How social structures, space, and behaviors shape the spread of infectious diseases using chikungunya as a case study

Henrik Salje, Justin Lessler, Paul Kishor, Andrew S. Azman, M. Waliur Rahman, Mahmudur Rahman, Derek Cummings, Emily S. Gurley, Simon Cauchemez

How social structures, space and behaviors shape the spread of infectious diseases: chikungunya as a case study

Henrik Salje[1,2,3,4], Justin Lessler[1], Kishor Paul[6], Andrew Azman[1], M. Waliur Rahman[5,6], Mahmudur Rahman[5], Derek Cummings[1,7], Emily S. Gurley[6*], Simon Cauchemez[2,3,4*]

1. *Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA*
2. *Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Paris 75015, France*
3. *Centre National de la RechercheScientifique, URA3012, Paris 75015, France*
4. *Center of Bioinformatics, Biostatistics and Integrative Biology, Institut Pasteur, Paris 75015, France*
5. *Institute of Epidemiology Disease Control & Research (IEDCR), Mohalkhali, Dhaka 1212, Bangladesh*
6. *Center for Communicable Diseases, icddr,b, 68 SaheedTajuddinAhamedSarani, Mohakhali, Dhaka 1212, Bangladesh*
7. *Department of Biology, University of Florida, Gainesville, FL 32603, USA*

*Joint contribution as senior authors

Major category: Biological sciences
Minor category: Ecology

Corresponding author: Henrik Salje, Johns Hopkins School of Public Health, 615 N. Wolfe Street, Baltimore, MD 21205, USA. Email: hsalje@jhu.edu. Tel: +33 (0)140 613 379

Keywords: epidemiology; mathematical modeling; infectious diseases; chikungunya

Abstract

Whether or not an individual becomes infected in an infectious disease outbreak depends on many interconnected risk factors, which may relate to characteristics of the individual (e.g. age, sex), their close relatives (e.g. household members) or the wider community. Studies monitoring individuals inhouseholds or schools have helped elucidate the determinants of transmission in small social structures thanks to advances in statistical modelling;butsuch approach has so far largely failed to consider individuals in the wider context they live in. Here, we used an outbreak of chikungunya in a rural community in Bangladesh as a case study to obtain a more comprehensive characterization of risk factors in disease spread. We developed Bayesian data augmentation approaches to account for uncertainty in the source of infection, recall uncertainty and unobserved infection dates. We found that the probability of chikungunya transmission was 11% (95% CI: 8%-15%) between household members but dropped to 0.3% for those living 50m away (95% 0.2-0.5%). Overall, the mean transmission distance was 95m (95% CI: 77–113m). Females were 1.5 times more likely to become infected than males (95% CI: 1.2–1.8), which was virtually identical to the relative risk of being at home estimated from an independent human movement study in the country. Reported daily use of anti-mosquito coils had no detectable impact on transmission. This study shows how the complex interplay between characteristics of the individual, their close and wider environment contributes to the shaping of infectious disease epidemics.

Significance Statement

While the determinants of infectious disease transmission have been extensively investigated in small social structures such as households or schools, impact of the wider environment (e.g. neighborhood) on transmission has received less attention. Here we use an outbreak of chikungunya as a case study where detailed epidemiological data werecollected and combine it with novel statistical approaches to characterize the multiple factors that influence the risk of infectious disease transmission and may depend on characteristics of the individual (e.g. age, sex), of their close relatives (e.g. household members) or of the wide neighborhood. Our findings highlight the role that integrating statistical approaches with in-depth information on the at-risk population can have on understanding pathogen spread.

/body

Introduction

Factors that affect the risk of pathogen infection are multiple and complex. They often intertwine features of the individual (e.g. age, behavior or mobility) with those of their social network, the wider population and, in some cases, the environment they live in. Assessing the relative contribution of these factors to transmission often proves difficult since, apart from few exceptions (1-3), it is rarely possible to directly measure individual exposures to potential sources of infection. However, recent advances in statistics and modeling now make it to possible to reconstruct such information from data gathered during outbreaks, allowing a more refined evaluation. These approaches have been extensively used to ascertain how the structure of the social network, behaviors, socio-demographic and biological factors affect the spread of pathogens in relatively small social communities such as households, hospitals or schools (2, 4-8).

Although these studies provide great details on transmission at the very local scale of a household, they have so far largely failed to consider individuals in the wider context they live in. For example, we still poorly understand how the risk of infection of an individual may be affected by the presence of cases in neighboring households or in households that are further away. It also remains unclear whether the heterogeneous mobility profiles observed in a population (e.g. children versus adults, women versus men) have any impact on individual risks of infection. As a consequence, it remains difficult to robustly calibrate spatial spread in simulation models that are used to inform policy making (9-11) resulting in predictions that may sometimes look at odds with the data (12).

