INTERFACE

rsif.royalsocietypublishing.org

Research

Cite this article: Prentice JC, Marion G, Hutchings MR, McNeilly TN, Matthews L. 2017 Complex responses to movement-based disease control: when livestock trading helps. *J. R. Soc. Interface* **14**: 20160531.
http://dx.doi.org/10.1098/rsif.2016.0531

Received: 5 July 2016
Accepted: 2 December 2016

Subject Category:
Life Sciences–Mathematics interface

Subject Areas:
biomathematics, computational biology

Keywords:
basic reproduction ratio, *Escherichia coli* O157, bovine viral diarrhoea virus, *Mycobacterium avium* ssp. *paratuberculosis*, heterogeneity, supershedder

Author for correspondence:
Jamie C. Prentice
e-mail: jamie.prentice@glasgow.ac.uk

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.3601649.

THE ROYAL SOCIETY PUBLISHING

Complex responses to movement-based disease control: when livestock trading helps

Jamie C. Prentice[1,2], Glenn Marion[3], Michael R. Hutchings[4], Tom N. McNeilly[5] and Louise Matthews[1,2]

[1]Institute of Biodiversity, Animal Health and Comparative Medicine, College of Medical, Veterinary and Life Sciences, Glasgow G61 1QH, UK
[2]Boyd Orr Centre for Population and Ecosystem Health, Institute of Biodiversity, Animal Health and Comparative Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK
[3]Biomathematics and Statistics Scotland, Edinburgh EH9 3FD, UK
[4]Disease Systems, SRUC, Edinburgh EH9 3JG, UK
[5]Moredun Research Institute, Penicuik EH26 0PZ, UK

JCP, 0000-0002-5672-2163

Livestock disease controls are often linked to movements between farms, for example, via quarantine and pre- or post-movement testing. Designing effective controls, therefore, benefits from accurate assessment of herd-to-herd transmission. Household models of human infections make use of R_*, the number of groups infected by an initial infected group, which is a metapopulation level analogue of the basic reproduction number R_0 that provides a better characterization of disease spread in a metapopulation. However, existing approaches to calculate R_* do not account for individual movements between locations which means we lack suitable tools for livestock systems. We address this gap using next-generation matrix approaches to capture movements explicitly and introduce novel tools to calculate R_* in any populations coupled by individual movements. We show that depletion of infectives in the source group, which hastens its recovery, is a phenomenon with important implications for design and efficacy of movement-based controls. Underpinning our results is the observation that R_* peaks at intermediate livestock movement rates. Consequently, under movement-based controls, infection could be controlled at high movement rates but persist at intermediate rates. Thus, once control schemes are present in a livestock system, a reduction in movements can counterintuitively lead to increased disease prevalence. We illustrate our results using four important livestock diseases (bovine viral diarrhoea, bovine herpes virus, Johne's disease and *Escherichia coli* O157) that each persist across different movement rate ranges with the consequence that a change in livestock movements could help control one disease, but exacerbate another.

1. Background

Livestock diseases have an important impact not only on the economy and animal welfare [1,2], and can also pose a zoonotic risk to humans [3–5]. Many are introduced into herds via movements of infected animals, e.g. bovine tuberculosis (bTB), brucellosis, bovine viral diarrhoea (BVD), scrapie, foot-and-mouth disease (FMD) and Johne's disease [5–11]. Livestock disease control is therefore often implemented at the point of between-farm movement [12,13]. Controls that target infected animals moving between farms, include vaccination, quarantine, restricting movement for farms found to have infected animals, or even international movement restrictions [14]. This leads us to the question of how hard a

disease is to control when control is directed at herd-to-herd spread and how disease control effort depends on the rate of livestock movement.

The usual metric for assessing the required degree of control is the basic reproduction ratio, R_0 [15–17]. R_0 is the number of secondary infectives following introduction of a single typical primary infected individual into an entirely susceptible population. If $R_0 > 1$, then a disease can invade, whereas if $R_0 \leq 1$ it cannot. The aim of disease intervention is often described in terms of reducing effective R_0 to below 1. For example, if a proportion q of the population is vaccinated, then the effective R_0 is $R = (1 - q)R_0$, giving the critical coverage to prevent disease spread of $q_c = 1 - 1/R_0$ [15,18].

However, R_0 is an individual-based rather than a group-based metric, and a system with high R_0 could have high within-group (i.e. within farm) transmission, but only low between-group (between farm) transmission [19]; R_0 can therefore be poor at describing transmission within metapopulations [20], such as the risk of disease in one farm spreading to others, e.g. via livestock movements.

There have been several approaches to address this deficiency. Early patch-based models proved analytically tractable, but only considered the infected status of a patch as a whole, and assumed that the timescale of reaching a quasi-stationary state was short relative to movement dynamics [21–23]. This sort of simple model has sometimes failed to predict more complex and unintuitive disease dynamics [24]. Household models examine disease persistence within a metapopulation of a large number of small groups (e.g. households), and typically assume that disease spreads between groups that share individuals (e.g. children mix with other children at school, the adults mix with other adults at work and both return to the household), or that proximity is sufficient (e.g. transmission between patches of plant populations) [25–27]. However, household models neglect more long-lived movements such as those from livestock moving between farms, or wildlife dispersing from their natal range [25,28,29], and in doing so ignore the depletion of infectives from the primary group.

In the context of household models, Ball & Neal [30–32] introduce R_*, a group-level analogue of R_0, describing the number of secondary infected groups generated by a primary infected group. As with R_0, $R_* = 1$ provides a threshold for disease spread within the metapopulation. Similarly, $1 - 1/R_*$ provides the degree of disease intervention necessary to prevent disease spread. In situations where disease control is at the group-to-group level, R_* provides a convenient alternative to R_0 for predicting levels of disease control required, especially in highly heterogeneous metapopulations.

Here, we derive R_* using the next-generation matrix (NGM) technique [33,34], for a generic metapopulation model with disease spread via explicit animal movements that *does* account for the depletion of susceptibles from the primary group. In the simplest case, R_* takes an intuitive form in terms of the movement rate, the within-herd prevalence and the within-herd persistent time. We demonstrate that R_* peaks at intermediate movement rates, revealing ranges of movement rates where disease intervention will be most difficult.

We illustrate our findings for four important livestock infections—bovine herpes virus (BHV), bovine viral diarrhoea virus (BVDV), *Mycobacterium avium* ssp *paratuberculosis* (paraTB) and *Escherichia coli* O157 (*E. coli* O157)—showing that a reduction in movement rates could counterintuitively result in an increase in disease prevalence, and moreover, that control to reduce one disease could exacerbate another.

1.1. The next-generation matrix approach

In a recent helpful overview of NGM approaches, Diekmann *et al*. cover their use over a wide range of single population disease models [34,35], but do not consider metapopulation models. The NGM provides a natural basis for the calculation of R_0. In brief, the approach is to obtain a matrix $\boldsymbol{K}$, where the entries K_{ij} represent the expected number of new cases with state-at-infection i, arising from one individual with state-at-infection j. R_0 is the dominant eigenvalue of this matrix.

To illustrate the technique, consider a single population with SEIR disease dynamics:

$$\dot{S} = +\mu N - \mu S - \frac{\beta SI}{N},$$

$$\dot{E} = -\mu E + \frac{\beta SI}{N} - \alpha E,$$

$$\dot{I} = -\mu I + \alpha E - \gamma I$$

and

$$\dot{R} = -\mu R + \gamma I,$$

where individuals are born into susceptible state S, following infection they enter exposed state E and incubate the disease, then progress to the infectious state I, and finally recover to the immune state R. N is the population size, μ is the *per capita* mortality rate, here set equal to the birth rate, β is the disease transmission coefficient, $1/\alpha$ is the average incubation period and γ is the *per capita* recovery rate. The state space is the vector $\boldsymbol{x}(t) = (S, E, I, R)^\top$.

To obtain $\boldsymbol{K}$, first linearize around the disease-free equilibrium, $\boldsymbol{x}^*_{\mathrm{DF}} = (N, 0, 0, 0)^\top$, giving for small E and I the linearized infectious subsystem

$$\dot{E} = +\beta I - (\mu + \alpha)E$$

and

$$\dot{I} = +\alpha E - (\mu + \gamma)I,$$

where only the production of new infectives and changes in the state of existing infectives are captured. The linearized subsystem is the form $\dot{\boldsymbol{y}} = \boldsymbol{A}\boldsymbol{y}$, where $\boldsymbol{y}(t) = (E, I)^\top$ and

$$\boldsymbol{A} = \begin{pmatrix} -\mu - \alpha & +\beta \\ +\alpha & -\mu - \gamma \end{pmatrix},$$

is the Jacobian matrix.

