A mechanistic model for superspreading

Suzanne O'Regan

Drake Lab Meeting, 11/12/2021

Heterogeneity in disease transmission arises frequently in epidemics

- ▶ Individuals can vary in their ability to transmit infectious agents through biological, behavioral and environmental factors
- ▶ Superspreading events, where one infected individual gives rise to a large number of secondary infections in a single generation, may be the source of most of the secondary cases in a population
- ▶ The first wave of the SARS CoV-2 pandemic was characterized by multiple superspreading events
- ▶ Understanding the role of superspreading individuals in fuelling transmission in an outbreak is important for epidemic containment.
- ▶ Superspreading often modeled using branching processes

Modeling the number of secondary infections per infectious individual using a branching process: assumptions

- ▶ Homogeneous population
- ▶ Infectious individuals are independent from each other
- ▶ Early in an epidemic when depletion of susceptibles can be ignored
- ▶ Negative binomial model commonly used for superspreading
- ▶ Negative binomial has two parameters: mean R_0 and dispersion parameter k
- ▶ Macro-level: discrete time, in units equal to the infectious period of an individual

The negative binomial offspring distribution exhibits hallmarks of transmission heterogeneity
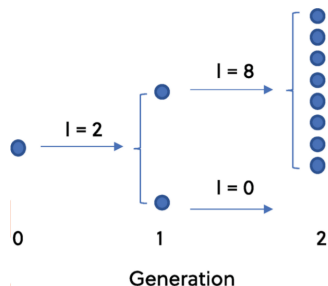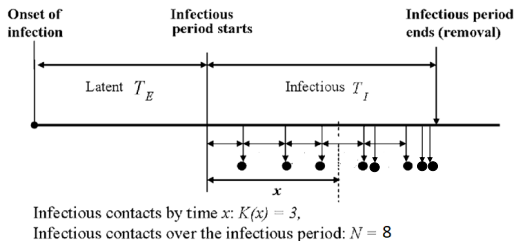

Figure 1: Negative binomial model. Source: Althouse et al. 2020

▶ Greater variability in the number of secondary infections (fat tailed)
▶ Smaller probability of major epidemics
▶ Greater variability in chain sizes
▶ Larger probability of observing no secondary infections and of observing small chains that go extinct

Discrete-time branching process embedded within continuous time branching process

Continuous time Crump-Mode-Jagers branching process

Onset of infection **Infectious period starts** **Infectious period ends (removal)**

Latent T_E Infectious T_I

x

Infectious contacts by time x: $K(x) = 3$,
Infectious contacts over the infectious period: $N = 8$

Discrete time Galton-Watson branching process

l = 2

l = 8

l = 0

0   1   2

Generation

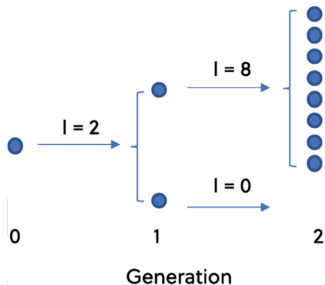Images modified from Yan (2008)
Althouse et al. 2020
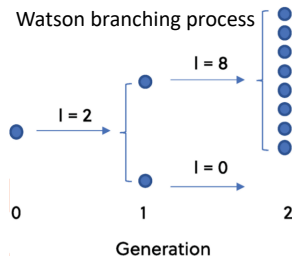
Figure 2: Negative binomial model. Source: Althouse et al. 2020

- ▶ Typically researchers simulate the number of secondary infections per individual in discrete time using an offspring distribution arising from a GW process
- ▶ Additionally, the theory of GW processes offers many analytical and computational advantages. The number of cases per generation can be simulated easily, the probability of a large outbreak can be calculated and the distribution of transmission chains that go extinct can often be obtained

