JOURNAL OF THE ROYAL SOCIETY INTERFACE

Inferring the reproduction number using the renewal equation in heterogeneous epidemics

Journal:	*Journal of the Royal Society Interface*
Manuscript ID	Draft
Article Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Green, William; Imperial College London, Infectious Disease Epidemiology Ferguson, Neil; Imperial College London, MRC Centre for Global Infectious Disease Analysis; Imperial College London, Abdul Latif Jameel Institute for Disease and Emergency Analytics Cori, Anne; Imperial College, MRC Centre for Global Infectious Disease Analysis
Categories:	Life Sciences - Mathematics interface
Subject:	Computational biology < CROSS-DISCIPLINARY SCIENCES
Keywords:	heterogeneous, epidemics, renewal equation, generation time, asymptomatic transmission

Author-supplied statements

Relevant information will appear here if provided.

Ethics

Does your article include research that required ethical approval or permits?:
This article does not present research with ethical considerations

Statement (if applicable):
CUST_IF_YES_ETHICS :No data available.

Data

It is a condition of publication that data, code and materials supporting your paper are made publicly available. Does your paper present new data?:
Yes

Statement (if applicable):
Data on Ebola cases and taken from: Heterogeneities in the case fatality ratio in the West African Ebola outbreak 2013â€“2016 (T. Garske, A. Cori et al, 2017)

Data on SARS-CoV-2 deaths in the UK taken from the Government dashboard:
https://coronavirus.data.gov.uk/

Both data and code are available in the github repository:
https://github.com/willgreen236/Heterogeneity_transmission.git

Conflict of interest

I/We declare a competing interest

Statement (if applicable):
William Green has received payment fromt he Joint Biosecurity Centre for services in relation to the UK's response to the COVID-19 pandemic.

Anne Cori has received payment from Pfizer for teaching of mathematical modelling of infectious disease transmission and vaccination

Authors' contributions

This paper has multiple authors and our individual contributions were as below

Statement (if applicable):
All authors were responsible for developing the project methodology. William Green developed the software, conducted the formal analysis and was responsible for output visualization and original draft preparation. Anne Cori and Neil Ferguson were jointly responsible for project conceptualization, project supervision and reviewing and editing the manuscript.

1

Inferring the reproduction number using the renewal equation in heterogeneous epidemics

Abstract

Real-time estimation of the reproduction number has become the focus of modelling groups around the world as the SARS-Cov-2 pandemic unfolds. One of the most widely adopted means of inference of the reproduction number is via the renewal equation, which uses the incidence of infection and the generation time distribution. In this paper, we derive a multi-type equivalent to the renewal equation which accounts for heterogeneity including asymptomatic transmission, symptomatic isolation, and vaccination. We demonstrate how use of the renewal equation that misses these heterogeneities can result in biased estimates of the reproduction number. While the bias is small with symptomatic isolation, it can be much larger with asymptomatic transmission or transmission from vaccinated individuals if these groups exhibit substantially different generation time distribution to unvaccinated symptomatic transmitters, whose generation time distribution is often well defined. The bias in estimate becomes larger with greater population size or infectiousness of the poorly characterised group, as well as if the population exhibits disassortative mixing. We apply our methodology to Ebola in West Africa in 2014 and the SARS-CoV-2 in the UK in 2020-21.

Introduction

The effective reproduction number, R, defined as the average number of secondary infections generated by each primary case, is of fundamental importance in infectious disease epidemiology. When R is above 1, infection prevalence is expected to increase exponentially, whereas if R is below 1, it will decline. As such, interventions for epidemic control generally aim to reduce the R to below unity.

Estimation of R has taken on particular significance over the past year in light of the global COVID-19 pandemic, which is so far responsible for over 100 million cases, and 2.6 million deaths worldwide (1). Given the importance of R in elucidating the extent of control measures required to suppress the epidemic, real-time estimation of R has been the focus of disease modelling groups and government health departments worldwide (2).

The effective reproduction number principally depends on the underlying infectiousness of the pathogen in a totally susceptible population, contacts between individuals in the population, and the level of immunity in the population. The underlying infectiousness is often represented by the basic reproduction number, R_0, defined as the average number of secondary infections arising from a primary case in a large totally susceptible population. R may be further modified by changes in the number, frequency, and closeness of contacts in a population, hygiene practices, seasonal variation, population demographics and pathogen evolution. Given the myriad of variables contributing to R, it is generally estimated from trends in infections, cases, hospitalizations or deaths over time, rather than from R_0 (3–6).

There are two distinct reproduction numbers than can be derived from data on infection incidence. The instantaneous reproduction number, henceforth denoted $R(t)$, represents the average number of individuals someone infected at time t would infect if conditions remained unchanged. Conversely, the case reproduction number, $R_c(t)$, represents the average number of people an individual infected at time t actually infects, which will depend on changes in policy or behaviour over the period of that

cohort's infection, and can thus only be estimated in retrospect (7). The work of this paper focusses on the former, which is better suited to track changes in transmissibility in real-time, and which will reduce immediately following the start of a successful intervention, unlike $R_c(t)$ which will change gradually and is only possible to calculate with hindsight (8).

A simple and widely used approach to estimate the effective reproduction number is via the renewal equation, which uses as inputs the generation time distribution, $\omega(\tau)$ (the distribution of times, τ, between infection in a case and infection of their infector) and the time-series of infection incidence (7). In this method, all infected individuals are assumed to have the same infectious profile over time, characterized by the generation time distribution, and are assumed to generate the same average number of secondary cases.

However, variation in infectious profiles may results from biological and behavioural differences between individuals. An example of biological differences is clinical presentation resulting in symptomatic or asymptomatic infection. Similarly, a subset of symptomatic individuals may change their behaviour to limit their social contacts (self-isolation), as is currently mandated in UK law for both confirmed SARS-CoV-2 cases and their immediate household. Interventions that alter viral dynamics, resulting in faster viral clearance or a reduced viral load, for example though deployment of novel antivirals (e.g. ART) or through vaccination priming the immune response in a subset of individuals, may also contribute to differences in infectious profile between cases.

The generation time distribution is hard to estimate directly, given that it is often difficult to identify the exact timing of an infection event, let alone the timing of two sequential infection events required for inference of the generation time distribution. One of two approaches is typically used:

1. Fitting a generation time distribution to infector-infectee relationships where the time of infection of the index case and infection of the secondary case are unambiguously known
2. Estimating via the serial interval distribution (the time between symptom onset in a case and symptom onset of their infector), which is much easier to measure. This is done either by assuming the generation time distribution and serial interval distribution are equivalent, or by convolution with the incubation period distribution (the time from infection to onset of symptoms). The optimal approach will depend upon the joint relationship of the infectious distribution and the incubation period (9). It is worth noting that the observed distribution of serial intervals is also affected by epidemic dynamics at the time of measurement. For example, in an exponentially growing outbreak, there will be many more recent infected who are potential infectors, so long serial intervals will be censored (10).

In a novel outbreak, initial estimates of the generation time distribution are typically based on an analysis of the "first few hundred cases" (FFHC) (11–15), with little emphasis on characterizing heterogeneities among infected individuals.

In this paper, we derive a multi-type equivalent of the renewal equation which accounts for heterogeneity in infectious profiles including variation in case isolation behaviour, symptomatic/asymptomatic infection, and heterogeneity introduced owing to vaccine roll-out. We then explore how much the corresponding estimated R differs from a "naïve" R derived from a single-type branching process based on a single generation time distribution that does not account for variation caused by case isolation, asymptomatic transmission, or vaccination.

We consider two applications: to Ebola Virus Disease in Guinea in 2014-15, and to SARS-CoV-2 in the UK between March 2020 and January 2021, to illustrate the potential impact on R estimates of neglecting heterogeneities due to case isolation and asymptomatic transmission.

Methods

Single-type renewal equation

For a single-type epidemic, the renewal equation gives the relationship between the expected incidence, or number of new infected individuals on time *t*, $I(t)$, the instantaneous reproduction number at time *t,* $R(t)$, and the generation time distribution, as function of time τ since infection, $\omega(\tau)$: $I(t) = R(t)\int_0^\infty \omega(\tau)I(t-\tau)d\tau$ (16). This assumes that the generation time distribution remains constant in the period up from time *t* - τ_{max} to time *t,* where τ_{max} is the maximum time after infection at which a case can onwardly infect.

Assuming exponential growth/decay for the incidence $I(t) = ke^{rt}$ gives the Wallinga-Lipsitch form of the relationship between the reproduction number and the growth rate *r* when the growth rate is a constant (Equation 1) (17).

Equation 1

$$R = \frac{1}{\int_0^\infty \omega(\tau)e^{-r\tau}d\tau}$$

This enables us to generate the relationship between the reproduction number and growth rate for various generation time distributions in a homogeneous epidemic. In the work which follows in this paper we use the gamma distribution for analytical tractability. Substitution of a gamma distribution with shape α and rate β into Equation 1 yields an analytic equation for the reproduction number in terms of the epidemic growth rate (Equation 2).

Equation 2

$$\omega(\tau) = \frac{\beta^\alpha}{\Gamma(\alpha)}\tau^{\alpha-1}e^{-\beta\tau} \quad ; \quad R = \frac{(r-\beta)^\alpha}{\beta^\alpha}$$

Multitype renewal equation

Moving to a paradigm where there are *n* groups with generation time distribution for a case in group *i* given by $\omega_i(\tau)$, we consider the next generation matrix, a *n* x *n* matrix where $R_{j\to i}$ represents the average number of secondary cases in group *i* resulting from one index case in group *j*. The values of $R_{j\to i}$ in turn will depend on the overall susceptibility and infectiousness of each group, and the extent of assortativity between groups.

In this case the renewal equation becomes multi-dimensional, and takes the form given in Equation 3. As above, this assumes that the generation time distribution for each group remains constant in the period up from time *t* - τ_{max} to time *t*.

