CONTACT RATES IN ECOLOGY: USING PROXIMITY LOGGERS TO EXPLORE DISEASE TRANSMISSION ON WYOMING'S ELK FEEDGROUNDS

by

Tyler Graydon Creech

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Biological Sciences

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2011

APPROVAL

of a thesis submitted by

Tyler Graydon Creech

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citation, bibliographic style, and consistency and is ready for submission to The Graduate School.

Dr. Scott Creel

Dr. Paul Cross

Approved for the Department of Ecology

Dr. David Roberts

Approved for The Graduate School

Dr. Carl A. Fox

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Tyler Graydon Creech

April 2011

ACKNOWLEDGMENTS

I am deeply grateful to Paul Cross and Scott Creel for their expert guidance during my time at MSU, and to Jay Rotella for serving on my committee. The contributions of many people at the Wyoming Game and Fish Department made this research possible; I would particularly like to recognize Brandon Scurlock, Eric Maichak, Jared Rogerson, John Henningsen, Amy Williams, and Kate Belinda. Special thanks are due to Mike Ebinger for his many efforts on behalf of this project. I appreciate the friendship and insightful comments of Angela Brennan, Vicki Patrek, Paul Schuette, Jake Ferguson, Jim Junker, and Dan Drinan. Lastly, I would like to thank my family for their support and encouragement.

Funding for this research was provided by the U.S. Geological Survey, Wyoming Game and Fish Department, Wyoming Wildlife Livestock Disease Partnership, and National Science Foundation and National Institutes of Health Ecology of Infectious Disease Grant DEB-1067129.

TABLE OF CONTENTS

TABLE OF CONTENTS - CONTINUED

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Infectious diseases are an important consideration in the management of wildlife populations, and contact rate is a key parameter for understanding the epidemiology of such diseases. In the first section of this thesis, I review current issues and challenges that researchers face when designing animal contact studies and analyzing contact data. I examine how characteristics of methods for collecting contact data affect inferences that can be drawn about contact structures; describe applications of social network analysis of contact data to disease ecology and animal behavior, focusing on sampling issues and dynamic networks; suggest how new technologies can be used to answer important questions about variation in individual contact rates within populations; and propose a new statistical approach for analyzing contact data in a linear modeling framework.

In the second section, I describe an experimental field study that used proximity loggers (a new technology for measuring contact rates) to understand transmission of *Brucella abortus* on elk feedgrounds in Wyoming. Brucellosis is a bacterial disease that causes abortions in elk and is transmitted by contact with infectious aborted fetuses. Supplemental feeding of elk on winter feedgrounds is believed to exacerbate *B. abortus* transmission by aggregating elk at high densities, increasing their chance of contacting infectious fetuses. I evaluated the effectiveness of a proposed low-density feeding strategy by comparing elk-fetus contact rates (as measured by proximity collars and video cameras) during high-density and low-density feeding treatments that provided the same total amount of food at different densities. Low-density feeding led to >50 percent reductions in the total number of contacts and the number of individuals contacting a fetus. Elk contacted fetuses and random control points equally, suggesting that elk were not attracted to fetuses but encountered them incidentally while feeding. The relationship between contact rate and disease prevalence is non-linear and simple disease models suggest that low-density feeding may result in dramatic reductions in brucellosis prevalence, though this depends on the amount of transmission that occurs on and off feedgrounds

INTRODUCTION

Wildlife and Disease

Infectious diseases present a formidable challenge to the conservation and management of wildlife populations worldwide. Infectious diseases are known to effect survival, reproduction, movement patterns, genetic structure, and age structure of wildlife populations (Scott 1988). In extreme cases, diseases have even threatened the persistence of entire species; for instance, the black-footed ferret (*Mustela nigripes*) was extirpated from the wild in the mid-1980s by an outbreak of canine distemper (Thorne and Williams 1988). In recent years, the rapid decline of thousands of amphibian species has been linked to infectious pathogens such as the chytrid fungus *Batrachochytrium dendrobatidis* (Daszak et al. 2003, Pounds et al. 2006).

Conservation motives aside, we should be concerned about wildlife disease because many infectious diseases of humans originate in or are maintained by animal populations ("zoonotic" diseases), such as HIV and lyme disease. Jones et al. (2008) found that 60 percent of emerging infectious diseases are zoonotic, the majority of which originate in wildlife populations (versus domestic animals). Thus, controlling infectious disease in wildlife is necessary to protect the health of animal and human populations.

Contact Rates in Ecology

The study of patterns and causes of disease in populations is known as epidemiology, and one of the central epidemiological parameters is contact rate.

Depending on the disease of interest, the relevant contact rate may be either the rate at which host individuals contact each other (for directly transmitted diseases like influenza), or the rate at which host individuals contact infectious materials in their environment (for indirectly transmitted diseases like anthrax). Contact rate is typically assumed to be proportional to transmission rate in disease models (McCallum et al. 2001), and because contacts are far easier to recognize and observe than transmission events, contact rates are a fundamental aspect of epidemiology.

Understanding the patterns of contact underlying disease transmission is critically important if we are to control the impacts of zoonotic disease on both human and wildlife populations. Knowledge of contact rates allows us to: 1) identify environmental conditions and other external factors (e.g., season or time of day) associated with high rates of contact; 2) determine which characteristics of individuals are associated with high contact rates; 3) parameterize models that help us predict the dynamics and understand the drivers of disease transmission; and 4) utilize all of these types of information to design strategies to minimize impacts of disease.

While particularly relevant to the study of epidemiology, contact rates have applications in many other scientific disciplines such as ecology, wildlife biology, animal behavior, and even sociology. Researchers from these diverse fields use contact data to answer different types of questions; a disease ecologist, for example, may want to know which individuals in a population are most likely to become infected and spread infection to others, while an animal behaviorist may be interested in how the rate of cooperative interaction varies with kinship. Still, the technological and statistical tools used to collect

and analyze contact data are often remarkably similar and adaptable across disciplines. A wealth of such tools exist thanks to many decades of research, and the pace of development has accelerated as new technologies have made it possible to collect large amounts of contact data with unprecedented spatial and temporal resolution. For a researcher in any of the aforementioned fields, it can be challenging to keep abreast of relevant developments in related fields and to avoid spending time tackling problems that may have already been addressed elsewhere. As an example, one of the most popular techniques for analyzing contact data is social network analysis (SNA), which explicitly accounts for the links between specific individuals or groups; SNA evolved along largely independent paths in the fields of sociology and physics (where it is known as "graph theory"), resulting in duplicated effort and multiple sets of scientific jargon describing identical concepts in different contexts. Yet even among more closely related disciplines that are concerned only with animal contacts, such problems of duplication and poor communication can occur.

Clearly, there is great need for the technological and statistical tools developed for studying animal contacts in one context (e.g., disease ecology) to be shared with and accessible to those working in other contexts (e.g., animal behavior). Consequently, the first manuscript in this thesis reviews recent developments in the use of animal contact data in ecological fields, with particular attention to disease ecology and animal behavior. I explore how characteristics of methods for collecting contact data can limit inference about contact structures, discuss challenges associated with sampling individuals and

contacts from social networks, and describe a new statistical approach for analyzing animal contact data using generalized linear mixed models.

Brucellosis and Supplemental Feeding

One of the most promising new technologies highlighted in the first manuscript is proximity loggers, which provide contact data by transmitting and receiving unique radio frequencies between logging devices in close proximity. Proximity loggers are affixed to individual animals, usually in the form of radiocollars, and collect high-resolution data on contacts with other loggers (which could be affixed to other individuals or associated with stationary environmental features such as infectious materials). In the second manuscript of this thesis, I use proximity loggers to study one of the most intractable wildlife disease issues in the Greater Yellowstone Ecosystem (GYE): the transmission of *Brucella abortus* bacteria among elk (*Cervus elaphus*) on supplemental feedgrounds in western Wyoming.

B. abortus is the causative agent of brucellosis, a disease that has infected elk and bison (*Bison bison*) populations in the GYE for nearly a century (Meagher and Meyer 1994, Murie 1951). Infected female elk and bison typically abort during their first pregnancy following infection and occasionally during subsequent pregnancies, and contact with infectious abortion materials (i.e., fetuses, placentas, and fetal fluids) is the primary route of disease transmission (Thorne et al. 1978). Because it can be transmitted to livestock and humans with serious health consequences, brucellosis is a disease of major concern in the GYE. Over $3.5 billion has been spent since 1934 in a federal

campaign to eradicate brucellosis from the U.S. (Cheville et al. 1998), but the disease maintains a strong foothold in the GYE despite being eliminated from the rest of the nation.

It is generally believed that supplemental feeding of elk in winter (a practice that began in western Wyoming in 1910 to compensate for loss of native winter range and minimize conflicts on agricultural lands; Smith 2001) is at least partly responsible for the high prevalence of brucellosis in the region. By aggregating elk at unnaturally high densities, feedgrounds increase the chance of elk contacting infectious aborted fetuses in their environment. State wildlife managers have proposed a new strategy to reduce disease-transmitting elk-fetus contacts on feedgrounds: low-density feeding, in which feed is distributed across an increased area to reduce elk densities along feedlines. The effectiveness of this strategy remains untested, however, and if elk are strongly attracted to fetuses, low-density may do little to prevent such contacts.

I use proximity loggers to test whether low-density feeding is likely to reduce *B. abortus* transmission on feedgrounds by recording and comparing rates of elk-fetus contact during periods of low-density and high-density (i.e., traditional) supplemental feeding. I also explore the variation in elk-fetus contact rates that exists among individuals in the population with the objectives of identifying classes of individuals (e.g., age or sex) that have high contact rates and using this information to focus disease control efforts on such high-risk individuals. Finally, I use observed contact rates during high- and low-density feeding to model the potential effects of broad implementation of low-density feeding on the prevalence of brucellosis in the region.

Literature Cited

Cheville, N.F., D.R. McCullough, and L.R. Paulson. 1998. Brucellosis in the Greater Yellowstone Area. National Academy Press, Washington, D.C., USA.

Daszak, P., A.A. Cunningham, and A.D. Hyatt. 2003. Infectious disease and amphibian population declines. Diversity and Distributions 9:141-150.

Jones, K.E., N.G. Patel, M.A. Levy, A. Storeygard, D. Balk, J.L. Gittleman, and P. Daszak. 2008. Global trends in emerging infectious diseases. Nature 451:990-994.

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

Meagher, M., and M.E. Meyer. 1994. On the origin of brucellosis in bison of Yellowstone National Park - a review. Conservation Biology 8:645-653.

Murie, O.J. 1951. The elk of North America. Stackpole, Harrisburg, Pennsylvania, USA and the Wildlife Management Institute, Washington, D.C, USA.

Pounds, J.A., M.R. Bustamante, L.A. Coloma, J.A. Consuegra, M.P.L. Fogden, P.N. Foster, E. La Marca, K.L. Masters, A. Merino-Viteri, R. Puschendorf, S.R. Ron, G.A. Sanchez-Azofeifa, C.J. Still, and B.E. Young. 2006. Widespread amphibian extinctions from epidemic disease driven by global warming. Nature 439:161-167.

Scott, M.E. 1988. The impact of infection and disease on animal populations: implications for conservation biology. Conservation Biology 2:40-56.

Smith, B.L. 2001. Winter feeding of elk in western North America. Journal of Wildlife Management 65:173-190.

Thorne, E.T., J.K. Morton, F.M. Blunt, and H.A. Dawson. 1978. Brucellosis in elk. II. Clinical effects and means of transmission as determined through artificial infections. Journal of Wildlife Diseases 14:280-291.

Thorne, E.T., and E.S. Williams. 1988. Disease and endangered species: the black-footed ferret as a recent example. Conservation Biology 2:66-74.

CONTACTS BETWEEN ANIMALS: EMERGING METHODS, QUESTIONS, AND CHALLENGES

Contributions of Authors and Co-Authors

Manuscript 1

Author: Tyler G. Creech
Contributions: conducted literature review of animal contact studies and prepared manuscript.

Co-author: Scott Creel
Contributions: provided conceptual input and comments on manuscript.

Co-author: Paul C. Cross
Contributions: provided conceptual input and comments on manuscript.

Manuscript Information Page

Manuscript 1

Title: Contacts Between Animals: Emerging Methods, Questions, and Challenges

Authors: Tyler G. Creech, Scott Creel, and Paul C. Cross

Journal: Methods in Ecology and Evolution

Status of manuscript:

__x__ Prepared for submission to a peer-reviewed journal
_____ Officially submitted to a peer-reviewed journal
_____ Accepted by a peer-reviewed jour
_____ Published in a peer-reviewed journal

Anticipated date of submission: April 2011

Summary

1. Animal contact rates are central to several ecological fields, most notably animal behavior and disease ecology.
2. Recent technological advances such as proximity loggers allow researchers to collect contact data from wild populations with unprecedented spatial and temporal resolution.
3. Animal contact data are increasingly analyzed using the social network analysis framework, but this approach has often failed to capture the dynamic nature of contact networks. Some inferences drawn from observed networks are questionable when only a fraction of the individuals in the true network is sampled, as in most studies of wild populations.
4. Variation in contact rate among individuals is a common feature of animal populations and can have large effects on disease transmission. Past research has sometimes attributed such variation to inherent differences among individuals, but we believe that the environmental context of contacts can be an equally important source of variation in contact rates.
5. We present a new approach for analyzing contact data of the type provided by proximity loggers that utilizes generalized linear mixed models to assess the individual, dyadic, and environmental factors contributing to variation in contact rates among individuals.

Introduction

The rate of contact between individuals in animal populations is a key parameter in several fields of ecology, including animal behavior and disease dynamics. Contact rate is often assumed to be directly proportional to transmission rate of infectious diseases (McCallum *et al.* 2001), and for studies of animal behavior, the type and frequency of interactions among individuals are the most basic description of social structure. This review identifies current issues and challenges that researchers face when designing animal contact studies and analyzing contact data. We examine how specific characteristics of contact data collected with different methods affect inferences that can be drawn about contact structures; describe applications of social network analysis to

contact data in disease ecology and animal behavior, focusing on sampling issues and dynamic networks; explore how new technologies can be used to answer important (but rarely addressed) questions about variation in individual contact rates within populations; discuss statistical challenges associated with analyzing contact data; and propose a new framework for analyzing contact data that assesses the individual, dyadic, and environmental factors contributing to the variation in contact rates. In particular, we aim to provide guidance on appropriate sampling methods in animal contact studies, to spur researchers to consider how their inferences may be affected by incomplete contact data, and to suggest new avenues of research emerging from technological and statistical advances.

Methods for Collecting Contact Data

Common methods for estimating contact rates include direct observation of animal groups, VHF and GPS telemetry, and mark-recapture studies (see Prange *et al.* [2006] and Real & Biek [2007] for information on contact studies using these methods and others). These methods have provided a wealth of contact data that has advanced our understanding of disease ecology and animal behavior, but the data often have poor spatial and temporal resolution, which has limited the questions that can be addressed. Recently, proximity loggers have emerged as a powerful tool for estimating contact rates. Proximity loggers use ultra-high frequency (UHF) transceivers to continuously record contacts between individuals within a user-specified distance (currently adjustable from 0.5 to 100 m). The date, time of initiation, duration, and unique ID codes of loggers

involved in each contact are stored to memory. Proximity loggers have been used to study intra-specific contact rates of brushtail possums (Ji *et al.* 2005), European wild rabbits (Marsh *et al.* 2010), Tasmanian devils (Hamede *et al.* 2009), and elk (T. Creech, unpublished data), and inter-specific contact rates between European badgers and cattle (Bohm *et al.* 2009). Because they provide a complete record of contacts between sampled individuals over long time periods, we expect that use of proximity loggers in disease ecology and animal behavior studies will increase; accordingly, we pay particular attention to new issues and opportunities presented by proximity loggers in this review.

Implications for Inference

To select appropriate methods for animal contact studies, it is critical to understand how the characteristics of contact data limit the questions that can be answered and the inferences that can be drawn. We identify key attributes of contact data, compare methods with respect to these attributes (Table 1-1), and discuss their implications for inference:

Temporal and Spatial Resolution: The frequency at which contacts are observed is jointly determined by the true rate of interaction and either researcher effort (e.g., frequency of direct observations or radio-tracking) or technological limitations (e.g., GPS fix rate). Coarse temporal resolution may be sufficient to detect stable social patterns, such as membership in stable groups or factors that consistently affect contact rates, such as age or sex (Lusseau & Newman 2004). Where rare contacts are important, for example in studies of disease dynamics, fine temporal resolution may be necessary and the majority of methods may fail to detect such events. For instance, infrequent inter-group

contacts that allow disease to spread through a population may be detected by proximity loggers but not traditional VHF telemetry (with typical intervals between fixes), GPS telemetry (because a small proportion of individuals are monitored), or direct observation (if observations are not extensive).