Common model for superspreading: the negative binomial model

- ▶ negative binomial model = Poisson-gamma mixture
- ▶ discrete time (individual infectious period = generation)
- ▶ a) Poisson contact process with intensity λx with gamma-distributed infectious period x with mean $1/\gamma$ and CV $1/\sqrt{k}$ gives rise to negative binomial offspring distribution for the number of secondary infections per infectious individual over the course of their infectious period with mean R_0 and dispersion parameter k
- ▶ b) Lloyd-Smith et al. (2005) assumed individual reproductive number ν is gamma-distributed and demographic stochasticity in individuals follows a Poisson process, yielding a negative binomial offspring distribution with mean R_0 and dispersion parameter k
- ▶ Limitation: The model does not take population risk structure into account. The population may be grouped by social, biological, behavioral or environmental factors

Mechanisms for superspreading transmission

Source of heterogeneity	Factor
Micro-level binary	
Proximity to susceptible individuals (remote worker vs. healthcare worker)	Environmental
Transmission mode (e.g., aerosol vs. droplet transmission)	Biological
Symptomatology (e.g. shedding at high rates vs. low rates)	Biological
Compliance behaviors (e.g., self-isolation when sick vs. no self isolation)	Behavioral
Vaccination status (i.e., vaccinated vs. not vaccinated)	Behavioral
Infectiousness (e.g., having underlying health conditions or not)	Biological or Behavioral
Micro-level continuous	
Symptomatology (infectiousness affecting probability of infection)	Biological
Symptomatology (severe longlasting symptoms)	Biological

Models for the distribution of secondary infections that combine
heterogeneous transmission patterns with realistic distributions of
infection duration are needed

Key questions

- ▶ Does the mechanistic addition of population structure induce qualitatively different outbreak patterns from a standard superspreading model?
- ▶ How does decreasing the level of superspreading by a) changing the population structure e.g., by shifting the contact structure away from opportunistic encounters/aerosol transmission and towards regular contacts/direct contact transmission, and b) decreasing the average number of successful contacts over the course of the average infectious period in the superspreading cohort affect heterogeneity in outbreak patterns, and what are the implications for containment?

Goals

- Derive a mechanistic branching process model
- Derive chain size distribution
- Compare mechanistic model with standard negative binomial model
- Use the model to explore impact of control activities

Mechanistic model assumptions

- ▶ We assume that infected individuals can be divided into two disjoint groups: a fraction p that contribute to transmission via superspreading, and the remaining fraction $1 - p$ that are characterized by regular transmission

- ▶ In the superspreading cohort, the mean cumulative number of contacts leading to transmission of infection per infected individual per unit time is high (β^S) whereas in the regular cohort it is low $(\beta < \beta_S)$

- ▶ The two cohorts contact others according to Poisson processes with different intensities $\beta < \beta_S$

- ▶ Letting C be a random variable denoting the cumulative number of transmission contacts (contacts that infect susceptibles) by time x, a finite mixture of Poisson distributions with probability generating function

$$G(s, x) = p \exp\left(\beta^S x(s - 1)\right) + (1 - p)\exp\left(\beta x(s - 1)\right), \quad s \in [0, 1]$$

describing the stochastic contact process in the population.
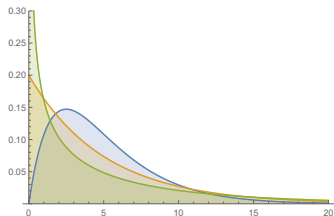
Gamma distributed infectious period assumption

Figure 3: Green: $k = 1/2$, Red: $k = 1$, Blue: $k = 2$

▶ In both groups, the infectious period is gamma distributed with mean $1/\gamma$ and coefficient of variation $1/\sqrt{k}$

▶ The gamma distribution is flexible in that it allows for heavily right-skewed distributions (i.e., $k < 1$), and distributions with a central tendency ($k > 1$).

▶ Strongly right-skewed distributions (i.e., $k < 1$) capture the property of there being a small proportion of individuals in the population with extremely long infectious period, who could therefore make many contacts leading to transmission over the course of being infected.

