A MECHANISTIC MODEL TO COMPARE THE IMPORTANCE OF INTERRELATED POPULATION MEASURES ON PATHOGEN RICHNESS: HOST POPULATION SIZE, DENSITY AND COLONY SIZE

TIM C.D. LUCAS[1*†], HILDE M. WILKINSON-HERBOTS[2] AND KATE E. JONES[3*]

Word count: 4647

[1]Centre for Biodiversity and Environment Research, Department of Genetics, Evolution and Environment, University College London, Gower Street, London, WC1E 6BT, United Kingdom. 0000-0003-4694-8107
[2]Department of Statistical Science, University College London, Gower Street, London, WC1E 6BT, United Kingdom
[3]Centre for Biodiversity and Environment Research, Department of Genetics, Evolution and Environment, University College London, Gower Street, London, WC1E 6BT, United Kingdom and Institute of Zoology, Zoological Society of London, Regent's Park, London, NW1 4RY, United Kingdon. 0000-0001-5231-3293
[†]Current address: Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, United Kingdom.
[*]Corresponding authors:
Tim C.D. Lucas: timcdlucas@gmail.com
Kate E. Jones: kate.e.jones@ucl.ac.uk

Abstract

Zoonotic diseases are an increasingly important source of human infectious diseases, and reservoir host pathogen richness is a critical driver of spill-over risk. Host population-level traits such as population size and density, geographic range size and population structure have all been shown to be important determinants of host pathogen richness. However, empirically identifying the independent influences of these traits has proven difficult as many of these traits directly depend on each other. Here we develop a mechanistic metapopulation susceptible-infected-recovered model to identify the independent life history trait influences on the ability of a newly evolved pathogen to invade and establish in host populations in the presence of an endemic pathogen, using a case study of bats; a highly social mammalian order. We show that larger population and group sizes had a greater influence on the chances of pathogen invasion and establishment than increased population densities (and therefore decreased population structure) and number of groups. As anthropogenic change affects these traits to different extents, this increased understanding of how traits independently determine pathogen richness will aid in predicting future zoonotic spill-over risk.

Keywords. Pathogen competition, zoonotic disease, metapopulations, host pathogen richness, bats, emerging infectious disease

1. Introduction

Zoonotic diseases are an increasingly important source of human infectious diseases [1–3]. The chance that a new zoonotic disease will come from a particular reservoir-host depends on a number of factors including the number of pathogen species it carries [4]. Much attention has been devoted to comparatively assessing the factors that are associated with high or low pathogen richness in wild animal species [9–11]. Many empirical comparative studies have examined morphological or life history traits [10,11], but factors related to reservoir-host population biology are also expected to affect disease dynamics and therefore affect pathogen richness. Population-level traits that have already been shown to correlate with pathogen richness include: increased host density [10,12–15], increased range size [10,13,14,16,17] and increased population structure [16,18]. Other traits such as group size have been well studied but results have been equivocal [14,19–22]. However, some important population-level traits, such as population size, have not been included in comparative analyses despite it being the natural way to describe epidemiological populations [38].

Collinearity between explanatory variables is a common problem in correlative studies. However, this issue is exacerbated when there are clear, causal relationships between explanatory variables. There are

two particularly clear relationships between the population-level factors associated with pathogen richness. Firstly, host density, d, host population size, N, and geographic range size, a, are, by definition, linked by $d = N/a$ (see Table S1 for all parameters used). Furthermore, the relationship $N \propto a$ has broad empirical support [23,24]. Secondly, host population size can be decomposed into two components, the number of groups, m, and the average size of a group, n, with $N = mn$. Therefore, correlative comparative studies will be especially poor at identifying which of these factors are closely correlated with pathogen richness. This lack of discrimintaroy power is particularly important with respect to global change and its effects on zoonotic disease emergence. Population-level factors such as host population size and geographic range size, although interrelated, will respond differently to global change and the response will be species specific. Some host species may suffer large range contractions, and therefore large falls in population size, while their density remains fairly constant [28]. Other host species might retain their distribution but have a depressed population density [29]. Therefore, only by knowing which of these interrelated factors controls pathogen richness will we be able to predict future changes in pathogen richness.

The alternative is mechanistic models. Firstly, mechanistic models provide a deeper understanding of the system than correlative approaches. Secondly, mechanistic models are expected to be more accurate when forecasting into the future or when extrapolating outside the range of the data it was fitted with. The ability of mechanistic models to extrapolate is Theoretical studies have established that a number of host population factors are important for epidemiological dynamics generally and for the maintenance of pathogen richness specifically. Host density and structure are well established as having central roles in pathogen dynamics [5–7,30–32]. Host group size is also known to strongly affect disease dynamics with disease spreading more quickly through populations made up of larger groups [7]. Fewer studies specifically study how these factors affect pathogen coexistence. A number of studies find that increased host population structure can promote pathogen coexistence [33–35]. While some theoretical studies have examined whether these population-level factors can promote pathogen richness, none have attempted to distinguish which might be the most important. Most studies have examined how pathogen coexistence depends on pathogen traits such as the transmission rate and virulence [34,36,37]. This focus is more relevant to studies of pathogens in humans and how different human pathogens may coexist. As wild animal populations are often very different to human populations these conclusions based on human populations cannot necessarily be transferred easily to wild animal species. However, it has been noted that host population size is a more natural measure than population density and that particularly in comparative settings, population size should be preferred [38]. This preference is due to the fact that host population size uniquely describes a property of the population while, for example, a high host density could be produced by a large population in a medium sized area or a medium sized population in a small area [38].