Now, decompose $\boldsymbol{A}$ into the sum of two matrices $\boldsymbol{T} + \boldsymbol{\Sigma}$, where

$$\boldsymbol{T} = \begin{pmatrix} 0 & +\beta \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} T_{EE} & T_{EI} \\ T_{IE} & T_{II} \end{pmatrix},$$

is the matrix of *transmissions*, where T_{EI} represents the rate at which newly infected individuals in state E are created by infectious individuals I, and

$$\boldsymbol{\Sigma} = \begin{pmatrix} -\mu - \alpha & 0 \\ +\alpha & -\mu - \gamma \end{pmatrix} = \begin{pmatrix} \Sigma_{EE} & \Sigma_{EI} \\ \Sigma_{IE} & \Sigma_{II} \end{pmatrix}$$

is the matrix of *transitions*, where, for example, Σ_{IE} is the rate at which individuals move into state I from state E. Negative entries represent a net flow out of the state in question; hence, Σ_{EE} shows the rate at which individuals that start in E leave

this class, without returning. The matrix

$$-\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \dfrac{1}{\mu+\alpha} & 0 \\ \dfrac{\alpha}{(\mu+\alpha)(\mu+\gamma)} & \dfrac{1}{\mu+\gamma} \end{pmatrix},$$

is interpreted biologically as the matrix of *sojourn times* [34]. Thus, the entries of the first column of matrix $-\boldsymbol{\Sigma}^{-1}$ are the expected time spent in states E and I conditional on starting in state E (likewise entries of the second column are the expected times conditional on starting in state I).

The NGM with large domain, $\boldsymbol{K}_{\mathrm{L}}$, is given by the matrix product of transmission rate and residence time, that is $\boldsymbol{K}_{\mathrm{L}} = -\boldsymbol{T}\boldsymbol{\Sigma}^{-1}$ [33], and so

$$\boldsymbol{K}_{\mathrm{L}} = \begin{pmatrix} \dfrac{\alpha\beta}{(\mu+\alpha)(\mu+\gamma)} & \dfrac{\beta}{(\mu+\gamma)} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} K_{EE} & K_{EI} \\ K_{IE} & K_{II} \end{pmatrix},$$

where, for example, K_{EI} is the number of infections of type E generated by an index case in the I class. R_0 is then the dominant eigenvalue of $\boldsymbol{K}_{\mathrm{L}}$

$$R_0 = \rho(\boldsymbol{K}_{\mathrm{L}}) = \frac{\alpha\beta}{(\mu+\alpha)(\mu+\gamma)}.$$

By including only the rows and columns of $\boldsymbol{K}_{\mathrm{L}}$ related to categories of state-at-infection (i.e. exposed E, but not infectious I), $\boldsymbol{K}_{\mathrm{L}}$ can be reduced to the NGM matrix $\boldsymbol{K}$

$$\boldsymbol{K} = \left(\frac{\alpha\beta}{(\mu+\alpha)(\mu+\gamma)} \right) = (\, K_{EE}\,),$$

which is smaller and mathematically easier to work with, and has a biological interpretation convenient for direct construction using epidemiological principles [34]. The dominant eigenvalue is the same for both $\boldsymbol{K}_{\mathrm{L}}$ and $\boldsymbol{K}$, and either may be used to calculate R_0.

In the scalar case above, $R_0 = \boldsymbol{K}_{\mathrm{EE}}$, which, on taking the limits $\alpha \to \infty$ and $\mu \to 0$, reduces to the familiar $R_0 = \beta/\gamma$ for the SIR model. This calculation for the SIR model also follows by identifying transmission $\boldsymbol{T} = \beta$ and the transition rate $\boldsymbol{\Sigma} = -\gamma$, whence the time spent infectious is $-\boldsymbol{\Sigma}^{-1} = 1/\gamma$, and the expected number of secondary infections from an index case in an otherwise susceptible population is $R_0 = -\boldsymbol{T}\boldsymbol{\Sigma}^{-1} = \beta/\gamma$. The NGM approach thus rigorously extends such arguments to more complex settings.

2. Next-generation matrix approach for homogeneous metapopulation dynamics with one disease category

We now apply the NGM approach to disease spread among a metapopulation of livestock herds, first illustrating the approach for a disease system with one disease category, and then showing how this may be naturally extended to more complex diseases. We show that R_* may be given by the intuitive formula

$$R_* = \kappa N P_{\mathrm{pos}} T_{\mathrm{inf}},$$

with *per capita* movement rate κ, herd size N, herd expected infectious lifetime T_{inf} and average prevalence of infectives during the infectious lifetime P_{pos}. This is conceptually similar to the SIR model formula $R_0 = \beta/\gamma$ if one considers substituting the rate at which new infectious individuals are formed, β, with the average rate at which infectives leave herds $\kappa N P_{\mathrm{pos}}$, and substituting in the expected infectious period, $1/\gamma$, with the expected time disease persists in the herd, T_{inf}.

2.1. Derivation of R_*

Consider an SIS disease dynamic in a metapopulation of herds each containing N individuals. In the absence of infection, individuals die and are replaced with susceptibles at *per capita* rate μ. We assume frequency-dependent disease transmission with transmission rate β, recovery at *per capita* rate γ and a *per capita* movement rate between herds of κ.

For analytic tractability, we maintain constant herd size by assuming that the birth and death processes are coupled. Thus, the status of each herd may be defined by just the number of infectives, i, because the number of susceptibles is $s = N - i$. We consider a homogeneous metapopulation where we assume undirected movement between herds, and that movements are equally likely between any herds.

We choose to represent the metapopulation dynamics using the master equation approach (also known as the Chapman–Kolmogorov forward equation, see [36] for a detailed explanation) that allows us to capture the probability of a herd being in a state with i infectives. We begin by considering the respective rates $l(i)$ and $g(i)$ at which infectives are lost and gained. We assume when an animal leaves a herd it is replaced with a susceptible or infected individual in proportion to their prevalence in the metapopulation. Consequently, $l(i)$ and $g(i)$ depend on P_{S} and P_{I}, the mean prevalence of susceptibles and infectives in the metapopulation, i.e.

$$P_{\mathrm{I}} := \frac{1}{N}\sum_{i=0}^{N} i p_i, \quad P_{\mathrm{S}} := 1 - P_{\mathrm{I}},$$

where p_i is the probability that a herd contains i infectives.

In a herd with i infectives, the net loss of infectives owing to movements is $\kappa P_{\mathrm{S}} i$ (because a proportion P_{S} of replacements are susceptible); therefore, the net loss of infectives via mortality, recovery and movement is

$$l(i) = \mu i + \gamma i + \kappa P_{\mathrm{S}} i.$$

Similarly, the net gain of infectives owing to movements is equal to the net loss of susceptibles, which is given by $\kappa P_{\mathrm{I}} s$. Therefore, the net gain of infectives via disease transmission and movement is

$$g(i) = \frac{\beta s i}{N} + \kappa P_{\mathrm{I}} s,$$

where $s = N - i$.

We may now write down the master equation governing the probability $p_i(t)$ of a herd containing i infectives at time t:

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = +g(i-1)p_{i-1} - [g(i) + l(i)]p_i + l(i+1)p_{i+1},$$

for $i = 1,2,\ldots,N$ and subject to

$$p_{N+1} = 0, \quad \text{and} \quad p_0 = 1 - \sum_{i=1}^{N} p_i.$$

Here, $\boldsymbol{p}(t) = (p_0,\ldots,p_N)^{\top}$ is a vector of length $N+1$. Removing the disease-free state $i = 0$, gives $\boldsymbol{q}$(t), a vector of length N describing the probability of i infectives in the infectious subsystem.

To determine R_*, we first linearize around the disease-free state $\boldsymbol{q}^*_{\mathrm{DF}} = (0,\ldots,0)^{\top}$. For q_i close to the disease-free state for

$i = 1, \ldots, N$, we obtain

$$\frac{\mathrm{d}q_1}{\mathrm{d}t} = +\kappa P_I N - \left[\frac{\beta s}{N} + \mu + \gamma + \kappa\right] q_1 + 2(\mu + \gamma + \kappa) q_2$$

and

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = +\frac{\beta(s+1)}{N}(i-1)q_{i-1} - \left[\frac{\beta s}{N} + \mu + \gamma + \kappa\right] i q_i + (\mu + \gamma + \kappa)(i+1) q_{i+1}.$$

This can be written in matrix form

$$\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t} = \boldsymbol{A}\boldsymbol{q} = (\boldsymbol{T} + \boldsymbol{\Sigma})\boldsymbol{q},$$

where $\boldsymbol{A}$ is the Jacobian matrix, and is decomposed into $(\boldsymbol{T} + \boldsymbol{\Sigma})$, where $\boldsymbol{T}$ is the matrix of transmissions, and $\boldsymbol{\Sigma}$ is the matrix of transitions. Here

$$T_{ij} = \kappa \delta_{i1} j,$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise, i.e.

$$\boldsymbol{T} = \begin{pmatrix} \kappa 1 & \kappa j & \kappa N \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

and

$$\Sigma_{ij} = \begin{cases} +\dfrac{\beta(s+1)}{N}(i-1) & \text{if } j = i-1 \\ -\left(\dfrac{\beta s}{N} + \mu + \gamma + \kappa\right) i & \text{if } j = i \\ +(\mu + \gamma + \kappa)(i+1) & \text{if } j = i+1 \\ 0 & \text{otherwise} \end{cases}.$$

For the metapopulation model, the interpretation of $\boldsymbol{T}$ and $\boldsymbol{\Sigma}$ differs from the single population model as follows. In a single herd model, $\boldsymbol{T}$ describes the production of new infections via within-herd transmission [34]; however, in the metapopulation model, it represents the production of new infected *herds* via movement of infected individuals from an infected herd to a susceptible one. In a single herd model, $\boldsymbol{\Sigma}$ represents transitions between different disease states; in the metapopulation model, it represents the transitions between different states (in this case, different numbers of infectives) of an infected herd via within-herd transmissions, recoveries or mortalities, and movement of infectives to already infected herds.

