and elimination. Specifically, we focussed on computing the probability at time s, z_s, that the total future incidence of the epidemic is zero.

This probability is directly responsible for determining how quickly an epidemic will end. In fact, if an outbreak is defined as surviving if it can propagate at least 1 future infection then $1 - z_s$ is precisely its survival function and is therefore rigorously linked to the future risk of cases. By taking a renewal process approach, we were able to derive an analytic and real-time measure of z_s that explicitly depends on up-to-date estimates of the effective reproduction number (see Eq. (9)). This result formed the main theorem of this paper and provided a clear and easily-computed link between epidemic transmission and elimination. To our knowledge, no previous work has directly obtained z_s. Specifically, [2] computed a simpler and more conservative quantity while [4] and [5] approximated something similar via simulation, and so cannot provide real-time formulae. The event-trigger for declaring an outbreak over with $\mu\%$ confidence is then the first time that z_s crosses a threshold of $\frac{\mu}{100}$.

To validate the correctness of our approach we considered several comparisons. We proved mathematically that our formulae recover the true elimination probability and event trigger given perfect knowledge of the epidemic. This provided theoretical justification for our approach (Eq. (11)). We verified practical performance by benchmarking our method against the known (true) declaration times from simulated outbreaks (Fig. 3) and on empirical data by comparing to [2] (Fig. 4). Fig. 3 also explained why time-triggered methods can be unreliable. Diseases with wider serial interval distributions engender more inherently variable declaration times, which cannot be summarised well by fixed or deterministic times.

A key motivation for developing our method was to gain rigorous insight into the tail dynamics of epidemics. We therefore explored two prevalent sources of noise in surveillance – unreported and imported cases. While [4, 5] both looked at the effect of constant under-reporting on declarations, general insight into the more realistic time-varying case is lacking. Further, no analyses have yet considered the influence of importation on the epidemic tail. By adapting z_s to various surveillance hypotheses we proved two key results. First, we showed that any type of random under-reporting will precipitate early declarations, which worsen as the fraction of unreported cases increases (Eq. (13)).

Second, we found that any random importation process will lead to conservative declarations. This effect

medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

is more exaggerated as the fraction of imports increase (Eq. (15)). We illustrated the biases of both unreported and imported cases using empirical data (Fig. 5). These results provide a clearer picture of how the epidemic tail is sensitive to imperfections in the collection or reporting of incidence data and highlights a need for continued, heightened surveillance both in the quality of data (e.g. intensive testing rates can reduce under-reporting or at least measure it) and associated metadata (i.e. this can prevent misidentification of cases which is the main issue with unknown or unrecognised imports).

While our method provides a straightforward framework for estimating the lifetime of an epidemic and for investigating various surveillance noise sources, it has several limitations. It assumes that the serial interval is stationary and that reporting delays can be ignored [10]. Moreover, we neglect transmission heterogeneity and do not assess interactions among sources of reporting noise. While these could bias z_s and alter declaration times, some of these more realistic dynamics can be included as future extensions. We can adjust for delays by applying nowcasting techniques [24] and include heterogeneity by using a negative binomial renewal model [1]. Future generalisations of our method will consider how data about reporting trends (e.g. from seroprevalence or case ascertainment studies) might be included to improve end-of-epidemic time estimates and compensate for biases.

Real-time assessments of epidemic dynamics are crucial for understanding and aptly responding to unfolding epidemics [20]. We hope that the analytic approach that we developed here will serve as a useful tool for gaining ongoing insight into the tail dynamics of an outbreak, motivate the adoption of more event-triggered decision making and provide clear impetus for improving and sustaining surveillance across all phases of an epidemic. Our method is freely available in both R and Matlab at https://github.com/kpzoo/end-of-epidemic.

Acknowledgments

KVP and CAD acknowledge joint centre funding from the UK Medical Research Council and Department for International Development under grant reference MR/R015600/1. RNT thanks Christ Church (Oxford) for funding via a Junior Research Fellowship. CAD thanks the UK National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Modelling Methodology at Imperial College London in partnership with Public Health England (PHE) for funding (grant HPRU-201210080). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

Conceptualization: KVP and RNT. Formal analysis, investigation, methodology, project administration, software, visualisation and writing (original draft preparation): KVP. Validation: KVP, RJ and RNT. Writing (review and editing): KVP, RNT, RJ and CAD.

