Disease Invasion and Control in Structured Populations: Bovine Tuberculosis in the Buffalo Population of the Kruger National Park

by

Paul Chafee Cross
B.A. (University of Virginia) 1998

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy
in
Environmental Science, Policy and Management
in the
GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Wayne M. Getz, Chair
Professor Steven R. Beissinger
Professor Cheryl J. Briggs
Professor Johan T. du Toit

May 2005

Disease Invasion and Control in Structured Populations: Bovine Tuberculosis in the Buffalo Population of the Kruger National Park

by

Paul Chafee Cross

Abstract

Disease Invasion and Control in Structured Populations: Bovine Tuberculosis in the Buffalo Population of the Kruger National Park

by

Paul Chafee Cross

Doctor of Philosophy in

Environmental Science, Policy and Management

University of California, Berkeley

Professor Wayne M. Getz, Chair

From 1991 to 2004, bovine tuberculosis (*Mycobacterium bovis*, BTB) moved north and increased in prevalence in the African buffalo (*Syncerus caffer*) population of the Kruger National Park, South Africa. I use this epidemic as a case study to understand how host population structure affects disease dynamics. Radio-tracking data indicated that all sex and age groups moved between herds, but males over eight years old had higher mortality and dispersal rates than any other sex or age category. BTB appeared to have only minor effects upon survival. Models incorporating these data suggest that the success of vaccination programs will depend strongly upon the duration that a vaccine grants protection, which is currently unknown. Even with a lifelong vaccine,

however, eradication is unlikely unless vaccination is combined with other control strategies.

To analyze the radio-tracking and association data, I proposed a new metric of association, the fission decision index (FDI) that significantly reduces the biases that exist in traditional association analyses in fission-fusion societies. Adult female and juvenile buffalo made non-random fission decisions while adult male choices were indistinguishable from a random coin toss. Incorporating the association data into a dynamic social network model suggested that the dynamic nature of the network has a strong influence on disease dynamics, particularly for diseases with shorter infectious periods. Buffalo herds were more tightly associated in 2002 than 2003, perhaps due to drier conditions in 2003 prompting additional movement that would facilitate the spread of disease among herds.

Using a metapopulation model, I illustrate how the group-level metric, R_*, which is the average number of groups infected by the initial group, is a better predictor of disease invasion than the traditional individual-level R_0 in structured host populations. R_* is a function of group size, movement rate, infection rate, and length of the infectious period. Chronic diseases allow for more host mixing between groups; thus they 'perceive' a more well-mixed host population. As a result, chronic diseases are more likely to invade structured populations than acute diseases, given the same R_0, and it is more important to incorporate the spatial structure of the host population for acute diseases than chronic diseases.

Table of Contents

Acknowledgements

First, I would like to thank my family for all of their support. Craig Hay, Justin Bowers, Julie Wolhuter, Khutani Bulunga, and Augusta Mabunda collected the vast majority of the field data included in this dissertation. Always professional and great companions, they made the last four years in Satara an amazing experience. I am grateful to the managers, scientists, and staff of the Kruger National Park for facilitating the project, and to the United States National Science Foundation Ecology of Infectious Disease Grant DEB-0090323 for funding this research. Drs. Markus Hofmeyr, Peter Buss, Lin-Mari de Klerk, Roy Bengis and Douw Grobler, as well as Marius Kruger, KNP Game Capture, State Veterinary technicians, and many University of Pretoria students assisted in buffalo capture operations. Roy Bengis, Steve Beissinger, Cherie Briggs, Justin Brashares, Charlie Nunn, and Anna Jolles all provided many helpful comments along the way. Martin Haupt and Elmarie Cronje at the University of Pretoria were invaluable in their assistance with equipment and finances. All the students in the Getz Lab including Peter Baxter, Chris Wilmers, Sadie Ryan, Andy Lyons, Allison Bidlack, Wendy Turner, Shirli Bar-David, John Eppley, George Wittemyer, James Lloyd-Smith, and Karen Levy, made it a fantastic environment to work in. In particular, Jamie Lloyd-Smith and George Wittemyer were a great support and source of ideas and inspiration throughout this dissertation. Lastly, I would like to thank Molly Smith, Johan du Toit, and Wayne Getz. To list all the ways that they helped would be its own dissertation.

Chapter One

Summary

P.C. Cross

Bovine tuberculosis (BTB), caused by *Mycobacterium bovis*, is an airborne bacterial pathogen that is re-emerging in wildlife and livestock worldwide. In the Kruger National Park (KNP) of South Africa, BTB is increasing in prevalence and moving northwards from its introduction from cattle along the southern border of the KNP in the early 1960s (Bengis et al. 1996, Bengis 1999, Rodwell et al. 2001). African buffalo (*Syncerus caffer caffer*) are a reservoir host, maintaining the disease at high prevalence (over 50% in some herds), while predators such as lions and leopards appear to be spill-over hosts (Keet et al. 1996, Rodwell et al. 2000). It remains unclear how the effects of BTB, with its wide range of potential hosts, will ripple through the KNP ecosystem (Caron et al. 2003). Furthermore, Kruger National Park is the largest reserve in South Africa, covering over 20,000 km^2 and is the source of many animal translocations. BTB may limit the ability of KNP managers to translocate animals from the KNP, thereby decreasing a potential source of revenue for South African National Parks and creating a conservation island.

This dissertation is one component of a larger research program on BTB in the buffalo population of the KNP, the goal of which is to develop a better understanding of disease dynamics in wildlife systems as well as to evaluate potential management strategies. Here I investigate the role of host social and spatial structure on the spread and control of disease by integrating empirical data with a number of different epidemiological models and using BTB in buffalo as a case study. Early models of disease dynamics assumed a homogenous host population that was well-mixed (Kermack and McKendrick 1927, Anderson and May 1979, May and Anderson 1979, Anderson and May 1991). Recent studies have begun to investigate the role of spatial

and social structure to disease invasion or persistence (e.g. Hess 1996, Swinton et al. 1998, Keeling 1999, Keeling and Gilligan 2000, Keeling and Grenfell 2000, Park et al. 2001, Eames and Keeling 2002, Fulford et al. 2002, Keeling and Rohani 2002, Newman 2002, Park et al. 2002, Read and Keeling 2003, Eames and Keeling 2004, Eubank et al. 2004, Hagenaars et al. 2004). In this dissertation, I investigate host structure in four different contexts. First, I investigate the effect of age and sex structure within a single herd on the efficacy of a proposed BTB vaccination program. Second, I illustrate how traditional association analyses may result in biased conclusions about host social structure in fission-fusion societies and propose a new metric of individual association that may be more appropriate in fission-fusion societies such as buffalo. Third, I illustrate how an empirically-derived social network affects disease dynamics for diseases of different infectious periods. Finally, I use a metapopulation model with explicit movement to illustrate how traditional metrics of disease invasion break down in structured populations.

As an exotic disease, KNP managers would like to control or eradicate BTB via culling, vaccination, or some combination of the two. Previous modeling work on BTB in buffalo, however, suggests that BTB may persist even when the buffalo population is reduced to low densities, making random culling, for the purpose of eradication, problematic (Rodwell 2000). Thus, vaccination, or some combination of vaccination and culling, is a more attractive management option. Early action is likely to be the most effective, however there are many uncertainties surrounding the potential impacts (or lack thereof) of BTB on the buffalo population and spill-over hosts, the efficacy and

duration of the vaccine in buffalo, and the potential effectiveness of a vaccination or test-and-remove management strategy.

In the second chapter, we use a sex and age structured epidemiological model to assess the potential efficacy a vaccination program on controlling or eradicating BTB in a single buffalo herd. The model incorporates dispersal of individuals between the focal herd and a constant background population as well as the non-random mixing of sex and age categories within a herd. The model is parameterized with survival and dispersal estimates from over 130 radio-collared buffalo that were intensively monitored over a three-year period. We use the model to assess the importance of between-herd mixing, the potential effectiveness of a vaccination strategy, and whether a vaccination strategy can be improved by focusing on particular sex and age groups.

Radio-tracking data indicated that all sex and age categories move between mixed herds, and males over eight years old had higher mortality and dispersal rates than any other sex or age category. In part due to the high dispersal rates of buffalo, sensitivity analyses indicate that disease prevalence in the background population accounts for the most variability in the BTB prevalence and quasi-eradication within the focal herd. Vaccination rate and the transmission coefficient were the second and third most important parameters of the sensitivity analyses. Further analyses of the model without dispersal suggest that the amount of vaccination necessary for quasi-eradication (*i.e.* prevalence < 5%) depends upon the duration that a vaccine grants protection. Vaccination programs are more efficient (*i.e.* fewer wasted doses) when they focus on younger individuals. However, even with a lifelong vaccine and a closed population, the model suggests that >70% of the calf population would have to be vaccinated every

year to reduce the prevalence to less than 1%. If the half-life of the vaccine is less than five years, even vaccinating every calf for 50 years may not eradicate BTB. Thus, although vaccination provides a means of controlling BTB prevalence it should be combined with other control measures if eradication is the objective.

In the third chapter, we investigate the social structure of buffalo in greater depth and develop new techniques for analyzing association patterns in fission-fusion societies. Previous studies of association patterns usually calculate an association index of the proportion of time or sightings that a pair of individuals is seen together. Using a simulation model we show how this method can yield biased results in fission-fusion societies. In particular, if sampling occurs more often than fission and fusion events, the proportion of time dyads spend together will show statistically significant clustering, even if individuals choose herds at random and independently of other individuals' decisions. This follows because multiple samples taken within the same inter-fission interval are autocorrelated with respect to individual choices. Thus, we propose a new metric, the fission decision index (FDI), which is the number of times a pair of individuals chose to remain together when a group separates. The FDI eliminates auto-correlated data and presents an unbiased estimate of individual choices.

Traditional association analyses suggested that the buffalo population we studied was spatially and temporally structured into four different groups that were statistically different from random. The FDI approach, however, illustrated that the probability of a dyad remaining in the same group during a fission event was not closely correlated with the amount of time they spent together. Therefore, the non-random group structure apparent in the traditional association analysis was due in small part to

non-random decisions made by individuals during all fission events, but in greater part to the variable lifetimes of the resulting fission groups. Consideration of FDI scores thus helped us to understand the mechanisms underlying results of traditional association analysis.

In the fourth chapter, we integrate the radio-tracking association data with a dynamic social network model of disease dynamics. Association data are often collected but, to the author's knowledge, this is the first use of empirical data in a network disease model in a wildlife population. Although there are still aspects of this approach that need to be further developed before it can be applied more broadly, we believe it provides a flexible method of capturing the social structure of wildlife populations that are not well-described by more traditional methods such as meta-population or cellular-automata models. Using this method, we investigate the importance of network topology and the amount of switching between groups to disease spread. We also analyze the effect of increasing the variance in association index. In other words, for disease spread does it matter if the individuals associate equally with their neighbors or if individuals have a set of close friends and loose associates?

The dynamic nature of the network had a strong influence on simulated disease dynamics. Further, acute diseases with shorter infectious periods were less likely to invade the social network than chronic diseases with longer infectious periods, even after scaling the transmission rate appropriately. This result is due to the fact that chronic diseases allow for more host mixing to occur, thus creating a more well-connected network. Cluster analyses of the association data demonstrated that buffalo herds were not as well defined as previously thought. Associations were more tightly

clustered in 2002 than 2003, perhaps due to drier conditions in 2003. As a result, diseases may spread faster during drought conditions due to increased population mixing.

In the fifth and final chapter, we generalize some of the patterns discovered during the analysis of the empirical social networks in chapter four using a stochastic, individually-based metapopulation model. This model was not intended to represent any particular system, but rather to illustrate how the interactions of several parameters dictate the probability of disease invasion in a structured host population. In particular, we show how the spread of disease in systems with small group sizes is largely a function of the ratio of dispersal and disease recovery rates. As a result, for a given movement and transmission rate, chronic diseases are more likely to invade structured populations than acute diseases. Further, acute diseases perceive a more structured host population, and it is more important to incorporate host social structure in models of these diseases. In systems with small groups, R_*, which is related to the ratio of host movement and disease recovery as well as the average group size, is a better indicator of disease invasion than the traditional R_0. Finally, although the behavioral susceptibility of the host is a function of group size and movement rate, the probability of a pandemic is not a function of either the host or the disease independently, but an emergent property of their interaction.

Chapter Two

Assessing vaccination as a control strategy in an ongoing epidemic: Bovine tuberculosis in African Buffalo

P.C. Cross and W.M. Getz

Introduction

Bovine tuberculosis (BTB, *Mycobacterium bovis*), an airborne bacterial pathogen, is re-emerging in wildlife and livestock worldwide. In the Kruger National Park of South Africa, BTB is increasing in prevalence and moving northwards from its introduction from cattle along the southern border of the KNP in the early 1960s (Bengis et al. 1996, Bengis 1999, Rodwell et al. 2001). Buffalo are a reservoir host, maintaining the disease at high prevalence (over 60% in some herds), while predators such as lions and leopards appear to be spill-over hosts (Keet et al. 1996, Rodwell et al. 2000). BTB is a chronic and progressive bacterial disease with a wide host range, and there is no evidence that animals recover from infection (Bengis 1999). As in cattle, most buffalo are infected with BTB via aerosol transmission (Bengis 1999). Vertical (intrauterine) and psuedovertical transmission (through infected milk) may occur but appear to be rare events (Bengis 1999). The pathology of BTB in lions suggests that they acquire the disease by consuming infected buffalo (Keet et al. 1996), and it remains unclear how the effects of BTB, with its wide range of potential hosts, will ripple through the KNP ecosystem.

Due to the potential effects of BTB on buffalo and alternative host species, KNP managers would like to control or eradicate BTB via culling, vaccination, or some combination of the two. Previous modeling work on BTB in buffalo suggests that BTB may persist even when the buffalo population is reduced to low densities, making random culling, for the purpose of eradication, problematic (Rodwell 2000). Thus, vaccination, or some combination of vaccination and culling, is a more attractive management option. Early action is likely to be the most effective, however there are

many uncertainties surrounding the potential impacts (or lack thereof) of BTB on the buffalo population, the efficacy and duration of the vaccine in buffalo, and the logistic difficulties of a vaccination program. Here we use an epidemiological model to tackle some of these questions surrounding management of the disease using the best available data. Analysis of the model provides the rapid answers required by managers and helps to focus further research projects.

Vaccination trials with Bacille Calmette-Guérin (BCG) and buffalo are currently underway in the KNP as the vaccine has been shown to give protection to cattle, deer, brushtail possums , ferrets, and badgers (for a review see Suazo et al. 2003). In cattle, BCG has been used in a number of trials and the amount of protection has varied widely (e.g. Francis 1958, Berggren 1977, Rodrigues et al. 1993, Colditz et al. 1994, Buddle et al. 1995c). More recent BCG vaccination trials in New Zealand have demonstrated up to 70% protection in cattle (Buddle *et al.* 1995c), data on the longevity of protection of the BCG vaccine, however, are non-existent in buffalo and limited in cattle (Berggren 1977). Here we use an SEI (Susceptible-Exposed-Infectious) epidemic model to assess the potential effectiveness of a buffalo vaccination program.

Several models of BTB in wildlife and cattle have been published previously (for a review see Smith 2001), but this is the first study to assess the efficacy of a vaccination program fore African buffalo. The effectiveness of control programs in African buffalo is likely to differ from other wildlife and cattle for several reasons. First, African buffalo tend to live longer than other wildlife species that are infected with BTB (e.g. badgers, possums, and deer). Second, BTB appears to have only minor effects upon the survival of buffalo (Rodwell et al. 2001). As a result, BTB can reach

high prevalence (60-92%) in buffalo herds (Bengis 1999, Rodwell et al. 2000, Jolles 2004). In contrast, a model of BTB transmission within cattle herds suggested that within-herd transmission alone was insufficient to maintain the infection (Barlow et al. 1997). In badgers and possums the prevalence of BTB has tended to be less than 20% (Smith 2001). Due to these differences with previously analyzed systems and the current effort to evaluate and implement a vaccination program to control BTB in African buffalo, we combine field data and mathematical modeling techniques to assess the likely success of a vaccination program.

Computer models exist along a continuum from detailed data-based models used for predictions about specific systems to models intended to improve understanding of general processes. We attempt to balance the specific and the general by using the limited data available to draw conclusions about the likely efficacy of a vaccination program, and the degree to which our conclusions depend on specific model parameters. Since, there has been comparatively little research conducted on BTB in African buffalo we keep our model relatively simple, including a minimal amount of structure. In contrast, research on BTB in badgers and possums has been ongoing for many years, and, as a result, the corresponding modeling analyses can be much more specific (e.g. Anderson and Trewhella 1985, Barlow 1991, Smith 2001). The data to construct and support this model of BTB in buffalo come from cross-sectional surveys in the KNP (Rodwell 2000) and Hluhluwe-Umfolozi Park (Jolles 2004), our ongoing longitudinal study in the KNP (Cross et al. 2004, Cross et al. 2005), and what is understood about the biology of BTB in buffalo and cattle. Research and management strategies, however, may need to be implemented prior to the collection of additional data and a

general model of an SEI disease in a long-lived social host may help to frame the problem and guide management decisions prior to their implementation.

We present preliminary analyses of field data to estimate dispersal and survival rates using data on over 130 radio-collared buffalo from November 2000 to December 2003. These estimates are then supplemented with published information on cattle and buffalo. The model is then used to assess the potential efficacy of a vaccination program and highlight the importance of several model parameters. For the model, we assume that the vaccine protects individuals in any sex and age category completely, but that protection may wane over time. Even though an effective vaccine with these characteristics is not yet available, our model can be used to assess what elements are necessary for a successful vaccination program.

The technical part of this paper begins with a presentation of a discrete time SEI model. This model builds upon previous models of bovine tuberculosis in cattle and wildlife (e.g. Bentil and Murray 1993, Ruxton 1996, Barlow et al. 1997, Kao and Roberts 1999, Smith et al. 2001) by incorporating dispersal, vaccination, sex, and age structure. Next, we outline the parameter estimation procedures and conduct sensitivity analyses of the model to determine which parameters explain the most variability in disease prevalence and eradication. We then further investigate the model in a single herd context (*i.e.* without dispersal to a background population). Finally, we assess the potential effectiveness of a vaccination program and whether it can be improved by focusing on different sex and age categories.

Methods

Simulation Model

The model presented here is an age- and sex-class elaboration of a discrete time SEI epidemic model for a focal buffalo herd in contact with a background population. We use a one-month time step to account for the annual reproduction and vaccination of buffalo and the within-year transmission dynamics of BTB. Let X, Y, Z and V represent the number of individuals respectively susceptible to infection, exposed but not yet infectious, infectious, and vaccinated, and let $N=X+Y+Z+V$ represent the total number of individuals. We assume that exposed individuals (Y) become infectious at a constant "incubation rate" γ. A proportion $\psi(t)$ of the population is vaccinated every June, and a constant proportion δ of the vaccinations fail in each month.

Currently available vaccines require the use of a helicopter to dart or drive the buffalo into pens prior to vaccination. With either method, an individual's disease status is unknown prior to capture. We account for this logistic difficulty by allowing for repeat vaccinations as well as the vaccination of already infectious individuals. Thus one measure of the efficiency of a vaccination program is the percentage of individuals captured that are successfully vaccinated for the first time (*i.e.* the percentage of individuals not infected nor previously vaccinated). We compare management strategies based on the ratio of successful to total vaccinations, assuming that vaccination protects only susceptible individuals.

We assume that a fixed proportion ε of individuals emigrate from the focal herd to a pool of individuals outside the herd each month. We do not consider the herd structure or demographics of this background population, but assume that during each

month the same number of individuals, I, immigrate to as emigrate from our focal herd. Further, we assume a constant proportion of the immigrants are susceptible (p_x), exposed (p_y), infectious (p_z), and vaccinated (p_v) for the duration of each simulation. This would be equivalent to vaccinating a focal herd that is embedded in a background population that is either in equilibrium or being maintained at some reduced prevalence. The model can then be used to inform managers about the effects of reducing movement between the focal herd and background population as well as reducing the prevalence of infection in the background population.

To incorporate sex and age structure, let $j = 1, 2$ represent the sex (1=male, 2=female) and $i = 1,...,18$ the age class. We assume that a smaller proportion (s^z) of infectious individuals (Z) survive each time period than healthy individuals and that the survival of calves depends upon the total population size. We control the relationship between annual and monthly time using the index r defined by the statement

IF mod(t,12)=0, THEN r=1 ELSE r=0.

Assuming events occur in the following order: vaccination (ψ), transmission (β), vaccine failure (δ) and disease incubation (γ), emigration and immigration (ε), and survival (s_{ij}), the monthly model equations conforming to the above assumptions are:

$$X_{i+r,j}(t+1) = s_{i,j}(N(t))\left((1-\varepsilon_{i,j})\left(\left(1-\frac{\beta\sum_{i=1}^{18}\sum_{j=1}^{2}Z_{i,j}(t)}{N(t)^{\theta}}\right)(1-\psi_{i,j}(t))X_{i,j}(t)+\delta V_{i,j}(t)\right)+p_x I_{i,j}(t)\right)$$

$$Y_{i+r,j}(t+1) = s_{i,j}(N(t))\left((1-\varepsilon_{i,j})\left(\left(\frac{\beta\sum_{i=1}^{18}\sum_{j=1}^{2}Z_{i,j}(t)}{N(t)^{\theta}}\right)(1-\psi_{i,j}(t))X_{i,j}(t)+(1-\gamma)Y_{i,j}(t)\right)+p_y I_{i,j}(t)\right)$$

$$Z_{i+r,j}(t+1) = s^{z}_{i,j}(N(t))\left((1-\varepsilon_{i,j})\left(\gamma Y_{i,j}(t)+Z_{i,j}(t)\right)+p_z I_{i,j}(t)\right)$$

$$V_{i+r,j}(t+1) = s_{i,j}(N(t))\left((1-\varepsilon_{i,j})(1-\delta)\left(V_{i,j}(t)+\psi_{i,j}(t)X_{i,j}(t)\right)+p_v I_{i,j}(t)\right)$$

$$I_{i,j}(t) = \varepsilon_{i,j}\left(X_{i,j}(t)+Y_{i,j}(t)+Z_{i,j}(t)+V_{i,j}(t)\right)$$

Annual reproduction is handled through the following statement that implicitly assumes an equal sex ratio at birth: IF mod(t,12)=0 THEN for j= 1 or 2,

$$X_{1,j}(t) = \sum_{i=1}^{18}\frac{b_i}{2}\left(X_{i,2}(t)+Y_{i,2}(t)+Z_{i,2}(t)+V_{i,2}(t)\right),\ Y_{1,j}(t)=0\,,Z_{1,j}(t)=0\,,\text{ and }V_{1,j}(t)=0\,,$$

where b_i is the expected number of offspring per female each calving cycle (approximated to be in January each year). Following Rodwell et al. (2001) in the KNP, we assume that BTB does not affect fecundity, although Jolles (2004) showed some effect of BTB on the fecundity of young and old buffalo, but not prime-aged buffalo from five to eight years old. Disease transmission in the model depends on $1/N^{\theta}$ (Anderson and May 1992), which is frequency-dependent when θ=1and density dependent when θ=0. Population size in this model is relatively constant, so that selecting θ= 0 or 1 is essentially equivalent to rescaling the transmission coefficient β. Thus, the value of θ is not critical to our analysis and we present results only for the case θ= 1.

The model incorporates density-dependent survival in the first age-class (0-1 year) because long-term studies of large herbivores suggest that survival of adults varies little compared to juveniles (Gaillard et al. 1998, but see Sinclair 1977). Density-dependent juvenile survival is regulated by an abruptness parameter ϕ, a scaling parameter κ, and the maximum survival rate s_0 (Getz 1996).

$$s_{i,j}\left(N(t)\right) = \frac{s_0}{1+\left(\frac{N(t)}{\kappa}\right)^{\phi}}, \quad i=1, \quad j=1,2$$

This form of density dependence results in a stable age structure and relatively constant herd size, which allowed us to investigate different vaccination strategies in the absence of large population fluctuations. Since we restrict analyses to a relatively stable population size, the form of density dependence is unlikely to play a major role in these simulations. We decreased the maximum survival rate of infectious calves by a constant α_0, where $\alpha_0 < s_0$, to obtain the survival function

$$s^{z}_{i,j}\left(N(t)\right) = \frac{s_0-\alpha_0}{1+\left(\frac{N(t)}{\kappa}\right)^{\phi}}, \quad i=1, \quad j=1,2$$

Finally, we assumed that adult survival is constant over time, but depended upon the age, sex, and disease status of the animal, whereby the survival of infectious individuals was reduced by α_1. We assumed that all buffalo in the 18th age-class died at the end of each year.

Parameter Estimation

We based estimates of buffalo survival and dispersal on preliminary analyses of over 130 radio-collared buffalo from ongoing research in the central region of the KNP (Caron et al. 2003, Cross et al. 2004, Cross et al. 2005). Depending upon the amount of herd fragmentation, the study area contained 4-12 buffalo herds and roughly 3000 buffalo. The majority of individuals were collared in four helicopter sessions: November 2000 ($N = 6$), April 2001 ($N = 27$), August 2001 ($N = 51$) and November 2001 ($N = 12$), while the remaining individuals were darted from ground vehicles throughout the study period. Animals were placed into age classes using incisor eruption patterns for those individuals under five years old (Pienaar 1969, Grimsdell 1973, Sinclair 1977). For those animals over five years we used horn development and wear to subjectively place individuals into two categories: 5-7 yrs and 8+ yrs. We re-sighted collared individuals, on foot and from vehicles, approximately 2-3 times per week throughout the study period. If an individual was missing for over one month we relocated them from fixed-wing aircraft. We tested buffalo for BTB using a modified gamma-interferon (IFNg) BOVIGAM^{TM} assay (Wood and Jones 2001), which has similar sensitivity (82-100%) and specificity (~99%) to the intradermal skin test (Wood and Jones 2001, Grobler et al. 2002). Negative individuals were retested at six or 12-month intervals.

Survival

We assumed that calf survival (s_{1j}) was density dependent and the maximum possible calf survival (s_0) was one (see above). Survival estimates for the other sex and age

categories were estimated from field data with known-fate models in program MARK using BTB-status, age, and gender as individual covariates (White and Burnham 1999). The dataset consisted of 132 radio-collared buffalo from May 2001 to November 2003. Twenty-two buffalo were BTB-positive on the first test. Thirteen animals that converted from BTB-negative to positive were reclassified as positive individuals at the time of their first positive test. We assumed buffalo were BTB-negative if they were a multiple, equal, or avian reactor on the gamma-interferon test. For model selection of the survival data, we used AICc, which is a modified version of Akaike's Information Criterion (AIC) that corrects for small sample sizes relative to the number of parameters (Burnham and Anderson, 2002). The AIC approach is a method of comparing the goodness-of-fit of nested and non-nested models and discourages the use of models with too many parameters that overfit the data. The minimum and maximum values shown in Table 1 are the 95% confidence intervals for each parameter using the delta method (White and Burnham 1999).

Dispersal (ε)

Previous studies suggest that only adult males move between herds (Sinclair 1977, Prins 1996). Using radio-tracking data of 120 buffalo in four herds from May 2001 to December 2003, we estimated the movement rate of buffalo between a focal herd and background population. We found lower cohesion amongst buffalo herds of the KNP than past studies (Cross et al. 2005). At the start of the study, the study area appeared to contain four herds that separated and re-fused over time. Three of those herds fragmented and fused with herds in other areas of the KNP, thus making it difficult to

define a dispersal event because the herd "dispersed" as a unit. As a minimum estimate of dispersal, we used the individuals in the remaining cohesive herd to estimate the probability of an individual moving from a focal herd to other herds in the background population. Individuals were assigned a value of zero or one respectively for every month they were present in the focal herd and either dispersed to another herd or remained in the focal herd. These data were then used in logistic regression analyses with sex and age as covariates. We excluded those dispersal events that lasted for less than one week because short duration events are unlikely to result in many disease transmission events. Animals were grouped into the following age categories: 1-2, 3-4, 5-7, and 8+ yrs old.

Disease transmission (β)

The transmission coefficient was estimated by fitting the model to the observed 1998 BTB prevalence values for the KNP (Rodwell et al. 2000). The model was parameterized with the baseline values in Table 1 and started with a disease prevalence of either 0.04 or 0.29 to correspond to the prevalence of BTB in the central and southern regions of the KNP in 1991. The initial population size was 250, to match the average herd size of the KNP (Whyte, unpublished data), and distributed among age classes to correspond to the stable age distribution predicted by the demographic component of the model. The model was run for 7 years and β was adjusted to minimize the sum of squared residuals between the predicted and observed prevalence in 1998. Transmission coefficient values that resulted in predicted 1998 prevalence values within the empirical 95% confidence intervals for the south and central regions were used to

establish the minimum and maximum values in Table 1. Simulations assumed a closed population and no vaccination. Data from the northern region of the KNP were not used to parameterize β because no BTB-positive individuals were sampled in that area prior to 1998.

