Inferring the reproduction number in heterogeneous epidemics

Background

- A simple and widely used approach to estimate the effective reproduction number is via the renewal equation, in which all infected individuals are assumed to have the same infectious profile over time
- However, variation in infectious profiles may results from biological and behavioral differences between individuals

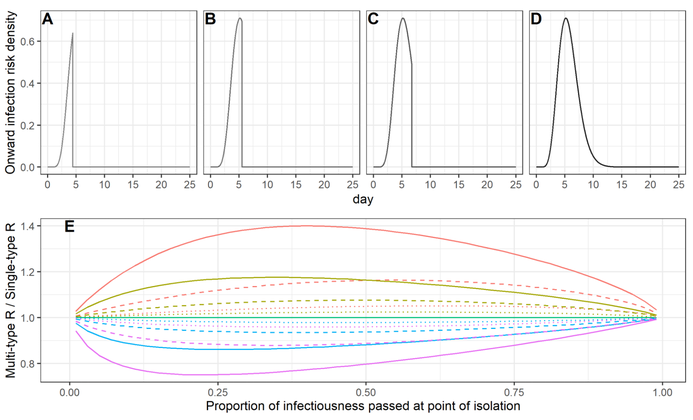
Symptomatic vs asymptomatic

Isolation vs Non-isolation

Vaccinated vs Non-vaccinated

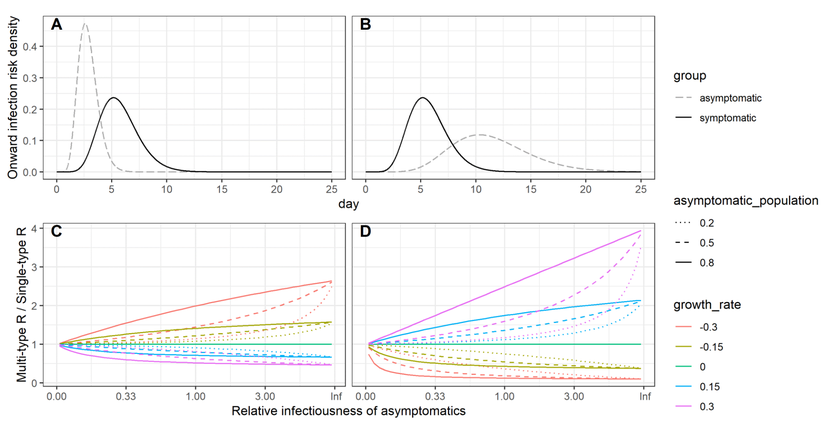
- In a novel outbreak, initial estimates of the generation time distribution are typically based on analysis of the "first few hundred cases" which are typically representative of symptomatic transmission only
- In this paper, we derive a multi-type equivalent of the renewal equation which accounts for heterogeneity in infectious profiles, and explore how much the corresponding estimated R differs from a "naïve" R derived from a single-type branching process
- We consider two applications to illustrate the potential impact on R estimates of neglecting heterogeneities: Ebola Virus Disease in Guinea in 2014-15, and to SARS-CoV-2 in the UK between March 2020 and January 2021

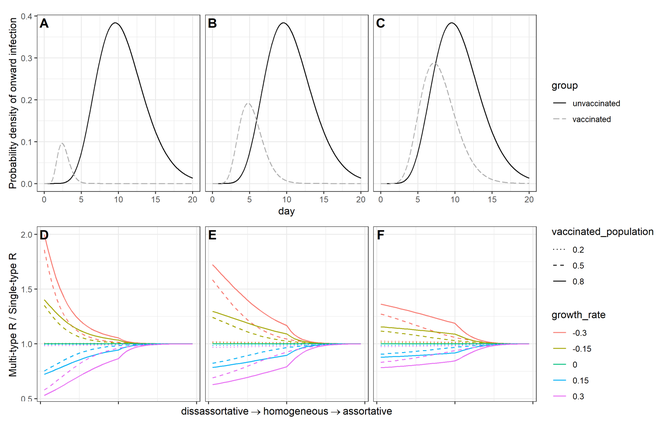
Results



- Isolation of symptomatic cases on onset will necessarily reduce the generation time distribution
- 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

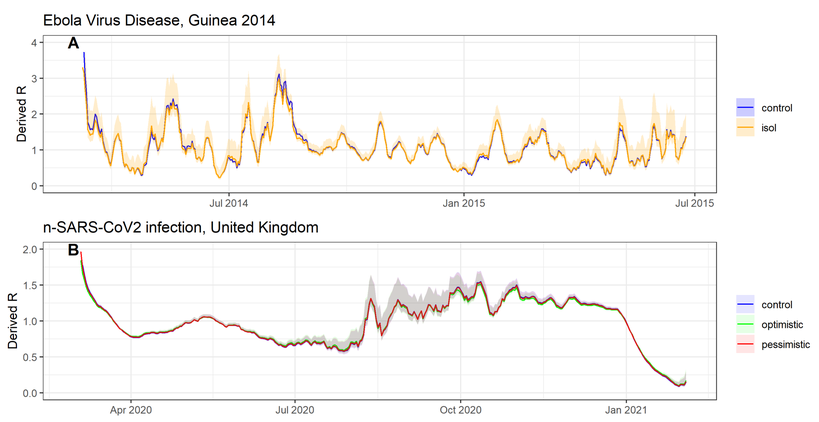
- If the generation time distribution of asymptomatic carriers is longer than that of symptomatic carriers, the true R derived using a multitype approach will exceed the R derived through the single-type approach in a growing epidemic