The minimum distance between individuals at which contact can be assessed also limits inference. For behavioral questions, the interaction of interest will define the appropriate spatial resolution – group membership, for instance, could be assessed at a distance of many meters, but if one were interested in grooming behavior, data would need to reflect direct contact. For questions about disease transmission, fine spatial resolution will often be required to make strong inferences about risk of transmission between individuals, although the exact scale depends on the mechanism of transmission (e.g., sexually-transmitted versus airborne disease, or pathogens that persist well outside the host versus those that do not).

Detectability Bias: Across methods, the ability to detect contacts may vary by season, time of day, or even irregularly for some methods. Direct observations and aerial radio-tracking are difficult or impossible at night or in conditions of poor visibility, which could bias estimates if contact rate covaries in time with sampling rate. If a species is either diurnal or nocturnal, and tends to aggregate or disaggregate during periods of activity or inactivity, direct observation or daytime telemetry is likely to present an incomplete description of contact patterns. Species may also exhibit seasonal variation in contact structure associated with changing resource distribution, activity levels, or

aggregation patterns (Altizer *et al.* 2006), which may be problematic if contact data are obtained only in some seasons, regardless of method.

In some cases, the ability to detect contacts may also vary spatially. For many species, we would expect that contact rates vary with the habitat type occupied, particularly if group sizes are correlated with habitat variables (e.g., Creel & Winnie 2005; Fortin *et al.* 2009); if observing contacts in certain habitat types is difficult or impossible, then contact data derived from observation may not accurately describe true contact rates.

Interactions Recorded: Methods vary widely in the type of information that is recorded (i.e., what constitutes a contact). In the best case, direct visual observations allow researchers to distinguish between types of interactions (e.g., grooming vs. territorial defense), to determine the direction of interactions (e.g., *A* grooming *B*, *B* grooming *A*, or mutual grooming), and to compare contact patterns based on different types of interactions. In the worst case, GPS and VHF telemetry reveal only that two individuals were in the same area at the same time, and often at coarse spatial and temporal resolution. For GPS telemetry in particular, a 'contact' is likely to be defined as two individuals being located less than a defined distance apart on the same day (or perhaps some shorter time interval), even if neither was aware of the other or affected by its presence. Behavioral studies in which specific types interactions are of interest will generally require direct observations, but the requirements of disease studies may be more flexible, depending on the host species and pathogen of interest. If proximity is a strong predictor of disease transmission risk, then telemetry and proximity loggers will be

suitable methods, but if a specific type of direct contact is necessary for contact (e.g., sexual contact), direct observation will be required.

Sampling of Individuals: Quantifying animal contact rates usually requires sampling some subset of individuals from the population of interest. Methods that require capturing and outfitting individuals with recording devices (proximity loggers, GPS or VHF telemetry) typically limit researchers to sampling a small fraction of the total population of interest because of the costs associated with purchasing and deploying these devices. Direct observation methods often allow a much greater fraction of the population to be sampled over the course of a study. The implications of proportion of the population sampled (hereafter referred to as "sampling intensity") are rarely discussed explicitly in animal contact studies, but may be critical when population-level inferences are desired. If contact rates are highly variable among individuals (which is often true; e.g., Creel *et al.* 1992, Woolhouse *et al.* 1997, Clay *et al.* 2009), careful stratification of sampled individuals is needed to obtain representative data from a small proportion of the population, and failing to sample individuals that play key roles in social structure or disease transmission can lead to incorrect conclusions. Such stratification presumably requires a priori knowledge of those characteristics that are correlated with contact rate, which may not be obvious.

Spatial Considerations in Proximity Logger Studies

Proximity loggers record contact data with very high temporal resolution, but provide no information about the context of contacts unless they are paired with GPS collars or direct observations. This can present a challenge when analyzing differences in

contact rates among sampled individuals. In many cases it will be misleading to assess variation among individuals in contact rate by comparing raw numbers of contacts for each sampled individual, because many populations are spatially structured such that some sampled individuals spend more time in the vicinity of other sampled individuals than others (and thus have greater opportunity for contacts to be recorded, regardless of true contact rate). To account for this, it will be necessary to have regular information on the spatial distribution of sampled individuals (i.e., which sampled individuals are grouped together). This could be a daunting task if individuals must be located in real-time using aerial surveys or ground-based tracking, but could be accomplished relatively easily by pairing proximity loggers with GPS collars; researchers could then retrospectively determine when sampled individuals were within some threshold distance at which interaction is possible. Alternatively, two proximity loggers (with different UHF frequencies) could be placed on each sampled individual – one logger calibrated to record proximity at coarse spatial resolution to determine which other individuals are in the vicinity, and the other logger set at fine spatial resolution to record contacts.

Controlling for the spatial distribution of loggers will be necessary when examining questions about social interaction or variation in contact rate among individuals, but for some purposes spatial distribution may be a non-issue. If we are simply interested in whether individual *A* is likely to contract an infection from individual *B*, for instance, it doesn't matter whether a low rate of contact between *A* and *B* is due to rarely being in the same neighborhood, or due to a low rate of interaction despite being in the same neighborhood. There are also two population structures that simplify the issue

of logger spatial distribution: 1) populations with group structures that are extremely stable, so that the initial distribution of sampled individuals remains for the period of interest, and 2) populations that are so well-mixed that all individuals have the opportunity to contact each other regularly within the period of interest.

Sampling Issues in Social Network Analysis

Social network analysis (SNA) has become a common method for analyzing patterns of contact in animal populations. Social networks represent individuals as nodes and the connections between them as edges, and use formal metrics to describe network properties at the level of the individual or the entire network (see Table 1-2 for definitions of network terminology). Because it explicitly accounts for both direct and indirect connections between individuals, SNA can be more informative than simpler analysis frameworks that consider only contacts between immediate neighbors. Recent studies have used SNA to explore a variety of topics in epidemiology and animal behavior, such as the evolution of cooperation, the role of policing behavior in social stability, seasonal variation in social structure, and the link between social status and parasite infection (Table 1-3). However, increasing use of SNA in ecology has brought with it new statistical and conceptual challenges, particularly for applications to animal populations.

Describing Contact Networks: Sampling Strategies and Tradeoffs

Virtually all descriptions of animal social networks (especially for wild populations) are incomplete because sampling all individuals in a population with high

frequency is rarely possible. Sampling strategies that maximize the proportion of individuals sampled from a population typically must minimize the frequency at which each sampled individual's contacts are observed (i.e., low temporal resolution), and vice versa. Low temporal resolution generally results in omission of edges, while incomplete sampling of individuals results in omission of nodes and the edges that would have been associated with them. Thus, researchers face a tradeoff between accurately describing network topology (i.e., the arrangement of nodes and edges in the network) and accurately estimating dyadic contact rates (i.e., the presence/absence or weight of edges). Choosing how to allocate resources to these competing interests is a critical consideration in animal contact studies, but guidance is practically non-existent in the literature.

One of the difficulties in determining how sampling intensity and temporal resolution affect the structure of the observed network is that multiple sampling strategies have almost never been simultaneously applied to a study population. Thus, we rarely know what the observed network would have looked like given a different sampling strategy. The only exception that we are aware of is recent work by Perkins *et al.* (2009), who compared rodent contact networks based on radiotelemetry data (with relatively fine temporal resolution and low sampling intensity) and mark-recapture data (with relatively coarse temporal resolution and high sampling intensity). The authors concluded that radiotelemetry data were more informative when population density was low and mark-recapture data were preferable at high population densities, and further noted that radiotelemetry may be more appropriate in studies of diseases with short infectious periods because finer temporal resolution is necessary to identify an infected individual's

contacts during the brief window of transmission, but even in this case there is no gold-standard method.

Basic social characteristics of the population of interest can provide some insight into the optimal partitioning of resources between sampling intensity and temporal resolution. Relevant characteristics of social behavior include:

Territoriality: Network topology should be predictable in territorial populations because interactions occur almost exclusively between individuals (or well-defined groups of individuals) from adjacent territories. While contact data are unlikely to provide new insights into topology beyond what can be deduced from the spatial arrangement of territories, they may be valuable for quantifying rates of contact.

Gregariousness: In solitary species, contacts are generally rare and sampling many individuals may not be productive if each individual is observed so infrequently that its contacts are unlikely to be recorded. In gregarious species with frequent contacts, more emphasis can be placed on sampling as many individuals as possible because even sparse temporal sampling should detect frequent dyadic contacts.

Dominance and Mating System: Polygynous mating and dominance hierarchies should increase the variance in contact rates among individuals, making it important to sample as many individuals as possible to avoid missing highly connected nodes (i.e., "superspreaders" in the context of epidemiology). The omission of such nodes can dramatically change observed network properties (James *et al.* 2009).

Group Stability: Gregarious species vary in the degree to which group membership changes through time. In species with very stable group structures (e.g.,

many social carnivores, primates, and cooperatively breeding birds), group membership may be apparent even with few observations, but more frequent sampling may be required to document inter-group dynamics. Species with unstable groups (e.g., many ungulates, shorebirds, and waterfowl) will benefit more from increased sampling of individuals to describe patterns of association that are not otherwise obvious.

While social characteristics can be helpful predictors of contact patterns, almost all populations contain individuals, or even classes of individuals, that do not fit these categories (e.g., nonterritorial ‘floaters’ in many highly social species). Moreover, social structures are not always constant through time (e.g., solitary species that become gregarious during a brief mating season, as in some amphibians). Thus, sampling strategies that work well for most individuals and time periods will sometimes work poorly for describing contact patterns of atypical individuals or time periods, which can be critical to processes such as disease transmission.

If there is little pre-existing knowledge of social structure of the study population, researchers may face a “catch-22” situation wherein the optimal sampling strategy depends on a contact structure that is unknown until the study has already begun. In such cases, it can be advantageous to employ an adaptive method that allows the sampling strategy to be modified as data on contact structure become available. Direct observations work well in this regard because the number of individuals observed and time spent observing each individual can sometimes be adjusted as data collection progresses. In contrast, proximity loggers and GPS collars are typically deployed at the outset of the

study and contact data are not available until the end of the study period, so the sampling strategy is essentially fixed.

Static versus Dynamic Networks

The coarse temporal resolution of many animal contact studies results in contact datasets that are sparse over short time intervals. Consequently, a common method for analyzing social networks is to collapse dyadic contact data over relatively long time periods in order to capture connections that would go undetected over shorter time periods due to sampling limitations. Static networks generated in this manner rely on the assumption that network structure is constant through time, or that temporal variation would not affect inferences about the question of interest, but as Wey *et al.* (2008) note, "Not all of the relationships represented may have existed at the same time, nor indeed may have all the individuals been together simultaneously."

While static networks may be appropriate for answering questions about long-term patterns of association (e.g., Lusseau *et al.* 2006), they can be problematic for answering disease-related questions because the timing of contacts matters for disease transmission. Static networks obscure information on the concurrency and order of contacts between dyads, which determine the possible pathogen transmission pathways in a network (Bansal *et al.* 2010). Static networks can be particularly problematic when contact data are collapsed over a time interval that is longer that the average duration of infection for an individual because the network structure will suggest a greater number of potential contacts between an infected node and its neighboring nodes than is actually possible during the infectious period (Cross *et al.* 2004). Several recent simulation studies

have confirmed that static networks can misrepresent patterns of transmission and epidemic thresholds in dynamic networks (Fefferman & Ng 2007; Volz & Meyers 2007; Volz & Meyers 2009; Risau-Gusman 2010). In systems where pathogens alter contact behavior of infected hosts (e.g., healthy house finches preferentially feeding next to conspecifics infected with *Mycoplasma gallisepticum*; Bouwman & Hawley 2010), static networks may be unable to identify resulting shifts in network structure. Static networks can also be misleading in analyses of social behavior; for instance, data on agonistic interactions are sometimes aggregated into a matrix to produce a dominance hierarchy that includes some dyads that were never present at the same time.

Proximity loggers avoid many of the difficulties associated with static networks because they capture network structures among sampled individuals continuously over long periods (up to several years, depending on battery life and available memory), allowing the temporal dynamics of networks to be fully explored. Shifts in network structure through time can be observed by comparing networks based on data from different subsets of the study period. The fine temporal resolution of proximity logger data comes at a cost of limited sampling of the network, however, because proximity loggers can rarely be deployed on all individuals within a population. Alternative methods such as direct observation may identify most or all individuals in the network, but yield sparser data on contacts between individuals. In general, when applied to dynamic networks, proximity loggers provide an accurate description of contact structure across all time periods but only for a small portion of nodes, while direct observations

provide a fuller description of overall network structure but miss contacts during any particular time period.

Effects of Incomplete Data in Social Networks

Effects of incomplete data on network properties have received considerable attention in the human social sciences literature. The most common approach has been to randomly remove nodes or edges from simulated model networks (with varying structures, e.g., random, scale-free, small world) and observe the effects on node-based centrality measures. Using this method, Borgatti *et al.* (2006) found predictable declines in the accuracy of centrality measures for random networks, suggesting that valid confidence intervals can be constructed when sampling intensity is known. However, Frantz *et al.* (2009) found large differences between five model networks in the robustness of centrality metrics to sampling, and concluded that network topology has a greater effect on metric accuracy than other network properties such as size or density. Both of these studies simulated error rates (i.e., percentages of omitted nodes and edges) of up to 50 percent, corresponding to sampling intensities of 50 percent or greater; field studies of wildlife populations rarely obtain sampling intensities as good as these studies' worst-case scenarios. Studies of incomplete data in empirical networks are relatively rare compared to simulation studies (but see Costenbader & Valente 2003 and Wey *et al.* 2008). Clearly, our understanding of effects of incomplete data remains limited, particularly for empirical networks.

The effects of incomplete data are even more poorly understood for animal social networks than for human social networks, for several reasons. First, sampling intensity is

sometimes not known in wildlife studies because precise estimates of population size are often difficult to obtain. Second, studies of incomplete data have typically removed data at random, but nodes and edges that are not sampled in the wild are probably not a random subset of the true network's nodes and edges. If an individual's probability of being sampled (and thus included in the network) is correlated with its rate of social interaction, then resulting network structure could be misleading. This could occur if individuals in groups are more easily detected and observed (or outfitted with collars) than solitary individuals.

Moreover, Lee *et al.* (2006) found that subsampling a dataset has different effects on network metrics such as average path length and clustering coefficient when sampling occurs via random selection of nodes versus random selection of edges. Lastly, it is unclear whether wildlife networks are well represented by the model networks that have been used to test effects of incomplete data in human social networks. Human networks are often scale-free and small-world networks (Amaral *et al.* 2000; Liljeros *et al.* 2001), in which most connections are local but rare long-distance connections (e.g., global air travel) greatly reduce the average distance between any two nodes. While some wildlife networks exhibit small-world properties (Croft *et al.* 2004, Lusseau *et al.* 2006), most are inherently more spatially structured than human networks because animal species are often constrained in their movements by landscape features and lack the capacity for rapid, long-distance movements that humans commonly make.

Effects of Sampling Intensity on Network Metrics

The structure of proximity logger data, with complete records of contacts between a set of known individuals, makes it an attractive candidate for SNA. Yet proximity logger studies will often be limited to a fixed sample of network nodes representing only a small fraction of the population of interest. Thus, understanding the implications of low sampling intensity for the accuracy of the network metrics commonly used in SNA is critical, particularly if we wish to make comparisons across studies or populations. In essence, we must understand how the properties of a sample 'scale up' to the entire network. In many cases this scaling is not well understood and sometimes depends on the properties of the network that one is trying to estimate in the first place (Stumpf *et al.* 2005).

Many metrics have been used to describe the properties of social networks (Table 1-2). Wey *et al.* (2008) describe three levels of organization for network metrics: individual-level metrics describing the properties of a focal node (e.g., node degree), intermediate-level metrics describing sub-group structure within a network (e.g., clustering coefficient, cliquishness), and group-level metrics describing properties of the entire network (e.g., density, diameter). It is also useful to distinguish between metrics that are influenced only by direct connections between nodes (e.g., node degree) and metrics that also account for indirect connections between nodes separated by more than one edge (e.g., average path length).

Some metrics will be biased in a predictable direction by subsampling a network. For instance, mean node degree will be equal or lower in a randomly sampled network

than a full network because a portion of each node's neighbors are omitted from the sampled network (e.g., Guimera & Sales-Pardo 2009, Stumpf *et al.* 2005). This may be an unfamiliar problem for many ecologists because typically sampling intensity does not affect estimates of the sample mean (only the variance). If we measured body mass for a sample of animals, for example, we would not expect the mean body mass of the sample to increase as more animals were measured.