Equation 3

$$\begin{pmatrix} I_1(t) \\ \vdots \\ I_n(t) \end{pmatrix} = \int_0^\infty \begin{pmatrix} R_{1\to1}(t)\omega_1(\tau) & \cdots & R_{n\to1}(t)\omega_n(\tau) \\ \vdots & \ddots & \vdots \\ R_{1\to n}(t)\omega_1(\tau) & \cdots & R_{n\to n}(t)\omega_n(\tau) \end{pmatrix} \begin{pmatrix} I_1(t-\tau) \\ \vdots \\ I_n(t-\tau) \end{pmatrix} d\tau$$

We can assume an exponential solution as for the single-type case, albeit with a vectorized scalar *k*, with elements k_i proportional to the fraction of new infections occurring in group *i* (Equation 4).

Equation 4

$$\begin{pmatrix} I_1(t) \\ \vdots \\ I_n(t) \end{pmatrix} = \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} e^{rt}$$

Substituting Equation 4 into Equation 3 yields an eigenvalue equation (Equation 5).

Equation 5

$$\begin{pmatrix} R_{1\to1}(t)\int_0^\infty \omega_1(\tau)e^{-r\tau}d\tau & \dots & R_{n\to1}(t)\int_0^\infty \omega_n(\tau)e^{-r\tau}d\tau \\ \vdots & \ddots & \vdots \\ R_{1\to n}(t)\int_0^\infty \omega_1(\tau)e^{-r\tau}d\tau & \dots & R_{n\to n}(t)\int_0^\infty \omega_n(\tau)e^{-r\tau}d\tau \end{pmatrix} \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} = \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}$$

To infer the overall reproduction number *R* we factorise the matrix into the product of the scalar reproduction number *R* and the normalized next generation matrix *M*, whose elements $M_{j\to i}$ give the relative risk posed to a member of group *i* by an infected member of group *j*. Rearranging shows that for a growth rate *r*, the corresponding *R* will be the reciprocal of the dominant eigenvalue of the elementwise product of the matrix *M* and the matrix of Laplace transforms of the generation time distributions (Equation 6).

Equation 6

$$R_{hetero}(t) = \frac{1}{\max\left\{eigen\begin{pmatrix} M_{1\to1}\int_0^\infty \omega_1(\tau)e^{-r\tau}d\tau & \dots & M_{n\to1}\int_0^\infty \omega_n(\tau)e^{-r\tau}d\tau \\ \vdots & \ddots & \vdots \\ M_{1\to n}\int_0^\infty \omega_1(\tau)e^{-r\tau}d\tau & \dots & M_{n\to n}\int_0^\infty \omega_n(\tau)e^{-r\tau}d\tau \end{pmatrix}\right\}}$$

In the following, we assume that the matrix *M* can be simply expressed using (i) the relative infectiousness of each group, ρ_i, (ii) the relative susceptibility of each group, ξ_i, and (iii) the assortativity between the groups, given by a matrix *A* whose elements A_{ij} give the proportion of group *j*'s contacts which are made with individuals in group *i* (Equation 7). The relative infectiousness and susceptibility of each group are denoted relative to the most infectious and most susceptible group respectively. We can consider ρ to be determined by biological factors and fixed through the course of an outbreak.

Equation 7

$$M = \frac{\rho\xi^T \odot \mathrm{A}}{|\rho\xi^T \odot \mathrm{A}|} \; ; \sum_i A_{ij} = 1$$

We explore the extent to which the central reproduction number estimated from the single-type renewal equation, using a homogeneous generation time distribution taken from the FFHC (Equation 1), differs from the central *R* estimate obtained when heterogeneous transmission is taken into account (Equation 6). The deterministic formulation above means we only consider the central estimates. In reality, there is stochastic variation that can be accounted for using a Poisson or negative binomial offspring distribution. In the applications we present in this paper we have used EpiEstim to capture this confidence interval, which uses a Poisson distributed offspring distribution (see below).

We consider three scenarios relevant to many infectious diseases, but in particular to SARS-Cov-2: heterogeneity due to i) the isolation of symptomatic cases on symptom onset; ii) the presence of

asymptomatic carriers; and iii) differential transmission potential of vaccinated individuals, which will be increasingly important as vaccination is rolled-out. For each scenario we assume the baseline generation time distribution is given by that of the "reference" group i.e. non-isolating, symptomatic, and non-vaccinated individuals. We consider epidemic growth rates of -0.3, -0.15, 0, 0.15, 0.3 day^{-1} which correspond to doubling times of -2.3 days, -4.6 days, steady state, 4.6 days and 2.3 days respectively[1]. We consider "reference" group sizes corresponding to 20%, 50% and 80% of the total population.

For isolating vs non-isolating cases we vary the timing of isolation, modelled as truncation of the generation time after a certain amount of the generation time distribution has passed. We assume isolation proportionately reduces the reproduction number of the isolating group. For example, if isolation occurs 40% into the generation time distribution of an otherwise non-isolating case, the infectiousness of a non-isolating individual is 40% that of an isolating individual. For symptomatic vs asymptomatic transmission, we vary the contribution of asymptomatic cases to transmission. In both of these instances we assume homogeneous mixing between groups, where row i of the assortativity matrix A will correspond to the proportion of the population that is in group i. For vaccinated vs non vaccinated cases, we also explore the impact on estimated R of assortative mixing – varied from fully disassortative mixing in which all contacts of the smaller group are with individuals in the larger group, to fully assortative mixing, in which all contacts are between individuals in the same group. We assume vaccination reduces susceptibility to infection by 70%.

We parameterize the assortativity matrix A in a similar way to (18), which described HIV transmission by considering mixing within and between sexual activity groups via the contact rates of members from each group c_1 and c_2; the population size in each group, p_1 and p_2, and an assortativity parameter δ. We do not explore the effect of heterogeneity in contact rate by individuals in each group; making the simplifying assumption that contact rates are uniform independent of group, resulting in the parameterization given in Equation 8. This parameterization requires p_1 to be the smaller group such that all matrix elements are less than or equal to 1.

Equation 8

$$A(\delta) = \begin{pmatrix} \delta & (1-\delta)\dfrac{p_1}{p_2} \\ 1-\delta & 1-(1-\delta)\dfrac{p_1}{p_2} \end{pmatrix}$$

For plotting, we vary δ in a two-part linear manner, from 0 to p_1 from the left of the x-axis to the middle, and from p_1 to 1 from the middle of the x-axis to the right-hand side. This standardizes homogeneous mixing at the centre of the x-axis, allowing comparisons between different values of p_1.

Equivalent single type formulation

The single-type formalism of the reproduction number provides a more straightforward means of inferring the reproduction number. Additionally, existing software packages used for epidemic analysis will typically only work with single-type renewal processes, so there is a benefit to expressing multi-type renewal processes as an equivalent single-type.

Equation 3 and Equation 4 can be re-written as Equation 9.

[1] A negative doubling time indicated the epidemic is in decline, with the magnitude representing the halving time

Equation 9

$$\begin{pmatrix} I_1(t) \\ \vdots \\ I_n(t) \end{pmatrix} = R(t) \int_0^\infty \begin{pmatrix} M_{1\to1}\omega_1(\tau) & \cdots & M_{n\to1}\omega_n(\tau) \\ \vdots & \ddots & \vdots \\ M_{1\to n}\omega_1(\tau) & \cdots & M_{n\to n}\omega_n(\tau) \end{pmatrix} \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} e^{r(t-\tau)} d\tau$$

The total number of newly infected individuals is then:

Equation 10

$$I(t) = \sum_i R(t) \int_0^\infty \sum_j k_j M_{j\to i} \omega_j(\tau) e^{r(t-\tau)}\, d\tau$$

$$= R(t) \int_0^\infty \tilde{\omega}(\tau) e^{r(t-\tau)} d\tau$$

Where $\tilde{\omega}(\tau)$ is given by

Equation 11.

Equation 11

$$\tilde{\omega}(\tau) = \sum_{i,j} k_j M_{j\to i} \omega_j(\tau) = \sum_j \left(k_j \omega_j(\tau) \sum_i M_{j\to i} \right)$$

Equation 10 shows that the multi-type renewal equation can be written as one-dimensional renewal equation with a weighted mean generation time distribution. The weighting is given by the overall relative infectiousness of the group j, and the j^{th} element of the eigenvector (which corresponds to the proportion of infections that occur in group j) (Equation 11). Given this *weighted single-type* approach derived in Equation 10 is equivalent to the *multi-type* approach derived in Equation 6, in what follows *weighted single-type* and *multi-type* are used interchangeably, depending on which provides the most straightforward means of explanation.

Use in EpiEstim for application to COVID-19 in the UK and Ebola Virus Disease in Guinea

Equation 1 and Equation 6 describe the relationship between the instantaneous reproduction number, the growth rate, and the generation time distribution of different groups. In practice, the growth rate is not directly observed, but can be estimated from the incidence time-series. This leads to uncertainty in the growth rate estimates, and in turn the corresponding reproduction number estimates, which are not represented in the equations above.

The EpiEstim package for R software implements estimation of the instantaneous reproduction number, based on an incidence time series and a discrete generation time distribution. EpiEstim uses a single-type renewal equation to estimate a full posterior distribution of the instantaneous reproduction number capturing full uncertainty in the estimates. We therefore use EpiEstim to compare the R estimated using a naïve single-type renewal equation (in which the generation time distribution of the FFHC is used without adjustment), with that estimated from an appropriately weighted single-type generation time distribution accounting for case isolation, asymptomatic transmission, or vaccination (equivalent to the multi-type approach). In all applications R values were estimated over sliding weekly windows.

We estimate the instantaneous reproduction number for Ebola Virus Disease (EVD) case data from Guinea between March 2014 and July 2016, with data taken from (19). The generation time distribution is assumed to be Gamma distributed with mean 15.3 days and standard deviation 9.3 days

following (20). This is assumed to be reflective of a non-isolating cohort. We assume that isolation occurs at the point of hospitalization at 14.9 days after infection (56.5% of the way into a non-isolated infectious course), based on the sum of the mean incubation period and the mean delay from symptoms to hospitalization given in (21). We assume that 34% of infected individuals are hospitalized (Unwin *et al, unpublished*), and that individuals who seek hospitalization and individuals who do not seek hospitalization mix homogeneously.