As above, the matrix $\boldsymbol{S} = -\boldsymbol{\Sigma}^{-1}$ is the matrix of *sojourn times*, where the entry S_{ij} is the expected time that a herd currently observed in state j will thereafter spend in state i. Because infected herds are assumed to begin with a single infective (i.e. $i = 1$), the total expected infectious period, T_{inf}, is the sum of the times spent in each state, i.e. the sum of the entries in *column* 1, gives

$$T_{\mathrm{inf}} := \sum_{i=1}^{N} S_{i1}.$$

As $\boldsymbol{T}$ is zero everywhere other than the first row, the NGM of large domain $\boldsymbol{K}_{\mathrm{L}} = -\boldsymbol{T}\boldsymbol{\Sigma}^{-1}$ is also zero everywhere except the first row (in this case, the dominant eigenvalue of $\boldsymbol{K}_{\mathrm{L}}$ is equal to the first entry of $\boldsymbol{K}_{\mathrm{L}}$). The only state-at-infection is I_1, and so $\boldsymbol{K} = [\boldsymbol{K}_{\mathrm{L}}]_{11}$. Thus, R_* is given by

$$R* = K_{11} = \kappa \sum_{i=1}^{N} i S_{i1}.$$

The expected proportion of time spent by a herd in state i, having started in state 1, is given by S_{i1}/T_{inf}. Using this, we now define the expected prevalence in an infected herd, P_{pos}, by

$$P_{\mathrm{pos}} := \frac{1}{N}\sum_{i=1}^{N} i S_{i1}/T_{\mathrm{inf}} = \frac{1}{N}\frac{R_*}{\kappa T_{\mathrm{inf}}}$$

Rearranging, we obtain

$$R* = \kappa N P_{\mathrm{pos}} T_{\mathrm{inf}}. \tag{2.1}$$

Therefore, R_*, the expected number of secondary infected herds, is (intuitively) given by the product of the expected rate at which infectives leaving a herd ($\kappa N P_{\mathrm{pos}}$), and the duration of the infection in a herd (T_{inf}). This form is instructive, both because of its close relation to the definition of R_0 via β and γ, and because calculating $\boldsymbol{\Sigma}^{-1}$ directly may be computationally infeasible for even moderately complicated models, but it can be relatively straightforward to calculate P_{pos} and T_{inf} numerically (see §2 of the electronic supplementary material).

2.2. Dependence of R_* on movement rate, R_0, heterogeneity and implications for control

2.2.1. Features of R_*

In this section, we illustrate the features of R_* within a metapopulation of herds using SIS model dynamics. We use the formulation of R_0 that reflects the primary infective's total capacity to generate secondary cases, irrespective of movement between herds (see §1.1). Then, for an underlying $R_0 > 1$, R_* is zero in the absence of movements, rises above 1 as the movement rate increases, peaks at an intermediate movement rate, and then declines to 1 from above (figure 1*a*). Note that, for an underlying $R_0 \leq 1$, R_* approaches 1 for large movement rates from below (not shown). R_* provides a threshold for persistence of infection in the metapopulation as indicated by the quasi-equilibrium proportion of infected herds: zero for $R_* \leq 1$, and greater than zero for $R_* > 1$ (figure 1*b*).

R_* initially rises, because the disease multiplies within the herd before infectives are exported to other herds via movement. However, R_* eventually declines as it becomes more likely that the primary infective leaves the herd before it has a chance to transmit infection within the herd (or recover or die). This results in an intermediate peak occurring when movement is low enough that the disease is sustained within the herd, but fast enough that it can reach other herds before being removed by stochastic extinction.

The peak in R_* increases in magnitude and shifts to lower movement rates as R_0 increases (figure 2*a*) with a corresponding shift in the threshold for persistence (figure 2*b*). In addition, for the same R_0, slowly progressing diseases (i.e. those with a low recovery rate; figure 2*a*, red curves) have a higher equilibrium proportion of infected herds at the same movement rate than a rapidly progressing disease (i.e. those with a high recovery rate; figure 2*a*, blue curves), with corresponding shifts in the threshold for persistence (figure 2*b*).

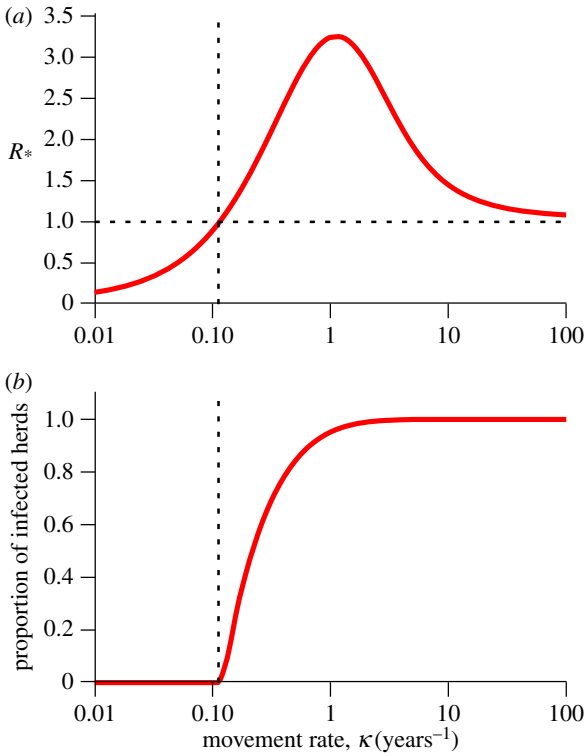

Figure 1. R_* is a threshold parameter for persistence in the metapopulation. (*a*) R_* versus movement rate κ in the metapopulation SIS model (see §1.1 of the electronic supplementary material). (*b*) The quasi-equilibrium proportion of infected herds, for the same model. The proportion of infected herds is 0 when $R_* \leq 1$, illustrating threshold behaviour and increases to 1 as movement rates increases. Parameters are $\mu = 1/3$, $\gamma = 10$ and $\beta = 14$.

From the expression for R_* (equation (2.1)), we see that herd size N contributes to R_* via the number of movements out of the herd and also via its potential effect on the prevalence when infected, P_{pos}, and the persistence time, T_{inf}. The net result is substantial nonlinear increases in R_* with increased herd size, but relatively little change in the position of the peak in R_* (figure 2*c*).

2.2.2. Implications for control

To simulate the effect of control measures, we define, for any set of individuals selected to move, p to be the proportion of infectious individuals that are treated or prevented from bringing the disease into another group. This interception occurs at the point of movement, therefore, only the remaining proportion $1 - p$ of individuals successfully carry the infection to another herd. We therefore obtain the effective reproduction ratio in the presence of control, $R_*(p)$, which is

$$R*(p) := (1 - p)R*.$$

This leads to an important result. Disease can spread in the presence of control only if $R_*(p)$ remains above 1, leading to 'islands' of persistence (figure 3*b*).

Using our metapopulation model (see §2 of the electronic supplementary material), we calculated $R_*(p)$ for a range of levels of disease intervention and the corresponding equilibrium proportion of infected herds in the metapopulation. Near the R_* peak (intermediate movement rate), even high levels of disease intervention may fail to control the disease (figure 3*a*, yellow curve), but when the movement rate is high even low levels of control may be sufficient to reduce $R_*(p)$ to below 1 and prevent the disease from spreading.

Consequently, for a range of intermediate values of the movement rate the infection persists in the meta-population for a given level of control, but at higher or lower movement rates, infection cannot persist under the same level of control (e.g. figure 3*b*, yellow curve). The range of values of the movement rate for which disease persists depends on the level of control applied.

3. Next-generation matrix approach for heterogeneous systems

In this section, we demonstrate how R_* may be constructed for the more complex disease systems and explore the impact of such heterogeneity on R_*.