Other Model Parameters

The pathology of BTB in buffalo and cattle suggests that BTB lesions are not encapsulated by the immune system and individuals are probably infectious within a few months of infection (De Vos et al. 2001). Neill *et al*. (1991) estimated the latency period of BTB in cattle to be 87 days, which translates to a monthly incubation rate γ of 0.21 (Table 1). Due to the limited data available on incubation rate, γ, we chose a wide range of 0.056 to one for the sensitivity analyses. Since ϕ is probably between two and six (see discussion in Getz 1996), we used a value of four and adjusted κ to yield a stable herd size of around 250 individuals. Given the limited data available on the duration of vaccine protection in buffalo and the indication in Berggren (1977) and Francis (1947) that protection wanes in vaccinated cattle between one and five years, we explored a wide range of vaccine failure rates from lifelong protection (*i.e.* δ=0) to a half-life of one year.

Sensitivity Analyses

Sensitivity analyses were conducted using Monte Carlo methods to assess the relative effects of several model parameters. Specifically, 10,000 random parameter sets were created by choosing ψ, δ, β, γ, $\varepsilon_{i,j}$, p_z, $s_{i,j}$, α_0 and α_1 from uniform distributions

bounded by the minimum and maximum values shown in Table 1. We assumed that the vaccine was 100% protective for this analysis. . Each parameter set was used to run the deterministic model once. The prevalence at year 50 was recorded for each run and used as the dependent variable in multiple linear regressions where model parameters were the explanatory variables (Wisdom and Mills 1997, Wisdom et al. 2000, Cross and Beissinger 2001). To facilitate comparisons between parameters measured on different scales all model parameters were transformed to percentage difference from the mean (*i.e.* $(x_i - \bar{x}) / \bar{x}$, where x_i is the value of the model parameter on run i and $\bar{x}$ is the mean) prior to statistical analysis. Subsampling this dataset and evaluation of standard errors revealed that 10,000 runs were sufficient to solidify the ranking hierarchy of the top six model parameters in the statistical analyses. The other eight parameters in the sensitivity analysis explained little variability in either disease prevalence or quasi-eradication, had large standard errors, and were omitted from Table 2.

Model parameters were ranked according to the magnitude of their standardized coefficients (*i.e.* the regression coefficient divided by its standard error), which is a unitless quantity expressing the unique contribution of that variable scaled by the estimation uncertainty (Selvin 1995). Model parameters that are good predictors of disease prevalence may be different from those that are good predictors of disease eradication. Logistic regression was used to identify the latter parameters (McCarthy et al. 1995, Cross and Beissinger 2001). For the logistic regression analysis, disease prevalence at year 50 was converted to a binary variable of disease persistence or eradication. We used a prevalence of 5% as the quasi-eradication threshold because very few parameter sets resulted in quasi-eradication at lower threshold values making

maximum likelihood estimation difficult in the logistic regression analysis. As in the linear regression analysis, each model parameter was transformed to percentage difference from the mean and individually used to predict disease eradication at year 50. All simulations started with 250 individuals in a stable-age distribution and an initial prevalence of 0.05, evenly distributed amongst all sex- and age-classes. The epidemiological model was coded in Matlab 5.3 (Mathsoft™) and statistical analyses were conducted using S-Plus 6 (Insightful Corp 2001).

Results

Field Data Parameter Estimates

Our survival data indicated that males and individuals over 8 yrs old had lower survival rates than females and those under 8 yrs old (Table 1). Age category (1-7 vs. 8+) and gender were statistically significant explanatory variables in likelihood ratio tests (LRT) where we included one variable at a time (p =0.004 and 0.001, respectively; df = 1) and were supported in AICc analyses. There was little statistical support for the inclusion of more refined age categories using either AIC or LRT methods. Our analyses did not indicate any additional mortality amongst the 35 BTB-positive animals (df = 1, p = 0.19). Even though this is based on only the first two years of data, this result, in combination with previous cross-sectional analyses, suggests that the annual disease induced mortality associated with BTB is probably between zero and 10 percent (Rodwell et al. 2000, Jolles 2004). Statistical model comparisons using AICc values and sex, age (1-7yrs, 8+), and BTB-status as explanatory variables indicated that the additive sex+age model provided the best fit to the survival data. A sex*age model was

the only other model that appeared to be supported by the data (ΔAICc = 0.869). Using the sex+age model we estimated the survival rates shown in Table 1.

In contrast to previous studies of buffalo, we found that all sex and age groups moved between herds, albeit adult males moved between herds more frequently than young males and females (Table 1, Fig. 1). Logistic regression analyses indicated a significant age*sex interaction ($p = 0.02$, $df = 3$, deviance = 9.88) whereby males over 8 yrs old were more likely to disperse (Table 1, Fig. 1). Due to the similar dispersal rates of females and juvenile males we aggregated these sex and age categories and reran the logistic model before calculating the dispersal rates in Table 1.

Sensitivity Analyses

Sensitivity analyses of this model around the baseline set of parameter values indicate that for the model and parameter space presented, disease prevalence in the background population, p_z, is the primary factor determining the BTB prevalence in the focal herd (Table 2, $r^2 = 0.95$). Vaccination rate ψ and the transmission coefficient β were the second and third parameters in the ranking hierarchy, while survival rates explained very little variability in BTB prevalence or eradication. There were few differences in the ranking of the top five parameters between the linear and logistic regression results indicating that parameters that regulate prevalence also determine the probability of eradication.

Model Results Without Dispersal

To assess the effects of other variables we analyzed the model in the context of a single herd without dispersal (such as a cattle herd or small reserve). Using the baseline values in Table 1, the model predicts that BTB should reach an asymptotic prevalence of ~0.44 in a herd of ~250 individuals, and as would be expected of a chronic disease without recovery, the model does not exhibit any cyclical behavior (Fig. 2). Optimal vaccine strategies depend on the distribution of vaccinations to different sex and age categories and the duration of vaccine protection. Assuming that the vaccine grants lifelong protection, the model suggests that control programs should focus upon younger individuals. In particular, focusing the control strategy on younger individuals produces a higher ratio of successful to total vaccinations (*i.e.* less wasted vaccinations) than vaccinating individuals in proportion to the age structure of the population (Fig. 3). As vaccine duration decreases, however, the advantages of vaccinating juveniles decreases and the lines on Fig. 3 approach the efficiency of the calf-only vaccination strategy. In other words, when the vaccine lasts for shorter amounts of time there are fewer redundant vaccinations of older individuals. Few differences exist between the efficacy of male or female vaccination programs of a focal herd without dispersal (Fig. 3). However, in model simulations with dispersal (data not shown) vaccination programs were more effective if they focused on females because vaccinations of dispersing males were, in effect, wasted with respect to reducing prevalence in the focal herd.

In the best-case scenario of a 100% effective vaccine that provides life-long protection, a vaccination program focusing upon calves would require the vaccination

of around 70% of the calves every year to eradicate (*i.e.* prevalence < 1%) BTB by year 50 (Fig. 4a). A program focused on the entire population would require a vaccination rate of less than 30% for eradication (Fig. 4a). However, 70% of the calf population translates into ~1300 vaccinations over a 50-year period; whereas, 30% of the total population translates into ~3100 vaccinations over the same time period in a herd of approximately 250 individuals. Thus a calf-only policy requires higher coverage of that age-class but fewer vaccinations in total. This is not surprising given that calves are approximately 16% of the total population and we have assumed lifelong vaccine protection.

Given that vaccination rate ψ and vaccine failure rate (also expressible in terms of the half-life of the vaccine) δ are two parameters that may be altered by managers or scientists, we calculated BTB prevalence at year 50 for different combinations of these parameters and calf-only vaccination program (Fig. 5). Model results suggest that if 70% of the calf population is vaccinated every year BTB would may be eradicated by year 50 assuming the vaccine granted lifelong protection and no migration between the focal herd and background population. If the half-life of the vaccine was less than five years, however, a vaccination program is unlikely to eradicate the disease by year 50 (Fig. 5).

Discussion

The management issues surrounding bovine tuberculosis in the Kruger National Park of South Africa are typical of many invasive species and emerging infectious diseases. Immediate actions are more likely to be effective, but limited data are

available upon which to base management and research decisions. We used field data to bound the range of possible parameter values and simulated an SEI disease model to assess the importance of different model parameters and the effectiveness of vaccination as a control strategy. Our results, based upon the first analysis of a dynamic epidemiological model of BTB in a buffalo population, indicate that vaccination alone is unlikely to be an effective management tool to eradicate BTB. The literature on modeling BTB in other species is extensive and conclusions vary, but, in general, our conclusions about the utility of vaccination is in contrast with more encouraging modeling studies on vaccination of BTB in possums, badgers, and cattle (Barlow 1991, White and Harris 1995, Roberts 1996, Kao et al. 1997, Tuyttens and MacDonald 1998, Smith 2001, Smith and Cheeseman 2002, Wilkinson et al. 2004). This difference may be due to the high prevalence of BTB in African buffalo (>50%; Rodwell et al., 2000) compared to <20% for badgers and possums (Krebs et al. 1997, Woodroffe et al. 1999, Coleman and Cooke 2001).

Our model predicts that BTB should reach an asymptotic prevalence of around 44% assuming a closed population and a relatively constant herd size of 250 individuals (Fig. 2). Data from one herd in the southeastern corner of the KNP suggests that this estimate of asymptotic prevalence may be low. The Mpanamana herd of the KNP had a stable prevalence of 67% in 1992 and 1996 (De Vos et al. 2001). In addition, Jolles (2004) estimated the asymptotic BTB prevalence to be 53% (95% CI = [49-58%]) in Hluhluwe-Umfolozi Park. Additional data on BTB prevalence and transmission rates would help discern whether our parameter estimates or model structure should be modified to match the higher prevalence seen in these two herds. If this is the case, then

our model results may represent conservative estimates of the vaccination effort necessary to control and eradicate BTB.

Assuming frequency-dependent transmission and a vaccine that is equally effective and protective across all sex- and age-classes, vaccination programs will be most effective in reducing prevalence in the focal population if they focus on calves (Figs. 3 and 4). Younger individuals are unlikely to be infected or previously vaccinated, thus increasing the ratio of successful to total vaccinations (Fig. 3). Also, younger individuals will, on average, remain longer in the population than older individuals. Therefore, vaccinated calves will contribute more to herd immunity than vaccinated adults. The BCG vaccine has been the most comprehensively tested BTB vaccine in wildlife and cattle (e.g. Ellwood and Waddington 1972, Waddington and Ellwood 1972, Berggren 1977, Aldwell et al. 1995, Buddle et al. 1995a, Buddle et al. 1995b, Buddle et al. 1995c, Buddle et al. 1997). Due to the fact that the efficacy of the BCG vaccine may be reduced by prior exposure to environmental mycobacteria (Buddle et al. 1995a, Fine 1998), BCG vaccination may also be most effective in the younger age groups. Thus the vaccination of younger rather than older individuals is supported for both biological and mathematical reasons (Fig. 3).

Even in the best-case scenario of a 100% effective vaccine with lifelong protection, a management program focused on calves would need to vaccinate around 70% of calves every year to eradicate BTB by year 50 (Fig. 4a). In a herd of 250 buffalo this translates to around 1500 vaccinations over a 50-year period (Fig. 4b). Our predictions may be overly optimistic since we assumed that all vaccinated individuals were protected against infection and our estimated asymptotic herd prevalence may be

lower than that observed in the field. Given that: (1) the current BCG vaccine is very unlikely to be 100% effective in buffalo, (2) that protection probably wanes over time (Berggren 1977, Buddle et al. 2000), (3) that drug and labor costs for vaccinating a buffalo are over $100 (US) per individual, and (4) a high percentage of individuals must be vaccinated to eradicate the disease, the eradication of BTB via vaccination alone is probably not an effective management strategy. Vaccination may be useful to control BTB at lower prevalence levels, but then the question arises as to whether the cost of an indefinite control program outweighs the benefit in reduced prevalence. At this point, we cannot answer this question because it rests upon the effect of BTB on lion populations and the relationship between the prevalence of BTB in lions and buffaloes. Additional work on the interaction of lions, buffaloes, and BTB, and the degree to which lions select for BTB-positive individuals would be enlightening. Further, additional modeling work is necessary to assess the potential effectiveness of vaccination, in combination with other control measures such as a test-and-remove program, in a spatial context.

Sensitivity analyses suggest that disease incubation (γ) and buffalo survival rates were relatively unimportant model parameters (*i.e.* they explained very little variability in the prevalence or probability of eradication of BTB). Surprisingly, the dispersal rate was also relatively unimportant in the sensitivity analyses. The importance of dispersal, however, is reflected in the importance of the background prevalence, whereby higher dispersal rates increase the importance of the background prevalence (data not shown). Vaccination rate and the transmission coefficient are a distant second and third in the ranking hierarchy compared to the disease prevalence in the background population

(Table 2). This result is obvious in hindsight given the model structure presented here and the high dispersal rates of buffalo in the KNP (Table 1, Fig. 1). It is important to note, however, that other studies of African buffalo indicate a more stable herd structure with less movement between herds in other areas (Sinclair 1977, Prins 1996). If these differences reflect real differences rather than an artifact of sampling intensity it would suggest that disease control is likely to be more effective in these populations.

We parameterized the epidemiological model using previously published parameter estimates as well as data from a longitudinal study of radio-collared individuals in the central region of the KNP. The dispersal and survival rate estimates presented here are the first estimates based upon longitudinal studies of known radio-collared individuals. Previous estimates were based upon cross-sectional life-table analyses and/or a few known, but unmarked individuals (Sinclair 1977, Prins 1996, Jolles 2004). Analysis of the longitudinal dataset of known individuals indicated that dispersal rate varied by sex and age, whereby males over eight years old were the most likely to disperse from the focal herd (Table 1, Fig. 1). In contrast to previous studies, we found females and juvenile males also moved between mixed herds via splinter groups when herds split and later fused with other herds (Cross et al. 2004, Cross et al. 2005). We began the study with four herds in the study area. Since 2001 these four herds have splintered into as many as 13 herds and only one of the original herds remained as a cohesive unit within the study region. As a result it is difficult to define dispersal events when the herd itself is changing. We used the one cohesive herd to estimate dispersal rates, which, given the fluid and mobile nature of the other herds,

should be seen as minimum estimates and highlights the importance of developing spatial models of disease spread.

Survival analyses indicated that survival was a function of both age and sex. Age categories could be collapsed to 1-7 vs 8+ yrs and females survived better than males (Table 1). Comparing models, both the additive (sex+age) and interactive (sex*age) survival models were supported by the data based upon AIC values and likelihood ratio tests. We did not find any increased mortality of BTB-positive individuals during the first two years of the study. Previous research by Rodwell et al. (2001), Jolles (2004), and Caron et al. (2003) suggests that the additional mortality due to BTB infection is around 11% or less. With our sample of 35 BTB-positive and 97 BTB-negative individuals, we would be unlikely to detect small differences in survival rates in a two-year timeframe. The low mortality rate of infected individuals and the lack of any known recovery suggest that buffalo will maintain BTB at high levels (Fig. 2). This conclusion, in combination with the difficulty of eradicating this disease with a vaccination or culling program, and the ability of BTB to spill-over into other hosts (Bengis et al. 1996) presents a worrying scenario. Since lions are the dominant predator of adult buffalo, and BTB can infect lions via the gastrointestinal tract (Keet et al. 1996), the largest effects of BTB may be in altering the competitive dynamics of the large predator guild in the KNP.

The model presented here is based upon best empirical data currently available and provides an objective view of the likely effectiveness of a vaccination strategy as well as highlights important research and management issues. The importance of the background population in the sensitivity analysis suggests that managers should view

herds, or parks, as open systems and incorporate the surrounding populations into their control strategies, and the degree to which such invasions must be controlled depends on the background prevalence of the disease. Further, researchers should focus on spatial disease models to incorporate individuals' movement patterns. The inclusion of spatial complexity into our model, and more data on transmission rates and the inter-herd movement patterns of individuals are needed to assess the likely efficacy of a combined vaccination and selective removal of infected individuals in containing the spread of BTB in African buffalo. Nonetheless, the model presented here suggests that even in the best case scenario, vaccination alone is unlikely to be an effective control strategy for BTB in buffalo, and thus research and managers should focus on other possible methods to control the spread of this exotic disease.

Table 1. Parameter estimates used in the buffalo vaccination model.

Parameter	Symbol	Minimum	Baseline	Maximum	Source
Annual buffalo survival					
Maximum calf survival	$s_{1,\,1\text{-}2}$	0.95	1.00	1.00	1
Young males	$s_{2\text{-}8,1}$	0.74	0.84	0.90	1
Old males	$s_{9\text{-}18,1}$	0.20	0.59	0.86	1
Young females	$s_{2\text{-}8,2}$	0.83	0.95	0.99	1
Old females	$s_{9\text{-}18,2}$	0.35	0.86	0.98	1
Scaling parameter	κ	--	400	--	see text
Abruptness parameter	ϕ	2	4	6	2
Annual buffalo reproduction					
Cows 3-4	r_3	--	0.51	--	3
Cows 4-5	r_4	--	0.64	--	3
Cows 5+	r_{5+}	--	0.68	--	3
Monthly dispersal					
Immature males	$\varepsilon_{1\text{-}6,1}$	0.01	0.02	0.04	1
Mature males	$\varepsilon_{7\text{-}9,1}$	0.24	0.09	0.03	1
Old males	$\varepsilon_{10+,1}$	0.45	0.26	0.13	1
Females	$\varepsilon_{1+,2}$	0.04	0.02	0.01	1
Monthly disease parameters					
Transmission coefficient	β	0.034	0.043	0.053	1
Incubation rate	γ	0.056	0.21	1	4
Reduction in maximum juvenile surviva	α_0	0	0.0043	0.0084	5
Reduction in adult survival	α_1	0	0.0043	0.0084	5
Transmission exponent	θ	0	--	1	see text
Vaccination rate	ψ	0	--	1	see text
Vaccine failure rate	δ	0	--	0.056	6
Background prevalence	p_z	0	--	0.7	see text

1) this study; 2) Getz 1996; 3) Funston 1999; 4) Neill et al.1991, de Vos et al. 2001; 5) Rodwell et al. 2001, Caron et al 2003, Jolles 2003; 6) unknown, but see Berggren 1977.

Table 2: Logistic and linear regression sensitivity analysis of 10,000 runs of the vaccination model using parameter values chosen from uniform distributions. All parameters were transformed to % change from the mean and only the six most important variable are shown.

		Linear regression[a]		Logistic regression[a]	
Model parameter	Variable	Coefficient	b/SE[b]	Coefficient	b/SE[b]
Background prevalence	p_z	0.222	476.5	-97.8	-13.0
Vaccination	ψ	-0.036	-77.1	3.0	9.3
Transmission coefficient	β	0.138	64.8	-16.8	-9.6
Reduction in adult survival	α_1	-0.019	-39.4	2.1	7.0
Female and juvenile dispersal	$\varepsilon_{1\text{-}5,1;1+,2}$	0.026	32.7	3.0	6.2
Vaccine failure	δ	0.013	27.1	-1.0	-4.1

[a] Multiple linear and logistic regressions with year 50 disease prevalence or eradication as the dependent variable, respectively.

[b] standardized coefficients were divided by the standard error to standardize measurements and yield a comparative measure of importance.

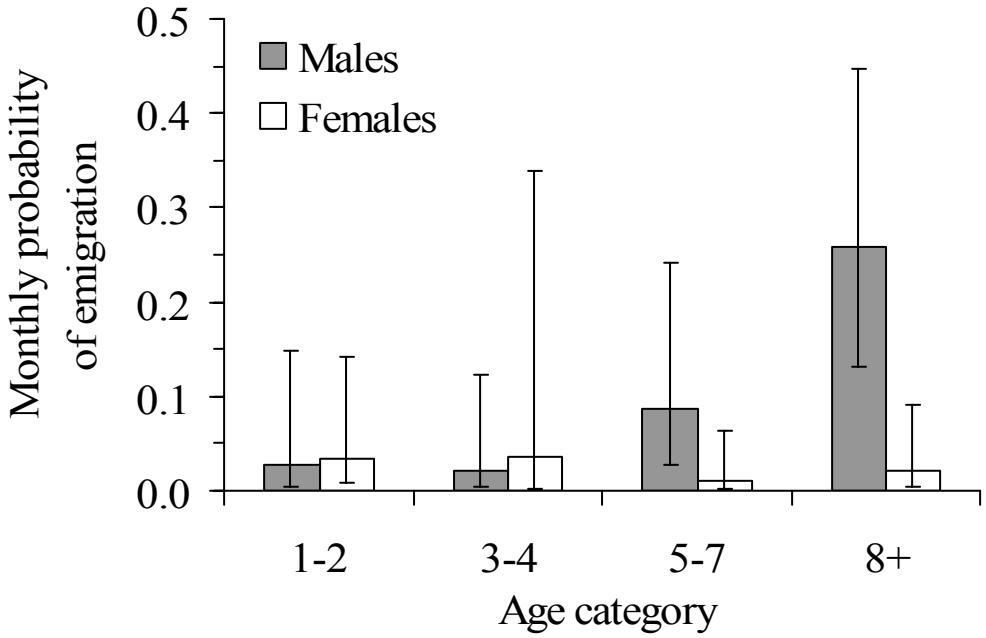


Figure 1. Buffalo emigration rates between a focal herd and the background population varied by sex and age in the Central Region of the Kruger National Park. Estimates were based on logistic regression analyses of the probability of dispersal per month. Bars represent the 95% confidence intervals.

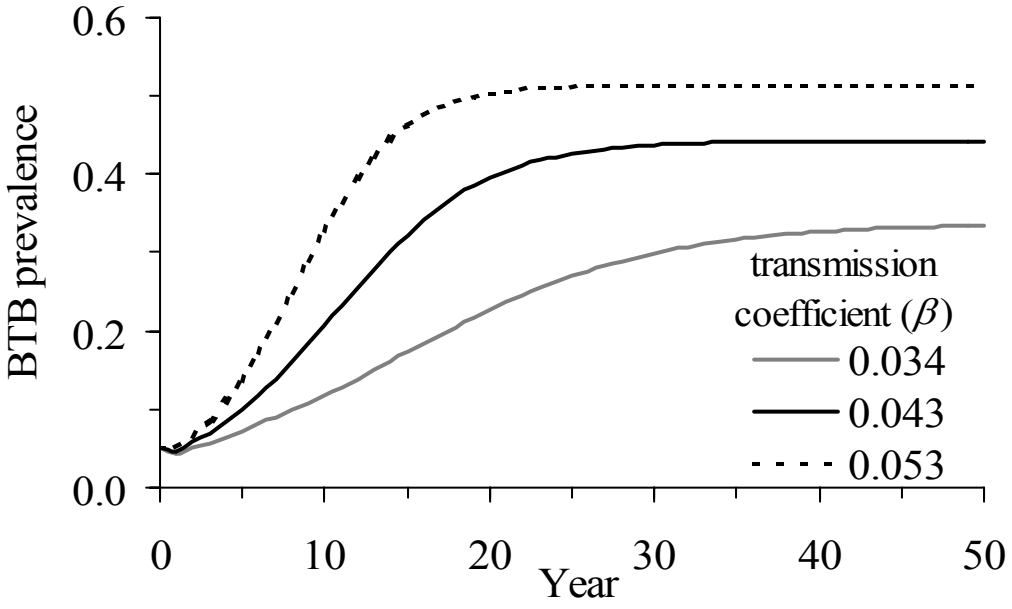


Figure 2. Predicted BTB prevalence plotted as a function of time using the upper, lower and baseline values of the transmission coefficient (β = 0.034, 0.043, 0.053). Simulations assume a closed population and no vaccination effort.

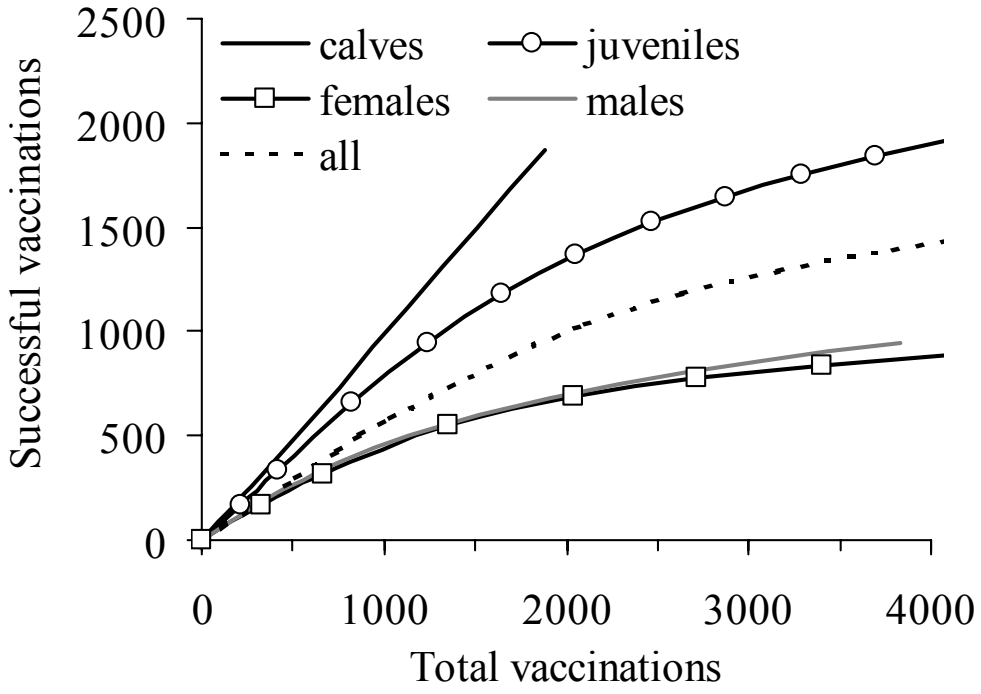
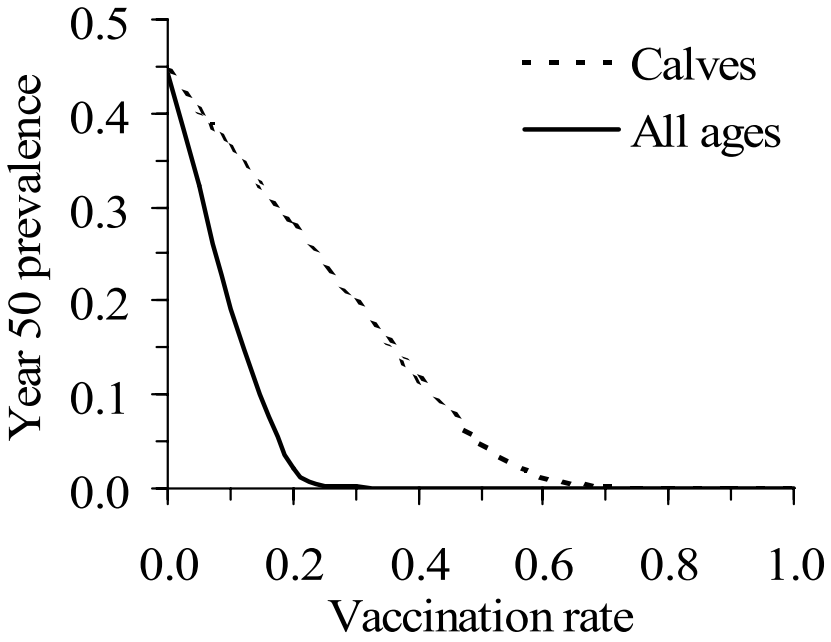


Figure 3. The number of successful vaccinations for the indicated vaccination strategies (*i.e.* focusing on calves, juveniles, males, females or the entire population) plotted as a function of the total vaccinations in a herd by year 50. "Successful" is defined as the vaccination of uninfected and previously unvaccinated individuals. "All" indicates a management strategy that vaccinates all individuals in proportion to their abundance in the population. Simulations assume a closed population and lifelong vaccine protection.

a)



b)

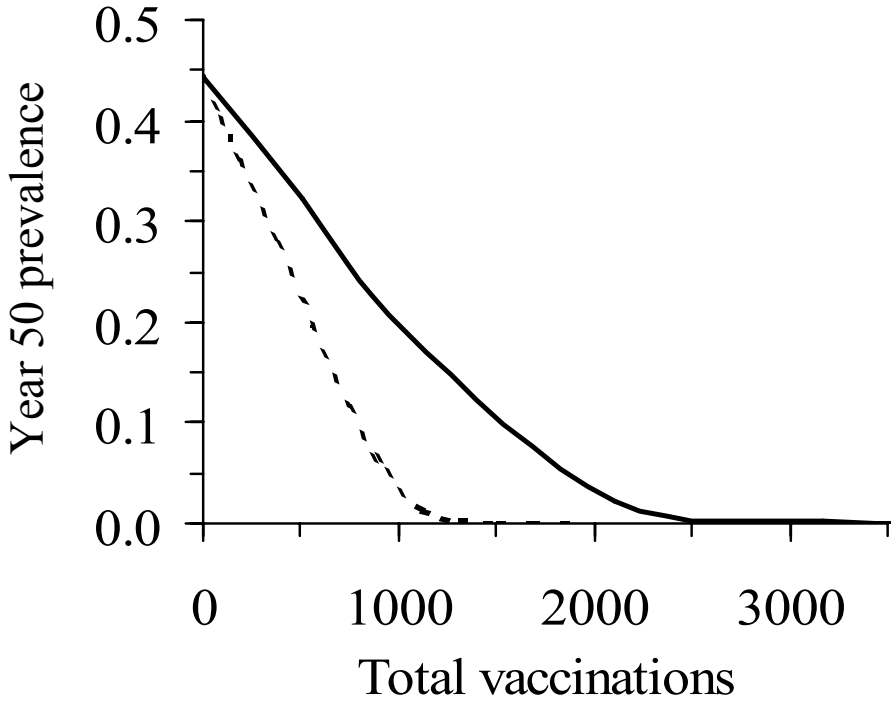


Figure 4. Prevalence of BTB at year 50 plotted as a function of (a) vaccination rate and (b) cumulative vaccinations to demonstrate the importance of focusing vaccination effort on calves. Simulations assume a closed population, vaccination in proportion to abundance in the focal population, and lifelong vaccine protection.