There is great need for mechanistic models that try to disentangle the interplay between these many factors: host density, population size, range size, population structure, group size and the number of groups. Here, we have used multipathogen, metapopulation models to individually vary these host population parameters. The metapopulations were parameterised to broadly mimic wild bat populations. We used bats as a case study as the size of bat groups (colonies) is very variable and bat colonies are often very stable [39,40]. Furthermore, bats are particularly relevant in the context of zoonotic disease as they are thought to be reservoirs for a number of important, recent outbreaks [41,42]. We examined how the interrelated population factors affect the ability of a newly evolved pathogen to invade and persist in a population in the presence of strong competition from an endemic pathogen strain. We used these simulations to test two specific hypotheses. First, we tested whether host population size or density more strongly promotes the invasion of a new pathogen. Secondly, we tested whether the invasion of a new pathogen is more strongly promoted by colony size or the number of colonies. We found that population size has a much stronger affect on the invasion of a new pathogen than host density and that increasing population size by increasing group size promotes pathogen invasion much more than increasing population size by increasing the number of groups.

2. Methods

(a). **Two pathogen SIR model.** We developed a multipathogen, SIR compartment model. We examined the ability of a second pathogen to invade into a population given the presence of an identical, endemic pathogen. Individuals were classed as susceptible, infected or recovered with immunity (Figure 1). Susceptible individuals are counted in class S (see Table S1 for a list of symbols and values used). There are three infected classes, I_1, I_2 and I_{12}, being individuals infected with Pathogen 1, Pathogen 2 or both respectively. Recovered individuals, R, are immune to both pathogens, even if they have only been infected with one (i.e. there is complete cross-immunity). Furthermore, recovery from one pathogen moves

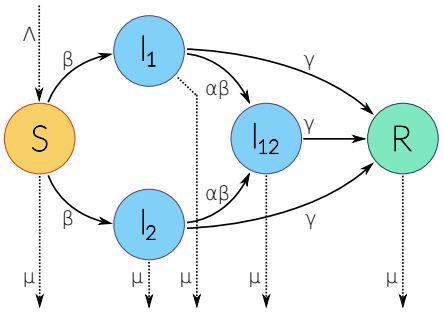

FIGURE 1. Schematic of the SIR model used. Individuals are in one of five classes, susceptible (orange, S), infected with Pathogen 1, Pathogen 2 or both (blue, I_1, I_2, I_{12}) or recovered and immune from further infection (green, R). Transitions between epidemiological classes occur as indicated by solid arrows. Vital dynamics (births and deaths) are indicated by dashed arrows. Parameter symbols for transitions are indicated. Note that individuals in I_{12} move into R, not back to I_1 or I_2. That is, recovery from one pathogen causes immediate recovery from the other pathogen.

an individual straight into the recovered class, even if the individual is infected with both pathogens (Figure 1). This modelling choice allows the model to be easily expanded to include more than two pathogens, though this study is restricted to two pathogens. The assumption of immediate recovery from all other diseases is likely to be reasonable. Any up-regulation of innate immune response will affect both pathogens equally. Furthermore, as the pathogens are identical, any acquired immunity would also affect both pathogens equally. The coinfection rate (the rate at which an infected individual is infected with a second pathogen) is adjusted compared to the infection rate by a factor α.

In the application of long term existence of pathogens it is necessary to include vital dynamics (births and deaths) as the SIR model without vital dynamics has no endemic state. Birth and death rates (Λ and μ) are set as being equal meaning the population does not systematically increase or decrease. New born individuals enter the susceptible class. Infection and coinfection were assumed to cause no extra mortality as for a number of viruses, bats show no clinical signs of infection [43,44].

The population is modelled as a metapopulation, being divided into a number of subpopulations (colonies). This model is an intermediate level of complexity between fully-mixed populations and contact networks. There is ample evidence that bat populations are structured to some extent. This evidence comes from the existence of subspecies, measurements of genetic dissimilarity and ecological studies [39,40,45,46]. Therefore a fully mixed population is a large oversimplification. However, trying to study the contact network relies on detailed knowledge of individual behaviour which is rarely available.

The metapopulation is modelled as a network with colonies being nodes and dispersal between colonies being indicated by edges (Figure 2). Individuals within a colony interact randomly so that the colony is fully mixed. Dispersal between colonies occurs at a rate ξ. Individuals can only disperse to colonies connected to theirs by an edge in the network. The rate of dispersal is not affected by the number of edges a colony has (known as the degree of the colony and denoted k). Therefore, the dispersal rate from a colony y with degree k_y to colony x is ξ/k_y. Note this rate is not affected by the degree and size of colony x.

We examined the model using stochastic, continuous-time simulations implemented in R [47]. The full details of the model are given in Supplementary Methods S1. The implementation is available as an R package on GitHub [48].

(b). **Determistic model.** We can get some initial insights into the behaviour of the system by examining a simplified, deterministic model with a single path. If we first consider the endemic pathogen (Pathogen 1) we have a typical SIR model with equilibrium values $S^* = \frac{\mu+\gamma}{\beta}$ and $I_1^* = \frac{\Lambda N}{\gamma+\mu}$. When Pathogen 2 is introduced, its rate of change can be written as

$$\frac{dI_2}{dt} = \beta S^* I_2 + \alpha\beta I_1^* I^2 - (\mu+\gamma)I_2 \quad (1)$$

which is greater than zero when $\frac{\alpha\Lambda N}{\gamma+\mu} > 0$ or $\alpha R_0 > 0$. As Pathogen 1 is endemic, and Pathogen 2 is identical to Pathogen 1, this inequality holds as long as α is greater than 0. That is, as long as cross-immunity is not complete, Pathogen 2 will always invade in this deterministic model.