Other metrics respond less predictably to sampling intensity. For instance, density may not decline in sampled networks because it depends on the percentage of possible edges in the observed network, but not on the number of nodes included in the network; in other words, having fewer nodes in the sampled network does not imply that these nodes should be less connected to their remaining neighbors than in the full network. Clustering coefficient is similar, in that only the fraction of possible edges between a node's neighbors matters, and not the actual number of neighbors. While low sampling intensity will produce more variable estimates of density and clustering coefficient, it may not produce biased estimates.

Indirect metrics may be especially vulnerable to sampling effects because the omission of a node or edge potentially affects many distant nodes. Failing to include even a single node, for instance, may dramatically increase the diameter of the observed network if the omitted node provided an important link between otherwise distantly-connected nodes. Some indirect metrics may be of no value at all in sampled networks - average path length, for example, is not calculable for networks consisting of

unconnected components (which may occur when sampling omits nodes that connect subgroups).

The effects of sampling intensity depend on whether absolute or relative differences between nodes are of interest. For instance, is it important to know that the most connected individual in a network has 20 neighbors and the least connected node has 5 neighbors, or is it sufficient to know that there is four-fold variation in node degree among individuals? If one wanted to predict the rate of disease spread through a population, the first type of information might be needed, which would require sampling all or nearly all nodes in the network. However, if one wanted simply to determine which age class is most connected, the second type of information would be adequate and could be obtained with fewer nodes sampled.

SNA is a potentially powerful approach partly because of its flexibility to handle many different social structures, but dealing with incomplete data is a critical and unresolved problem, particularly when sampling intensity is limited. Except in cases where the majority of a population can be outfitted with proximity loggers, the uncertainties associated with sampling intensity may outweigh the advantages conferred by proximity loggers in a network context. Consequently, we believe the strength of proximity loggers lies outside of the network paradigm. Here we propose an alternative framework for analyzing contact data, like those provided by proximity collars, that assesses the individual, dyadic, and environmental factors contributing to the variation in contact rates.

Explaining Individual Variation in Contact Rate

Variation in contact rates among categories of individuals based on characteristics such as age, sex, or social rank is a well-studied phenomenon (e.g., Pereira 1988, Creel *et al.* 1992, Bradley *et al.* 2004; Wolf *et al.* 2007). Yet variation in contact rates among individuals within these categories remains poorly understood, despite being a common feature of human and animal populations (Bansal *et al.* 2007; Clay *et al.* 2009; Marsh *et al.* 2010). Woolhouse *et al.* (1997) proposed a "20/80 rule" as a general feature of animal populations, whereby 20 percent of individuals are responsible for 80 percent of disease transmission in a population, and disease ecologists have conventionally thought of superspreading events (in which large numbers of individuals become infected over a short time) as being associated with particular individuals. Similar to Lloyd-Smith *et al.* (2005), however, we think it is more informative to think about super-spreading events rather than super-spreading individuals. Super-spreading events are due to a combination of three broad factors: individual variation in infectiousness and susceptibility to infection, individual variation in contact rates, and the environmental context. We do not address the first factor here, as studies of infectiousness and susceptibility are better suited for controlled laboratory environments. The last two factors, however, can be tackled with field studies of animal contact rates, and proximity loggers are allowing us to distinguish between these factors in ways not previously possible.

Disease ecologists and modelers are keenly interested in the sources of variation in contact rate (and thus transmission rate) in animal populations because heterogeneity in disease transmission is associated with rarer but more explosive disease outbreaks and

higher estimates of R_0, a measure of a pathogen's potential for epidemic spread (Lloyd-Smith *et al.* 2005; May 2006; Porphyre *et al.* 2008). Targeted interventions will vastly improve the efficiency of disease control (May 2006), but designing such measures requires an understanding of whether the variation in transmission is due to individual characteristics (i.e., "intrinsic" sources of heterogeneity) or the environmental context (i.e., "extrinsic" sources). Lloyd-Smith *et al.* (2005) documented many cases of human super-spreading event and many were related to specific events (e.g., church gathering, fraternity party) or locations (e.g., airplanes). Thus, the individuals involved may be less important than the context.

Extrinsically-driven heterogeneity may be a common feature of animal populations in which individuals are dispersed across spatially variable environments or vary greatly in group size and social organization (e.g., Arctic hares). Potential extrinsic sources of variation include weather conditions, climate, habitat type, and topography, as well as biotic factors such as predation pressure, social structure, and inter-specific competition. Some would argue that differences in the environmental conditions experienced by individuals are a reflection of differences in habitat preference, an individual attribute; although the distinction between intrinsic and extrinsic sources of variation is sometimes unclear, identifying environmental factors contributing to variation in contact rate is fundamental to management efforts, and this distinction is largely an academic one.

We now describe a novel statistical approach for exploring intrinsic and extrinsic sources of heterogeneity in individual contact rate by pairing contact data of the type

provided by proximity loggers with environmental data provided by GPS collars or other means.

Statistical Approaches

The complex dependencies inherent in many contact and network datasets are not easily addressed by traditional statistical approaches. As a result, many ecological network analyses have been conducted using randomization tests (e.g., Mantel and partial Mantel tests; Table 1-3) that compare the properties of the observed network to a random null model of association between nodes. Mantel tests are often used to determine whether network structure is correlated with some other characteristic of dyads, such as their genetic relatedness or difference in age. For instance, Croft *et al.* (2006) used this method to test whether the strength of association of guppy (*Poecilia reticulata*) dyads was correlated with their tendency to investigate predators together. We find many of these analyses unsatisfying because differentiating the empirical data from a random distribution often does not help to identify the mechanisms responsible for the departure, and it is often unclear what the null model should be (e.g., Cross *et al.* 2004). More recently, exponential random graph models (ERGMs) have been developed to analyze network data (Snijders *et al.* 2006, Robins *et al.* 2007). These models estimate the probability of a contact (or edge) between individuals/nodes as a function of network parameters such as degree distribution and transitivity. The development of estimation procedures (maximum likelihood, pseudo-likelihood, Bayesian Markov Chain Monte Carlo) and specifications to better fit empirical network data are active areas of research. ERGMs have typically been used for static network analyses, but Snijders (2005) and

others have extended these approaches to dynamic networks. These approaches are usually applied to known networks with complete data, and many network estimates using ERGMS are highly biased by incomplete data (Huisman 2009). Here we propose an alternative approach that may be applicable to many field settings where the network is relatively weakly sampled.

In our approach, we redefine the question and ask: What factors are associated with contact rate or the probability of contact between individuals *A* and *B,* given that they are located within the same group? By focusing our analysis on within-group associations we remove some of the higher-order network dependencies that are not easily modeled with traditional approaches. Separating within-group processes (contact) from among-group processes (dispersal, fission-fusion) is perhaps a more intuitive approach for many wildlife ecologists and animal behavior researchers. While we focus here on statistically characterizing contact rates within a group, a full understanding of contact structure will also require information on how individuals move among groups and how groups themselves interact.

Our approach utilizes generalized linear mixed models (GLMMs), which are increasingly applied to ecological datasets (Bolker *et al.* 2009). Some aspects of contact data are well-suited for GLMM analyses because these models can accommodate non-normally distributed data arising from large variation in contact rates and hierarchical data structures, where contact data are grouped by dyad or individual. Random effects models are often used in the analysis of ecological data to account for the non-independence of multiple samples taken from the same individual (Breslow & Clayton

1993; Gillies *et al.* 2006), and often these random effects are viewed as a statistical nuisance. Individual and dyadic effects, however, are of central interest to us.

Let y_i represent the number of contacts within a given time period between a pair of individuals that were in the same group. Using the notation of Gelman and Hill (2007), we can write a simple multi-level model as

$$y_i \sim N(\beta x_i + \alpha_{j[i]}, \sigma_y^2), \text{ for } i = 1,...,n$$

where x_i refers to some environmental covariate and $\alpha_{j[i]}$ is the random effect of the j^{th} dyad. In the absence of additional information, we let

$$\alpha_j \sim N(\mu_\alpha, \sigma_\alpha^2)$$

and assume that contact rate is normally distributed around a predicted mean of $\beta x_i + \alpha_{j[i]}$ and the dyad effects are also normally distributed. This formulation appears simple, but represents a parameter-rich model due to the *J* dyad effects, where *J* is not the total number of possible dyads, but the total number of dyads that were observed in the same group at least once in the dataset. The multiple observations of dyads over time are critical to estimating the dyad effects. μ_α indicates the mean contact rate over all the dyads given some covariate effects βx_i, and the heterogeneity among dyads is captured by σ_α^2. The individual α_j's can be used to investigate which dyads were more or less likely to make contact, and the model could then be elaborated upon to predict why some pairs were more likely to make contact by assuming that the dyad effects are themselves a function of covariates. For example, suppose that individuals of the same sex were more likely to make contact. We could then assume that

$$\alpha_j \sim N(\gamma_0 + \gamma_1 z_j, \sigma_\alpha^2)$$

where z_j is an indicator variable representing whether the pair was of the same sex or not. By including more x_j and z_j covariates, we can decompose the variance in contact rate between environmental and dyadic effects, reducing the amount of variation in α_j (i.e., σ_α^2) as important covariates are included whenever there is confounding (e.g., dyad *j* was only observed in a high-contact environmental context) (Fig. 1-1). A perfect estimate of the proportion of the variation in contact rate attributable to individual versus environmental causes is unlikely because one will never know all the important covariates, but even rough estimates would represent an important advance in our understanding of many social and disease processes. Models of particular importance to disease ecologists will probably include a covariate of group size or local density. If dyadic contact rates are independent of one another within a group, we could then use this model to scale up to how the total contact rate among all individuals in the group varies with group size or density.

Because our approach models contact rates between pairs of individuals given common group membership, it requires information beyond the contact data provided by proximity loggers; we must also know when dyads were in the same group and thus potentially interacting. Sometimes this information will be obtainable by direct observation of groups, noting which sampled individuals are present in each observed group and the time of observation. Contact rate for each potential dyad in each observed group can then be assessed within a predefined time interval bracketing the observation by summing the number of dyadic contacts in that interval. The width of time interval

over which to sum contacts has important ramifications. Short time intervals allow examination of fine-scale changes in contact patterns, but as the interval width decreases, the number of contacts approaches a binomial distribution of 'within contact' or 'not within contact' at a single instant in time, which requires different link functions and distributional assumptions about y_i. The optimal interval width will depend on the question and study system being addressed.

An alternative method for obtaining group membership information would be pair GPS collars and proximity loggers on sampled individuals. GPS locations could be used to retrospectively determine when individuals were in same group (or at least close enough to be potentially interacting) and the environmental context of contacts (e.g., habitat type occupied when contact occurred), without requiring field observations. This method should dramatically increase the amount of useable contact data because contacts from the entire study period could be analyzed, not just those contacts that occurred within relatively brief time intervals bracketing field observations. A frequent GPS fix rate may be necessary to tightly link contacts recorded by proximity loggers with geographic locations of those contacts, especially for highly mobile species.

As currently presented, our approach models number of contacts at the level of the dyad, with a random effect term to explain variation in contact rate among dyads that is not accounted for by environmental covariates. Our model can be used to ask, "What characteristics of this specific pairing of animals contribute to its observed contact rate?" It is also valuable, however, to consider contact rates at the level of the individual and ask, "What characteristics of this specific animal contribute to its observed contact rate?"

Heterogeneity in contact rate among individuals is arguably more interpretable than heterogeneity among dyads, and for purposes of disease control, it is easier to individuals than dyads for management actions. One way to modify our statistical approach to incorporate individual effects would be to replace the dyad effect term (α_j) in our model with two individual effect terms, one for each individual in the dyad. These individual effects could be then be modeled as a function of covariates such as individual age or sex, similar to our model for dyads. Comparing the fit of this individual model to the fit of our dyad model could be informative: a better fit of the individual model would suggest that properties of the individual determine contact rates (after accounting for environmental effects), whereas a better fit of the dyad model would suggest that contact rates are more influenced by the relationships between individuals (i.e., "friends" or "enemies" within the population).

Limitations and Unresolved Issues

Although our approach is promising, several statistical challenges remain for this type of analysis. We have presented our approach using a normal model for simplicity, but often this will not be the most appropriate distribution for modeling contact data. If contact rates are measured as counts of contact events, a discrete probability distribution such as the Poisson or negative binomial will be a better choice. A continuous probability distribution like the normal or gamma distribution could be used, however, if contact rates are measured as durations of contact rather than counts. In either case, correcting for data overdispersion may be necessary because contact data commonly exhibit greater variance (often in the form of excess zeros) than predicted by distributions like the

Poisson. Zero-inflated models (Lambert 1992, Hall 2000) and hurdle models (Mullahy 1986, Gurmu 1998) may be good options for modeling overdispersed contact data (e.g., Martin *et al.* 2005).

Our approach as currently presented does not differentiate between true heterogeneity caused by dyad effects and apparent heterogeneity caused by differences in proximity logger performance. Could an individual have a high contact rate (as recorded by proximity loggers) not because it makes more contacts, but because its proximity logger receives incoming UHF signals at a greater distance than other loggers? While careful calibration should minimize variation among loggers in threshold contact distance, field conditions will sometimes present unavoidable issues (e.g., UHF antenna damage that limits the transmitting/receiving range of a collar). The ratio of an individual's number of contacts as recorded by its own logger divided by its number of contacts as recorded by other loggers should provide a simple measure of logger performance, with ratios much higher or lower than 1 indicating potential issues. We have not yet considered how such information should be integrated into our multi-level modeling approach.

Interpreting and comparing dyadic effects will be difficult when dyads are not well sampled. Dyads could be poorly sampled because some pairs of individuals rarely occurred in the same group, or because the groups in which they were found together were rarely observed (if field observations were used to determine common group membership). In our multi-level model framework, dyad effects are estimated via partial pooling, a tradeoff between excluding a dyad-level predictor from the model (i.e.,

complete pooling, which ignores variation among dyads) and estimating separate models for each dyad (i.e., no pooling, which overstates variation among dyads; Gelman and Hill 2007). The less data available for a particular dyad, the more strongly that dyad's estimated effect is pulled toward the overall mean. Thus, for the same underlying contact structure, sparse datasets with few observations per dyad will produce relatively homogeneous estimates of dyad effects, while richer datasets will suggest more heterogeneous dyad effects. We must therefore be careful when comparing dyad effects, both within and between populations. Within a population, do outlying dyad effect estimates represent true outliers, or just better-sampled dyads? And between populations, do distributions of estimated dyad effects differ because of true differences in heterogeneity, or because of differences in observation effort?

Given these limitations, our approach will be most useful when dyads can be sampled numerous times across a broad range of the environmental covariates. Studies that rely on field observations to determine group membership may have trouble meeting this standard, but the pairing of GPS collars and proximity loggers is a powerful approach that should maximize the ability to link contacts to their environmental context. Still, the utility of our approach will sometimes be limited by the temporal and spatial resolution of environmental data. Consider, for example, the effect of snow depth on contact rate. Relatively precise locations of each dyad's contacts might be available from GPS and proximity logger data, but snow depths are typically measured at a very coarse spatial scale (e.g., one weather station per several hundred square miles). Consequently, fine-scale variation in snow depths experienced by dyads in the population would be masked

by the poor spatial resolution of environmental data. Many environmental variables, however, should perform much better – for instance, habitat type is often defined at fine enough scale to place contacts within well-defined patches and allow a meaningful examination of the effects of habitat type on variation in dyadic contact rates. Spatiotemporal resolution should become less limiting in the future as remotely-sensed environmental data continue to improve.

Conclusions

Technological advances such as proximity loggers allow researchers to collect animal contact data with much greater resolution and efficiency than in the past, providing new opportunities but also ushering in new theoretical and statistical challenges. We have identified some of the issues that animal behaviorists and disease ecologists are likely to encounter, in hopes that recognition of these issues will increase productivity in these fields. Part of our motivation for writing this paper is to encourage thoughtful application of proximity loggers; we find there is a tendency to deploy new technologies without clear goals or research questions, leading to inefficient data collection and post-hoc analyses (as we would argue has been the case with GPS collars). There is much to be learned from increasingly available animal contact data, but a judicious approach will be necessary.

We have provided a novel method for analyzing contact data in a linear modeling framework that avoids many of the difficulties associated with networks, particularly sparsely sampled networks that are common in studies of animal contact rates. Despite

several outstanding statistical issues, we hope our approach is a useful stepping stone for future advances that will allow researchers to understand the factors affecting variation in contact rate, which is likely to create many new insights in multiple fields.