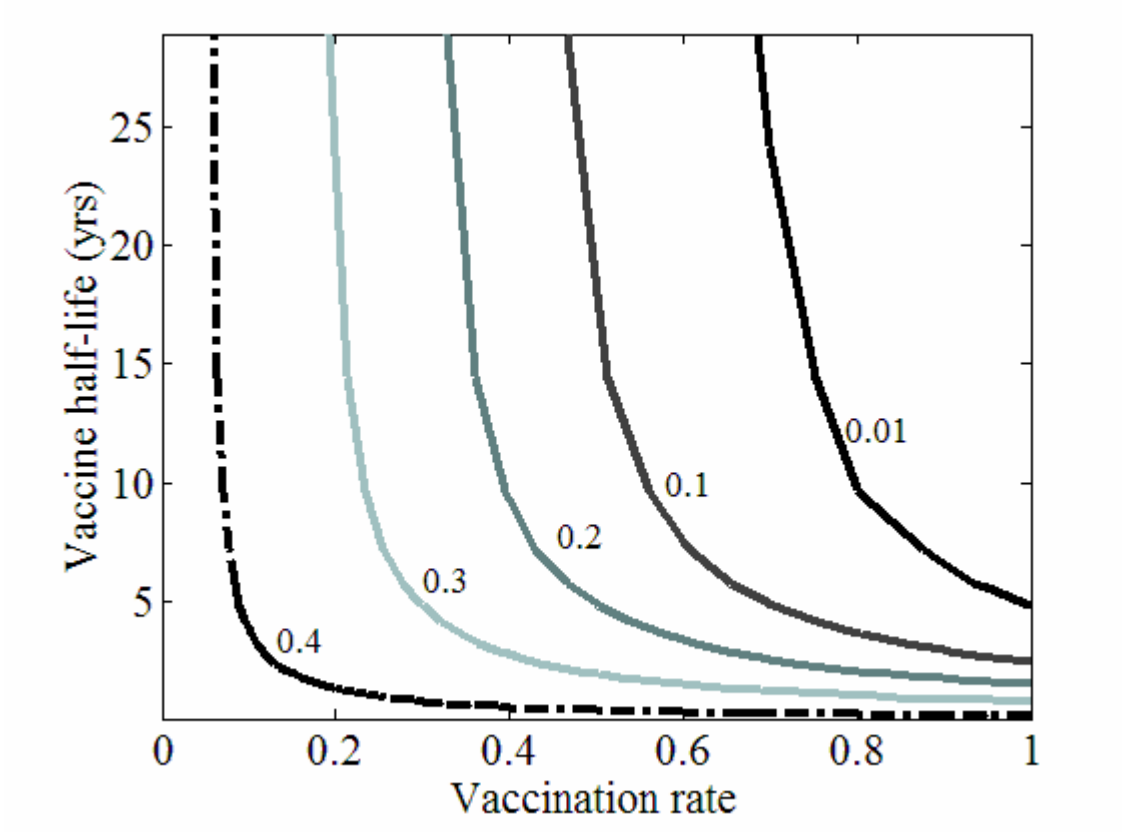


Figure 5. Prevalence isopleths at year 50 as a function of vaccination rate and vaccine half-life. Simulations assume a closed population and an annual calf-only vaccination program. Vaccine half-life refers to the amount of time before half of the vaccinated individuals are susceptible again.

Chapter Three

Disentangling association patterns in fission-fusion societies using African buffalo as an example

P.C. Cross, J.O. Lloyd-Smith and W.M. Getz

Introduction

A description of the social network of a population aids us in understanding dispersal, the spread of disease, and genetic structure in that population. Many animal populations can be classified as fission-fusion societies, whereby groups form and separate over time. Examples discussed in the literature include ungulates, primates, and cetaceans (Lott and Minta 1983, Whitehead et al. 1991, Henzi et al. 1997, Christal et al. 1998, Chilvers and Corkeron 2002). In this study, we use a heuristic simulation model to illustrate potential problems in applying traditional techniques of association analysis to fission-fusion societies and propose a new index of association: the fission-decision index (FDI). We compare the conclusions resulting from traditional methods with those of the FDI using data from African Buffalo (*Syncerus caffer*) in the Kruger National Park (KNP). The traditional approach suggested that the buffalo population was spatially and temporally structured into four different "herds" with adult males only peripherally associated with mixed herds. Our FDI method indicated that association decisions of adult males appeared random but those of other sex and age categories were non-random, particularly when we included the fission events associated with adult males. Further, the amount of time individuals spent together was only weakly correlated with their propensity to remain together during fission events. We conclude with a discussion of the applicability of the FDI to other studies.

Researchers attempting to quantify individual association patterns in fission-fusion societies often use group membership as an indicator of association, calculating an index of association for all pairs of individuals (or dyads) based on the proportion of time spent in the same group (Cairns and Schwager 1987, Ginsberg and Young 1992,

Whitehead and DuFault 1999). Traditional association indices that are based upon the proportion of time together, however, may be the product of two underlying processes: the fission and fusion of groups, and the choices of individuals as to which subgroup to join during a fission event. Most studies calculate association indices using the entire dataset–without paying attention to the timing of fission events. While these indices represent the proportion of time two individuals spend together, they may be poor estimates of the propensity for dyads to remain together during future fission events. Each fission event provides only one data point on a pair's likelihood of remaining together when the group separates, and additional samples within a fission-fusion event are auto-correlated with samples occurring on the same interval between fission and fusion events.

We define a fission event as when one group of individuals separates into two or more distinct subgroups. For our study system, distinct groups are readily identified spatially: in KNP, African buffalo exist in herds of ~200-1200 individuals, and individuals within herds are usually separated by a few meters. Herds are typically separated by one to 40 kilometers and are rarely within visual contact of one another. Thus, we defined a group as the set of individuals that were within ~1km of one another. This definition was only problematic when groups were relatively close (e.g. within a diameter of the group size itself), which occurred rarely; and when groups were close to one another they were either in the process of splitting apart or rejoining. More generally, we might define groups as distinct if they are sufficiently separated for individuals to incur significant cost (e.g. greater exposure to predation) if they were to move between groups. Very mobile species, or those that can communicate over large

distances, may perceive groups at a larger spatial scales than species that use groups as a form of predator defense. Ultimately, the definition of a group will vary between study systems and the appropriate definition may depend on the questions being addressed.

The distinction between the fission-fusion of groups and individual decisions during fission events has not been made in traditional association analyses, and each measurement of group membership is assumed implicitly to be the result of independent individual choices. For many species, however, dispersal between fission groups may be limited by predation, reduced foraging efficiency, and/or hostility from conspecifics (e.g. Waser et al. 1994, Alberts and Altmann 1995, Isbell and Van Vauren 1996, Ferreras et al. 2004). Thus, an individual's association with other individuals may be constrained because it is unable to independently move between groups. Factors influencing the fission and fusion of groups are often not well understood. In the absence of information, one might assume that either individuals have no control over the timing of group fission events or group fission events are the product of individuals' dissatisfaction with their associates.

Here we consider the case in which the timing of fission and fusion events are beyond any individual's control. From this viewpoint, traditional association indices are a function of both individual choices during a fission event and the duration that a splinter group remains separate after the fission event. We propose a modified pair-wise association index, the fission-decision index (FDI), which is the proportion of fission events involving both individuals in which they choose the same post-fission subgroup. This limits sampling to only one point during the interval between fission and any

subsequent fusion events. Let T_{ij} be the number of times individuals i and j were together after fission events, and A_{ij} be the number of times i and j separated during fission events. The FDI, which we denote by δ_{ij}, is given by the formula:

$$\delta_{ij} = \frac{T_{ij}}{T_{ij} + A_{ij}}.$$

Note that the symmetry in the roles of individuals i and j implies that $\delta_{ij}=\delta_{ji}$.

In this study, we use simulation models in two contexts. First we use a heuristic model to illustrate some of the deficiencies of traditional analyses by generating simulated association data for particular fission-fusion processes and sampling regimes. In these simulations, individuals choose to join post-fission subgroups at random. We analyze the simulated data using standard methods and show that auto-correlated data lead to statistically significant association patterns that do not reflect the random individual choices made during fission events. Repeating the analysis using the FDI reveals the random structure of the data. In the second section we apply our FDI method to an empirical dataset of 123 radio-collared buffalo, and use a simulation model to generate the expected distribution of FDI values for the observed fission-fusion history if all buffalo had chosen subgroups at random. We then compare the conclusions that are drawn from the traditional and FDI approaches. We conclude with a discussion of scenarios where the FDI should be used in conjunction with traditional methods of analysis to gain a more comprehensive understanding of animal associations.

Methods

Heuristic Simulation Model

Our first simulation model is intended to clarify conceptually the need for, and utility of, the FDI. The model generates simulated association data for particular fission-fusion histories, subject to the rule that individuals choose subgroups at random. Using this model, we assessed how conclusions about social structure depend on the association index used and on sampling protocol and intensity. We analyzed the simplest possible fission-fusion society: one group, which sometimes separates into two groups and later fuses back together. This simplification is intended to expose potential bias with a minimal level of complexity rather than quantify the bias present in more complicated field situations.

To make the model, we first generated fission-fusion histories (e.g. Fig. 1a) that described when the group splits apart and regroups. Fission and fusion were treated as Poisson processes, occurring randomly with a constant probability per unit time. This is a simple way of modeling the fission process, which agrees with the roughly negative exponential distribution of group lifetimes for African buffalo in the KNP (Cross, unpublished data; i.e. the probability of a group ceasing to exist remains constant and independent of how long it has existed in the past). Fission and fusion rates were set equal, so groups spent half of their time apart, on average. For a given fission-fusion history we simulated the movements of 20 individuals as they randomly chose subgroups during fission events. We sampled group membership at different intensities to generate data from which association indices could be calculated for all dyads. Sampling events occurred at regular intervals (or once following each fission event, for the FDI; Fig. 1a), and all individuals were recorded during a sampling event.

African Buffalo in the KNP: Field Methods

Field data were collected during an ongoing study of bovine tuberculosis in the Satara Region of the Kruger National Park from November 2000 to November 2003 (Caron et al. 2003). The study area contained 4-12 buffalo herds, depending upon the amount of herd fragmentation, and roughly 3000 buffalo. The majority of study individuals were fitted with radio-collars in four helicopter sessions: November 2000 ($N = 6$), April 2001 ($N = 27$), August ($N = 51$) and November 2001 ($N = 12$). The remaining individuals were darted from ground vehicles throughout the study period. Animals were placed into age classes using incisor eruption patterns (Pienaar 1969, Grimsdell 1973, Sinclair 1977). All individuals over 5 years old were classified as adult. Although data for some individuals were available from November 2000, we restricted the dataset to sightings of 123 radio-collared individuals that were seen more than five times from January 2002 through October 2003. During this period we had relatively complete information about the fission-fusion process. We monitored the buffalo herds approximately 2-3 times per week (917 herd sightings on 351 days) from distances ranging from 50 to 1000m. If an individual was missing for over one month we located it from aircraft. Group membership was recorded only once per day per individual. Since all marked individuals had radio-collars and herds were usually separated by several kilometers, we could determine which individuals were in a herd without visually sighting all individuals.

Although fission events occurred when groups of individuals separated, we identified fission events in the buffalo dataset as any time two radio-collared individuals were together on one sighting and then recorded in different groups, separated by

several kilometres, in their next sighting. Individuals other than adult males only separated as a result of a larger group-level fission event. Adult males, however, often moved to smaller bachelor groups (~2-30 individuals) from mixed groups (Sinclair 1977, Prins 1996), and thus the definition of a fission event can be made dependent on the class of animals being considered (in our case adult males versus females and juveniles). By analysing datasets with or without adult males included we show how the FDI approach accurately captures this aspect of buffalo biology, whereby females are rarely seen moving with males into bachelor groups.

On some occasions single radio-collared individuals were absent from the dataset for a short period of time and then returned to the same herd where they were last seen. These data records may occur due to actual fission events that involved only one marked individual, or due to data-entry errors. We required that an individual be absent for at least two successive observations of its last-known herd before we considered it a fission event. According to this definition, fission events involving only one marked individual occurred 16 times out of a total of 185 fission events for the dataset excluding adult males, and 38 times out of 375 fission events when adult males were included. We analyzed the FDI values for all dyads that were involved in two or more fission events together (N = 1093 and 834 with and without male fission events, respectively).

Buffalo Data Randomization

We compared the results of the buffalo FDI analysis to 1000 simulated random datasets. To generate these we collected the following data from the real dataset: identification of

individuals involved in each fission event, the number of fragments sighted after the fission event, and the number of individuals in each fragment. To simulate random decisions, we conserved the number and size of post-fission fragments but distributed the individuals at random between them. This eliminated the variability associated with different herd sizes and isolated the variability associated with individual fission-decisions. We simulated fissions on an event-by-event basis, so that each simulated fission event was begun with the same individuals on-hand as in the buffalo data. This served to maintain the same number of fission events per dyad in the simulated and buffalo datasets, which has important implications for the distribution and variance of FDI values.

We encountered one problem when comparing the simulated and real FDI values. In real datasets, dyads will either separate from one another during their last shared fission event or the dyad will still be together when one animal dies or the study ends. Because fission events were frequent, the FDI values for dyads with only two events were either 0 or 0.5 (since their final event must have been a separation, or else they would have been involved in further fission events). However, the simulated FDI values could equal 0, 0.5 or 1 for dyads with only two events, because the individuals could randomly choose to remain together both times (and yet not undergo further shared fissions, since simulated fission events used the on-hand individuals from the real data). Taking this approach, simulated FDI values would be biased upward relative to the data-based FDI values for these two-event dyads. The same reasoning applies to dyads with greater numbers of fission events. To correct for this bias and to make the

simulated and real datasets comparable, we did not include the last decision of each dyad in the analysis of either the buffalo or simulated datasets.

Statistical Analysis

Following Whitehead and DuFault (1999), we considered two individuals to be associating if they were located in the same group. This one-zero metric of association was used to calculate the proportion of samples in which two individuals were seen together (*i.e.* the simple ratio index). For the case presented here, where all individuals have radio-collars, the probability of locating a pair of animals is unlikely to be related to whether they are together or apart, and as a result the simple ratio index yields an unbiased estimate of the proportion of time they spent together (Cairns and Schwager 1987, Ginsberg and Young 1992). To test the statistical significance of the traditional simple ratio we applied the permutation methods described by Bejder et al. (1998), using programs modified from the SOCPROG 1.3 package (http://is.dal.ca/~hwhitehe/social.htm) to shuffle individuals within samples to test the null hypothesis that no preferred companions exist between sampling periods. We considered the null hypothesis rejected if fewer than 5% of the permuted datasets had a standard deviation greater than that of the original dataset. The standard deviation of the distribution of all pair-wise association indices is expected to be higher when certain individuals preferentially associate with others (Whitehead 1999). For datasets generated by our heuristic simulations, we found that *p*-values stabilized around 40,000 permutations. Statistical significance of the buffalo FDI values was determined using a χ^2 goodness-of-fit test to compare the empirical distribution of FDI values to that of the

mean of 1000 simulations (see Buffalo Data Randomization above). Dendrograms were generated using the unpaired group averaging method (UPGMA), which had the highest cophenetic correlation coefficient compared to other cluster analysis methods (Romesburg 1984). Simulations were coded in MATLAB 6.1, while cluster analyses were conducted using the MATLAB 6.1 Statistics Toolbox (MathWorks, Inc.).

Results

Heuristic Simulations

To demonstrate how the method of analysis can affect conclusions about social structure, we analyzed data from the same fission-fusion history (shown in Fig. 1a) using standard methods and the FDI. With 50 regularly spaced sampling events, the simple ratio index appears to show non-random population structure (Fig. 1b) even though the model separated individuals at random during each fission event. This apparent structure arises because groupings that lasted longer were sampled more times. Specifically, the main division in the population (the top node of the dendrogram in Fig. 1b) is defined by the second fission event because it lasts for the longest period of time. The next two nodes of the dendrogram (at linkage distance ~0.2) are defined by the sixth fission event, which produces the next longest-lived subgroups. The Monte Carlo randomization procedure indicates that the association indices that led to Fig 1b are highly unlikely to occur at random ($p < 0.0001$), even though individual decisions were simulated to be random. In contrast, the FDI analysis did not show any spurious non-random structure (Fig. 1c, $p = 0.71$).

The above results show that association patterns derived from the traditional simple ratio index (or other association indices based on proportion of time spent together) can be biased toward non-randomness, due to over-sampling of longer-lived fission subgroups. To assess the effects of study design on this bias, we next simulated the model for a range of sampling and fission rates and analyzed the results using the simple ratio index. For ratios of sampling rate to fission rate between 0.5 and 3, the reported non-random structure gains significance (*i.e.* the *p*-value decreases) as the sampling to fission ratio increases (Fig. 1d). Thus, if sampling events occur more frequently than fission events then the simple ratio index tends to be significantly non-random. The apparent non-random structure of the simple ratio indices is not influenced as much by the number of fission events observed as by the ratio of sampling to fission events (Fig. 1d).

Buffalo Association Patterns

Analysis using the simple ratio index indicated that the association patterns of buffalo in the Satara region were significantly non-random according to the permutation methods described by Bejder et al (1998, $p < 0.001$). UPGMA cluster analysis (Fig. 2) suggests four main buffalo groups, with adult males less tightly clustered both among themselves and to other sex and age categories. Cluster analyses of the FDI results did not show the same structuring of the sampled population into four groups as in figure 2 (data not shown).

For dyads involving females and juveniles, the FDI and the traditional association index were not closely correlated, indicating that the proportion of time

spent together may not be a good predictor of the probability of remaining in the same group during a fission event (Fig. 3). The mean FDI score for females and juveniles was significantly higher when adult male fission events were included (0.805 ± SE 0.005) than when they were excluded (0.603 ± SE 0.006; Fig 3). Further, the distribution of FDI values was further from random with the inclusion of adult male fission events compared to when only those events associated with females and juveniles were used ($\chi^2 =$ 2040 compared to $\chi^2 = 151$ for $df = 6$ and $p < 0.001$ for both; Fig 4a and b). The fission-decisions of adult male dyads were not significantly different from what would be expected given random decisions ($\chi^2 = 10.5$, $df = 6$, $p = 0.11$, Fig 4c).

Discussion

Fission-fusion societies lie on a spectrum. At one end of the spectrum individuals are free to move between groups at any time. At the other end, individuals only move between groups when subgroups separate from one group and join another, and individuals do not control the fission-fusion process. Traditional association indices and the fission-decision index apply best to opposite extremes of this spectrum, but in many situations they offer complementary information. Traditional association indices assume that the proportion of time a pair of individuals spends together indicates the strength of their association. In some species or ecosystems, however, individuals may be reluctant to switch between groups on their own and the proportion of time individuals spend together may reflect aspects of the group-level fission-fusion history rather than individual-level preferences. Our results demonstrate that traditional

association indices are poor descriptors of individual choices in such settings, and suggest our FDI as a more appropriate index to study individual choices.

Our simulations illustrate that, if sampling occurs faster than fission and fusion events, the proportion of time dyads spend together may show statistically significant clustering (Fig. 1b), even if individuals choose herds at random and independently of other individuals' decisions. This follows because multiple samples taken within the same inter-fission interval are autocorrelated with respect to individual choices. Further, the statistical significance of this effect is dependent upon the ratio of sampling and fission events rather than the absolute number of fission events. The FDI eliminates auto-correlated data and presents an unbiased estimate of individual choices. In this study we used the simple ratio index as the traditional metric of association. The potential biases shown in this study, however, apply to other association indices that are based upon the total number of samples taken (e.g. twice-weight, half-weight, simple ratio, square-root) rather than the number of fission events.

Traditional association analyses suggest that the buffalo population we studied was spatially and temporally structured into four different groups (Fig. 2). This result matches our intuition from collecting the field data, because (like the association indices) our intuition places greater weight on group compositions with longer lifetimes. Since it is based on association indices that incorporate all observations from the entire study period, the cluster analysis shows the hierarchical structure of the population integrated over time. On the other hand, an aerial survey representing a snapshot in time typically would show,, anywhere between 4-12 herds in our study area. As one would expect from previous research, adult males were less tightly clustered than other

sex and age groups because they often moved between mixed herds and bachelor groups (Fig. 2).

Not surprisingly, dyads that had high association indices (e.g. > 0.8) also had high FDI values because, in order for a pair to spend all of their time together, they would have to choose to remain in the same groups during fission events (Fig 3). For dyads with lower association indices, however, the probability of a dyad remaining in the same group during a fission event was not closely correlated with the amount of time they spent together (Fig. 3). Therefore, the non-random group structure apparent in the traditional association analysis (Fig. 2) is due in small part to non-random decisions made by individuals during all fission events (the weak effect in Fig. 3), but in greater part to the variable lifetimes of the resulting fission groups. Consideration of FDI scores has thus helped us to understand the mechanisms underlying results of traditional association analysis.

FDI scores reflected the qualitatively different fission-fusion behaviour of adult male buffalo versus females and juveniles. Mean FDI scores of dyads involving females and juveniles were significantly higher when fission events associated with adult males were included (Fig. 3 and Fig. 4a). This arises from the frequent occasions when a few adult males leave a mixed group (constituting a fission event)—all other pairs of individuals are counted as having stayed together, thus increasing their FDI. The distribution of FDI values of female and juvenile pairs was significantly different from what would be expected given random decisions (Fig. 4). This difference was magnified when adult male fission events were included in the analysis (Fig. 4a),

compared to when they were excluded (Fig 4b). Finally, the fission-decisions of adult males were not statistically different from random decisions (Fig. 4c).

While this distinct behaviour of females and juveniles versus adult males confirms and elaborates earlier results, some of our findings differ from the previous studies of African buffalo by Sinclair (1977), Mloszewski (1983), and Prins (1996). First, fission events seem to happen frequently in the KNP (185 female and juvenile fission events over a 2-year period), whereas in previous studies it is not clear how often herds were splitting apart, perhaps due to the smaller number of marked individuals in those studies. During this study we saw only 36 groups in the study area that did not have radio-collars (compared to 917 sightings of collared groups) due to the high density of marked individuals (~90 radio-collars in 4 to 12 groups). Secondly, both Mloszewski (1983), and Prins (1996) suggested that certain individuals always remain together in fission events due to either dominance and intra-herd competition (Prins 1996) or family-group structure within herds (Mloszewski 1983). In this study, we showed that although there were some non-random patterns in the FDI of female and juvenile pairs (Fig 4b) the pattern is not as strong as what might be expected from previous studies. Further work on how fission-decisions may be affected by body condition, reproductive status, and genetic relatedness would be enlightening.

Application of the FDI

The fission-decision index may not be applicable to all studies. The FDI is best applied when individuals choose between subgroups only during fission events. This may be reasonable for species that incur high dispersal costs, perhaps due to high

predation rates or lowered foraging efficiency, but if individuals are highly mobile and move often between subgroups then traditional indices reflecting the proportion of time dyads spend together are reasonable measures of association. The study of Szykman et al. (2001) on hyenas is an example where individuals may be relatively unconstrained in their choice of subgroups: as top predators, the cost of dispersing short distances between subgroups within a clan may be minimal for hyenas. The necessity of observing multiple fission events per dyad may also limit application of the FDI method. Finally, the lifetime of fission subgroups may reflect the degree of satisfaction with that group composition. In this case, our initial assumption that group fission and fusion dynamics are not controlled by individuals no longer applies, and thus the greater weight that is assigned to longer-lived groupings in traditional association indices may be more appropriate.

The FDI requires rich data regarding the underlying fission-fusion process, and the probability of detecting fission events is related to the proportion of individuals that are marked. Studies with fewer marked individuals will have downwardly biased FDI values because they are more likely to miss fission events when all the marked animals stay together, which would increase their FDI. This presents difficulties if individuals choose differently in short-duration fission events, as these are most likely to be missed. It also presents difficulties when comparing across studies. However, if a researcher has a random sample of fission events, then the FDI should represent an unbiased estimate of the probability that two individuals will remain together during a fission event. Further, one could structure the data according to the duration a splinter group existed to investigate whether the FDI was more random during short-term splits (which may

be unintentional and due to predation and/or lack of communication) than long-term splits (which may be due to intra-group competition). Previous studies may not have had the data resolution necessary for our FDI approach. We believe, however, that this technique will become increasingly valuable as improved technology facilitates the tracking of more individuals with greater spatial and temporal resolution.

Analysis of an aggregate parameter (e.g. proportion of time spent together by a pair) that is the product of two underlying processes limits our ability to understand those underlying processes. We believe that separating the process of individual choice from the process of group fission or fusion leads to an improved understanding of the mechanisms of association. In addition, this separation helps avoid apparently arbitrary definitions of classes of companionship based upon the amount of time individuals spend together (e.g. Weinrich 1991, Whitehead et al. 1991). Analyses using traditional association indices implicitly combine the effects of the fission-fusion process and the choices made by individuals, and although their conclusions about the proportion of time individuals spend with one another remain valid, they may not accurately reflect individual preferences.

