

Yes, it matters who is spreading monkeypox: Supplemental Information

Jonathan Smith, PhD, MPH

September 2, 2022

Contents

1	Theoretical Overview	2
2	Proportion of secondary cases attributed to the proportion of infectious cases	2
3	Final Outbreak Size	3
4	References	4

1 Theoretical Overview

Following previous studies, we use branching process theory and assume the number of secondary cases produced from each infectious case is distributed according to a negative binomial distribution with mean R_0 (the basic reproductive number) and dispersion parameter k . The parameter k quantifies the degree of individual heterogeneity by measuring overdispersion in the distribution (e.g., higher than expected variation). For a given R_0 , smaller values of k ($k \ll 1$) suggest increased heterogeneity in the number of secondary cases between individual index cases. As k increases, transmission becomes more uniform and centered around the mean, R_0 ; as k approaches infinity the negative binomial distribution converges to the Poisson distribution, wherein all differences in the number of secondary cases is attributed to demographic stochasticity.

Each infectious case is assigned an individual reproductive number, ν , which describes the expected number of secondary cases produced from that individual. The observed number of secondary cases, z , is a product of both the generating function of the branching process (the underlying mechanism driving transmission) and demographic stochasticity. We assume ν is gamma distributed with mean R_0 and dispersion k and demographic stochasticity is classically modeled using a Poisson process, i.e., $z \sim Poi(Gamma(R_0, k))$. This gamma-Poisson mixture results in a negative binomial offspring distribution (Lloyd-Smith, 2005).

Code to implement the included methods and recreate figure is available at the following GitHub repository:

<https://github.com/jpsmithuga/Comment-Supplementary>

2 Proportion of secondary cases attributed to the proportion of infectious cases

Lloyd-Smith *et al.* (in [in supplementary information 2.2.5](#)) propose a distribution for describing transmission based from the distribution of ν :

$$F_{\text{trans}}(x) = \frac{1}{R_0} \int_0^x u f_{\nu}(u) du$$

This gives the cumulative distribution function (CDF) in terms of the density f_{ν} , as the proportion of all transmission due to infectious individuals with reproduction number $\nu < x$. If we make the classical assumption that ν is gamma distributed with shape parameter $k > 0$ and rate parameter k/R_0 , where R_0 is the mean (expected) individual reproduction number, we have

$$f_{\nu}(u) = \frac{k^k}{R_0^k \Gamma(k)} u^{k-1} e^{-ku/R_0}$$

It follows that

$$\begin{aligned}
f_{\text{trans}}(x) = F'_{\text{trans}}(x) &= \frac{1}{R_0} x f_{\nu}(x) \\
&= \frac{k^k}{R_0^{k+1} \Gamma(k)} x^k e^{-kx/R_0} \\
&= \frac{k^{k+1}}{R_0^{k+1} \Gamma(k+1)} x^k e^{-kx/R_0} \\
&\sim \text{Gamma}(k+1, k/R_0)
\end{aligned}$$

The rate parameter remains the same and the shape parameter increases by 1. To calculate t_p , the expected proportion of transmission due to the most infectious 100 p % of cases, we first find the $(1-p)$ th centile of the individual reproduction number, $x_p = F_{\nu}^{-1}(1-p)$, then calculate $t_p = 1 - F_{\text{trans}}(x_p)$. As both random variables are gamma distributed, this is implemented in R via the ‘qgamma()’ (F^{-1}) and ‘pgamma()’ (F) functions (see [here](#)).

3 Final Outbreak Size

Let Y represent the sum of all cases in a transmission chain (including the index case and all secondary, tertiary, etc, cases resulting from that index case), resulting in a final outbreak size of Y . Using branching process theory, the distribution of the total number of secondary infections in a single transmission chain (including the index case at generation) can be defined by the probability generating function (pgf) yielding the implicit relationship (Yan 2008):

$$G_Y(s) = sG_Z((G_Y(s)))$$

We can treat this as $G_Y(s) = \sum_{y=1}^{\infty} q_y s^y$, where $q_y = P(Y = y)$, allowing us to extract the probability (Becker 1974; Blumberg 2014 (in [Supplemental Text S1](#)):

$$P(Y = y) = \frac{1}{y!} \left. \frac{d_y G_Y(s)}{ds^y} \right|_{s=0}$$

Where $Y = 1, 2, 3 \dots$. Substituting the negative binomial distribution yields the following probability of a outbreak: (see Blumberg 2012 ($R < 1$), Nishiura 2012 ($R \geq 1$), and Smith 2022 (multiple index cases, n , where $n \geq 1$))

$$P(Y = y) = \binom{n}{y} \frac{\Gamma(ky + y - n)}{\Gamma(ky)(y - n)!} \frac{(\frac{R_0}{k})^{y-n}}{(1 + \frac{R_0}{k})^{ky+y-n}}$$

Since when $R > 1$ we have $P(Y = \infty|n) > 0$, the distribution is improper for $R > 1$, though becomes proper for an $R > 1$ when we define $P(Y = \infty|n) = 1 - q(\theta)^s$, where $q(\theta)^s$ is the probability of extinction (see Nishiura 2012 (section 2.3, eq. 17); Farrington 2003 (pg 284)). This is implemented in R software (see [here](#)).

4 References

- Becker, N. On parametric estimation for mortal branching processes. *Biometrika*, 61(3):393, 1974.
- Blumberg S and Lloyd-Smith JO. Inference of r_0 and transmission heterogeneity from the size distribution of stuttering chains. *PLoS Comput Biol.*, 5(9):393, 2013.
- Dwass, M. The total progeny in a branching process and a related random walk. *Journal of Applied Probability*, 6(3):682, 1969.
- Farrington CP, Kanaan MN, Gay NJ. Branching process models for surveillance of infectious diseases controlled by mass vaccination. *Biostatistics*. 2003 Apr;4(2):279-95.
- Lloyd-Smith, JO, et al. Superspreading and the effect of individual variation on disease emergence, *Nature* 438, 355–359 (2005)
- Nishiura H, Yan P, Sleeman CK, Mode CJ. Estimating the transmission potential of supercritical processes based on the final size distribution of minor outbreaks. *J Theor Biol.* 2012 Feb 7;294:48-55. doi: 10.1016/j.jtbi.2011.10.039. Epub 2011 Nov 7. PMID: 22079419; PMCID: PMC3249525.
- Smith, JP, et al. A cluster-based method to quantify individual heterogeneity in tuberculosis transmission. *Epidemiology* 33, 217–227. 2022.
- Yan, P. *Distribution Theory, Stochastic Processes and Infectious Disease Modelling*, pages 229. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.