Mechanistic model is a finite mixture of negative binomials

To find the probability distribution for the cumulative number of transmission contacts generated by an infectious individual throughout its entire infectious period the probability generating function is given by

$$G_N(s) = \int_0^\infty G(s, x) f_I(x) dx$$
$$= \int_0^\infty \left(p e^{\beta^S x(s-1)} + (1-p) e^{\beta x(s-1)} \right) \frac{(\gamma k)^k}{\Gamma(k)} x^{k-1} e^{-k\gamma x} dx.$$

Letting $\beta/\gamma = R_0^R$ and $\beta^S/\gamma = R_0^S$, evaluating the integral yields

$$G_N(s) = \frac{p}{(1 + \frac{\beta^S}{\gamma k}(1-s))^k} + \frac{(1-p)}{(1 + \frac{\beta}{\gamma k}(1-s))^k}$$
$$= \frac{p}{(1 + \frac{R_0^S}{k}(1-s))^k} + \frac{(1-p)}{(1 + \frac{R_0^R}{k}(1-s))^k}.$$

This is a finite mixture of negative binomial models with means $R_0^R < R_0^S$ and dispersion parameter k.

Basic reproduction number for the mixture branching process

▶ The mean number of secondary infections per infectious individual per generation, is

$$R_0 = G'_N(1) = p\frac{\beta^S}{\gamma} + (1-p)\frac{\beta}{\gamma} = pR_0^S + (1-p)R_0^R.$$

▶ If $R_0 < 1$ outbreaks are small and stutter to extinction; if $R_0 > 1$ either the number of infectious individuals grows exponentially or outbreaks are minor and go extinct with probability s^* found by solving $s = G_N(s)$.

Probability mass function of mechanistic mixture model

Evaluating $\frac{1}{j!}\frac{d^j}{ds^j}G_N(0)|_{s=0}$ $j = 0, 1, 2 \ldots$ yields the probability mass function for the number of secondary infections with parameters p, k, R_0^S and R_0^R,

$$P(N = j) = \frac{\Gamma(j+k)}{j!\Gamma(k)}\left[p\left(\frac{k}{k+R_0^S}\right)^k\left(\frac{R_0^S}{k+R_0^S}\right)^j + (1-p)\left(\frac{k}{k+R_0^R}\right)^k\left(\frac{R_0^R}{k+R_0^R}\right)^j\right].$$

The model allows for a variety of infectious histories including having extremely high risk of superspreading transmission to others (e.g., high contact rate and long infectious period), high risk of superspreading transmission to others (e.g., high contact rate and fast recovery rate), moderate risk of being a superspreader (e.g., low contact rate and long infectious period) and being characterized by regular transmission (e.g., low contact rate and fast recovery rate).

Probability mass function of mechanistic model compared to standard model
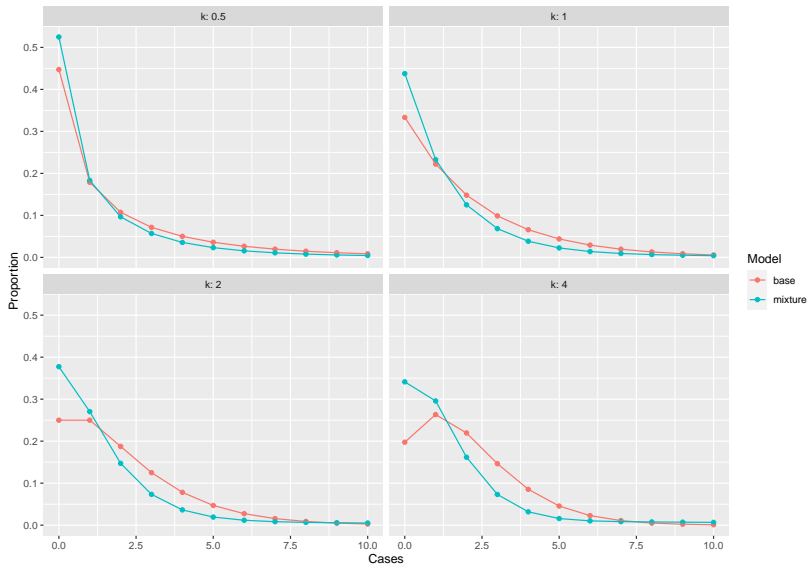
Transmission chains (sizes of small outbreaks that go extinct)

