Characterizing spatiotemporal variation in transmission heterogeneity during the 2022 mpox outbreak in the USA

Jay Love*[1], Cormac R. LaPrete*[2], Theresa R. Sheets[2], George G. Vega Yon[1], Alun Thomas[1], Matthew H. Samore[1], Lindsay T. Keegan[1], Frederick R. Adler[2,3], Rachel B. Slayton[4], Ian H. Spicknall[5], Damon J.A. Toth[1]

*Contributed equally

[1] Division of Epidemiology, University of Utah, Salt Lake City, Utah
[2] Department of Mathematics, University of Utah, Salt Lake City, Utah
[3] School of Biological Sciences, University of Utah, Salt Lake City, Utah
[4] Division of Healthcare Quality and Promotion, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia
[5] Division of STD Prevention, Centers for Disease Control and Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD and TB Prevention, Atlanta, Georgia

Abstract

Transmission heterogeneity plays a critical role in the dynamics of an epidemic. During an outbreak of an emerging infectious disease, efforts to characterize transmission heterogeneity are generally limited to quantifications during a small outbreak or a limited number of generations of a larger outbreak. Understanding how transmission heterogeneity itself varies over the course of a large enduring outbreak not only improves understanding of observed disease dynamics but also informs public health strategy and response. In this study, we employ a method, adaptable to other emerging infectious disease outbreaks, to quantify spatiotemporal variation in transmission heterogeneity for the 2022 mpox outbreak in the United States. Based on past research on mpox and following reports of potential superspreading events early in this outbreak, we expected to find high transmission heterogeneity as quantified by the dispersion parameter of the offspring distribution, k. Our methods use maximum likelihood estimation to fit a negative binomial distribution to transmission chain offspring distributions informed by a large mpox contact tracing dataset. We find that, while estimates of transmission heterogeneity varied across the outbreak with spatiotemporal pockets of higher heterogeneity, overall transmission heterogeneity was low. When testing our methods on simulated data from an outbreak with high transmission heterogeneity, k estimate accuracy depended on the contact tracing data completeness. Because the actual contact tracing data had high incompleteness, our values of k estimated from the empirical data may be artificially high. However, it is also possible that our estimates accurately reflect low transmission heterogeneity for the United States mpox outbreak.

Introduction

In an infectious disease outbreak, transmission heterogeneity occurs when there is variation in the number of secondary infections per infected host, which can be caused by underlying

mechanistic heterogeneity. This variation can manifest at spatial (May and Anderson 1984), temporal (Bjørnstad, Finkenstädt, and Grenfell 2002), and/or individual (Lloyd-Smith et al. 2005; Ke et al. 2022) scales. These different components of transmission heterogeneity interact with one another, creating the potential for complex patterns of transmission heterogeneity over the course of an outbreak. Because transmission heterogeneity plays a critical role in epidemic spread, characterizing how it changes over an outbreak can help to illuminate the causes behind observed epidemic dynamics, informing measures taken by public health decision makers.

Prior methods have simultaneously estimated the mean transmission (R) and the dispersion (k) parameters from data on the number of secondary infections per primary infection (*i.e.*, the offspring distribution) in an outbreak (Lloyd-Smith et al. 2005; Lloyd-Smith 2007; Blumberg and Lloyd-Smith 2013; Toth et al. 2015). These methods have been applied both retrospectively and during ongoing outbreaks (Toth et al. 2016; Kucharski and Althaus 2015; Nishiura et al. 2015) for several diseases. Past research on transmission heterogeneity has largely focused on small outbreaks or has used data from early transmission generations in larger outbreaks (Lloyd-Smith et al. 2005; Lloyd-Smith 2007; Blumberg and Lloyd-Smith 2013; Toth et al. 2015; 2016; Kucharski and Althaus 2015; Nishiura et al. 2015). Some research has explored how the first transmission generation may differ from subsequent generations in R and k (Toth et al. 2016), but work to understand how individual transmission heterogeneity may vary over the component outbreaks of a large epidemic is more rare (Lau et al. 2020; Tkachenko et al. 2021; Miyama, Jung, and Nishiura 2022).

Among those diseases for which transmission heterogeneity has been characterized in the past is mpox (formerly 'Monkeypox' (World Health Organization 2022b)). Monkeypox virus (MPXV) is a member of the *Orthopoxvirus* genus of pathogens and is endemic in West and Central Africa. Mpox was first detected in humans in 1970, shortly after the eradication of smallpox (Breman et al. 1980), and has subsequently caused occasional outbreaks thought to result from repeated spillover events (Nalca et al. 2005; Yinka-Ogunleye et al. 2018; Bunge et al. 2022; Durski et al. 2018). While the exact reservoir host is still unknown, the virus has been found in several rodent species native to central and west African countries (Durski et al. 2018). Mpox is known to spread through direct contact with mpox rash, scabs, or bodily fluids; touching objects that have been used by someone with mpox; contact with respiratory secretions; and sexual or intimate contact with an infected person (Centers for Disease Control and Prevention 2022).

Past outbreaks have been largely contained to a modest number of cases. However, regional outbreaks have been growing in recent years (Sklenovská and Van Ranst 2018; Beer and Bhargavi Rao 2019; Rimoin et al. 2010). Past work to define the basic reproductive number (R_0, the expected mean transmissions per infection in a large, fully susceptible population), for mpox outbreaks have produced estimates below 1, the critical threshold above which an infectious disease will continue to spread and below which an outbreak will eventually die out. However, more recent work has accounted for declining cross-protective immunity from smallpox to reach an estimate of $R_0 = 2.13$ (uncertainty bounds: 1.46-2.67) for mpox in the Democratic Republic of Congo (Grant, Nguyen, and Breban 2020). Past estimates of transmission heterogeneity, quantified by the inversely varying dispersion parameter (low values of k indicate high transmission heterogeneity) of the offspring distribution (k), for mpox outbreaks have been 0.58 (90% confidence interval [CI]: 0.32-3.57) (Lloyd-Smith et al. 2005), 0.36 (0.16-1.47), and 0.33

(0.19-0.64) (Blumberg and Lloyd-Smith 2013), indicating high transmission heterogeneity.

In April 2022, MPXV was identified outside of the countries where the virus is endemic and quickly spread to over 100 countries (World Health Organization 2022a). The World Health Organization (WHO) declared this mpox epidemic a Public Health Emergency of International Concern on July 23, 2022. This rapid global spread was not characteristic of an infectious disease with a basic reproductive number below 1. Contact tracing conducted during the 2022 outbreak indicated that the disease was initially spreading by sexual contact between gay and bisexual men, and nonbinary and transgender people, often at large gatherings (Philpott et al. 2022). In light of this, public health messaging focused on reducing spread among this group of people, especially at large gatherings (Delaney et al. 2022). Potentially due to a novel, sexually mediated, dominant transmission route, the discrepancy between estimates of the basic reproductive number for mpox outbreaks became globally relevant. Estimates of the 2022 outbreak vary by region and range from 1.56 to 3.8 (Branda, Pierini, and Mazzoli 2022) or potentially higher (Endo et al. 2022).