Here, we take chikungunya, a mosquito-borne virus that causes fever and joint pain (13, 14), as a case study. We analyze detailed data describing a chikungunya

outbreak in a rural community in Bangladesh to obtain a more comprehensive view of infection risk factors considering the different environments an individual interacts with: from their household, to their neighborhood and the wider community. We evaluate the influence of spatial proximity on the risk of transmission and, comparing our findings with nationally representative human mobility data, evaluate whether different mobility profiles may correlate with different individual's risk of infection. The analysis requires the development of sound Bayesian data augmentation statistical techniques (6, 15) to account for uncertainty in the source of infections, recall uncertainty and unobserved infection dates. Such uncertainties are typical in outbreak scenarios.

Results

In 2012 an outbreak of chikungunya was reported in the village of Palpara in Tangail district, 100km northwest of the capital, Dhaka. An outbreak investigation team was deployed at the end of November by the governmental outbreak response team at the Institute for Epidemiology, Disease Control and Research in collaboration with the icddr,b. The outbreak investigation team visited every single household in the outbreak village and interviewed 1933 individuals from 460 households. 364 (18%) individuals reported having suffered from symptoms consistent with chikungunya infection (the case definition was fever with either joint pain or a rash) between 29 May and 1 December 2012. Chikungunya infection was confirmed using serology in a subset of 175 cases. The mean age of cases was 30 (range: 0 – 80) and 958 (57%) of cases were female (Figure 1). Sixty four per cent of individuals (N=1,238) lived in households that reported using anti-mosquito coils on a daily basis.

We built a transmission model to ascertain transmission risk factors. All individuals that met the case definitionwere included as cases in the analysis. Data augmentation techniques were used to incorporate both onset date uncertainty and

the unobserved infection dates. We used an exponentially distributed kernel to characterize transmission distancesfor between-household transmissions (i.e., for pairs of individuals that live in different households) and used a separate parameter for within-household transmission (i.e. for pairs of individuals that live in the same household). We found that the probability of transmission was 12% (95% CI: 8%-17%) between household members (Figure 2A) but dropped to 0.3% for those living 50m away (95% 0.2-0.5%) and 0.2% for those 100m away (95% CI: 0.1% - 0.2%) (Figure 2B), indicating that transmission was highly focal. A sensitivity analysis using a power-law distribution resulted in almost an identical transmission kernel (Figure 2B). Females were 1.5 (95% CI 1.2 – 1.8) times more likely to get infected than males (Figure 2C). Children (defined as those under 16 years) were at similar risk as adults (relative risk of 0.9, 95% CI: 0.8 – 1.2) (Figure 2C). Reported daily use of anti-mosquito coils had no impact on transmission risk (1.0, 95% CI: 0.8 – 1.2) (Figure 2E).

To ascertain the contribution of these different factors to the overall epidemic, we probabilistically reconstructed 200 fully resolved transmission trees consistent with the data (Figure 3A). Analysis of these trees indicates that household transmissions represented 27% of all transmission events (95% CI: 23% - 31%) (Figure 3B). Fifty-eight per cent of transmissions (95% CI: 51% - 65%) occurred at the neighborhood level (defined here as within 200m of a home, an area that consisted of 27% of the population on average) while only 15% of transmission (95% CI: 9% - 21%) occurred in the wider community (>200m) despite 73% of the population living this far away from cases. Overall the mean transmission distance was 95m (95% CI: 77m – 113m). Neighborhood transmission was the largest contributor to the effective reproductive number (Figure 3C). We calculated the basic reproductive number for each individual based on where they lived and the individual characteristics of the community. We then mapped how the basic reproductive number differed over the study area. We found significant spatial heterogeneity that was consistent with where the majority of infections occurred (Figure S5).As the transmissibility of a pathogen may change over time, especially with vector-mediated pathogens that

may have strong seasonal drivers, we allowed a step-change in transmissibility and estimated both the timing and the magnitude of the change. We estimated that on the 10 October 2012 (95% CI: 5 October – 13 October), the probability of transmission fell by 74% (95% CI: 63% – 84%).

To assess model performance, we simulated epidemics starting from 1 August using our estimated parameters for the outbreak. At this time, eight cases had occurred. We found that both the temporal trajectory (Figure 4A) and the spatial spread of infections (Figure 4B and Figure 4C) were consistent with that observed. The simulations resulted in a mean of 475 cases (95% CI: 258-670) compared to 364 observed cases.

