

RESEARCH ARTICLE

Differential mobility and local variation in infection attack rate

David J. Haw¹, Derek A. T. Cummings^{2,3}, Justin Lessler⁴, Henrik Salje^{4,5,6,7}, Jonathan M. Read⁸, Steven Riley^{1*}

1 MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, United Kingdom, **2** Department of Biology, University of Florida, Gainesville, Florida, United States of America, **3** Emerging Pathogens Institute, University of Florida, Gainesville, Florida, United States of America, **4** Department of Epidemiology, Johns Hopkins University, Baltimore, Maryland, United States of America, **5** Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Paris, France, **6** CNRS, URA3012, Paris, France, **7** Center of Bioinformatics, Biostatistics and Integrative Biology, Institut Pasteur, Paris, France, **8** Centre for Health Informatics Computing and Statistics, Lancaster Medical School, Lancaster University, Lancaster, United Kingdom

* s.riley@imperial.ac.uk



OPEN ACCESS

Citation: Haw DJ, Cummings DAT, Lessler J, Salje H, Read JM, Riley S (2019) Differential mobility and local variation in infection attack rate. *PLoS Comput Biol* 15(1): e1006600. <https://doi.org/10.1371/journal.pcbi.1006600>

Editor: Ian Hall, UNITED KINGDOM

Received: April 25, 2018

Accepted: October 30, 2018

Published: January 22, 2019

Copyright: © 2019 Haw et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: For funding we acknowledge: Medical Research Council (UK, Project MR/J008761/1, Centre MR/R015600/1); Wellcome Trust (UK, Investigator Award 200861/Z/16/Z, Collaborator Award 200187/Z/15/Z); National Institute of Health (US, U01GM110721-01, U54GM088491, R01 AI102939, R01AI114703) National Institute for Health Research (UK, for Health Protection Research Unit funding). The funders had no role in

Abstract

Infectious disease transmission is an inherently spatial process in which a host's home location and their social mixing patterns are important, with the mixing of infectious individuals often different to that of susceptible individuals. Although incidence data for humans have traditionally been aggregated into low-resolution data sets, modern representative surveillance systems such as electronic hospital records generate high volume case data with precise home locations. Here, we use a gridded spatial transmission model of arbitrary resolution to investigate the theoretical relationship between population density, differential population movement and local variability in incidence. We show analytically that a uniform local attack rate is typically only possible for individual pixels in the grid if susceptible and infectious individuals move in the same way. Using a population in Guangdong, China, for which a robust quantitative description of movement is available (a travel kernel), and a natural history consistent with pandemic influenza; we show that local cumulative incidence is positively correlated with population density when susceptible individuals are more connected in space than infectious individuals. Conversely, under the less intuitively likely scenario, when infectious individuals are more connected, local cumulative incidence is negatively correlated with population density. The strength and direction of correlation changes sign for other kernel parameter values. We show that simulation models in which it is assumed implicitly that only infectious individuals move are assuming a slightly unusual specific correlation between population density and attack rate. However, we also show that this potential structural bias can be corrected by using the appropriate non-isotropic kernel that maps infectious-only code onto the isotropic dual-mobility kernel. These results describe a precise relationship between the spatio-social mixing of infectious and susceptible individuals and local variability in attack rates. More generally, these results suggest a genuine risk that mechanistic models of high-resolution attack rate data may reach spurious conclusions if the precise implications of

study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

spatial force-of-infection assumptions are not first fully characterized, prior to models being fit to data.

Author summary

We know that some places have higher rates of infectious disease than others. At the moment, we usually only measure these differences for large towns and cities, though modern data allows us to track movement at much higher resolution. In this paper, we used a computer simulation of an epidemic to propose ways that rates of incidence in small local areas might be related to population density. We found that if infectious people are better connected than non-infectious people, perhaps because they receive visitors, then, on average, higher density areas would have lower rates of infection. If infectious people were less connected than non-infectious people then higher density areas would have higher rates of infection. As data get more accurate, this type of analysis will allow us to propose and test ways to optimize interventions such as the delivery of vaccines and antivirals during a pandemic.

Introduction

The spatial heterogeneity of infectious disease incidence at large scales presents numerous intervention opportunities and challenges. Maps of malaria prevalence [1] have been used to target additional surveillance and to prioritize countries and geographical regions for additional intervention investment, resulting in substantial decreases in numbers of infections [2]. Over shorter timescales, spatial asynchrony in the northern hemisphere during the 2009 influenza pandemic likely led to variable effectiveness of vaccination when eventually deployed because of prior infections [3]. The epidemiological implications of substantial spatial heterogeneity in both incidence and transmission are topics of active research for most human pathogens [4].

These spatial heterogeneities must be influenced by two key human behaviours: where people choose to live and how they move. Because the home location of an individual is primarily used as the geographic location when cases are recorded, absolute spatial incidence is driven by population density: where more people live in a given unit area, there is greater potential for cases. Accurate high resolution estimates of population density [5, 6] and travel [7] have helped refine global absolute estimates of disease incidence and prevalence [8–11]. In order for a directly transmitted human pathogen to move through space, at least one person must travel away from home and meet another person. Even for vector borne pathogens such as malaria and Zika virus, typical distances traveled by the vector are much shorter than those traveled by human hosts. Human movement is captured by survey data on journeys to work [12], questionnaire-based surveys [13] and location logging of mobile devices [14–16].

Although spatial heterogeneity has been measured at larger scales (e.g. serological attack rates for influenza [17]), modern pathogen surveillance enables more finely resolved incidence data sets, with details such as precise geographical location captured with increasing frequency by modern digital and biological technology. For example, the full genome of a pathogen can be made available in almost real time directly from clinical samples taken in the community [18], and the home location of everyone attending a health care facility can be extracted from clinical episode data [19]. Because this level of geographical precision for high quality

incidence data has not previously been available, both epidemiological and disease-dynamic studies of infectious disease have focused on predicting and explaining incidence patterns measured at larger spatial scales, often with all cases within an administrative unit reported together. Additional insights are likely being lost during this aggregation process.

Available evidence and intuition suggests that infectious and non-infectious individuals have different social interactions during an outbreak [20], with plausible scenarios in which either one or the other may be more connected in space. For example, susceptible individuals are more likely to travel more than are infectious individuals with mild symptoms [21]. However, family members and friends providing care for infectious individuals may often not behave in the same way as an average susceptible individual. Also, infectious individuals themselves may travel long distances away from transmission hotspots to seek medical care during outbreaks of highly pathogenic infections [22].

Disease dynamic models are often used to study infection incidence and are defined primarily by their force-of-infection (FOI) term: a precise mathematical specification of how the risk of infection experienced by a susceptible individual is driven by the number of currently infectious individuals and by their characteristics. For example, the ages of infectious and susceptible individuals must sometimes affect the risk of infection, as must the distance between their home addresses. Disease dynamic models that represent space [23] are now used routinely to understand large-scale spatial heterogeneity in incidence: to estimate the relative effectiveness of spatially heterogeneous interventions (given the observed incidence); to reveal underlying social mechanisms of transmissions; and, with increasing frequency, to forecast future spatial incidence patterns [24]. All transmission models that represent space include some kind of spatial kernel—a formal definition of the way in which individuals from different locations distribute their influence over the whole of geographical space.