Because the degree distribution of sexual contact networks can be overdispersed, having greater variability than expected, transmission through sexual contact networks can result in heterogeneous patterns of secondary transmission across the infected population (Handcock and Jones 2006). Spreading of the disease through large gatherings, termed 'superspreader events', can also cause high levels of transmission heterogeneity (Althouse et al. 2020). Such heterogeneous transmission has been shown to influence outbreak dynamics (Toth et al. 2015; Elie, Selinger, and Alizon 2022) and therefore may have contributed substantially to the rapid global spread of mpox. Modeling studies have provided support for the idea that heterogeneous transmission through overdispersed contact networks could partially explain the dynamics of the 2022 mpox outbreak (Endo et al. 2022; Spicknall et al. 2022). To date, empirical estimates of transmission heterogeneity for the present outbreak have not been published.

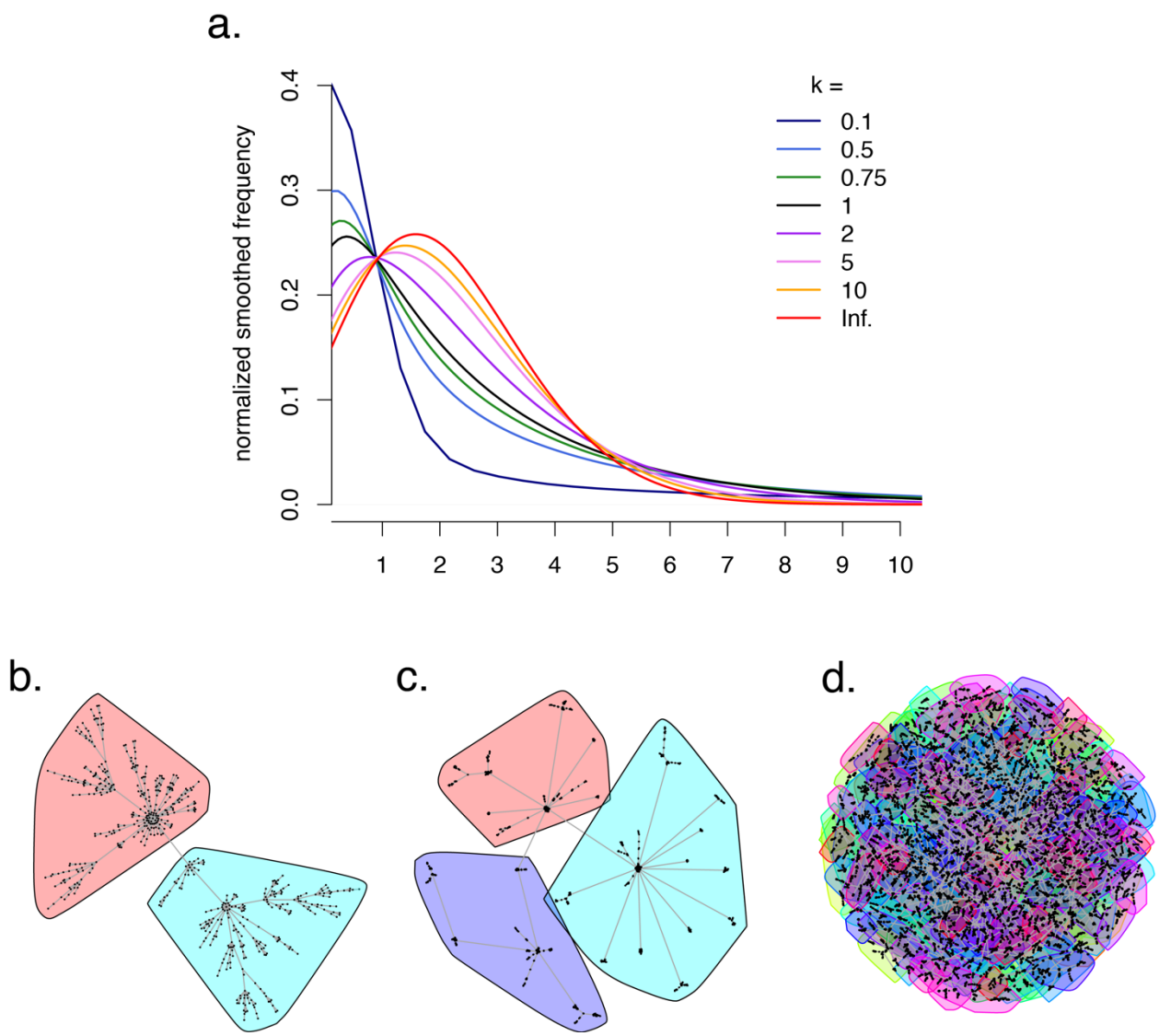
Here, we estimate the mean transmission and transmission heterogeneity for the 2022 mpox epidemic in the United States using contact tracing data fortified with algorithmically determined probable case pairs. We estimate these parameters directly from the empirical data and, through simulation, assess the sensitivity of our approach. We stratify our analysis by geographic location and time. In this way, we characterize transmission heterogeneity on spatial, temporal, and individual scales. Our aims are to (1) better understand the dynamics of the 2022 mpox outbreak and (2) advance the theory of transmission heterogeneity.

Methods

Overview

We quantified the mean transmission using R, defined as the mean number of secondary transmissions per infected individual at a given time period and location and quantified as the mean of the offspring distribution generated from a series of transmissions within that period and location. We quantified transmission heterogeneity using k, the dispersion parameter of the negative binomial distribution, with lower values indicating greater heterogeneity (Fig. 1a). We generated estimates of R and k for the United States and its public health jurisdictions at different periods of time based on offspring distributions of algorithmically generated numbers of

secondary infections per case. We then used a maximum likelihood approach to fit negative binomial distributions to the offspring distributions of given regions and time periods and estimated the 95% confidence intervals (CI) for both parameters (*detailed below*). Transmission trees were constructed using methods similar to those in previous works (e.g., Haydon et al 2003; Wallinga and Teunis 2004; Heijne et al 2012), though our method differs in notable ways, as detailed below.