3.1. Multiple disease categories

Consider a disease with two possible infectious states: types A and B (e.g. a regular shedder and a supershedder). As above, we assume the herd size N is constant, so the herd has potential disease states $x_{a,b}$, where a and b correspond to the number of individuals in a herd in categories A and B, respectively; here $0 \leq a + b \leq N$, and the disease-free state is $S = x_{0,0}$. The state space $\boldsymbol{x}(t)$ is obtained by enumerating over all the possible infected herd states, and then the NGM with large domain $\boldsymbol{K}_{\text{L}}$ needed to calculate R_* may be constructed by proceeding as before. Here, we illustrate the process.

Because infection in a herd is initiated by one individual, a disease with a single infectious category has one entry point x_1, whereas with two infectious categories, there are two entry points: $x_{1,0}$ and $x_{0,1}$, depending on which type of infective first enters the disease-free herd. Therefore, in this case, the transmission matrix $\boldsymbol{T}$ (and hence also $\boldsymbol{K}_{\text{L}}$) has two rows with non-zero entries, and therefore, $\boldsymbol{K}$ is a 2×2 matrix.

Any given herd state can be reached by a limited number of adjacent states via the various event types, and each row in $\boldsymbol{\Sigma}$ will have as many entries as possible transitions (e.g. here under the constraint of fixed herd size N there are four, corresponding to increases in A and B due to infection, and decreases owing to recovery or mortality). The matrix $\boldsymbol{S} = -\boldsymbol{\Sigma}^{-1}$ of sojourn times is dense, but as above we exploit the fact that the columns corresponding to the entry points determine the total infection duration T_{inf}, which now depends on which entry point is reached (i.e. we must now consider both T^A_{inf} and T^B_{inf}).

Extracting the elements of $\boldsymbol{K}_{\text{L}}$ relating to the entry points, we obtain the reduced NGM $\boldsymbol{K}$, which has entries:

$$\boldsymbol{K} = \begin{pmatrix} \kappa N P^A_{\text{pos}}(A) T^A_{\text{inf}} & \kappa N P^B_{\text{pos}}(A) T^B_{\text{inf}} \\ \kappa N P^A_{\text{pos}}(B) T^A_{\text{inf}} & \kappa N P^B_{\text{pos}}(B) T^B_{\text{inf}} \end{pmatrix} = \begin{pmatrix} K_{AA} & K_{AB} \\ K_{BA} & K_{BB} \end{pmatrix}$$

where, for example, $P^A_{\text{pos}}(B)$ means the expected prevalence of B given entry point A, and T^A_{inf} is the expected duration of the infection in the herd given entry point A.

Here, K_{BA} is the number of secondary herds initially infected by a class B individual that are caused by a primary

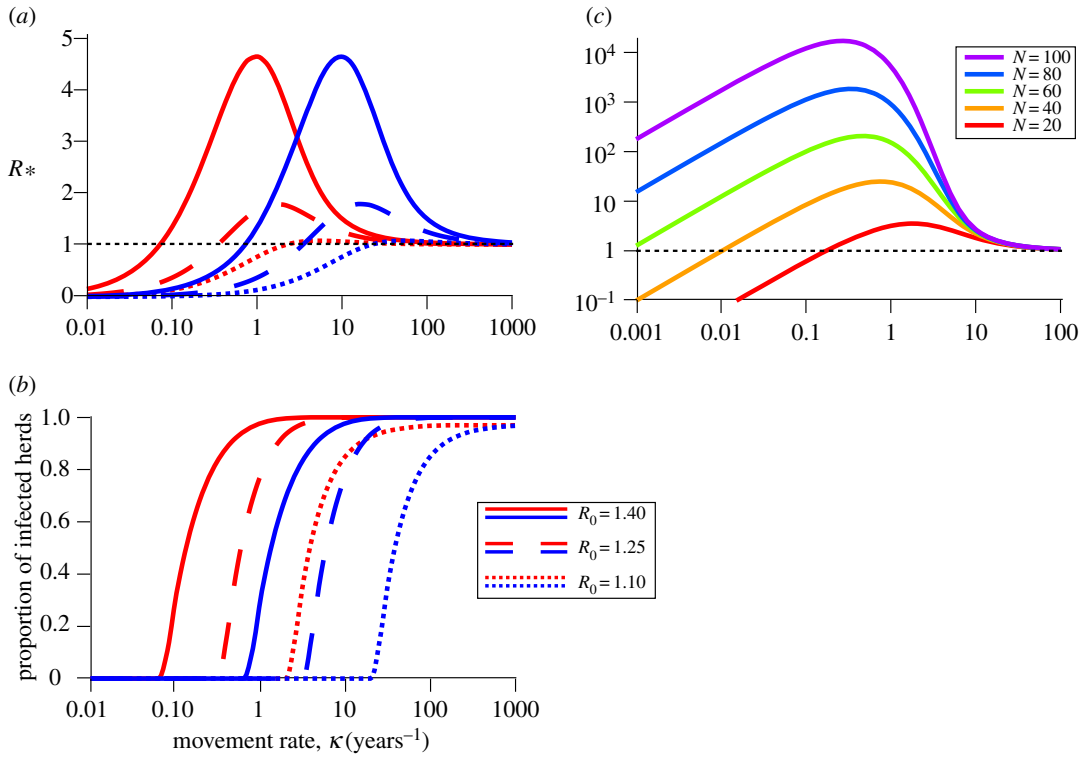

Figure 2. Impact of R_0, herd size and movement rate on R_*. (*a*) R_* in the metapopulation SIS model (see §1.1 of the electronic supplementary material). (*b*) Quasi-equilibrium proportion of infected herds in the same model, versus movement rate κ for slowly progressing (red) and rapidly progressing (blue) diseases, and for varying R_0. The R_* peak occurs for lower movement rate in the slowly progressing disease, and increases rapidly with R_0. Parameters are $\mu = 1/3$, $\gamma = 10$ and $\beta = 11.4, 12.9, 14.5$. (*c*) R_* (log scale) versus increasing herd size N in the same model. The R_* peak increases roughly exponentially with N. Parameters are $\mu = 1/3$, $\gamma = 10$ and $\beta = 17.5$.

infected herd initially infected by a class A individual. This is convenient, as it means that the entries in $\boldsymbol{K}$ may be computed via simulation of a single herd, by seeding an infection with an infective of category j, and approximating each entry K_{ij} as the number of category i infectives moving to susceptible herds (averaged over a sufficiently large number of repeated simulations).

Note that because R_* is a function of all entries in $\boldsymbol{K}$, it is possible that a disease with multiple infectious categories may have multiple R_* peaks (this phenomenon is just distinguishable in the curve for BHV in figure 5).

3.1.1. Example illustrating the impact of within-herd heterogeneity in infectiousness

Consider a livestock infection such as *E. coli* O157, which exhibits substantial heterogeneity between individuals in transmissibility [37–39]. Here, we characterize this heterogeneity using a simple low shedder–high shedder version of an SIS model which we call the SLHS model (see §1.2 of the electronic supplementary material). We assumed that susceptibles S become either supershedders H (high) or regular infectives L (low), with probabilities p and $1 - p$, respectively, and that supershedders are η times more infectious than regular infectives. To illustrate the effect of heterogeneity on R_*, we chose η, p and a normalizing constant (see §1.2 of the electronic supplementary material) to ensure that R_0 remains constant as we vary the relative contributions to transmission from the low and high shedders.

We calculate R_* by simulating herds where the initial infection is either a low or high shedder, and each case populates a column in the NGM

$$\boldsymbol{K} = \begin{pmatrix} \kappa N P^L_{\mathrm{pos}}(L) T^L_{\mathrm{inf}} & \kappa N P^H_{\mathrm{pos}}(L) T^H_{\mathrm{inf}} \\ \kappa N P^L_{\mathrm{pos}}(H) T^L_{\mathrm{inf}} & \kappa N P^H_{\mathrm{pos}}(H) T^H_{\mathrm{inf}} \end{pmatrix},$$

as described in §3.1.

The highest R_* comes from the most homogeneous disease transmission (figure 4*a*). The explanations for this are a combination of (i) susceptible depletion, i.e. while R_0 (which ignores susceptible depletion) remains constant, the initial supershedder, when highly infectious, is unable to reach its full potential owing to a lack of susceptibles and (ii) an increased chance of stochastic extinction when the majority of the transmission is due to the relatively rare supershedders.

3.2. Between-herd heterogeneity

We now consider the case of an SIS disease with heterogeneity in herd size N and movement rate κ. Suppose the population consists of n herd types where a proportion p_j of herds have N_j individuals and *per capita* movement rate κ_j. The mean herd size is $\langle N \rangle = \sum_j p_j N_j$, and the mean movement rate is $\langle \kappa N \rangle = \sum_j p_j \kappa_j N_j$. Here the state vector is of size $M = \sum_j N_j$, representing the numbers of infectives for each herd size $\{I^1_1, \ldots, I^1_{N_1}, \ldots, I^n_1, \ldots, I^n_{N_n}\}$, where I^j_i is the number of herds of type j with i infectives.