(c). **Parameter selection.** The fixed parameters were chosen to roughly reflect realistic wild bat populations. The death rate μ was set as 0.05 per year giving a generation time of 20 years. The birth rate Λ was set to be equal to μ. This yields a population that does not systematically increase or decrease. However, the size of each colony changes as a random walk. Given the length of the simulations, colonies were very unlikely to go extinct. The starting size of each colony was set to 3000. This is appropriate for many bat species [49], especially the large, frugivorous *Pteropodidae* that have been particularly associated with recent zoonotic diseases.

The recovery rate γ was set to one, giving an average infection duration of one year. This is therefore a long lasting infection but not a chronic infection. It is very difficult to directly estimate infection durations in wild populations but it seems that these infections might sometimes be long lasting [50,51]. However, other studies have found much shorter infectious periods [52]. These shorter infections are not studied further here.

The coinfection adjustment parameter, α, was set to 0.1. Therefore, an individual with a single infection is 90% less likely to gain a second infection. Given Section (b), $\alpha = 0$ and $\alpha > 0$ are the two qualitatively different conditions. The case where Pathogen 2 does not invade and spread ($\alpha = 0$) is unlikely to be important for pathogen richness so we chose a small, non-zero value for α.

(d). **Population factors.** The effect of range size on disease dynamics occurred through changes in the metapopulation network. The metapopulation network was created for each simulation by randomly placing colonies in a square space (Figure 2). This square space was considered to be the species geographic range, the size of which was varied. Range size was varied between 2.5×10^3 and 4×10^4 km^2. This corresponds to square areas with sides of 50 to 200 km. Dispersal was only allowed to occur between two colonies if they were within 100km of each other i.e. they were then counted as connected nodes in the metapopulation network. The number of connections each colony has is called its degree, k. The mean degree, $\bar{k}$ is a measure of how well connected the metapopulation network is overall. The metapopulation network was not necessarily connected (i.e. made up of a single connected component) as the network was created by randomly placing colonies and only connecting colonies within 100 km of each other. To ensure connected metapopulation networks would have required repeatedly resampling of the colony locations until a connected metapopulation population occurred. However, this would bias $\bar{k}$. Therefore, it was considered preferential to keep the unconnected networks. The threshold of 100 km was arbitrary but we aimed to maximise the range of $\bar{k}$ (Figure S1) while not having many simulations with networks that were unconnected. Given this setup, populations with low densities had relatively unconnected metapopulation networks while high density populations had fully connected networks.

(e). **Experimental setup.** We let two identical pathogens compete: an endemic pathogen (Pathogen 1) and an invading pathogen (Pathogen 2). We used persistence (coded as 1) or extinction (coded as 0) of Pathogen 2 as a binomial response variable. We tested whether host population size is more important than host density. We then tested whether colony size or the number of colonies is the more important component of population size. We used three values of the transmission rate β, 0.1, 0.2 and 0.3. This yields very high values of R_0 which was required so that a reasonable number of simulations experienced invasion of Pathogen 2. All simulations were run under all three transmission rates.

In each simulation the host population was seeded with 20 individuals infected with Pathogen 1 in each colony. Pathogen 1 was then allowed to spread and reach equilibrium. After 6×10^5 events, five host individuals infected with Pathogen 2 were added to one randomly selected colony. The simulation was then run until a further 75 years had elapsed. The invasion of Pathogen 2 was considered successful if any individuals infected with Pathogen 2 remained at the end of the simulation.

Three sets of simulations were run. This set of three simulations was used to compare two pairs of population factors: *i*) population size and host density, *ii*) colony size and the number of colonies. The population parameters that were directly varied were colony size, the number of colonies and range size.

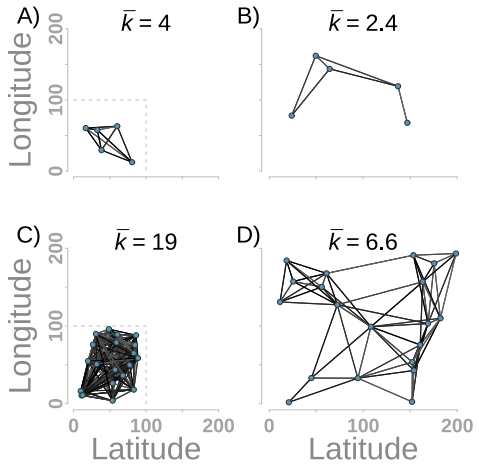

FIGURE 2. The relationship between range size and metapopulation network structure. Colonies are shown by circles. Colonies that are close enough for animals to disperse between (i.e. within 100 km of each other) are joined by a line. Colonies are placed randomly in spaces of various sizes (grey dashed lines). A and C) the default range size (10000 km^2). B and D) the largest range size (40000 km^2). A and B) the smallest number of colonies (five). C and D) the default number of colonies (20). The mean number of connections per subpopulation, $\bar{k}$, is shown for each metapopulation.

In each case these parameters were assigned their default value multiplied by 0.25, 0.5, 1, 2 and 4. The default colony size was 400, the default number of colonies was 20 and the default range size was 10000 km^2.

In the first set of simulations, host density was varied by keeping population constant while varying range size. Colony size was kept at a constant value of 400 while the number of colonies was fixed at 20 giving a population size of 8000. The values of range size used were 40000, 20000, 10000, 5000 and 2.5×10^3 km^2 which gave density values of 0.2, 0.4, 0.8, 1.6 and 3.2 animals per km^2.

In the second set of simulations, population size was varied by changing colony size while the number of colonies was kept constant. To keep host density constant, range size was reduced as population size increased. The values of colony size used were 100, 200, 400, 800 and 600 while range size was set to 40000, 20000, 10000, 5000 and 2.5×10^3 km^2. This gave population size values of 2000, 4000, 8000, 16000 and 32000 while host density remained at 0.8 hosts per km^2.