To explore whether the increased risk of infection for females was due to spending more time at home, we compared our results to that from a separate, nationally representative, human movement study that we conducted of 52 rural populations in Bangladesh using GPS monitors (see methods). Overall 380 individuals' monitors returned useable data. Individuals spent an average of 56% of their time between the hours of 8am and 8pm within or around their homes (defined as within 50m of the central coordinates of their home). However, this differed greatly by sex. We found that females were 1.5 (95% CI: 1.4 – 1.6) times more likely to be in and around their home compared to males (66% of time at home for females versus 45% for males) (Figure 2D). Children (those under 16 years) were 0.9 (95% CI: 0.8–1.0) times as likely to be in and around their home than adults (Figure 2D). These findings are completely consistent with the findings of relative risk of infection in our model (Figure 2E), suggesting the increased time females spent in and around the home may have been responsible for their increased risk of infection.

Not all infection events are likely to have been detected. Infections may not have resulted in symptoms that met the case definition or may have caused no symptoms at all (16, 17). Further, individuals may have forgotten more mild febrile episodes.

To assess the impact of these undetected infections to our estimates, we simulated outbreaks based on the spatial structure of our study population and randomly assigned 0% (to reflect outbreaks with no undetected infections), 20%, 40%, 60% or 80% of cases as unobserved infections. We then estimated the parameters using only observed cases. We found that in these scenarios, all model parameters could be accurately estimated except the mean transmission distance, which was slightly overestimated (mean estimate of 170m when 40% of cases were undetected compared to a true value of 140m) and the household force of infection (resulting in a mean estimate of 9% of infections as household infections when 40% of cases were detected for a true value of 13%) (Table S1). To explore the impact of over-estimating the transmission kernel, we compared the spatial spread of cases in simulations that used kernels with mean transmission distances ranging from 125m to 200m. We found that for the range of kernels explored the spatial and temporal distribution remained similar (Figure S1).

Where the proportion of undetected infections is known, reversible-Jump MCMC (RJ-MCMC) methods can be used to account for undetected infections when estimating parameters (18).Using this approach, we found that in scenarios where up to 40% of cases were undetected we could accurately estimate parameters, including both the transmission kernel parameter and the household force of infection (Table S1). The performance of the model diminished when a greater proportion of cases were undetected. The RJ-MCMC model was able to accurately estimate the transmission kernel parameter across a range of simulated values (Table S2). Applying RJ-MCMC to the outbreak data where 20% were assumed to be undetected resulted in a shorter mean transmission distance of 80m (70-100m) with 32% of infections occurring within the home. Increasing the number of undetected infections to 40% gave a mean transmission distance of 70m (60-90m) with 36% of infections occurring in the home. All other parameter estimates were essentially unchanged (Table S3).

Discussion

Epidemic spread is driven by a complex interplay of individual actions and local environment. Statistical methods developed to reconstructtransmission trees from incomplete outbreak dataprovide an invaluable tool to help disentangle these factors. Previous attempts to reconstruct infectious disease transmission trees have been largely restricted to highly structured communities such as schools, hospitals or households (2, 6, 19). Here, we incorporated the wider context of their local environment. Using chikungunya as a case study, we have shown that we can combine detailed epidemiological dataand mathematical models to gain insight into detailed dynamics of disease spread in a wider community. We have demonstrated that individual characteristics (e.g., sex) and local environment, in particular where individuals live relative to cases, has a critical impact on risk of infection. Further, we have shown through an independent human mobility dataset that these risk differences are entirely consistent with individual level differences in movement behavior. This highlights the importance of incorporating local context into assessments of outbreak spread.

This study illustrates the many challenges epidemiologists studying infectious disease transmission are confronted with when working on real-world outbreak data. During outbreak investigations, it is common that transmission pathways or dates of infection cannot be documented; or that cases misremember when they were sick. The data augmentation strategies we relied on make it possible to properly account for these uncertainties in the inferential framework, therefore greatly enhance our ability to analyze outbreak data in a robust fashion.

The collection of fine scale location data can greatly aid outbreak investigations. A major strength of our approach is that we do not have to rely on the assumption that individuals are uniformly distributed on the landscape but instead do take into account the exact locations where individuals reside to estimate the spatial kernel. It is important to note that we cannot infer the exact location of any transmission event, for example whether it occurred indoors or outdoors.