The transmission matrix $\boldsymbol{T}$ is of size $M \times M$, but has only n entry points, corresponding to I^1_1 to I^n_1 and is therefore composed entirely of zeros except for n rows. The transition matrix $\boldsymbol{\Sigma}$ is an $M \times M$ block matrix, where each diagonal block is a tridiagonal submatrix of size $N_j \times N_j$ (because the only state change is to increase or decrease the number of

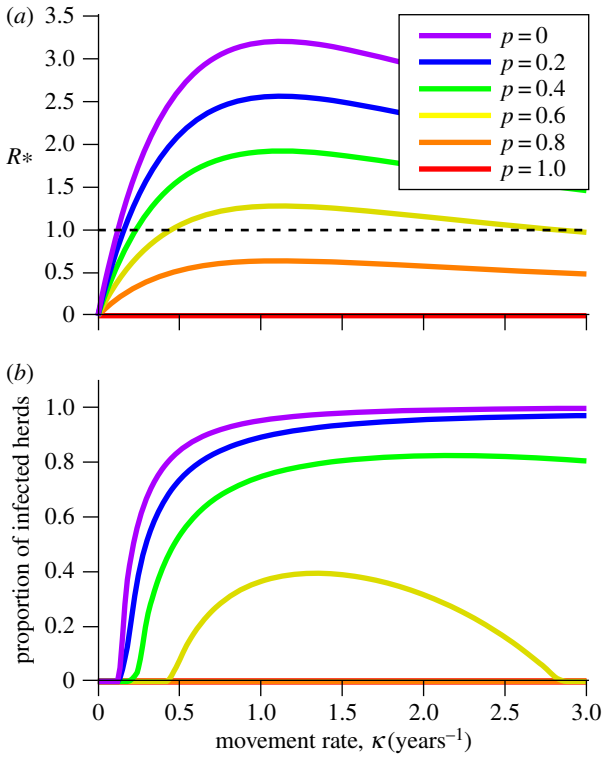

Figure 3. Effect of movement-based controls on R_*. (*a*) R_* in the meta-population SIS model (see §1.1 of the electronic supplementary material) and (*b*) equilibrium proportion of infected herds, in the same model, plotted against movement rate κ, and for increasing disease prevention p, from $p = 0$ (violet) to $p = 1$ (red), shown at intervals of 0.2. Effective $R_*(p) \to 1 - p$ for large movement rates and the disease can persist whenever effective $R_*(p) > 1$. This creates an intermediate range of movement rates for which the disease is able to persist ('islands' of persistence) at the specified level of control. Parameters are $\mu = 1/3$, $\gamma = 10$ and $\beta = 14$.

infectives by 1), and each off-diagonal block is an $N_i \times N_j$ zero submatrix (because herds do not change size). $\boldsymbol{\Sigma}$ is tri-diagonal, and so $\boldsymbol{S} = -\boldsymbol{\Sigma}^{-1}$ has dense diagonal blocks. Consequently, $\boldsymbol{K}_{\mathrm{L}}$ is size $M \times M$, and $\boldsymbol{K}$ is $n \times n$.

As above, we avoid calculation of $\boldsymbol{T}$ and $\boldsymbol{S} = -\boldsymbol{\Sigma}^{-1}$ and proceed by direct calculation of the elements of $\boldsymbol{K}$. Because herd 'susceptibility' and 'transmissibility' are independent of who is infecting whom, each entry K_{ij} only requires computation of the expected persistence time and the expected prevalence when infected for each herd type j denoted by T^j_{inf} and P^j_{pos}, respectively. Then, each K_{ij} is the number of secondary infections in a herd type i corresponding to entry from a herd of type j. Because

$$t^j_{\mathrm{out}} := \kappa_j N_j P^j_{\mathrm{pos}} T^j_{\mathrm{inf}},$$

infectives leave herds of type j, and enter disease-free herds of type i with probability

$$s^i_{\mathrm{in}} := \frac{p_i \kappa_i N_i}{\langle \kappa N \rangle},$$

this gives

$$K_{ij} = s^i_{\mathrm{in}} t^j_{\mathrm{out}} = \frac{p_i \kappa_i N_i}{\langle \kappa N \rangle} \kappa_j N_j P^j_{\mathrm{pos}} T^j_{\mathrm{inf}},$$

and so

$$\boldsymbol{K} = \begin{pmatrix} s^1_{\mathrm{in}} t^1_{\mathrm{out}} & \cdots & s^1_{\mathrm{in}} t^n_{\mathrm{out}} \\ \vdots & \ddots & \vdots \\ s^n_{\mathrm{in}} t^1_{\mathrm{out}} & \cdots & s^n_{\mathrm{in}} t^n_{\mathrm{out}} \end{pmatrix} = \begin{pmatrix} s^1_{\mathrm{in}} & & \\ \vdots & t^1_{\mathrm{out}} & \cdots & t^n_{\mathrm{in}} \\ s^n_{\mathrm{in}} & & \end{pmatrix}.$$

Because $\boldsymbol{K}$ is the outer product of two vectors, and so all rows of $\boldsymbol{K}$ are linear multiples of each other, there is just one non-zero eigenvalue, given by the sum of the diagonal elements of $\boldsymbol{K}$, i.e.

$$\begin{aligned} R_* &= \mathrm{Trace}(\boldsymbol{K}) \\ &= \sum_{i=1}^{n} s^i_{\mathrm{in}} \cdot t^i_{\mathrm{out}} \\ &= \frac{1}{\langle \kappa N \rangle} \sum_{i=1}^{n} p_i \kappa_i N_i \cdot \kappa N_i P^i_{\mathrm{pos}} T^i_{\mathrm{inf}} \\ &= \frac{\langle \kappa N \cdot \kappa N P_{\mathrm{pos}} T_{\mathrm{inf}} \rangle}{\langle \kappa N \rangle}. \end{aligned} \tag{3.1}$$

This form has natural parallels with the expression for R_0 on a random network:

$$R_0 = \frac{\langle k_{\mathrm{in}} k_{\mathrm{out}} \rangle}{\langle k_{\mathrm{in}} \rangle},$$

where k_{in} and k_{out} refer to the number of infectious in and out links per node [40]. In our expression, the number of outward infectious links also captures the within-node disease dynamics via the terms P^i_{pos} and T^i_{inf} for the expected on farm prevalence while infected and the expected duration of infection.

Note that to maintain herd sizes, we assume that movement in κ_{in} equals movement out κ_{out}. However, in more complex scenarios, such as asymmetric cattle movement, this restriction may be relaxed, relying on within herd dynamics to maintain herd size. This would lead to more complex expressions for s^i_{in} and t^j_{out}, however this is beyond the scope of this paper.

3.3. Heterogeneity in herd size N and movement rate κ

Now using the NGM method described in §3.2, we examine how R_* in the SIS model depends on heterogeneity in herd size N and movement rate κ. Heterogeneity is created by keeping a fixed mean, but varying the variance-to-mean ratio of a gamma distribution (discretized in the case of herd size).

R_* is higher in populations with greater heterogeneity in herd size (figure 4*b*), but lower in populations with greater heterogeneity in movement rate (figure 4*c*), which can be explained heuristically as follows. Larger herds are associated with a lower chance of stochastic extinction [41] and therefore, a larger T_{inf}. Thus, larger herds will have a greater N and T_{inf} and therefore contribution disproportionately to R_*.

If each herd has its own *per capita* movement rate κ_i, then each herd will contribute differently to R_*. As there is a movement rate κ^* that maximizes R_*, the highest R_* should occur in the homogeneous case where $\kappa_i = \kappa^*$ for all herds. Any heterogeneity in κ_i should reduce R_*, as some herds will contribute less to R_*. Consider the extreme case, where the population is composed of two groups, a small number of herds with high