However, there is substantial variability in the underlying FOI assumptions made in these models, which are often not discussed explicitly and have likely only rarely made material differences to model-based results aggregated at larger spatial scales. Nonetheless, we hypothesise that these different FOI assumptions represent important alternate hypotheses for the mechanisms of transmission and may lead to substantial structural biases in the predictions of attack rates at smaller spatial scales. Here, we propose a general theoretical framework for the study of infectious disease incidence at arbitrarily small spatial scales and, in particular, we look at the relative mobility of infectious individuals relative to susceptible individuals as a potential driver of heterogeneity in incidence.

Results

Algebraic analyses show that differential spatial connectivity of susceptible and infectious individuals can lead to variability in local attack rates (S1 Protocol). Firstly, we showed that if susceptible and infectious individuals are assumed to be connected in the same way across all points in space, then local attack rates are uniform for any population density distribution or grid resolution. For lower resolution grids with large individual spatial elements, where the amplitude of connectivity of individuals outside their home pixel is small, the impact of differential connectivity between susceptible and infectious individuals is still negligible, even to the point that it is reasonable to assume that infectious individuals have no connectivity at all outside their home location. However, as the resolution of the grid increases and pixels become smaller, individuals have a substantial number of connections outside their home pixel. Under this scenario, it was no longer possible to prove analytically that differences in the connectedness of susceptible and infectious individuals would not lead to local variation in attack rates. These analytical results were not affected by the presence of age stratification in the

transmission process, so long as the behavior and distribution of age groups was assumed to be uniform across space.

We established a baseline numerical scenario consistent with a 1918-like influenza pandemic by implementing the underlying transmission model (see [Methods](#)) as ordinary differential equations (ODEs). Using: a 1km by 1km gridded population density (55km by 33km to the east and north of Guangzhou, China); a spatial contact kernel estimated in the same population [25]; a basic reproductive number $R_0 = 1.8$ [26] and recovery rate $1/2.6 \text{ days}^{-1}$ [27]; we recovered a global uniform attack rate of $z = 0.73$, consistent with the homogeneous mixing model SIR model [28]. We also introduced age-stratified populations and transmission using parameters estimated in this population [13]. For this population, accurate high-resolution data on local age distributions were not available, therefore, we assumed that all pixels had populations with the same age distribution, even though the total number of individuals in a single pixel varied substantially. This addition of age effects in the transmission process did not introduce spatial variation but did reduce the uniform global attack rate to $z = 0.43$, consistent with analysis of the 2009 influenza pandemic [29]. We validated the precision of attack rates obtained from the ODEs using age- and space-stratified refinements [23] of the standard implicit equation relating attack rate (final size) z to R_0 : $z = 1 - e^{-R_0 z}$ [28].

We hypothesized that both population density and the gradient of population density may influence small-scale attack rates in these models. [Fig 1A and 1B](#) show the uniform attack rate when mobility is independent of infection status (henceforth referred to as “dual mobility”) with four age classes, plotted against log of population density and gradient of log population density respectively (with log gradient defined as the average difference between the log of a location’s resident population and that of its 8 immediate neighbors).

When only non-infectious individuals were assumed to be mobile (S-mobility), location-specific attack rates were positively correlated with log population density, correlation coefficient $c = 0.75$ ([Fig 1C](#)). Attack rates varied between a minimum of 33.72% to a maximum of 45.76%, an absolute range of 12.04%. Location-specific attack rates were slightly less correlated with the log gradient of population density (correlation coefficient $c = 0.73$, [Fig 1D](#)). Locations with higher attack rates tended to be densely populated relative to neighboring locations ([Fig 2A and 2B](#)). Note that the term “S-mobility” includes mobility in the recovered population.

Conversely, when only infectious individuals were assumed to be mobile (I-mobility), pixel attack rates were negatively correlated with log population density ($c = -0.7707$, [Fig 1E](#)) and even more strongly negatively correlated with log density gradient ($c = -0.8816$, [Fig 1F](#)). Attack rates varied over a greater range than for susceptible-only mobility: from a minimum of 32.61% to a maximum of 90.73%, with an absolute range of 58.12%. High attack rate pixels tended to be sparsely populated relative to neighboring locations ([Fig 2A and 2C](#)). The reader is referred to the discussion for an evaluation of the applicability of this assumption to epidemic models.

Measures of spatial variation are inherently dependent on the resolution of the model grid and even the strong variability outlined above would be missed by most surveillance systems. The absolute range of attack rates for the susceptible-only movement was reduced to 1.67% when aggregated to 8km by 8km pixels. Even though the effect of infectious-only movement was stronger than for susceptible-only mobility, it was rapidly hidden by the aggregation of pixels, with the absolute range dropping to 3.78% when aggregated to 8km by 8km pixels. Results of aggregation using S-mobility is shown in [Fig 3](#), and the corresponding result using I-mobility is shown in [S1 Fig](#).

The direction of association between FOI assumptions and local attack rate was preserved and the amplitude remained substantial for intermediate scenarios in which both susceptible and infectious individuals were mobile but to differing degrees. If infectious individuals had