Acknowledgments

This work was partially supported by the National Science Foundation and National Institutes of Health Ecology of Infectious Disease Grant DEB-1067129. The views and opinions in this article are those of the authors and should not be construed to represent any views, determinations, or policies of the U.S. Government. Also, any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Literature Cited

Altizer, S., Dobson, A., Hosseini, P., Hudson, P., Pascual, M. & Rohani, P. (2006) Seasonality and the dynamics of infectious diseases. *Ecology Letters*, **9**, 467-484.

Altmann, J. (1974) Observational study of behavior: sampling methods. *Behavior*, **49**, 227-267.

Amaral, L.A.N., Scala, A., Barthelemy, M. & Stanley, H.E. (2000) Classes of small-world networks. *Proceedings of the National Academy of Sciences*, **97**, 11149-11152.

Bansal, S., Grenfell, B.T. & Meyers, L.A. (2007) When individual behavior matters: homogeneous and network models in epidemiology. *Journal of the Royal Society Interface*, **4**, 879-891.

Bansal, S., Read, J., Pourbohloul, B. & Meyers, L.A. (2010) The dynamic nature of contact networks in infectious disease epidemiology. *Journal of Biological Dynamics*, **4**, 478-489.

Bohm, M., Hutchings, M.R.& White, P.C.L. (2009) Contact networks in a wildlife-livestock host community: identifying high-risk individuals in the transmission of bovine TB among badgers and cattle. *PLoS ONE*, **4**, e5016.

Bolker, B., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H. & White, J.S. (2009) Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution*, **24**, 127-135.

Borgatti, S.P., Carley, K.M. & Krackhardt, D. (2006) On the robustness of centrality measures under conditions of imperfect data. *Social Networks*, **28**, 124-136.

Bouwman, K.M. & Hawley, D.M. (2010) Sickness behavior as an evolutionary trap? Male house finches preferentially feed near diseased conspecifics. *Biology Letters*, **6**, 462-465.

Bradley, B.J., Doran-Sheehy, D.M., Lukas, D., Boesch, C. & Vigilant, L. (2004) Dispersed male networks in western gorillas. *Current Biology*, **14**, 510-513.

Breslow, N.E. (1984) Extra-Poisson variation in log-linear models. *Applied Statistics*, **33**, 38–44.

Breslow, N.E. & Clayton, D.G. (1993) Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, **88**, 9-25.

Carslake, D., Bennett, M., Bown, K., Hazel, S., Telfer, S. & Begon, M. (2005) Space-time clustering of cowpox virus infection in wild rodent populations. *Journal of Animal Ecology*, **74**, 647-655.

Clay, C.A., Lehmer, E.M., Previtali, A., St. Jeor, S. & Dearing, M.D. (2009) Contact heterogeneity in deer mice: implications for Sin Nombre virus transmission. *Proceedings of the Royal Society B*, **276**, 1305-1312.

Costenbader, E. & Valente, T.W. (2003) The stability of centrality measures when networks are sampled. *Social Networks*, **25**, 283-307.

Creel, S., Creel, N., Wildt, D.E., & Monfort, S.L. (1992) Behavioural and endocrine mechanisms of reproductive suppression in Serengeti dwarf mongooses. *Animal Behaviour*, **43**, 231-245.

Creel, S. & Winnie, J.A. (2005) Responses of elk herd size to fine-scale spatial and temporal variation in the risk of predation by wolves. *Animal Behaviour*, **69**, 1181-1189.

Croft, D.P., Krause, J., & James, R. (2004) Social networks in the guppy (*Poecilia reticulata*). *Proceedings of the Royal Society B*, **271**, 516-519.

Croft, D.P., James, R., Thomas, P.O.R., Hathaway, C., Mawdsley, D., Laland, K.N. & Krause, J. (2006) Social structure and co-operative interactions in a wild population of guppies (*Poecilia reticulata*). *Behavioral Ecology and Sociobiology*, **59**, 644-650.

Croft, D.P., Krause, J., Darden, S.K., Ramnarine, I.W., Faria, J.J. & R. James (2009) Behavioural trait assortment in a social network: patterns and implications. *Behavioral Ecology and Sociobiology*, **63**, 1495-1503.

Cross, P.C., Lloyd-Smith, J.O., Bowers, J.A., Hay, C.T., Hofmeyr, M. & Getz, W.M. (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.

Drewe, J.A. (2010) Who infects whom? Social networks and tuberculosis transmission in wild meerkats. *Proceedings of the Royal Society B*, **277**, 633-642.

Fefferman, N.H. & Ng, K.L. (2007) How disease models in static networks can fail to approximate disease in dynamic networks. *Physical Review E*, **76**, 031919.

Fenner, A.L., Godfrey, S.S. & Bull, C.M. (2011) Using social networks to deduce whether residents or dispersers spread parasites in a lizard population. Journal of Animal Ecology, DOI:10.1111/j.1365-2656.2011.01825.x.

Fortin, D., Fortin, M.E., Beyer, H.L., Duchesne, T., Courant, S. & Dancose, K. (2009) Group-size mediated habitat selection and group fusion-fission dynamics of bison under predation risk. *Ecology*, **90**, 2480-2490.

Frantz, T.L., Cataldo, M. & Carley, K.M. (2009) Robustness of centrality measures under uncertainty: examining the role of network topology. *Computational and Mathematical Organization Theory*, **15**, 303-328.

Gelman, A. & Hill, J. (2007) *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, New York, New York, USA.

Gillies, C.S., Hebblewhite, M., Nielsen, S.E., Krawchuk, M.A., Aldridge, C.L., Frair, J.L., Saher, D.J., Stevens, C.E. & Jerde, C.L. (2006) Application of random effects to the study of resource selection by animals. *Journal of Animal Ecology*, **75**, 887-898.

Godfrey, S.S., Bull, C.M., James, R. & Murray, K. (2009) Network structure and parasite transmission in a group living lizard, the gidgee skink, *Egernia stokesii*. *Behavioral Ecology and Sociobiology*, **63**, 1045-1056.

Gurmu, S. (1998) Generalized hurdle count data regression models. *Economics Letters*, **58**, 263-268.

Hall, D.B. (2000) Zero-inflated Poisson binomial regression with random effects: a case study. *Biometrics*, **56**, 1030-1039.

Hamede, R.K., Bashford, J., McCallum, H. & Jones, M. (2009) Contact networks in a wild Tasmanian devil (*Sarcophilus harrisii*) population: using social network analysis to reveal seasonal variability in social behaviour and its implications for transmission of devil facial tumour disease. *Ecology Letters*, **12**, 1147-1157.

Henzi, S.P., Lusseau, D., Weingrill, T., van Schaik, C.P. & Barrett, L. (2009) Cyclicity in the structure of female baboon social networks. *Behavioral Ecology and Sociobiology*, **63**, 1015-1021.

Huisman, M. (2009). Imputation of missing network data: some simple procedures. Journal of Social Structure, **10.1**.

James, R., Croft, D.P. & Krause, J. (2009) Potential banana skins in animal social network analysis. *Behavioral Ecology and Sociobiology*, **63**, 989–997.

Ji, W., White, P.C.L. & Clout, M.N. (2005) Contact rates between possums revealed by proximity data loggers. *Journal of Applied Ecology*, **42**, 595-604.

Kjær, L.J., Schauber, E.M. & Nielsen, C.K. (2008) Spatial and temporal analysis of contact rates in female white-tailed deer. *Journal of Wildlife Management*, **72**, 1819-1825.

Lambert, D. (1992) Zero-inflated Poisson regression with an application to defects in manufacturing. *Technometrics*, **34**, 1-14.

Lee, S.H., Kim, P. & Jeong, H. (2006) Statistical properties of sampled networks. *Physical Review E*, **73**, 016102.

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

Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E. & Getz, W.M. (2005) Superspreading and the effect of individual variation on disease emergence. *Nature*, **438**, 355-359.

Lusseau, D. & Newman, M.E.J. (2004) Identifying the role that individual animals play in their social network. *Proceedings of the Royal Society B*, **271**, S477-S481.

Lusseau, D., Wilson, B., Hammond, P.S., Grellier, K., Durban, J.W., Parsons, K.M., Barton, T.R. & Thompson, P.M. (2006) Quantifying the influence of sociality on population structure in bottlenose dolphins. *Journal of Animal Ecology*, **75**, 14-24.

Madden, J.R., Drewe, J.A., Pearce, G.P. & Clutton-Brock, T.H. (2009) The social network structure of a wild meerkat population: 2. Intragroup interactions. *Behavioral Ecology and Sociobiology*, **64**, 81-95.

Marsh, M.K., Hutchings, M.R., McLeod, S.R. & White, P.C.L. (2010) Spatial and temporal heterogeneities in the contact behaviour of rabbits. *Behavioral Ecology and Sociobiology*, DOI:10.1007/s00265-010-1024-y.

Martin, A.D. (2003) Bayesian inference for heterogeneous event counts. *Sociological Methods and Research*, **32**, DOI:10.1177/0049124103253500.

May, R.M. (2006) Network structure and the biology of populations. *Trends in Ecology and Evolution*, **21**, 394-399.

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

Mullahy, J. (1986) Specification and testing of some modified count data models. *Econometrics*, **3**, 341-365.

Otterstatter, M.C & Thomson, J.D. (2007) Contact networks and transmission of an intestinal pathogen in bumble bee (*Bombus impatiens*) colonies. *Oecologia*, **154**, 411-421.

Pereira, M. E. (1988) Agonistic interactions of juvenile savanna baboons. *Ethology*, **79**, 195–217.

Perkins, S.E., Cagnacci, F., Stradiotto, A., Arnoldi, D. & Hudson, P.J. (2009) Comparison of social networks derived from ecological data: implications for inferring infectious disease dynamics. *Journal of Animal Ecology*, **78**, 1015-1022.

Porphyre, T., Stevenson, M., Jackson, R. & McKenzie, J. (2008) Influence of contact heterogeneity on TB reproduction ratio R_0 in a free-living brushtail possum *Trichosurus vulpecula* population. *Veterinary Research*, **39**, DOI:10.1051/vetres:2008007.

Prange, S., Jordan, T., Hunter, C. & Gehrt, S.D. (2006) New radiocollars for the detection of proximity among individuals. *Wildlife Society Bulletin*, **34**, 1333-1344.

Ramsey, D., Spencer, N., Caley, P., Efford, M., Hansen, K., Lam, M. & Cooper, D. (2002) The effects of reducing population density on contact rates between brushtail possums: implications for transmission of bovine tuberculosis. *Journal of Applied Ecology*, **39**, 806-818.

Real, L.A. & Biek, R. (2007) Infectious disease modeling and the dynamics of transmission. *Wildlife and Emerging Zoonotic Diseases: The Biology, Circumstances and Consequences of Cross-Species Transmission* (eds J. E. Childs, J. S. Mackenzie, and J. A. Richt), pp. 33-50. Springer, New York, New York, USA.

Richomme, C., Gauthier, D. & Fromont, E. (2006) Contact rates and exposure to inter-species disease transmission in mountain ungulates. *Epidemiology and Infection*, **134**, 21-30.

Risau-Gusman, S. (2010) Influence of network dynamics on the spread of sexually transmitted diseases. arxiv.org, 1004.1378.

Robins, G., Pattison, P., Kalish, Y. & Lusher, D. (2007) An introduction to exponential random graph (p^*) models for social networks. *Social Networks*, **29**, 192-215.

Schauber, E.M., Storm, D.J. & Nielsen, C.K. (2007) Effects of joint space use and group membership on contact rates among white-tailed deer. *Journal of Wildlife Management*, **71**, 155-163.

Snijders, T.A.B. (2005) Models for longitudinal network data. *Models and methods in social network analysis* (eds. J.S.P. Carrington and S. Wasserman), Ch. 11. Cambridge University, New York, New York, USA.

Snijders, T.A.B., Pattison, P.E., Robins, G.L., & Handcock, M.S. (2006) New specifications for exponential random graph models. *Sociological Methodology*, **36**, 99-153.

Volz, E. & Meyers, L.A. (2007) Susceptible–infected–recovered epidemics in dynamic contact networks. *Proceedings of the Royal Society B*, **274**, 2925-2933.

Volz, E. & Meyers, L.A. (2009) Epidemic thresholds in dynamic contact networks. *Journal of the Royal Society Interface*, **6**, 233-241.

Wey, T., Blumstein, D.T., Shen, W. & Jordan, F. (2008) Social network analysis of animal behavior: tools for the study of sociality. *Animal Behavior*, **75**, 333-344.

Wey, T.W. & Blumstein, D.T. (2010) Social cohesion in yellow-bellied marmots is established through age and kin structuring. *Animal Behavior*, **79**, 1343-1352.

Wolf, J.B.W., Mawdsley, D., Trillmich, F. & James, R. (2007) Social structure in a colonial mammal: unravelling hidden structural layers and their foundations by network analysis. *Animal Behavior*, **74**, 1293-1302.

Woolhouse, M.E.J., Dye, C., Etard, J.-F., Smith, T., Charlwood, J.D., Garnett, G.P., Hagan, P., Hii, J.L.K., Ndhlovu, P.D., Quinnell, R.J., Watts, C.H., Chandiwana, S.K. & Anderson, R.M. (1997) Heterogeneities in the transmission of infectious agents: implications for the design of control programs. *Proceedings of the National Academy of Sciences*, **94**, 338-342.

Table 1-1. Comparison of methods for acquiring contact data.

Method	Example studies	Interaction(s) recorded	Temporal Resolution	Spatial resolution	Proportion of population sampled	Potential biases
Direct observation	Drewe 2010 Lusseau *et al.* 2006 Richomme *et al.* 2006 Wey & Blumstein 2010	Any of interest (e.g., grooming, aggression, sexual contact)	Depends on observation frequency; usually coarse	Ranges from coarse (e.g., common group membership) to very fine (e.g., direct contacts)	Highly variable; depends on size of study population and observer effort	Observability could be biased by habitat type, time of day, weather conditions, animal behavior, or group size
VHF telemetry	Ramsey *et al.* 2002 Cross *et al.* 2004 Perkins *et al.* 2009	Usually spatial proximity or common group membership	Depends on tracking frequency; usually coarse to intermediate	Usually coarse	Usually low	Observability could be biased by habitat type, time of day, weather conditions
GPS telemetry	Schauber *et al.* 2007 Kjær *et al.* 2008	Spatial proximity	Depends on fix frequency; usually intermediate to fine	Fine (limited by triangulation error)	Usually low	GPS fix success may vary by habitat type
Mark-recapture	Carslake *et al.* 2005 Porphyre *et al.* 2008 Perkins *et al.* 2009	Spatial proximity	Depends on trapping frequency; usually coarse	Usually very coarse	Varies by trapping effort; usually low to intermediate	If capture probability is correlated with contact rate, may overestimate mean contact rate for population
Proximity loggers	Ji *et al.* 2005 Bohm *et al.* 2009 Hamede *et al.* 2009 Marsh *et al.* 2010	Spatial proximity	Finest possible (constantly recording)	Very fine (currently adjustable from 0.5 – 100 m)	Usually low	Detection range may be reduced in habitats with dense vegetation

Table 1-2. Glossary of social network analysis terminology.

Betweeness centrality: the proportion of shortest paths between nodes along which a focal node lies.

Closeness centrality: the average shortest path length between a focal node and all other network nodes

Clustering coefficient: a measure of the connectedness of a node's neighbors; calculated as the number of existing edges between a focal node's neighbors divided by the maximum possible number of such edges.

Edge: represents a relationship (e.g., contact) between two nodes; may be weighted (i.e., scaled according to strength of association) or unweighted, and directed (i.e., indicating direction of interaction from one node to the other) or undirected

Degree: the number of edges directly connected to a focal node; in directed networks, in-degree is the number of edges directed into a node, and out-degree is the number of edges directed out of a node.

Density: the proportion of all possible edges between nodes that is present in the observed network

Diameter: the longest path length in a network.

Dynamic network: a network with properties that vary through time

Efficiency: a metric describing how well information or disease flows through a network consisting of multiple, unconnected components

Network: a description (in graphical or matrix form) of the relationships among a set of nodes.

Node: represents an actor within the network, typically an individual or group

Path length: the shortest distance (number of edges) between two nodes

Random network: a network structure in which edges are formed at random between nodes, irrespective of their spatial distribution

Scale-free network: a network structure characterized by large heterogeneity in node degree, with many low-degree nodes and few high-degree nodes

Small world network: a network structure characterized by high clustering and short average path lengths, with many short-distance connections and few long-distance connections between nodes.

Static network: a network with properties that are assumed to remain constant through time

Topology: the physical arrangement of nodes and edges in a network

Table 1-3. Selected animal contact studies from past 5 years.