We found that in this outbreak, viral spread was largely driven by transmissions at distances not much further than neighboring households. Human mobility in rural Bangladesh is very limited with individuals spending >50% of the time in and around the home. Females in particular spend the vast majority of their day around their homes. These human mobility patterns were consistent with our estimates of the spread of chikungunya and could explain the higher risk of infection observed in females. Release-recapture experiments have demonstrated that the *Aedes* mosquito, responsible for chikungunya and dengue transmission, does not travel very far and often stays within the same residence for days (20). For viral infections to spread over small distances as observed here may require human movement.

We did not find evidence of protection from the use of anti-mosquito coils.The coils used by this community may not sufficiently reduce mosquito levels to prevent transmission. This is consistent with a recent meta-analysis that found that anti-mosquito coils did not reduce the risk of dengue infection, another virus spread by the same vector(21).However, both the meta-analysis and a similar review of vector-based strategies concluded that the evidence base for the impact of coils and other forms of vector control remained weak (22). More field-based studies are required to properly understand the potential of coil-based and other forms of vector control in different settings. Where more effective insecticides or other spatially targeted interventions are available, our findings suggest that deploying them in neighboring households of cases may be sufficient to reduce viral spread. This requires early detection of the outbreak.

We estimated that transmission decreased substantially in the beginning of October. This coincided with a steep change in mean temperatures, which dropped from 29°C at the end of September to 22°C by early November and 17°C by the start of December (Figure S1). Rainfall also decreased substantially in October (Figure S1). This is consistent with previous findings of a key role of temperature and rainfall on

chikungunya risk (23). In addition to the role of climate, the build up in immunity in asymptomatic individuals may have contributed to this fall in transmissibility.

The outbreak investigation was conducted two months after the peak of the epidemic. Individuals are unlikely to precisely remember when they started to have symptoms. However, by using data augmentation techniques we were able to incorporate recall uncertainty into our estimates. The case definition we used was specific for chikungunya. While we cannot rule out false positive cases, these are likely to be minimal and not impact our parameter estimates. The case definition may have resulted in missed cases. However, we have demonstrated the robustness of our model to substantial misspecification. Households may have increased their use of anti-mosquito coils since the outbreak. Any such change would potentially falsely hide any impact of the coils. We also do not know how households used the coils or the precise type.Human mobility data was not collected in the outbreak community. Future outbreak investigations could incorporate movement diaries or GPS monitors into their investigations to better understand the role of human movement in pathogen spread. It is noteworthy that the patterns observed at the national level were consistent with our model estimates.

To characterize the complex interplay of the multifaceted risk factors that shape the spread of infectious diseases, modern epidemiology needs to move away from simple case counting. Instead, it must take an integrative approach where thorough field investigations benefit from technological advances such as global positioning systems and where data interpretation is considerably strengthened by the use of innovative statistical and modelling techniques. These technological and methodological advances open a new exciting era for infectious disease epidemiologists that can and should use the framework proposed here to study the spread of other pathogens.

Materials and Methods

Data collection

An outbreak investigation team was deployed at the end of November by the governmental outbreak response group in collaboration with the icddr,b. The team visited each household in all the villages and interviewed all household members that agreed to participate. The study team recorded whether individuals reported symptoms consistent with chikungunya (fever with either joint pain or a rash) and the date of fever onset. In addition, they recorded the age and gender of all household members and whether the household reported the use of anti-mosquito coils on a daily basis. The GPS location of all homes was also recorded. To confirm that the outbreak was due to chikungunya, infection was confirmed using IgM ELISA in a subset of 175 cases (SD BIOLINE, Korea).

Statistical model

Assuming that individuals who reported symptoms had been infected with chikungunya virus, we built a statistical model to ascertain risk factors for transmission (6, 24). In particular, the model was used to estimate the role that the location and structure of households, sex, age and anti-mosquito coils had on transmission dynamics.

The force of infection exerted on individual *i* at time *t* is:

$$\lambda_i(t) = \sum_{j:t_j<t}^{N} \lambda_{j\to i}(t|x_j, x_i)$$

where $\lambda_{j\to i}(t|x_j, x_i)$ is the hazard that individual *j* transmits to individual *i* at time *t*.