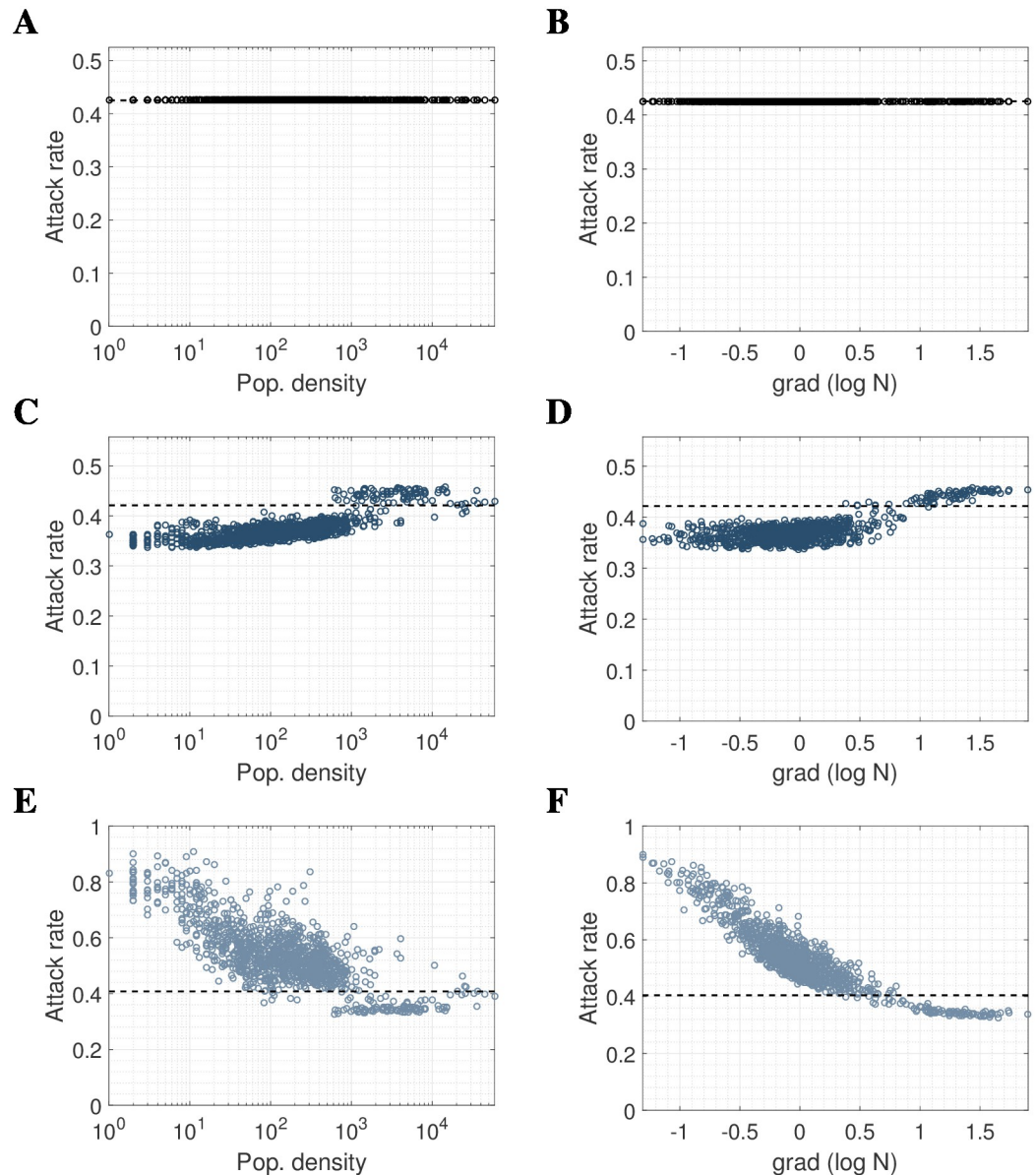


Fig 1. The relationship between force-of-infection (FOI) assumptions, local attack rates, population density and population density gradient, for a pandemic-influenza-like epidemic. The LHS shows the relationship between population density N (people/km²) and attack rate for (A) mobility independent of infection status (dual mobility), (C) mobility in non-infectious population only (S-mobility) and (E) mobility in infectious population only (I-mobility). The RHS shows the relationship between the gradient of $\log_{10}N$ and attack rate for (B) dual mobility, (D) S-mobility and (F) I-mobility. We used a 33km by 55km grid of 1km by 1km pixels to the North-East of Guangzhou, with kernel parameters $\alpha = 0.52$, $a = 0.58$, $p = 2.72$ and influenza natural history parameters $R_0 = 1.8$, $\gamma = 1/2.6$. Population gradient was defined as the difference between the log population density of a pixel and the average log population density of the 8 surrounding pixels.

<https://doi.org/10.1371/journal.pcbi.1006600.g001>

any more contacts than susceptible individuals then attack rates were negatively correlated with population density, and vice versa (Fig 4). When infectious individuals reduced their travel by a factor of 0.5, the absolute range of attack rates was 5.38% and when susceptible individuals reduced their mixing by the same degree (with infectious agents fully mobile), the absolute range was 12.89%.

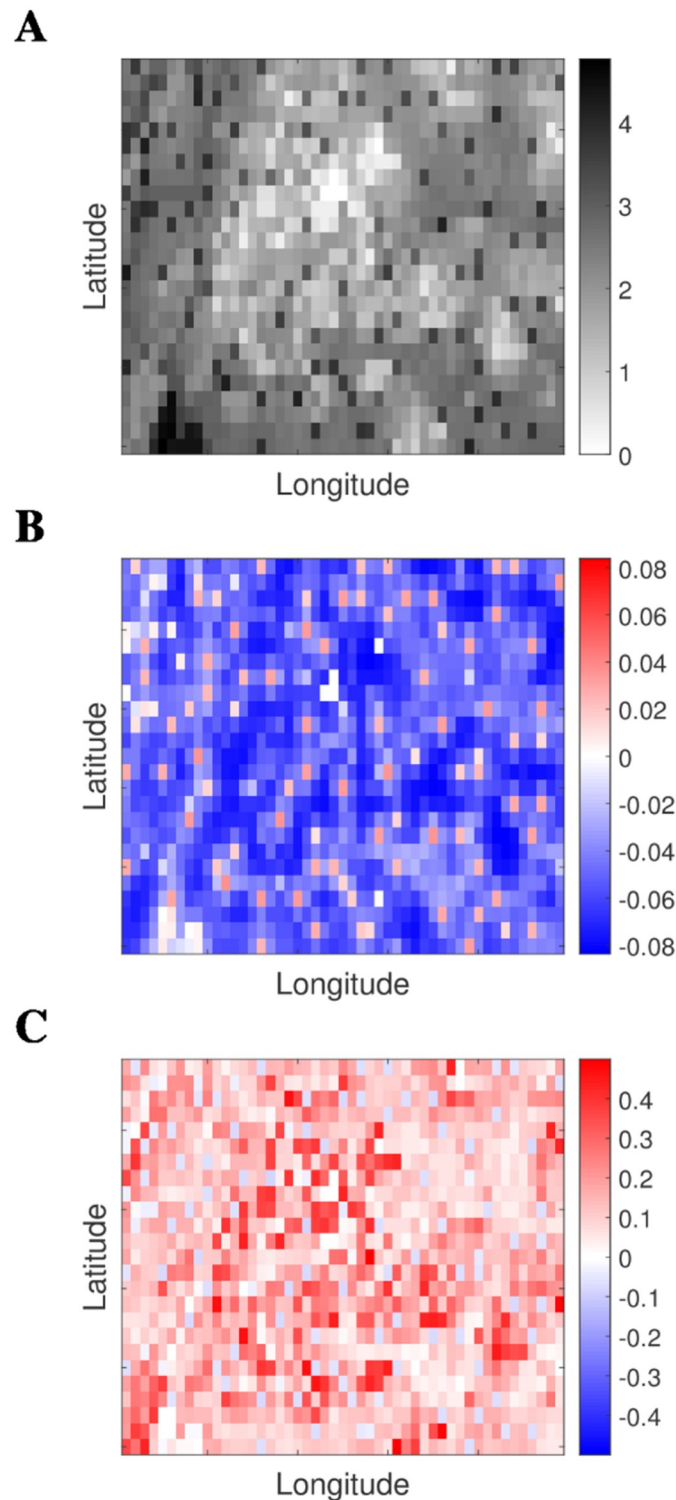


Fig 2. Spatial illustration of population density and non-uniform attack rates generated using different mobility assumptions. (A) Log_{10} population density (people/ km^2). **(B)** Difference between location-specific attack rates and global attack rate for S-mobility and **(C)** difference between location-specific attack rates and global attack rate for I-mobility. We change color scale between plots to better illustrate the emergent patterns. A total of 4 pixels are unpopulated and so attack rates are necessarily always zero in these locations.