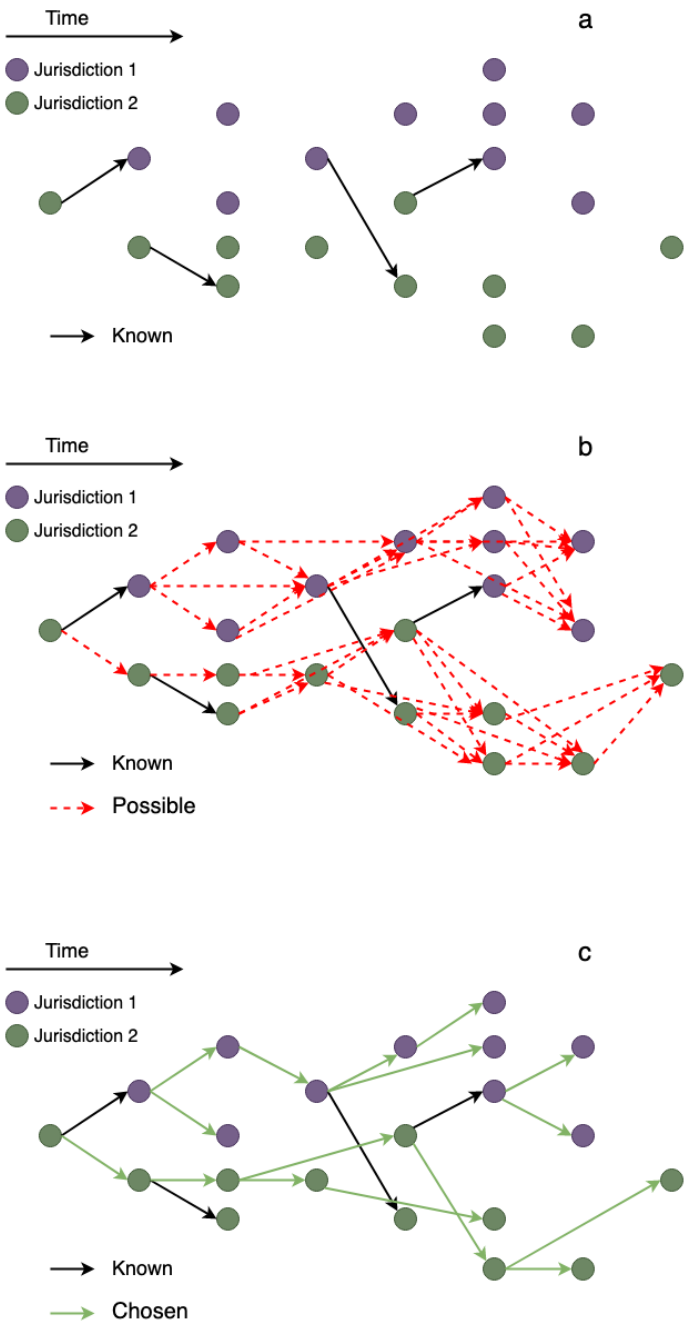
TEXT BOX: Figure 1. (a) Frequencies of negative binomially distributed discrete data, smoothed with a bandwidth of .75 and normalized by total samples for visual clarity. These negative binomial pseudo-distributions have R (mean) = 2 and varying k (dispersion). When k is infinity (red), the negative binomial distribution is equivalent to the Poisson. (b-d) Visualizations of the contact network used in simulations, with simulated jurisdictions shown in colored polygons. See text for description of methods used to generate networks (b) shows two jurisdictions, (c) shows three jurisdictions, inclusive of the two in (b), and (d) shows the full contact network. Overlapping polygons in (d) are visual artifacts of the large network – each node is associated with only one jurisdiction.

Data

We used data collected through U.S. Centers for Disease Control and Prevention (CDC) case detection and contact tracing in the United States from May 1st to November 7th, 2022. This dataset includes, for each probable and confirmed case, information on date of symptom onset, date of reporting, date of positive test, reporting jurisdiction, any recent travel history, and any known contacts with other probable or confirmed cases.

Processing

To compile the offspring distribution, we started by including contact-tracing identified probable transmission events and their associated cases. For each case where a probable infector has not been identified by contact tracing, we used dates of symptom onset included in the dataset and published serial interval estimates to compile a pool of possible infectors (Fig. 2). We defined a possible infector of a case as one with overlapping location (US county) and whose date of symptom onset is within one maximum serial interval, 21 days, of the case's date of symptom onset (UK Health Security Agency 2022). Because the pool of possible infectors generally includes more than one possible infector, we then sampled one possible infector from each pool to serve as the chosen infector for each case. In this sampling process, the serial interval was calculated for each possible transmission, and the probability that a serial interval was sampled is equal to the gamma distribution density, parameterized according to published estimates of the serial interval distribution for mpox (shape = 2.9, scale = 2.9; (Madewell et al. 2022)). The case producing the sampled serial interval was assigned to be the most probable infector for a case. From this dataset, containing both contact-tracing identified and algorithmically determined transmission pairs, we calculated the number of secondary infections produced for each case and fit a negative binomial distribution to the resulting offspring distribution, finding the best-fit parameter values and their 95% CIs. This entire process was repeated 100 times with replacement, producing a sample of the possible offspring distributions of an outbreak.



TEXT BOX: Figure 2. A representation of data processing to build a complete transmission tree. (a) Pre-processed case data points from two public health jurisdictions with four contact tracing identified links. (b) Arrows added to show pools of possible infectors for each case, drawn based on jurisdiction and time (each column of points represents one time step where, for example, possible infectors come from only the most recent two time steps). (c) Chosen infectors sampled and other possible infectors disregarded. This is the post-processing data set used for analysis.

We compiled a separate set of offspring distributions using only the contact tracing-identified transmission pairs. Because this set of data is much smaller than the hybrid dataset described above (see results), we employ it only at the national level. Parameter estimates from the two datasets are compared as a validity check.

Parameter estimation

To fit a negative binomial distribution to the offspring distribution, we used a maximum likelihood approach, following established methods (Lloyd-Smith et al. 2005; Toth et al. 2016). This approach is sensitive to the number of individuals with mpox assumed to transmit to no other individuals, which we identify as terminal cases. Because a portion of the individuals in our dataset with no reported offspring may have actually transmitted the disease to others without being detected, including these false terminal cases would artificially inflate the number of zeros in the offspring distribution. We used the zero-inflated negative binomial to test the sensitivity of our methods to this type of potential bias, complementing past work that accounts for underreporting of terminal cases (Lloyd-Smith 2007).

The zero-inflated negative binomial differs from the negative binomial distribution in its inclusion of a parameter p_0, the probability that an individual's onward transmission data are absent, independent of the number of individuals they infected, leaving 1 - p_0 as the probability that the individual's transmission data are complete. A lower p_0 indicates greater confidence in the accuracy of the data, and the zero-inflated negative binomial reduces to the negative binomial when $p_0 = 0$. Similar to past work that fits modified negative binomial distributions to discrete data that undercounts zeros (Ridout, Demetrio, and Hinde 1998; Lloyd-Smith 2007), accounting for p_0 helps to correct for missingness in the data and make more reliable estimates of R and k. The probability mass function (pmf) for the zero-inflated negative binomial can be defined as follows, with the reported number of offspring for each infected individual distributed based on whether or not the data point, y, is a zero, where we assume the true number of offspring is negative binomially distributed:

$$f(y) = \begin{cases} p_0 + (1-p_0)\left(\frac{k}{k+R}\right)^k & \text{if } y = 0 \\ (1-p_0)\begin{pmatrix} k+y-1 \\ y \end{pmatrix}\left(\frac{R}{k+R}\right)^y\left(\frac{k}{k+R}\right)^k & \text{if } y \neq 0 \end{cases}$$

In the negative binomial estimates, the logarithms of negative binomially distributed probabilities of each data point are a function of R and k, while in the zero-inflated estimates the logarithms of zero-inflated negative binomially distributed probabilities of each data point are a function of R, k, and p_0 . We used maximum likelihood estimation by taking the log-likelihood function of the entire data set. We then used multivariate optimization to find R and k for the negative binomial estimates and R, k, and p_0 for the zero-inflated estimates. We used the likelihood ratio test to find 95% CIs for these parameters.