$$\lambda_{j\to i}\left(t\middle|x_j, x_i\right) = \beta\left(x_i, x_j\right) \cdot f(t - t_j) \cdot \delta(t) \quad [1]$$

where$\beta\left(x_i, x_j\right)$represents the transmission rate between individuals *j* and *i.* Where *i*and *j* reside in the same household:

$$\beta\left(x_i, x_j\right) = \beta_H \cdot \beta_{sex}\left(x_i\right) \cdot \beta_{age}\left(x_i\right) \cdot \beta_{control}\left(x_i\right)$$

whereβ_{sex} characterizes the role of sex on risk of infection (male is the reference group), β_{age} characterizes the role of age on risk of infection (individuals over the age of 16 are the reference group), $\beta_{control}$ characterizes whether the household reported daily use of anti-mosquito coils (no coil use is the reference group). Where *i*and *j*reside in different households:

$$\beta\left(x_i, x_j\right) = \beta_c \cdot g(x_i, x_j) \cdot \beta_{sex}\left(x_i\right) \cdot \beta_{age}\left(x_i\right) \cdot \beta_{control}\left(x_i\right)$$

where$g(x_i, x_j)$ characterizes the transmission kernel for individuals living in different households and is a function of the distance between the households. We used an exponential distribution to characterize the transmission kernel. In addition, we performed a sensitivity analysis using a power law kernel that allowed a fatter-tailed distribution:

$$g(x_i, x_j) = \frac{1}{\left(1 + \frac{d_{ij}}{1000}\right)^{\alpha}}$$

whered_{ij} is the distance (in meters) of individuals *i* and *j* andα was estimated; $f(t - t_j)$ represents the infectivity of individual *j* over time and can be approximated by the generation time distribution (the time between two successive infections). In chikungunya it is made up of the incubation time in the individual, the duration during which the individual can transmit to a mosquito and the duration of infectiousness in the mosquito. We derived a generation time distribution with mean of 14 days and variance of 41 days. Details of the derivation can be found in the supplementary materials. Finally, we consider the possibility that

transmissibility may have changed over time as may occur where local climate (or other) conditions alter the transmissibility of the pathogen. We estimate both the timing (through a change-point parameter τ) of a change and the magnitude (through parameter β_{late}). Coefficient $\delta(t)$ is equal to one before change point τ and to β_{late} after change-point τ.

The effective reproductive number R for individual *j* early in the epidemic (i.e., before change point τ) is the sum of the beta terms:

$$R_j = \sum_{i \neq j} \beta(x_i, x_j)$$

Estimation

Parameters were estimated within a Bayesian Markov chain Monte Carlo (MCMC) framework. We only observed dates of symptom onset, not when infections occurred. In addition there may have been uncertainty in the recollection of precise dates of symptom onset. In order to account for these limitations, Bayesian data augmentation techniques were used (6, 15) whereby true dates of symptom onset and dates of infection were considered as augmented data (i.e. nuisance parameters) of the inferential framework. The joint posterior distribution of augmented data and model parameters is proportional to

$$P(z, \theta|y) \propto P(y|z) \cdot P(z|\theta) \cdot P(\theta)$$

where$y$ the observed data, z is the augmented data, and θ the parameter vector. $P(y|z)$is the observation model and assumes that (1) the error with which individuals estimated their date of symptom onset was normally distributed with mean zero and standard deviation of three days and (2) that the incubation period of chikungunya was exponentially distributed with a mean of three days (25).$P(z|\theta)$represents the transmission model characterized by Equation 1. Finally,

the prior distribution of the parameters is provided by$P(\theta)$. The joint posterior distribution is explored using Markov chain Monte Carlo sampling. Additional details about the model and estimation are given in SI Materials and Methods.

Prior distributions

For all parameters except for the transmission kernel parameter, we used a lognormal prior distribution with a log(mean) equal to zero and a log(variance) equal to one. For the transmission kernel parameter we used an exponential prior distribution with parameter of 0.0001.

MCMC sampling scheme

The MCMC sampling scheme we implemented consisted of (a) Metropolis-Hastings update for the parameters in the model; (b) Independence sampler for the infection day for fifty randomly chosen cases and (c) Independence sampler for the true onset date (to account for recall uncertainty) for fifty randomly chosen individuals. Metropolis-Hastings updates were performed on a log-scale with the step size adjusted to achieve an acceptance probability between 20% and 30%.

Climate data

We obtained 3-hourly temperature data for Tangail district from the national meteorological department of Bangladesh. From these data we calculated daily mean temperature. We also collected daily rainfall data. From this we calculated the mean number of rainfall in each two week period over the study period.

Collection of human movement data

In order to quantify the time individuals spend in and around their homes, we separately conducted a separate field study in 52 randomly selected rural

communities from throughout Bangladesh. In each community, up to ten individuals of all ages were randomly selected and asked to carry a small GPS device (IgotU GT-600 (http://www.i-gotu.com/)) that collected their location every two minutes for a period of up to four days. We also collected the home location of each participant. For each reading from the GPS device, we calculated the distance a participant was from their home. Further details on the collection of human movement data can be found in SI Materials and Methods.