<https://doi.org/10.1371/journal.pcbi.1006600.g002>

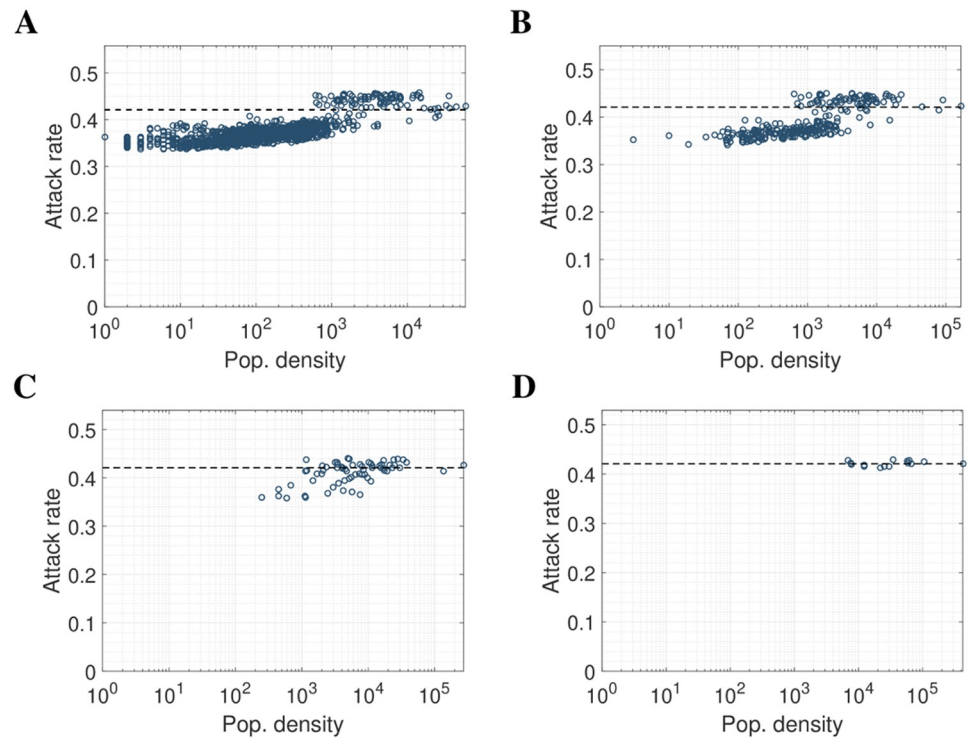


Fig 3. Aggregation of result using S-mobility. Plots show (A) initial result, aggregated into (B) 2km by 2km, (C) 4km by 4km, and (D) 8km by 8km pixels.

<https://doi.org/10.1371/journal.pcbi.1006600.g003>

The underlying mobility choice kernel K was defined by the relative probability of making a contact in a population at a distance r and of population size N . It was parameterized by an offset distance a , a distance power p and destination population power α ; $K = N^\alpha(1 + r/a)^{-p}$, with values obtained by fitting to data from this population [25]. Qualitatively, our conclusions about the impact of differential contact rates by susceptible individuals were not sensitive to values for the offset distance a nor the distance power p (Fig 5A–5D). However, they were sensitive to values of the destination power α for which we have used the best fit value of 0.53 (for results up to this point) (Fig 5E&5F). Intriguingly, with the often-assumed default value $\alpha = 1$, the correlation between local attack rates and population density or gradient have the opposite sign (S2 and S3 Figs). Moreover, $\alpha = 1$ induces weaker correlations with local population gradient. It is therefore essential to provide an accurate estimate for α , which does not require infection-related data, before attempting to infer infection-dependent mobility.

Stochastic solutions to the meta-population models suggest that attack rate variation driven by asymmetric mobility would not be dominated by demographic stochasticity (Fig 6). Variation in attack rate for the extreme cases of S- and I-mobility was dominated by stochastic effects only in sparsely populated areas. For pixels with the smallest population, the amplitude of variation expected to arise from asymmetric mobility is similar to that which may arise by chance due to stochastic effects. However, the expected amplitude of stochastic variation diminishes as population density increases, and variation in attack rate due to mobility assumption becomes apparent (S4 Fig). For example, using susceptible-only mobility for 1km by 1km pixels with populations between 1 and 85,163, the standard deviation in attack rate due to stochasticity is 9.45% while the standard deviation of expected attack rates due to asymmetric mobility is 2.61%.

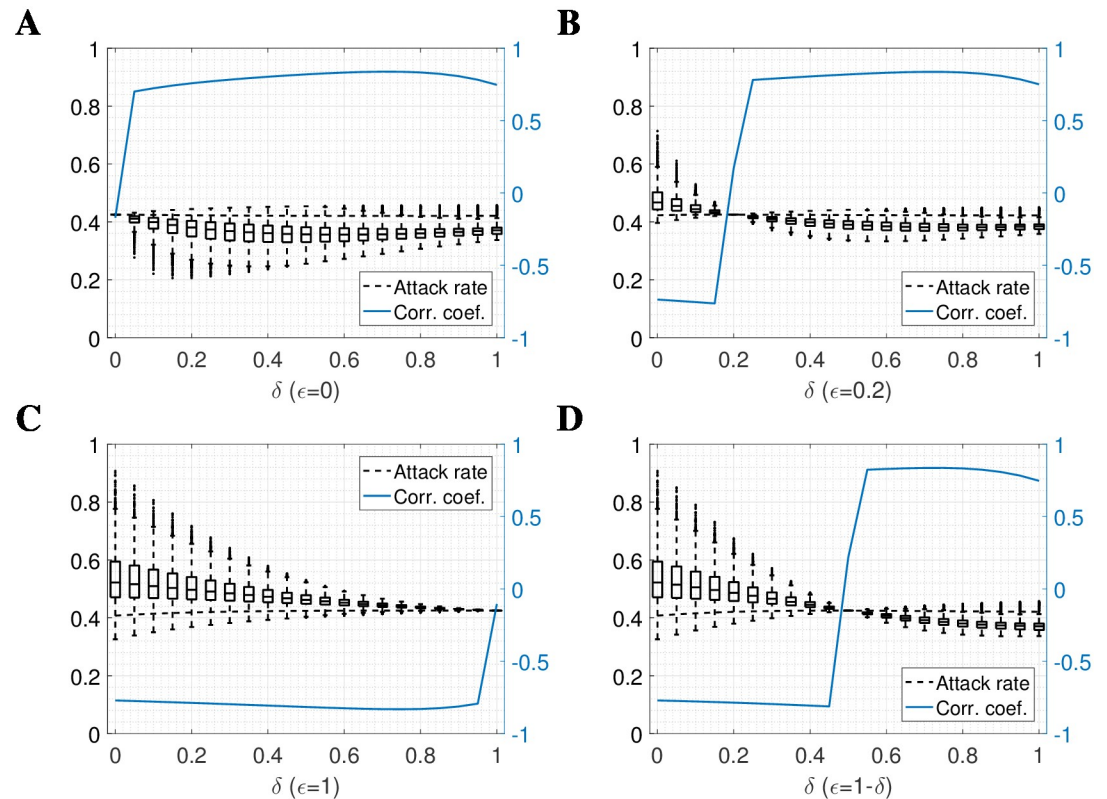


Fig 4. Limiting mobility of susceptible/recovered and immune agents according to parameters δ and ϵ . Mobility of the non-infective population is described by δ such that $\delta = 0$ yields no mobility, $\delta = 1$ yields mobility described by the kernel K , and transformation between these 2 extremes in linear. Similarly, ϵ describes the mobility of the infective population. Any values of $\delta = \epsilon$ thus yield (reduced) dual mobility, and so attack rates are uniform in space. Plots show (A) infectious population immobile, non-infectious mobility ranging from $\delta = 0$ to $\delta = 1$, moving from dual mobility to S-mobility, (B) constant reduced mobility in the infectious population ($\epsilon = 0.2$), possibly accounting for mobility in asymptomatic cases only, (C) full mobility in the infectious population, moving from I-mobility to dual mobility, and (D) $\epsilon = 1 - \delta$, illustrating the transition from I-mobility to S-mobility. Dashed lines show the global attack rate, and solid blue lines show correlation coefficient with log population density.