Sliding and telescoping window approaches

To evaluate change in parameter estimates over time, we conducted our fitting methods on offspring distributions generated across both sliding and telescoping windows of time. In the

sliding window approach, offspring distributions were generated consecutively over 14-day periods, with each consecutive window starting 7 days after the start of the previous window. The sliding window approach generates estimates that can be likened to parameters R_t and k_t described elsewhere (Schneckenreither et al. 2022; Zhang, Britton, and Zhou 2022; Wallinga and Teunis 2004), though to avoid confusion regarding definitions and assumptions that differ from those of our estimated parameters, we simply use R and k. In the telescoping window approach, consecutive windows are inclusive of the previous window; the end date for each consecutive window is advanced 7 days while the start date remains the same, May 1st, 2022. Estimates of R and k generated from the telescoping window approach describe the mean and dispersion of secondary transmission from the beginning of the outbreak up to the end date of a given time window. In this way, they can be thought of as the effective R and k parameters, but again we simply use our previously defined R and k terms for the mean and dispersion of the offspring distribution.

Simulations

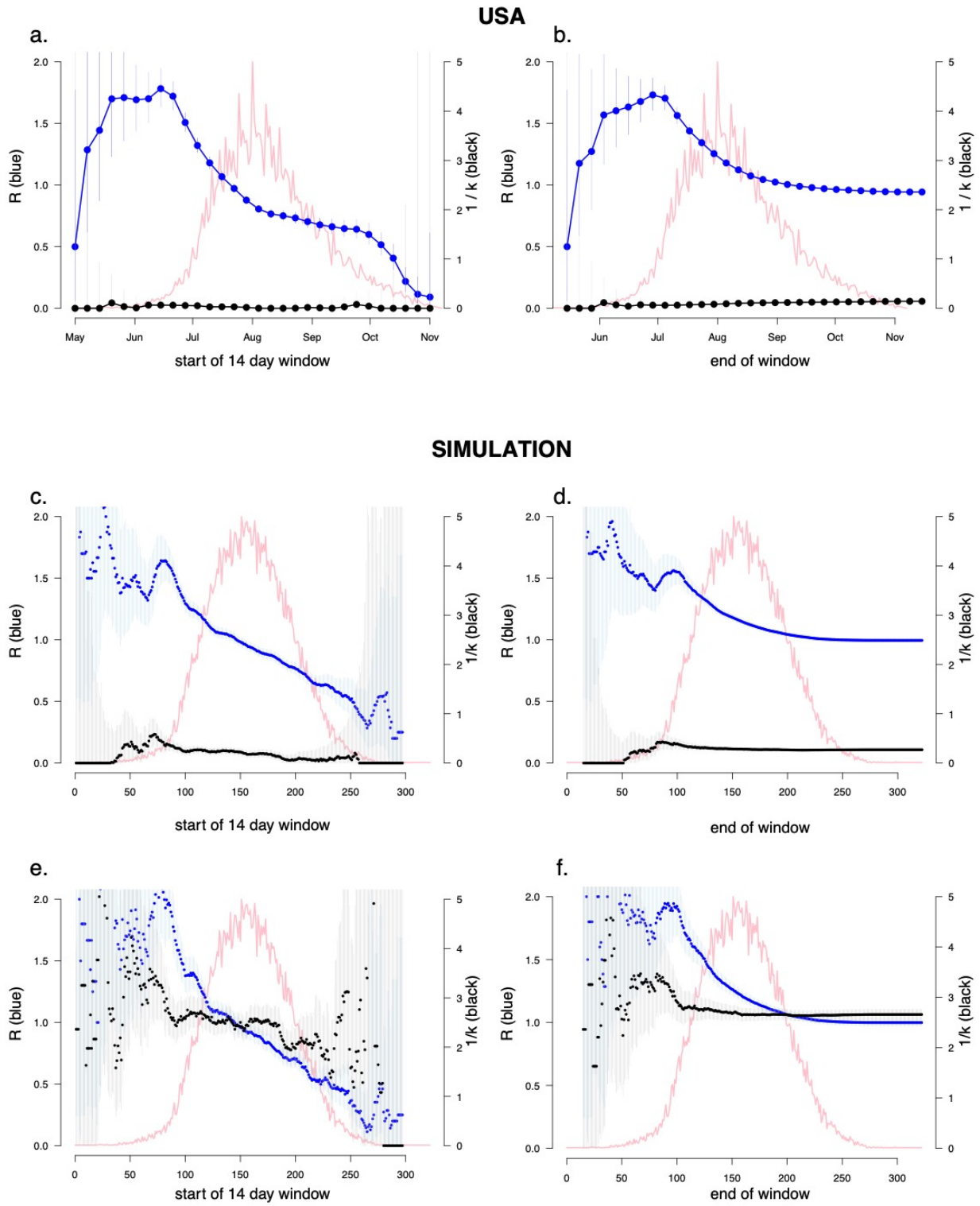
To characterize the reliability of our methods to estimate transmission heterogeneity, we repeated them on simulated data generated with a model of infectious disease spread over a static contact network. We conducted simulations on undirected scale-free networks, shown to produce relatively heterogeneous transmission (Bansal, Grenfell, and Meyers 2007). We initialized the scale-free network with a Barabási-Albert algorithm of preferential attachment (Barabási and Albert 1999), allowing for the network to serve as a proxy for a population with highly heterogeneous contact rates where simulated outbreaks can be expected to result in an overdispersed offspring distribution. The network was generated using the *barabasi.game* function from the *igraph* package in the *R* programming environment (Csardi and Nepusz 2006), using a population of 50,000 and power of 1.

Each outbreak was seeded with a randomly chosen single infection. Disease dynamics were modeled following a *Susceptible – Exposed – Infectious – Recovered* paradigm, parameterized for mpox. The latent period was a random variable drawn from a lognormal distribution with log mean = 1.5 and log standard deviation = 0.7 (Madewell et al. 2022). For lack of detailed information on the duration of infectiousness for mpox, the infectious period in our simulations was a random variable drawn from a gamma distribution with shape = 21 days and scale = 1. This distribution is consistent with the reported serial interval distribution (Madewell et al. 2022), clinical guidance indicating symptoms last 2-4 weeks (Word Health Organization 2022c), and investigations reporting that some hospitalized cases show symptoms for much longer periods (Adler et al. 2022). However, it should be noted that more detailed study may reveal that the distribution we used does not describe the infectious period for mpox well. Contacts of infectious nodes were infected with probability of 0.15 per contact day.

In order to replicate our methods on the simulated data, we assigned each node a "jurisdiction" using a modularity optimization algorithm that identifies community structure in networks (Blondel et al. 2008), as depicted in Figure 1b. We then identified the most probable infector for each case, as detailed above, and estimated the parameters R and k by fitting a negative binomial distribution, as detailed above. In order to test the sensitivity of our methods to incompleteness in detailed contact information, we also conducted a batch of simulations with the same level of incompleteness as the empirical dataset.

All computational methods were conducted in R statistical software (R Core Team 2017).