In the third set of simulations, population size was varied by changing the number of colonies while colony size was kept constant. Again, to keep host density constant, range size was reduced as population size increased. The numbers of colonies used were 5, 10, 20, 40 and 80 while range size was set to 40000, 20000, 10000, 5000 and 2.5×10^3 km^2. Again, this gave population size values of 2000, 4000, 8000, 16000 and 32000 while host density remained at 0.8 hosts per km^2.

To compare colony size and the number of colonies, only the second and third set of simulations above were used. However, colony size and the number of colonies were directly used as independent variables instead of using the derived values for host population size or density. It can be seen that population density and range size are equivalent in the two sets of simulations. Therefore, the only difference between these two sets of simulations is the factor used to increase population size: this factor being either colony size or the number of colonies.

(f). **Statistical analysis.** We tested two hypotheses. Firstly we tested the hypothesis that an increase in host population size creates a stronger increase in invasion probability (of the second pathogen) than an equal increase in host density. Secondly, we tested the hypothesis that an increase in colony size creates a

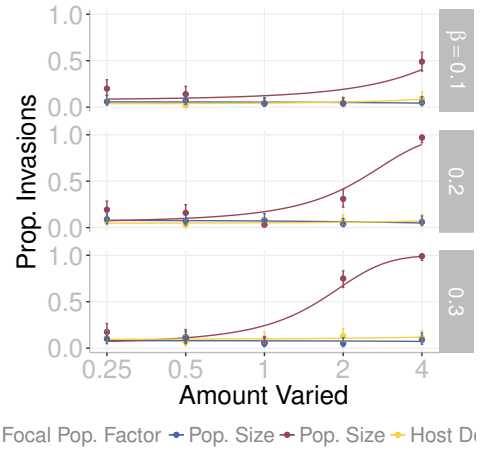

FIGURE 3. Comparison of the effect of colony size, number of colonies and host density on probability of invasion. The x-axis shows the change ($\times 0.25$, 0.5, 1, 2 and 4) in the each of these factors relative to the default value. Default values are: colony number = 20, colony size = 400 and density = 0.8 animals per km^2. Red lines: population size is altered by changing colony number. Blue lines: population size is altered by changing colony size. Yellow lines: population density is altered by changing range size. Each point is the mean of 100 simulations and bars are 95% confidence intervals. Curves are simple logistic regression fits for each independent variable. Relationships are shown separately for each transmission value, β.

stronger increase in invasion probability than a proportionally equal increase in number of colonies. To statistically test these hypotheses we combined the results from different simulations and fitted multiple logistic regressions, centering and scaling the independent variables. Specifically, we fitted the model

$$\text{Invasion} = b_1 d + b_2 n + b_3 m + c + \epsilon \tag{2}$$

where d, n and m are density, colony size and number of colonies respectively and b_i are the regression coefficients. c is a fitted intercept and ϵ is a binomially distributed error term. To test the first hypothesis we compared the size (and 95% confidence intervals) of b_1 to b_2 and b_3. To test the second hypothesis we compared b_2 to b_3.

Logistic regression was also used to test for an effect of transmission rate at the default parameter setting. Finally, in a small number of simulations both pathogens went extinct. Logistic regression was used to test whether transmission rate was associated with these events.

3. RESULTS

At the default parameter settings, the probability of invasion and establishment of the second pathogen, $P(I)$, was rare (Figure 3 and Tables S2–4). These proportions significantly increased with transmission rate (GLM: $b = 0.15$, $p = 0.1$).

In 9 simulations, both of the pathogens went extinct. The number of simulations where both pathogens went extinct did not depend on transmission rate (GLM: $b = -10 \times 10^{-3}$, $p = 0.22$). However all of the simulations with extinction of both pathogens had either the smallest colony size (colony size = 100, 9 simulations) or the fewest number of colonies (5 colonies, 0 simulations). Results from these simulations were removed before further analyses.

TABLE 1. Regression results comparing effects of colony size, colony number and density. Coefficients are from multiple logistic regressions with invasion as the dependent variable and all independent variables being scaled and centred. Colony size and colony number were varied while keeping density equal while density was varied by changing range size while keeping population size equal. p is for the test against the null hypothesis that $b = 0$.

β	*Variable*	*Estimate* (b)	(95 *CI*)	p
0.1	Intercept	-2.44	$(-2.64, -2.25)$	$< 10^{-5}$
	Colony Size	0.53	(0.33, 0.75)	$< 10^{-5}$
	Colony Number	-0.19	$(-0.48, 0.09)$	0.19
	Density	0.12	$(-0.2, 0.4)$	0.44
0.2	Intercept	-2.05	$(-2.23, -1.87)$	$< 10^{-5}$
	Colony Size	1.31	(1.07, 1.56)	$< 10^{-5}$
	Colony Number	-0.11	$(-0.39, 0.16)$	.3f
	Density	-0.07	$(-0.39, 0.22)$	0.67
0.3	Intercept	-1.54	$(-1.71, -1.38)$	$< 10^{-5}$
	Colony Size	1.92	(1.61, 2.26)	$< 10^{-5}$
	Colony Number	-0.12	$(-0.35, 0.11)$	0.31
	Density	-3.56×10^{-3}	$(-0.27, 0.24)$	0.98

(a). **Host population size or density.** To test whether host population size or density had a stronger effect on invasion probability we compared the regression coefficients of the multiple regressions fitted to simulation results (Figure 3). Increasing host population size, either by increasing colony size or number of colonies, increased the probability of invasion (Table 1). The relationship between colony size and invasion is strong and significant at all transmission rates, while the relationship between colony number and invasion is weaker and more marginally significant. In contrast, varying host density does not alter invasion probability. Therefore the simulations support the hypothesis that population size affects invasion more strongly than host density.