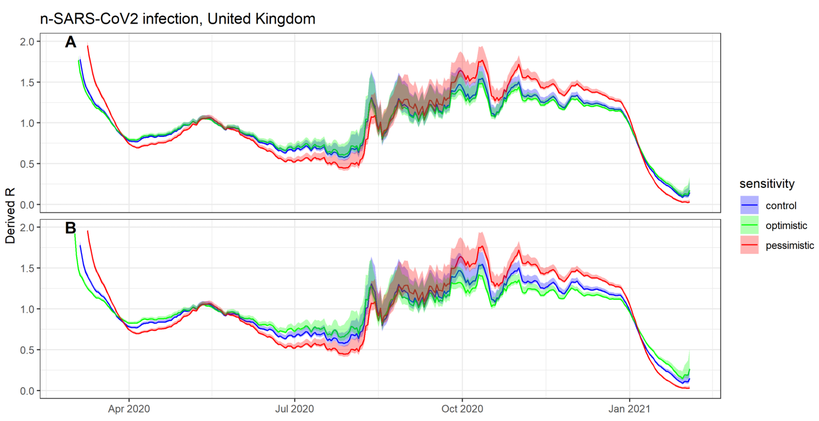




- If the generation time distribution of asymptomatic carriers is longer than that of symptomatic carriers, the true R derived using a multitype approach will exceed the R derived through the single-type approach in a growing epidemic

- The difference in inferred R is small when considering symptomatic isolation and applied to Ebola Virus Disease in Guinea or SARS-CoV-2 in the UK





- There is a larger difference between estimates of R in the case of a different generation time distribution for asymptomatic individuals, as can be seen here when the generation time distribution is assumed to be half- or twice-as long

Methods

- The renewal equation links the incidence of infection at time *t* with the reproduction number *R(t) and* the generation time distribution *ω(τ)*

$$I(t) = R(t)\int_0^\infty \omega(\tau)I(t-\tau)d\tau$$

- With heterogeneity in transmission, this can be adapted to a matrix equation

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

- We assume a solution of the form below

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

- This yields an eigenvalue equation

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

- The reproduction number will be the dominant eigenvalue of the next generation matrix constructed with the elements $R_{i\to j}$. As such, we can factorise out R, leaving matrix elements M which represent the relative risk of transmission between groups i and j.

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

- The reproduction number will be the dominant eigenvalue of the next generation matrix constructed with the elements $R_{i\to j}$. As such, we can factorise out R, leaving matrix elements M which represent the relative risk of transmission between groups i and j.

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

$$= R(t)\int_0^\infty \tilde{\omega}(\tau)e^{r(t-\tau)}d\tau \quad where \quad \tilde{\omega}(\tau) = \sum_{i,j} k_j M_{j\to i}\omega_j(\tau) = \sum_j \left(k_j\omega_j(\tau)\sum_i M_{j\to i} \right)$$

- We parameterize assortativity in a two-part linear scale between 0 and p_1 and p_1 and 1 using the following form

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

Discussion

- Heterogeneity in the generation time distribution can distort estimates of the reproduction number
- The impact is small in the case of heterogeneity caused by symptomatic case isolation, but can be high if asymptomatic or vaccinated individuals have particularly different generation time distributions from the well characterized group
- In using the renewal equation, we assumed the generation time distribution depends only on time since infection, τ. This is open to challenge: behavior is likely to change as an epidemic progresses, for instance through a reduction in out-of-household contacts
- For the multi-group case, we also assumed that relative infectiousness and susceptibility between groups remained constant through time. However, interventions that reduce susceptibility and infectiousness may have differential uptake between groups (e.g. compliance to isolation on symptom onset may be correlated with compliance of the group to mask-wearing and handwashing)
- As vaccines against SARS-CoV-2 is rolled out, understanding the impact of the vaccine and of variants on the susceptibility and infectious profile will be increasingly important for an accurate inference of R
- Given the vaccine priority schedule broadly follows an age-based approach, mixing between the vaccinated and unvaccinated groups 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
- While estimation of the generation time distribution is necessarily a time-consuming endeavor, 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.

1) Eisele TP et al. Short-term Impact of Mass Drug Administration With Dihydroartemisinin Plus Piperaquine on Malaria in Southern Province Zambia: A Cluster-Randomized Controlled Trial. J Infect Dis 2016 Dec 15, 214(12):1831–9
2) Huber JH et al. Quantitative, model-based estimates of variability in the generation and serial intervals of Plasmodium falciparum malaria. Malaria Journal 2016 Sep 22 15(1):490