Study	Species	Question(s) addressed	Contact data	Contact metric(s) analyzed	Method of analysis
Clay *et al.* 2009	Deer mouse (*Peromyscus maniculatus*)	How heterogeneous are individual contact rates in deer mouse populations? Do contact rates vary with sex or breeding condition?	Frequency and durations of dyadic contacts (defined by spatial proximity and/or exchange of colored powder via direct contact)	Contact frequencies and durations	Examined distribution of contact frequency within population; used generalized linear mixed models to test relationship between serostatus and contact rates
Croft *et al.* 2006	Guppy (*Poecilia reticulate*)	How are guppy populations socially structured? Does network structure predict patterns of cooperation?	Number of times individuals were observed in same shoal or inspected a predator together	Association strength (binary measure based on if pair seen together ≥3 times)	Used randomization test to compare numbers of persistent pairs in observed data to a null model of shoal membership; used Mantel tests to test association between networks based on association and predator inspection
Croft *et al.* 2009	Guppy (*Poecilia reticulata*)	Is social structure shaped by individuals' behavioral phenotype?	Number of days individuals were seen in same shoal	Association index, node degree	Examined correlation between behavioral score (a measure of behavioral phenotype) and node degree
Drewe 2010	Meerkat (*Suricata suricatta*)	Is disease status related to rate of intra- and inter-group social interaction?	Frequency of grooming, aggression, and eviction behaviors	Out-degree, in-degree, flow-betweenness	Used OLS regression to test for association between individuals' centrality measure scores and change in disease status
Fenner *et al.* 2011	Pygmy bluetongue lizard (*Tiliqua adelaidensis*)	Are patterns of parasite infection best explained by transmission via contact with resident or disperser lizards?	Spatial proximity of lizard burrows	Node 'strength' (sum of edge weights connected to individual)	Used randomization test to determine whether infected individuals were more connected to resident lizards or disperser lizards than uninfected individuals
Flack *et al.* 2006	Pigtailed macaque (*Macaca nemestrina*)	How does policing behavior affect the stability of social networks?	Frequency of behavioral interactions (grooming, play) and spatial proximity	Mean degree, reach, assortative mixing, clustering	Used repeated measures and randomization to compared contact metrics for empirical networks with and without experimental removal of high-status (i.e., policing) individuals
Godfrey *et al.* 2009	Gidgee skink (*Egernia stokesii*)	Does observed contact network structure explain patterns of parasite infection?	Presence/absence of contact for each dyad (based on common rock crevice use)	Node degree (population means)	Used linear mixed models to test relationship between node degree and infection status

Table 1-3 (continued). Selected animal contact studies from past 5 years.

Study	Species	Question(s) addressed	Contact data	Contact metric(s) analyzed	Method of analysis
Hamede *et al.* 2009	Tasmanian devil (*Sarcophilus harrisii*)	Do contact networks vary seasonally? Are there certain age or sex classes that are highly connected?	Frequency and durations of dyadic contacts (defined by spatial proximity)	Assortativity coefficient, node degree, betweenness, centrality, transitivity	Compared network-based metrics during mating and non-mating season; compared empirical networks to random networks generated using Monte Carlo process to test for sex- or age-based interaction patterns
Henzi *et al.* 2009	Chacma baboon (*Papio ursinus*)	How stable are social structures in the face of environmental variability?	Timing of spatial proximity or grooming events	Lagged association rates of dyads; node-based centrality measures	Used AIC to compare models of social structure stability; used randomization tests to determine if node centrality varied seasonally
Marsh *et al.* 2010	European wild rabbit (*Oryctolagus cuniculus*)	How variable are contact rates spatially and temporally? How do intra- and inter-group contact rates compare?	Frequency and durations of dyadic contacts (defined by spatial proximity)	Frequency and durations of dyadic contacts	Used linear mixed effects models to compare intra-group contact rates between seasons
Madden *et al.* 2009	Meerkat (*Suricata suricatta*)	How do group size, sex ratio, dominance, and parasite load relate to network structure?	Rates of allogrooming, dominance interaction, and foraging competition	Degree centrality, average path length, compactness, clustering coefficient, density	Used Wilcoxon signed-rank test and Fisher's test to relate network structure to group attributes; used quadratic assignment procedure (a type of Mantel test) to compare networks based on different interactions
Otterstatter & Thomson 2007	Bumble bee (*Bombus impatiens*)	Does risk of protozoan infection vary with rate of physical contact, activity level, or activity type?	Dyadic contact rate (as recorded by automated video-tracking	Degree centrality	Used linear mixed models to test whether an individual's infection risk varies with its degree centrality or rate of contact with infected conspecifics
Perkins *et al.* 2009	Yellow-necked mouse (*Apodemus flavicollis*)	How do social networks derived from radiotelemetry and mark-recapture data differ?	Presence/absence of contacts for each dyad during study	Average contact rate, closeness centrality, betweenness centrality, connectedness	Tested for differences in contact metrics of networks based on radiotelemetry or mark-recapture methods using generalized linear models

Table 1-3 (continued). Selected animal contact studies from past 5 years.

Study	Species	Question(s) addressed	Contact data	Contact metric(s) analyzed	Method of analysis
Wey & Blumstein 2010	Yellow-bellied marmot (*Marmota flaviventris*)	Do patterns of affiliative and agonistic interaction vary with age, sex, or kinship?	Frequency of contacts (broken down by type of interaction, e.g. grooming or fighting)	Attractiveness, expansiveness, closeness	Used linear mixed effects model to test for effects of age and sex on node-based metrics; used Mantel tests to determine if network was structured by age, sex, or kinship.
Wolf *et al.* 2007	Galapagos sealion (*Zalophus wollebaeki*)	Does sealion population have substructure? How important are sex and age class, site-fidelity, and male territory distribution in determining network structure?	Dyadic contact rate during course of study (with threshold applied to make binary)	Newman's assortativity coefficient, modularity	Used factorial ANOVA to test for sex- or age-based differences in degree; used randomization to test for community structure

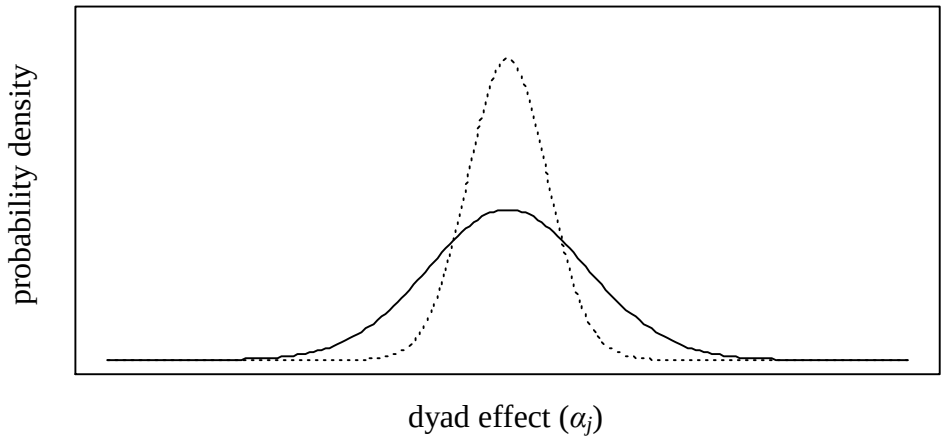

Figure 1-1. Theorized reduction in variance of dyad effects, σ_α^2, when extrinsic sources of variation are incorporated in multi-level models of contact rate. Solid line: distribution of dyad effects (α_j) for model without environmental covariates. Dotted line: distribution for model including environmental covariates.

LOW-DENSITY FEEDING REDUCES ELK CONTACT RATES AND *BRUCELLA* TRANSMISSION ON FEEDGROUNDS

Contributions of Authors and Co-Authors

Manuscript 2

Author: Tyler G. Creech
Contributions: designed and carried out experimental study, analyzed and interpreted data, and prepared manuscript.

Co-author: Paul C. Cross
Contributions: provided comments and conceptual input at all stages, including study design, data analysis, interpretation of results, and manuscript preparation.

Co-author: Brandon M. Scurlock
Contributions: assisted with data collection and commented on manuscript drafts.

Co-author: Eric J. Maichak
Contributions: assisted with data collection and commented on manuscript drafts.

Co-author: Jared D. Rogerson
Contributions: assisted with data collection and commented on manuscript drafts.

Co-author: John C. Henningsen
Contributions: assisted with data collection and commented on manuscript drafts.

Co-author: Scott Creel
Contributions: provided comments and conceptual input at all stages, including study design, data analysis, interpretation of results, and manuscript preparation.

Manuscript Information Page

Manuscript 2

Title: Low Density Feeding Reduces Elk Contact Rates and *Brucella* Transmission on Feedgrounds

Authors: Tyler G. Creech, Paul C. Cross, Brandon M. Scurlock, Eric J. Maichak, Jared D. Rogerson, John C. Henningsen, and Scott Creel

Journal: Journal of Wildlife Management

Status of manuscript:

_____ Prepared for submission to a peer-reviewed journal
__x__ Officially submitted to a peer-reviewed journal
_____ Accepted by a peer-reviewed jour
_____ Published in a peer-reviewed journal

Date of submission: October 16, 2010

Abstract

Supplemental feeding is one of many ways that humans alter wildlife aggregations. Aggregation patterns often have cascading effects on parasite transmission and disease dynamics by altering the rate of contact between individuals. However, directly estimating contact rates, particularly at the level of the individual, has been difficult in field settings. We used proximity loggers and video cameras to estimate rates of disease-relevant elk-to-fetus contacts (the primary source of infection by *Brucella abortus*) during winter supplemental feeding. We compared contact rates during high-density and low-density feeding treatments that provided the same total amount of food at different densities. Low-density feeding led to >50 percent reductions in total number of contacts and number of individuals contacting a fetus. Proximity loggers and cameras provided similar estimates of elk-fetus contact rates. Elk contacted fetuses and random control points equally, suggesting that elk were not attracted to fetuses but encountered them incidentally while feeding. The relationship between contact rate and disease prevalence is non-linear and low-density feeding may result in dramatic reductions in brucellosis prevalence, though this depends on the amount of transmission that occurs on and off feedgrounds.

Introduction

The rate at which individuals contact other individuals or infectious materials in their environment has a strong effect on the transmission and persistence of diseases. Contact rate is directly related to disease transmission rate (whereby transmission rate is proportional to contact rate × probability of transmission given contact), but is difficult to quantify for many wildlife populations, and may often be correlated with population density (McCallum et al. 2001). When contact rates are directly measured, sparse datasets are common because contact data are often obtained through time-consuming direct observation, and in many cases contacts cannot be assigned to unique individuals. Simple disease models often assume that the mean contact rate applies to all individuals in a population, but individual heterogeneity in contact rate is now recognized as an important

determinant of disease dynamics (Dwyer et al. 1997, Lloyd-Smith et al. 2005, Bolzonil et al. 2007). Proximity loggers (Prange et al. 2006) have tremendous potential in wildlife disease ecology studies because they can provide continuous individual-level contact data for a large sample without requiring direct observation by researchers (e.g., Hamede et al. 2009, Marsh et al. 2010). Here, we present a study of contacts in an elk (*Cervus elaphus*) population using both proximity loggers and video cameras to record contacts.

Many wildlife disease studies are observational and relate existing patterns of disease prevalence to environmental covariates, establishing correlates of high prevalence but typically not testing direct causation. Experimental studies in which researchers actively manipulate some variable and measure resulting changes in disease levels (e.g., Jacobson and Hurst 1979, Hudson et al. 1998) allow more direct inferences about causation. Because of the inherent difficulties in applying experimental treatments in natural habitats, opportunities for experimental studies of disease in wild populations are relatively rare. This study takes advantage of a system well-suited to experimental manipulation, Wyoming's elk feedgrounds, to directly test the effect of animal feeding density on contact rates.

Supplemental feeding of wildlife ranges from residential bird feeders to large-scale, long-term feeding programs administered by wildlife management agencies. Justifications for supplemental feeding include increasing survival of threatened populations, manipulating the geographical distribution of populations to avoid conflict with human land uses, increasing wildlife viewing opportunities, and maintaining populations on private lands for hunting. Among other impacts, supplemental feeding

may increase disease transmission by aggregating animals at high density around concentrated food sources. There are strong theoretical underpinnings for an effect of animal density on disease transmission (Anderson and May 1979, McCallum et al. 2001) and empirical studies of mammals have positively linked population density to disease prevalence (Cross et al. 2010) and parasite abundance (Arneberg et al. 1998), although results from meta-analyses have been mixed (Côté and Poulin 1995, Ezenwa 2004). Supplemental feeding has been associated with increased disease prevalence in elk (Thorne and Herriges 1992, Scurlock and Edwards 2010) and white-tailed deer (Schmitt et al. 1997, Miller et al. 2003, Rudolph 2006).

Wyoming's elk feedgrounds comprise the largest and best-known feeding program in the U.S. Since 1910, elk in western Wyoming have been supplementally fed during winters to compensate for loss of native winter range and to minimize conflicts on agricultural lands (Smith 2001). Today, approximately 22,000 elk are fed each year on 21 state-maintained feedgrounds and the National Elk Refuge (WGFD 2010). Currently, the most problematic disease on these feedgrounds is brucellosis, a chronic bacterial disease caused by *Brucella abortus.* Within the U.S., *B. abortus* in wildlife is limited to the Greater Yellowstone Ecosystem (GYE), where it was likely introduced to bison from imported European cattle prior to 1917 (Meagher and Meyer 1994) and subsequently spread to elk by 1930 (Murie 1951). The primary symptom of brucellosis in elk is abortion during the first pregnancy following infection and occasionally during subsequent pregnancies (Thorne et al. 1978). The disease is not considered a major

mortality factor in elk herds (Cheville et al. 1998) but may reduce herd reproductive potential by as much as 12 percent (Thorne et al. 1991).

Because *B. abortus* can infect humans through contaminated dairy products, a U.S. Department of Agriculture (USDA) program to eradicate the disease from cattle herds was instituted in 1934, and by 1998 an estimated $3.5 billion dollars had been spent on eradication efforts (Cheville et al. 1998). The USDA declared all U.S. cattle herds brucellosis-free in 2008, but since then infections have been reported in Montana and Wyoming. Cattle infections have economic consequences for state cattle industries due to increased testing requirements, stricter regulations on in-state cattle movement, and refusal by some states to allow importation of cattle from infected states (Healey et al. 1997). Genetic analysis of *B. abortus* strains from elk, bison, and cattle indicates that elk are the most probable source of the recent cattle infections (Beja-Pereira et al. 2009), increasing the pressure on state and federal wildlife management agencies to control elk brucellosis.

Brucellosis is typically transmitted among elk via direct contact with infectious abortion materials including fetuses, placentas, and fetal fluids (Thorne et al. 1978). Vertical transmission from mother to calf through milk has been reported (Cheville et al. 1998) and aerosol transmission may also be possible (Nicoletti 1980), but both are believed to be uncommon routes of transmission. Contacts between elk and naturally-aborted fetuses on feedgrounds are rarely observed, but elk have been seen investigating fetuses placed on feedgrounds (Cook et al. 2004, Maichak et al. 2009), and supplemental feeding is believed to play a key role in *B. abortus* transmission. Historically, average

seroprevalence of elk management units with feedgrounds has been several times that of units without feedgrounds (Scurlock and Edwards 2010), although recently seroprevalence has been increasing in areas distant from feedgrounds (Cross et al. 2010).

Given the economic impacts of brucellosis, there is strong incentive to reduce intraspecific transmission on feedgrounds through management actions such as enhancing surrounding habitat, shortening feeding seasons, feeding on fresh snow, protecting scavengers, and reducing feeding density (Cross et al. 2007, WGFD 2008). Feed has historically been distributed on elk feedgrounds along continuous, high-density (HD) feedlines (Fig. 2-1), but in 2008 the Wyoming Game and Fish Department (WGFD) began a low-density (LD) feeding technique at five feedgrounds (WGFD 2008). LD feeding involves distributing feed in small, discrete units over a larger area, encouraging elk to disperse evenly across the feedground (Fig. 2-1) and reducing animal densities along feedlines (Patrek 2009). Here, we assess whether LD feeding reduces the rate of contact between elk and aborted fetuses on feedgrounds. If fetuses are a strong attractant to elk, then reducing feeding density may not prevent elk-fetus contacts, but if elk are not attracted to fetuses, then reducing density may reduce contacts both with other elk and with fetuses.

