

Supplement: Assessing the plausibility of supercritical transmission for an emerging or re-emerging pathogen

Seth Blumberg, Phoebe Lu, Thomas M Lietman, Travis C Porco

November 3, 2021

All calculations and simulations are conducted in R version 4.1.1. The code is provided at

https://github.com/proctor-ucsf/supercritical_plausibility.

1 Inference of R

Based on prior research, we assume that the every new infection causes subsequent infection according to a negative binomial distribution with mean, R and dispersion parameter k .¹ With this assumption, the likelihood that i cases cause j cases in a single generation of transmission is,^{2,3}

$$L_{i \rightarrow j} = \frac{\Gamma(j + ki)}{\Gamma(j + 1)\Gamma(ki)} \left(\frac{k}{R + k}\right)^{ki} \left(\frac{R}{R + k}\right)^j. \quad (1)$$

The likelihood that m primaries causes n total cases over multiple generations is then,^{2,4,5}

$$L_{m \rightarrow n}^C = \frac{m}{n} L_{n \rightarrow (n-m)}. \quad (2)$$

Maximizing the likelihood reveals that the maximum likelihood value of R is $\frac{n-m}{n}$, which is simply the proportion of cases due to secondary transmission (i.e. non-imported cases).⁶ The confidence interval for R is obtained by the likelihood ratio test, which models the log of the likelihood function as following a chi-squared distribution.⁷ The threshold number of transmissions indicating concern that R may have exceeded a specific value (Figure 1C and 1D of the main text) was determined by incrementally increasing the number of observed transmissions and observing when the inferred confidence interval for R crossed the threshold value.

2 Notable caveats

Notable caveats to these equations are that the R and k do not change from generation to generation, and that all cases are observed. Processes such as depletion of susceptible individuals or isolation can decrease R changes in late generations of a transmission chain. Such changes would lead to an overestimation of the capacity for sustained transmission, and in this context our estimates for the probability of sustained transmission are conservative.

Imperfect case observation has more complicated implications for the bias of our R estimates. A particular concern is if, at the time of data collection, many cases remain infectious. This would imply that the full potential full endemic transmission may not yet have been demonstrated and could lead to underestimation of the capacity for sustained transmission. A second concern is that cases do not always reach the attention of surveillance systems. We address this possibility in the next section.

3 Simulation of imperfect observation

We model two possible ways that cases may be observed. The first is that there is an independent probability of $p_{passive}$ that any given case will be observed by the surveillance system. Once one or more cases in a chain is observed by this ‘passive’ process, we assume the other cases in the transmission chain can be observed via an ‘active’ contact tracing process with an independent probability of p_{active} .^{6,8} Perfect observation corresponds to $p_{passive} = 1$. Absence of contact tracing corresponds to $p_{active} = 0$, and perfect contact tracing corresponds to $p_{active} = 1$.

To conduct our analysis of imperfect observation, we run 1,000 simulations of observation scenarios corresponding to specific values of $p_{passive}$ and p_{active} . The output of each simulation is an observed number of secondary transmission events for a pre-specified number of observed importations. For each simulation run, we infer the corresponding confidence interval for R . We use the distribution of these inferred confidence intervals to examine the impact of imperfect observation via Table 2 of the main text.

REFERENCES:

1. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature*. 2005;438(7066):355-359. doi:10.1038/nature04153
2. Blumberg S, Funk S, Pulliam JR. Detecting differential transmissibilities that affect the size of self-limited outbreaks. *PLoS pathogens*. 2014;10(10):e1004452.
3. Lange K. *Applied Probability*. Second. Springer; 2010.
4. Dwass M. The total progeny in a branching process and a related random walk. *Journal of Applied Probability*. 1969;6(3):682-686.

5. Nishiura H, Yan P, Sleeman CK, Mode CJ. Estimating the transmission potential of supercritical processes based on the final size distribution of minor outbreaks. *Journal of Theoretical Biology*. 2012;294:48-55. doi:10.1016/j.jtbi.2011.10.039
6. Blumberg S, Lloyd-Smith JO. Inference of R_0 and transmission heterogeneity from the size distribution of stuttering chains. *PLoS Computational Biology*. 2013;9(5):e1002993.
7. Bolker BM. *Ecological Models and Data in R*. Princeton University Press; 2008.
8. Ferguson NM, Fraser C, Donnelly CA, Ghani AC, Anderson RM. Public health risk from the avian H5N1 influenza epidemic. *Science*. 2004;304(5673):1-5. doi:10.1126/science..