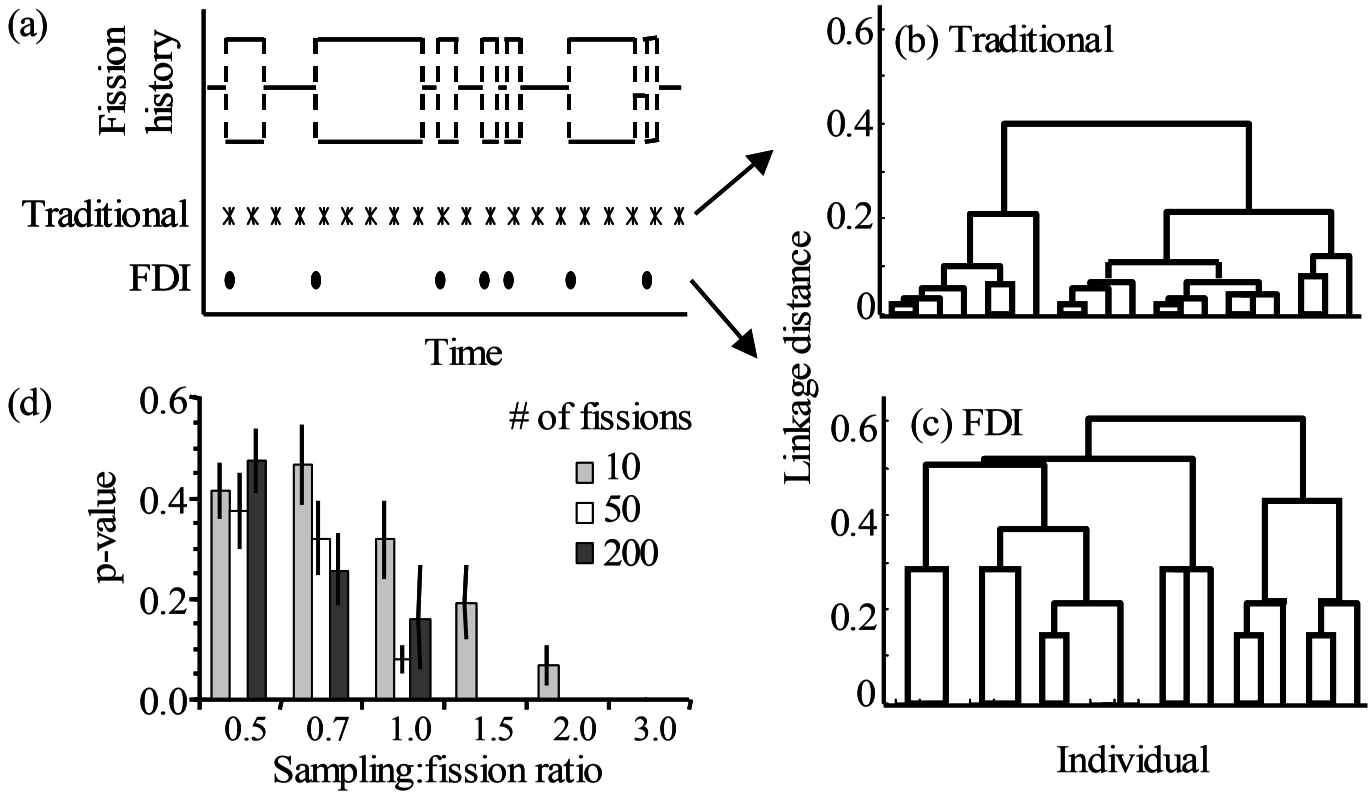


Figure 1(a-d). We used a heuristic model of random individual choices to simulate different fission-fusion histories and sampling regimes. The fission-fusion history shown in Fig. 1a depicts one group separating into two groups seven times, subject to traditional sampling occurring at regular intervals, and to FDI sampling after fission events. Simulated data were used to calculate the UPGMA dendrogram for the traditional simple ratio association index (b, $p < 0.001$) and FDI (c, $p = 0.71$). The statistical significance of the traditional simple ratio index increases with sampling intensity regardless of the number of fission events. For Fig. 1d, each column represents 20 simulations of the heuristic model with a fixed number of fission and sampling events, and vertical lines represent the standard deviation of the *p*-value

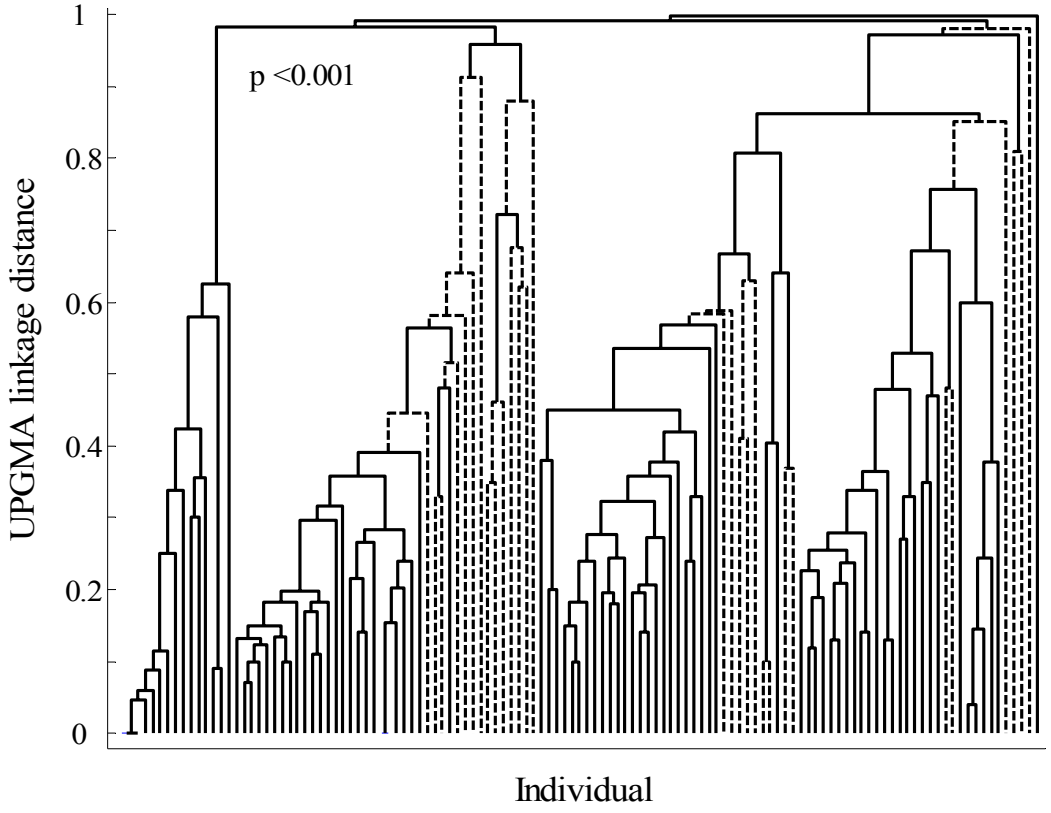


Figure 2. UPGMA cluster diagram of the traditional simple ratio index for 123 radio-collared buffalo using data from January 2002 through October 2003. Adult males are represented by dotted lines. Statistical significance was determined using the randomization procedure described in the methods.

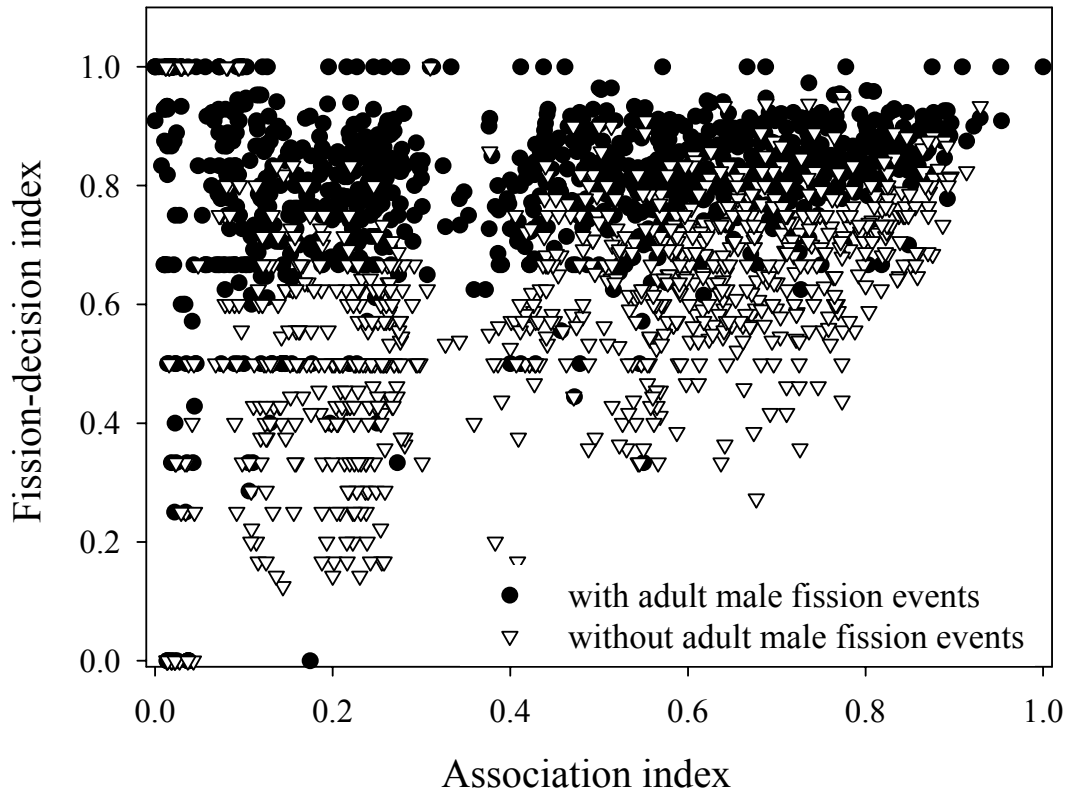


Figure 3. The relationship between the traditional simple ratio association index and the FDI for all pairs of radio-collared buffalo that had two or more fission events together and were not adult males. Filled circles indicate indices calculated from data that included fission events involving only adult males (1093 pairs), while open triangles did not include adult male fission events (834 pairs).

.

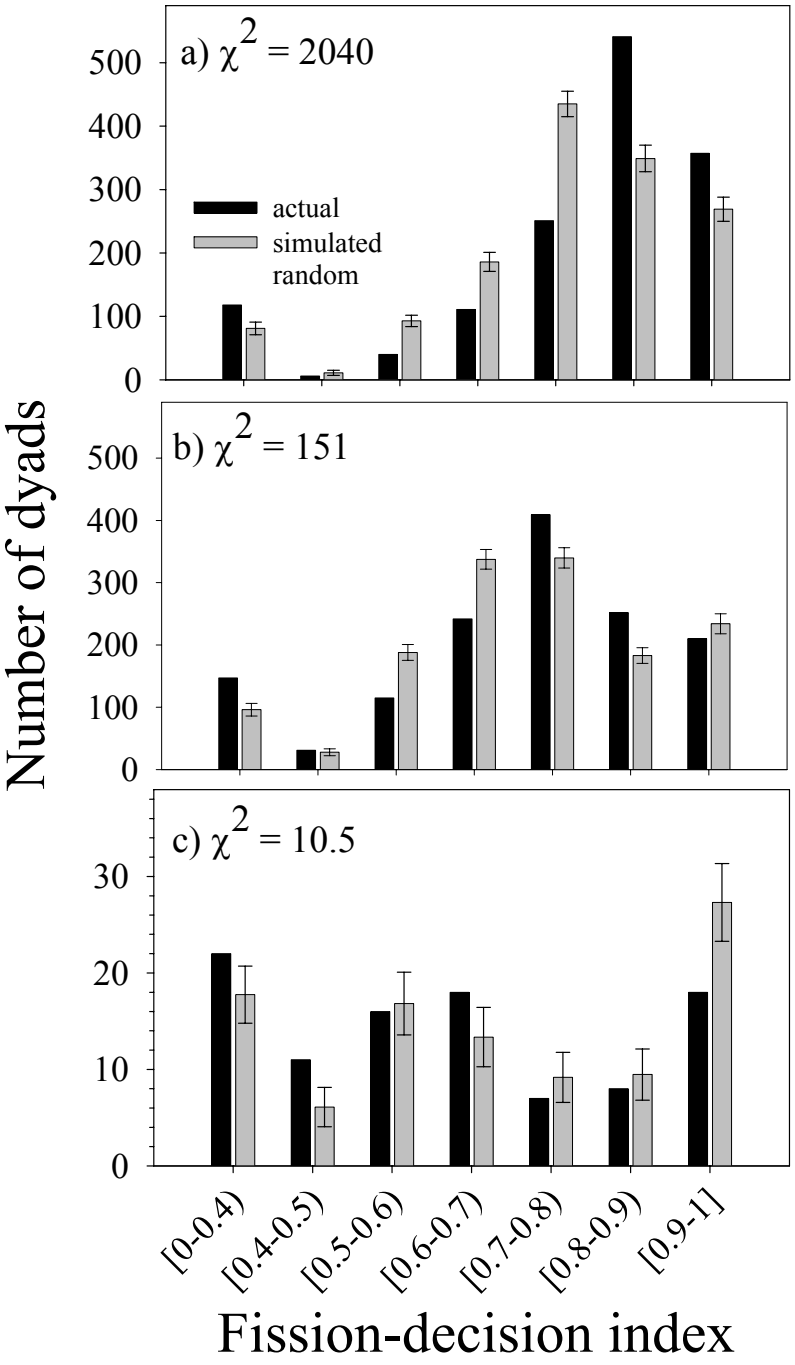


Figure 4. FDI values calculated from the buffalo data and from simulations assuming random fission-decisions, using (a) female and juvenile dyads with all fission events (b) female and juvenile dyads without adult male fission events and (c) dyads involving adult males only. Bars represent the standard deviations from 1000 simulations of random decisions.

Chapter Four

Integrating association data and disease dynamics in a social ungulate: bovine tuberculosis in African buffalo in the Kruger National Park

P.C. Cross, J.O. Lloyd-Smith, J.A. Bowers, C.T. Hay,

M. Hofmeyr and W.M. Getz[1,2]

Introduction

Group structure is the hallmark of social species. The size and integrity of groups reflects their function, which may include vigilance against predators, sequestration and protection of resources, and alloparenting (for review see Dugatkin 1997). Increased susceptibility to disease is generally believed to be a cost of sociality (Alcock 1998). If the movement between groups is limited, however, group structure can act to contain disease spread. Here we investigate how the movement of individuals between groups affects disease spread using data on African buffalo (*Syncerus caffer*) and relatively simple models of disease processes.

African buffalo exist in a fission-fusion society where groups often separate and rejoin (Prins 1996, Cross et al. In press). The ability to recognize others is a prerequisite for non-random association patterns in fission-fusion societies (Cross et al. In press). In this context we do not differentiate between the evaluator and cue-bearer in the recognition process, but analyze the proportion of time individuals spend together and assume that it is a function of individuals perceiving and acting upon the cues expressed by one another. We use pair-wise association data that we have collected over a two-year period in the Kruger National Park, South Africa, to characterize buffalo population structure. These data are then combined with disease models to illustrate how association patterns affect disease dynamics. Behavioral researchers often collect data on the association between individuals, but these data are seldom used in disease models. We illustrate the importance of incorporating behavioral data to disease dynamics, and discuss several of the unresolved questions which limit our ability to integrate data on recognition and association with disease models.

Models of disease dynamics have become increasingly important to understanding and managing disease invasion (e.g. Ferguson *et al.* 2001, Keeling *et al.* 2003, Lloyd-Smith *et al.* 2003). They reduce the complexity of a system and allow for the investigation of specific factors in ways that would not be possible using experimental methods. A tradeoff exists between realism and generality of models, however, and factors that are omitted for the sake of simplicity may play important roles in the real system. Traditionally, the network of connections between individuals has not been included in disease models (e.g. Anderson and May 1991). Modelers assumed that association was random, i.e. that every individual was equally likely to contact every other individual. More recent studies suggest, however, that the way individuals contact one another (i.e. the network structure and topology) plays an important role in determining the probability of disease invasion, the total number infected, and the speed of disease spread (Keeling 1999, Watts 1999, Newman 2002). Both traditional disease models that assume random mixing and spatial disease models that assume limited dispersal between fixed groups poorly characterize some socially structured populations, such as the African buffalo. Dynamic network models more accurately reflect connections within and between groups and the spread of disease between associating individuals. In this paper, we attempt to narrow the gap between the fields of behavior, recognition, and disease ecology by illustrating the importance of non-random association to the spread of disease using empirical data from an ongoing study of African buffalo and integrating them with simulation models of disease spread.

Social networks have been visualized in a number of different ways. Network graphs depict individuals as points and their contacts as connecting lines (Fig. 1). In the

behavioral literature of animal association patterns, researchers often use cluster analyses, such as the Ward's or UPGMA methods, to describe the network (e.g. Whitehead 1999). Underlying these visual techniques is a contact or association matrix **A** where each matrix element a_{ij} describes the amount or type of contact between individuals *i* and *j*. Several properties of this matrix will affect the overall rate of disease spread as well as each individual's risk of infection. Most obvious are the average number of connections per individual and the strength of those connections. Less obvious, and the focus of recent work, is the way that the topology of connections (*i.e.* who is connected to whom) affects disease processes (e.g. Anderson et al. 1990, Kretzschmar and Morris 1996, Morris and Kretzschmar 1997, Boots and Sasaki 1999, Keeling 1999, Watts 1999, Kretzschmar 2000, Newman 2002, Eames and Keeling 2003, Meyers et al. 2003, Read and Keeling 2003).

More traditional disease models assume that an individual's risk of infection depends upon the global state of the population. From a network perspective, however, an individual's risk of infection depends on the number of connections they have and which of those connected individuals are infected. Network models of disease have largely focused on sexually transmitted diseases (STDs), often simulated on static networks with uniform connection strengths (e.g. Kretzschmar and Morris 1996, Keeling 1999, Watts 1999), though recent work has included analytic approximations (Ferguson and Garnett 2000, Bauch and Rand 2001, Eames and Keeling 2002), exact solutions (Newman 2002), non-sexually transmitted diseases (Meyers *et al.* 2003), an STD network with changing connections (Eames and Keeling In press), and an empirically-derived urban social network (Eubank et al. 2004). However, our extension

of this network approach to cover airborne diseases in animal systems, with a network based on empirical data, raises two novel issues. Is variance in the connection strengths an important factor? How does the dynamic nature of the network affect disease spread?

Early attempts to model stochastic disease dynamics in networks were probably limited by computational power. Current efforts are impeded by scarcity of empirical data on the structure of human and animal networks and the limited communication between behavioral researchers and epidemiologists. In practice, several critical issues make it difficult to determine the network structure of a population: (1) contacts may be difficult to define and differ between diseases, (2) people are usually bad at estimating their contacts while animals are often difficult to observe, (3) large populations require researchers to choose a sample of individuals or a small portion of the network and then extrapolate, (4) connections between individuals change over time, and (5) on longer time scales individuals enter and leave the network due to birth and death processes or migration. All of these issues affect our ability to produce an unbiased estimate of the network structure and none of them are easily solved. For this reason, relatively few empirically-based social networks exist in the disease ecology literature (but see, Woodhouse et al. 1994, Edmunds et al. 1997, Wallinga et al. 1999, Liljeros et al. 2001, Jolly and Wylie 2002). In this study, we address issues two and four above, in the context of data that we have collected on associations among individuals in an African buffalo (*Syncerus caffer*) population in Kruger National Park that is in the midst of a bovine tuberculosis (BTB) epidemic (Rodwell et al. 2000). More importantly, though,

we demonstrate that association data collected routinely by behavioral researchers are applicable and valuable to studies of disease dynamics.

African buffalo typically occur in breeding herds of approximately 30 to 1000 individuals, and adult males move between breeding herds in bachelor groups of 2--30 individuals. Previous researchers concluded that buffalo herds were relatively stable units, and although herds sometimes separated they did not associate with neighboring groups (Sinclair 1977, Mloszewski 1983, Prins 1996) . Furthermore, females, subadults, and juveniles were assumed not to move between herds (Sinclair 1977, Mloszewski 1983, Prins 1996) . Recent studies based upon larger sample sizes of radio-collared individuals, however, suggest that females and subadults do move between groups and the structure of herds may not be stable (Halley et al. 2002, Cross et al. In press). Radio-tracking data from our study in the KNP of over 123 radio-collared individuals since November 2000 indicate that herds frequently separate and reunite, and females move to different areas via splinter groups. Further, because herd membership changes, the definition of a herd becomes more nebulous over time. The framework developed in this study presents a rigorous definition of a herd based upon association data.

In this study, we begin by describing the association pattern and network structure of 64 radio-tracked buffalo from November 2001 to October 2003. We then use the association data in a stochastic disease model to investigate three questions: (1) How does the incorporation of non-random association data affect our predictions about the speed and intensity of a disease outbreak in the buffalo population in Kruger National Park? (2) Does the variance in the frequency of contact between individuals

affect disease dynamics?, and (3) How does duration of infectiousness affect the degree of population structure experienced by the disease process? At the moment, we do not have empirical disease data to compare with the model predictions, but work on this aspect is ongoing. To the authors' knowledge, this study is the first application of a network disease model to a wildlife population using empirical data to create the social network.

Methods

Association data

Field data were collected during an ongoing study in the Satara Region of the Kruger National Park. The study area contained 4--12 buffalo herds, depending upon the amount of herd fragmentation, and roughly 3000 buffalo. Buffalo were darted from helicopters and fitted with radio-collars in four sessions: November 2000 ($N = 6$), April 2001 ($N = 27$), August ($N = 51$) and November 2001 ($N = 12$). To simplify the analysis, we restricted the data to sightings of 64 radio-collared buffalo that survived from November 2001 to October 2003. This restriction allowed for a 'complete' dataset where individuals were present for the duration of the study period.

We monitored buffalo herds, on foot and from vehicles, approximately 2--3 times per week throughout the year from distances ranging from 50--1000m. If an individual was missing for over one month, we located it from an aircraft. If a herd split during the day, only the first sighting was used for that day. Since all marked individuals had radio-collars and herds were usually separated by several kilometers, we

could determine which individuals were in a herd without visually sighting all individuals.

Following Whitehead and DuFault (1999), we considered two individuals to be associating if they were located in the same herd. This one/zero metric of association was used to calculate the proportion of samples in which two individuals were seen together (*i.e.* the simple ratio index). Ideally, when using association data in disease models the distance cutoff used to determine whether two individuals are in the same group should depend on how infectivity decreases with increasing distance. We assumed that buffalo were associating when they were in the same herd and the probability of infection between a particular susceptible-infected dyad was proportional to the time they spent in the same herd. This definition implicitly assumes that herds are sufficiently well mixed that within-herd transmission is equal between all dyads and sufficiently separated that between-herd transmission is nonexistent. Non-random association, however, may also play a role within herds when herds are large and diseases are transmitted only over very short distances. Unpublished data suggest that buffalo frequently move between the front, middle, and back portions of the herd, but it remains to be determined whether this is sufficient for our assumption of a well-mixed herd to be valid. Recognition and its role in determining association patterns may be important both within and between herds depending upon the distance over which individuals are infectious. In this study, however, we investigate the role of different association patterns at the herd-level only.

We constructed association indices between all possible dyads using the simple-ratio association index calculated from varying windows of time. Except where noted

otherwise, association indices were calculated using monthly data or all the data from November 2001 through October 2003. Over the 24-month study period, there were 16 occasions when a pair of individuals was not seen during a month. In these cases, we assumed that the individuals were not associating during this time period. Since this is less than 0.04% of the association indices used, this assumption probably had a negligible effect on our numerical simulations. When yearly estimates for 2002 and 2003 were calculated they were based upon data from November 2001 through October 2002 and November 2002 through October 2003, respectively, to coincide with the onset of the wet season.

We analyzed the association data in three ways, independent of the disease model. First we constructed dendrograms of the association matrices using the unpaired group averaging method (UPGMA), which had the highest cophenetic correlation coefficient compared with dendrograms that we constructed using Ward's weighted, complete linkage, and single linkage clustering methods (Romesburg 1984). Second, we calculated a measure of the network clustering using the following metric, proposed by Keeling (1999):

> From the association matrix the total number of connected triples (three individuals sharing at least two connections) is calculated. Some proportion of these triples will be closed loops, or triangles, where all three individuals are connected.

Keeling's clustering coefficient, ϕ, is the ratio of triangles to triples in the network. When ϕ equals one then all triples form closed loops and one's neighbor's neighbor is always one's own neighbor as well. When ϕ is small individuals have few associates in

common. We calculated ϕ by first converting the association matrix to a matrix of zeros and ones where all non-zero associations were changed to one. Finally, we calculated the percentage of connections between individuals that changed from one month to the next, where a connection was defined as any association index greater than zero. All cluster analyses were coded in MATLAB 6.1 (MathWorks, Inc.).

Disease Modelling

The model presented here is a stochastic individually-based elaboration of a discrete time Susceptible-Infected-Recovered (SIR) epidemic model (Anderson and May 1991, Kermack and McKendrick 1991), with time-steps of one month. A variable $x_i(t)$ is used to represent the state of an individual *i* at time *t* and equals zero when the individual is susceptible and one when the individual is infected. We assume that the probability for the disease to be transmitted between an infected individual *i* and a susceptible individual *j* is a function of the transmission coefficient β and the association coefficient $a_{ij}(t)$, where $a_{ij}(t)$ equals the proportion of time individuals *i* and *j* spent together over the period $[t-1,t]$. Additionally, we assume that infected individuals recover with constant probability γ per time-step and do not become susceptible again. Under these assumptions, our model takes the form:

$$\mathrm{Prob}\{x_i(t+1)=1 \mid x_i(t)=0\} = 1-\exp\left(-\beta\sum_{j=1}^{n} a_{ij}(t)x_j(t)\right) \quad t=0,1,2,\ldots,T-1 \quad (1),$$

$$\mathrm{Prob}\{x_i(t+1)=0 \mid x_i(t)=1\} = \gamma \quad (2),$$

where *T* is the length of the simulation in months and *n* is the total number of individuals in the population. In all simulations, we calculated $a_{ij}(t)$ values using the radio-tracking association data from the 64 buffalo for which we had a complete dataset

during the study period. We started each simulation with one infected individual. Qualitative results were insensitive to the choice of this individual, except for a few individuals who were isolated from the others during the initial phase of the study period. In most cases, we simulated the model for 24 months to match the amount of radio-tracking data that was available. For those cases where we simulated the model for longer periods of time, we used the same association data repeatedly. We compared the results of this model to a mean-field equivalent to highlight the importance of incorporating the association data. The mean-field model assumed that all individuals were associated according to the grand mean of the association matrix for that month, so the overall force of infection for the mean-field model equaled that of the models using association data.

Using the buffalo association data and SIR model described above we investigated the impact on disease dynamics of two aspects of association patterns: topology (i.e. who is connected to whom) and variation in the frequency of connections. Further, we investigated how the effects of association patterns depended on disease characteristics by varying the transmission coefficient β and probability of recovery γ to simulate "fast" versus "slow" diseases, while holding the basic reproductive number (R_0) constant. The basic reproductive number is the expected number of infections caused by the first case in a completely susceptible population, and is a measure of the growth potential of an epidemic (Anderson and May 1991).

We investigated the importance of topology by randomly rewiring connections in both static and dynamic network simulations, using the following algorithm. Let δ represent the proportion of network connections that were randomly re-assigned, such

that δ=0 corresponds to the original network structure and δ=1 corresponds to a network with all connections randomly reassigned. Both the static and dynamic simulations started with the association matrix from November 2001. In the static network simulations, a proportion δ of the connections were rewired at the beginning of the simulation, then associations between individuals were assumed to be constant over time. In the dynamic network simulations, a proportion δ of the connections were rewired before every time-step.

Next, we investigated the importance of variation in connection frequency or strength. Previous research on network disease models typically assumed that the strengths of connections between individuals are equal (i.e. an 'unweighted' network) and thus the variance of connection strengths is zero. The variance in the amount of time spent together, however, may be biologically meaningful and have important consequences for disease dynamics. In dynamic networks, connection strength may vary within a dyad over time or among dyads. Assuming that individuals contact one another with a constant probability that is related to their time-averaged association index $\overline{a}_{ij}$, then the variance in contact frequency within the dyad is related to a binomial probability density function and decreases as $\overline{a}_{ij}$ approaches zero or one. We investigate the effects of increasing the variance of the time-averaged association indices $\overline{a}_{ij}$ among dyads, which results in lower variance within dyads. High variance among dyads would be analogous to some individuals having two sets of associates, those that they spend time with often and those that they do not. One might hypothesize that weak connections are less significant for disease transmission, and hence that systems with a high variance in connection strength may be less permeable to disease spread. On the

other hand, the disease may spread more rapidly amongst those individuals that are tightly associated.

Let $\mathbf{A}_t$ represent the matrix of association coefficients a_{ij} that is based on data from the period $[t-1, t]$ and $\mathbf{A}_{\text{avg}}$ equal the time-average of $\mathbf{A}_t$ over the study period. Elements of $\mathbf{A}_{\text{avg}}$ between (but not including) zero and one indicate that a pair of individuals spent only a portion of the study period together. We introduced a parameter α to represent the relative amount to increase or decrease the time-averaged connection strength for each pair of individuals. For each matrix element $\bar{a}_{ij}$ in $\mathbf{A}_{\text{avg}}$, we drew a random variable z from a uniform distribution between zero and one, and calculated a new element $\bar{a}'_{ij}$ as:

$$\bar{a}'_{ij} = \begin{cases} \bar{a}_{ij} + \alpha\left(1 - \bar{a}_{ij}\right) & IF \quad \bar{a}_{ij} > z, \\ \bar{a}_{ij} - \alpha\left(\bar{a}_{ij}\right) & OTHERWISE. \end{cases} \quad (3)$$

Thus, we randomly increased or decreased each element of $\mathbf{A}_{\text{avg}}$ some proportion α of the distance between $\bar{a}_{ij}$ and zero or one. Specifically, when α=0 the time-averaged connection strength has its original value, and when α=1 all connections in the time-averaged association matrix are either zero or one. Using this algorithm we could increase the variance in time-averaged connection strengths while maintaining the expected mean connection strength. We also preserved the topology of connections from the original matrix $\mathbf{A}_{\text{avg}}$, except when α equals one and some connections were removed entirely. We applied this algorithm once to $\mathbf{A}_{\text{avg}}$ at the beginning of each simulation to create $\mathbf{A}'_{\text{avg}}$. Then for each timestep of the simulation we created an

association matrix $\mathbf{A}_t'$ of zeros or ones, using $\bar{a}_{ij}'$ in the altered association matrix $\mathbf{A}'_{\text{avg}}$ as the probability that each connection exists.