Results



TEXT BOX: Figure 3. Temporal variation in transmission parameters. Left column shows

estimates made with a sliding window, right column shows estimates made with a telescoping window, starting on May 1st, 2022. Time points for the sliding window approach (left column) are the start of the 14-day sliding windows, time points for the telescoping window approach (right column) are the end of the telescoping windows. Top row shows empirical parameter estimates, middle and bottom rows show estimates made from simulated outbreaks. In the middle and bottom rows, the x axis is units of days since the start of the outbreak simulation. In all plots, black points represent values of $1/k$, plotted on the right vertical axes, whereas blue points represent values of R, plotted on the left vertical axes. (a,b) Estimated median and 95% CIs for $1/k$ (black, right-hand vertical axis) and R (blue, left-hand vertical axis), displayed over daily new mpox cases(pink) for the 2022 mpox epidemic in the United States. (c-f) Parameter values estimated on simulated data from a simulated outbreak which exhibited strongly heterogeneous transmission. (c-d) Estimates from simulated data where probable transmission pairs were determined with complete information on true individual contacts for 3.6% of the population, with the remainder being determined only with simulated dates and geographic locations. This reflects the condition of the empirical data in a,b. (e-f) Estimates from simulated data where all pairs were determined with complete information on true individual contacts. In all plots, values of $1/k$, instead of k, are displayed to ease visualization.

National estimates

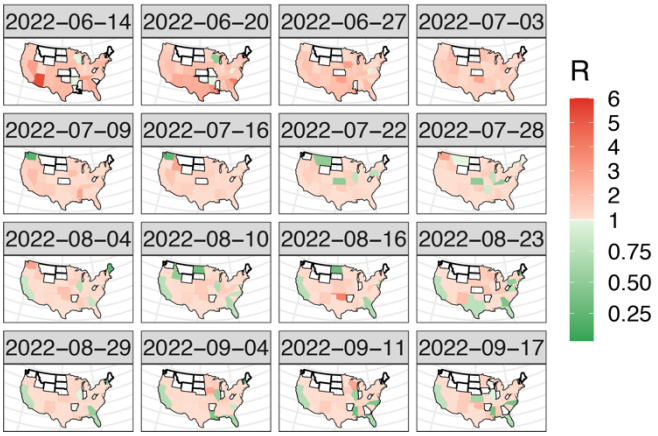
There were 22,391 probable and confirmed mpox cases in the empirical dataset, after excluding 5,074 (18.47%) cases with missing data on symptom onset date. In the exclusive dataset, 820 (3.66%) of these cases had contact tracing-identified connections to other probable and confirmed cases, compared to 12 (0.24%) of those excluded cases.

R and k estimated at the national level changed over time. Using a sliding 2-week window (Fig. 3a), we observe that, after early increase to a peak, estimated R declines over the course of the epidemic, dropping below 1 around the peak of new infections on approximately August 1st, as is expected to occur when an outbreak changes from growth to decline. Median estimated k remains above 1 for all time windows, though with high variability – a pattern which does not indicate high transmission heterogeneity.

Applying a telescoping window approach allows us to evaluate R and k taken cumulatively at different points during the epidemic (Fig 3b). Using the time window beginning at the start of the outbreak, May 1st, and ending at the peak of new cases, August 1st, we find that $R = 1.25$ (95% CI: 1.21-1.28) and $k = 12.18$ (95% CI: 8.18-21.27). Using a zero-inflated negative binomial distribution returned the same values of R and k for this period, with a p_0 value of 0.001 (95% CI: 0.00-0.02) (see supplemental results).

TEXT BOX: Figure 4: Spatiotemporal variation in estimated values of R across US states during the middle of the 2022 mpox epidemic. Unshaded regions indicate insufficient data to conduct

parameter estimations for that region and time period.



Jurisdiction estimates

Estimates of *R* varied across U.S. State and over time (Fig. 4). Estimates of *k* also vary across public health jurisdictions and over time (Table S1). While median estimates of *k* are over 1 at all time points for every jurisdiction, the variation in *k* estimates across iterations is substantial for some points. Estimates of p_0 made using the zero-inflated negative binomial were highly variable across jurisdictions and timepoints (see supplemental results).

Simulations

Conducting the methods on data resulting from an outbreak simulation over a scale-free contact network showed reliability in detecting overdispersed offspring distributions when detailed contact information was used, reflected by estimates of *k* below 1 (Fig. 3e,f).When using only network structure-based jurisdiction clusters for 96.4% of the population, and using detailed contact information for the remaining 3.6%, estimates of *k* were much higher, staying above 1 (Fig. 3c,d).

Discussion

Our results suggest that transmission of mpox during the 2022 outbreak in the United States was not highly heterogeneous. Instead, estimates of $k >> 1$ across many jurisdictions and time periods indicate that transmission was approximately what would be expected in an evenly mixing population of people with equal inherent transmissibility. We do find some indication of overdispersed transmission in certain locations during certain time periods (Table S1). Such a pattern could be consistent with explanations that invoke individual level transmission heterogeneity, resulting from individuals who spread to far more than the average infected individual and/or event-based superspreading. Whatever the case, our methods do not return evidence that such heterogeneity has been widespread. However, conducting our methods on simulated data indicates that we only expect to recover a strong pattern of heterogeneity if we have very high-quality data. Because the data quality in our empirical dataset likely varied by case, geographic location, and through time, we cannot rule out the possibility that our estimates of *k* for the 2022 US mpox epidemic are artificially high, and care must be taken with any interpretation.

There are points during the data collection process when errors may be introduced and underreporting may occur, resulting in incomplete and/or inaccurate contact tracing data. Our methods operated under a number of assumptions related to missing data. First, it was assumed that the data are a representative sample of all mpox cases. If there is an association between case detection and secondary infections produced, then this assumption could impact our estimates. For example, diagnostic testing can be biased towards cases of more severe disease (Tsang et al. 2021), which could produce more secondary infections through an increased duration of viral shedding (Fielding et al. 2014). Second, we excluded the cases lacking symptom onset dates, and we assumed that the cases for which date of symptom onset exist constitute a random sample from the population. If there is a relationship between having the symptom onset date for a case and some other factor (e.g., information on contacts), then our results could be impacted.

Incomplete contact tracing data can be the result of various errors introduced during the data collection process. For the 2022 mpox outbreak we focused on the undercounting of secondary transmissions manifesting as extra zeros in the offspring distribution, which are mainly introduced in the following three ways. First, inconsistencies in the dates associated with cases could impact our analyses. Date of reporting could occur at any time during the course of an infection, even after the patient is no longer infectious, and often includes “data dumps” where particular jurisdictions may have posted large portions of their data on a given day. While we use date of symptom onset, it is possible that these dates were influenced by behavioral or systematic processes in healthcare and disease progression . Even for cases with a date of symptom onset we cannot be sure those dates are consistent with the start of the infectious period. For the most recent cases in the dataset, it may be too close to the end of the dataset to know whether any of their contacts had become infected. Second, there are barriers to a suspected case becoming a probable or confirmed case – tests may only be provided to verified contacts of another probable or confirmed case or to those in high-risk subpopulations. In the extreme, this feature of the contact tracing process can result in missing entire transmission chains. Third, identifying case-by-case contacts and their exact dates can be difficult in sexual/social networks such as those where mpox has been shown to spread. Infected individuals may not know or choose to disclose the identities of their close contacts and additionally may delay seeking care due to historical discrimination and stigmatization of the LGBTQ+ community in the healthcare setting (Ennab et al. 2022; Rodríguez et al. 2022; The Lancet Regional Health – Europe 2022). The spatiotemporal variability in estimates of p_0 may reflect variation in data incompleteness related to this potential zero inflation. However, overall at the national level, estimates of p_0 were consistently low, and our sensitivity analysis using the zero-inflated negative binomial distribution returned very similar parameter estimates to those returned using the negative binomial distribution. In a large population, the influence of missing case data on estimates of R and k may be minimized or obscured.