Ethical approval

The outbreak investigation was exempt from IRB review. The Government of Bangladesh reviewed and approved of the investigation protocol and participants provided informed consent for participation. For the human mobility study, informed consent was obtained from all individuals and their parents or guardians for those under the age of 18. The study was approved the institutional review board of the icddr,b.

Acknowledgement

This study was funded by Centers for Disease Control and Prevention (CDC), Atlanta, USA under the cooperative agreement (grant number 5U01CI000628). HS, JL and DC would also like to recognize funding from the National Institutes of Health (Grant number R01 AI102939-01A1). The authors would like to thank Dr. FarhanaHaque from icddr,b and IEDCR for helping gain access to climate data. icddr,b is grateful to the Governments of Bangladesh, Canada, Sweden and the UK for providing core/unrestricted support.icddr,b acknowledges with gratitude the commitment of CDC and NIH to its research efforts.

References

1. Faye O, et al. (2015) Chains of transmission and control of Ebola virus disease in Conakry, Guinea, in 2014: an observational study. *Lancet Infect Dis* 15(3):320–326.

2. Assiri A, et al. (2013) Hospital outbreak of Middle East respiratory syndrome coronavirus. *N Engl J Med* 369(5):407–416.

3. Vijayaraghavan R, Chandrashekhar R, Sujatha Y, Belagavi CS (2006) Hospital outbreak of atypical mycobacterial infection of port sites after laparoscopic surgery. *J Hosp Infect* 64(4):344–347.

4. Tsang TK, et al. (2014) Association between antibody titers and protection against influenza virus infection within households. *J Infect Dis* 210(5):684–692.

5. Lau MSY, Cowling BJ, Cook AR, Riley S (2015) Inferring influenza dynamics and control in households. *Proc Natl Acad Sci USA* 112(29):9094–9099.

6. Cauchemez S, et al. (2011) Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza. *Proc Natl Acad Sci USA* 108(7):2825–2830.

7. Lessler J, et al. (2009) Outbreak of 2009 pandemic influenza A (H1N1) at a New York City school. *N Engl J Med* 361(27):2628–2636.

8. Zelner JL, King AA, Moe CL, Eisenberg JNS (2010) How infections propagate after point-source outbreaks: an analysis of secondary norovirus transmission. *Epidemiology* 21(5):711–718.

9. Ferguson NM, et al. (2005) Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* 437(7056):209–214.

10. Longini IM, et al. (2005) Containing pandemic influenza at the source. *Science* 309(5737):1083–1087.

11. Ferguson NM, Donnelly CA, Anderson RM (2001) The foot-and-mouth epidemic in Great Britain: pattern of spread and impact of interventions. *Science* 292(5519):1155–1160.

12. Gog JR, et al. (2014) Spatial Transmission of 2009 Pandemic Influenza in the US. *PLoS Comput Biol* 10(6):e1003635.

13. Weaver SC, Lecuit M (2015) Chikungunya virus and the global spread of a mosquito-borne disease. *N Engl J Med* 372(13):1231–1239.

14. Staples JE, Breiman RF, Powers AM (2009) Chikungunya fever: an epidemiological review of a re-emerging infectious disease. *Clin Infect Dis* 49(6):942–948.

15. Cauchemez S, et al. (2006) Investigating Heterogeneity in Pneumococcal Transmission. *Journal of the American Statistical Association* 101(475):946–958.

16. Queyriaux B, et al. (2008) Clinical burden of chikungunya virus infection. *Lancet Infect Dis* 8(1):2–3.

17. Sissoko D, et al. (2007) Seroprevalence and risk factors of chikungunya virus infection in Mayotte, Indian Ocean, 2005-2006: a population-based survey. *PLoS ONE* 3(8):e3066–e3066.

18. Green PJ (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*.

19. Cauchemez S, et al. (2014) Determinants of influenza transmission in South East Asia: insights from a household cohort study in Vietnam. *PLoS Pathog* 10(8):e1004310.

20. Harrington LC, et al. (2005) Dispersal of the dengue vector Aedes aegypti within and between rural communities. *The American Journal of Tropical Medicine and Hygiene* 72(2):209–220.

21. Bowman LR, Donegan S, McCall PJ (2016) Is Dengue Vector Control Deficient in Effectiveness or Evidence?: Systematic Review and Meta-analysis. *PLoS Negl Trop Dis* 10(3):e0004551.