Thus we reconstructed an association history for the population in which pairs of individuals were assumed to be completely associated or isolated within time-steps, but the time-averaged association strength for each pair was determined by $\mathbf{A}'_{\text{avg}}$. High values of α corresponded to increased variance among pairs in time-averaged connection strength, but decreased variability in the existence of particular connections through time. Low values of α, which yielded more intermediate values of $\bar{a}_{ij}'$ in the altered association matrix $\mathbf{A}'_{\text{avg}}$, corresponded to lower variation between pairs but higher temporal variation for each pair. Note that this approach is distinct from altering the variation in association indices within each time-step. To explore the role of variance in connection strength, we experimented with different $\mathbf{A}_{\text{avg}}$ matrices (e.g. using all data or just the data from particular months). All simulations were conducted in MATLAB 6.1 (MathWorks, Inc.).

Results

Association data

The time frame used to calculate association indices has a large effect on the apparent structure of the system and thus conclusions about the ability of a disease to spread through that system. We visualize the network structure using 3-D network graphs and dendrograms (Figs. 1 and 2). The network graph of May 2002 (Fig. 1A) indicates two distinct groups and two outlying individuals all of whom become well connected when considering contacts over the whole period from November 2001 to October 2003 (Fig.

1B). In general, association networks become more connected over longer time frames due to the movement of individuals between groups, thereby allowing diseases to spread among groups. Dendrograms illustrate the hierarchy of associations in the buffalo network (Fig. 2). Individuals that are more often together, similar to highly related species on a phylogeny, are joined at lower linkage distances on the dendrogram. As in the network graphs, dendrograms based upon a month of data were generally more tightly clustered than those based upon a year of data (Fig. 2).

Buffalo appeared more highly clustered in 2002 than 2003 (cf. Fig. 2B and C). In 2002, three groups are apparent, but two of these appeared to have merged in 2003. Further, in 2003 a lower proportion of nodes (where two individuals/groups are joined in the dendrogram) were included for a given linkage distance compared to 2002, indicating looser associations between individuals and groups (Fig. 2D). Also, considerable variation exists between months, with May 2002 and January 2003 respectively representing the low and high connectivity extremes of monthly association data (Fig. 2A, D). Although the network graph for May 2002 (Fig. 1A) shows only two groups, the dendrogram, which accounts for the relative strength of associations, appears to show three groups (Fig. 1A). Over the entire study period the frequency of mixing events between herds resulted in a well-connected network (Fig. 1B).

Simulation Results

To show the effects of the clustering patterns in Fig. 2, we simulated SIR disease dynamics ($\beta = 0.3$, $\gamma = 0.1$) using monthly association matrices from either the entire dataset, one year of data, or a mean-field model (Fig. 3). The mean-field model where

all individuals were connected predicts much faster spread of disease even though the force of infection was the same as the model using all the association data (cf. closed circles to open circles; Fig. 3). When we used only November 2001 through October 2002 (i.e. year 2002), then repeated the same values to simulate months 13 through 24, the disease was limited to only one herd and did not infect as many individuals compared to using data from 2003 (Fig. 3). Thus, the tighter clustering of buffalo in 2002 compared to 2003 (Fig. 2) translated into a population that is less permeable to disease invasion. When we used the monthly data for the whole period, there was a second pulse of infections starting around month 14, presumably due to the disease moving into a new herd (Fig. 3). Coincident with the second pulse of infections was a wet season with below-average rainfall (November 2002 to February 2003). This also coincided with a marked decrease in the clustering of the association data in November 2002 as indicated by Keeling's clustering coefficient ϕ (Fig. 4).

Using all the monthly association data we simulated two SIR-type diseases with the same basic reproductive number ($R_0 = \beta/\gamma$) but different infectious periods. A faster-moving disease, with high infectiousness and rapid recovery ($\beta = 0.4$, $\gamma = 0.3$), is more likely to fade out in populations where the frequency of movement between groups is low, because it may burn out in the local population before sufficient connections to other groups are made (Fig. 5). A slower-moving disease with the same R_0 ($\beta = 0.04$, $\gamma = 0.03$) is less infectious but persists longer, increasing the probability of transmission to other groups of individuals(Fig. 5). Thus, slower diseases effectively integrate over a longer time period and the network becomes more fully connected (Fig. 1).

Next, to assess the impact of network structure on the spread of disease, we analyzed the system using only particular months of association data and either randomly rearranging network connections or increasing the variance in connection strengths. Random rearrangement (or "rewiring") of the network connections involves establishing a new contact between two randomly chosen individuals and removing a contact between two other individuals. Random rewiring had a non-linear effect upon disease dynamics (Fig. 6). Only a small amount of rewiring ($\delta < 0.2$) creates a network that behaves like a randomly wired network ($\delta = 1.0$). Furthermore, because only a small amount of rewiring had a large impact the static simulations (i.e. rewiring the matrix once at the beginning of the simulation) yielded similar results to the dynamic simulations (i.e. cumulative random rewiring of the matrix over time), though for very small δ the dynamic rewiring showed greater effects as expected (Fig. 6). Other output variables, such as the number of susceptible individuals remaining at the end of the simulation, showed results similar to those in Fig. 6.

To place these random rewiring simulations into context, we plot the percentage of the topology that remains the same over increasing time lags for the monthly empirical data and dynamic rewiring simulations (Fig. 7). The % similar topology of the empirical data appears to be similar to randomly rewiring 10% of the connections (*i.e.* $\delta \approx 0.1$) in the November 2001 data (i.e. about 10% of the buffalo network changed per month; Fig. 7). However, for reasons we discuss later, the disease dynamics of simulations using $\delta = 0.1$ differ greatly from those based on the actual association data, which yield results more similar to runs with $\delta \leq 0.01$ runs (Fig. 6).

The variance in the connection strength among dyads was only important under certain circumstances. Using data from September 2003 and a fast SIR disease ($\beta = 0.3$, $\gamma = 0.1$) as an example, Fig. 8 illustrates how the total number of individuals infected over the entire epidemic was a function of α (a proxy for the variance in connection strength; see Methods) and the connectivity of the network. We conducted the same analysis as in Fig. 8 for other disease characteristics and found that α had less effect upon disease dynamics for slower diseases (*e.g.* $\beta = 0.03$, $\gamma = 0.01$). Further, α had little effect when we used data from months where buffalo herds were either very well connected or very weakly connected (data not shown). Finally, dropping a certain proportion of network connections had a much larger impact than increasing α (Fig. 8).

Discussion

Early work on disease modeling assumed instantaneous random mixing between all individuals. More recently researchers have begun to account for non-random host mixing patterns, often using a static network (e.g. Watts 1999, Newman 2003) although one recent study incorporated changing connections within a background network that is static (Eames and Keeling 2004). Pair-formation models of sexually transmitted diseases include dynamic contacts but usually not the fully non-random mixing of network structure (Dietz and Hadeler 1988, Kretzschmar and Morris 1996, Morris and Kretzschmar 1997, Lloyd-Smith et al. 2004). To the authors' knowledge, empirically-derived, fully dynamic contact networks have not been used as a substrate for investigating disease dynamics for STDs or non-sexually transmitted diseases. Our analysis shows how disease dynamics depend on the topology of connections between

individuals, the dynamic nature of these connections, and the variance in the frequency of contacts between individuals. Association data similar to those presented here are often collected by behavioral researchers (e.g. Myers 1983, Smolker et al. 1992, Brager et al. 1994, Whitehead 1999, Szykman et al. 2001), but have not been combined with models of disease dynamics even though they may have important impacts upon the spread of disease (Fig. 3).

Analysis of the association data suggests that the buffalo population was more tightly clustered in 2002, but with more connections between groups in 2003 (Fig. 2). From the perspective of a disease, this translates into a more permeable population in 2003 (Fig. 3). More specifically, many buffalo were moving to new areas and groups in November 2002, which resulted in a less tightly clustered population at the same time when rainfall was well below average (Fig. 4). Rainfall totals for November 2002 through January 2003 were only 34% of the long-term average for those months. Results from the model then present a testable prediction that dry conditions facilitate more rapid spatial spread of disease in the buffalo population due to increased herd-switching.

This presents a worrying scenario for southern Africa, where precipitation is predicted to remain stable or decline while temperatures are likely to increase due to global climate change (Hulme *et al.* 2001). As a result, vegetation conditions are likely to decline, which may result in an increase in the amount of population mixing as wildlife is forced out of previously habitable areas. Many studies have linked climate change with altered disease distributions or dynamics (e.g. Epstein 2002, Patz and Khaliq

2002). However, evidence for the impact of climate change on animal behavior, which in turn affects disease dynamics, is rare.

Previous studies of African buffalo have suggested that herds are relatively static (Sinclair 1977, Prins 1996). We show, however, that buffalo herds are very dynamic in the KNP and that over time the population becomes well-connected (Fig. 1). As a result, over time it becomes difficult to define what is a herd. As suggested by a reviewer, we propose a definition of a herd based upon the association data, such that a herd is defined as the set of maximal complete subgraphs of the network where all individuals, or vertices, are associated with a_{ij} greater than some threshold (for a particular time period). Here a complete subgraph is defined as a subset of the vertices of the network such that for every pair of vertices in the subset, there is an edge connecting them and the set of those edges and vertices are a complete network and a subset of the original network. We will explore the implications of this definition for our buffalo data in a future publication.

In considering effects of association patterns on disease, one issue of critical importance is accounting for the relative time scales of the disease and host mixing patterns. This interplay has important implications for the spread of disease, the way association data should be collected, and the evolution of both host and parasite. Fast diseases, such as rabies, measles, and Ebola hemorrhagic fever, with a short duration of infectivity encounter a more structured population because the amount of mixing between groups decreases as the time frame decreases (Fig. 2D, 5). Thus, association data should have higher time resolution if they are intended to give insight into the spread of diseases with fast dynamics. More specifically, association data should be

collected at least as frequently as the duration of infection for the disease of interest, to capture population mixing on timescales relevant to the disease. Furthermore, association data should not be averaged over too long a period. If association matrices are constructed using data from a time frame longer than the duration of infection within an individual, then the network is biased in favor of too many connections between individuals. As an extreme example compare the connectivity of the networks based upon one or 24 months of data (Fig. 1).

Similar to Watts's work on static networks (1999) we found that only a small amount of random rewiring is necessary to make the empirical buffalo network behave as a randomly wired network. Further, there was little difference between rewiring the network once at the beginning of the simulation or once each time-step (Fig. 6). This is probably due to limited number of groups in the network. Only a small number of connections are necessary to make it a 'small-world' graph where there are very few degrees of separation between any two individuals and additional changes to the network have little additional impact (Fig. 6). Interestingly, the model predicts that fewer individuals would be infected when using empirical association data compared to simulated data that approximates the amount of change in the association matrix ($\delta \approx 0.1$) seen in the empirical dataset (Fig. 6). This suggests that although the empirical network is changing every time step (Fig. 4), these changes are not random and a proportion of the population remains inaccessible to the disease (Fig. 6B). Thus the population is less permeable to disease invasion than one would expect if movement of individuals between groups was random. This amount of movement is higher than what might have been expected from previous studies of African buffalo (Sinclair 1977, Prins

1996), however, given the intensive radio-tracking conducted in this study probably more accurately describes the fluid structure of the population in the KNP (Cross et al. In press).

Somewhat surprisingly, we found that increasing the variance in connection strength among dyads had only minor effects upon disease dynamics in this system (Fig. 8). This suggests that for the range of networks we investigated, disease dynamics are very similar in a system where individuals spend their time equally with all associates or spend a large portion of their time with only a few individuals and a little time with many others. This conclusion, however, may be limited to well-connected networks with a limited number of groups. Further, we did not investigate the effects of increasing the variation in contact frequency within a pair of individuals. More work is necessary to determine the generality of these results and in what contexts networks of weighted connections behave differently than unweighted networks.

Future directions

Despite the known importance of association patterns to disease dynamics, empirical data are lacking, especially for airborne diseases. First of all, disease ecology is a relatively new field, and although a number of studies have investigated how association patterns may affect disease dynamics (e.g. Keeling 1999, Eames and Keeling 2003, Boots et al. 2004), few studies have attempted to integrate empirical data and disease models. Secondly, an accurate depiction of a social network is very difficult to generate. We believe that future research should focus on the following questions, which currently limit our ability to apply empirical data to disease models:

(1) What defines a contact for airborne diseases? (2) What are the appropriate time and spatial scales to sample a network of animals? (3) How does one scale-up a sample of a network to represent an entire population?, and (4) Given that population dynamics are an important factor in disease dynamics, how does one allow for births, deaths and changes of association patterns while maintaining the overall properties of a network? Studies on the tradeoffs faced by individuals in social systems with regard to disease and behavior could also be enlightening (e.g. Adamo et al. 2001, Boots and Knell 2002). In particular, individuals may modify their behavior to decrease their contact with others when they are at higher disease risk. If empirical data were available, they could be incorporated into the network model framework by adjusting the association indices between infectious and susceptible individuals.

In conclusion, the way that individuals associate with one another has a large effect upon the spread and dynamics of a disease. Although there are a few questions that need to be answered before empirically-derived social networks can be widely applied, the methods presented here provide a flexible framework for combining behavioral data with models of disease dynamics. Finally, our results suggest that a critical interplay exists between the time scales over which social interactions take place and those associated with a particular disease. An assessment of how likely it is that a given disease become an epidemic requires that we pay attention to social interaction and disease time scales, which then determines the appropriate temporal resolution for the collection of data determining the structure of social networks.

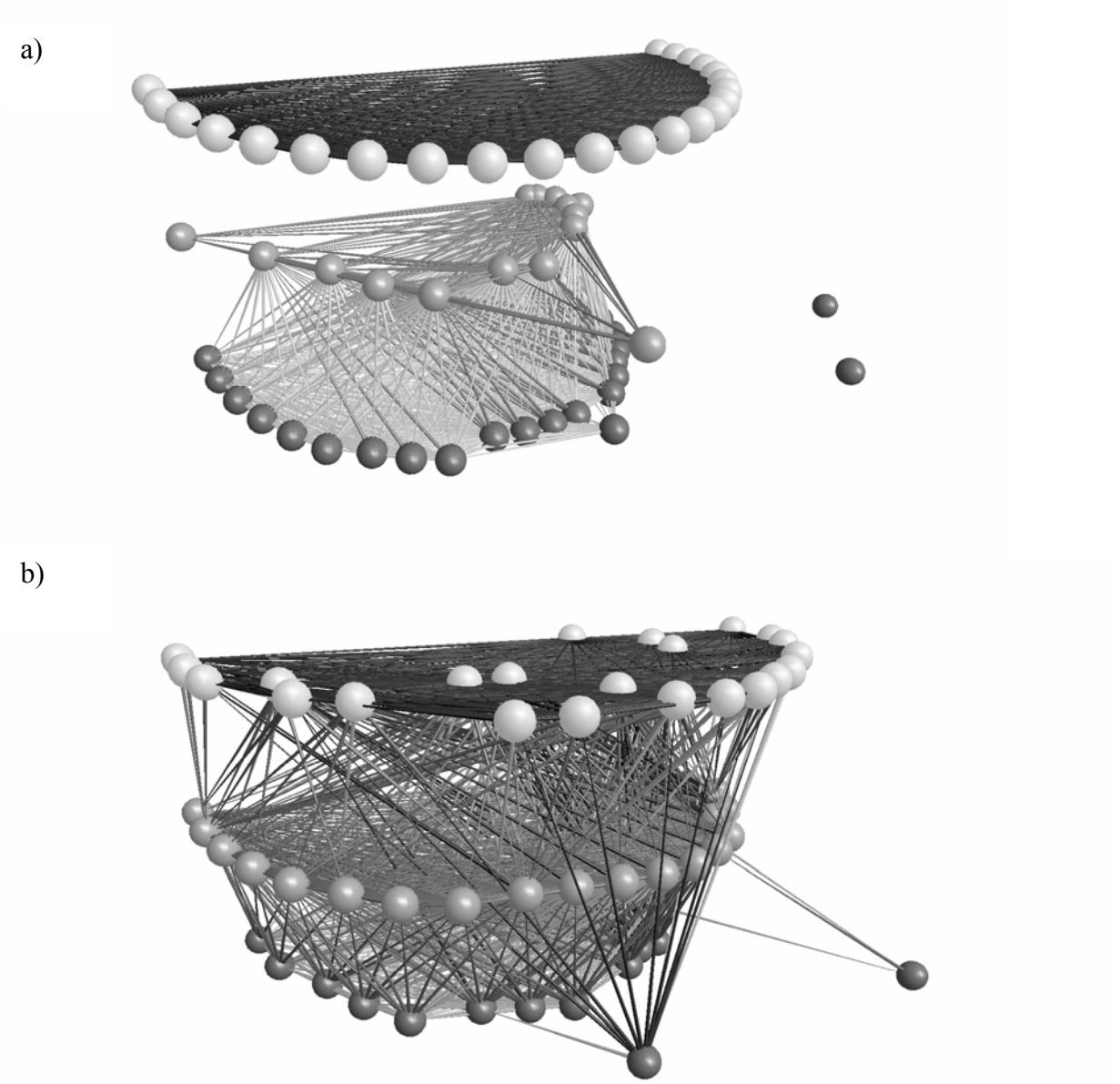


Figure 1. Network graphs of the buffalo association data for May 2002 (A) and November 2001 through October 2003 (B). Balls represent individual buffalo and the lines represent all non-zero association values. Individuals are distributed vertically according to herd membership. Herd membership was determined by cluster analysis (e.g. the solid, dotted and dot-dashed lines in Fig. 2A refer to the black, grey and light grey, vertices respectively in Fig. 1A).

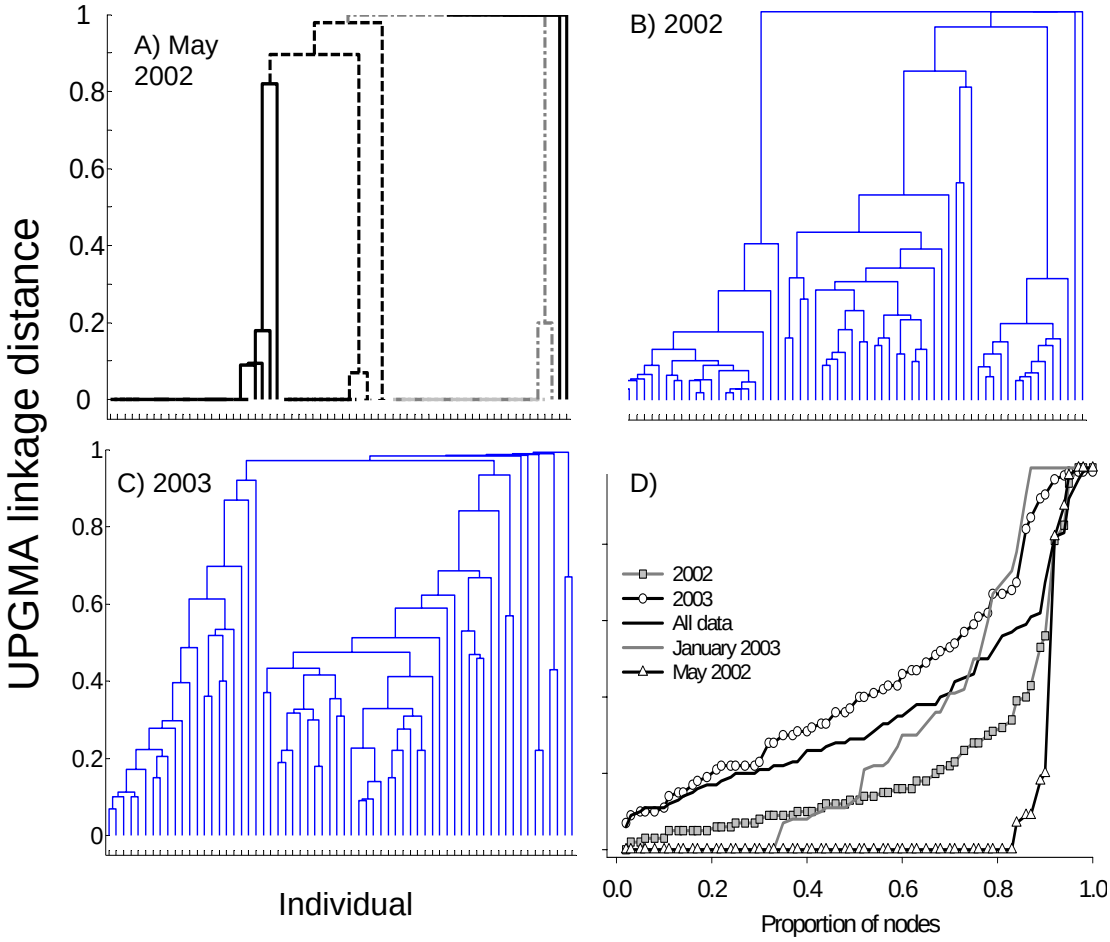


Figure 2. UPGMA cluster analyses of 64 radio-collared buffalo in the Kruger National Park for different periods of data collection: May 2002 (A), Nov 2001 through Oct. 2002 (B), and Nov. 2002 through Oct. 2003 (C). Buffalo with higher association indices are linked at lower linkage distances. In panel A the solid, dashed, and dot-dashed lines show three distinct herds present in May 2002. Panel D compares the overall structure of the dendrograms by showing the linkage distance required to include a given proportion of nodes (see text).

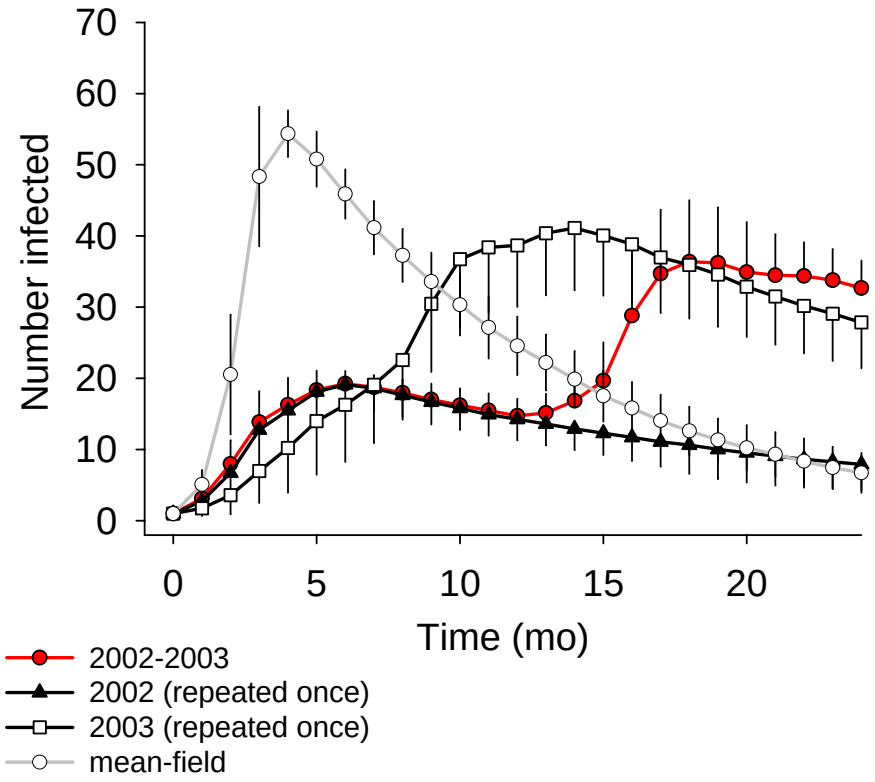


Figure 3. Mean and standard deviations of the number of infected individuals for 50 runs of the disease model using monthly association data from the entire study period (closed circles), 2002 (closed triangles), 2003 (open squares), or a mean-field model (open circles) where all individuals were connected but the force of infection per month was the same. All simulations used a transmission coefficient β of 0.3 and recovery probability γ of 0.1. For simulations using one year of data, the same association matrices were used again the second year.

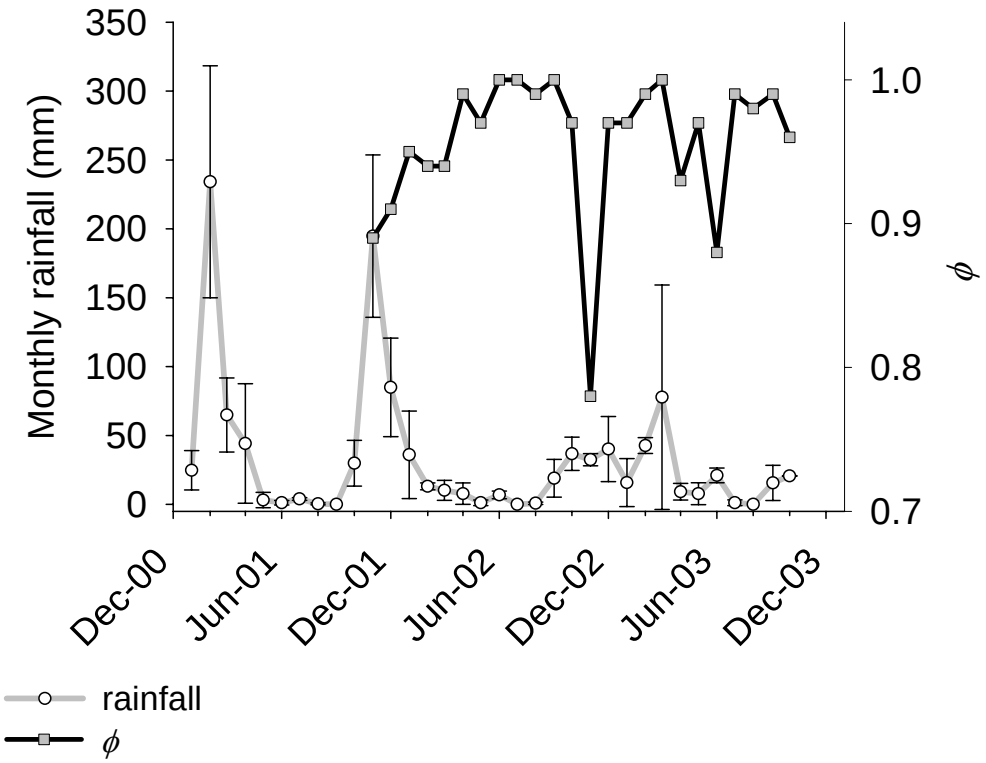


Figure 4. Monthly rainfall (mm) and Keeling's clustering coefficient ϕ of the association data during the course of the study. Mean and standard deviations of the monthly rainfall was calculated using four rainfall stations in the Satara region of the Kruger National Park. Rainfall data from 2000 to 2001 are shown to contrast the below-average wet season in 2002-03.

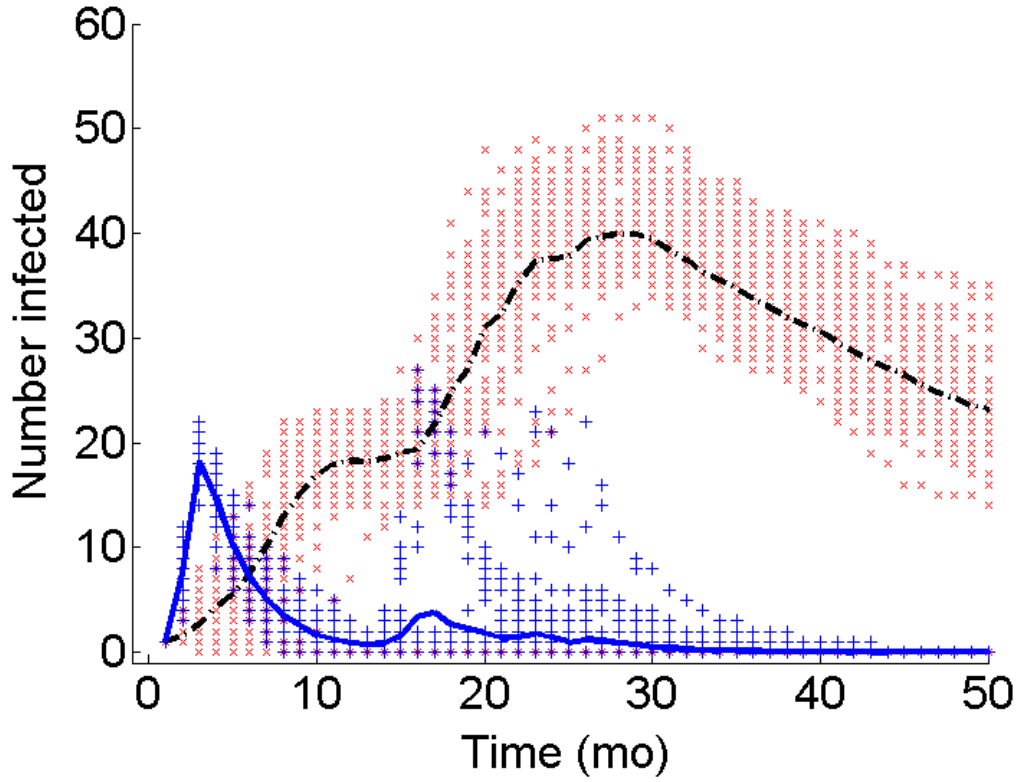


Figure 5. The number of individuals infected as a function of the speed of the disease. Slow diseases (dotted line and ×'s; $\beta = 0.04$, $\gamma = 0.03$) allow for more switching of infectious individuals between groups than faster diseases (solid line and +'s; $\beta = 0.4$, $\gamma = 0.3$), and hence for greater overall disease spread. Simulations used monthly association values from November 2001 to October 2003. Symbols show particular model runs, and lines represent the mean of 50 runs.