We define a chain that goes extinct at time t by

$$Y = \sum_{i=0}^{t-1} x_i$$

with x_i denoting the cumulative number of offspring in the ith generation, and $x_0 = 1$. The final size Y upon extinction is a random variable with probability distribution $P(Y = y)$, $y = 1, 2, \dots$.

Deriving the chain size distribution

To derive the chain size distribution for the Poisson mixture, we use the result from Blumberg and Lloyd-Smith (2014) and therefore require the derivatives of powers of the generating function. Let

$$T_y(z) = \frac{1}{y}(G_N(z))^y, \quad y = 1, 2, \ldots$$

Then the probability of a chain having size y is

$$P(Y = y) = \frac{1}{(y-1)!} T_y^{(y-1)}(z)\Big|_{z=0} \qquad (1)$$

To evaluate the derivatives of

$$(G_N(z))^y = \left(\frac{p}{(1 + \frac{R_0^S}{k}(1-s))^k} + \frac{(1-p)}{(1 + \frac{R_0^R}{k}(1-s))^k} \right)^y, \qquad (2)$$

we need to apply the chain rule for derivatives $y - 1$ times.

Derivation continued

The nth derivative of the inner function g_n of equation (2), $n = 1, 2, \ldots, y-1$, evaluated at $z = 0$ is

$$g_n^{(n)} = p \frac{(R_0^S)^n}{k^{n-1}} \prod_{i=1}^{n-1}(k+i)\left(1 + \frac{R_0^S}{k}\right)^{-k-n} + (1-p)\frac{(R_0^R)^n}{k^{n-1}} \prod_{i=1}^{n-1}(k+i)\left(1 + \frac{R_0^R}{k}\right)^{-k-n}.$$

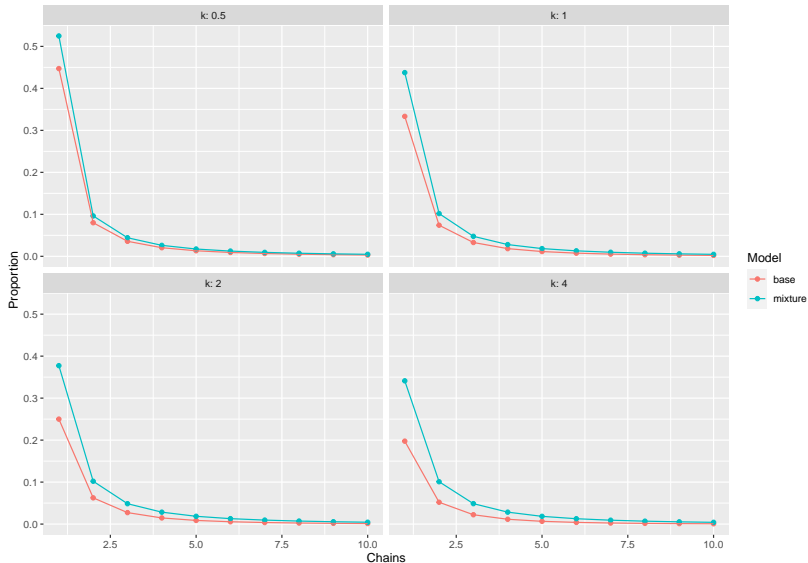The nth derivative of the outer function f_n of equation (2) evaluated at $z = 0$ is

$$f_n^{(n)} = \frac{y!}{(y-n)!}\left(\frac{p}{(1+\frac{R_0^S}{k})^k} + \frac{(1-p)}{(1+\frac{R_0^R}{k})^k}\right)^{y-n}, \quad n = 1, 2, \ldots, y-1.$$

The generalized chain rule (Faa di Bruno's formula) yields

$$T_y^{(y-1)}\Big|_{z=0} = \sum_{n=1}^{y-1} f_n B_{y,n}(g_1, g_2, \ldots, g_{y-1-n}) \tag{3}$$

where $B_{y,n}$ are Bell polynomials of the g_n.

