Appendix:
Unravelling Transmission in Epidemiological Models and its Role in the Disease-Diversity Relationship.

Marjolein E.M. Toorians[1,*], Ailene MacPherson[2], and T. Jonathan Davies[1,3,4]

[1]Department of Botany, Biodiversity Research Centre, University of British Columbia, 2212 Main Mall, Vancouver, BC V6T 1Z4, Canada
[2]Department of Mathematics, Simon Fraser University, Burnaby, Canada
[3]African Centre for DNA Barcoding, University of Johannesburg, Johannesburg 2092, South Africa
[4]Department Forest & Conservation Sciences, University of British Columbia, 2212 Main Mall, Vancouver, BC V6T 1Z4, Canada
[*]Corresponding author

December 2021

1 Appendix

1.1 Derivation transmission, $\beta = \kappa \cdot c$

From the original equation for transmission, from Keeling and Rohani (2011) :

$$\beta = -\kappa \cdot \ln(1-c) \tag{1}$$

assuming probability of successful transmission after contact ($c \to 0$), we can perform a second order Taylor expansion of κ around c=0:

$$\begin{aligned} \beta \approx -\kappa \cdot \ln(1-c)\Big|_{c=0} + \frac{d}{dc}\Big(-\kappa \cdot \ln(1-c)\Big)\Big|_{c=0} \cdot c \\ = 0 + \frac{\kappa}{1-c}\Big|_{c=0} \cdot c \\ \beta \approx \kappa \cdot c \end{aligned} \tag{2}$$

Therefore, $\beta = \kappa \cdot c$ is the approximation of Eqn 1 at low probability of successful transmission, $c \approx 0$.

1.2 Contact network model distributions

Many animals are aggregated in flocks, herds or other social networks. For directly transmitted diseases in populations with strong social and spatial structuring, we need to revisit the assumption of homogeneous and random host mixing, assumed within density-dependent transmission, as animals interact with a small network, not the entire population

(Ferrari et al., 2011; Craft, 2015; White et al., 2017). Contact models are based on graph-theory, named *epidemic network models*, where the realized per capita transmission rate $\hat{\beta}$ is scaled to the mean neighbourhood size (Ferrari et al., 2011; White et al., 2017). It has been shown that the degree of connections within a population has a large effect on the disease dynamics. For example, the existence of triangular contacts (3 connected individuals) reduces both the initial spread and final disease outbreak (Keeling, 1999). The transmission across a connection is weighted against its connections, using parameter $< k >$, the mean degree number of contacts, or dispersion parameter, which can be scaled in 3 ways (Lloyd-Smith et al., 2005; Ferrari et al., 2011):

1. Independent of total population size, N
   This is similar to the assumptions of the classic frequency-dependent transmission function
2. Increasingly slower than linear $C \cdot \sqrt{N}$, C being a constant
   An intermediate between case 1 and 3.
3. Increasing linearly with population size, N
   This is analogous to the density-dependent transmission function, assuming a constant area.

In epidemic network models, the R_0 calculation is different from the original compartmental models, as it is assumed that some individuals contribute more to the spread of the disease as others, such as *superspreaders* (Craft, 2015; White et al., 2017; Lloyd-Smith et al., 2005). This is similar to some of the previously introduced multi-host models, as in both cases the population consists of individuals contributing differently to the disease spread. However, the nature of transmission described above is homogeneous, therefore not taking into account the heterogeneity of individuals, group formation and the extend of mixing within these groups. It has been shown that in heterogeneous populations the limited spread of invading pathogens creates bigger intraspecific competition than in homogeneously mixed population, resulting in a reduced success invasion (Keeling, 1999). This has a big effect on the R_0 and so should be taken into account. Fortunately, epidemic network models, that are often used to track contact in human epidemics (Childs et al., 2007), have the advantage of being able to take into account the heterogeneous nature of contacts between hosts and the topology of the contact network, by using the mean and variance of contacts. This can be done by choosing from a number of degree distributions. A couple of common examples are listed below (Ferrari et al., 2011).