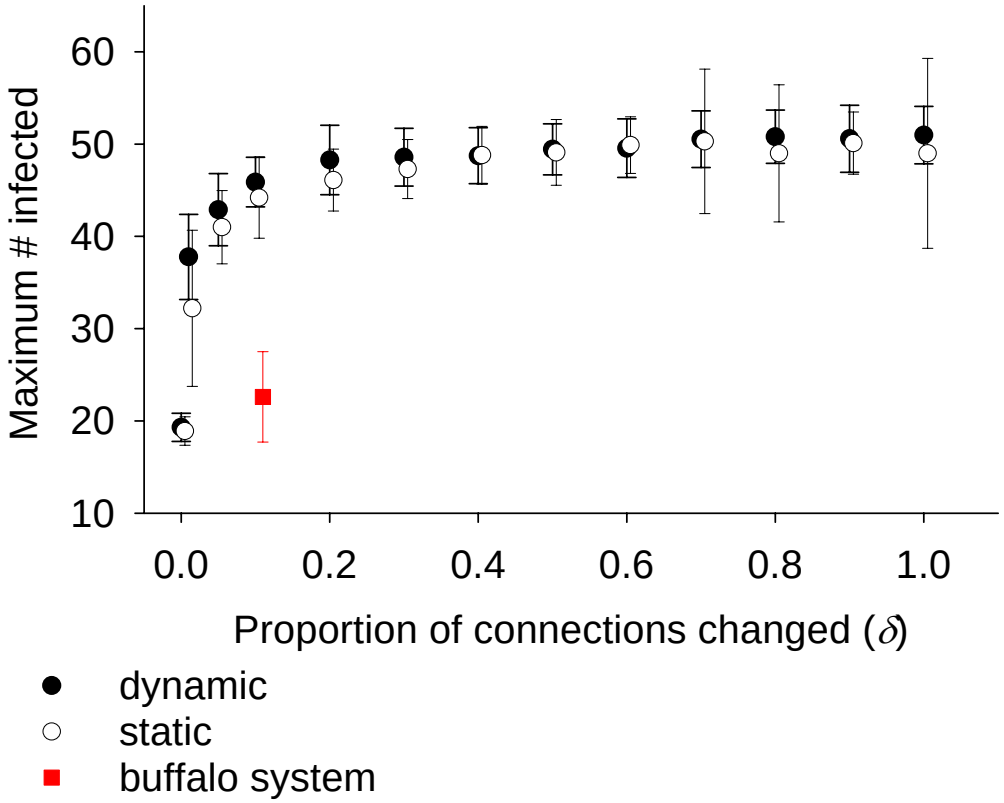


Figure 6. The maximum number of individuals infected at any point in time after 50 time steps depends upon the amount δ of random rewiring of the association network at the beginning of each simulation (static) or cumulatively every time-step (dynamic). Dynamic and static simulations started with association data from November 2001; the buffalo system data point was based on all of the association data (i.e. unmanipulated). Disease parameters were $\beta = 0.3$, $\gamma = 0.2$. Error bars represent the standard deviations from 50 stochastic simulations. (For clarity, δ values of the static simulations were increased slightly before plotting.)

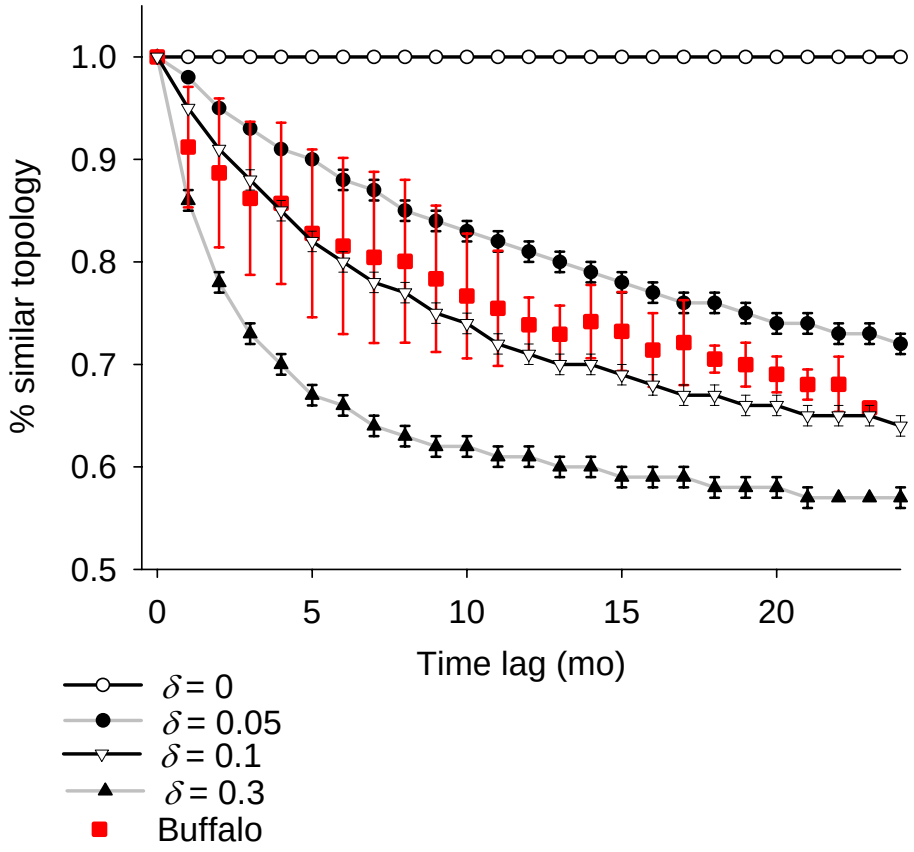


Figure 7. Random rewiring of a proportion of connections δ in the association matrix each time-step decreases the similarity between association matrices over time. Buffalo data (squares) represent the mean and SD of the similarity between all available association matrices that are separated by a time lag of one to 23 months, whereas other lines represent the mean and SD of 50 simulations of the model starting with November 2001 association data using different δ values.

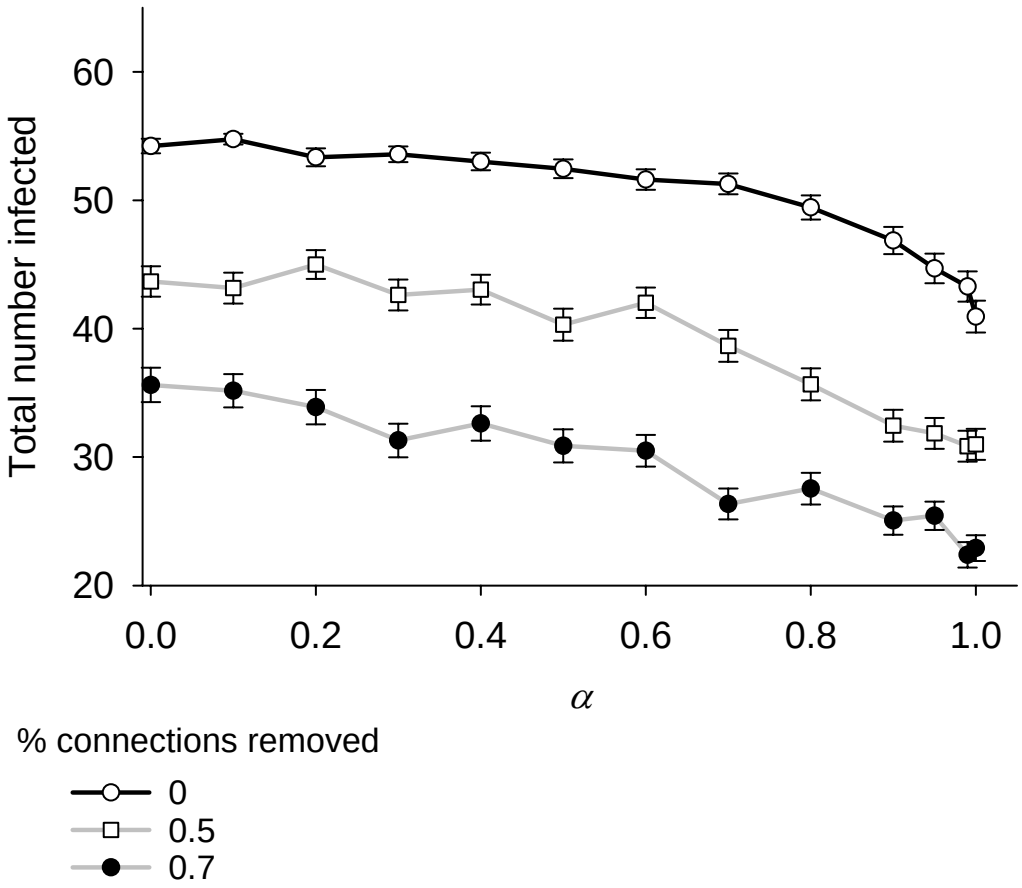


Figure 8. The total number of individuals infected after 50 time-steps decreases with increasing variability in the time-averaged connection strength between pairs and decreasing temporal variability of connections within pairs (α) and the proportion of connections that are removed from the network. See methods for a description of how α increases the variance in association indices. Error bars indicate the standard errors of 200 simulations using September 2003 as the association matrix, $\beta = 0.3$, and $\gamma = 0.1$.

Chapter Five

A disease-eye view of population structure:

dueling timescales of host mixing and disease recovery

P.C. Cross, J.O. Lloyd-Smith, P.L.F Johnson and W.M. Getz

Introduction

At the turn of the 20th century, rinderpest swept through Africa, devastating populations of African buffalo and wildebeest (Sinclair 1977, Plowright 1982, Anderson 1995). From 1929 to 1983 recurrent rinderpest outbreaks occurred in the buffalo and eland populations of Central and Eastern Africa, while many other ungulate species, such as duikers, steenbok, oribi, roan, sable and gerenuk, were relatively unaffected (Anderson 1995). Why were some hosts affected more than others? Traditionally, this may have been explained by immunological differences in susceptibility. We illustrate, however, a significant component of behavioral susceptibility exists that is not a simple function of group size or population density, but rather the interaction of group size and host movement.

Risk of disease is assumed to be a significant cost of group living (Freeland 1976, Moller et al. 1993), yet recent comparative analyses that investigated the effect of group size on the immune system or parasite diversity have had mixed results (Cote and Poulin 1995, Nunn et al. 2000, Nunn 2002, Stanko et al. 2002, Tella 2002, Nunn et al. 2003a, Nunn et al. 2003b). These mixed results may be due, in part, to the interaction of movement and group size, whereby reduced movement rates can mitigate some costs associated with larger group sizes. Specifically, large groups will be exposed to fewer introductions of disease if movement between groups is sufficiently rare.

Several recent theoretical studies have investigated the role of host population structure in the invasion or persistence of disease (Hess 1996, Swinton et al. 1998, Keeling 1999, Keeling and Gilligan 2000, Keeling and Grenfell 2000, Thrall et al. 2000, Park et al. 2001, Fulford et al. 2002, Keeling and Rohani 2002, Park et al. 2002,

Hagenaars et al. 2004). These studies incorporated host movement into structured disease models either phenomenologically or mechanistically. Models with mechanistic host movement explicitly move individuals from one group to another (e.g. Hess 1996, Thrall et al. 2000, Keeling and Rohani 2002). Models with phenomenological host mixing assume that hosts do not move between groups but can infect others within and among groups simultaneously (e.g. Ball et al. 1997, Swinton et al. 1998, Hagenaars et al. 2004). The phenomenological approach may be appropriate for plant-pathogen systems (e.g. Park et al. 2001, Park et al. 2002), but can obscure the relationships between host movement, group size, and disease recovery in mobile host populations. For example, in a system where between-group movements are rare, an epidemic of an acute, highly-transmissible disease may run to completion within a group before any individual moves and spreads infection to a new group. A mechanistic model more readily captures this possibility, while a model with phenomenological mixing between groups does not. In this study, we investigate how the interactions of group size, movement and recovery affect the probability of invasion by disease into structured populations using a mechanistic mixing model.

Lloyd-Smith *et al.* (2004) showed that transmission of sexually-transmitted diseases is well-described by a phenomenological mixing model when partner exchange is very rapid relative to the infectious period, and otherwise a mechanistic pair-formation model is required. Keeling and Rohani (2002) reached similar conclusions for a two-patch system where host mixing was frequent. We expand upon these analyses by exploring a broad range of relative timescales of movement and disease recovery as well as group and population size. Our analysis is motivated by questions

regarding the invasion of disease in wildlife populations where host movement between groups can be rare or relatively frequent (e.g. natal dispersal versus frequent fission and fusion of entire groups), infectious periods range from several days to several years (e.g. rabies versus bovine tuberculosis), and group sizes range from monogamous pairs to thousands of individuals. We focus on directly-transmitted diseases where hosts may move among groups, but contacts that are sufficient for disease transmission occur only within a group. These groups may reflect either social or spatial structure in the host population.

The basic reproductive number, R_0, is the expected number of infections caused by a typical infectious individual in a completely susceptible population (Anderson and May 1991). The R_0 statistic has been the traditional standard by which epidemiologists and disease ecologists quantify the potential growth of a disease (Anderson and May 1991, Diekmann and Heesterbeek 2000). In stochastic models, a disease cannot invade the entire system when $R_0 \leq 1$ and has a non-zero probability of invading only when $R_0>1$. In the simplest case of an SIR (Susceptible-Infected-Recovered) disease (Anderson and May 1991), R_0 is the ratio of two rates, or timescales: the infection rate and the recovery rate. If transmission is density independent, with rate parameter β, and the recovery rate γ is constant, then $R_0 = \beta/\gamma$ (McCallum *et al.* 2001). Further, the average length of the infectious period is $1/\gamma$. We use a stochastic metapopulation model to illustrate the importance of another ratio of two timescales, specifically the ratio of the rates at which hosts move between groups (μ) and recover from disease (γ). For the simple case of constant recovery and no mortality, this ratio, μ/γ, is the expected number of times an infectious individual will move between groups.

First, we describe the simulation model and explore how the interactions of group size, host movement, and infectious period affect the probability of invasion by a disease. Next, we describe a relatively new metric of disease invasion, R_*, which is the number of groups that are expected to become infected from the initially infected group (Ball *et al.* 1997). In other words, R_* is the group-level analogue of R_0. We then use the simulation model to estimate R_0 and R_* and demonstrate that R_* is a better predictor of disease invasion in structured populations with mechanistic host movement between groups. We conclude with a number of testable predictions that follow from the ideas presented here.

Demonstration of Dueling Timescales Effect

We use an individual-based, stochastic, discrete-time SIR model to investigate how the dueling timescales of host movement and disease recovery affect the ability of a directly-transmitted disease to invade a spatially, or socially, structured population. The total host population is evenly divided into an array of groups. The host population is further subdivided into susceptible, infected, and recovered classes where S, I, and R are respectively the number of hosts in each category. Three processes are described in the model: movement between groups, infection, and recovery of infected hosts. Since the intent of the model is a qualitative description of different interactions, we have simplified these processes as much as possible. For the case presented here, we consider a successful invasion to have occurred when the disease becomes a pandemic and infections occur within all groups of a structured population. This narrow definition does not count disease establishment within a single patch as an invasion, but

instead emphasizes the spread of the disease among groups which is the phenomenon of interest here.

We assume that movement between groups is density independent, and all individuals have a constant probability, μ, of leaving their current group each time step. Groups are organized on a square lattice and individuals can only move to their four nearest-neighboring groups. To avoid boundary effects, opposite edges of the array are connected to create a torus. In the supplementary material, we expand the analysis to include a loop structure, where each group has only two nearest-neighbors, and a spatially implicit array, where individuals can move to any other group within a time step (equivalent to the 'island' model used previously (Hess 1996, Fulford et al. 2002)). We assume that infected individuals recover to an immune class with a constant probability, γ, per time step.

To isolate the effects of host movement and facilitate the comparison of disease dynamics in a range of population structures, we assume disease transmission is frequency-dependent (Getz and Pickering 1983). Thus the probability of infection per time step for each susceptible in group i is given by the expression $1 - \exp\left(-\beta \frac{I_i}{N_i}\right)$, where β is the transmission coefficient, I_i is the number of infected individuals in group i, and N_i is the total number of individuals in group i. Since we do not incorporate host demographic dynamics, the assumption of frequency-dependent rather than density-dependent transmission represents a rescaling of the transmission coefficient β. If contact or transmission rates increase with population density (McCallum *et al.* 2001), then disease invasion would be even less likely when the population is divided into

many small groups than indicated by results presented here, but our overall conclusions about the interaction of movement and recovery timescales would still hold.

In a continuous-time model with frequency-dependent transmission and a constant recovery rate, $R_0 = \beta/\gamma$ (Anderson and May 1991, McCallum et al. 2001). For the discrete-time model used here, β/γ is an approximation of R_0, which works well when the probability of infection per timestep is small and group sizes are relatively large. The approximation does not change our qualitative conclusions, however, so for clarity we refer to the ratio β/γ as R_0. We also assume that disease invasion is fast relative to birth and death rates, so the total population size is constant. Each simulation starts with one infected individual, and all groups begin with the same number of individuals. Since our spatial model was symmetric, group sizes remained relatively constant during the course of each run.

We begin by comparing the dynamics of two diseases with the same R_0 value (β/γ) but where one disease is slow (i.e. a chronic disease with a relatively long infectious period; $\beta = 0.05$, $\gamma = 0.01$) while the other is an order of magnitude faster ($\beta = 0.5$, $\gamma = 0.1$). For time steps of 1 day, these parameters correspond to mean infectious periods of 100 or 10 days, respectively. We simulated the invasion of these two diseases in three different host population structures: 1 group of 1000 individuals (equivalent to the common "mean-field approximation" of random mixing among all individuals), 25 groups of 40 individuals, and 100 groups of 10 individuals.

As expected, subdividing the population into more groups decreases the probability of pandemic (Wilson and Worcester 1945) because it decreases the average group size and increases the number of between-group jumps the disease must make to

penetrate the entire population. A less obvious effect is that slower diseases are more likely to invade a structured population, even if they have the same R_0 as a faster disease. For the case of 100 groups of ten individuals, the slow disease ($\gamma = 0.01$, $\beta = 0.05$) infected, on average, far more individuals than the fast disease ($\gamma = 0.1$, $\beta = 0.5$) before the disease died-out (658 ± 45 SE compared to 19 ± 2.3 SE; Fig. 1). This typifies the interaction of the host movement and disease recovery timescales: diseases with longer infectious periods allow more time for host mixing to occur and thus experience populations that are effectively larger. When the movement rate is zero, neither a fast nor slow disease will invade the entire population regardless of the value of R_0. When movement is very frequent, both the fast and the slow disease are likely to invade the structured population provided that $R_0>1$.

The simulated epidemics in the one-group and 100-group populations differ markedly, but were less different for the slow disease compared with the fast disease. In other words, approximating a structured population by a mean-field model (i.e. a single group with homogenous mixing of hosts) is more appropriate for slow diseases than fast diseases (Fig. 1; Cross 2004). The faster the disease, the more important it is to incorporate the spatial/social structure into any analysis. The mean infectious period ($1/\gamma$) defines the natural disease timescale, and when movement occurs on this timescale or slower then movement should be incorporated mechanistically, rather than phenomenologically.

Next we examine a range of host movement and recovery probabilities. The proportion of the population infected over the course of an epidemic depends on the expected number of group changes per infectious lifetime (μ/γ), on R_0 (β/γ), and on

group size (Fig. 2). When movement is infrequent relative to the recovery rate, R_0 has little predictive ability because few infections result for all values of R_0 (Fig. 2a). If movement is frequent relative to recovery, increasing R_0 increases the average proportion of the population that becomes infected (Fig. 2a), consistent with predictions from mean-field models (Anderson and May 1991, Diekmann and Heesterbeek 2000). Increasing the host group size decreases the amount of host movement required for the disease to invade the entire population (Fig. 2b). Larger groups experience larger within-group outbreaks, and hence more infected individuals dispersing from each infected group (given density-independent movement). For our model with frequency-dependent transmission, the total number of infected individuals is, on average, a fixed proportion of group size; for density-dependent transmission, the proportion infected would increase with group size, causing greater increases in the number of infected dispersers.

For high values of R_0, the ratio μ/γ yields a sharp threshold for invasion (Fig. 2b). As a rule of thumb, a disease will invade the metapopulation if μ/γ is greater than the reciprocal of the expected number of individuals that will be infected within a single group. This makes intuitive sense because in this model system μ/γ is the expected number of between-group movements made by each infectious individual. Thus, μ/γ multiplied by the expected number of infected individuals is the expected number of infected dispersers per group—which must exceed one for a pandemic. When R_0 is high, almost all individuals in a group will be infected, so for a pandemic μ/γ should be greater than the reciprocal of the average group size. For example, if the group size is

200, then, on average, more than one infectious individual in 200 will need to move between groups (i.e. $\mu/\gamma > 0.005$) for a pandemic to occur.

The mean proportion of the population that becomes infected, shown in Fig. 2, obscures the underlying distribution of the number of infections per epidemic (i.e. over different runs of the stochastic simulation). It is incorrect to assume that the mean of this distribution is similar to the median or mode, because in many cases the distribution is bimodal with peaks centered on zero and one or close to one (Fig. 3c, f, i). When movement is very infrequent relative to recovery, the probability of a pandemic is close to zero because the disease almost always dies out within the initial group (Fig. 3a,d,g). When movement is frequent, then the disease tends to either die out stochastically within the initial group or invade most or all of the metapopulation (Fig. 3 c, f, i), with the relative frequencies of die-out versus invasion determined by R_0 as in mean-field models (Diekmann and Heesterbeek 2000). With intermediate movement rates variation is considerable with regard to the extent to which the disease penetrates the population (Fig. 3b).

These results (Fig. 2a, Fig. 3) agree with previous studies when μ/γ is either much greater than one or close to zero. When μ/γ is large, then group structure of the population is less important and β/γ is a good predictor of disease invasion (Fig. 2a). When μ/γ is close to zero then β/γ is a good predictor of disease invasion within the initial group, but the probability of the spread of disease between groups is rather small. For the intermediate scenarios we analyzed, however, the ratio of movement to recovery rate (μ/γ) has greater influence on the invasion of a disease than β/γ.

Predictors of a Pandemic

Recent theoretical work has extended the R_0 concept to account for depletion of the susceptible pool (Keeling and Grenfell 2000), host spatial structure (Keeling 1999, Fulford et al. 2002), and populations with heterogeneous infectiousness or susceptibility (Diekmann and Heesterbeek 2000). However, even after incorporating these effects R_0 may be misleading in metapopulations with limited mixing, since R_0 as it is traditionally used, is an individual-based measure. Ball and colleagues demonstrated that the individual-based R_0 is not the best predictor of disease invasion in structured host populations and introduced a group-level reproductive number, R_*, which is the average number of groups infected by the initially infected group (Ball et al. 1997, Ball 1999, Ball and Lyne 2001, Ball and Neal 2002, Ball et al. 2004). This finding has been echoed in the context of reproductive fitness of a new mutant in a metapopulation (Gyllenberg and Metz 2001, Metz and Gyllenberg 2001). In a model with phenomenological mixing, $R_* > 1$ is the formal threshold criterion for invasion of a disease into an infinite number of finite-sized groups (Ball *et al.* 1997).

The phenomenological mixing model used by Ball *et al*. facilitates analysis, but to demonstrate the utility of the R_* metric in the context of interacting timescales of host movement and recovery, we applied our simulation model with explicit host movement between groups. For the model described above, we estimated R_0 and R_* by tracking the mean number of infections caused by the initially infected individual or group, respectively. Then we averaged these estimates of R_0 and R_* over many runs of the stochastic model. When a susceptible individual was infected and two or more infected individuals were present within the group, we randomly allocated the infection to only

one of those infectious individuals. To estimate R_*, we tracked the number of groups that were infected by individuals that were themselves infected within the index group. Infected individuals had to move to a completely susceptible group and cause infection there in order to contribute to R_*. When individuals from multiple groups moved to a susceptible group and caused an infection, we randomly allocated the infection to one of the individuals (and thus its source group). The mean estimates over many simulations, denoted as $\hat{R}_0$ and $\hat{R}_*$, are "empirical" in the sense that they are based on data collected from simulated epidemics mimicking epidemiological contact-tracing data. As estimates from model output, they incorporate the effects of spatial structure, host movement, and depletion of the susceptible pool. Thus they will differ from traditional analytical R_0 and R_* values, which assume an infinite susceptible population and, hence, overestimate the value of these parameters when populations have a finite size.

We simulated the model using a range of transmission (β) and movement (μ) probabilities, a fixed recovery probability $\gamma = 0.1$, and an 11x11 toroidal array with ten individuals per group and nearest neighbor movement. Each parameter set was simulated 1000 times to generate mean values of $\hat{R}_0$ and $\hat{R}_*$. We then plotted the relationship between the model output variables $\hat{R}_0$, $\hat{R}_*$, and the proportion of the population infected (Fig. 4, Fig. S2.). Each line in Figure 4 corresponds to a fixed within-group transmission rate (β) and a range of movement probabilities increasing from left to right.

The empirical individual-level $\hat{R}_0$ is not a good predictor of the mean proportion infected: even when $\hat{R}_0$ is much greater than one, the mean proportion infected may be

low depending upon the movement probability (Fig. 4a, Fig. S2). Also for different β, the proportion infected appears to show a threshold at different values of $\hat{R}_0$. The group-level $\hat{R}_*$, on the other hand, is a much better predictor of a pandemic in a structured population (Fig. 4b). In an idealized metapopulation, $R_* > 1$ is the threshold above which there is a finite probability of disease invasion (Ball *et al.* 1997). In our simulations, the proportion infected begins climbing at $\hat{R}_* \approx 1$ and rises most steeply around $\hat{R}_* \approx 2$ (Fig. 4b). This gradual transition around the threshold is typical of stochastic epidemic models, particularly with spatially-constrained mixing, because the invasion has many chances to die out before invading the entire population. When transmission rates (and hence R_0) are low, $\hat{R}_*$ is small for all values of movement (Fig. 4b, $\beta = 0.1$). When transmission is intermediate and movement is frequent, the disease will either stochastically die out in the initial group or invade the entire population (Fig.3d, e, f), resulting in intermediate values of $\hat{R}_*$ and mean proportion infected (Fig. 4b, $\beta = 0.5$ and 1). Finally, when both movement and transmission rates are high, $\hat{R}_*$ and the mean proportion infected are also high.

Future Empirical Research

These findings suggest important directions for empirical studies, as well as a number of testable predictions. Previous analyses of disease presence/absence in different host social structures have not considered the interaction between movement rates and the duration of the infectious period (e.g. Cote and Poulin 1995, Nunn et al. 2000, Altizer et al. 2003, Nunn et al. 2003a, Nunn et al. 2003b). Our results illustrate,

however, that it is the relative timescales of movement, recovery, and infection that determine the probability of a pandemic. A slow, chronic disease may "perceive" a host to be relatively well-mixed with frequent movement of individuals among groups. An acute disease will perceive that same host population to be more structured because movements between groups are less frequent relative to the timescale of the infectious period (Cross *et al.* 2004). We hypothesize that all else being equal, chronic diseases will be more likely to penetrate structured populations than acute diseases. Conversely, we hypothesize that behaviorally susceptible host species, with large groups and frequent movement, are likely to be more heavily impacted by acute diseases than hosts with small groups and infrequent movement. Thus the ratio of acute to chronic diseases found in different host populations should increase as a function of group size and movement rate.