(b). **Colony size or number of groups.** To test whether colony size or the number of colonies is the more important component of population size, we compared the regression coefficients, b_2 and b_3, of the multiple regressions fitted to simulation results (Figure S2). Increasing either colony size or the number of colonies increased the probability of invasion but this effect was much stronger and more statistically significant for colony size (Table 1). Therefore the simulations support the hypothesis that colony size is the more important component of population size. This is supported by the deterministic model that showed that Pathogen 2 can invade even in a single subpopulation. As R_0 of the invading pathogen depends on the number of individuals in a subpopulation, invasion probability will increase with colony size.

4. DISCUSSION

Overall, our results suggest that population size promotes pathogen richness significantly more than host density in the context of metapopulations or group living. Furthermore, the results suggest that the most important component of population size is colony size; a large population made up of large colonies rather than many, small colonies promotes pathogen invasion to the greatest degree.

These results lead to a number of other conclusions. All else being equal, increasing range size (with density remaining constant) will not strongly increase pathogen richness unless the increased range size promotes larger groups. Furthermore, social species that live in large groups are likely to harbour more pathogen species, even if the larger groups require more space and therefore dispersal between groups is reduced.

The results suggest that, for related, strongly competing strains, the factor that most strongly allows new pathogens to invade is the number of susceptible individuals in the local group. As long as there are enough susceptible individuals that the new pathogen species does not go extinct during the stochastic, early stages of the epidemic, the new pathogen will persist. As dispersal is a very slow process compared to infection, the global pool of susceptibles is not important. This is probably why increasing the number of colonies did not increase pathogen invasion rate as quickly as the size of a colony did. Similarly, the global

host density of the species had little effect on pathogen invasion rate. In these simulations, increasing density without increasing population size was only achieved by reducing range size; this simply increased the number of connections between colonies in the metapopulation network. This, in turn, increased the pool of susceptibles that were within one dispersal of the invading pathogen. However, again, this effect was very weak compared to the strong changes in local disease dynamics caused by increasing colony size.

(a). **Global change.** It is clear that many species are suffering strong population changes due to global change [28]. However these changes might affect range size [28], population size [29], population connectivity [54–56] or group size [57–60] to different extents. Our results suggest that pathogen communities will respond differently depending on which factors are most strongly affected by global change. In short, species suffering reductions in groups size [57–60] are predicted to experience decreases in pathogen richness in the long term and there is some evidence that this process is occurring [16,61]. Species that are experiencing increases in group size [57] would be expected to gain new pathogen species. In contrast, species suffering range contractions [28] and decreases in population size [29] are expected to experience smaller changes in pathogen richness. However, it should be noted that these conclusions apply only to the specific mechanism studied here; that is, the invasion of newly evolved pathogens.

(b). **Comparative studies.** Many comparative studies measure some aspect of a species population size or structure, yet it is rarely discussed how these are related. Instead most studies use the data that are available, without considering *a priori* how the explanatory variables are causally related (though statistical correlations between independent variables are usually considered and dealt with using PCA or by removing collinear variables). Host density is often measured [12–15] yet density is directly associated with population size. Our results suggest that it is in fact population size that is important (in the context of social species as studied here). This leads to the suggestion that the density measures in these comparative studies are in fact proxies for population size rather than the true causal factor. Similarly, our results suggest that host range size does not promote pathogen richness by the mechanism studied here, yet a number of studies have found evidence of this relationship [10,14]. These differences suggest that either the relationship found in comparative studies is in fact due to a correlation with another factor, or that mechanisms other than probability of invasion of new pathogens are important. Range size has been suggested to affect pathogen richness by a number of mechanisms such as increasing the number of sympatric species and these other mechanisms should be specifically tested [11].

The studies that have specifically tested the effect of group size have in fact found both positive [19] and negative associations [20] or no relationship [21]. Meta-analyses suggest that the relationship between social group size and pathogen richness is weak [22]. However, we have found that group size is the most important population factor. This suggests that the mechanism studied here — invasion of recently evolved pathogens — may not be the major mechanism by which pathogen richness is created in wild populations.

We used the simple relationships between demographic factors — $d = N/a$ for example — to illustrate that these are tightly linked. In order to isolate the effects of these factors we assumed that there are no further correlations between factors; to examine density without altering population size, we fixed population size and manipulated range size. However, in reality, these factors are likely to covary in more complex ways, both within species across time and also between species. Therefore, while these factors are certainly linked, they cannot be assumed to have simple, linear relationships and should not be used as proxies of each other without further examination. For example, rates and distances of dispersal — which affect the influence of space — may be related to local density [62]. Similarly it is unlikely that a species whose range size decreases will not experience a decrease in total population size as well; the range contraction is likely to occur over generations rather than a simple squeezing of the existing individuals into a smaller area.

(c). **Conclusions.** Overall we have shown that while a number of demographic factors are intrinsically linked, they have different effects on the rate at which new pathogens will invade. We found that population size, not density, has the stronger impact on the ability of a pathogen to invade. Furthermore, species with large groups are likely to harbour more pathogens than species with many, smaller groups.

Data accessibility

The implementation of the model is available as an R package on GitHub [48]. This can be found at https://github.com/timcdlucas/MetapopEpi. All code and simulation output data is available on GitHub at https://github.com/timcdlucas/Abundance-Density-Manuscript.

Competing interests

We have no competing interests.

Author's contributions

TCDL wrote the simulations and performed the analysis. TCDL, HMH and KEJ all helped design the study. TCDL drafted the manuscript TCDL, HMH and KEJ all edited the manuscript and gave final approval for publication.