Because proximity collars provide individual-level contact data, we investigate whether variation exists among individual elk in rate of contact with fetuses. Individual heterogeneity is commonly ignored in wildlife disease management, potentially leading to poor models of disease dynamics and ineffective control strategies (Dwyer et al. 1997, Bolzonil et al. 2007). Brucellosis management stands to benefit from any information on

individual heterogeneity in infection risk, one component of which is elk-fetus contact rate. If classes of individuals with higher-than-average elk-fetus contact rate (and thus elevated infection risk) can be identified based on characteristics such as age or pregnancy status, then management strategies could be improved by focusing efforts on these high-risk animals.

The objectives of our study were: 1) to evaluate the effectiveness of LD feeding in reducing elk-fetus contact rates relative to traditional HD feeding practices, 2) to determine the magnitude of individual heterogeneity in elk-fetus contact rates among feedground elk, 3) to explore potential changes in *Brucella* seroprevalence if LD feeding were implemented throughout the feedground system, and 4) to compare two alternative technologies used to measure elk-fetus contact rates. We placed *Brucella*-negative fetuses along feedlines during LD- and HD-feeding periods and recorded elk-fetus contact events using both proximity loggers and video cameras. We used a simple hierarchical model to examine individual heterogeneity in elk-fetus contact rate and a simple *SIR* disease model to translate changes in contact rate into altered seroprevalence.

Study Area

We conducted our study at the WGFD-administered Soda Lake feedground (SLF) near Pinedale, Wyoming in the Wind River Range foothills (42°95’N, 109°81’W, elevation 2314 m). The regional climate is characterized by long, cold winters and brief, warm summers with most precipitation falling as winter snow. Regional vegetation comprises sagebrush (*Artemesia* spp.) communities at lower elevations and mixed conifer

(*Pinus* spp., *Abies lasiocarpa, Picea engelmannii*) forests at higher elevations. SLF is dominated by herbaceous species with limited shrubs. Depending on winter severity, 500-900 elk are fed at SLF between December and April.

Methods

We used proximity loggers (Sirtrack Ltd., Havelock North, New Zealand) to record elk contacts with aborted fetuses along feedlines. Proximity loggers transmit and receive unique UHF signals and record the date, time, duration, and individual logger identities when they are within a user-defined contact distance of each other (see Prange et al. 2006 for technical details). In a disease context, contact rate will often be multiplied by the probability of transmission given contact, which in our case was an unknown function of distance. Smaller contact distances may result in relatively few contacts to analyze while larger distances may be less relevant for disease transmission. We calibrated proximity loggers to record contacts within 2 m to reflect that brucellosis transmission requires relatively close contact with infectious materials, while still recording a sufficient number of recorded contacts for analysis.

In January and February 2009, we captured 30 cow elk (≥1.5 yrs old) at SLF and fitted them with proximity-logging collars. Bulls were excluded from the study because they likely do not contribute to brucellosis transmission (Thorne et al. 1978, Cheville et al. 1998). Elk were captured via chemical immobilization using 1.5-mL darts loaded with carfentanil (0.01 mg/kg) and xylazine (0.1 mg/kg; Kreeger et al. 2002). Captures were performed in accordance with approved Montana State University Animal Care and Use

Protocol (no. 2010-02). We obtained 14 *Brucella* culture-negative fetuses, placentas, and fetal fluids (hereafter collectively termed "fetuses") from elk killed at Muddy Creek and Fall Creek feedgrounds in 2008 during the Test and Slaughter pilot project (Scurlock 2010). Fetuses were processed and cultured as described in Maichak et al. (2009) and Alton et al. (1988).

We conducted experimental feeding trials during 14 days at SLF in late February and early March 2009. Elk were fed daily from a horse-drawn sled with an attached Global Positioning System (GPS) unit to record the path of the sled. Feed was distributed in seven pairs of HD and LD feeding days; quantity of hay fed daily varied during the study, but was constant within HD-LD feeding day pairs. Each day, we placed a single fetus at a random point along the feedline with a proximity logger (the "fetus logger") buried in snow 15 cm below the fetus. Another proximity logger without an accompanying fetus (the "control logger") was buried at a second random point along the feedline to determine a baseline rate of contact against which to measure the attractiveness of the fetus to elk. We placed loggers directly along the path of the sled because contacts with fetuses >2 m from a feedline occur very rarely (Maichak et al. 2009). We set up digital video cameras with infrared lights (Model X100, Sandpiper Technologies, Inc., Manteca, CA) approximately 10 m from fetus and control loggers to visually record contact events. Cameras allowed us to count elk-fetus contacts for feedground cow elk without proximity collars and to distinguish between incidental contacts (e.g., elk feeding next to a fetus) and investigations (sniffing, licking, or other

physical contact with the fetus). Elk were conditioned to the presence of cameras for 2 weeks prior to the experiment.

Proximity data were downloaded from the fetus and control loggers at the end of the study period. Prior to analysis, we removed all 1-second duration contacts from the dataset as recommended by Prange et al. (2006) to avoid counting interactions outside the 2-m continuous detection zone of the proximity collars. We assigned contacts to "feeding days" commencing at the time of logger placement during feed distribution and ending 18 hours later; all contacts recorded >18 hours after logger placement but prior to the initiation of the next feeding day were censored to standardize by the shortest period between consecutive feedings during the experiment. Approximately 90 percent of contacts occurred within 18 hours of logger placement.

We reviewed video camera footage and recorded for each feeding day the number of cow elk (with or without proximity collars) that approached within 2 m of the fetus or control logger, and how many of these elk investigated the fetus. Because elk were not individually identifiable in camera footage, we could determine the total number of contact events but not the number of unique individuals making contacts. We also did not record durations of contacts for elk in camera footage due to time limitations. Thus, only data on total numbers of contacts were available from camera footage for analysis.

We analyzed proximity logger and camera data independently, but used the same statistical methods for the two datasets. For each feeding day, we calculated three contact rate metrics for fetus contacts and control contacts: 1) *total contacts* - the number of contact events recorded; 2) *unique contacts* - the number of individuals that contacted the

logger at least once during the feeding day; and 3) *mean duration* - the average duration of a contact. (Note that unique contacts and mean duration were not calculated for camera data because these data were not collected). To analyze the effect of LD feeding on elk-fetus contact rates, we treated density as a categorical variable (i.e., HD or LD) and determined the percent reduction in each contact metric for LD feeding relative to HD feeding for each feeding day pair. Similarly, we compared fetus logger contacts to control logger contacts within each feeding day for each contact metric to determine relative attractiveness of fetuses. We collapsed contact data among all individuals for each feeding day, leading to conservative estimates of the significance of the treatment effect. We used non-parametric, two-tailed Wilcoxon signed-rank tests (Zar 1999) to determine if mean percent reductions from HD to LD feeding were different than zero, and to determine if the ratios of fetus contacts to control contacts were different than 1.

To explore the extent of individual heterogeneity in elk-fetus contact rate among feedground elk, we modeled the number of fetus contacts (y) by individual j on feeding day i using a generalized linear mixed model. We used a negative binomial distribution whereby μ_{ij} (the expectation of y_{ij}) is predicted by a fixed effect of feeding density (X_i) and a random effect of individual (α_j):

$$\log(\mu_{ij}) = \beta_0 + \beta X_i + \alpha_j$$

This is a variable-intercept model in which all individuals exhibit the same slope for the relationship between contact rate and density, but at a given density the mean contact rate is allowed to vary by individual. For this analysis, feeding density was treated as a continuous variable and calculated by applying a 2-m buffer to GPS sled tracks and

dividing the mass of hay fed by the area of the resulting polygon; densities were then scaled by the number of elk present on the feedground each day to account for slight fluctuations in feedground population size during the study. By this measure, density was 85 percent lower on average during LD feeding than during HD feeding. We compared within-individual and between-individual variability of estimated random effect coefficients to determine the importance of individual heterogeneity in elk-fetus contacts. Modeling was conducted using the R2WinBUGS package to call WinBUGS version 1.4.3 (Gilks et al. 1994) from R version 2.7.2 (R Development Core Team 2008). We assumed diffuse normal priors for β_0 and β_1 with a mean of zero and a precision of 0.01. We assigned the random effect α_j a normal prior distribution with a mean of zero and a standard deviation that was uniformly distributed from zero to 10. The model was run for 500,000 iterations on three different Markov chains and the first half of each chain was discarded. We assessed convergence using the Gelman-Rubin-Brooks statistic, where $\hat{R}$ < 1.1 for all parameters indicated that relatively little variation was associated with specific MCMC chain (Gelman and Hill 2007).

Results

Proximity Loggers

Proximity loggers recorded a total of 168 fetus contacts and 142 control contacts (after removing 96 <1-s contacts and 30 contacts occurring outside the feeding day). Twenty-nine of 30 elk made at least one fetus contact during the study (mean: 5.7 contacts during study, range: 0-18). We found large, statistically significant reductions in

total contact rate and unique contact rate on LD feeding days relative to HD feeding days (Fig. 2-2): total contacts were reduced by an estimated 86 percent ($P = 0.022$, W = 28, n = 7) and unique contacts by 83 percent ($P = 0.022$, W = 28, n = 7). Mean duration was reduced by 46 percent on average, but this reduction was not statistically significant ($P = 0.27$, W = 21, n = 7). However, this result was highly influenced by a single LD feeding day on which several lengthy fetus contacts were recorded, and thus the percent reduction in mean duration for that HD-LD feeding day pair was highly negative; excluding this outlier, mean duration was reduced by 84 percent on average ($P = 0.035$, W = 21, n = 6). We found no difference between rate of contact with the fetus logger and the control logger for total contacts ($P = 0.780$, W = 41, n = 12; Fig. 2-3), unique contacts ($P = 0.794$, W = 92, n = 12), or mean duration ($P = 0.305$, W = 69, n = 12).

Cameras

Camera footage was less complete than proximity logger data due to battery failures, elk tampering with cables and lighting, and heavy snowfall obscuring the camera lens. Complete camera footage was available for 8 of 14 feeding days for fetus contacts and 6 of 14 feeding days for control contacts. We limited our analyses to portions of feeding days for which footage was available for both days in a pair (when comparing contact rates on HD and LD days) or for both loggers simultaneously (when comparing fetus and control logger contact rates), totaling 273 hrs and 200 hours of fetus logger and control logger footage, respectively. We found large reductions in total contact rate on LD feeding days relative to HD feeding days for both contacts within 2 m and investigation contacts (Fig. 2-2). The mean reduction in contacts was 59 percent for <2-m

contacts (P = 0.031, W = 21, n = 6) and 70 percent for investigations (P = 0.031, W = 21, n = 6; Fig. 2-2). Fetuses and control points did not differ in total contact rate (P = 0.652, W = 69, n = 9; Fig. 2-3). Of those elk approaching within 2 m of the fetus, 24 percent investigated the fetus. The percentage of elk investigating the fetus at <2 m did not differ between HD and LD feeding days (P = 0.393, W = 24, n = 6).

Individual Heterogeneity

We found inconclusive evidence of individual heterogeneity in elk-fetus contact rates. There was complete overlap of 95 percent credible intervals for random effect coefficient estimates (corresponding to individual elk) with the overall mean coefficient estimate, suggesting limited heterogeneity within the feedground population (Fig. 2-4). However, point estimates for random effect coefficients showed some variation: the individual with highest contact rate had an estimated 1.75 times as many daily contacts as the individual with the lowest contact rate. Traditionally, the importance of a random effect term would be evaluated by comparing AIC scores, R^2 values, or other goodness-of-fit measures for models with and without the random effect. For generalized linear mixed models (such as the negative binomial model we use), however, goodness-of-fit testing procedures are still being developed and debated (Zuur et al. 2009). We therefore did not attempt to formally assess the significance of the random effect in our model. Given the weak support for a random effect of individual, we did not model the random effect as a function of individual covariates such as elk age or pregnancy status.

Discussion

LD feeding dramatically reduced elk-fetus contact rates and may reduce brucellosis transmission events among elk on the feedgrounds. We found large and significant reductions in total contact rate and unique contact rate despite the small sample size (n = 7 pairs) and conservative statistical methods used in our analysis. Proximity collars allowed us to collect information on unique individuals that was unavailable from camera data, but the associated need to capture many individuals at a given site limited our ability to replicate the study across many sites. However, camera data from HD and LD feeding in 2008 at five additional Wyoming feedgrounds (Bench Corral, South Park, Greys River, Franz, and Muddy Creek) corroborate this study's finding of reduced numbers of elk-fetus contacts with LD feeding (WGFD, unpublished data). Assuming that LD feeding is causally responsible for the observed reductions in contact rates, we expect our results from Soda Lake would apply to other feedgrounds, depending on the extent to which each site allows for dispersed feeding.

We used multiple measures of contact rate because the probability of disease transmission depends on several characteristics of contact events. The total number of contact events with a fetus is a logical indicator of transmission risk, but how those contacts are distributed among the population is also important. For instance, five elk-fetus contacts could be distributed as five contacts by a single individual or one contact by each of five different individuals, with differing disease dynamics expected for these scenarios, particularly if the probability of transmission from a single contact is large. The unique contact rate thus provides additional information critical to assessing

transmission. Duration of contact events is also informative, as lengthy contact events may be more likely to result in transmission. The observed reductions in total contacts and unique contacts, even without a significant change in mean duration of a contact, provide convincing evidence that LD feeding could reduce transmission rate.

The impacts of LD feeding on transmission risk may be larger than estimated by our study. We only examined changes in contact rate, but recent research suggests that probability of transmission given contact may also vary as a function of feeding density. Patrek (2009) found higher stress levels among feedground elk than among unfed elk, and stress has been linked to immune system suppression and increased disease susceptibility in some species (Barnard et al. 1994, Oppliger et al. 1998). Thus, LD feeding may decrease probability of transmission given contact via reductions in stress levels as well as directly reducing contact rate.

We calibrated proximity loggers to record contacts at a distance of <2 m in order to ensure enough contacts for a sufficiently powerful analysis. While it could be argued that *B. abortus* transmission is very unlikely from 2 m away, we still found contact rates at this distance to be informative regarding the effect of LD feeding on contact rate. Analysis of camera data suggests that while only a quarter of <2-m contacts were investigations, the percent reductions in <2-m contacts and investigations due to LD feeding were similar. In fact, the point estimate of percent reduction was greater for investigations than for <2-m contacts, so a 2-m distance may actually have provided a conservative estimate of percent reduction in close contact rates. We acknowledge, however, that our study did not directly assess reductions in actual transmission. If the

probability of transmission given close contact with a fetus is high, then contact rate should be a strong predictor of transmission, but the relationship may be weak if transmission probability is very low.

In our study, 32 ± 6 (mean ± SE) percent of elk approached within 2 m of an aborted fetus during a typical HD feeding day. If we assume that all elk are equally likely to contact a fetus, which seems reasonable given the limited heterogeneity we observed, then almost 80 percent of the population would be exposed after only four abortion events along feedlines. Yet the average seroprevalence on feedgrounds in 2010 was only 19 percent (WGFD, unpublished data). This suggests that the probability of infection given contact at <2 m is small, reinforcing that transmission typically requires direct contacts with abortion materials, not all of which may provide the infectious dose of *B. abortus* required for transmission. Alternatively, this discrepancy could be explained if antibody titer loss is common (and seroprevalence estimates at a given point in time therefore underestimate the proportion of individuals having been infected during their life), or if fetuses are aborted very rarely along feedlines. We discuss these possibilities in more detail below.

Although previous studies have documented elk investigating fetuses, we are unaware of any that have rigorously assessed the attractiveness of fetuses to elk. It is widely believed that cow elk are behaviorally predisposed to investigate fetuses (Geist 1982, Maichak et al. 2009), but similar contact rates with fetus and control loggers in our study suggest that fetuses are encountered incidentally during feeding and are not actively sought out by elk. Camera footage of contact events showed that less than a quarter of elk

passing within 2 m of a fetus actively investigated it, further supporting findings that elk are minimally inclined to investigate abortions. However, we have examined fetus attractiveness only in the feedground environment using previously frozen fetuses; it is possible that elk respond differently to fetuses on native winter ranges or to fetuses that have not been frozen.

Our data provided only weak evidence for individual heterogeneity in elk-fetus contact rates. The day-to-day variation in contact rate for any particular individual was much larger than the variation in mean daily contact rates among individuals in the population. However, the nearly two-fold difference in estimated contact rates for the individuals with the highest and lowest contact rates suggests that there could be biologically meaningful heterogeneity amongst individuals, but our study lacked sufficient statistical power to detect it. Additional research on individual elk-fetus contact rates is needed to clarify these results. Individual heterogeneity in animal contact rates is an under-studied aspect of wildlife ecology, but proximity logger technology should make many more of these studies possible.