A major focus of recent disease ecology has been how transmission or contact rates depend on population density (e.g. Bouma et al. 1995, Begon et al. 2003), but for metapopulations, we have shown that movement rates are also critical to understanding disease invasion (Figs. 2-4). Despite the importance of host movement, very few studies have been published that examine the amount of mixing between groups of many wildlife species. Our results can help to guide the design of field studies intended to estimate host movement for disease models. The proportion of individuals that should be tracked and the duration of the study will depend upon the infectious period of the disease as well as the average group size of the host. As group sizes and infectious periods increase, the amount of movement required for a pandemic to occur decreases. Low movement rates, however, will require researchers to track more

individuals to accurately estimate the amount of movement among groups. If group sizes are large, say 200, and the disease is highly infectious, say $R_0 > 5$, then only 1 in approximately 200 infectious individuals needs to switch between groups for a pandemic to become likely (Fig. 2b).

Researchers may estimate movement between groups using genetic data or tracking of known individuals (Waser et al. 1994, Koenig et al. 1996, Cain et al. 2003, Nathan et al. 2003). Radio-tracking or re-sighting data are more effective than genetic methods as long as individuals frequently move between groups relative to the duration of the study. Genetic methods of estimating movement will be relevant only for chronic diseases in large groups, and their use in a disease context involves at least three major assumptions: 1) past movement accurately reflects current movement; 2) short-term movements that are likely to be missed in genetic signatures (e.g. foraging rather than mating) are unimportant to disease dynamics; and 3) moving individuals are as reproductively successful as non-moving individuals (Waser *et al.* 1994).

Future Theoretical Research

Our findings emphasize that the group-level reproductive number R_* is a critical determinant of invasion success in structured populations. Analytical formulations of R_* in systems with explicit host movement may clarify the important interaction between timescales of host movement and disease recovery, and help to formalize the rule of thumb proposed above. Previous work on R_* has focused on models with phenomenological host mixing and an infinite number of groups such that all new infections are in susceptible groups (Ball et al. 1997, Ball and Neal 2002, Ball et al.

2004). Analogous to recent developments in the theory of R_0 (Keeling 1999, Keeling & Grenfell 2000), further work on R_* should consider finite populations with spatial constraints on movement. Longer dispersal distances and spatial configurations that increase the number of neighboring groups (Supplementary Material) will mitigate the depletion of susceptible groups and facilitate the invasion of a disease. These effects are implicitly incorporated into our $\hat{R}_0$ and $\hat{R}_*$ estimates, but analytical exposition would help advance our understanding.

In our stochastic model, we made a number of simplifying assumptions that could be relaxed to make our simulations more realistic. First, the assumption of a constant probability of recovery per time step, γ, results in a geometrically distributed infectious period. The effects of alternative infectious period distributions on R_* are unclear (cf. Keeling & Grenfell 2000). For instance, with a fixed infectious period, time spent in the home group while infectious will increase the number of local infections, but will also diminish the infectious period in any new group, thereby decreasing the number of infections elsewhere. A fixed infectious period would also cause fewer individuals to recover before moving, compared to the geometric infectious period (with its mode at one timestep). Second, we have assumed that movement between groups occurs instantaneously and without mortality, but if individuals spend time or die during movement their infectious lifetime within the next group is reduced, thereby decreasing R_*. Finally, we assumed that disease invasion was fast relative to the timescale of host birth and death. This is less likely to hold for chronic diseases, or for acute diseases that lead to rapid mortality rather than recovery. Both natural and disease-induced mortality shorten the infectious period and thus reduce R_0 (Anderson

and May 1991) and R_*. Our broad conclusions about the interaction of host movement and disease recovery timescales should still apply, but investigating the effects of host demographics and disease mortality on R_* would be an important extension of this study.

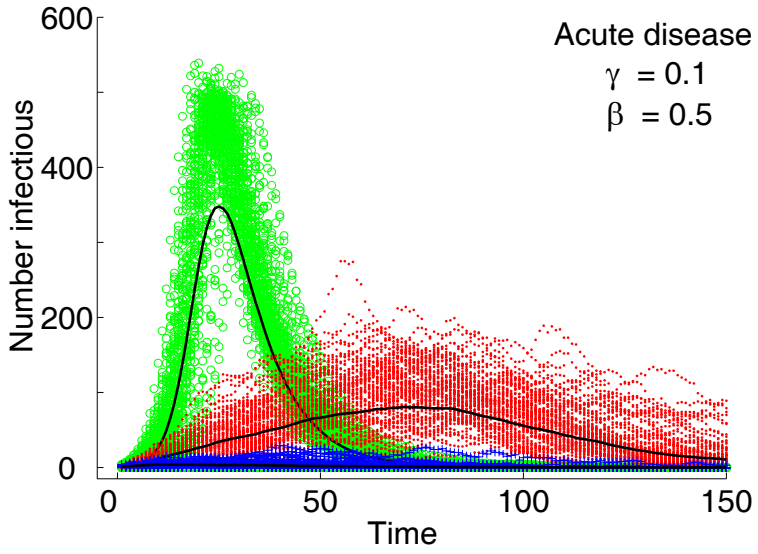
Conclusion

Traditionally, epidemiologists and disease ecologists have focused on $R_0 > 1$ as a threshold for disease invasion (e.g. Anderson and May 1991, Diekmann and Heesterbeek 2000). We have shown that in metapopulations the relationship between invasion of disease and an individual-level R_0 is often weak. Even for very large values of R_0, a pandemic is unlikely if the expected number of times an individual will move between groups during their infectious lifetime (μ/γ) is low (Fig. 2). Pandemics in structured populations require both within-group and between-group transmission, and the group-level reproductive number R_* is a better predictor than the individual-level R_0 for these systems (Fig. 4). Results from our individually-based stochastic model support the analytical results of Ball et al. (Ball et al. 1997, Ball and Neal 2002, Ball et al. 2004), which proved that $R_* > 1$ is the threshold for disease invasion in a population with group structure. As a general rule of thumb, the individual-level R_0 must be greater than one and the expected number of group changes while infectious (μ/γ) multiplied by the average group size must be greater than one for a pandemic to occur (Fig. 2b).

Chronic diseases with longer infectious periods allow for more mixing to occur between groups. As a result, chronic diseases will perceive more thoroughly mixed host populations and exhibit dynamics that are closer to those predicted by mean-field

models than acute diseases (Fig. 1). For the same R_0, chronic diseases are more likely to invade structured populations than slow diseases. 'Slow' and 'fast' diseases are relative terms: a fast, acute disease in a host population with frequent movement between groups may well behave like a relatively slow disease in a population with less frequent movement. The probability of a pandemic in a structured population is thus an emergent property of the interaction of host and parasite demography and behavior, incorporating a dimension of host behavioral susceptibility arising from group size and movement rates. The results presented here, and in a recent paper by Lloyd-Smith et al. (2004), suggest that when contact, movement, birth and death processes occur on a timescale similar to that of the disease (i.e. the infectious period), these processes should be incorporated mechanistically into disease models.

a)



b)

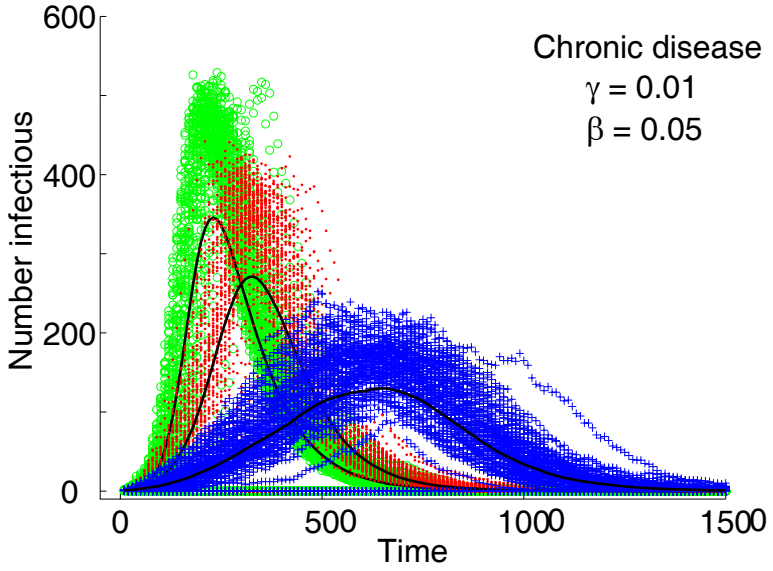
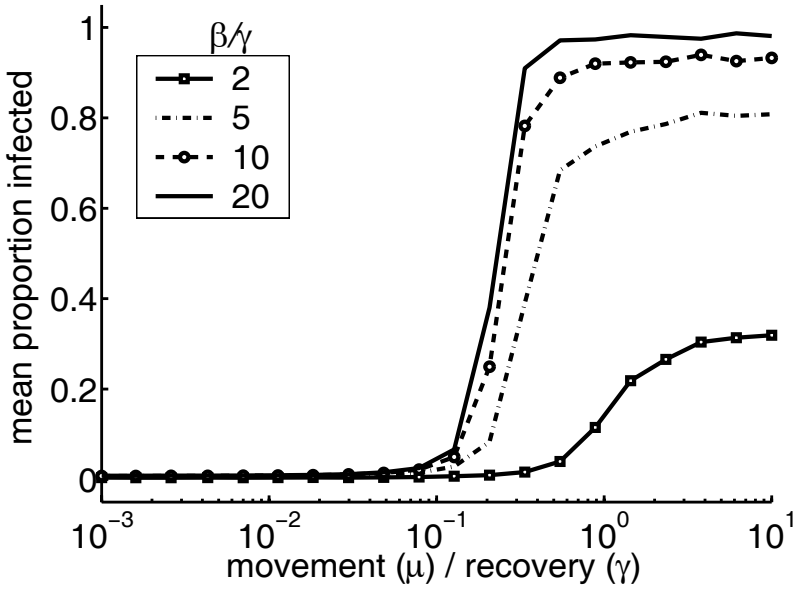


Figure 1. Disease invasion depends upon population structure (green circles: 1 group of 1000 individuals; red points: 25 groups of 40 individuals; blue crosses: 100 groups of 10 individuals) and the duration of the infectious period. A mean-field model of one group (green circles) is a worse approximation of a structured population for an acute disease with $\gamma = 0.1$ (a) than a chronic disease with $\gamma = 0.01$ (b). For both diseases $\beta/\gamma = 5$, but the slow disease causes more infections in the structured population. Lines represent the mean of 100 simulations. Simulations with 25 or 100 groups were run on a toroidal spatial structure with a movement probability μ of 0.01, such that $\mu/\gamma = 0.1$ (a) or 1 (b).

a)



b)

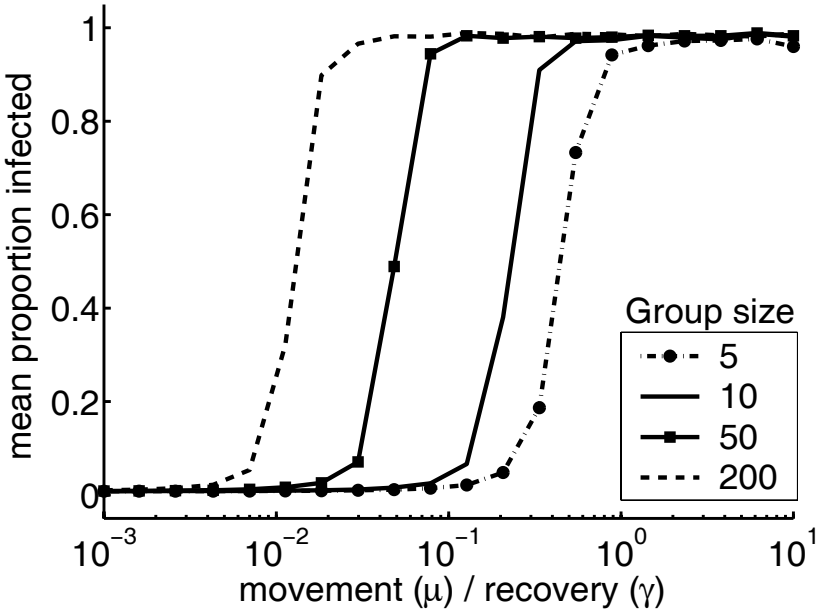


Figure 2. The interaction of movement (μ), transmission (β), and group size determines the mean proportion of the population that becomes infected. In (a) β varied from 0.2 to 2 while group size was fixed at 10. In (b) group size was increased from 5 to 200 while β was fixed at 2 (β/γ=20). Increasing β/γ only affected the proportion infected when movement was frequent (a). Larger group sizes require less movement for the disease to invade (b). Each parameter set was simulated 1000 times on an 11x11 toroidal array of groups with a constant recovery probability γ of 0.1.

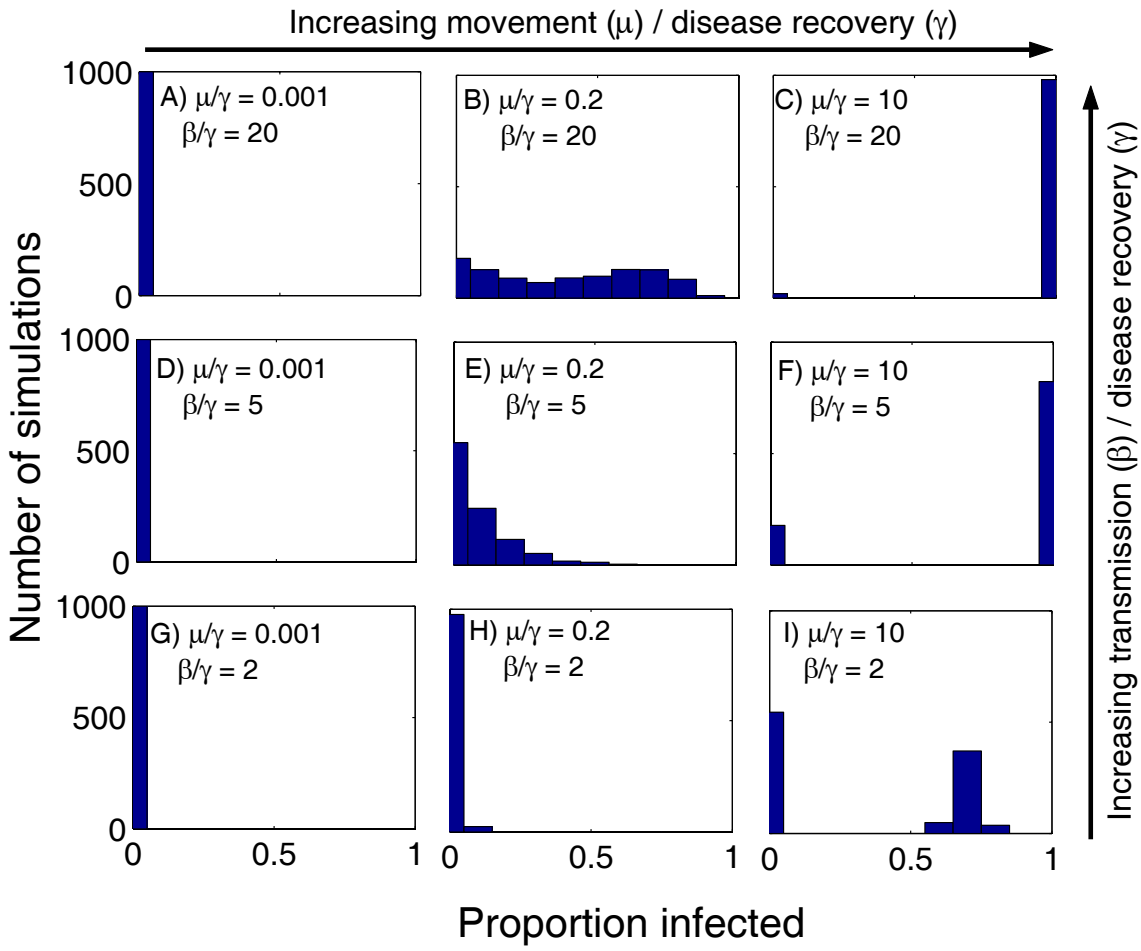
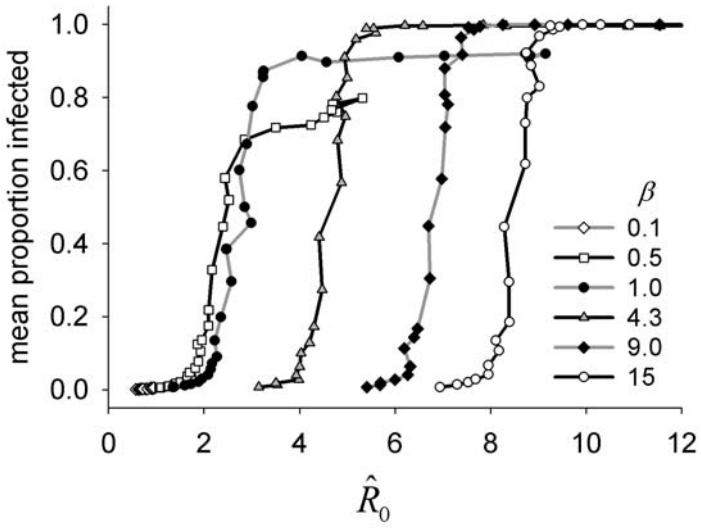


Figure 3. Histograms of the proportion of individuals infected during an epidemic for different transmission (β) and movement (μ) values scaled by the probability of disease recovery (γ). Each parameter set was simulated 1000 times on an 11x11 toroidal array of groups with 10 individuals each and a recovery probability γ of 0.1.

a)



b)

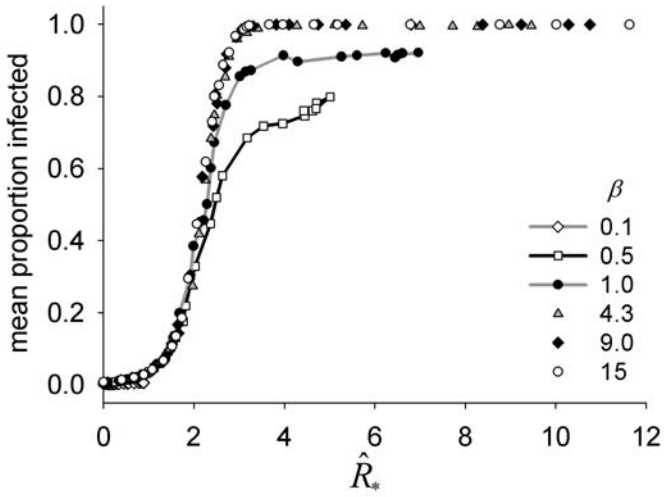


Figure 4. $\hat{R}_0$ can be substantially greater than one and yet not cause a pandemic (a), whereas $\hat{R}_*$ is a much better single predictor of the mean proportion of individuals infected (b). Each line represents a fixed transmission parameter β and a range of movement probabilities from zero to one (increasing from left to right) sampled on a log scale. Each parameter set was simulated 1000 times on an 11x11 toroidal array of groups with 10 individuals each and a recovery probability γ of 0.1.

Supplementary Material

In the main text we focused primarily on an 11x11 toroidal array of groups with differing group sizes, and hence varying population sizes. Here we expand our analysis to investigate the effect of total population size while group size is kept constant by varying the number of groups. The effect of population size is minimal provided the group size remains fixed (Fig. S1a). At low values of μ/γ though, the proportion infected is higher for the 5x5 array than the 11x11 or 21x21 array simply because the initial herd represents a greater proportion of the population. The difference shown in Fig. 2b thus can be interpreted as a group size effect rather than an effect of increasing total population size.

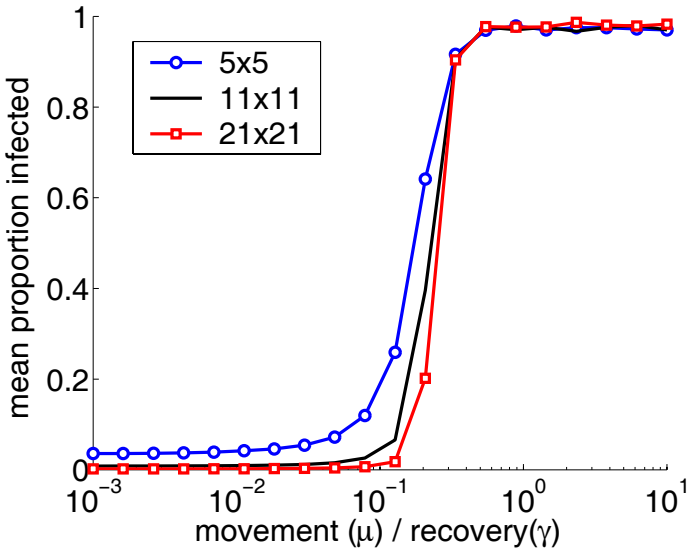
We also investigated the effect of two other spatial configurations: a loop with nearest-neighbor movement and a spatially implicit array where an individual could move to any other group in one time step. As expected, disease invasion of the loop array requires a higher μ/γ ratio than invasion of the torus due to the depletion of susceptible groups (Fig. S1b). On a loop an infected individual dispersing from any given group has only two neighboring groups (compared to four in the torus) that may have been infected already by a previously dispersing individual. In contrast, invasion of the spatially implicit metapopulation requires less host movement than the torus because all groups are neighbors, thus minimizing the depletion effect by maximizing the number of neighboring groups.

Finally, we present a more complete picture of the relationships among $\hat{R}_0$, $\hat{R}_*$ and the mean proportion infected than shown in Fig. 4. Similar to Fig. 4, we simulated range of transmission and movement probabilities on an 11x11 toroidal array of groups

with 10 individuals each. In contrast to Fig. 4, in Fig. S2 each line represents a fixed movement probability, but a range of transmission coefficients. When movement is infrequent, higher transmission rates cause an increase in $\hat{R}_0$, but $\hat{R}_*$ and the mean proportion infected are relatively unaffected. Increasing the movement probability increases $\hat{R}_*$, for a given value of β. Note that when $\hat{R}_0$ is near one then both $\hat{R}_*$ and the average proportion infected are low: if the disease is likely to die out within a patch ($\hat{R}_0 \approx 1$) more frequent movement will not lead to a pandemic. When $\hat{R}_0$ is greater than one it is not a strong predictor of the mean proportion infected because the penetration of the disease into the metapopulation depends upon $\hat{R}_*$ and the probability of host movement. For a pandemic to occur both $\hat{R}_0$ and $\hat{R}_*$ must be significantly greater than one (Fig. S2).

Note that there are regions of the $\hat{R}_0$ - $\hat{R}_*$ parameter space where lines are absent. If $\hat{R}_0 \leq 1$ then $\hat{R}_*$ is not greater than one 1 because the disease is unlikely to invade multiple groups if extinction is likely within the initial group. Further, for the simulations we conducted with a recovery rate of 0.1, $\hat{R}_0$ was usually less than the group size because of the competition amongst infectious individuals to infect the remaining susceptibles within the group. $\hat{R}_0$ can exceed the group size, however, if movement rates are high and recovery rates are low.

a)



b)

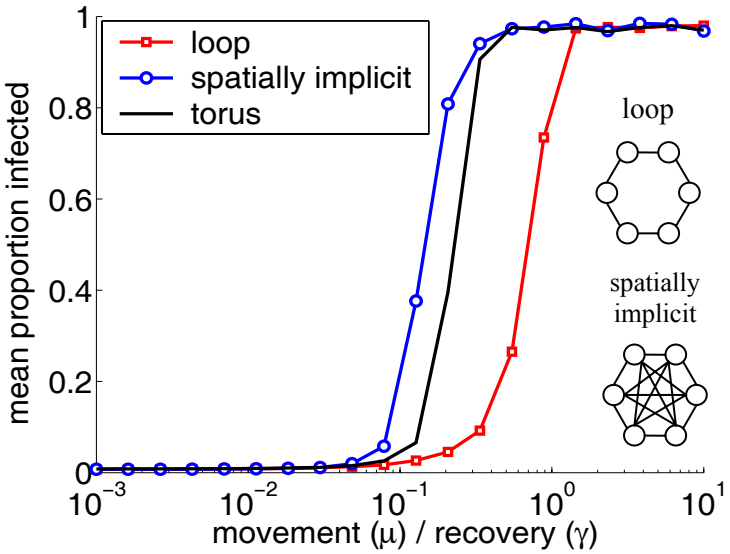


Fig. S1. The mean proportion of the population that becomes infected hardly depends on the number of groups, or individuals, in the population (a), but does depend on the spatial array (b). In (a) the spatial configuration is a torus, but the number of groups (and hence total population size) varies. In (b) all simulations have 121 groups in different spatial configurations. Each parameter set was simulated 1000 times, with a group size of ten and a recovery probability γ of 0.1.

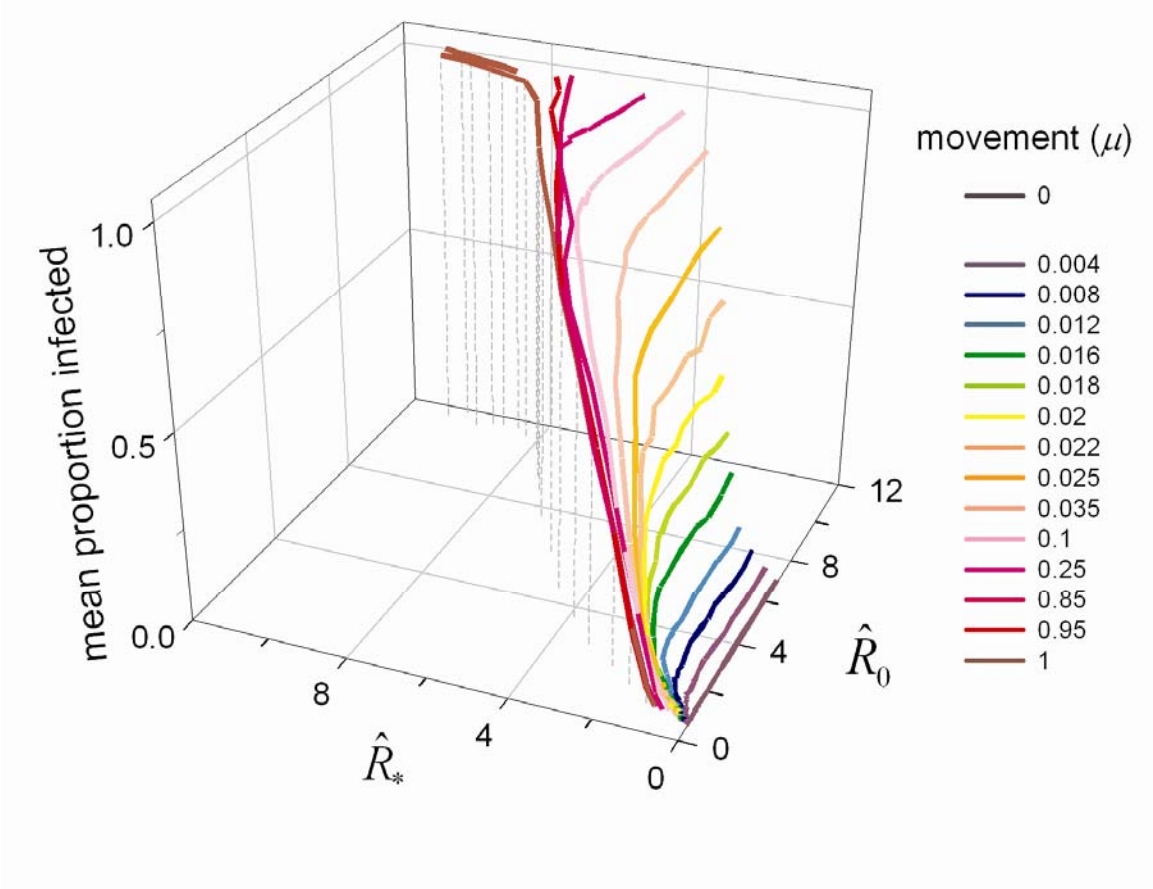


Figure S2. The mean proportion of individuals that are infected is a function of the average $\hat{R}_*$ and average $\hat{R}_0$. Each line represents averages of 1000 simulations with a fixed movement probability and varying transmission rates. Note $\hat{R}_0$ can be substantially greater than one and yet not cause a pandemic if movement is infrequent. All parameter sets were simulated on an 11x11 toroidal array of groups with 10 individuals each and a recovery probability γ of 0.1.

References

Adamo, S. A., M. Jensen, and M. Younger. 2001. Changes in lifetime immunocompetence in male and female *Gryllus texensis*: trade-offs between immunity and reproduction. Animal Behaviour **62**:417-425.

Alberts, S. C., and J. Altmann. 1995. Balancing costs and opportunities - dispersal in male baboons. American Naturalist **145**:279-306.

Alcock, J. 1998. Animal Behavior: An Evolutionary Approach, 6 edition. Sinauer Associates, Inc., Sunderland, Massachusetts.