We estimate the instantaneous reproduction number of SARS-CoV-2 using timeseries of COVID-19 deaths in the UK from March 2020 to January 2021. We use incidence of deaths rather than cases because ascertainment of deaths is significantly better than that of cases, especially during the early months of the epidemic when testing capacity was highly constrained. However, due to a delay between infection and death, the resulting estimates of R will be lagged and smoothed – the average delay from infection to symptom onset is estimated to be around 5.5 days in the UK (22), while estimates of the delay from symptom onset to death range from around 13 days (23) to 18 days (24). Other papers have demonstrated inference of R accounting for lagged metrics, for example (25). Daily UK COVID-19 deaths were taken from the Government's coronavirus data repository, at (26). We assumed a generation time distribution in the absence of isolation gamma distributed with mean 5.71 days and standard deviation 3.09 days as in (27). We consider two scenarios: an optimistic scenario in which 75% of symptomatic individuals (47% of all infected) isolate 30% of the way into their infection, and a pessimistic scenario in which 25% of symptomatic individuals (16% of all infected) isolate 70% of the way into their infection.

Results

Symptomatic isolation and non-isolation

Isolation of symptomatic cases on onset will necessarily reduce the generation time distribution for those who isolate (for example, Figure 1A-C – corresponding to isolation 25%, 50% and 75% into the generation time distribution, compared to Figure 1D – corresponding to no isolation). In a growing epidemic, isolation will mean the true R is lower than that derived assuming generation time distribution representative only of non-isolating cases, and vice-verse for shrinking epidemics (Figure 1E). This error is higher with higher growth/reduction rates, as well as with higher isolating populations (Figure 1E). The error is zero at the extremes: if isolation occurs immediately upon infection; or if isolation occurs following all infection (Figure 1E). This is because if isolation occurs immediately, isolators do not contribute to the infectious pool, so the weighting of the non-isolating group in Equation 11 is zero. Likewise, if isolation occurs following the infectivity period, it is equivalent to no isolation occurring.

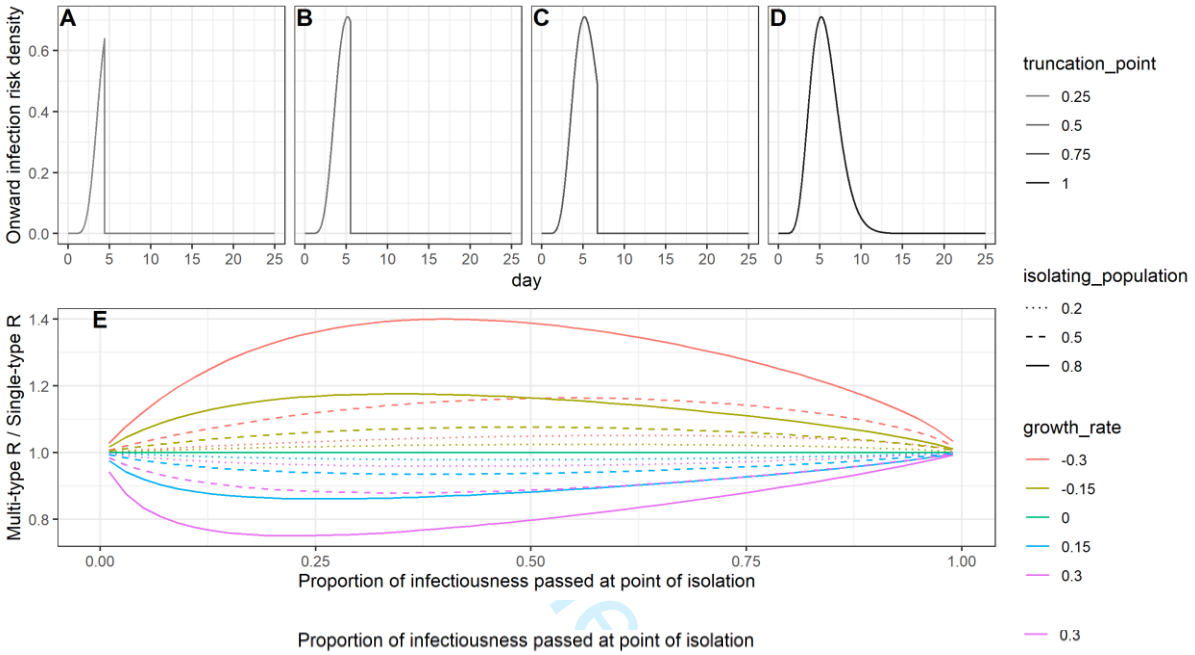
Overall, the true R with the reference generation time distribution explored is a maximum of 1.4x higher than that inferred using a single-type approach when the epidemic growth rate is -0.3 day^{-1}, and isolation occurs among 80% of the population 39% of the way into the infectious period. The true R is 0.75x lower than that inferred using a single-type approach when the epidemic growth rate is 0.3 day^{-1}, and isolation occurs among 80% of the population 23% of the way into the infectious period (Figure 1E).

One can understand why the error remains relatively small across the range by considering the one-dimensional approach presented in

Equation 11. When constructing the weighted generation time distribution, early isolation will lead to a low weighting of the isolating group in the overall generation time distribution; on the other

hand, later isolation will result in little difference in the generation time distributions. This inherently constrains the impact isolation has on the derived R.

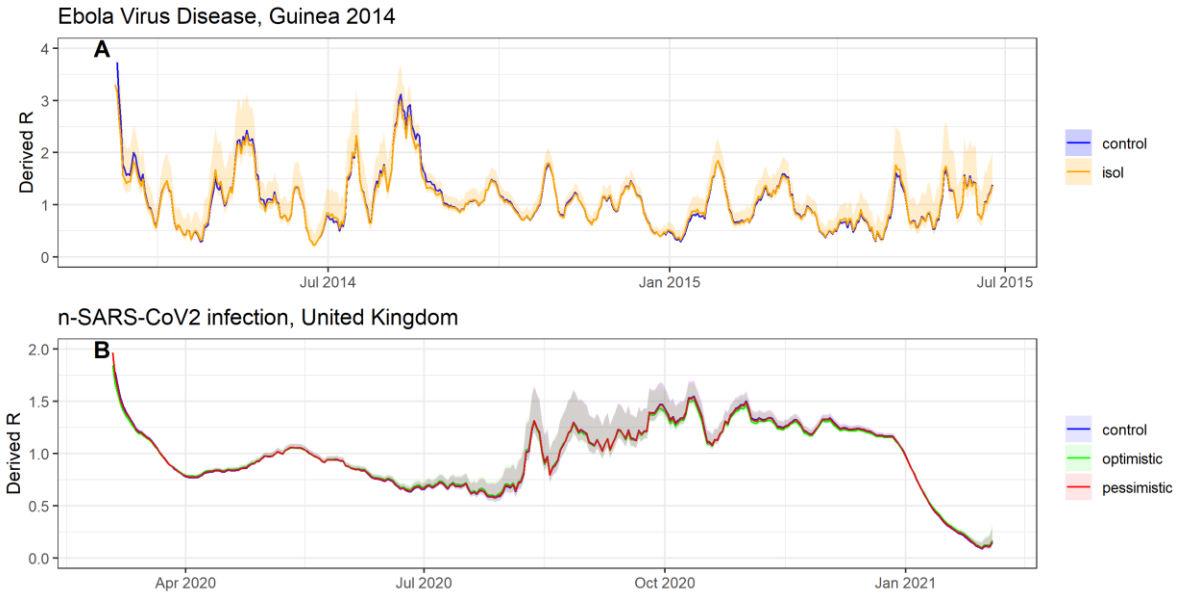
Figure 1: (A-C) Infectivity curves for individuals isolating 25%, 50%, 75% through their infection respectively. (D) Infectivity curve for non-isolating individuals. (E) The relative difference in R inferred by considering a multi-type branching process as compared to a naïve single-type branching process, accounting for isolation of a subset of cases by (i) the amount of infectivity that has passed at the time of isolation (x-axis), (ii) the size of the isolating population (linetype) and (iii) the growth rate of the epidemic (colour). All graphs correspond to a reference generation time distribution which is gamma distributed with a mean of 5.71 days and a standard deviation of 3.09 days.



Case isolation heterogeneity was considered in the context of the Ebola Virus Disease outbreak in Guinea between March 2014 and July 2015 in Figure 2A, and for SARS-CoV-2 between March 2020 and February 2021 in Figure 2B. Both applications confirm that case isolation has limited impact on the overall derived R.

The reason for the limited impact can be seen by considering the equivalent single-type renewal process airing from Equation 10 and Equation 11. The isolating group have a lower infectiousness than the non-isolating group which means their contribution to the weighted generation time distribution is often more muted. Additionally, the resulting weighted generation time distribution undergoes a Laplace transform (Equation 10) which reduces the relative contribution of late transmission in the inference of R – and it is only the late-stage transmission that case isolation has an impact on.

Figure 2: (A) Time varying R values for the Ebola Virus Disease outbreak from March 2014 to July 2015 in Guinea from EpiEstim, derived using an unadjusted generation time distribution (black), and a weighted generation time distribution (equivalent to a multi-type branching process) accounting for case isolation (orange). Credible intervals are shown for the weighted generation time distribution, demonstrating that consideration of case isolation makes little difference to the overall estimated R. The reference generation time distribution is assumed to be gamma distribution with a mean of 15.3 days and a standard deviation of 9.3 days. We assume 36% of cases isolate with isolation occurring 55% of the way into the infectious distribution. (B) Time varying R values for the SARS-CoV-2 outbreak from March 2020 to February 2021 in the United Kingdom, based on optimistic and pessimistic assumptions around isolation. In both cases we assume 63% of infections are symptomatic and the reference generation time distribution (corresponding to non-isolating symptomatic individuals) is gamma distributed with a mean of 5.71 days and a standard deviation of 3.10 days. Optimistic assumptions are that isolation occurs among 75% of symptomatic infection after 30% of infectivity has passed. Pessimistic assumptions are that isolation occurs in 25% of symptomatic infection after 70% of infectivity has passed. In both cases R estimates are based on sliding weekly windows.