Acknlowledgements

We thank Andy Fenton and David Murrel for comments on the manuscript and help with the analytical model. The Dynamic Drivers of Disease in Africa Consortium, NERC project no. NE-J001570-1 was funded with support from the Ecosystem Services for Poverty Alleviation Programme (ESPA). The ESPA programme is funded by the Department for International Development (DFID), the Economic and Social Research Council (ESRC) and the Natural Environment Research Council (NERC)

Funding

This study was funded through a CoMPLEX PhD studentship at University College London supported by BBSRC and EPSRC (TCDL). KEJ was funded by the Ecosystem Services for Poverty Alleviation Programme (ESPA) (NE-J001570-1).

References

[1] KE Jones, NG Patel, MA Levy, A Storeygard, D Balk, JL Gittleman, and P Daszak. "Global trends in emerging infectious diseases". *Nature* 451.7181 (2008), pp. 990–993. DOI: 10.3201/eid0903.020336.

[2] ME Woolhouse and S Gowtage-Sequeria. "Host range and emerging and reemerging pathogens". In: *Ending the War Metaphor: The Changing Agenda for Unraveling the Host-Microbe Relationship: Workshop Summary*. Vol. 192. National Academies Press. Washington, D.C., 2006.

[3] LH Taylor, SM Latham, and E Mark. "Risk factors for human disease emergence". *Phil. Trans. R. Soc. B* 356.1411 (2001), pp. 983–989. DOI: 10.1093/acprof:oso/9780198525738.003.0035.

[4] ND Wolfe, MN Eitel, J Gockowski, PK Muchaal, C Nolte, A Tassy Prosser, J Ndongo Torimiro, SF Weise, and DS Burke. "Deforestation, hunting and the ecology of microbial emergence". *Global Change & Human Health* 1.1 (2000), pp. 10–25. DOI: 10.1023/A:1011519513354.

[5] RM Anderson and RM May. "Population biology of infectious diseases: Part I." *Nature* 280 (1979), pp. 361–7. DOI: 10.1038/280361a0.

[6] RM May and RM Anderson. "Population biology of infectious diseases: Part II". *Nature* 280.5722 (1979), pp. 455–461. DOI: 10.1038/280455a0.

[7] V Colizza and A Vespignani. "Invasion threshold in heterogeneous metapopulation networks". *Phys. Rev. Lett.* 99.14 (2007), p. 148701. DOI: 10.1103/physrevlett.99.148701.

[8] RM May and AL Lloyd. "Infection dynamics on scale-free networks". *Phys. Rev. E* 64.6 (2001), p. 066112. DOI: 10.1103/PhysRevE.64.066112.

[9] R Poulin and S Morand. "The diversity of parasites". *Q. Rev. Biol.* (2000), pp. 277–293.

[10] T Kamiya, K O'Dwyer, S Nakagawa, and R Poulin. "What determines species richness of parasitic organisms? A meta-analysis across animal, plant and fungal hosts". *Biol. Rev.* 89.1 (2014), pp. 123–134. DOI: 10.1111/brv.12046.

[11] AD Luis et al. "A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special?" *Proc. R. Soc. B* 280.1756 (2013), p. 20122753. DOI: 10.1098/rspb.2012.2753.

[12] S Morand and R Poulin. "Density, body mass and parasite species richness of terrestrial mammals". *Evol. Ecol.* 12.6 (1998), pp. 717–727. DOI: 10.1023/A:1006537600093.

[13] P Lindenfors, CL Nunn, KE Jones, AA Cunningham, W Sechrest, and JL Gittleman. "Parasite species richness in carnivores: effects of host body mass, latitude, geographical range and population density". *Global. Ecol. Biogeogr.* 16.4 (2007), pp. 496–509. DOI: 10.1111/j.1466-8238.2006.00301.x.

[14] CL Nunn, S Altizer, KE Jones, and W Sechrest. "Comparative tests of parasite species richness in primates". *Am. Nat.* 162.5 (2003), pp. 597–614. DOI: 10.1086/378721.

[15] P Arneberg. "Host population density and body mass as determinants of species richness in parasite communities: comparative analyses of directly transmitted nematodes of mammals". *Ecography* 1 (2002), pp. 88–94. DOI: 10.1034/j.1600-0587.2002.250110.x.

[16] AS Turmelle and KJ Olival. "Correlates of viral richness in bats (order Chiroptera)". *EcoHealth* 6.4 (2009), pp. 522–539. DOI: 10.1007/s10393-009-0263-8.

[17] S Huang, JM Drake, JL Gittleman, and S Altizer. "Parasite diversity declines with host evolutionary distinctiveness: a global analysis of carnivores". *Evolution* 69.3 (2015), pp. 621–630. DOI: 10.1111/evo.12611.

[18] GD Maganga, M Bourgarel, P Vallo, TD Dallo, C Ngoagouni, JF Drexler, C Drosten, ER Nakouné, EM Leroy, and S Morand. "Bat distribution size or shape as determinant of viral richness in african bats". *PLoS One* 9.6 (2014), e100172. DOI: 10.1371/journal.pone.0100172.

[19] ND Vitone, S Altizer, and CL Nunn. "Body size, diet and sociality influence the species richness of parasitic worms in anthropoid primates". *Evol. Ecol. Res.* 6.2 (2004), pp. 183–199.

[20] N Gay, KJ Olival, S Bumrungsri, B Siriaroonrat, M Bourgarel, and S Morand. "Parasite and viral species richness of Southeast Asian bats: fragmentation of area distribution matters". *Int. J. Parasitol. Parasites Wildl.* 3.2 (2014), pp. 161–170. DOI: 10.1016/j.ijppaw.2014.06.003.