22. Achee NL, et al. (2015) A critical assessment of vector control for dengue prevention. *PLoS Negl Trop Dis* 9(5):e0003655.

23. Perkins TA, Metcalf CJE, Grenfell BT, Tatem AJ (2015) Estimating drivers of autochthonous transmission of chikungunya virus in its invasion of the americas. *PLoS Curr* 7. doi:10.1371/currents.outbreaks.a4c7b6ac10e0420b1788c9767946d1fc.

24. Cauchemez S, Ferguson NM (2012) Methods to infer transmission risk factors in complex outbreak data. *J R Soc Interface* 9(68):456–469.

25. Rudolph KE, Lessler J, Moloney RM, Kmush B, Cummings DAT (2014) Incubation periods of mosquito-borne viral infections: a systematic review. *J Trop Med Hyg* 90(5):882–891.

Figures

Figure 1: (A) Location of outbreak within Bangladesh. (B) Location of all households in the village. Orange triangles indicate at least one individual in the household with symptoms. (C) Epidemic curve separated by sex.

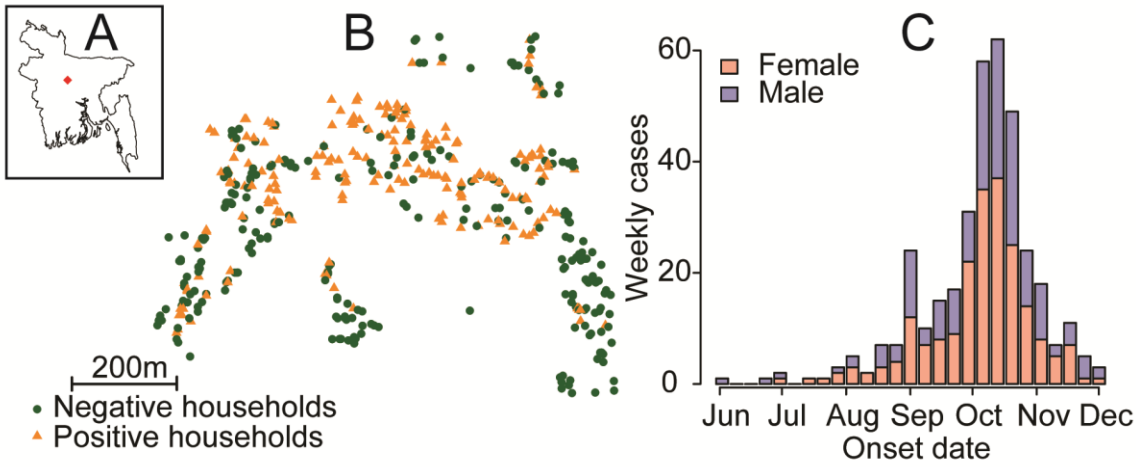

Figure 2: Parameter estimates from transmission model and results of human mobility study. (A) Probability of transmission for an individual living in the same household as a case. (B) Transmission kernel. (C) Relative susceptibility for children (those under 16 years) versus adults and females versus males. (D) Relative risk of being in or around the home (defined as within 50m of home location) from a nationwide mobility study. (E) Relative susceptibility for individuals from households that used anti-mosquito coils versus individuals from households that do not.