Symptomatic and asymptomatic transmission

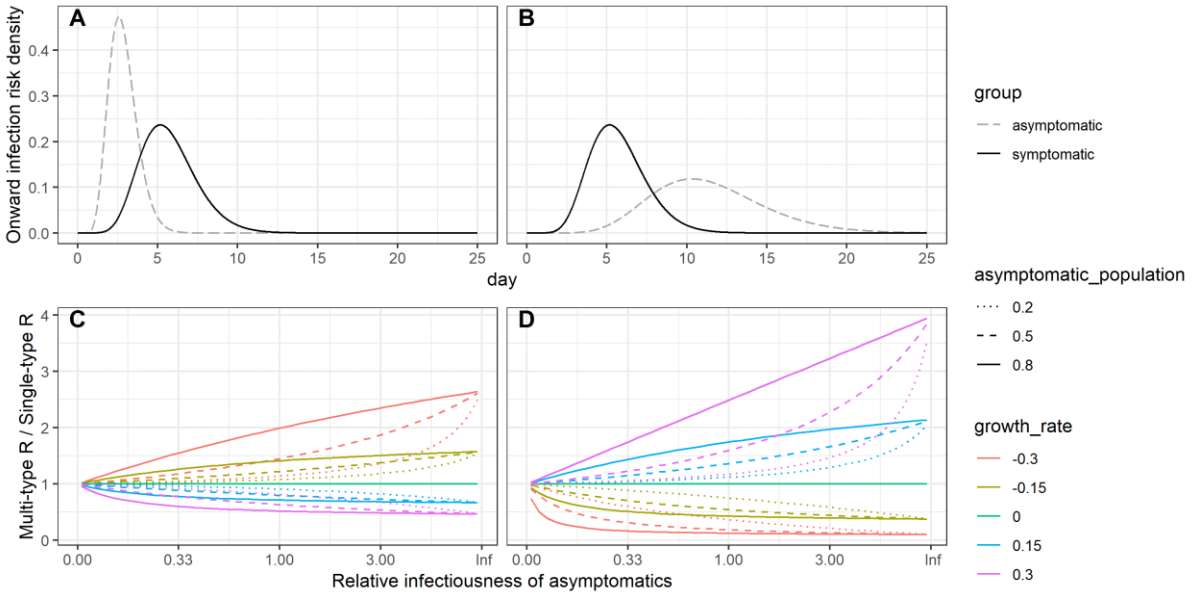
As discussed previously, for a pathogen with symptomatic and asymptomatic infection courses, the generation time distribution is generally only easily obtainable from symptomatic transmission. We therefore compare R estimates from the single-type model using the generation time distribution of the symptomatic individuals with the R estimates from the multi-type model.

If the generation time distribution of asymptomatic carriers is longer than that of symptomatic carriers (Figure 3 - right), the true R derived using a multitype approach will exceed the R derived through the single-type approach in a growing epidemic and will be lower than the R derived through the single-type approach in a declining epidemic. This trend is reversed for dynamics in which the generation time distribution of asymptomatic carriers is shorter than that of symptomatic carriers (Figure 3 -left). The error in inferred R becomes greater at higher absolute values of growth rate, with higher asymptomatic infection rates, and with higher relative infectiousness of asymptomatic individuals (Figure 3C and Figure 3D).

With the extent of variation explored the true R-rate exceeded the R derived from a single-type branching process by up to 4x, when the generation time distribution of asymptomatics was twice as long as that of symptomatics, and asymptomatics were responsible for all onward infection (Figure 3D). While this represents a relatively extreme scenario, it may be relevant for pathogens with early onset of symptoms among symptomatic cases but late onset of infectiousness, by which point

symptomatic individuals may have reduced their contacts substantially, meaning asymptomatic individuals would be more responsible for onward transmission.

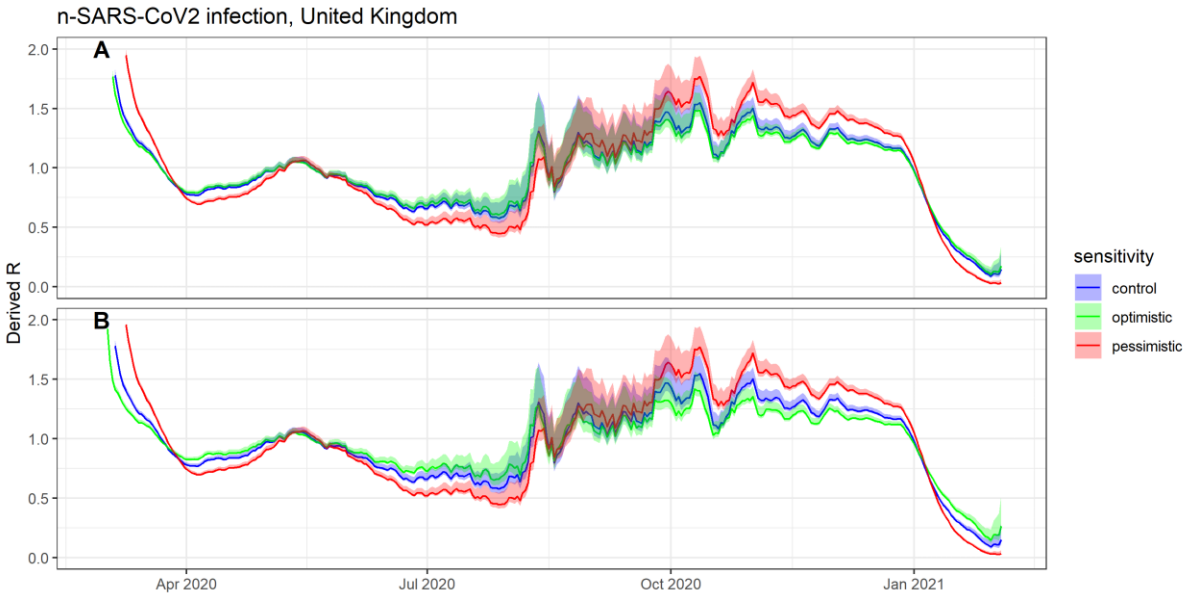
Figure 3: (A-B) Explored generation time distributions for symptomatic and asymptomatic individuals. In (A) the generation time distribution of asymptomatics is half that of symptomatics, whereas in (B) the generation time distribution of symptomatics is half that of asymptomatics. In both cases the symptomatic (reference) generation time distribution has a mean of 5.71 days and a standard deviation of 3.09 days. (C-D) The relative difference in R obtained using a multi-type branching proves vs a naïve single-type branching process, accounting for asymptomatic transmission by (i) the relative infectiousness of asymptomatics (x-axis), (ii) the size of the asymptomatic population (linetype) and (iii) the growth rate of the epidemic (colour). (C) corresponds to the generation time distributions given in (A) while (D) corresponds to those given in (B).



Potential asymptomatic transmission of SARS-CoV-2 in the UK is considered in Figure 4A. A different asymptomatic generation time distribution can result in a substantial difference in the inferred R. We explore an optimistic case, in which asymptomatic transmitters are half as infectious and have half the generation time distribution as symptomatic counterparts, and a pessimistic case, in which asymptomatic transmitters have a prolonged generation time distribution and are twice as infectious as their symptomatic counterparts.

In Figure 4B we consider a three-type branching process consisting of asymptomatic carriers, symptomatic carriers who isolate and symptomatic carriers who do not isolate.

Figure 4: Derived value of R based on UK deaths using a multi-type and naïve branching process for (A) asymptomatic transmission, (B) asymptomatic transmission and symptomatic isolation together. We assume homogeneous mixing for both cases. Optimistic assumptions correspond to asymptomatics with half the generation time distribution of symptomatics and half the infectivity. Pessimistic assumptions correspond to asymptomatics having twice the generation time distribution of symptomatics and twice the infectivity. For case isolation which is included in (B) the optimistic and pessimistic assumptions given in Figure 3 apply. We assume 63% of infections are symptomatic and the reference generation time distribution (corresponding to non-isolating symptomatic individuals) is gamma distributed with a mean of 5.71 days and a standard deviation of 3.10 days. In both cases R estimates are based on sliding weekly windows.