One reason that transmission heterogeneity is of potential interest to public health decision making is that it can indicate that within-host processes during infection result in inherent differences in the probability of infection given contact across infected people (Ke et al. 2022). In such a situation, equal rates of contact between infected and susceptible people will result in different numbers of secondary infections, and interventions, when implemented, may differ in their efficacy relative to expectations. Additionally, we should expect to find consistently low values of k across spatiotemporal dimensions. In the case of the 2022 mpox outbreak in the United States, we did not find evidence that this sort of individual-level transmission

heterogeneity was broadly present. Instead, the rare occurrences of highly heterogeneous transmission that we observe likely resulted from differences in contact rates, either due to standing individual variation in contact rates or to the occurrence of superspreading events. In this case, targeted interventions could be employed to reduce the spread of disease. Without high heterogeneity in transmission, the rate of disease spread is well-characterized by an R_0 value greater than the threshold 1, as has been suggested by other recent studies of R_0 for the 2022 mpox epidemic (Branda, Pierini, and Mazzoli 2022; Endo et al. 2022). These results are consistent with the theory that cross-protective immunity from smallpox may be waning, resulting in higher rates of overall transmission (Grant, Nguyen, and Breban 2020).

In light of past work which estimated low values of k for mpox outbreaks, it is surprising that we did not recover strong evidence of high transmission heterogeneity for this mpox epidemic. The population structure, timing, environmental characteristics, and primary mechanisms of transmission differed between the 2022 mpox outbreak in the United States and the outbreaks described in other studies. It is possible that these differences explain the differences in estimates of k between our study and the others. The timing of the US outbreak relative to the beginning of the global epidemic may also have afforded individuals with a higher potential risk of infection the time necessary to take precautions against infection. This may be especially true as it relates to large events, which were implicated in the initial spread outside the United States and were the focus of targeted public health messaging in the United States (Delaney et al. 2022). Indeed, survey results suggest that many US gay and bisexual men, and nonbinary and transgender people took active measures to reduce the risk of exposure to mpox (Delaney et al. 2022), and so this could explain the values of k we observed. It is important to recognize that, while it is plausible that the k estimates recovered by our methodology accurately reflect real patterns of transmission heterogeneity, the limited ability to detect cases and track their contacts during the 2022 mpox outbreak in the United States could impact our estimates of both R and k.

Molecular, phylogenetic, and in-vivo studies have indicated that monkeypox virus samples from the outbreak in this study are significantly divergent from those of past outbreaks (Isidro et al. 2022, Americo et al. 2023). The outbreak investigated in the current study is caused by a virus that might have quite different epidemiological characteristics than the virus associated with outbreaks upon which past estimates of epidemiological parameters have been based. Our methods can be extended to clarify the differences between our estimates of k and those of past studies of the same disease.

Our study contributes to a small but growing literature on spatiotemporal variation in transmission heterogeneity of human infectious diseases. Past research on transmission heterogeneity has largely focused on small outbreaks or has used data from the first generations of infections in larger outbreaks. Related research on the influence of contact networks on the spread of infectious diseases has illuminated important concepts regarding transmission heterogeneity. For example, as the variation in the degree distribution of the infected population changes over the course of an epidemic simulated on a network with a particular underlying structure (Bansal, Grenfell, and Meyers 2007), transmission heterogeneity should be expected to change over time.

Our methods used a straightforward algorithm to determine likely infectors for cases without a

contact tracing-identified infector, and we subjected these methods to a limited sensitivity analysis using simulated data. In the process, we determined that our methods can be expected to accurately estimate *k* with high-quality data on the direct contacts of each case. In some instances, our empirical data may meet these qualifications, but in other instances they likely do not. So, it is possible for our methods to mask high transmission heterogeneity under some conditions. One study has found that different models used to measure transmission heterogeneity may result in different estimates with starkly different interpretations, with their instant-individual heterogeneity model recovering high heterogeneity when alternate models do not (Zhang, Britton, and Zhou 2022). The methods we present here are adaptable to incidence time series data, holding promise that transmission heterogeneity may be reliably and easily estimated, but care must be taken to contextualize findings within the caveats of the data and the potential limitations of data gathered during emerging outbreaks.

Acknowledgements

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

Literature Cited

Adler H, Gould S, Hine P, Snell LB, Wong W, Houlihan CF, et al. 2022 "Clinical features and management of human monkeypox: a retrospective observational study in the UK." *The Lancet Infectious Diseases.* 22(8):1153–62.

Americo JL, Earl PL, Moss B. 2023 "Virulence differences of mpox (monkeypox) virus clades I, IIa, and IIb.1 in a small animal model." *Proc Natl Acad Sci USA*. 21;120(8):e2220415120.

Althouse, Benjamin M., Edward A. Wenger, Joel C. Miller, Samuel V. Scarpino, Antoine Allard, Laurent Hebert-Dufresne, and Hao Hu. 2020. "Superspreading Events in the Transmission Dynamics of SARS-CoV-2: Opportunities for Interventions and Control." *PLoS Biology* 18 (11): 1–13. https://doi.org/10.1371/journal.pbio.3000897.

Bansal, Shweta, Bryan T. Grenfell, and Lauren Ancel Meyers. 2007. "When Individual Behaviour Matters: Homogeneous and Network Models in Epidemiology." *Journal of the Royal Society Interface* 4 (16): 879–91. https://doi.org/10.1098/rsif.2007.1100.

Barabási, Albert-László, and Réka Albert. 1999. "Emergence of Scaling in Random Networks." *Science* 286 (5439): 509–12. https://doi.org/10.1126/science.286.5439.509.

Beer, Ellen M., and V. Bhargavi Rao. 2019. "A Systematic Review of the Epidemiology of Human Monkeypox Outbreaks and Implications for Outbreak Strategy." *PLoS Neglected Tropical Diseases* 13 (10): 1–20. https://doi.org/10.1371/journal.pntd.0007791.

Bjørnstad, Ottar N., Bärbel F. Finkenstädt, and Bryan T. Grenfell. 2002. "Dynamics of Measles Epidemics: Estimating Scaling of Transmission Rates Using a Time Series SIR Model." *Ecological Monographs* 72 (2): 169–84. https://doi.org/10.1890/0012-9615(2002)072[0169:DOMEES]2.0.CO;2.