<https://doi.org/10.1371/journal.pcbi.1006600.g004>

These results are robust to our choice of illustrative population density and to alternate natural history parameters. The same effects are observed when using population density of Puerto Rico with influenza natural history parameters (S5 Fig) and with parameters that approximate vector-borne transmission, such as those of Zika or Chikungunya (S6 Fig). Summary statistics for these and all other deterministic model variants we have presented in this study are shown in S1 Table.

Discussion

We have shown that, under the assumption that an individual's total contact is independent of home location and where they travel, substantial heterogeneity in local attack rates could arise if mobility is dependent on infection status. Moreover, the direction of the relationship between attack rate and population density is dependent on the contribution of population density to the relative attractiveness of a location. For the estimate of that scaling for our sample population ($\alpha = 0.52$), and when susceptible individuals are more mobile than infectious individuals, attack rates are positively correlated with population density. Conversely, when

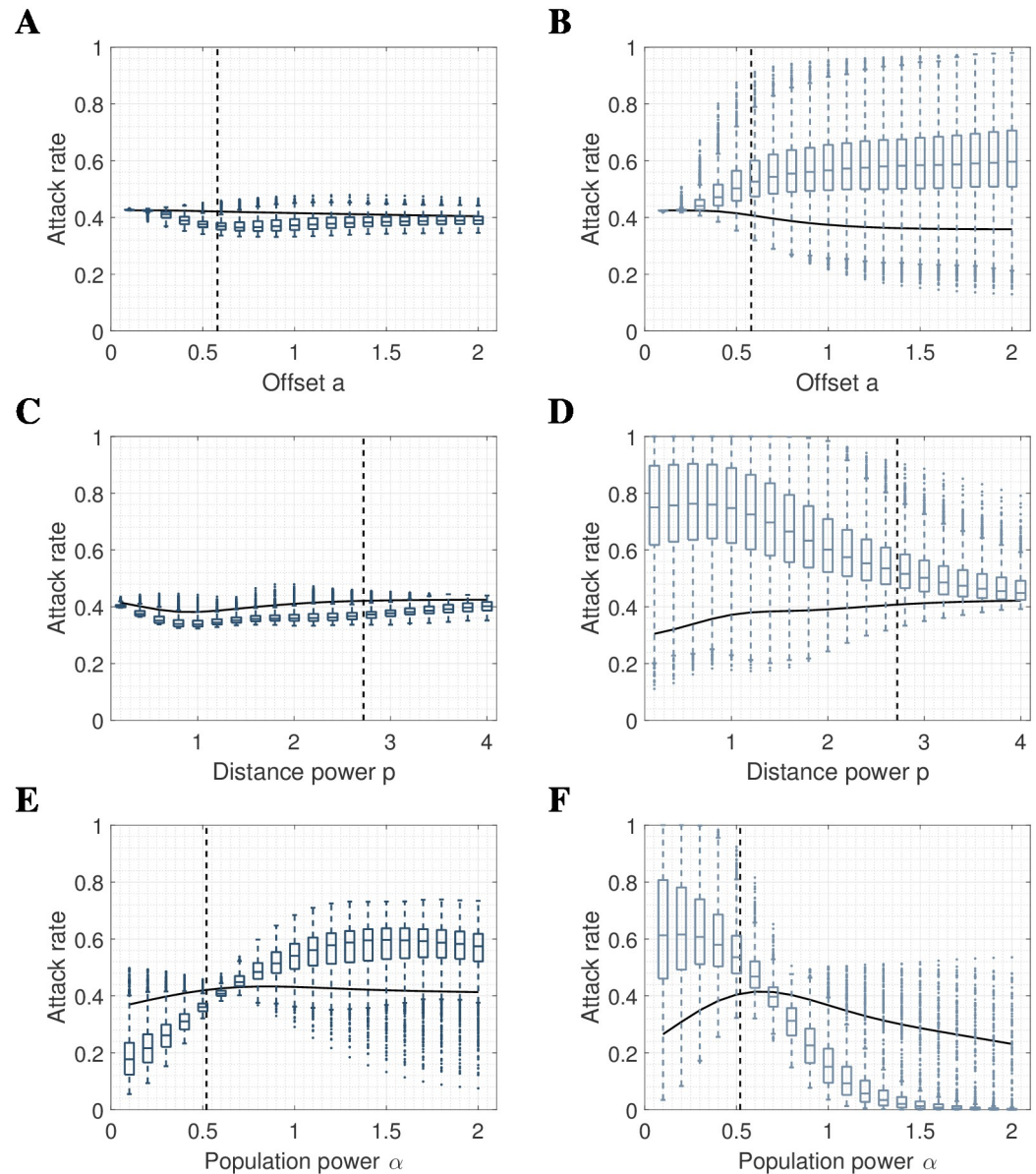


Fig 5. Sensitivity analysis. Distribution of local attack rates with respect to (A) offset a using S-mobility, (B) offset a using I-mobility, (C) distance power p using S-mobility, (D) distance power p using I-mobility, (E) population power α using S-mobility, and (F) population power α using I-mobility. Box plots show standard percentiles and outliers, solid lines show global attack rate, and dashed lines show parameter values used in the main result. When fixed, all parameters are as in main result, i.e. $a = 0.58$, $p = 2.72$, $\alpha = 0.52$. Dual mobility are omitted as they are flat with variance $\sigma^2 = 0$. Empty pixels yield attack rate zero and are omitted from calculations.

<https://doi.org/10.1371/journal.pcbi.1006600.g005>

using the often implicit assumption that the kernel is directly proportional to population density ($\alpha = 1$), this correlation is negative.

Though increased mobility in infectious agents may seem less likely than reduced mobility, there do exist potential scenarios where this may be the case in both human and animal systems. For example, humans may travel to access health care in the case of severe symptom onset as has been the case anecdotally during the 2003/4 SARS outbreak and the 2013/14 Ebola outbreak. Also infectious opiate users in the USA may be more mobile than less infectious

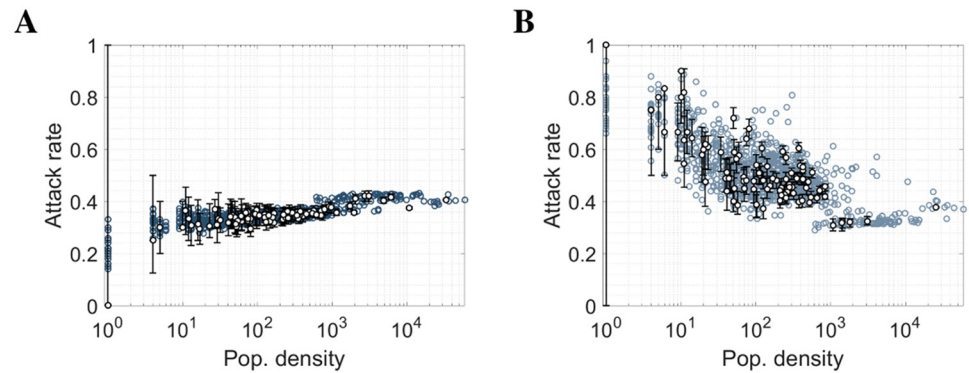


Fig 6. Mean attack rate over 100 iterations of stochastic equivalent of main result. We use (A) S-mobility and (B) I-mobility. 25-, 50- and 75-percentiles are shown for a sample of 100 locations.