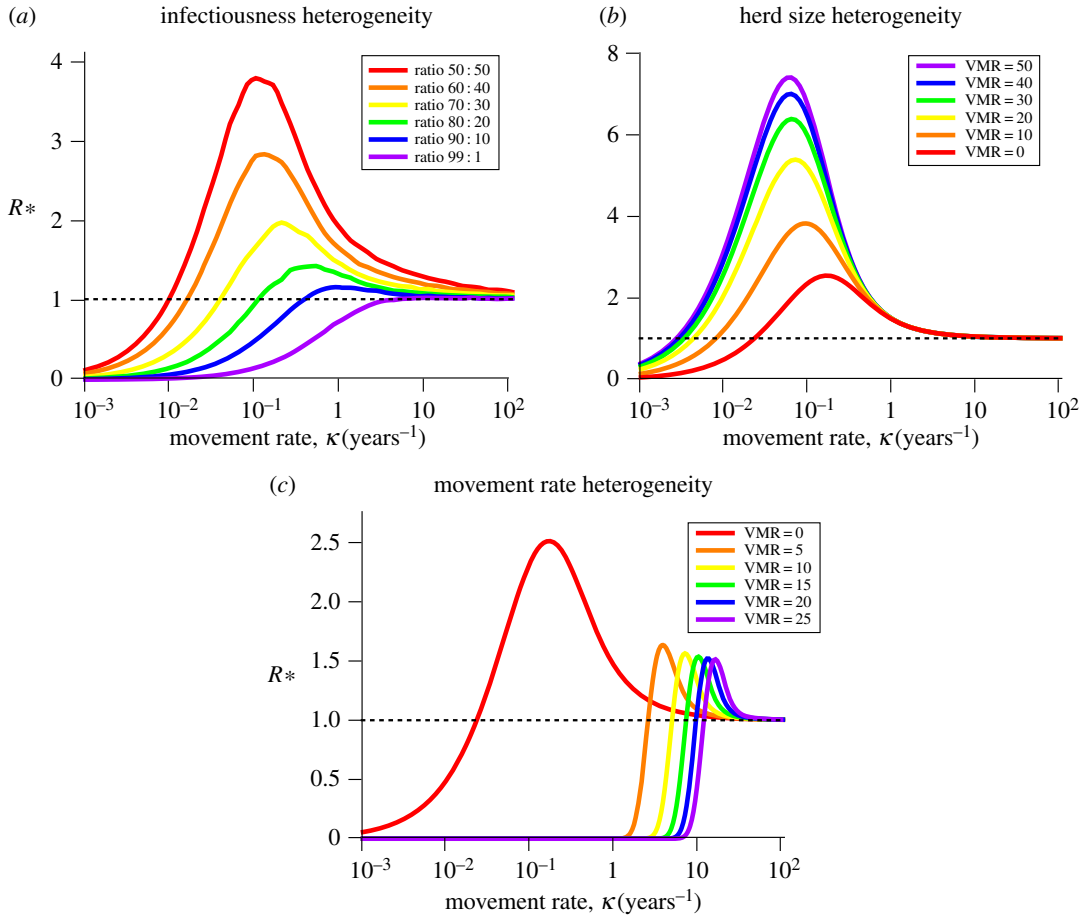

Figure 4. Impact of heterogeneity on R_*. R_* versus movement rate κ with (*a*) increasing heterogeneity between low and high shedders. Parameters are as the *E. coli* O157 model (see §1.2 of the electronic supplementary material); (*b*) increasing heterogeneity in herd size N in the SIS model (see §1.1 of the electronic supplementary material) and (*c*) increasing heterogeneity in movement rate κ in the same model. Parameters are $\mu = 1/3$, $\gamma = 1$, $\beta = 1.75$. The homogeneous case in each plot is shown in red, moving to purple with increasing heterogeneity. R_* is maximized by homogeneous infectiousness and movement, but maximized by heterogeneous herd size, as larger herds contribute disproportionately more to transmission.

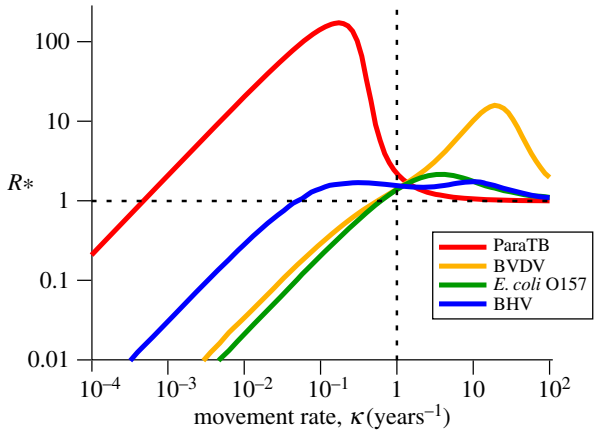

Figure 5. R_* versus movement rate κ for four different cattle diseases: BVDV, BHV, ParaTB, and *E. coli* O157 (see §1 of the electronic supplementary material for full details). Around typical cattle movement rates of $\kappa = 1$, all diseases here have $R_* > 1$, and hence are able to spread between herds, however R_* is maximized for higher κ in BVDV, and lower κ in ParaTB. Owing to long persistence times of infection, some simulations for ParaTB were truncated, and so the value of R_* presented is actually a lower bound on the true value.

κ_i, which contribute $R_* \approx 1$, and a large number of herds with low κ_i, which contribute $R_* \approx 0$. Therefore, R_* is maximized by homogeneous movement, and this result is indeed shown in figure 4*c*.

4. R_* in four important livestock disease systems

We consider four important and epidemiologically different cattle diseases: BVDV, BHV, *Mycobacterium avium* ssp *paratuberculosis* (ParaTB, the pathogen responsible for Johne's disease), and *Escherichia coli* O157 (*E. coli* O157). Models and parameters for the first three are based on non-spatial deterministic models described by Carslake [42], whereas those for *E. coli* O157 are based on [37,43], (see §1 of the electronic supplementary material).

We calculated R_* for each model by populating the NGM $\boldsymbol{K}$ directly, obtaining each entry K_{YX} by simulation, introducing a single individual of infectious type X to a susceptible herd, and counting the number of infectious type Y leaving the herd via movement until the infection died out in the primary herd. To find the associated quasi-equilibrium proportion of infected herds, we also simulated a metapopulation of $n = 100$ herds each with $N = 50$ individuals. Our assumption of a homogeneous metapopulation means that we assume undirected movement between herds, and that movements are equally likely between any herds.

We considered movement rates κ between 0.0001 and 100 per year (see §2 of the electronic supplementary material for details on the methods used). Cattle typically move around one to four times during their lifetime, which has a mean of around 3 years [44]. Consequently, the range of movements of most interest is around $\kappa = 1$ (one movement per

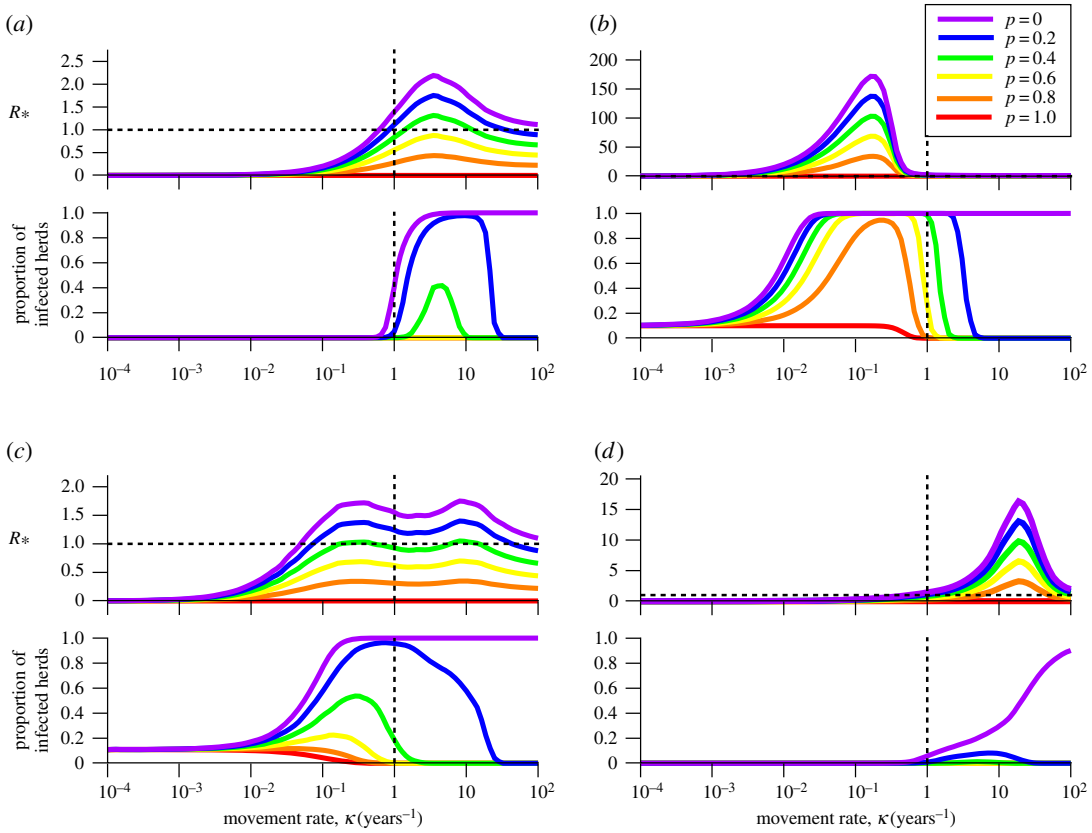

Figure 6. R_* and proportion of infected herds against movement rate κ for (*a*) *E. coli* O157, (*b*) ParaTB, (*c*), BHV and (*d*) BVDV. $n = 100$ herds were simulated (see §1 of the electronic supplementary material for full details). While not intended as an exact representation of reality, the vertical dashed line at $\kappa = 1$ represents the area roughly closest to real life movement rates. Higher κ would make *E. coli* O157 and BVDV more persistent, while lower κ would favour ParaTB. The highest R_* is seen in ParaTB (T_{inf} is extremely high for low κ, and the value give for R_* here is only a lower bound), and this corresponds to ParaTB being difficult to treat when κ is low. Note the double peak for R_* in BHV (*c*), and the green line dips below 1 around $\kappa = 1$; while the proportion of infected herds is calculated at $t = 20$, the disease may ultimately be unable to persist for longer time periods.

animal per year). However, these models are intended to expose the range of behaviours of R_*, rather than make precise predictions, and thus we consider a wider range of movement rates than is typically recorded.