In light of the strong evidence that LD feeding reduces elk-fetus contact rates, we explore how seroprevalence might change with the implementation of LD feeding throughout the feedground system. Contact rate and seroprevalence are nonlinearly related, thus the observed reduction in contact rate cannot be assumed to cause an equivalent reduction in seroprevalence. We relate seroprevalence to contact rate using a simple *SIR* disease model that partitions individuals into susceptible (*S*), infected (*I*), and recovered (*R*) disease classes based on their exposure to *B. abortus*, assuming density-

dependent transmission (Anderson and May 1991). For this model, the equilibrium seroprevalence (P^*) is a function of three parameters: 1) β, a transmission coefficient incorporating both contact rate and probability of transmission given contact; 2) γ, the annual recovery rate for infected individuals (and the inverse of the mean time to recovery in years); and 3) δ, the demographic turnover, which reflects the rate at which individuals in the population are replaced at equilibrium through offsetting births and deaths. For simplicity, we use a non-age structured model, and as a result δ is not equivalent to a traditional fecundity or mortality rate. It can be better conceptualized as a birth or death rate averaged across all age classes and weighted by the proportion of the population in each age class. Our equation for equilibrium seroprevalence is thus:

$$P^* = 1 - \frac{\delta + \gamma}{\beta}$$

This model formulation is typically applied to pathogens transmitted directly from animal to animal, rather than indirectly (e.g., from infectious fetus to elk). However, we believe the model is adaptable to our case if we assume the number of infected individuals is an index of the number of infectious fetuses in the environment, with the coefficient β appropriately rescaled to reflect this distinction (and because fetuses are typically scavenged within 48 hours on feedgrounds, it is unnecessary to model the buildup of infectious materials in the environmental). We introduce one additional metric, the basic reproductive number R_0, defined as the average number of infections caused by an infectious individual in a completely susceptible population. Diseases cannot invade host populations when $R_0 < 1$. For our basic *SIR* model,

$$R_0 = \frac{\beta}{\delta + \gamma}$$

We consider a range of P^* in our model because seroprevalence varies by feedground and year. We let P^* = 0.10 or 0.30, roughly corresponding to the lowest and highest long-term seroprevalence estimates among Wyoming feedgrounds (WGFD, unpublished data). We let δ = 0.13, based on estimates of female elk mortality from twelve western U.S. elk populations (Raithel et al. 2007). This value is higher than a typical cow elk mortality rate and lower than a typical cow elk fecundity rate because it incorporates calves and senescent individuals with relatively high mortality and low fecundity. We let γ = 0.5 based on limited evidence from artificial infections (Thorne et al. 1978). While the recovery rate of elk has never been well estimated in the literature, in our model the predicted reduction in seroprevalence due to LD feeding is unaffected by the specific value of γ chosen; only β is affected because it is scaled by γ and P^*.

Substituting in these parameter values and solving for β yields a transmission coefficient of 0.7 or 0.9 (for P^* = 0.10 or 0.30, respectively) for our system. Because β is the product of contact rate and probability of transmission given contact, a proportional reduction in contact rate causes an equivalent proportional reduction in β. Using the more conservative of our two point estimates of reduction in total contact rate (59 percent from the camera data), β would be reduced to 0.287 (for P^* = 0.10) or 0.369 (for P^* = 0.30) if LD feeding were implemented on feedgrounds; both β values would drop R_0 below 1 and eventually eliminate the disease. However, this assumes that all transmission occurs along feedlines. The fraction of transmission events occurring along feedlines, θ, is the subject of ongoing research but is certainly less than one, so the observed reduction in β

should be rescaled by θ to account for off-feedline transmission events that contribute to regional seroprevalence.

Seroprevalence declines non-linearly with transmission, and predicted reductions are dependent on the initial seroprevalence (Fig. 2-5). When initial seroprevalence is low, even small values of θ correspond to major reductions in seroprevalence. Our model predicts seroprevalence would drop to zero if $\theta > 0.17$, given initial seroprevalence of 10 percent. If initial seroprevalence is higher, reductions in seroprevalence expected from LD feeding are more modest for a given value of θ. With initial seroprevalence of 30 percent, our model predicts seroprevalence would drop to zero if $\theta > 0.51$. Thus, we expect that LD feeding could potentially lead to dramatic declines in seroprevalence, with the greatest reductions in areas where initial seroprevalence is low.

These results are unexpected in the context of recent research on brucellosis seroprevalence among Wyoming elk populations. Cross et al. (2010) found that seroprevalence of some unfed elk populations is now comparable to seroprevalence of some feedground populations, yet animal densities in unfed groups are generally thought to be lower than densities during LD feeding. How can we reconcile the high seroprevalence of these unfed elk with our model results suggesting that large reductions in seroprevalence of feedground elk could be achieved with a relatively modest decrease in feeding density? We propose four possible explanations for this apparent contradiction.

First, θ may be very small and only minor reductions in seroprevalence may occur with LD feeding because most transmission occurs away from feedlines. Feedline transmission may be rare if elk are behaviorally predisposed to move away from

feedlines when aborting (which has never been examined), or alternatively if fetuses are scavenged much more quickly on feedgrounds than on native winter ranges (as found by Maichak et al. 2009), or perhaps due to a combination of these mechanisms. Second, the form or parameterization of our *SIR* model may not accurately represent *B. abortus* transmission dynamics in elk. Three notable assumptions of our model are: 1) the rate of recovery is constant through time (i.e., exponentially distributed infectious period); 2) there is no variation in transmission rate amongst individuals in the population (based on lack of individual heterogeneity in our mixed modeling results); and 3) the population has no age structure. Each of these assumptions has been shown to affect to affect transmission dynamics in disease models (Lloyd-Smith et al. 2005, Wearing et al. 2005, Brooks-Pollack et al. 2010) and a more sophisticated model that does not rely on these assumptions could produce quantitatively different results from our model (though we suspect they would be qualitatively similar). Lastly, our *SIR* model could be appropriate and θ could be large, but transmission dynamics in unfed elk populations could be fundamentally different from dynamics in populations that utilize feedgrounds. Fed and unfed populations experience different climatic conditions, human land use, and hunting pressures, all of which could affect elk aggregation patterns and alter transmission dynamics. In this case, our model predictions could be accurate and LD feeding may result in dramatic seroprevalence reductions. Regardless of its effect on overall seroprevalence in the region, it appears that LD feeding will decrease transmission on feedgrounds and should be considered.

It is worth briefly considering the costs and benefits of proximity collars and video cameras. The primary advantage of proximity loggers was their ability to distinguish among individuals when recording contact events, allowing us to explore individual heterogeneity in contact rates. Proximity loggers also collected data more reliably than cameras and did not require daily attention as cameras did. Data from proximity loggers were available in a convenient format upon download, and although logger calibration was time consuming, even more time was spent viewing and counting elk numbers in camera footage. Camera footage, however, allowed us to examine elk-fetus behavioral interactions.

The main drawback of proximity loggers for most researchers will be the substantial costs of purchasing loggers and capturing animals for collaring. Most proximity logger studies will require sampling of individuals from the population of interest – a problem that camera studies will not face. In this study we outfitted approximately 7 percent of cow elk on our study feedgrounds with proximity collars but found reasonable correspondence between contact rate estimates from proximity logger data and camera data, suggesting that our low sampling intensity was sufficient to reveal population-level effects. However, it is unclear whether this sampling intensity would be sufficiently powerful for studies in different environments or of different species. In situations where contact events are rare or vary strongly by individual, a much greater proportion of the population of interest may need to be sampled.

Management Implications

Managers should consider implementing LD feeding throughout the feedground system where possible. Most of Wyoming's 21 state-run feedgrounds are large enough to accommodate LD feeding at a density similar to SLF, and the remainder might benefit from an intermediate-density feeding strategy. LD feeding offers a simple and cost-effective strategy for reducing *B. abortus* transmission; we estimate that switching to LD feeding represents a negligible increase in the total cost of operating a feedground (WGFD, unpublished data). Other management actions designed to reduce *B. abortus* transmission have had mixed results (e.g., vaccination of feedground elk; Herriges et al. 1989, Roffe et al. 2004) or have been scrutinized for potentially redistributing elk from feedgrounds to agricultural and public grazing lands (e.g., habitat enhancements, shortened feeding seasons). In contrast, we expect that LD feeding would accomplish notable reductions in *B. abortus* transmission and its expansion to additional feedgrounds would be well-received. LD feeding could also be implemented on feedgrounds outside of Wyoming. Regular or emergency feeding of elk still occurs in several other western states, where transmission of diseases other than brucellosis could be exacerbated by the increased animal-to-animal contact and accumulation of pathogens in the environment that occur on feedgrounds.

We recommend proximity collars for future studies of animal contacts with infectious materials in the environment but note that their greater potential lies in documenting direct contacts among animals that occur unpredictably in time or space and thus cannot be easily captured with cameras.

Acknowledgments

We thank M. Ebinger, K. Belinda, A. Williams, A. Brennan, and B. Lozier for assistance with field work and logistics. A. Fluegel, T. Cornish, C. Tate, A. Brown, K. Buckle, and Z. Turnbull collected and processed elk fetuses used in this study. C. Kochanny assisted with proximity logger technical issues. Funding was provided by the U.S. Geological Survey, WGFD, and the Wyoming Wildlife Livestock Disease Partnership. The views and opinions in this article are those of the authors and should not be construed to represent any views, determinations, or policies of the U.S. Government. Also, any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Literature Cited

Alton, G. G., L. M. Jones, R. D. Angus, and J. M.Verger. 1988. Techniques for the brucellosis laboratory. Institut National de la Recherche Agronomique, Paris, France.

Anderson, R. M., and R. M. May. 1979. Population biology of infectious diseases: Part I. Nature 280:361-367.

Anderson, R. M., and R. M. May. 1991. Infectious diseases of humans: dynamics and control. Oxford University, New York, New York, USA.

Barnard, C. J., J. M. Behnke, and J. Sewell. 1994. Social behaviour and susceptibility to infection in house mice (*Mus musculus*): effects of group size, aggressive behavior and status-related hormonal responses prior to infection on resistance to *Babesia microti*. Parasitology 108:487-496.

Beja-Pereira, A., B. Bricker, S. Chen, C. Almendra, P. J. White, and G. Luikart. 2009. DNA genotyping suggests that recent brucellosis outbreaks in the Greater Yellowstone Area originated from elk. Journal of Wildlife Diseases 45:1174-1177.

Bolzonil, L., L. Real, and G. De Leol. 2007. Transmission heterogeneity and control strategies for infectious disease emergence. PLoS ONE 2:e747.

Brooks-Pollock, E, T. Cohen, and M. Murray. 2010. The impact of realistic age structure in simple models of tuberculosis transmission. PLoS ONE 5(1):e8479.

Cheville, N. F., D. R. McCullough, and L. R. Paulson. 1998. Brucellosis in the Greater Yellowstone Area. National Academy Press, Washington, D.C., USA.

Cook, W. E., E. S. Williams, and S. A. Dubay. 2004. Disappearance of bovine fetuses in northwestern Wyoming. Wildlife Society Bulletin 32:254-259.

Côté, I. M., and R. Poulin. 1995. Parasitism and group size in social animals: a meta-analysis. Behavioral Ecology 6:159-165.

Cross, P. C., W. H. Edwards, B. M. Scurlock, E. J. Maichak, and J. D. Rogerson. 2007. Effects of management and climate on elk brucellosis in the Greater Yellowstone Ecosystem. Ecological Applications 17:957–964.

Cross, P. C., E. K. Cole, A. P. Dobson, W. H. Edwards, K. L. Hamlin, G. Luikart, A. D. Middleton, B. M. Scurlock, and P. J. White. 2010. Probable causes of increasing

brucellosis in free-ranging elk of the Greater Yellowstone Ecosystem. Ecological Applications 20:278-288.

Dwyer, G., J. S. Elkinton, and J. P. Buonaccorsi. 1997. Host heterogeneity in susceptibility and disease dynamics: tests of a mathematical model. American Naturalist 150:685-707.

Ezenwa, V. O. 2004. Host social behavior and parasitic infection: a multifactorial approach. Behavioral Ecology 15:446-454.

Geist, V. 1982. Adaptive behavioral strategies. Pages 219-278 *in* Thomas, W., and D. E. Toweill, editors. Elk of North America: ecology and management. Stackpole, Harrisburg, Pennsylvania, USA.

Gelman, A., and J. Hill. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York, New York, USA.

Gilks, W. R, A. Thomas, and D. J. Spiegelhalter. 1994. A language and program for complex Bayesian modelling. The Statistician 43:169–178.

Hamede, R. K., J. Bashford, H. McCallum, and M. Jones. 2009. Contact networks in a wild Tasmanian devil (*Sarcophilus harrisii*) population: using social network analysis to reveal seasonal variability in social behaviour and its implications for transmission of devil facial tumour disease. Ecology Letters 12:1147-1157.

Healey, B., G. Weber, A. Keating, J. L. Alley, L. G. Adams, and G. Slack. 1997. The importance of brucellosis eradication to the national cattle industry. Pages 195-197 *in* E. T. Thorne, M. S. Boyce, P. Nicoletti, and T. J. Kreeger, editors. Brucellosis, bison, elk and cattle in the Greater Yellowstone Area: defining the problem, exploring solutions. Pioneer, Cheyenne, Wyoming, USA.

Herriges, J. D., Jr., E. T. Thorne, S. L. Anderson, and H. A. Dawson. 1989. Vaccination of elk in Wyoming with reduced dose strain 19 *Brucella*: controlled studies and ballistic implant field trials. Proceedings of the US Animal Health Association 93:640–655.

Hudson, P. J., A. P. Dobson, D. Newborn. 1998. Prevention of population cycles by parasite removal. Science 282:2256-2258.

Jacobson, H. A., and G. A. Hurst. 1979. Prevalence of parasitism by *Amblyomma americanum* on wild turkey poults as influenced by prescribed burning. Journal of Wildlife Diseases 15:43-47.

Kreeger, T. J., J. M. Arnemo, and J. P. Raath. 2002. Handbook of wildlife chemical

immobilization, International edition. Wildlife Pharmaceuticals, Fort Collins, Colorado, USA.

Lloyd-Smith, J. O., S. J. Schreiber, P. E. Kopp, and W. M. Getz. 2005. Superspreading and the effect of individual variation on disease emergence. Nature 438:355-359.

Maichak, E. J., B. M. Scurlock, J. D. Rogerson, L. L. Meadows, A. E. Barbknecht, W. H. Edwards, and P. C. Cross. 2009. Effects of management, behavior, and scavenging on risk of brucellosis transmission in elk of western Wyoming. Journal of Wildlife Diseases 45:398-410.

Marsh, M. K., M. R. Hutchings, S. R., McLeod, and P. C. L. White. 2010. Spatial and temporal heterogeneities in the contact behaviour of rabbits. Behavioral Ecology and Sociobiology DOI 10.1007/s00265-010-1024-y.

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

Meagher, M., and M. E. Meyer. 1994. On the origin of brucellosis in bison of Yellowstone National Park - a review. Conservation Biology 8:645-653.

Miller, R., J. B. Kaneene, S. D. Fitzgerald, and S. M. Schmitt. 2003. Evaluation of the influence of supplemental feeding of white-tailed deer (*Odocoileus virginianus*) on the prevalence of bovine tuberculosis in the Michigan wild deer population. Journal of Wildlife Diseases 39:84-95.

Murie, O. J. 1951. The elk of North America. Stackpole, Harrisburg, Pennsylvania, USA and the Wildlife Management Institute, Washington, D.C, USA.

Nicoletti, P. 1980. The epidemiology of bovine brucellosis. Advances in Veterinary Science and Comparative Medicine 24:69-98.

Oppliger, A., J. Clobert, J. Lecomte, P. Lorenzon, K. Boudjemadi, and H. B. John-Alder. 1998. Environmental stress increases the prevalence and intensity of blood parasite infection in the common lizard *Lacerta vivipara*. Ecology Letters 1:129-138.

Patrek, V. E. 2009. The effects of supplemental feeding on stress hormone concentrations in elk. Thesis, Montana State University, Bozeman, USA.

Prange, S., T. Jordan, C. Hunter, and S. D. Gehrt. 2006. New radiocollars for the detection of proximity among individuals. Wildlife Society Bulletin 34:1333-1344.

R Development Core Team. 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raithel, J. D., M. J. Kauffman, and D. H. Pletscher. 2007. Impact of spatial and temporal variation in calf survival on the growth of elk populations. Journal of Wildlife Management 71:795-803.

Roffe, T. J., L. C. Jones, K. Coffin, M. L. Drew, S. J. Sweeny, S. D. Hagius, P. H. Elzer, and D. Davis. 2004. Efficacy of single calfhood vaccination of elk with *Brucella abortus* strain 19. Journal of Wildlife Management 68:830–836.