Vaccinated and unvaccinated groups with assortative and disassortative mixing

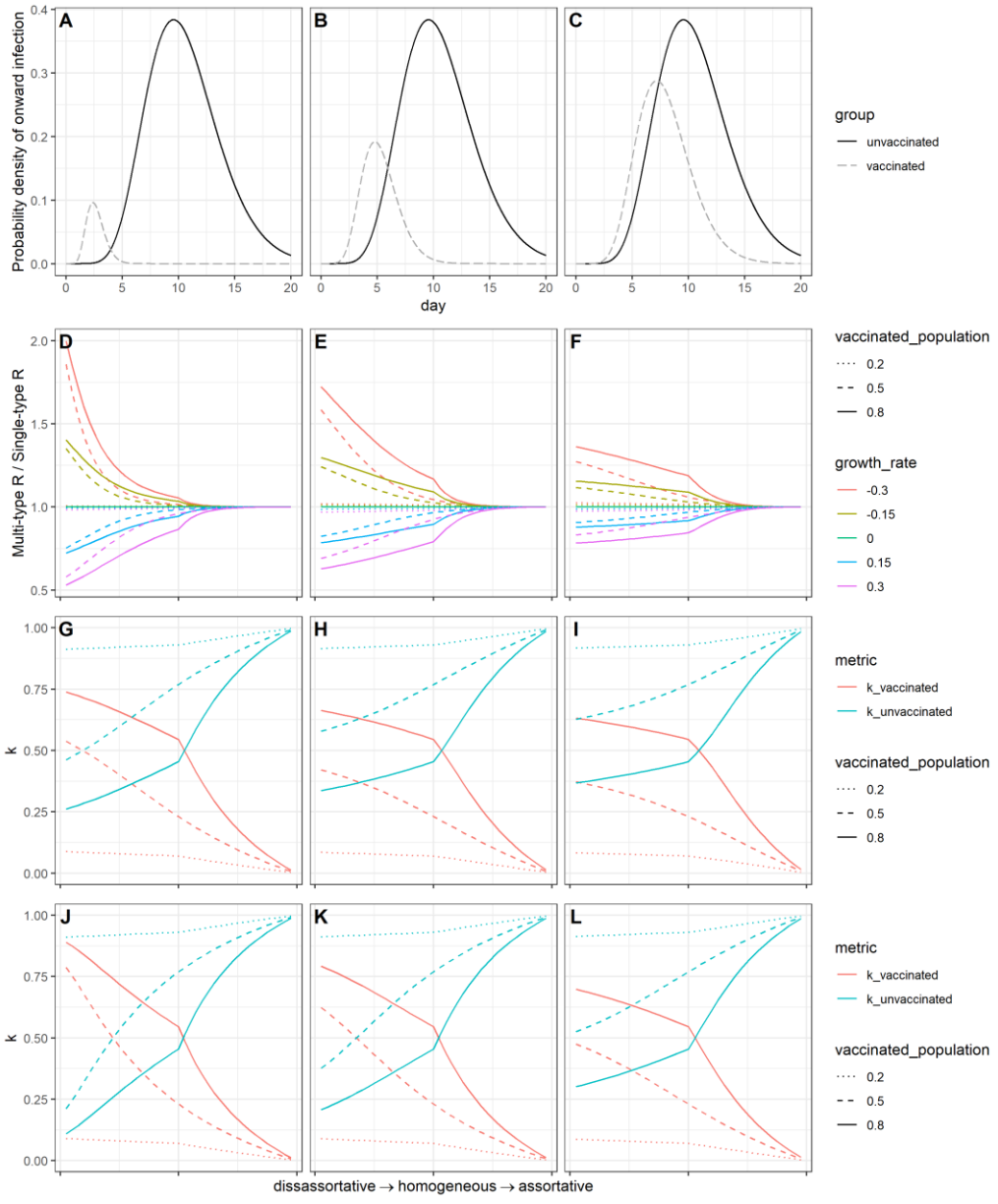
While the impact of vaccination against SARS-CoV-2 on viral load and shedding is as yet unknown, as a vaccine-primed immune response is generally quicker than an unprimed response we would anticipate that both the peak viral load, and the time of viral clearance are reduced. For example, this has been demonstrated with oral and inactivated poliovirus vaccine (28,29). As such, we may expect the effect of SARS-CoV-2 vaccination to be a reduction in both generation time distribution and overall infectiousness. Additionally, data gathered from routine testing of healthcare workers has shown vaccination reduces individuals' susceptibility to SARS-CoV-2 infection. As such, in what follows, we explore scenarios in which vaccination reduces susceptibility to infection, and results in a simultaneous and equal reduction in both the generation time distribution and the infectiousness among the subset of vaccinated individuals who still get infected.

In previous examples we have considered homogeneous mixing between groups. However, vaccination is a clear case in which we may also have to consider the assortativity of mixing. This could be because of vaccination policy such as targeting of age cohorts (30,31) or because of differential vaccine uptake (32,33).

We assume vaccination reduces individuals' susceptibility to infection by 70%. We trial three different reduction factors in the y=x direction for the infectivity curves of vaccinated individuals: 0.25, 0.5, and 0.75, corresponding to transmission reductions of 93.75%, 75%, and 44% from the vaccinated group (Figure 5A-C). The difference in R inferred through a heterogeneous approach vs a simple application of the baseline generation time distribution is highest for disassortative mixing, given with disassortative mixing a significant share of transmission passes through the vaccinated group, whose generation time distribution is not included in the unweighted single-type approach. The difference in inferred R reduces to zero in the limit of totally assortative mixing, as this represents two isolated

outbreaks, for which the epidemic growth rate is totally driven by the unvaccinated group (Figure 5D-F). The difference in inferred R reduces especially quickly as the contribution of the unvaccinated group (for whom the baseline generation time distribution is well characterized) increases. This can be seen by considering the relative sizes of the elements of the eigenvector in Equation 5 corresponding to the proportion of infections that are in the vaccinated and unvaccinated groups (shown in Figure 5G-I for a growing epidemic, and in Figure 5J-L for a shrinking epidemic). The greater the k value corresponding to the unvaccinated group, the closer the inferred R from the multi-type process is to the inferred R from the naïve single-type process.

*Figure 5: (A-C) Explored scenarios for vaccine impact on the infectivity distribution; (D-F) The relative difference in R obtained using a multi-type branching proves vs a naïve single-type branching process with unvaccinated reference group, accounting for vaccinated individuals by (i) the extent of assortativity (x-axis), (ii) the size of the asymptomatic group (linetype) and (iii) the growth rate of the epidemic (colour). We assume that vaccination reduces susceptibility by 70% in all explored scenarios; (G-I) elements of the eigenvector corresponding to the vaccinated and unvaccinated groups in a growing epidemic with r=0.3, normalised such that their sum is 1; (J-L) elements of the eigenvector corresponding to the vaccinated and unvaccinated groups in a declining epidemic with r=-0.3, normalised such that their sum is 1. For graphs D to L, the x-axis is a two-part discontinuous linear scale from disassortative to homogeneous and from homogeneous to assortative (**Error! Reference source not found.**).*



Discussion

In this paper we have shown how and when heterogeneity in the generation time distribution can distort estimates of the reproduction number. While the impact on the inferred reproduction number is limited in the case of symptomatic case isolation, it can be considerable if asymptomatic or vaccinated individuals have particularly different generation time distributions from unvaccinated, symptomatic individuals (for whom the generation time distribution will be best characterized). The difference in inferred R will be smaller for lower growth rates; where the poorly characterized groups represent a small part of the population; or where there is highly assortative mixing between groups.

In using the renewal equation, we assumed infectiousness could be separated into a reproduction number, depending only on calendar time t, and a generation time distribution depending only on time since infection, τ. However, behaviour is likely to change as an epidemic progresses, for instance through a reduction in out-of-household contacts, which may cause the generation time distribution to change independently of case-isolation.

For the multi-group case, we assumed that relative infectiousness and susceptibility between groups remained constant through time. This too is a simplifying assumption: interventions such as the adoption of face coverings may alter the relative susceptibility and infectiousness between groups, especially if there is a correlation between the group and the adoption of certain behaviours. For example, isolation on symptom onset may well be correlated with compliance to mask-wearing and handwashing.

Multiple pathogens have been demonstrated to have both symptomatic and asymptomatic clinical courses, including important epidemic viruses SARS-CoV-2 (34), Influenza (35), Ebola Virus Disease (36) and Middle East Respiratory Syndrome (MERS) (37). Previous work has shown that a difference in the generation time distribution of asymptomatic vs symptomatic carriers of COVID can lead to biased estimates of the effective reproduction number (38).

The infectious profile of an individual depends principally on the extent and duration of viral shedding, and their effective contact rate. Symptom presentation may impact both variables.

Viral shedding will itself depend on individual viral load, and the efficiency and duration of viral expulsion. Viral load studies for influenza infection have shown that asymptomatic and paucisymptomatic cases had 1-2 $\log_{10}$ fewer copies of viral RNA than symptomatic cases, and shorter shedding times (39). Similarly, studies on MERS found the duration of PCR-positivity increased with disease severity (40). Symptoms themselves also increase viral expulsion: a cough can produce an estimated 3,000 droplets, and a sneeze an estimated 40,000 (41); both far more efficient shedding processes than breathing or talking (42).

There have been varying conclusions from studies on the difference in viral load between symptomatic and asymptomatic infections in SARS-CoV-2 infection. Where some studies have found viral load to be similar between symptomatic and asymptomatic SARS-CoV-2 patients (43,44), others have found statistically significant differences in viral load (45,46) and clearance time (47,48), or that shedding duration increases with disease severity (49). A further study in Catalonia has found severity to be positively correlated with viral load, and that higher viral loads led to a greater extent of onward transmission (50). A recent literature review including 79 studies on SARS-CoV-2 concluded that the sum of evidence suggests viral load is similar between symptomatic and asymptomatic individuals, most studies 'demonstrate faster viral clearance among asymptomatic than those who are symptomatic' (51). A further systematic review of the reproduction number and secondary attack rate

suggested asymptomatic cases were around one seventh as infectious as symptomatic individuals (53).

Conversely, symptomatic SARS-CoV-2 infecteds are likely to reduce their contacts following onset: in the UK, isolation of 10 days is mandated for individuals developing symptoms (and subsequently receiving a positive test for) of COVID-19, and for their households (52).

Vaccination often has an important impact on viral load and/or infectiousness over time, for instance with oral poliovirus vaccine (OPV) and inactivated poliovirus vaccine (IPV) (28,29). The importance of understanding the generation time distribution of multiple groups becomes increasingly important with disassortative mixing. This may be particularly important when estimating the reproduction number of sexually transmitted infections in heterosexual contact networks, for instance with HPV, for which vaccination uptake was previously limited to females.

As vaccines against SARS-CoV-2 are rolled out over the coming months, understanding the impact of the vaccine on the susceptibility and infectious profile will be increasingly important for accurate inference of R. Given the vaccine schedule is broadly age-prioritised, mixing by vaccination status will be more assortative.

Estimating the contemporaneous generation time distribution should be regarded as similarly important to estimation of the reproduction number itself, which currently occupies the work of academic modelling groups worldwide for SARS-CoV-2. Better capturing the heterogeneities of the generation time distribution will become increasingly important as vaccination is rolled out, as well as with the emergence of new strains which may exhibit different infectious profiles. Upcoming SARS-CoV-2 challenge trials in the UK should enable detailed analysis of viral load profiles and symptomatic rates, which can inform updated generation time distributions (52).

While estimation of the generation time distribution is necessarily a time-consuming endeavour, testing systems should integrate additional epidemiological information in tandem with their test and trace protocols. Updated estimates of the serial interval could be obtained by requiring test applicants to supply their symptom onset date, with linkage to traced contacts should they also enter the testing system. For a more direct means to estimate changes in the generation time distribution, or indeed the incubation period, individuals could be asked for dates of contact with known infected in the previous week, and this too linked with contacts who enter the test and trace system.

Funding

This study is partially funded by the National Institute for Health Research (NIHR) Health Protection Research Unit in Modelling and Health Economics, a partnership between Public Health England, Imperial College London and LSHTM (grant code NIHR200908); and acknowledges funding from the MRC Centre for Global Infectious Disease Analysis (reference MR/R015600/1), jointly funded by the UK Medical Research Council (MRC) and the UK Foreign, Commonwealth & Development Office (FCDO), under the MRC/FCDO Concordat agreement and is also part of the EDCTP2 programme supported by the European Union; and acknowledges funding from the Abdul Latif Jameel Institute for Disease and Emergency Analytics jointly funded by Community Jameel.