1.2.1 Poisson distribution

Here, the variance to mean ratio is close to one. Due to this relatively low variance it is closest to homogeneous mixing. A Poisson distribution is based on the assumption that events happen in a certain time interval at a constant rate, irrespective of the time that has passed. An individual contacts another individual at a constant contact rate, with a constant probability of being infected, independent of the number of previously infected individuals. Increasing the number of nodes has either a slightly decreasing or more decreasing effect on $\hat{\beta}$, depending on whether $< k >$ increases linearly with population size (case 3), increases with $\sqrt{N}$ (case 2) or if it is constant with population size (case 1), respectively (Ferrari et al., 2011).

1.2.2 Power-law truncated distribution

This distribution consists of two power law functions, joint at a cutoff-group size value, at which the behavior of the group changes. This cut-off value can be an optimal group size (in mammals) or a certain minimal group size to bring group advantage (Sjöberg et al., 2000). Here, variance to mean ratio $>> 1$, such that in a group of individuals where some have many contacts, whereas others almost none, the phenomenon of super-spreaders can be captured (Ferrari et al., 2011). This distribution assumes extreme heterogeneity in local contacts. $\hat{\beta}$ slightly increased (case 3), slightly decreased (case 2) or decreased more (case 1) with the number of nodes.

1.2.3 An exponential distribution

This is the intermediate case between the Poisson and truncated power-law distributions. It is another case of a memoryless distribution, like the Poisson process, as it models the time between 2 events in a Poisson process. This distribution has, just like the power-law distribution, a greater proportion of superspreaders.
$\hat{\beta}$ more or less remained the same (case 3), slightly decreased (case 2) or linearly decreased (case 1) with an increasing number of nodes.

1.3 Empirical estimations of disease transmission

As described in box 2, there are a number of ways to empirically determine the probability of successful transmission, c. A couple of interesting ones that can be used in the definition of transmission according to this review are listed below.

1.3.1 Secondary Attack Rate to estimate c

The Secondary Attack Rate, or SAR, is a common approximation of c of transmission within a population (Childs et al., 2007). It uses contact tracing, where the secondary cases are defined as the cases that arise within the time-span of incubation and infection time of a primary case.

$$\text{SAR} = \frac{\text{total secondary cases}}{\text{total susceptibles}} \tag{3}$$

This probability ratio can be used as probability of successful transmission, c.

1.3.2 Binomial model of transmission for c

Assume the probability of transmission at contact is p and the probability of escaping infection is $q = 1 - p$. Suppose the susceptible host makes n contacts with an infected host. The probability of escaping infection will be $q^n = (1-p)^n$. Therefore, we can define the probability of getting infection after contact as:

$$1 - q^n = 1 - (1 - p)^n \tag{4}$$

This is the binomial distribution. Here, p can be empirically estimated with the maximum likelihood:

$$\hat{p} = \frac{\text{number of individuals who become infected}}{\text{total number of contacts with infected individuals}} \tag{5}$$

The difference between this method and SAR, is that here the number of contacts is determined relative to the contacts with infectious hosts, whereas SAR weighs the probability of infection against the number of susceptible and exposed hosts. This function is often used for sexually transmitted diseases, so for pathogens showing FD behaviour (Childs et al., 2007).

References

Childs, J. E., Mackenzie, J. S., and Richt, J. A. (2007). *Wildlife and emerging zoonotic diseases: the biology, circumstances and consequences of cross-species transmission*, volume 315. Springer Science & Business Media.

Craft, M. E. (2015). Infectious disease transmission and contact networks in wildlife and livestock. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1669).

Ferrari, M. J., Perkins, S. E., Pomeroy, L. W., and Bjrnstad, O. N. (2011). Pathogens, social networks, and the paradox of transmission scaling. *Interdisciplinary Perspectives on Infectious Diseases*, 2011.

Keeling, M. J. (1999). The effects of local spatial structure on epidemiological invasions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 266:859–867.

Keeling, M. J. and Rohani, P. (2011). *Modeling infectious diseases in humans and animals*. Princeton university press.

Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E., and Getz, W. M. (2005). Superspreading and the effect of individual variation on disease emergence. *Nature*, 438(7066):355–359.

Sjöberg, M., Albrectsen, B., and Hjältén, J. (2000). Truncated power laws: A tool for understanding aggregation patterns in animals? *Ecology Letters*, 3(2):90–94.

White, L. A., Forester, J. D., and Craft, M. E. (2017). Using contact networks to explore mechanisms of parasite transmission in wildlife. *Biological Reviews*, 92(1):389–409.