[21] VO Ezenwa, SA Price, S Altizer, ND Vitone, and KC Cook. "Host traits and parasite species richness in even and odd-toed hoofed mammals, Artiodactyla and Perissodactyla". *Oikos* 115.3 (2006), pp. 526–536. DOI: 10.1111/j.2006.0030-1299.15186.x.

[22] JL Rifkin, CL Nunn, and LZ Garamszegi. "Do animals living in larger groups experience greater parasitism? A meta-analysis". *Am. Nat.* 180.1 (2012), pp. 70–82. DOI: 10.1086/666081.

[23] TM Blackburn, P Cassey, and KJ Gaston. "Variations on a theme: sources of heterogeneity in the form of the interspecific relationship between abundance and distribution". *J. Anim. Ecol.* 75.6 (2006), pp. 1426–1439. DOI: 10.1111/j.1365-2656.2006.01167.x.

[24] MK Borregaard and C Rahbek. "Causality of the relationship between geographic distribution and species abundance". *Q. Rev. Biol.* 85.1 (2010), pp. 3–25. DOI: 10.1086/650265.

[25] DG Jenkins, M Carey, J Czerniewska, J Fletcher, T Hether, A Jones, S Knight, J Knox, T Long, M Mannino, et al. "A meta-analysis of isolation by distance: relic or reference standard for landscape genetics?" *Ecography* 33.2 (2010), pp. 315–320. DOI: 10.1111/j.1600-0587.2010.06285.x.

[26] JF Le Galliard, A Rémy, RA Ims, and X Lambin. "Patterns and processes of dispersal behaviour in arvicoline rodents". *Mol. Ecol.* 21.3 (2012), pp. 505–523. DOI: 10.1111/j.1365-294X.2011.05410.x.

[27] RL Schooley and LC Branch. "Enhancing the area-isolation paradigm: habitat heterogeneity and metapopulation dynamics of a rare wetland mammal". *Ecol. Appl.* 19.7 (2009), pp. 1708–1722. DOI: 10.1890/08-2169.1.

[28] CD Thomas, A Cameron, RE Green, M Bakkenes, LJ Beaumont, YC Collingham, BF Erasmus, MF De Siqueira, A Grainger, L Hannah, et al. "Extinction risk from climate change". *Nature* 427.6970 (2004), pp. 145–148. DOI: 10.1038/nature02121.

[29] ID Craigie, JE Baillie, A Balmford, C Carbone, B Collen, RE Green, and JM Hutton. "Large mammal population declines in Africa's protected areas". *Biol. Cons.* 143.9 (2010), pp. 2221–2228. DOI: 10.1016/j.biocon.2010.06.007.

[30] M Barthélemy, C Godreche, and JM Luck. "Fluctuation effects in metapopulation models: percolation and pandemic threshold". *J. Theor. Biol.* 267.4 (2010), pp. 554–564.

[31] Q Wu, H Zhang, M Small, and X Fu. "Threshold analysis of the susceptible-infected-susceptible model on overlay networks". *Commun. Nonlinear Sci. Numer. Simul.* (2013).

[32] QM Webber and CK Willis. "Sociality, Parasites, and Pathogens in Bats". In: *Sociality in Bats*. Springer, 2016, pp. 105–139. DOI: 10.1007/978-3-319-38953-0_5.

[33] Z Qiu, Q Kong, X Li, and M Martcheva. "The vector–host epidemic model with multiple strains in a patchy environment". *J. Math. Anal. Appl.* 405.1 (2013), pp. 12–36. DOI: 10.1016/j.jmaa.2013.03.042.

[34] LJ Allen, N Kirupaharan, and SM Wilson. "SIS epidemic models with multiple pathogen strains". *J. Differ. Equ. Appl.* 10.1 (2004), pp. 53–75. DOI: 10.1080/10236190310001603680.

[35] A Nunes, MT da Gama, and M Gomes. "Localized contacts between hosts reduce pathogen diversity". *J. Theor. Biol.* 241.3 (2006), pp. 477–487. DOI: 10.1016/j.jtbi.2005.12.010.

[36] RM May and MA Nowak. "Superinfection, metapopulation dynamics, and the evolution of diversity". *J. Theor. Biol.* 170.1 (1994), pp. 95–114. DOI: 10.1006/jtbi.1994.1171.

[37] S Alizon. "Decreased overall virulence in coinfected hosts leads to the persistence of virulent parasites". *Am. Nat.* 172.2 (2008), E67–E79. DOI: 10.1086/588077.

[38] M Begon, M Bennett, RG Bowers, NP French, S Hazel, and J Turner. "A clarification of transmission terms in host-microparasite models: numbers, densities and areas". *Epidemiol. Infect.* 129.1 (2002), pp. 147–153. DOI: 10.1017/S0950268802007148.

[39] G Kerth, N Perony, and F Schweitzer. "Bats are able to maintain long-term social relationships despite the high fission–fusion dynamics of their groups". *Proc. R. Soc. B* 278.1719 (2011), pp. 2761–2767.

[40] GF McCracken and JW Bradbury. "Social organization and kinship in the polygynous bat *Phyllostomus hastatus*". *Behav. Ecol. Sociobiol.* 8.1 (1981), pp. 11–34. DOI: 10.1007/bf00302840.

[41] C Calisher, J Childs, H Field, K Holmes, and T Schountz. "Bats: important reservoir hosts of emerging viruses". *Clin. Microbiol. Rev.* 19.3 (2006), pp. 531–545. DOI: 10.1128/CMR.00017-06.

[42] W Li et al. "Bats are natural reservoirs of SARS-like coronaviruses". *Science* 310.5748 (2005), pp. 676–679. DOI: 10.1126/science.1118391.