Disclaimer: "The views expressed are those of the author(s) and not necessarily those of the NIHR, Public Health England or the Department of Health and Social Care."

15

References

1. COVID-19 Map - Johns Hopkins Coronavirus Resource Center [Internet]. [cited 2021 Mar 10]. Available from: https://coronavirus.jhu.edu/map.html

2. Billah MA, Miah MM, Khan MN. Reproductive number of coronavirus: A systematic review and meta-analysis based on global level evidence. Flacco ME, editor. PLoS One [Internet]. 2020 Nov 11 [cited 2021 Feb 10];15(11):e0242128. Available from: https://dx.plos.org/10.1371/journal.pone.0242128

3. Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. Am J Epidemiol [Internet]. 2004 Sep 15 [cited 2019 Oct 30];160(6):509–16. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15353409

4. Thompson RN, Stockwin JE, van Gaalen RD, Polonsky JA, Kamvar ZN, Demarsh PA, et al. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. Epidemics. 2019;

5. Cori A, Ferguson NM, Fraser C, Cauchemez S. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. Am J Epidemiol [Internet]. 2013 Nov 1 [cited 2021 Feb 8];178(9):1505–12. Available from: https://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kwt133

6. Riley S, Fraser C, Donnelly CA, Ghani AC, Abu-Raddad LJ, Hedley AJ, et al. Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions. Science (80-) [Internet]. 2003 Jun 20 [cited 2021 Feb 8];300(5627):1961–6. Available from: http://science.sciencemag.org/

7. Fraser C. Estimating Individual and Household Reproduction Numbers in an Emerging Epidemic. Galvani A, editor. PLoS One [Internet]. 2007 Aug 22 [cited 2020 Nov 24];2(8):e758. Available from: https://dx.plos.org/10.1371/journal.pone.0000758

8. The R value and growth rate in the UK - GOV.UK [Internet]. [cited 2021 Feb 10]. Available from: https://www.gov.uk/guidance/the-r-number-in-the-uk

9. Lehtinen S, Ashcroft P, Bonhoe↵er S. On the relationship between serial interval, infectiousness profile and generation time. [cited 2020 Nov 24]; Available from: https://doi.org/10.1101/2020.09.18.20197210

10. Britton T, Tomba GS. Estimation in emerging epidemics: Biases and remedies. J R Soc Interface [Internet]. 2019 Jan 1 [cited 2020 Nov 24];16(150). Available from: /pmc/articles/PMC6364646/?report=abstract

11. Cheng HY, Jian SW, Liu DP, Ng TC, Huang WT, Lin HH. Contact Tracing Assessment of COVID-19 Transmission Dynamics in Taiwan and Risk at Different Exposure Periods before and after Symptom Onset. JAMA Intern Med [Internet]. 2020 Sep 1 [cited 2021 Mar 30];180(9):1156–63. Available from: https://jamanetwork.com/

12. Du Z, Xu X, Wu Y, Wang L, Cowling BJ, Meyers LA. Serial interval of COVID-19 among publicly reported confirmed cases [Internet]. Vol. 26, Emerging Infectious Diseases. Centers for Disease Control and Prevention (CDC); 2020 [cited 2021 Mar 30]. p. 1341–3. Available from: /pmc/articles/PMC7258488/

13. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. N Engl J Med [Internet]. 2020 Mar 26 [cited 2021 Mar 30];382(13):1199–207. Available from: http://www.nejm.org/doi/10.1056/NEJMoa2001316

16

14. Ma S, Zhang J, Zeng M, Yun Q, Guo W, Zheng Y, et al. Epidemiological parameters of coronavirus disease 2019: A pooled analysis of publicly reported individual data of 1155 cases from seven countries [Internet]. medRxiv. medRxiv; 2020 [cited 2021 Mar 30]. p. 2020.03.21.20040329. Available from: https://doi.org/10.1101/2020.03.21.20040329

15. Bi Q, Wu Y, Mei S, Ye C, Zou X, Zhang Z, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. Lancet Infect Dis. 2020 Aug 1;20(8):911–9.

16. Fraser C. Estimating individual and household reproduction numbers in an emerging epidemic. PLoS One. 2007 Aug 22;2(8).

17. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. Proc R Soc B Biol Sci [Internet]. 2007 Feb 22 [cited 2020 Nov 24];274(1609):599–604. Available from: /pmc/articles/PMC1766383/?report=abstract

18. Sunetra Gupta, Roy M. Anderson, Robert M. May. Networks of sexual contacts: implications for the pattern of spread of HIV | Ovid [Internet]. 1989 [cited 2021 Mar 5]. Available from: https://oce.ovid.com/article/00002030-198912000-00005/HTML

19. Garske T, Cori A, Ariyarajah A, Blake IM, Dorigatti I, Eckmanns T, et al. Heterogeneities in the case fatality ratio in the West African Ebola outbreak 2013–2016. Philos Trans R Soc B Biol Sci [Internet]. 2017 May 26 [cited 2019 Jul 16];372(1721):20160308. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28396479

20. Chowell G, Nishiura H. Transmission dynamics and control of Ebola virus disease (EVD): A review [Internet]. Vol. 12, BMC Medicine. BioMed Central Ltd.; 2014 [cited 2021 Feb 23]. p. 1–16. Available from: /pmc/articles/PMC4207625/

21. Faye O, Boëlle PY, Heleze E, Faye O, Loucoubar C, Magassouba N, et al. Chains of transmission and control of Ebola virus disease in Conakry, Guinea, in 2014: An observational study. Lancet Infect Dis [Internet]. 2015 Mar 1 [cited 2021 Feb 23];15(3):320–6. Available from: http://dx.doi.org/10.1016/

22. Challen R, Brooks-Pollock E, Tsaneva-Atanasova K, Danon L. Meta-analysis of the SARS-CoV-2 serial interval and the impact of parameter uncertainty on the COVID-19 reproduction number [Internet]. medRxiv. medRxiv; 2020 [cited 2021 Mar 8]. p. 2020.11.17.20231548. Available from: https://doi.org/10.1101/2020.11.17.20231548

23. Linton NM, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov AR, Jung S-M, et al. Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data. [cited 2021 Mar 8]; Available from: https://doi.org/10.1101/2020.01.26.20018754

24. Byrne AW, Mcevoy D, Collins AB, Hunt K, Casey M, Barber A, et al. Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. BMJ Open [Internet]. 2020 [cited 2021 Mar 8];10:39856. Available from: http://bmjopen.bmj.com/

25. Abbott S, Hellewell J, Thompson RN, Sherratt K, Gibbs HP, Bosse NI, et al. Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. Wellcome Open Res [Internet]. 2020 Dec 8 [cited 2021 May 14];5:112. Available from: https://doi.org/10.12688/wellcomeopenres.16006.1

26. PHE & NHSX. Daily summary: Coronavirus in the UK [Internet]. https://coronavirus.data.gov.uk/. 2020 [cited 2021 Feb 23]. Available from:

17

https://coronavirus.data.gov.uk/

27. Hu S, Wang W, Wang Y, Litvinova M, Luo K, Ren L, et al. Infectivity, susceptibility, and risk factors associated with SARS-CoV-2 transmission under intensive contact tracing in Hunan, China [Internet]. medRxiv. medRxiv; 2020 [cited 2021 Feb 8]. p. 2020.07.23.20160317. Available from: https://doi.org/10.1101/2020.07.23.20160317

28. Laassri M, Lottenbach K, Belshe R, Wolff M, Rennels M, Plotkin S, et al. Effect of different vaccination schedules on excretion of oral poliovirus vaccine strains. J Infect Dis [Internet]. 2005 Dec 15 [cited 2021 Jan 20];192(12):2092–8. Available from: https://academic.oup.com/jid/article-lookup/doi/10.1086/498172

29. Alexander JP, Gary HE, Pallansch MA. Duration of poliovirus excretion and its implications for acute flaccid paralysis surveillance: A review of the literature. J Infect Dis [Internet]. 1997 Feb 1 [cited 2021 Jan 25];175(2 SUPPL.):8176–82. Available from: https://academic.oup.com/jid/article/175/Supplement_1/S176/878744

30. Meehan MT, Cocks DG, Caldwell JM, Trauer JM, Adekunle AI, Ragonnet RR, et al. Age-targeted dose allocation can halve COVID-19 vaccine requirements [Internet]. medRxiv. medRxiv; 2020 [cited 2021 Feb 9]. p. 2020.10.08.20208108. Available from: https://doi.org/10.1101/2020.10.08.20208108

31. Joint Committee on Vaccination and Immunisation: advice on priority groups for COVID-19 vaccination. 2020.

32. Robertson E, Reeve KS, Niedzwiedz CL, Moore J, Blake M, Green M, et al. Predictors of COVID-19 vaccine hesitancy in the UK Household Longitudinal Study. medRxiv [Internet]. 2021 Jan 2 [cited 2021 Feb 9];2020.12.27.20248899. Available from: https://doi.org/10.1101/2020.12.27.20248899

33. Barclay VC, Smieszek T, He J, Cao G, Rainey JJ, Gao H, et al. Positive Network Assortativity of Influenza Vaccination at a High School: Implications for Outbreak Risk and Herd Immunity. McVernon J, editor. PLoS One [Internet]. 2014 Feb 5 [cited 2021 Feb 9];9(2):e87042. Available from: https://dx.plos.org/10.1371/journal.pone.0087042

34. Ward H, Atchison CJ, Whitaker M, Ainslie KEC, Elliot J, Okell LC, et al. Antibody prevalence for SARS-CoV-2 in England following first peak of the pandemic: REACT2 study in 100,000 adults. medRxiv [Internet]. 2020 Aug 14 [cited 2021 Jan 20];2020.08.12.20173690. Available from: https://doi.org/10.1101/2020.08.12.20173690

35. Leung NHL, Xu C, Ip DKM, Cowling BJ. The fraction of influenza virus infections that are asymptomatic: A systematic review and meta-analysis [Internet]. Vol. 26, Epidemiology. Lippincott Williams and Wilkins; 2015 [cited 2021 Jan 20]. p. 862–72. Available from: /pmc/articles/PMC4586318/?report=abstract

36. Dean NE, Halloran ME, Yang Y, Longini IM. Transmissibility and pathogenicity of Ebola virus: A systematic review and meta-analysis of household secondary attack rate and asymptomatic infection. Clin Infect Dis [Internet]. 2016 May 15 [cited 2021 Jan 20];62(10):1277–86. Available from: /pmc/articles/PMC4845791/?report=abstract

37. Al-Tawfiq JA, Gautret P. Asymptomatic Middle East Respiratory Syndrome Coronavirus (MERS-CoV) infection: Extent and implications for infection control: A systematic review [Internet]. Vol. 27, Travel Medicine and Infectious Disease. Elsevier USA; 2019 [cited 2021 Jan 20]. p. 27–32. Available from: /pmc/articles/PMC7110966/?report=abstract

38. Park SW, Cornforth DM, Dushoff J, Weitz JS. The time scale of asymptomatic transmission

affects estimates of epidemic potential in the COVID-19 outbreak. Epidemics. 2020 Jun 1;31:100392.