For each of the exemplar diseases, R_* reaches an intermediate peak above 1 for some intermediate movement rate (figure 5). The slowly progressing ParaTB has a peak at low movement rates, whereas the rapidly progressing BVDV has a peak at high movement. BHV and *E. coli* O157 have intermediate transmission rates, and thus peak at intermediate movements; however, the two categories of infective for BHV lead to a double peak.

Comparing R_* with the proportion of infected herds (figure 6) shows that while all diseases achieve the maximum proportion of infected herds for high movement rates in the absence of disease intervention, even a relatively weak control effort ($p = 0.2$, indicating that 20% of infected individuals are identified and treated, blue lines) is sufficient to control the disease at high movement rates. With low movement rates, ParaTB is difficult to control (even a high control effort fails to control infection), but with greater movement effective control becomes easier.

5. Discussion

The work reported here is motivated by the desire to control disease in regional and national livestock populations and addresses the lack of suitable metrics for determining the level of effort required when movement-based disease control is used to reduce disease transmission that is primarily driven by movement (trading) of livestock.

We describe a novel formulation of the threshold for disease spread in a structured population, R_*, that explicitly captures group to group transmission via animal movements. While a number of previous studies have addressed the impact of group structure on disease invasion, some analytically [29,30,45] and some via statistical and simulation methods [19,24], this is the first demonstration of a threshold parameter for disease invasion in a metapopulation that captures within-group stochastic dynamics coupled with the explicit movement of infected individuals between groups.

Following Diekmann and Heesterbeek, we use an NGM approach to calculate R_* and show how this may be used for disease systems with heterogeneities and multiple infectious states. We show for a simple disease system that R_* is given by the intuitive expression $R_* = \kappa N P_{\text{pos}} T_{\text{inf}}$, where κ is the movement rate, N is the herd size, T_{inf} is the expected persistence time and P_{pos} is the expected prevalence in an infected herd. Note that this factorization of R_* is non-trivial and accounts for the fact that prevalence and persistence time may be correlated. Pellis *et al.* [46] make a similar observation about their factorization of R_* for household models.

A key feature is the presence of a peak in R_* at intermediate movement rates. This novel observation arises, because we have explicitly modelled the herds' gain and loss in infectives that occurs when disease is spread by livestock movements. In household models where the contact process

resulting in disease transmission is captured phenomenologically, there would be no such peak [30,32].

The R_* peak depends on the interaction between movement rate κ, the within-herd disease persistence time T_{inf}, and the expected prevalence P_{pos} in an infected herd. Movement contributes directly to R_*, but crucially also removes infectives from the herd, and therefore can reduce T_{inf} and P_{pos}. It is this trade-off that leads to the characteristic intermediate peak. Theoretically, for very high movement rates, an infective animal would arrive on farm and then immediately leave with virtually no opportunity to make infectious contacts, recover or die; for this reason R_* tends to 1 at high movement rates.

The peak in R_* has important consequences for control directed at livestock moving between herds. The degree of control effort required also peaks at intermediate movement rates, and consequently a given level of control may be sufficient to prevent persistence at low or high movement rates, but be insufficient over a range of intermediate movement rates. This phenomenon arises, because increased movement exposes more animals to testing, with the consequence that controls need to be less effective at identifying infected animals at high movement rates to achieve a given reduction in prevalence.

R_* increases dramatically with increased herd sizes that substantially increase the persistence of infection. In addition, rather modest values of R_0 can, depending on the disease system, be associated with values of R_* that are orders of magnitude larger. This finding indicates that for some disease systems control directed at reducing R_0 may be more effective than controls directed at animals moving between holdings.

We also demonstrated that R_* is maximized when there is the least heterogeneity between farms in movement rates and when there is the least individual variation in infectiousness; conversely, increasing heterogeneity in herd size increases R_*.

Our exemplar disease models and their parametrizations were selected, not to give precise predictions, but to provide a range of R_* behaviours across four important livestock diseases. The different disease dynamics result in quite different R_* profiles, leading to potential trade-offs between the control of different diseases. All the exemplar diseases have low R_* near the intermediate *per capita* movement rate of one movement per year, but our predictions indicate that ParaTB and BVDV would have much higher R_* at lower movements for ParaTB and at higher movements for BVDV. ParaTB (Johne's disease), a slowly progressing disease which persists in a herd for a long time, has an R_* that peaks at low movement rates indicating that it might prove difficult to control if movement rates were reduced; however, increasing movement rates slightly could expose it to sufficient intervention that it would be unable to spread between herds.

In contrast, *E. coli* O157, a rapidly progressing disease with an R_* peak at higher movement rates may be better able to persist in the face of movement-based controls at higher movement rates. BHV, which can also persist in herds for long periods is able to invade at lower movement rates than would be needed for invasion by *E. coli* O157 or BVDV. These findings concur with the observations that chronic diseases are more likely to invade than acute diseases with the same R_0 [29].

The consequence of the differing R_* profiles is that if, for example, movement restrictions were put in place to reduce *E. coli* O157, ParaTB could become more difficult to control via movement-based controls. On the other hand, if movement rates increased, ParaTB could be easier to control via movement-based controls, at the cost of increased prevalence of *E. coli* O157. Overall, our results indicate that at current livestock movement rates, disease control implemented at the point of between-farm movement alone can be sufficient to control some pathogens, but for infections such as ParaTB control at herd level is likely to be needed in addition.

Inevitably, the models analysed in this paper include a number of simplifying assumptions; nevertheless, our methodology (a key result of this paper) is applicable to more realistic scenarios. The extensive explorations presented in this paper indicate that the following results will hold in more complex scenarios: R_* will peak and decline, leading to 'islands' of persistence when control is implemented. In addition, we anticipate that different diseases will have different R_* profiles, potentially leading to conflicting requirements when controlling multiple diseases.

In summary, our formulation of R_* provides novel theoretical insights into the likely effectiveness of alternative control strategies and an important addition to the selection of tools available to epidemiologists to be used in conjunction with R_0 for disease control in livestock systems.

Competing interests. We declare we have no competing interests.

Funding. We thank the Scottish Government for funding through the Strategic Partnership in Animal Science Excellence (SPASE). J.C.P. is supported by FSS project FS101055. L.M. is supported by National Science Foundation DEB1216040, BBSRC grants BB/K01126X/1, BB/L004070/1, BB/L018926/1, BB/L018926/1, BB/N013336/1. L.M. and G.M. are supported by the Scottish Government Rural and Environment Science and Analytical Services Division, as part of the Centre of Expertise on Animal Disease Outbreaks (EPIC). G.M., M.R.H. and T.N.M. are supported by the Strategic Research programme of the Scottish Government's Rural and Environment Science and Analytical Services Division (RESAS).

Acknowledgements. The authors also thank three anonymous referees and Rowland Kao for their helpful comments.

References

1. Houe H. 2003 Economic impact of BVDV infection in dairies. *Biologicals* **31**, 137–143. (doi:10.1016/S1045-1056(03)00030-7)
2. Hasonova L, Pavlik I. 2006 Economic impact of paratuberculosis in dairy cattle herds: a review. *Vet. Med. – Czech*. **51**, 193–211.
3. Daszak P, Cunningham AA, Hyatt AD. 2000 Emerging infectious diseases of wildlife — threats to biodiversity and human health. *Science* **287**, 443–449. (doi:10.1126/science.287.5452.443)
4. Cleaveland S, Laurenson MK, Taylor LH. 2001 Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergency. *Phil. Trans. R. Soc. Lond. B* **356**, 991–999. (doi:10.1098/rstb.2001.0889)
5. Févre EM, Bronsvoort BMdC, Hamilton KA, Cleaveland S. 2006 Animal movements and the spread of infectious diseases. *Trends Microbiol*. **14**, 3. (doi:10.1016/j.tim.2006.01.004)
6. Gibbens JC, Sharpe CE, Wilesmith JW, Mansley LM, Michalopoulou E, Ryan JB, Hudson M. 2001 Descriptive epidemiology of the 2001 foot-and-mouth disease epidemic in Great Britain: the first five months. *Vet. Rec*. **149**, 729–743. (doi:10.1136/vr.149.24.729)