References

1. Lee H, Nishiura H. Sexual transmission and the probability of an end of the Ebola virus disease epidemic. J Theor Biol. 2019;471:1–12.
2. Nishiura H, Miyamatsu Y, Mizumoto K. Objective determination of end of MERS outbreak, South Korea. Emerg Infect Dis. 2016;22:146–8.
3. WHO. WHO recommended criteria for declaring the end of the Ebola virus disease outbreak; 2020. Available from: https://www.who.int/who-documents-detail/who-recommended-criteria-for-declaring-the-end-of-the-ebola-virus-disease-outbreak.
4. Thompson R, Morgan O, Jalave K. Rigorous surveillance is necessary for high confidence in end-of-outbreak declarations for Ebola and other infectious diseases. Phil Trans R Soc B. 2019;374:20180431.
5. Djaafara B, Imai N, Hamblion E, et al. A quantitative framework to define the end of an outbreak: application to Ebola Virus Disease. medRxiv. 2020;(20024042).
6. White L, Pagano M. Reporting errors in infectious disease outbreaks, with an application to Pandemic Influenza A/H1N1. Epidemiol Perspec Innov. 2010;7(12).
7. Churcher T, Cohen J, Ntshalintshali N, et al. Measuring the path toward malaria elimination. Science. 2014;344(6189):1230–32.
8. Yang P, Chowell G. Quantitative Methods for Investigating Infectious Disease Outbreaks. vol. 70 of Texts in Applied Mathematics. Cham, Switzerland: Springer; 2019.
9. Fraser C, Cummings D, Klinkenberg D, et al. Influenza Transmission in Households During the 1918 Pandemic. Am J Epidemiol. 2011;174(5):505–14.
10. Cori A, Ferguson N, Fraser C, et al. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. Am J Epidemiol. 2013;178(9):1505–12.
11. Astrom K, Bernhardsson B. Comparison of periodic and event based sampling for first order systems. Proc IFAC World Conf. 1999:301–6.
12. Parag K. On signalling and estimation limits for molecular birth-processes. J Theor Biol. 2019;480:262–73.
13. Rabi M, Moustakides G, Baras J. Adaptive Sampling for Linear State Estimation. SIAM Journal of Control and Optimization. 2012;50(2):672–702.
14. Parag K, Vinnicombe G. Point Process Analysis of Noise in Early Invertebrate Vision. PLOS Comput Biol. 2017;13(10):e1005687.
15. Lemmon M. Event-Triggered Feedback in Control, Estimation, and Optimization. vol. 406 of Networked Control Systems. London: Springer; 2010. p. 293–358.
16. Bhatia S, Cori A, Parag K, et al.. Short-term forecasts of COVID-19 deaths in multiple countries.; 2020. Available from: https://mrc-ide.github.io/covid19-short-term-forecasts.
17. Pybus O, Rambaut A, du Plessis L, Zarebski A, et al.. Preliminary analysis of SARS-CoV-2 importation & establishment of UK transmission lineages; 2020. Available from: https://virological.org/t/preliminary-analysis-of-sars-cov-2-importation-establishment-of-uk-transmission-lineages [cited 13 June 2020].
18. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. Proc R Soc B. 2007;274:599–604.
19. Fraser C. Estimating Individual and Household Reproduction Numbers in an Emerging Epidemic. PLOS One. 2007;8:e758.
20. Cauchemez S, Boelle P, Thomas G, et al. Estimating in Real Time the Efficacy of Measures to Control Emerging Communicable Diseases. Am J Epidemiol. 2006;164(6):591–7.
21. Parag K, Donnelly C. Using information theory to optimise epidemic models for real-time prediction and estimation. PLOS Comput Biol. 2020;16(7):e1007990.
22. De Serres G, Gay N, Farrington P. Epidemiology of Transmissible Diseases after Elimination. Am J Epidemiol. 2000;151(11).
23. Parag K, Donnelly C. Adaptive Estimation for Epidemic Renewal and Phylogenetic Skyline Models. Syst Biol. 2020;(syaa035).
24. Azmon A, Faes C, Hens N. On the estimation of the reproduction number based on misreported epidemic data. Stats Med. 2014;33:1176–92.
25. Raikov D. On the decomposition of Poisson laws. Dokl Acad Sci URSS. 1937;14:9–11.
26. Roberts M, Nishiura H. Early Estimation of the Reproduction Number in the Presence of Imported Cases: Pandemic Influenza H1N1- 2009 in New Zealand. PLOS One. 2011;6(5):e17835.
27. Nouvellet P, Cori A, Garske T, et al. A simple approach to measure transmissibility and forecast incidence. Epidemics. 2018;22:29–35.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.13.20152082; this version posted July 14, 2020. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

28. White L, Pagano M. A likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic. Stats Med. 2008;27:2999–3016.
29. Champredon D, Dushoff J, Earn D. Equivalence of the Erlang-distributed SEIR Epidemic Model and the Renewal Equation. SIAM J Appl Math. 78;6(3258-78).
30. Parag K, du Plessis L, Pybus O. Jointly inferring the dynamics of population size and sampling intensity from molecular sequences. Mol Biol Evol. 2020;msaa016.
31. Thompson R, Stockwin J, van Gaalen R, et al. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. Epidemics. 2019;29:100356.
32. Brauer F, van den Driessche P, Wu J, editors. Mathematical Epidemiology. Lecture Notes in Mathematics. Berlin, Germany: Springer-Verlag; 2008.