Blondel, Vincent D, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. "Fast Unfolding of Communities in Large Networks." *Journal of Statistical Mechanics: Theory and Experiment* 2008 (10): P10008. https://doi.org/10.1088/1742-

5468/2008/10/P10008.
Blumberg, Seth, and J. O. Lloyd-Smith. 2013. "Inference of R0 and Transmission Heterogeneity from the Size Distribution of Stuttering Chains." *PLoS Computational Biology* 9 (5): 1–17. https://doi.org/10.1371/journal.pcbi.1002993.
Branda, Francesco, Massimo Pierini, and Sandra Mazzoli. 2022. "Monkeypox: Early Estimation of Basic Reproduction Number R 0 in Europe." *Journal of Medical Virology*. https://doi.org/10.1002/jmv.28270.
Breman, J. G., K. Ruti, M. V. Steniowski, E. Zanotto, A.I. Gromyko, and I. Arita. 1980. "Human Monkeypox, 1970-79." *Bulletin of the World Health Organization* 58 (2): 165–82.
Bunge, Eveline M., Bernard Hoet, Liddy Chen, Florian Lienert, Heinz Weidenthaler, Lorraine R. Baer, and Robert Steffen. 2022. "The Changing Epidemiology of Human Monkeypox—A Potential Threat? A Systematic Review." *PLoS Neglected Tropical Diseases* 16 (2): 1–20. https://doi.org/10.1371/journal.pntd.0010141.
Centers for Disease Control and Prevention. 2022. "CDC Monkeypox Response: Transmission - Media Statement," 1–2.
_____. 2022. "Case Definitions for Use in the 2022 Mpox Response" https://www.cdc.gov/poxvirus/mpox/clinicians/case-definition.html
Csardi G, Nepusz T. 2006. "The igraph software package for complex network research." InterJournal, Complex Systems, 1695. https://igraph.org.
Delaney, Kevin P., Travis Sanchez, Marissa Hannah, O. Winslow Edwards, Thomas Carpino, Christine Agnew-Brune, Kaytlin Renfro, et al. 2022. "Strategies Adopted by Gay, Bisexual, and Other Men Who Have Sex with Men to Prevent Monkeypox Virus Transmission — United States, August 2022." *MMWR. Morbidity and Mortality Weekly Report* 71 (35): 1126–30. https://doi.org/10.15585/mmwr.mm7135e1.
Durski, Kara N, Andrea M Mccollum, Yoshinori Nakazawa, Brett W Petersen, and Mary G Reynolds. 2018. "Emergence of Monkeypox in West Africa and Central Africa, 1970–2017." *Releve Epidemiologique Hebdomadaire* 93 (11): 125–32.
Elie, Baptiste, Christian Selinger, and Samuel Alizon. 2022. "The Source of Individual Heterogeneity Shapes Infectious Disease Outbreaks." *Proceedings of the Royal Society B: Biological Sciences* 289 (1974). https://doi.org/10.1098/rspb.2022.0232.
Endo, Akira, Hiroaki Murayama, Sam Abbott, Ruwan Ratnayake, Carl A. B. Pearson, W. John Edmunds, Elizabeth Fearon, and Sebastian Funk. 2022. "Heavy-Tailed Sexual Contact Networks and the Epidemiology of Monkeypox Outbreak in Non-Endemic Regions, May 2022." *MedRxiv*, no. May: 2022.06.13.22276353.
Ennab, Farah, Faisal A. Nawaz, Kapil Narain, Goodluck Nchasi, and Mohammad Yasir Essar. 2022. "Rise of Monkeypox: Lessons from COVID-19 Pandemic to Mitigate Global Health Crises." *Annals of Medicine and Surgery* 79 (June): 14–17. https://doi.org/10.1016/j.amsu.2022.104049.
Grant, Rebecca, Liem Binh Luong Nguyen, and Romulus Breban. 2020. "Modelling Human-to-Human Transmission of Monkeypox." *Bulletin of the World Health Organization* 98 (9): 638–40. https://doi.org/10.2471/BLT.19.242347.
Handcock, Mark S., and James Holland Jones. 2006. "Interval Estimates for Epidemic Thresholds in Two-Sex Network Models." *Theoretical Population Biology* 70 (2): 125–34. https://doi.org/10.1016/j.tpb.2006.02.004.
Haydon, D.T., M. Chase-Topping, D. J. Shaw, L. Matthews, J. K. Friar, J. Wilesmith and M. E. J. Woolhouse. 2003. "The Construction and Analysis of Epidemic Trees with Reference

549 to the 2001 UK Footand-Mouth Outbreak." *Proceedings of the Royal Society: B*
550 270(1511): 121-127. https://doi.org/10.1098/rspb.2002.2191
551 Heijne, Jannek C.M., Marc Rondy, Linda Verhoef, Jacco Wallinga, Mirjam Kretzschmar, Nicola
552 Low, Marion Koopmans and Peter F. M. Teunis. 2012. "Quantifying Transmission of
553 Norovirus During an Outbreak." *Epidemiology* 23:277-284.
554 https://doi.org/10.1097/EDE.0b013e3182456ee6
555 Isidro J, Borges V, Pinto M, Sobral D, Santos JD, Nunes A, et al. 2022. "Phylogenomic
556 characterization and signs of microevolution in the 2022 multi-country outbreak of
557 monkeypox virus." *Nat Med*. 28(8):1569–72.
558 Ke, Ruian, Pamela P. Martinez, Rebecca L. Smith, Laura L. Gibson, Agha Mirza, Madison
559 Conte, Nicholas Gallagher, et al. 2022. "Daily Longitudinal Sampling of SARS-CoV-2
560 Infection Reveals Substantial Heterogeneity in Infectiousness." *Nature Microbiology* 7
561 (5): 640–52. https://doi.org/10.1038/s41564-022-01105-z.
562 Kucharski, A. J., and C. L. Althaus. 2015. "The Role of Superspreading in Middle East
563 Respiratory Syndrome Coronavirus (MERS-CoV) Transmission." *Eurosurveillance* 20
564 (25): 14–18. https://doi.org/10.2807/1560-7917.ES2015.20.25.21167.
565 Lau, Max S.Y., Bryan Grenfell, Michael Thomas, Michael Bryan, Kristin Nelson, and Ben
566 Lopman. 2020. "Characterizing Superspreading Events and Age-Specific Infectiousness
567 of SARS-CoV-2 Transmission in Georgia, USA." *Proceedings of the National Academy*
568 *of Sciences of the United States of America* 117 (36): 22430–35.
569 https://doi.org/10.1073/pnas.2011802117.
570 Lloyd-Smith, J. O. 2007. "Maximum Likelihood Estimation of the Negative Binomial
571 Dispersion Parameter for Highly Overdispersed Data, with Applications to Infectious
572 Diseases." *PLoS ONE* 2 (2): 1–8. https://doi.org/10.1371/journal.pone.0000180.
573 Lloyd-Smith, J. O., S. J. Schreiber, P. E. Kopp, and W. M. Getz. 2005. "Superspreading and the
574 Effect of Individual Variation on Disease Emergence." *Nature* 438 (7066): 355–59.
575 https://doi.org/10.1038/nature04153.
576 Madewell, Zachary J, Kelly Charniga, Nina Masters, and Et. Al. 2022. "Serial Interval and
577 Incubation Period Estimates of Monkeypox Virus Infection in 12 U.S. Jurisdictions, May
578 – August 2022." *MedRxiv*. https://doi.org/10.1101/2022.10.26.22281516.
579 May, Robert M., and Roy M. Anderson. 1984. "Spatial Heterogeneity and the Design of
580 Immunization Programs." *Mathematical Biosciences* 72 (1): 83–111.
581 https://doi.org/10.1016/0025-5564(84)90063-4.
582 Miyama, Takeshi, Sung-mok Jung, and Hiroshi Nishiura. 2022. "Decrease in Overdispersed
583 Secondary Transmission of COVID-19 over Time in Japan." *Epidemiology and Infection*,
584 November, 1–29. https://doi.org/10.1017/S0950268822001789.
585 Nalca, Aysegul, Anne W. Rimoin, Sina Bavari, and Chris A. Whitehouse. 2005. "Reemergence
586 of Monkeypox: Prevalence, Diagnostics, and Countermeasures." *Clinical Infectious*
587 *Diseases* 41 (12): 1765–71. https://doi.org/10.1086/498155.
588 Nishiura, H., Y. Miyamatsu, G. Chowell, and M. Saitoh. 2015. "Assessing the Risk of Observing
589 Multiple Generations of Middle East Respiratory Syndrome (MERS) Cases given an
590 Imported Case." *Eurosurveillance* 20 (27): 6–11. https://doi.org/10.2807/1560-
591 7917.ES2015.20.27.21181.
592 Philpott, David, Christine M Hughes, Karen A Alroy, Janna L Kerins, Jessica Pavlick, and
593 Lenore Asbel. 2022. "Epidemiologic and Clinical Characteristics of Monkeypox Cases
594 — United States, May 17-July 22, 2022." *Morbidity and Mortality Weekly Report* 71