7. Kao RR. 2002 The role of mathematical modelling in the 2001 foot-and-mouth disease epidemic in the United Kingdom. *Trends Microbiol*. **10**, 279–286. (doi:10.1016/S0966-842X(02)02371-5)
8. England T, Kelly L, Jones RD, MacMillan A, Wooldridge M. 2004 A simulation model of brucellosis spread in British cattle under several testing regimes. *Prev. Vet. Med*. **63**, 63–73. (doi:10.1016/j.prevetmed.2004.01.009)
9. Christley RM, Robinson SE, Lysons R, French NP. 2005 Network analysis of cattle movement in Great Britain. In *Proc. Soc. Vet. Epidemiol. Prev. Med*. (eds DM Mellor, AM Russell, JLN Wood), pp. 234–243
10. Truyers IGR, Mellor DJ, Norquay R, Gunn GJ, Ellis KA. 2010 Eradication programme for bovine viral diarrhoea virus in Orkney 2001 to 2008. *Vet. Rec*. **167**, 566–570. (doi:10.1136/vr.c4944)
11. Beaunée G, Vergu E, Ezanno P. 2015 Modelling of paratuberculosis spread between dairy cattle farms at a regional scale. *Vet. Res*. **46**, 1. (doi:10.1186/s13567-014-0124-5)
12. Ayele WY, Macháčková M, Pavlik I. 2001 The transmission and impact of paratuberculosis infection in domestic and wild ruminants. *Vet. Med. Czech*. **46**, 205–224.
13. Gilbert M, Mitchell A, Bourn D, Mawdsley J, Clifton-Hadley R, Wint W. 2005 Cattle movements and bovine tuberculosis in Great Britain. *Nature* **435**, 491–496. (doi:10.1038/nature03548)
14. DEFRA. 2012 Controlling disease in farm animals—detailed guidance; 2012. https://www.gov.uk/guidance/controlling-disease-in-farm-animals. (accessed 30 Sep 2015).
15. Anderson RM, May RM. 1992 *Infectious diseases of humans*. Oxford, UK: Oxford University Press.
16. Heesterbeek JAP, Dietz K. 1996 The concept of R_0 in epidemic theory. *Stat. Neerl*. **50**, 89–110. (doi:10.1111/j.1467-9574.1996.tb01482.x)
17. van den Driessche P, Watmough J. 2002 Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci*. **180**, 29–48. (doi:10.1016/S0025-5564(02)00108-6)
18. Keeling MJ, Rohani P. 2007 *Modelling infectious diseases in humans and animals*. Princeton, NJ: Princeton University Press.
19. Cross PC, Johnson PLF, Lloyd-Smith JO, Getz WM. 2007 Utility of R_0 as a predictor of disease invasion in structured populations. *J. R. Soc Interface* **4**, 315–324. (doi:10.1098/rsif.2006.0185)
20. Hanski I, Gilpin ME. 1997 *Metapopulation biology*. San Diego, CA: Academic Press.
21. Hess G. 1996 Disease in metapopulation models: implications for conservation. *Ecology* **77**, 1617–1632. (doi:10.2307/2265556)
22. Gog J, Woodroffe R, Swinton J. 2002 Disease in endangered metapopulations: the importance of alternative hosts. *Proc. R. Soc. Lond. B* **269**, 671–676. (doi:10.1098/rspb.2001.1667)
23. Jesse M, Ezanno P, Davis S, Heesterbeek JAP. 2008 A fully coupled, mechanistic model for infectious disease dynamics in a metapopulation: movement and epidemic duration. *J. Theor. Biol*. **254**, 331–338. (doi:10.1016/j.jtbi.2008.05.038)
24. Nickbakhsh S, Matthews L, Dent JE, Innocent GT, Arnold ME, Reid SWJ, Kao RR. 2013 Implications of within-farm transmission for network dynamics: consequences for the spread of avian influenza. *Epidemics* **5**, 67–76. (doi:10.1016/j.epidem.2013.03.001)
25. Keeling MJ, Rohani P. 2002 Estimating spatial coupling in epidemiological systems: a mechanistic approach. *Ecol. Lett*. **5**, 20–29. (doi:10.1046/j.1461-0248.2002.00268.x)
26. Park AW, Gubbins S, Gilligan CA. 2002 Extinction times for closed epidemics: the effects of host spatial structure. *Ecol. Lett*. **5**, 747–755. (doi:10.1046/j.1461-0248.2002.00378.x)
27. Hagenaars TJ, Donnelly CA, Ferguson NM. 2004 Spatial heterogeneity and the persistence of infectious diseases. *J. Theor. Biol*. **229**, 349–359. (doi:10.1016/j.jtbi.2004.04.002)
28. Fulford GR, Roberts MG, Heesterbeek JAP. 2002 The metapopulation dynamics of an infectious disease: tuberculosis in possums. *Theor. Popul. Biol*. **61**, 15–29. (doi:10.1006/tpbi.2001.1553)
29. Cross PC, Lloyd-Smith JO, Johnson PLF, Getz WM. 2005 Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations. *Ecol. Lett*. **8**, 587–595. (doi:10.1111/j.1461-0248.2005.00760.x)
30. Ball F. 1999 Stochastic and deterministic models for SIS epidemics among a population partitioned into households. *Math. Biosci*. **156**, 41–67. (doi:10.1016/S0025-5564(98)10060-3)
31. Ball F, Neal P. 2002 A general model for stochastic SIR epidemics with two levels of mixing. *Math. Biosci*. **180**, 73–102. (doi:10.1016/S0025-5564(02)00125-6)
32. Ball F, Neal P. 2008 Network epidemic models with two levels of mixing. *Math. Biosci*. **212**, 69–87. (doi:10.1016/j.mbs.2008.01.001)
33. Diekmann O, Heesterbeek JAP. 2000 *Mathematical epidemiology of infectious diseases: model building, analysis, and interpretation*. Chichester, UK: John Wiley & Sons.
34. Diekmann O, Heesterbeek JAP, Roberts MG. 2010 The construction of next-generation matrices for compartmental epidemic models. *J. R. Soc. Interface* **7**, 873–885. (doi:10.1098/rsif.2009.0386)
35. Diekmann O, Heesterbeek JAP, Mets JAJ. 1990 On the definition and the computation of the basic reproduction ratio R_0. *J. Math. Biol*. **28**, 365–382. (doi:10.1007/BF00178324)
36. Keeling MJ, Ross JV. 2008 On methods for studying stochastic disease dynamics. *J. R. Soc. Interface* **5**, 171–181. (doi:10.1098/rsif.2007.1106)
37. Matthews L *et al.* 2016 Heterogeneous shedding of *Escherichia coli* O157 in cattle and its implications for control. *Proc. Natl Acad. Sci. USA* **103**, 547–552. (doi:10.1073/pnas.0503776103)
38. Matthews L, McKendrick IJ, Ternent HE, Gunn GJ, Synge BA, Woolhouse MEJ. 2006 Super-shedding cattle and the transmission dynamics of *Escherichia coli* O157. *Epidemiol. Infect*. **134**, 131–142. (doi:10.1017/S0950268805004590)
39. Matthews L *et al.* 2013 Predicting the public health benefit of vaccinating cattle against *Escherichia coli* O157. *Proc. Natl Acad. Sci. USA* **110**, 16 265–16 270. (doi:10.1073/pnas.1304978110)
40. Kao RR, Danon L, Green DM, Kiss IZ. 2006 Demographic structure and pathogen dynamics on the network of livestock movements in Great Britain. *Proc. R. Soc. B* **273**, 1999–2007. (doi:10.1098/rspb.2006.3505)
41. Nåsell I. 1996 On the time to extinction in recurrent epidemics. *J. R. Stat. Soc. B, Met*. **61**, 309–330. (doi:10.1111/1467-9868.00178)
42. Carslake D, Grant W, Green LE, Cave J, Greaves J, Keeling M, McEldowney J, Weldegebriel H, Medley GF. 2011 Endemic cattle diseases: comparative epidemiology and governance. *Phil. Trans. R. Soc. B* **366**, 1975–1986. (doi:10.1098/rstb.2010.0396)
43. Zhang XS, Woolhouse MEJ. 2011 *Escherichia coli* O157 infection on Scottish cattle farms: dynamics and control. *J. R. Soc. Interface* **8**, 1051–1058. (doi:10.1098/rsif.2010.0470)
44. Vernon MC. 2011 Demographics of cattle movements in the United Kingdom. *BMC Vet. Res*. **7**, 1–16. (doi:10.1186/1746-6148-7-31)
45. Becker NG, Dietz K. 1995 The effect of household distribution on transmission and control of highly infectious diseases. *Math. Biosci*. **127**, 207–219. (doi:10.1016/0025-5564(94)00055-5)
46. Pellis L, Ferguson NM, Fraser C. 2009 Threshold parameters for a model of epidemic spread among households and workplaces. *J. R. Soc. Interface* **6**, 979–987. (doi:10.1098/rsif.2008.0493)