Chain size distribution of mechanistic model compared to standard model

# Statistics that show hallmarks of transmission heterogeneity

Hallmarks of heterogeneous transmission include:

▶ Greater variability in the number of secondary infections (fat tailed)

▶ Smaller probability of major epidemics

▶ Greater variability in chain sizes

▶ Larger probability of observing no secondary infections and of observing small chains that go extinct

Numerical study to compare mechanistic mixture model with standard model

▶ Here we study the coefficient of variation of the number of secondary infections, the probability of a major outbreak, the probability of observing a small transmission chain of less than or equal to 10 cases, and the coefficient of variation of small chain sizes (conditioned on extinction).

▶ In each of the following, p and δ are varied (which alters R_0^S) but $R_0 = pR_0^S + (1-p)R_0^R = R_0^R + p\delta$ is fixed at $R_0 = 2$.
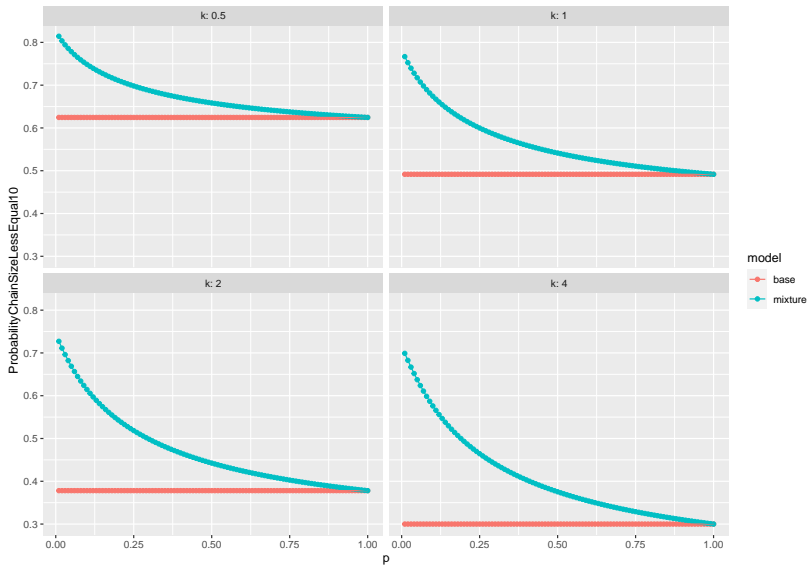
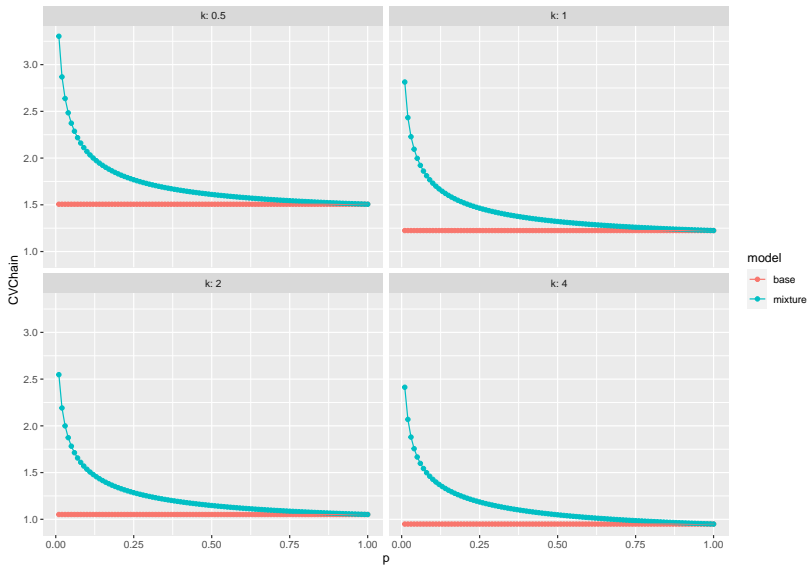# Coefficient of variation of secondary infections
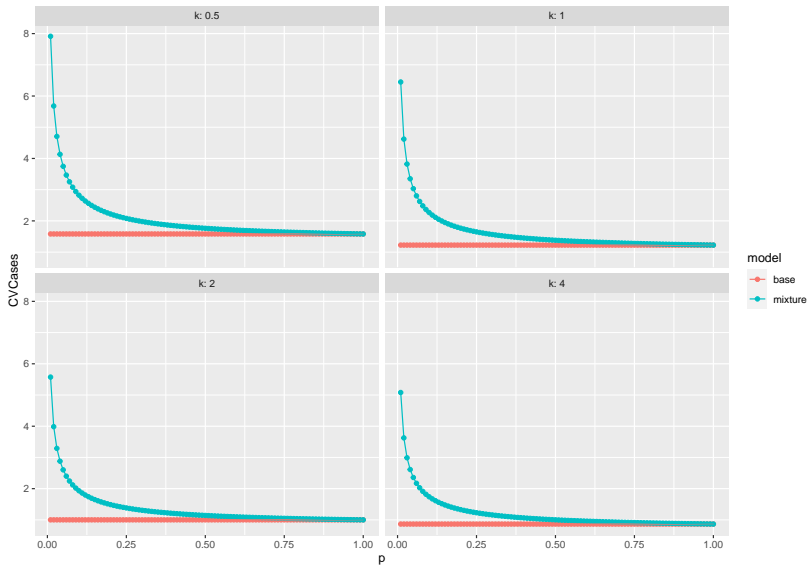
Probability of major outbreak
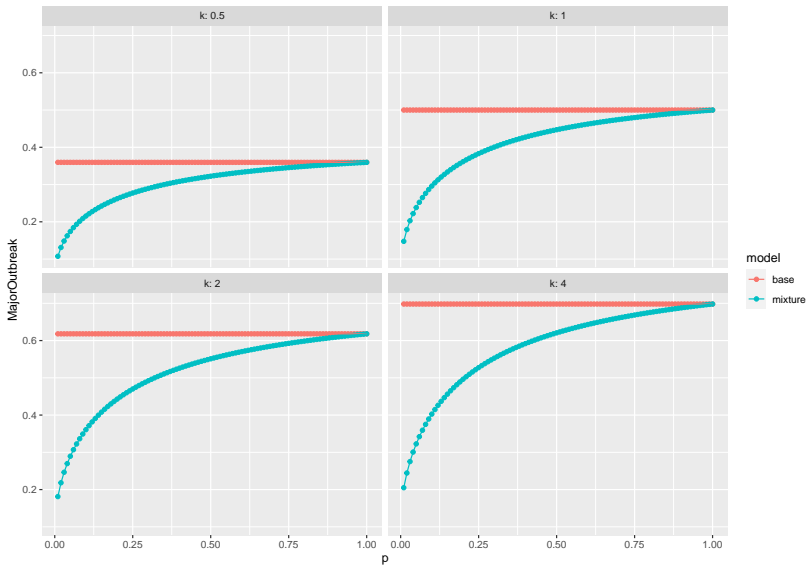
Probability of observing a transmission chain of size ≤ 10

Coefficient of variation of chain sizes

Preliminary conclusions from this study

▶ Having a small proportion of superspreaders with high R_0^S (i.e., smaller values of p and larger values of δ) lead to more heterogeneous epidemics than the standard model, even if the dispersion parameter $k > 1$.

▶ Having a dispersion parameter k that is less than one is not necessary for the mixture model to exhibit hallmarks of superspreading transmission, suggesting that heterogeneity in the contact process is sufficient.