<https://doi.org/10.1371/journal.pcbi.1006600.g006>

opiate users [30]. I-mobility may in fact be more relevant in the epidemiology of non-human infections, for example increased mobility in rabid dogs [31] and Gypsy moth caterpillars infected with baculovirus *forfeit* [32].

Our study has a number of limitations. We have not considered spatial variation in the age distribution of people, because these data were not available for our study population. Variability in local attack rates will very likely also be driven in non-trivial ways by spatial correlation in the proportion of the population in different age classes. This may be of particular significance in larger Chinese cities such as Guangzhou, in which urban areas are home to relatively few children and many rural locations have few working-age adults. There is also scope for the inclusion of an urban/rural distinction in the parametrization of the travel kernel [25], and the simulation of multiple years of transmission, which would extend the applicability of our results beyond pandemic scenarios for influenza and other emergent pathogens. The refinement of this framework to include the above phenomena is a priority for future work and we would expect differential movement patterns with age and population to impact our findings.

Though this study was limited to a standard SIR model, we would not expect the inclusion of a latent period, waning, or natural births and deaths to show make substantial differences to these findings. The primary results can be obtained using renewal equations which are only dependent on the probability of one individual escaping infection.

Our sensitivity analysis with respect to kernel population power α provides some insight into the underlying mechanisms that give rise to the observed correlations between attack rate and population density under different mobility assumptions. For example, consider the special case where only infectious people are mobile and α tends to large values, making mobility dependent only on population density of location, and not on geographical distance. Under this scenario, high density pixels will draw in more and more infectious people and therefore generate higher attack rates. Conversely, if $\alpha = 0$, then mobility is dependent only on distance. Under this scenario, we can think of the infectious populations spilling out of their home locations into neighboring ones. Thus, any sparsely populated location that is adjacent to a densely populated location will see an influx of infectious individuals resulting in a greater *proportion* infectious in that location, and therefore a stronger FOI and subsequent attack rate. A schematic for the latter case is given in [S7 Fig](#).

These results illustrate the potential knock-on effects of little or no dependence between transmissibility and population density: that infectious people from more densely populated areas go to nearby sparsely populated areas and in some sense “seek out” people in those areas to infect so they can reach their quota (I-mobility). Within the realm of parameters that are

supported by studies of human movement and infectious processes, the behaviors implied by the models we presented here seem valid.

Individual-based models have a number of advantages over other approaches. They can be coded in a generic way and adapted rapidly to different pathogen systems and specific scientific or policy questions. Even though they are often more substantially computationally burdensome than comparable meta-population approaches, they will likely be used with increasing frequency to address questions related to local attack rates. We have shown that mobility assumptions have implications for the interpretation of attack-rates derived from individual-based models, some of which assume implicitly that the spread of infection is driven by the movement of individuals. We have shown that, whichever mobility assumption is made in a given model, it is possible to modify this assumption by replacing isotropic K by a convoluted kernel L that accounts for the change in mobility assumption (and so L may not be a stochastic matrix and hence functions as a non-isotropic kernel). In particular, the low-prevalence assumption makes this transformation achievable with minimal modification to existing computer programs. Therefore, developers of individual-based models may wish to consider alternate connectivity matrices for their simulations so as to explicitly reflect different spatial assumptions about the force of infection.

We have also shown that the implications of typical assumptions that are made in spatially explicit FOI terms, including approximations to this crucial normalization, are non-trivial at small spatial scales. Such assumptions are, however, often not addressed explicitly and so may contribute unknowingly to results. We hope to offer clarity in the interpretation of FOI in spatial models, and to have provided a comprehensive framework from which we can gain a deeper understanding of the role of spatial mobility in disease transmission dynamics as infectious disease incidence data become available at higher and higher spatial resolution.

Methods

Spatial kernels

Data taken from populations we study here show that total contacts made per day, and contact durations, do correlate with population density ($p < 0.001$, [13]), but that the strength of the relationship is weak. This is in part due to working-age adults dominating the population of urban areas, but also to the phenomenon of urban isolation [33]. When investigating only the effect of mobility assumption in force of infection, our main results made the baseline assumption that total contact and duration of contact is independent of home location.

The way in which these contacts are distributed in space does, however, depend on distance and population density, and is described via a spatial kernel K . In matrix notation, K_{ij} is defined as the proportion of time spent by an agent from location i in location j . The assumption of uniformity of total contact therefore means that the rows of K sum to unity. Our model employs the offset gravity kernel, defined as follows:

$$K_{ij} \propto \frac{N_i N_j^\alpha}{1 + (r_{ij}/a)^p} \quad (1)$$

with baseline parameters of $a = 0.58$, $p = 2.72$, $\alpha = 0.52$, where r_{ij} denotes the geodesic distance between the center-points of pixels i and j . Of the kernel structures studied in [25], offset gravity is shown to best represent contact data. Imposing the constraint that K is stochastic renders redundant the factor N_i in the numerator (owing to row-normalization).

Population density map

We used rectangular excerpts from the Landscan dataset [34] with the lower left corner of the rectangle located on the center of the city of Guangzhou, China. The rectangle is 55km from east to west and 33km from north to south, and a 4km boundary area was excluded after simulation.

The boundary area was chosen according to the following rationale: when population density data for large suburban areas is truncated for the purpose of simulation, it is equivalent to imposing empty space outside of the boundary, and this modification may effect the attack rates calculated in pixels close to that boundary. We ran simulations on a large area of 1km by 1km pixels, and on smaller areas contained within this larger area. We found that attack rates agree on all pixels on the interior of the smaller area once a 4km perimeter is removed.

Force-of-infection

Let A denote the S-mobility kernel and B the I-mobility kernel. Then the age-independent generalized FOI equation is given by:

$$\lambda_i = \beta \sum_j A_{ij} \frac{\sum_k B_{jk}^T I_k}{\sum_l [A_{jl}^T (N_l - I_l) + B_{jl}^T I_l]} \tag{2}$$

For reduced mobility, movement of the non-infectious population is governed by a parameter δ such that $A = (1 - \delta)E + \delta K$, where E is an identity matrix representing absence of spatial mobility. Similarly, we describe mobility of infective individuals by ϵ such that $B = (1 - \epsilon)E + \epsilon K$. S-mobility thus corresponds to $\delta = 1, \epsilon = 0$ and I-mobility to $\delta = 0, \epsilon = 1$.

If K is the $n \times n$ spatial kernel, indexed by i, j, k, l , and C the 4×4 age-mixing matrix, indexed by a, b, c, d , then the age-explicit dual-mobility equation is given by:

$$\lambda_{(a,i)}^D = \beta \sum_{b,j} K_{ij} \delta_{ab} \frac{\sum_{c,k} K_{jk}^T C_{bc} I_{(c,k)}}{\sum_{d,l} K_{jl}^T N_{(d,l)}} \tag{3}$$

This can be combined with Eq (2) to give the age-dependent system with reduced mobility.

In all simulations presented in this study, we use the pointwise product of the matrices defining number of contacts and duration of contact between age groups 0–4, 5–19, 20–64 and 65+ derived in [13]. These age-mixing matrices were constructed from contact surveys conducted in the region of Guangzhou used in our results.

Model solutions

We define the gridded transmission model as ordinary differential equations. However, we also implement a stochastic compartmental version of the model and we calculate attack rates using recursive equations.