39. Ip DKM, Lau LLH, Leung NHL, Fang VJ, Chan KH, Chu DKW, et al. Viral Shedding and Transmission Potential of Asymptomatic and Paucisymptomatic Influenza Virus Infections in the Community. Clin Infect Dis. 2017 Mar 15;64(6):736–42.

40. Al Hosani FI, Pringle K, Al Mulla M, Kim L, Pham H, Alami NN, et al. Response to emergence of middle east respiratory syndrome coronavirus, Abu Dhabi, United Arab Emirates, 2013-2014. Emerg Infect Dis [Internet]. 2016 Jul 1 [cited 2021 Jan 25];22(7):1162–8. Available from: /pmc/articles/PMC4918155/?report=abstract

41. Cole EC, Cook CE. Characterization of infectious aerosols in health care facilities: An aid to effective engineering controls and preventive strategies. Am J Infect Control [Internet]. 1998 [cited 2021 Jan 26];26(4):453–64. Available from: /pmc/articles/PMC7132666/?report=abstract

42. DUGUID JP. The numbers and the sites of origin of the droplets expelled during expiratory activities. Edinb Med J [Internet]. 1945 Nov 1 [cited 2021 Jan 26];52(11):385–401. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5286249/

43. Walsh KA, Jordan K, Clyne B, Rohde D, Drummond L, Byrne P, et al. SARS-CoV-2 detection, viral load and infectivity over the course of an infection [Internet]. Vol. 81, Journal of Infection. W.B. Saunders Ltd; 2020 [cited 2021 Jan 25]. p. 357–71. Available from: https://doi.org/10.1016/j.jinf.2020.06.067

44. Ra SH, Lim JS, Kim GU, Kim MJ, Jung J, Kim SH. Upper respiratory viral load in asymptomatic individuals and mildly symptomatic patients with SARS-CoV-2 infection. Thorax [Internet]. 2021 Jan 1 [cited 2021 Jan 25];76(1):61–3. Available from: https://www.

45. Zhou R, Li F, Chen F, Liu H, Zheng J, Lei C, et al. Viral dynamics in asymptomatic patients with COVID-19. Int J Infect Dis [Internet]. 2020 Jul 1 [cited 2021 Jan 25];96:288–90. Available from: https://doi.org/10.1016/j.ijid.2020.05.030.

46. Lavezzo E, Franchin E, Ciavarella C, Cuomo-Dannenburg G, Barzon L, Vecchio C Del, et al. Suppression of COVID-19 outbreak in the municipality of Vo, Italy. medRxiv. 2020 Apr 18;2020.04.17.20053157.

47. Van Vinh Chau N, Lam VT, Dung NT, Yen LM, Minh NNQ, Hung LM, et al. The Natural History and Transmission Potential of Asymptomatic Severe Acute Respiratory Syndrome Coronavirus 2 Infection. Clin Infect Dis [Internet]. 2020 Dec 17 [cited 2021 Jan 25];71(10):2679–87. Available from: https://academic.oup.com/cid/article/71/10/2679/5851471

48. Yang R, Gui X, Xiong Y. Comparison of Clinical Characteristics of Patients with Asymptomatic vs Symptomatic Coronavirus Disease 2019 in Wuhan, China. JAMA Netw open [Internet]. 2020 May 1 [cited 2021 Jan 25];3(5):e2010182. Available from: https://jamanetwork.com/

49. Chen X, Zhu B, Hong W, Zeng J, He X, Chen J, et al. Associations of clinical characteristics and treatment regimens with the duration of viral RNA shedding in patients with COVID-19. Int J Infect Dis [Internet]. 2020 Sep 1 [cited 2021 Jan 25];98:252–60. Available from: https://doi.org/10.1016/j.ijid.2020.06.091

50. New legal duty to self-isolate comes into force today - GOV.UK [Internet]. [cited 2021 Feb 10]. Available from: https://www.gov.uk/government/news/new-legal-duty-to-self-isolate-comes-into-force-today

51. Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A. SARS-CoV-2, SARS-CoV, and MERS-

CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. The Lancet Microbe. 2021 Jan 1;2(1):e13–22.

52. Stay at home: guidance for households with possible or confirmed coronavirus (COVID-19) infection - GOV.UK [Internet]. [cited 2021 Jan 26]. Available from: https://www.gov.uk/government/publications/covid-19-stay-at-home-guidance/stay-at-home-guidance-for-households-with-possible-coronavirus-covid-19-infection

53. Thompson HA, Mousa A, Dighe A, Fu H, Arnedo-Pena A, Barrett P, et al. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Setting-specific Transmission Rates: A Systematic Review and Meta-analysis. Clin Infect Dis [Internet]. 2021 Feb 9 [cited 2021 May 14]; Available from: https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciab100/6131730

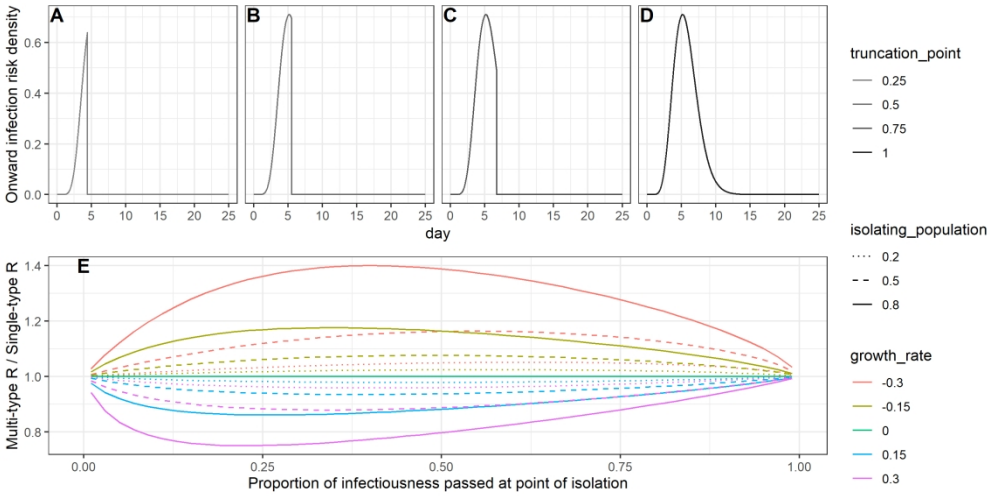


Figure 1: (A-C) Infectivity curves for individuals isolating 25%, 50%, 75% through their infection respectively. (D) Infectivity curve for non-isolating individuals. (E) The relative difference in R inferred by considering a multi-type branching process as compared to a naïve single-type branching process, accounting for isolation of a subset of cases by (i) the amount of infectivity that has passed at the time of isolation (x-axis), (ii) the size of the isolating population (linetype) and (iii) the growth rate of the epidemic (colour). All graphs correspond to a reference generation time distribution which is gamma distributed with a mean of 5.71 days and a standard deviation of 3.09 days.

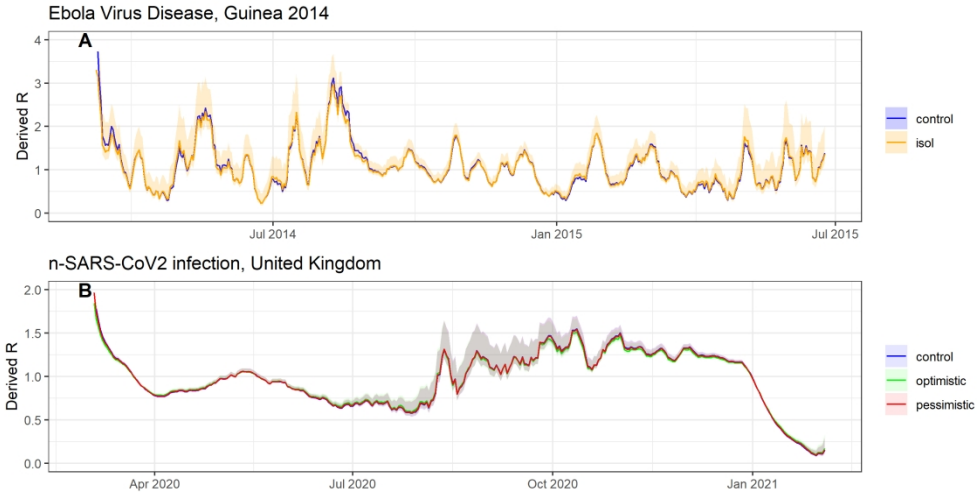


Figure 2: (A) Time varying R values for the Ebola Virus Disease outbreak from March 2014 to July 2015 in Guinea from EpiEstim, derived using an unadjusted generation time distribution (black), and a weighted generation time distribution (equivalent to a multi-type branching process) accounting for case isolation (orange). Credible intervals are shown for the weighted generation time distribution, demonstrating that consideration of case isolation makes little difference to the overall estimated R. The reference generation time distribution is assumed to be gamma distribution with a mean of 15.3 days and a standard deviation of 9.3 days. We assume 36% of cases isolate with isolation occurring 55% of the way into the infectious distribution. (B) Time varying R values for the SARS-CoV-2 outbreak from March 2020 to February 2021 in the United Kingdom, based on optimistic and pessimistic assumptions around isolation. In both cases we assume 63% of infections are symptomatic and the reference generation time distribution (corresponding to non-isolating symptomatic individuals) is gamma distributed with a mean of 5.71 days and a standard deviation of 3.10 days. Optimistic assumptions are that isolation occurs among 75% of symptomatic infection after 30% of infectivity has passed. Pessimistic assumptions are that isolation occurs in 25% of symptomatic infection after 70% of infectivity has passed. In both cases R estimates are based on sliding weekly windows.