Aldwell, F. E., D. L. Keen, V. C. Stent, A. Thomson, G. F. Yates, G. W. De Lisle, and B. M. Buddle. 1995. Route of BCG administration in possums affects protection against bovine tuberculosis. New Zealand Veterinary Journal **43**:356-359.

Altizer, S., C. L. Nunn, P. H. Thrall, J. L. Gittleman, J. Antonovics, A. A. Cunningham, A. P. Dobson, V. Ezenwa, K. E. Jones, A. B. Pedersen, M. Poss, and J. R. C. Pulliam. 2003. Social organization and parasite risk in mammals: Integrating theory and empirical studies. Annual Review of Ecology Evolution and Systematics **34**:517-547.

Anderson, E. C. 1995. Morbillivirus infections in wildlife (in relation to their population biology and disease control in domestic animals). Veterinary Microbiology **44**:319-332.

Anderson, R. M., S. Gupta, and W. Ng. 1990. The Significance of Sexual Partner Contact Networks for the Transmission Dynamics of Hiv. Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology **3**:417-429.

Anderson, R. M., and R. M. May. 1979. Population biology of infectious diseases: part I. Nature **280**:361-366.

Anderson, R. M., and R. M. May. 1991. Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, Oxford.

Anderson, R. M., and W. Trewhella. 1985. Population dynamics of the badger (*Meles meles*) and the epidemiology of bovine tuberculosis (*Mycobacterium bovis*). Philosophical Transactions of the Royal Society of London B **310**:327-381.

Ball, F. 1999. Stochastic and deterministic models for SIS epidemics among a population partitioned into households. Mathematical Biosciences **156**:41-67.

Ball, F., and O. D. Lyne. 2001. Stochastic multi-type SIR epidemics among a population partitioned into households. Advances in Applied Probability **33**:99-123.

Ball, F., D. Mollison, and G. Scalia-Tomba. 1997. Epidemics with two levels of mixing. The Annals of Applied Probability **7**:46-89.

Ball, F., and P. Neal. 2002. A general model for stochastic SIR epidemics with two levels of mixing. Mathematical Biosciences **180**:73-102.

Ball, F. G., T. Britton, and O. D. Lyne. 2004. Stochastic multitype epidemics in a community of households: Estimation of threshold parameter R* and secure vaccination coverage. Biometrika **91**:345-362.

Barlow, N. D. 1991. Control of endemic bovine Tb in New Zealand possum populations: results from a simple model. Journal of Applied Ecology **28**:794-809.

Barlow, N. D., J. M. Kean, G. Hickling, P. G. Livingston, and A. B. Robson. 1997. A simulation model for the spread of bovine tuberculosis within New Zealand cattle herds. Preventive Veterinary Medicine **32**:57-75.

Bauch, C., and D. A. Rand. 2001. A moment closure model for sexually transmitted disease transmission through a concurrent partnership network. Proceedings of the Royal Society of London Series B **268**:2614-2614.

Begon, M., S. M. Hazel, S. Telfer, K. Bown, D. Carslake, R. Cavanagh, J. Chantrey, T. Jones, and M. Bennett. 2003. Rodents, cowpox virus and islands: densities, numbers and thresholds. Journal of Animal Ecology **72**:343-355.

Bejder, L., D. Fletcher, and S. Bräger. 1998. A method for testing association patterns of social animals. Animal Behaviour **56**:719-725.

Bengis, R. G. 1999. Tuberculosis in free-ranging mammals. Pages 101-114 *in* M. E. Fowler and R. E. Miller, editors. Zoo and Wild Animal Medicine. W. B. Sauders Company, Philadelphia.

Bengis, R. G., N. P. Kriek, D. F. Keet, J. P. Raath, V. De Vos, and H. Huchzermeyer. 1996. An outbreak of bovine tuberculosis in a free-living African buffalo (*Syncerus caffer* Sparrman) population in the Kruger National Park: A preliminary report. Onderstepoort Journal of Veterinary Research **63**:15-18.

Bentil, D. E., and J. D. Murray. 1993. Modelling bovine tuberculosis in badgers. Journal of Animal Ecology **62**:239-250.

Berggren, S. A. 1977. Incidence of tuberculosis in BCG vaccinated and control cattle in relation to age distribution in Malawi. British Veterinary Journal **133**:490-494.

Boots, M., P. J. Hudson, and A. Sasaki. 2004. Large shifts in pathogen virulence relate to host population structure. Science **303**:842-844.

Boots, M., and R. J. Knell. 2002. The evolution of risky behaviour in the presence of a sexually transmitted disease. Proceedings of the Royal Society of London Series B **269**:585-589.

Boots, M., and A. Sasaki. 1999. 'Small worlds' and the evolution of virulence: infection occurs locally and at a distance. Proceedings of the Royal Society of London Series B **266**:1933-1938.

Bouma, A., M. C. M. de Jong, and T. G. Kimman. 1995. Transmission of pseudorabies virus within pig-populations is independent of the size of the population. Preventive Veterinary Medicine **23**:163-172.

Brager, S., B. Wursig, A. Acevedo, and T. Henningsen. 1994. Association patterns of bottlenose dolphins (*Tursiops truncatus*) in Galvestion Bay, Texas. Journal of Mammalogy **75**:431-437.

Buddle, B. M., F. E. Aldwell, D. L. Keen, N. A. Parlane, G. F. Yates, and G. W. De Lisle. 1997. Intraduodenal vaccination of brushtail possums with bacille Calmette-Guérin enhances immune responses and protection against *Mycobacterium bovis* infection. International Journal of Tuberculosis and Lung Disease **1**:377-383.

Buddle, B. M., F. E. Aldwell, and D. N. Wedlock. 1995a. Vaccination of cattle and possums against bovine tuberculosis. Pages 113-115 *in* F. Griffin and G. W. De Lisle, editors. Tuberculosis in wildlife and domestic animals. University of Otago Press, Dunedin.

Buddle, B. M., G. W. De Lisle, A. Pfeiffer, and F. E. Aldwell. 1995b. Immunological responses and protection against *Mycobacterium bovis* in calves vaccinated with a low dose of BCG. Vaccine **13**:1123-1130.

Buddle, B. M., D. Keen, A. Thomson, G. Jowett, A. R. McCarthy, J. Heslop, G. W. De Lisle, J. L. Stanford, and F. E. Aldwell. 1995c. Protection of cattle from bovine tuberculosis by vaccination with BCG by the respiratory or subcutaneous route, but not by vaccination with killed *Mycobacterium vaccae*. Research in Veterinary Science **59**:10-16.

Buddle, B. M., M. A. Skinner, and M. A. Chambers. 2000. Immunological approaches to the control of tuberculosis in wildlife reservoirs. Veterinary Immunology and Immunopathology **74**:1-16.

Cain, M. L., R. Nathan, and S. A. Levin. 2003. Long-distance dispersal. Ecology **84**:1943-1944.

Cairns, S. J., and S. J. Schwager. 1987. A comparison of association indices. Animal Behaviour **35**:1454-1469.

Caron, A., P. C. Cross, and J. T. du Toit. 2003. Ecological implications of bovine tuberculosis in African Buffalo herds. Ecological Applications **13**:1338-1345.

Chilvers, B. L., and P. J. Corkeron. 2002. Association patterns of bottlenose dolphins (*Tursiops aduncus*) off Point Lookout, Queensland, Australia. Canadian Journal of Zoology **80**:973-979.

Christal, J., H. Whitehead, and E. Lettevall. 1998. Sperm whale social units: variation and change. Canadian Journal of Zoology **76**:1431-1440.

Colditz, G. A., T. F. Brewer, C. S. Berkey, M. E. Wilson, E. Burdick, H. V. Fineburg, and F. Mosteller. 1994. Efficacy of BCG vaccine in the prevention of tuberculosis: meta-analysis of published literature. Journal of American Medical Association **271**:698-702.

Coleman, J. D., and M. M. Cooke. 2001. Mycobacterium bovis infection in wildlife in New Zealand. Tuberculosis **81**:191-202.

Cote, I. M., and R. Poulin. 1995. Parasitism and group-size in social animals - a metaanalysis. Behavioral Ecology **6**:159-165.

Cross, P. C., and S. R. Beissinger. 2001. Using logistic regression to analyze the sensitivity of PVA models: a comparison of methods based on African Wild Dog models. Conservation Biology **15**:1335-1346.

Cross, P. C., J. O. Lloyd-Smith, J. Bowers, C. Hay, M. Hofmeyr, and W. M. Getz. 2004. Integrating association data and disease dynamics in a social ungulate: bovine tuberculosis in African buffalo in the Kruger National Park. Annales Zoologici Fennici **41**:879-892.

Cross, P. C., J. O. Lloyd-Smith, and W. M. Getz. 2005. Disentangling association patterns in fission-fusion societies using African buffalo as an example. Animal Behaviour **69**:499-506.

De Vos, V., R. G. Bengis, N. P. J. Kriek, A. Michel, D. F. Keet, J. P. Raath, and H. F. K. A. Huchzermeyer. 2001. The epidemiology of tuberculosis in free-ranging African buffalo (*Syncerus caffer*) in the Kruger National Park, South Africa. Onderstepoort Journal of Veterinary Research **68**:119-130.

Diekmann, O., and J. A. P. Heesterbeek. 2000. Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation. John Wiley & Sons Ltd., Chichester, England.

Dietz, K., and K. P. Hadeler. 1988. Epidemiological Models for Sexually-Transmitted Diseases. Journal of Mathematical Biology **26**:1-25.

Dugatkin, L. A. 1997. Cooperation Among Animals. Oxford University Press, Oxford.

Eames, K. T. D., and M. J. Keeling. 2002. Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. Proceedings of the National Academy of Sciences of the United States of America **99**:13330-13335.

Eames, K. T. D., and M. J. Keeling. 2003. Contact tracing and disease control. Proceedings of the Royal Society of London Series B-Biological Sciences **270**:2565-2571.

Eames, K. T. D., and M. J. Keeling. 2004. Monogamous networks and the spread of sexually transmitted diseases. Mathematical Biosciences **189**:115-130.

Edmunds, W. J., C. J. Ocallaghan, and D. J. Nokes. 1997. Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. Proceedings of the Royal Society of London Series B-Biological Sciences **264**:949-957.

Ellwood, D. E., and F. G. Waddington. 1972. A second experiment to challenge the resistance to tuberculosis in B.C.G. vaccinated cattle in Malawi. British Veterinary Journal **128**:619-626.

Epstein, P. R. 2002. Biodiversity, climate change, and emerging infectious diseases. Pages 27-39 *in* A. A. Aguirre, R. S. Ostfeld, G. M. Tabor, C. House, and M. C. Pearl, editors. Conservation Medicine: Ecological Health in Practice. Oxford University Press, Oxford.

Eubank, S., H. Guclu, V. S. A. Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. 2004. Modelling disease outbreaks in realistic urban social networks. Nature **429**:180-184.

Ferguson, N. M., C. A. Donnelly, and R. M. Anderson. 2001. Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. Nature **413**:542-548.

Ferguson, N. M., and G. P. Garnett. 2000. More realistic models of sexually transmitted disease transmission dynamics - Sexual partnership networks, pair models, and moment closure. Sexually Transmitted Diseases **27**:600-609.

Ferreras, P., M. Delibes, F. Palomares, J. M. Fedriani, J. Calzada, and E. Revilla. 2004. Proximate and ultimate causes of dispersal in the Iberian lynx *Lynx pardinus*. Behavioral Ecology **15**:31-40.

Fine, P. E. 1998. The effect of heterologous immunity upon the apparent efficacy of (e.g. BCG) vaccine. Vaccine **16**:1923-1928.

Francis, J. 1958. Tuberculosis in Animal and Man. A Study in Comparative Pathology. Cassell, London.

Freeland, W. J. 1976. Pathogens and the evolution of primate sociality. Biotropica **8**:12-24.

Fulford, G. R., M. G. Roberts, and J. A. P. Heesterbeek. 2002. The metapopulation dynamics of an infectious disease: Tuberculosis in possums. Theoretical Population Biology **61**:15-29.

Funston, P. J. 1999. Predator-prey relationships between lions and large ungulates in the Kruger National Park. Ph.D. University of Pretoria, Pretoria.

Gaillard, J. M., M. Festa-Bianchet, and N. G. Yoccoz. 1998. Population dynamics of large herbivores: variable recruitment with constant adult survival. Trends in Ecology and Evolution **13**:58-63.

Getz, W. M. 1996. A hypothesis regarding the abruptness of density dependence and the growth rate of populations. Ecology **77**:2014-2026.

Getz, W. M., and J. Pickering. 1983. Epidemic models: thresholds and population regulation. American Naturalist **121**:892-898.

Ginsberg, J. R., and T. P. Young. 1992. Measuring association between individuals or groups in behavioural studies. Animal Behaviour **44**:377-379.

Grimsdell, J. J. 1973. Age determination of the African buffalo, *Syncerus caffer* Sparrman. East African Journal of Wildlife **11**:31-53.

Grobler, D. G., A. L. Michel, L. De Klerk, and R. G. Bengis. 2002. The gamma-interferon test: its usefulness in a bovine tuberculosis survey in African buffaloes (*Syncerus caffer*) in the Kruger National Park. Onderstepoort Journal of Veterinary Research **69**:221-227.

Gyllenberg, M., and J. A. J. Metz. 2001. On fitness in structured metapopulations. Journal of Mathematical Biology **43**:545-560.

Hagenaars, T. J., C. A. Donnelly, and N. M. Ferguson. 2004. Spatial heterogeneity and the persistence of infectious diseases. Journal of Theoretical Biology **229**:349-359.

Halley, D. J., M. E. Vanderwalle, M. Mari, and C. Taolo. 2002. Herd-switching and long distance dispersal in female African buffalo Syncerus caffer. East African Wildlife Society **40**:97-99.

Henzi, S. P., J. E. Lycett, and S. E. Piper. 1997. Fission and troop size in a mountain baboon population. Animal Behaviour **53**:525-535.

Hess, G. 1996. Disease in metapopulation models: Implications for conservation. Ecology **77**:1617-1632.

Hulme, M., R. Doherty, T. Ngara, M. New, and D. Lister. 2001. African climate change: 1900-2100. Climate Research **17**:145-168.

Insightful Corp. 2001. S-plus for Windows Guide to Statistics, Seattle, WA.

Isbell, L. A., and D. Van Vauren. 1996. Differential costs of locational and social dispersal and their consequences for female group-living primates. Behaviour **133**:1-36.

Jolles, A. 2004. Disease ecology of tuberculosis in African buffalo. Ph.D. Thesis. Princeton University.

Jolly, A. M., and J. L. Wylie. 2002. Gonorrhoea and chlamydia core groups and sexual networks in Manitoba. Sexually Transmitted Infections **78**:I145-I151.

Kao, R. R., and M. G. Roberts. 1999. A comparison of wildlife control and cattle vaccination as methods for the control of bovine tuberculosis. Epidemiology and Infection **122**:505-519.

Kao, R. R., M. G. Roberts, and T. J. Ryan. 1997. A model of bovine tuberculosis control in domesticated cattle herds. Proceedings of the Royal Society of London B **264**:1069-1076.

Keeling, M. J. 1999. The effects of local spatial structure on epidemiological invasions. Proceedings of the Royal Society of London Series B **266**:859-867.

Keeling, M. J., and C. A. Gilligan. 2000. Metapopulation dynamics of bubonic plague. Nature **407**:903-906.

Keeling, M. J., and B. T. Grenfell. 2000. Individual-based perspectives on R-0. Journal of Theoretical Biology **203**:51-61.

Keeling, M. J., and P. Rohani. 2002. Estimating spatial coupling in epidemiological systems: a mechanistic approach. Ecology Letters **5**:20-29.

Keeling, M. J., M. E. J. Woolhouse, R. M. May, G. Davies, and B. T. Grenfell. 2003. Modelling vaccination strategies against foot-and-mouth disease. Nature **421**:136-142.

Keet, D. F., N. P. Kriek, M. L. Penrith, A. Michel, and H. Huchzermeyer. 1996. Tuberculosis in buffaloes (*Syncerus caffer*) in the Kruger National Park: spread of the disease to other species. Onderstepoort Journal of Veterinary Research **69**:239-244.

Kermack, W. O., and A. G. McKendrick. 1927. Contributions to the mathematical theory of epidemics. Proceedings of the Royal Society of Edinburgh **115**:700-721.

Kermack, W. O., and A. G. McKendrick. 1991. Contributions to the Mathematical-Theory of Epidemics .1. (Reprinted from Proceedings of the Royal Society, Vol 115a, Pg 700-721, 1927). Bulletin of Mathematical Biology **53**:33-55.

Koenig, W. D., D. VanVuren, and P. N. Hooge. 1996. Detectability, philopatry, and the distribution of dispersal distances in vertebrates. Trends in Ecology & Evolution **11**:514-517.

Krebs, J. R., R. M. Anderson, T. Clutton-Brock, I. Morrison, D. Young, and C. A. Donnelly. 1997. Bovine Tuberculosis in Cattle and Badgers. Ministry of Agriculture, Fisheries and Food, London.

Kretzschmar, M. 2000. Sexual network structure and sexually transmitted disease prevention - A modeling perspective. Sexually Transmitted Diseases **27**:627-635.

Kretzschmar, M., and M. Morris. 1996. Measures of concurrency in networks and the spread of infectious disease. Mathematical Biosciences **133**:165-195.

Liljeros, F., C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Aberg. 2001. The web of human sexual contacts. Nature **411**:907-908.

Lloyd-Smith, J. O., A. P. Galvani, and W. M. Getz. 2003. Curtailing transmission of severe acute respiratory syndrome within a community and its hospital. Proceedings of the Royal Society of London Series B **270**:1979-1989.

Lloyd-Smith, J. O., W. M. Getz, and H. V. Westerhoff. 2004. Frequency-dependent incidence in models of sexually transmitted diseases: portrayal of pair-based transmission and effects of illness on contact behaviour. Proceeding of the Royal Society of London Series B **271**:625-634.

Lott, D. F., and S. C. Minta. 1983. Random individual association and social group instability in American bison (*Bison bison*). Zeitschrift fur Tierpsychologie **61**:153-172.

May, R. M., and R. M. Anderson. 1979. Population biology of infectious diseases: part II. Nature **280**:455-461.

McCallum, H., N. Barlow, and J. Hone. 2001. How should pathogen transmission be modelled? Trends in Ecology and Evolution **16**:295-300.

McCarthy, M. A., M. A. Burgman, and S. Ferson. 1995. Sensitivity analysis for models of population viability. Biological Conservation **75**:93-100.

Metz, J. A. J., and M. Gyllenberg. 2001. How should we define fitness in structured metapopulation models? Including an application to the calculation of evolutionarily stable dispersal strategies. Proceedings of the Royal Society of London Series B **268**:499-508.

Meyers, L. A., M. E. J. Newman, M. Martin, and S. Schrag. 2003. Applying network theory to epidemics: Control measures for Mycoplasma pneumoniae outbreaks. Emerging Infectious Diseases **9**:204-210.

Mloszewski, M. J. 1983. The behavior and ecology of the African buffalo. Cambridge University Press, Cambridge.

Moller, A. P., R. Dufva, and K. Allander. 1993. Parasites and the evolution of host social-behavior. Pages 65-102 *in* Advances in the Study of Behavior, Vol 22.

Morris, M., and M. Kretzschmar. 1997. Concurrent partnerships and the spread of HIV. AIDS **11**:641-648.

Myers, J. P. 1983. Space, time and pattern of individual associations in a group-living species: sanderlings have no friends. Behavioral Ecology and Sociobiology **12**:129-134.

Nathan, R., G. Perry, J. T. Cronin, A. E. Strand, and M. L. Cain. 2003. Methods for estimating long-distance dispersal. Oikos **103**:261-273.

Neill, S. D., J. J. O'Brien, and J. Hanna. 1991. A mathematical model for *Mycobacterium bovis* excretion from tuberculous cattle. Veterinary Microbiology **28**:103-109.

Newman, M. E. J. 2002. Spread of epidemic disease on networks. Physical Review E **66**.

Newman, M. E. J. 2003. The structure and function of complex networks. Siam Review **45**:167-256.

Nunn, C. L. 2002. A comparative study of leukocyte counts and disease risk in primates. Evolution **56**:177-190.

Nunn, C. L., S. Altizer, K. E. Jones, and W. Sechrest. 2003a. Comparative tests of parasite species richness in primates. American Naturalist **162**:597-614.

Nunn, C. L., J. L. Gittleman, and J. Antonovics. 2000. Promiscuity and the primate immune system. Science **290**:1168-1170.

Nunn, C. L., J. L. Gittleman, and J. Antonovics. 2003b. A comparative study of white blood cell counts and disease risk in carnivores. Proceedings of the Royal Society of London Series B-Biological Sciences **270**:347-356.

Park, A. W., S. Gubbins, and C. A. Gilligan. 2001. Invasion and persistence of plant parasites in a spatially structured host population. Oikos **94**:162-174.

Park, A. W., S. Gubbins, and C. A. Gilligan. 2002. Extinction times for closed epidemics: the effects of host spatial structure. Ecology Letters **5**:747-755.

Patz, J. A., and M. Khaliq. 2002. Global Climate Change and Health: Challenges for Future Practitioners. JAMA **287**:2283-2284.

Pienaar, U. 1969. Observations of developmental biology, growth and some aspects of the population ecology of African buffalo in Kruger National Park. Koedoe **12**:29-52.

Plowright, W. 1982. The effects of rinderpest and rinderpest control on wildlife in Africa. Symp. Zool. Soc. Lond. **50**:1-28.

Prins, H. H. 1996. Ecology and behaviour of the African buffalo. Chapman and Hall, London.

Read, J. M., and M. J. Keeling. 2003. Disease evolution on networks: the role of contact structure. Proceedings of the Royal Society of London Series B-Biological Sciences **270**:699-708.

Roberts, M. G. 1996. The dynamics of bovine tuberculosis in possum populations, and its eradication or control by culling or vaccination. Journal of Animal Ecology **65**:451-464.

Rodrigues, L. C., V. K. Diwan, and J. G. Wheeler. 1993. Protective effect of BCG against tuberculous meningitis and miliary tuberculosis: a meta-analysis. International Journal of Epidemiology **22**:1154-1158.

Rodwell, T. C. 2000. Epidemiology of bovine tuberculosis in the buffalo (*Syncerus caffer*) of the Kruger National Park. University of California, Davis.

Rodwell, T. C., N. P. Kriek, R. G. Bengis, I. J. Whyte, P. C. Viljoen, V. De Vos, and W. M. Boyce. 2000. Prevalence of bovine tuberculosis in African buffalo at Kruger National Park. Journal of Wildlife Diseases **37**:258-264.

Rodwell, T. C., I. J. Whyte, and W. M. Boyce. 2001. Evaluation of population effects of bovine tuberculosis in free-ranging African buffalo (*Syncerus caffer*). Journal of Mammalogy **82**:231-238.

Romesburg, H. C. 1984. Cluster Analysis for Researchers. Lifetime Learning Publications, Belmont, CA.

Ruxton, G. D. 1996. The effects of stochasticity and seasonality on model dynamics: bovine tuberculosis in badgers. Journal of Animal Ecology **65**:495-500.

Selvin, S. 1995. Practical Biostatistical Methods. Wadsworth Publishing Company, Belmont.

Sinclair, A. R. E. 1977. The African Buffalo. University of Chicago Press, Chicago.

Smith, G. C. 2001. Models of Mycobacterium bovis in wildlife and cattle. Tuberculosis **81**:51-64.

Smith, G. C., and C. L. Cheeseman. 2002. A mathematical model for the control of diseases in wildlife populations: culling, vaccination and fertility control. Ecological Modelling **150**:45-53.

Smith, G. C., C. L. Cheeseman, and D. Wilkinson. 2001. A model of bovine tuberculosis in the badger Meles meles: the inclusion of cattle and the use of a live test. Journal of Applied Ecology **38**:520-535.

Smolker, R. A., A. F. Richards, R. C. Connor, and J. W. Pepper. 1992. Sex differences in patterns of association among Indian Ocean bottlenose dolphins. Behaviour **123**:38-69.

Stanko, M., D. Miklisova, J. G. de Bellocq, and S. Morand. 2002. Mammal density and patterns of ectoparasite species richness and abundance. Oecologia **131**:289-295.

Suazo, F. M., A. M. A. Escalera, and R. M. G. Torres. 2003. A review of M-bovis BCG protection against TB in cattle and other animals species. Preventive Veterinary Medicine **58**:1-13.

Swinton, J., J. Harwood, B. T. Grenfell, and C. A. Gilligan. 1998. Persistence thresholds for phocine distemper virus infection in harbour seal *Phoca vitulina* metapopulations. Journal of Animal Ecology **67**:54-68.

Szykman, M., A. L. Engh, R. C. Van Horn, S. M. Funk, K. T. Scribner, and K. E. Holekamp. 2001. Association patterns among male and female spotted hyenas (*Crocuta crocuta*) reflect male mate choice. Behavioral Ecology and Sociobiology **50**:231-238.

Tella, J. L. 2002. The evolutionary transition to coloniality promotes higher blood parasitism in birds. Journal of Evolutionary Biology **15**:32-41.

Thrall, P. H., J. Antonovics, and A. P. Dobson. 2000. Sexually transmitted diseases in polygynous mating systems: prevalence and impact on reproductive success. Proceedings of the Royal Society of London Series B **267**:1555-1563.

Tuyttens, F. A., and D. W. MacDonald. 1998. Sterilization as an alternative strategy to control wildlife disease: Bovine tuberculosis in European badgers as a case study. Biodiversity and Conservation **7**:705-723.

Waddington, F. G., and D. C. Ellwood. 1972. An experiment to challenge the resistance to tuberculosis in BCG vaccinated cattle in Malawi. British Veterinary Journal **128**:541-552.

Wallinga, J., W. J. Edmunds, and M. Kretzschmar. 1999. Perspective: human contact patterns and the spread of airborne infectious diseases. Trends in Microbiology **7**:372-377.

Waser, P. M., S. R. Creel, and J. R. Lucas. 1994. Death and disappearance - estimating mortality risks associated with philopatry and dispersal. Behavioral Ecology **5**:135-141.

Watts, D. J. 1999. Small Worlds: the Dynamics of Networks Between Order and Randomness. Princeton University Press, Princeton.

Weinrich, M. T. 1991. Stable social associations among humpback whales (*Megaptera novaeangliae*) in the southern Gulf of Maine. Canadian Journal of Zoology **69**:3012-3018.

White, G. C., and K. D. Burnham. 1999. Program MARK: survival estimation from populations of marked animals. Bird Study **46**:S120-S139.

White, P. C., and S. Harris. 1995. Bovine tuberculosis in badger (*Meles meles*) populations in southwest England: An assessment of past, present and possible future control strategies using simulation modelling. Philosophical Transactions of the Royal Society of London B **349**:415-432.

Whitehead, H. 1999. Testing association patterns of social animals. Animal Behaviour **57**:F26-F29.

Whitehead, H., and S. DuFault. 1999. Techniques for analyzing vertebrate social structure using identified individuals: review and recommendations. Advances in the Study of Behavior **28**:33-74.

Whitehead, H., S. Waters, and T. Lyrholm. 1991. Social organization of female sperm whales and their offspring: constant companions and casual acquaintances. Behavioral Ecology and Sociobiology **29**:385-389.

Wilkinson, D., G. C. Smith, R. J. Delahay, and C. L. Cheeseman. 2004. A model of bovine tuberculosis in the badger Meles meles: an evaluation of different vaccination strategies. Journal of Applied Ecology **41**:492-501.

Wilson, E. B., and J. Worcester. 1945. The spread of an epidemic. Proceedings of the National Academy of Sciences **31**:327-333.

Wisdom, M. J., and L. S. Mills. 1997. Sensitivity analysis to guide population recovery: prairie-chickens as an example. Journal of Wildlife Management **61**:302-312.

Wisdom, M. J., L. S. Mills, and D. F. Doak. 2000. Life-stage simulation analysis: estimating vital-rate effects on population growth for conservation. Ecology **81**:628-641.

Wood, P. R., and S. L. Jones. 2001. BOVIGAM^{TM}: an in vitro cellular diagnostic test for bovine tuberculosis. Tuberculosis **81**:147-155.

Woodhouse, D. E., R. B. Rothenberg, J. J. Potterat, W. W. Darrow, S. Q. Muth, A. S. Klovdahl, H. P. Zimmerman, H. L. Rogers, T. S. Maldonado, J. B. Muth, and J. U. Reynolds. 1994. Mapping a Social Network of Heterosexuals at High-Risk for Hiv-Infection. AIDS **8**:1331-1336.

Woodroffe, R., S. D. Frost, and R. S. Clifton-Hadley. 1999. Attempts to control tuberculosis in cattle by removing infected badgers: constraints imposed by live test sensitivity. Journal of Applied Ecology **36**:494-501.