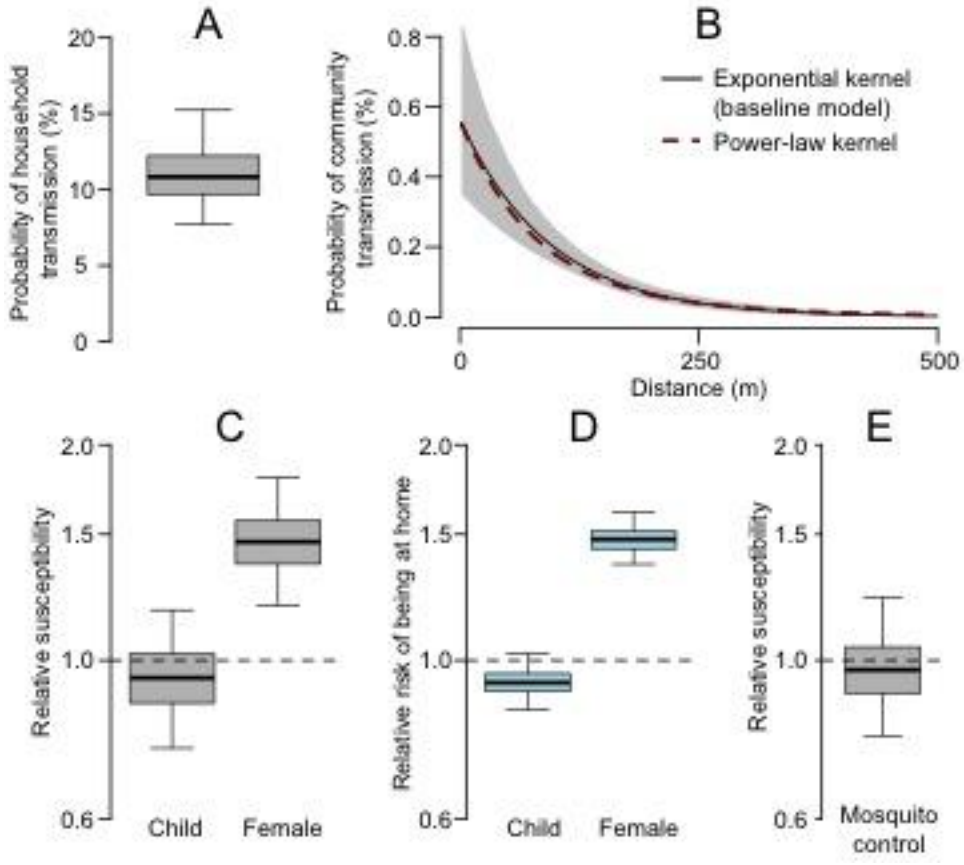

Figure 3. (A) Single model realization. (B) Proportion of transmission events within households and at different distance from a case household. (C) R(t) over time where each point represents the estimated reproductive number for the 30 day window centered at that time point. HH: Household.

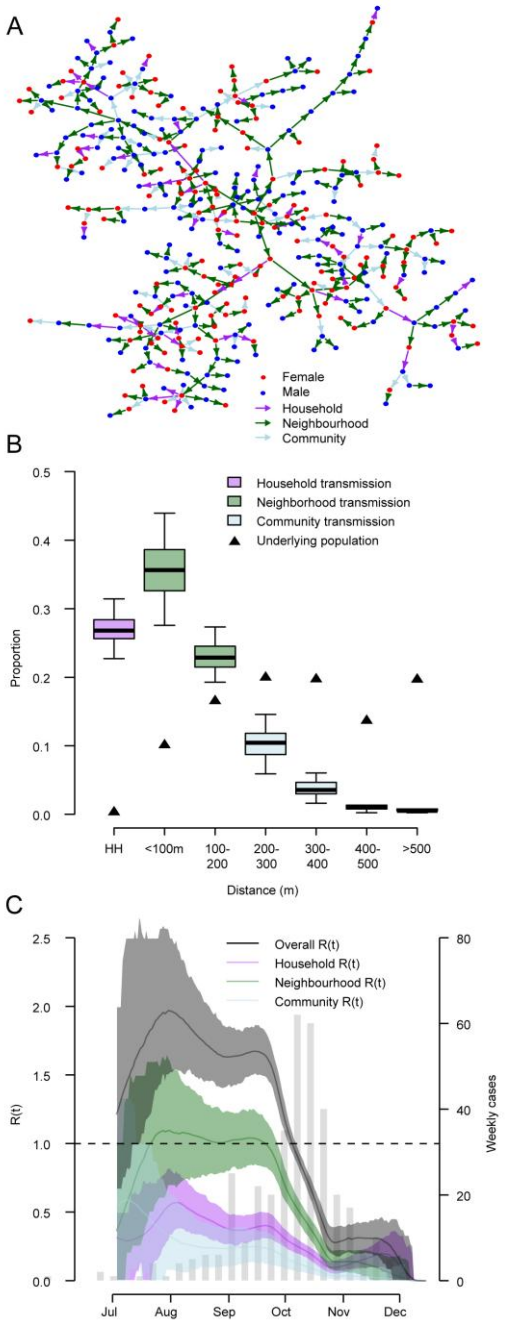

Figure 4. Model fit. To assess model fit we simulated epidemics starting from day 80 (8 infections had occurred by that time point). (A) Sets out the epidemic curves of the simulated epidemics as compared to the observed epidemic. (B) Sets out the spatial distribution of cases from the outbreak where a point is orange if at least 50% of model simulations have at least one case in that household. This figure should be compared to Figure 1B. (C) We divided the outbreak area into 50m by 50m grid cells and compared the mean proportion of individuals that lived in that cell that were positive across the simulations with the observed proportion that were positive.