(32): 1018–22.
R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.
Ridout, Martin, Clarice G B Demetrio, and John Hinde. 1998. "Models for Count Data with Many Zeros." In *Proceedings of the XIXth International Biometric Conference*, 19:179–92. International Biometric Society Invited Papers. Cape Town, South Africa.
Rimoin, Anne W., Prime M. Mulembakani, Sara C. Johnston, James O. Lloyd Smith, Neville K. Kisalu, Timothee L. Kinkela, Seth Blumberg, et al. 2010. "Major Increase in Human Monkeypox Incidence 30 Years after Smallpox Vaccination Campaigns Cease in the Democratic Republic of Congo." *Proceedings of the National Academy of Sciences of the United States of America* 107 (37): 16262–67. https://doi.org/10.1073/pnas.1005769107.
Rodríguez, Berta Suárez, Bernardo R Guzmán Herrador, Asunción Díaz Franco, María Paz Sánchez-seco Fariñas, Amo Valero, Adrián Hugo, Aginagalde Llorente, et al. 2022. "Epidemiologic Features and Control Measures during Monkeypox" 28 (9): 1847–51.
Schneckenreither, Günter, Lukas Herrmann, Rafael Reisenhofer, Nikolas Popper, and Philipp Grohs. 2022. "Assessing the Heterogeneity in the Transmission of Infectious Diseases from Time Series of Epidemiological Data." *Na* NA (NA): NA-NA.
Sklenovská, Nikola, and Marc Van Ranst. 2018. "Emergence of Monkeypox as the Most Important Orthopoxvirus Infection in Humans." *Frontiers in Public Health* 6 (September): 1–12. https://doi.org/10.3389/fpubh.2018.00241.
Spicknall, Ian H., Emily D. Pollock, Patrick A. Clay, Alexandra M. Oster, Kelly Charniga, Nina Masters, Yoshinori J. Nakazawa, Gabriel Rainisch, Adi V. Gundlapalli, and Thomas L. Gift. 2022. "Modeling the Impact of Sexual Networks in the Transmission of Monkeypox Virus Among Gay, Bisexual, and Other Men Who Have Sex with Men — United States, 2022." *MMWR. Morbidity and Mortality Weekly Report* 71 (35): 1131–35. https://doi.org/10.15585/mmwr.mm7135e2.
The Lancet Regional Health – Europe. 2022. "Lessons from COVID-19 Are Shaping the Response to Monkeypox Outbreak." *The Lancet Regional Health - Europe* 18: 100463. https://doi.org/10.1016/j.lanepe.2022.100463.
Tkachenko, Alexei V., Sergei Maslov, Ahmed Elbanna, George N. Wong, Zachary J. Weiner, and Nigel Goldenfeld. 2021. "Time-Dependent Heterogeneity Leads to Transient Suppression of the COVID-19 Epidemic, Not Herd Immunity." *Proceedings of the National Academy of Sciences* 118 (17): e2015972118. https://doi.org/10.1073/pnas.2015972118.
Toth, Damon J.A., Molly Leecaster, Warren B.P. Pettey, Adi V. Gundlapalli, Hongjiang Gao, Jeanette J. Rainey, Amra Uzicanin, and Matthew H. Samore. 2015. "The Role of Heterogeneity in Contact Timing and Duration in Network Models of Influenza Spread in Schools." *Journal of the Royal Society Interface* 12 (108). https://doi.org/10.1098/rsif.2015.0279.
Toth, Damon J.A., Windy D. Tanner, Karim Khader, and Adi V. Gundlapalli. 2016. "Estimates of the Risk of Large or Long-Lasting Outbreaks of Middle East Respiratory Syndrome after Importations Outside the Arabian Peninsula." *Epidemics* 16: 27–32. https://doi.org/10.1016/j.epidem.2016.04.002.
Tsang, Tim K., Can Wang, Bingyi Yang, Simon Cauchemez, and Benjamin J. Cowling. 2021. "Using Secondary Cases to Characterize the Severity of an Emerging or Re-Emerging Infection." *Nature Communications* 12 (1): 6372. https://doi.org/10.1038/s41467-021-

26709-7.
UK Health Security Agency. 2022. "Investigation into Monkeypox Outbreak in England: Technical Briefing 1." https://www.gov.uk/government/publications/monkeypox-outbreak-technical-briefings/investigation-into-monkeypox-outbreak-in-england-technical-briefing-1.
Wallinga, J., and P. Teunis. 2004. "Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal Similar Impacts of Control Measures." *American Journal of Epidemiology* 160 (6): 509–16. https://doi.org/10.1093/aje/kwh255.
World Health Organization. 2022a. "No Title." 2022 Mpox Outbreak: Global Trends. 2022. https://worldhealthorg.shinyapps.io/mpx_global/.
———. 2022b. "WHO Recommends New Name for Monkeypox Disease." 2022. https://www.who.int/news/item/28-11-2022-who-recommends-new-name-for-monkeypox-disease.
_____. 2022c. "Monkeypox." 2022 Fact Sheets. https://www.who.int/news-room/fact-sheets/detail/monkeypox
Yinka-Ogunleye, Adesola, Olusola Aruna, Dimie Ogoina, Neni Aworabhi, Womi Eteng, Sikiru Badaru, Amina Mohammed, et al. 2018. "Reemergence of Human Monkeypox in Nigeria, 2017." *Emerging Infectious Diseases* 24 (6): 1149–51. https://doi.org/10.3201/eid2406.180017.
Zhang, Y, T Britton, and X Zhou. 2022. "Monitoring Real-Time Transmission Heterogeneity from Incidence Data," *PLoS Comput Biol 18(12): e1010078. https://doi.org/10.1371/journal.pcbi.1010078*