[43] K Halpin, AD Hyatt, R Fogarty, D Middleton, J Bingham, JH Epstein, SA Rahman, T Hughes, C Smith, HE Field, et al. "Pteropid bats are confirmed as the reservoir hosts of henipaviruses: a comprehensive experimental study of virus transmission". *Am. J. Trop. Med. Hyg.* 85.5 (2011), pp. 946–951. DOI: 10.4269/ajtmh.2011.10-0567.

[44] B de Thoisy et al. "Bioecological Drivers of Rabies Virus Circulation in a Neotropical Bat Community". *PLoS Negl. Trop. Dis.* 10.1 (Jan. 2016). Ed. by M Kasper, e0004378. DOI: 10.1371/journal.pntd.0004378.

[45] LE Burns and HG Broders. "Correlates of dispersal extent predict the degree of population genetic structuring in bats". *Conserv. Genet.* 15.6 (2014), pp. 1371–1379.

[46] DE Wilson and DM Reeder. *Mammal species of the world: a taxonomic and geographic reference*. Vol. 12. JHU Press, 2005. ISBN: 9780801882210.

[47] R Development Core Team. *R: A language and environment for statistical computing*. R Foundation For Statistical Computing. Vienna, Austria, 2010. URL: http://www.R-Project.org.

[48] TCD Lucas. "MetapopEpi: Functions to run multipathogen, metapopulation epidemiological simulations" (2015). R package version 0.0.1. DOI: 10.5281/zenodo.48942. URL: https://github.com/timcdlucas/metapopepi.

[49] KE Jones, J Bielby, M Cardillo, SA Fritz, J O'Dell, CDL Orme, K Safi, W Sechrest, EH Boakes, C Carbone, et al. "PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals". *Ecology* 90.9 (2009), pp. 2648–2648. DOI: 10.1890/08-1494.1.

[50] A Peel, K Baker, G Crameri, J Barr, D Hayman, E Wright, C Broder, A Fernández-Loras, A Fooks, L Wang, et al. "Henipavirus neutralising antibodies in an isolated island population of African fruit bats". *PLoS One* 7.1 (2012), e30346. DOI: 10.1371/journal.pone.0030346.

[51] RK Plowright, P Eby, PJ Hudson, IL Smith, D Westcott, WL Bryden, D Middleton, PA Reid, RA McFarlane, G Martin, et al. "Ecological dynamics of emerging bat virus spillover". *Proc. R. Soc. B* 282.1798 (2015), p. 20142124. DOI: 10.1098/rspb.2014.2124.

[52] B Amengual, H Bourhy, M López-Roig, and J Serra-Cobo. "Temporal dynamics of European bat Lyssavirus type 1 and survival of *Myotis myotis* bats in natural colonies". *PLoS One* 2.6 (2007), e566. DOI: 10.1371/journal.pone.0000566.

[53] SJ Anthony et al. "A strategy to estimate unknown viral diversity in mammals". *MBio* 4.5 (2013), e00598–13. DOI: 10.1128/mBio.00598-13.

[54] TN Wasserman, SA Cushman, JS Littell, AJ Shirk, and EL Landguth. "Population connectivity and genetic diversity of American marten (*Martes americana*) in the United States northern Rocky Mountains in a climate change context". *Conserv. Genet.* 14.2 (2013), pp. 529–541. DOI: 10.1007/s10592-012-0336-z.

[55] F Rivera-Ortíz, R Aguilar, M Arizmendi, M Quesada, and K Oyama. "Habitat fragmentation and genetic variability of tetrapod populations". *Anim. Conserv.* 18.3 (2015), pp. 249–258. DOI: 10.1111/acv.12165.

[56] FE Fontúrbel, AB Candia, DA Salazar, J Malebrán, C González-Browne, and C Botto-Mahan. "How forest marsupials are affected by habitat degradation and fragmentation? A meta-analysis". *Naturwissenschaften* 101.7 (2014), pp. 599–602. DOI: 10.1007/s00114-014-1193-z.

[57] J Lehmann, AH Korstjens, and RI Dunbar. "Apes in a changing world — the effects of global warming on the behaviour and distribution of African apes". *J. Beogeogr.* 37.12 (2010), pp. 2217–2231. DOI: 10.1111/j.1365-2699.2010.02373.x.

[58] GE Zunino, MM Kowalewski, LI Oklander, and V Gonzalez. "Habitat fragmentation and population size of the black and gold howler monkey (*Alouatta caraya*) in a semideciduous forest in northern Argentina". *Am. J. Primatol.* 69.9 (2007), pp. 966–975. DOI: 10.1002/ajp.20389.

[59] R Manor and D Saltz. "Impact of human nuisance disturbance on vigilance and group size of a social ungulate". *Ecol. Appl.* 13.6 (2003), pp. 1830–1834. DOI: 10.1890/01-5354.

[60] TC Atwood. "The influence of habitat patch attributes on coyote group size and interaction in a fragmented landscape". *Can. J. Zool.* 84.1 (2006), pp. 80–87. DOI: 10.1139/z05-180.

[61] S Altizer, CL Nunn, and P Lindenfors. "Do threatened hosts have fewer parasites? A comparative study in primates". *J. Anim. Ecol.* 76.2 (2007), pp. 304–314. DOI: 10.1111/j.1365-2656.2007.01214.x.

[62] PH Marjamäki, AL Contasti, TN Coulson, and PD McLoughlin. "Local density and group size interacts with age and sex to determine direction and rate of social dispersal in a polygynous mammal". *Ecol. Evol.* 3.9 (2013), pp. 3073–3082. DOI: 10.1002/ece3.694.