We used a standard SIR model with $\dot{S}_i = -S_i \lambda_i, \dot{I}_i = S_i \lambda_i - \gamma I_i, \dot{R}_i = \gamma I_i$. ODE models were seeded proportional to population density ($\sigma = 10^{-8} \times N / \sum_i N_i$), and agreed with final size calculations (which assume infinitesimal seeding). Integration of ODEs with full FOI in the S- and I-mobility case, i.e. with $I_i(t)$ in denominators, showed low-prevalence approximations to be good. For example, in the main S-mobility result, the mean difference in pixel attack rates between the full FOI and low prevalence approximation was 6.22×10^{-4} with maximum difference 3.3×10^{-3} occurring in a pixel with population 726. Therefore, numerical solutions for all figures were obtained using the low prevalence approximation (c.f. S1 Protocol). A selection of smaller examples agreed when checked using the full FOI.

The stochastic compartmental variant of our model selected the number of agents to infect from binomial distribution with parameters $S_{(a,i)}$ and $1 - \exp(-\lambda_{(a,i)})$. This method requires specification of a time-step, and we found $\Delta t = 1/6$ days to be sufficiently small (results did not change when Δt was doubled, and results were consistent with the corresponding deterministic model).

Supporting information

S1 Table. Summary statistics for different model parameters, populations and mobility assumptions. Results for different grid sizes involve aggregation of result obtained at 1km by 1km resolution. In all cases, empty pixels are omitted from calculations. It is therefore possible to obtain a smaller minimum value of attack rate after aggregation.

(PDF)

S1 Fig. Aggregation of result using I-mobility. Plots show (A) initial result, aggregated into (B) 2km by 2km, (C) 4km by 4km, and (D) 8km by 8km pixels.

(TIFF)

S2 Fig. Sensitivity analysis: Correlation coefficient of attack rate with population density for different values of kernel parameters. We vary (A) α with $a = 0.58$ and $p = 2.72$ fixed, comparing S-mobility with I-mobility (B) a and α , using S-mobility with $p = 2.72$ fixed, (C) a and α , using I-mobility with $p = 2.72$ fixed, (D) p and α , using S-mobility with $a = 0.58$ fixed, and (E) p and α , using I-mobility with $a = 0.58$ fixed. All fixed parameter values are those used in main result.

(TIF)

S3 Fig. Repeating our main result with $\alpha = 1$. We use (A) S-mobility, with attack rates plotted against population density, (B) S-mobility/gradient, (C) I-mobility/density, and (D) I-mobility/gradient. Other parameters remain as in main result, i.e. $a = 0.58$, $p = 2.72$.

(TIF)

S4 Fig. Ratio R of location-specific standard deviation over 100 iterations of stochastic model to standard deviation of corresponding deterministic model result over all pixels.

We use (A) S-mobility and (B) I-mobility. All parameters as in main result, i.e. $a = 0.58$, $p = 2.72$, $\alpha = 0.52$.

(TIF)

S5 Fig. Simulated attack rates using population density of North-East Puerto-Rico: Influenza. We use a 60km by 60km grid of 1km by 1km pixels, and influenza-like natural history parameters $R_0 = 1.8$, $\gamma = 1/2.6$, with (A) S-mobility plotted against population density, (B) S-mobility plotted against log population gradient, (C) I-mobility/density, and (D) I-mobility/gradient. Kernel parameters as in main result, i.e. $a = 0.58$, $p = 2.72$, $\alpha = 0.52$.

(TIF)

S6 Fig. Simulated attack rates using population density of North-East Puerto-Rico: Zika.

We use a 60km by 60km grid of 1km by 1km pixels, and natural history parameters $R_0 = 4$, $\gamma = 1/10$ approximating vector-borne transmission (e.g. Zika, Chikungunya), with (A) S-mobility plotted against population density, (B) S-mobility plotted against log population gradient, (C) I-mobility/density, and (D) I-mobility/gradient. Kernel parameters as in main result, i.e.

$a = 0.58$, $p = 2.72$, $\alpha = 0.52$.

(TIF)

S7 Fig. Schematic illustration of the process by which the observed trends arise. As an example, assume infectious-only mobility and let location x be locally densely populated, with disease prevalence initially proportional to population density (initial infective populations are shown in light blue). If the travel kernel K is dominated by distance (α small, c.f. S3 Fig), then some of the infectious population in each pixel will relocate to neighboring pixels (white). The result is a higher prevalence in locally sparsely populated pixels. Moreover, a larger local population gradient will allow this phenomenon to persist. Moreover, infection status is recorded by home location, which, under the I-mobility assumption, is equivalent to location when susceptible/recovered. The result is a negative correlation between local population density and attack rate.

(TIF)

S1 Protocol. Additional algebraic analyses. Algebraic analyses of: uniform local attack rates for dual mobility assumptions; the relationship between our results and other approximations in the literature [35–37]; convoluted kernel formulations; and calculation of the global transmissibility coefficient.

(PDF)

Author Contributions

Conceptualization: David J. Haw, Steven Riley.

Formal analysis: David J. Haw.

Funding acquisition: Steven Riley.

Investigation: David J. Haw.

Methodology: David J. Haw.

Software: David J. Haw.

Supervision: Steven Riley.

Validation: David J. Haw.

Visualization: David J. Haw.

Writing – original draft: David J. Haw.

Writing – review & editing: David J. Haw, Derek A. T. Cummings, Justin Lessler, Henrik Salje, Jonathan M. Read, Steven Riley.

References

1. Hay SI, Snow RW. The malaria Atlas Project: developing global maps of malaria risk. *PLoS Med.* 2006; 3(12):e473. <https://doi.org/10.1371/journal.pmed.0030473> PMID: 17147467
2. Bhatt S, Weiss DJ, Cameron E, Bisanzio D, Mappin B, Dalrymple U, et al. The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature.* 2015; 526(7572):207–211. <https://doi.org/10.1038/nature15535> PMID: 26375008
3. Charu V, Zeger S, Gog J, Bjørnstad ON, Kissler S, Simonsen L, et al. Human mobility and the spatial transmission of influenza in the United States. *PLoS Comput Biol.* 2017; 13(2):e1005382. <https://doi.org/10.1371/journal.pcbi.1005382> PMID: 28187123
4. Lessler J, Azman AS, McKay HS, Moore SM. What is a Hotspot Anyway? *Am J Trop Med Hyg.* 2017; 96(6):1270–1273. <https://doi.org/10.4269/ajtmh.16-0427> PMID: 28719289
5. Sutton P, Elvidge C, Obremski T. Building and Evaluating Models to Estimate Ambient Population Density. *Photogrammetric Engineering and Remote Sensing.* 2003; 69(5):545–553. <https://doi.org/10.14358/PERS.69.5.545>