Rudolph, B. A., S. J. Riley, G. J. Hickling, B. J. Frawley, M. S.Garner, and S. R. Winterstein. 2006. Regulating hunter baiting for white-tailed deer in Michigan: biological and social considerations. Wildlife Society Bulletin 34:314-321.

Schmitt, S. M., S. D. Fitzgerald, T. M. Cooley, C. S. Bruning-Fann, L. Sullivan, D. Berry, T. Carlson, R. B. Minnis, J. B. Payeur, and J. Sikarskie. 1997. Bovine tuberculosis in free-ranging white-tailed deer from Michigan. Journal of Wildlife Diseases 25:749-758.

Scurlock, B. M., and W. H. Edwards. 2010. Status of brucellosis in free-ranging elk and bison in Wyoming. Journal of Wildlife Diseases 46:442-449.

Scurlock, B.M., W. H. Edwards, T. Cornish, and L. L. Meadows. 2010. Using Test and Slaughter to reduce prevalence of brucellosis in elk attending feedgrounds in the Pinedale elk herd unit of Wyoming; Results of a 5-year pilot project. WGFD. Cheyenne, Wyoming, USA.

Smith, B. L. 2001. Winter feeding of elk in western North America. Journal of Wildlife Management 65:173-190.

Thorne, E. T., J. K. Morton, F. M. Blunt, and H. A. Dawson. 1978. Brucellosis in elk. II. Clinical effects and means of transmission as determined through artificial infections. Journal of Wildlife Diseases 14:280-291.

Thorne, E. T., M. Meagher, and R. Hillman. 1991. Brucellosis in free-ranging bison: Three perspectives. Pages 275-287 *in* R. B. Keiter and M. S. Boyce, editors. The Greater Yellowstone Ecosystem: redefining America's wilderness heritage. Yale University, New Haven, Connecticut, USA.

Thorne, E. T., and J. D. Herriges, Jr. 1992. Brucellosis, wildlife and conflicts in the Greater Yellowstone Area. Pages 453-465 *in* Transactions of the 57th North American Wildlife & Natural Resources Conference. Montana State University, Bozeman, Montana, USA.

Wearing, H. J., P. Rohani, and M. J. Keeling. 2005. Appropriate models for the management of infectious disease. PLoS Medicine 2:e174.

WGFD. 2008. Target Feedground Project. WGFD, Cheyenne, Wyoming, USA.

WGFD. 2010. Annual big game herd unit job completion reports (E102-E108), Jackson/Pinedale Regions. WGFD, Cheyenne, Wyoming, USA.

Zar, J. H. 1999. Biostatistical analysis. Prentice Hall, Upper Saddle River, New Jersey, USA.

Zuur, A. F., E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith. 2009. Mixed effects models and extensions in ecology with R. Springer, New York, New York, USA.

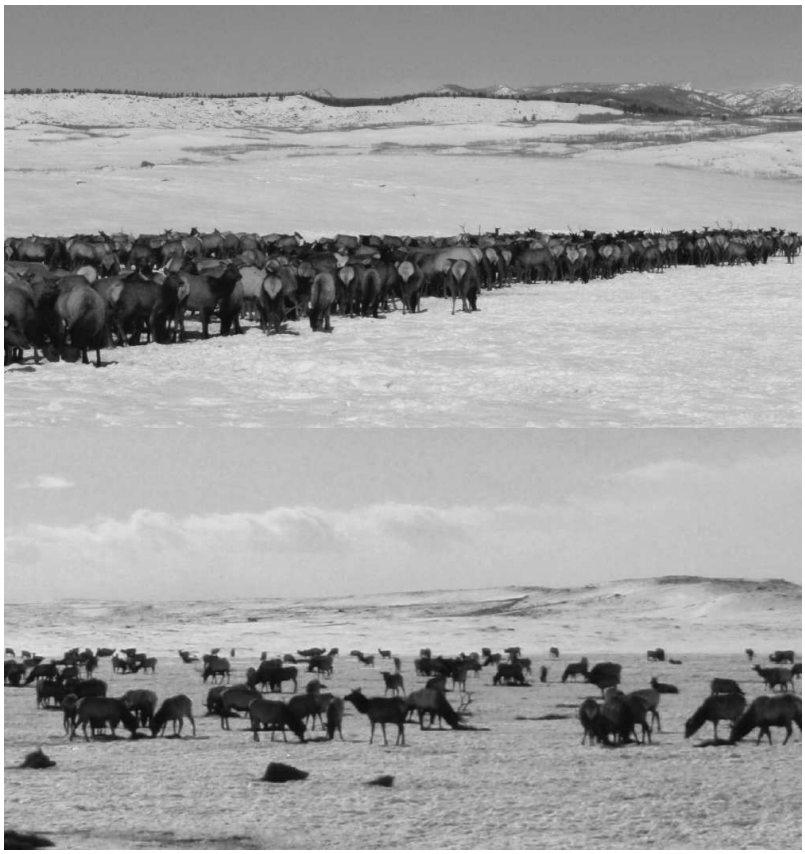

Figure 2-1. Typical elk distribution patterns on Wyoming feedgrounds during high-density feeding (top photo) and low-density feeding (bottom photo).

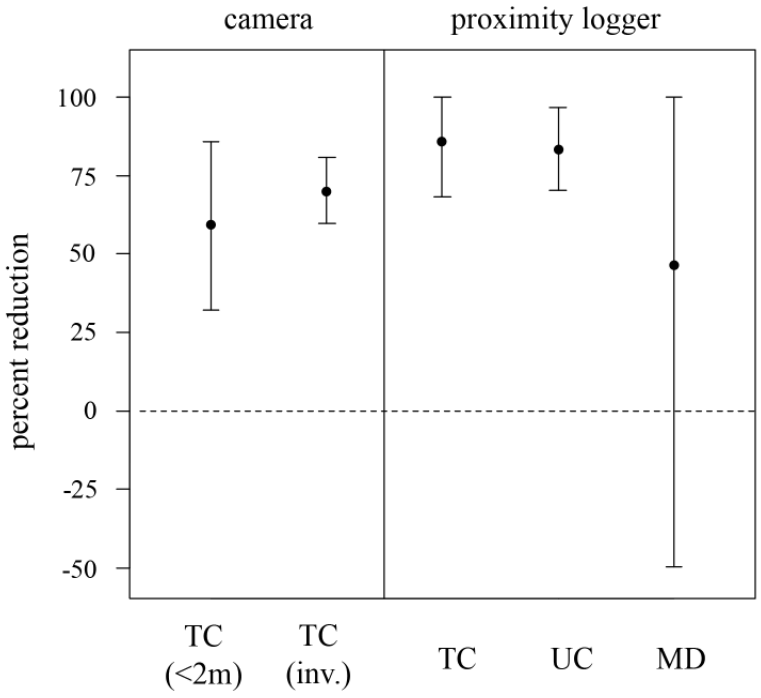

Figure 2-2. Reduction in elk-fetus contact rates for LD feeding relative to HD feeding on Soda Lake feedground, Wyoming. Left: percent reduction in total contacts for <2-m contacts and investigations (inv.) based on camera data. Right: percent reduction in total contacts, unique contacts, and mean duration (all <2-m) based on proximity logger data. Error bars are 95 percent confidence intervals. *n* = 6 for estimates from camera data, *n* = 7 for estimates from proximity logger data. TC = total contacts; UC = unique contacts; MD = mean duration.

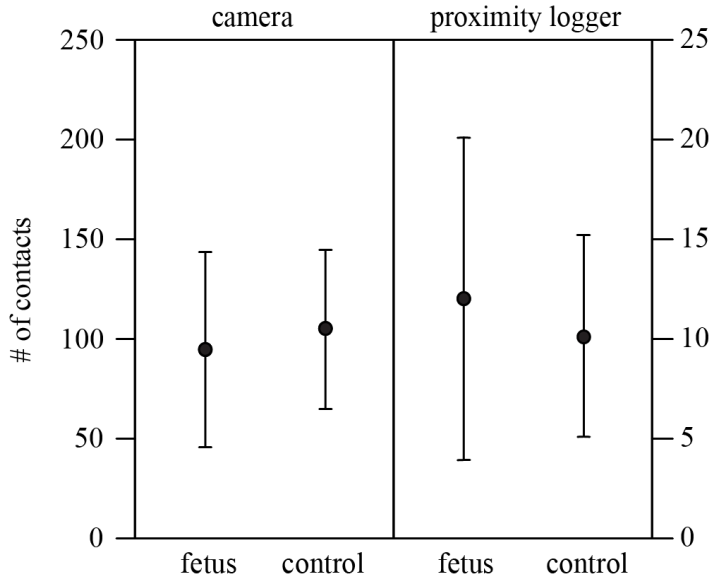

Figure 2-3. Comparison of total contacts with fetus and control based on camera data (left) and proximity logger data (right). Error bars are 95 percent confidence intervals. *n* = 9 for estimates from camera data, *n* = 14 for estimates from proximity logger data.

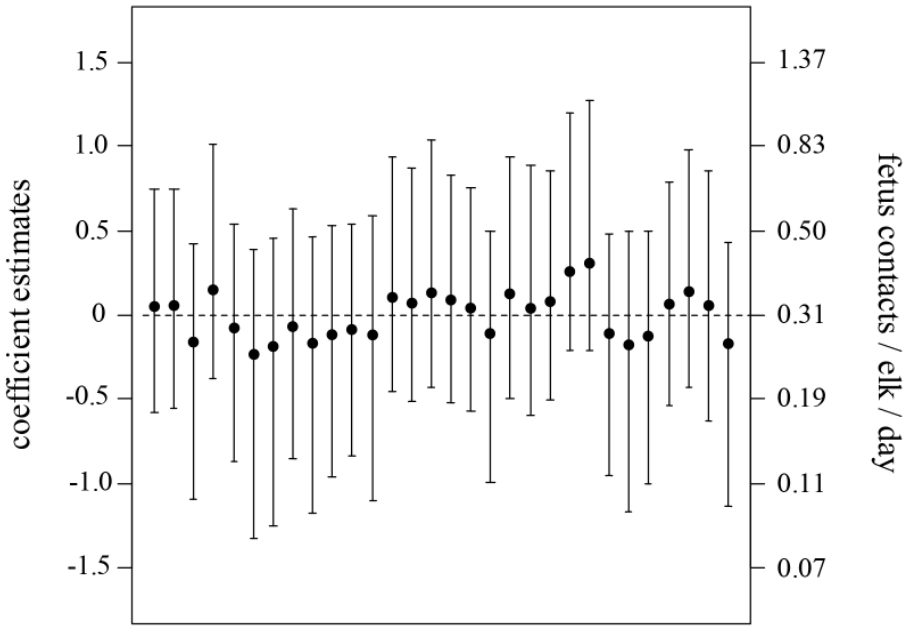

Figure 2-4. Estimated coefficients (α_j) for random effect of individual in mixed model of elk-fetus contact rates. Horizontal dashed line shows overall mean coefficient estimate. Right axis shows corresponding daily elk-fetus contact rates (note non-linear scale). Error bars are 95 percent credible intervals. n = 12 for each estimate of individual effect.

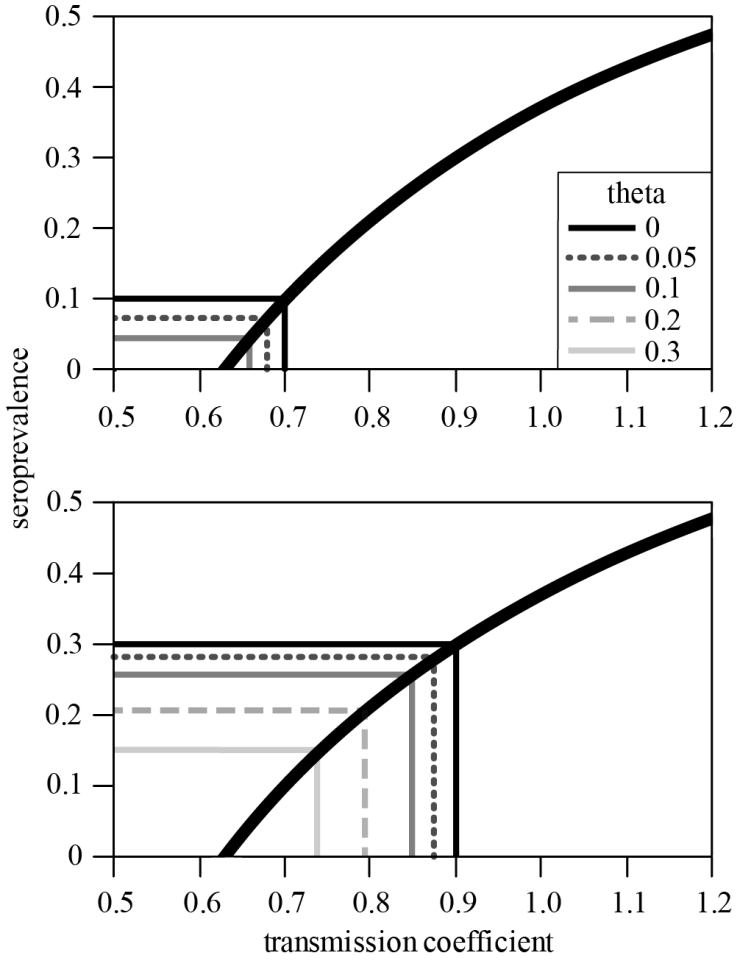

Figure 2-5. Potential reductions in *B. abortus* seroprevalence of Wyoming elk from LD feeding given initial seroprevalence of 0.10 (top) or 0.30 (bottom). Thick line shows relationship between transmission coefficient (β) and equilibrium seroprevalence (P^*) when $\delta = 0.13$ and $\gamma = 0.5$. Thin lines show predicted values of β and corresponding P^* if LD feeding is implemented and the proportion of transmission events occurring along feedlines (θ) is 0, 0.05, 0.1, 0.2, and 0.3. $\theta = 0$ represents initial P^* and β. Lines for $\theta = 0.2$ and $\theta = 0.3$ do not appear in top figure because $P^* = 0$ for these θ values.

CONCLUSION

This research provides strong support for the implementation of low-density feeding throughout Wyoming's elk feedground system, and an example of how information on animal contact rates can lead to improvements in the management of wildlife populations. In the case of Wyoming's feedgrounds, a very simple and cost-effective modification to the current feeding technique is likely to pay large dividends in the effort to reduce brucellosis in elk populations; in many other wildlife disease scenarios, currently-lacking information on animal contact rates could similarly provide the basis for effective disease control measures.

While my research has shown that low-density feeding can dramatically reduce *B. abortus* transmission on feedgrounds, there remains a pressing need for data on transmission away from feedgrounds on native winter range. It is generally believed that the majority of transmission events occur on feedgrounds, but this has never been rigorously examined, and designing region-wide strategies for brucellosis control requires an understanding of all potential sources of transmission. One possible method to obtain data on elk-fetus contact rates on native winter range would be to combine proximity logger technology with vaginal implant transmitter (VIT) technology. VITs are radio transmitters implanted in pregnant elk and expelled during parturition or abortion, allowing researchers to pinpoint the location and timing of such events. If VITs were designed to include a proximity-logging function, and unfed elk were outfitted with proximity loggers, then data on rates of elk-fetus contact away from feedgrounds could be acquired. With such data, we could improve our understanding of the relative

importance of on- and off-feedground *B. abortus* transmission in the region and the likely impacts of feedground management actions like low-density feeding.

This thesis has focused on quantifying elk-to-fetus contact rates in the context of *B. abortus* transmission, but ongoing research in this study system is employing proximity loggers to explore elk-elk contact rates, which are more relevant for diseases transmitted by direct contact between individuals. While directly-transmitted diseases have not captured as much attention in the GYE as brucellosis, they have the potential to seriously affect regional elk populations, and because of Wyoming's supplemental feeding program, impacts could be more severe than those seen in other regions for diseases like chronic wasting disease or bovine tuberculosis. One aspect of the current proximity logger research focuses on estimating the extent to which supplemental feeding elevates elk-elk contact rates above levels experienced by non-fed elk on native winter range. Ongoing research is also addressing an important theoretical question in wildlife epidemiology: what is the relationship between animal group size and contact rate? This question is particularly relevant to disease management of in the GYE given recent evidence that elk group sizes are increasing, but has never been adequately addressed with data from wildlife populations.

Use of proximity loggers in ecological studies is increasing because of the improved data resolution they afford, and it seems likely that they will become a popular tool for studying animal contact rates. It is my hope that the insights and examples provided in this thesis will be helpful to other researchers conducting contact studies using proximity loggers and other emerging technologies.