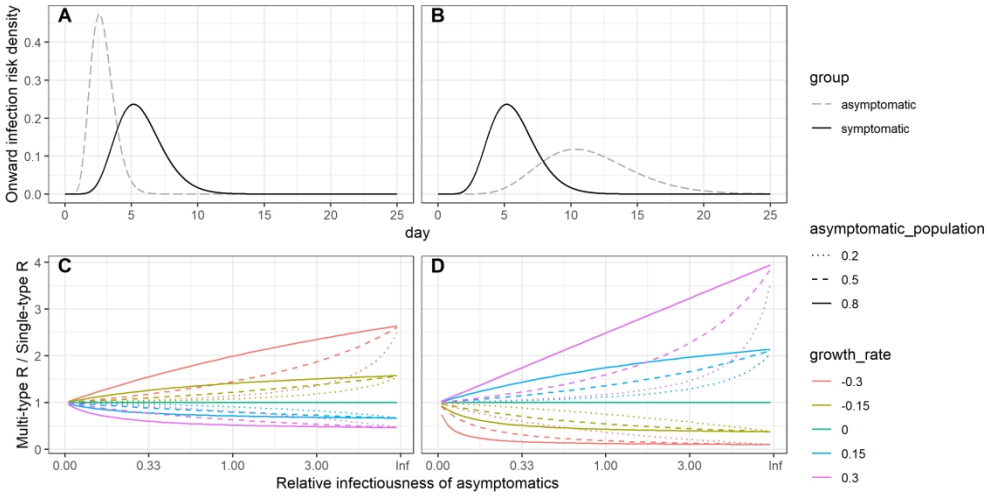


Figure 3: (A-B) Explored generation time distributions for symptomatic and asymptomatic individuals. In (A) the generation time distribution of asymptomatics is half that of symptomatics, whereas in (B) the generation time distribution of symptomatics is half that of asymptomatics. In both cases the symptomatic (reference) generation time distribution has a mean of 5.71 days and a standard deviation of 3.09 days. (C-D) The relative difference in R obtained using a multi-type branching proves vs a naïve single-type branching process, accounting for asymptomatic transmission by (i) the relative infectiousness of asymptomatics (x-axis), (ii) the size of the asymptomatic population (linetype) and (iii) the growth rate of the epidemic (colour). (C) corresponds to the generation time distributions given in (A) while (D) corresponds to those given in (B).

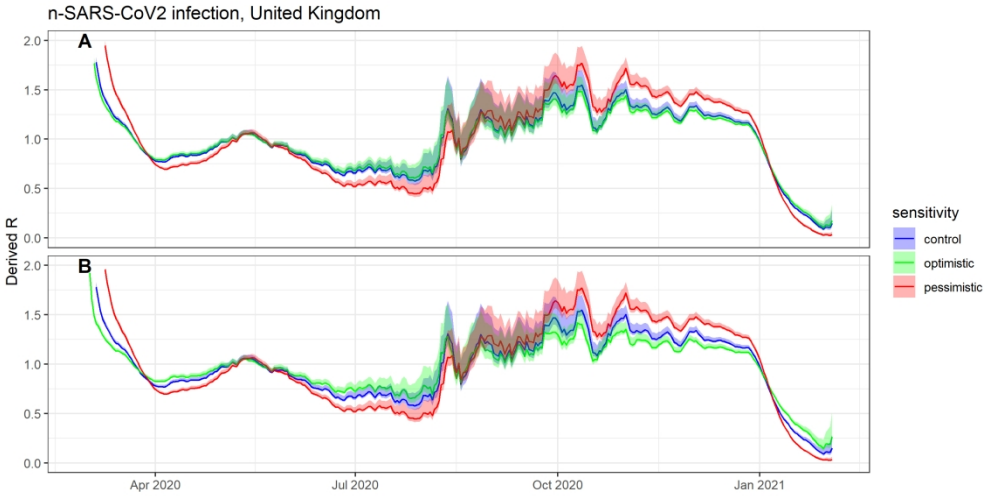


Figure 4: Derived value of R based on UK deaths using a multi-type and naïve branching process for (A) asymptomatic transmission, (B) asymptomatic transmission and symptomatic isolation together. We assume homogeneous mixing for both cases. Optimistic assumptions correspond to asymptomatics with half the generation time distribution of symptomatics and half the infectivity. Pessimistic assumptions correspond to asymptomatics having twice the generation time distribution of symptomatics and twice the infectivity. For case isolation which is included in (B) the optimistic and pessimistic assumptions given in Figure 3 apply. We assume 63% of infections are symptomatic and the reference generation time distribution (corresponding to non-isolating symptomatic individuals) is gamma distributed with a mean of 5.71 days and a standard deviation of 3.10 days. In both cases R estimates are based on sliding weekly windows.

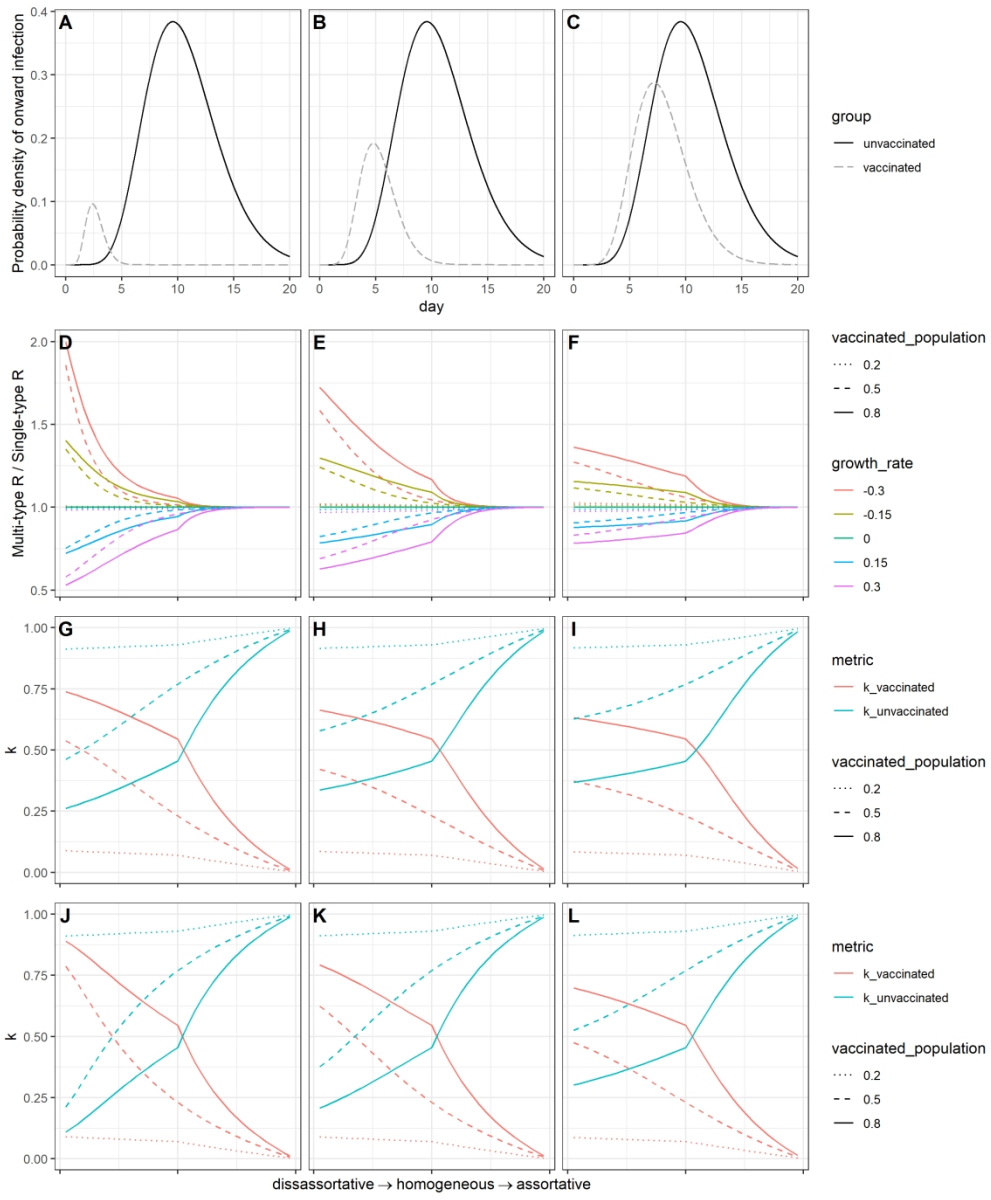


Figure 5: (A-C) Explored scenarios for vaccine impact on the infectivity distribution; (D-F) The relative difference in R obtained using a multi-type branching proves vs a naïve single-type branching process with unvaccinated reference group, accounting for vaccinated individuals by (i) the extent of assortativity (x-axis), (ii) the size of the asymptomatic group (linetype) and (iii) the growth rate of the epidemic (colour). We assume that vaccination reduces susceptibility by 70% in all explored scenarios; (G-I) elements of the eigenvector corresponding to the vaccinated and unvaccinated groups in a growing epidemic with r=0.3, normalised such that their sum is 1; (J-L) elements of the eigenvector corresponding to the vaccinated and unvaccinated groups in a declining epidemic with r=-0.3, normalised such that their sum is 1. For graphs D to L, the x-axis is a two-part discontinuous linear scale from disassortative to homogeneous and from homogeneous to assortative (Equation 12).