6. Deville P, Linard C, Martin S, Gilbert M, Stevens FR, Gaughan AE, et al. Dynamic population mapping using mobile phone data. *Proceedings of the National Academy of Sciences*. 2014; 111(45):15888–15893. <https://doi.org/10.1073/pnas.1408439111>
7. Wesolowski A, O'Meara WP, Eagle N, Tatem AJ, Buckee CO. Evaluating Spatial Interaction Models for Regional Mobility in Sub-Saharan Africa. *PLoS Computational Biology*. 2015; 11(7):1–16. <https://doi.org/10.1371/journal.pcbi.1004267>
8. Hay SI, Snow RW. The Malaria Atlas Project: Developing global maps of malaria risk. *PLoS Medicine*. 2006; 3(12):2204–2208. <https://doi.org/10.1371/journal.pmed.0030473>
9. Stanaway JD, Flaxman AD, Naghavi M, Fitzmaurice C, Vos T, Abubakar I, et al. The global burden of viral hepatitis from 1990 to 2013: findings from the Global Burden of Disease Study 2013. *The Lancet*. 2016; 388(10049):1081–1088. [https://doi.org/10.1016/S0140-6736\(16\)30579-7](https://doi.org/10.1016/S0140-6736(16)30579-7)
10. Rey J, Stanaway D, Shepard DS, Undurraga EA, Halasa YA, Coff LE, et al. The global burden of dengue: an analysis from the Global Burden of Disease Study 2013. *WwwTheLancetCom/Infection*. 2016; 16(6):712–723.
11. Wesolowski A, Zu Erbach-Schoenberg E, Tatem A, Lourenco C, Viboud C, Charu V, et al. Multinational patterns of seasonal asymmetry in human movement influence infectious disease dynamics. *Nature Communications*. 2017; 8(1). <https://doi.org/10.1038/s41467-017-02064-4> PMID: 29234011
12. Keeling MJ, Danon L, Ford AP, House T, Jewell CP, Roberts GO, et al. Networks and the epidemiology of infectious disease. *Interdisciplinary Perspectives on Infectious Diseases*. 2011; 2011. <https://doi.org/10.1155/2011/284909> PMID: 21437001
13. Read JM, Lessler J, Riley S, Wang S, Tan LJ, Kwok KO, et al. Social mixing patterns in rural and urban areas of southern China. *Proceedings of the Royal Society of London B: Biological Sciences*. 2014; 281(1785). <https://doi.org/10.1098/rspb.2014.0268>
14. Gonzalez MC, Hidalgo CA, Barabasi AL. Understanding individual human mobility patterns. 2008; 453 (June).
15. Vazquez-Prokopec GM, Bisanzio D, Stoddard ST, Paz-Soldan V, Morrison AC, Elder JP, et al. Using GPS Technology to Quantify Human Mobility, Dynamic Contacts and Infectious Disease Dynamics in a Resource-Poor Urban Environment. *PLoS ONE*. 2013; 8(4):1–10. <https://doi.org/10.1371/journal.pone.0058802>
16. Perkins TA, Garcia AJ, Paz-Soldan VA, Stoddard ST, Reiner RC, Vazquez-Prokopec G, et al. Theory and data for simulating fine-scale human movement in an urban environment. *Journal of The Royal Society Interface*. 2014; 11(99):20140642–20140642. <https://doi.org/10.1098/rsif.2014.0642>
17. Van KM, Hirve S, Koukounari A, Mounts A. Estimating age-specific cumulative incidence for the 2009 influenza pandemic: a meta-analysis of A(H1N1)pdm09 serological studies from 19 countries. *INFLUENZA AND OTHER RESPIRATORY VIRUSES*. 2013; 7:872–886. <https://doi.org/10.1111/irv.12074>
18. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat Protoc*. 2017; 12(6):1261–1276. <https://doi.org/10.1038/nprot.2017.066> PMID: 28538739
19. Paul P, Heng BH, Seow E, Molina J, Tay SY. Predictors of frequent attenders of emergency department at an acute general hospital in Singapore. *Emerg Med J*. 2010; 27(11):843–848. <https://doi.org/10.1136/emj.2009.079160> PMID: 20472703
20. Funk S, Bansal S, Bauch CT, Eames KTD, Edmunds WJ, Galvani AP, et al. Nine challenges in incorporating the dynamics of behaviour in infectious diseases models. *Epidemics*. 2015; 10:21–25. <https://doi.org/10.1016/j.epidem.2014.09.005> PMID: 25843377
21. Funk S, Salathé M, Jansen VAA. Modelling the influence of human behaviour on the spread of infectious diseases: a review. *J R Soc Interface*. 2010; 7(50):1247–1256. <https://doi.org/10.1098/rsif.2010.0142> PMID: 20504800
22. Campbell GL, Hughes JM. Plague in India: a new warning from an old nemesis. *Ann Intern Med*. 1995; 122(2):151–153. <https://doi.org/10.7326/0003-4819-122-2-199501150-00014> PMID: 7992992
23. Riley S. Models of Infectious Disease. *Science*. 2007; 316(5829):1298–1301. <https://doi.org/10.1126/science.1134695> PMID: 17540894
24. Riley S, Eames K, Isham V, Mollison D, Trapman P. Five challenges for spatial epidemic models. *Epidemics*. 2015; 10:68–71. <https://doi.org/10.1016/j.epidem.2014.07.001> PMID: 25843387
25. Read J, Mills H. Is this on bioRxiv yet?; 2017.
26. Biggerstaff M, Cauchemez S, Reed C, Gambhir M, Finelli L. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature. *BMC Infectious Diseases*. 2014; 14(1):480. <https://doi.org/10.1186/1471-2334-14-480> PMID: 25186370

27. Ferguson NM, Cummings DAT, Cauchemez S, Fraser C, Riley S, Meeyai A, et al. Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature*. 2005; 437(7056):209–214. <https://doi.org/10.1038/nature04017> PMID: 16079797
28. Diekmann O, Heesterbeek JAP. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*. 1st ed. Wiley; 2000.
29. Glatman-Freedman A, Portelli I, Jacobs SK, Mathew JI, Slutzman JE, Goldfrank LR, et al. Attack Rates Assessment of the 2009 Pandemic H1N1 Influenza A in Children and Their Contacts: A Systematic Review and Meta-Analysis. *PLOS ONE*. 2012; 7(11):1–17. <https://doi.org/10.1371/journal.pone.0050228>
30. Burke DS. Forecasting the opioid epidemic. *Science*. 2016; 354(6312):529–529. <https://doi.org/10.1126/science.aal2943> PMID: 27811241
31. Gompper ME. *Free-Ranging Dogs and Wildlife Conservation*. Oxford; 2015.
32. Hoover K, Grove M, Gardner M, Hughes DP, McNeil J, Slavicek J. A Gene for an Extended Phenotype. *Science*. 2011; 333(6048):1401–1401. <https://doi.org/10.1126/science.1209199> PMID: 21903803
33. Kinenberg E. *Dying Alone*; 2001.
34. Dobson JE, Bright EA, Coleman PR, Durfee RC, Worley BA. LandScan: A global population database for estimating populations at risk. *Photogrammetric Engineering and Remote Sensing*. 2000; 66(7):849–857.
35. Ma J, Earn DJD. Generality of the final size formula for an epidemic of a newly invading infectious disease. *Bulletin of Mathematical Biology*. 2006; 68(3):679–702. <https://doi.org/10.1007/s11538-005-9047-7> PMID: 16794950
36. Clancy D, Pearce CJ. The effect of population heterogeneities upon spread of infection. *Journal of Mathematical Biology*. 2013; 67(4):963–987. <https://doi.org/10.1007/s00285-012-0578-x> PMID: 22941453
37. Sattenspiel L, Dietz K, Sattenspiel L D K. A structured epidemic model incorporating geographic-mobility among regions. *Mathematical Biosciences*. 1995; 128(1-2):71–91. [https://doi.org/10.1016/0025-5564\(94\)00068-B](https://doi.org/10.1016/0025-5564(94)00068-B) PMID: 7606146