Retrovirus transmission dynamics in the Florida panther (*Puma concolor coryi*)

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE

UNIVERSITY OF MINNESOTA

BY

Marie Lynn Jones Gilbertson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Meggan E. Craft

January 2021

Acknowledgements

I would like to begin by thanking my advisor, Meggan Craft, who first took me on as a clueless veterinary student and inspires me every day through her empowering mentorship. I would further like to thank my committee members, Dominic Travis, Kim VanderWaal, Tiffany Wolf, and Sue VandeWoude, who all provided me with instrumental guidance both during veterinary and graduate school. Next, I must thank past and current members of the Craft lab who have provided crucial feedback on my work and boundless friendship, including but not limited to: Nick Fountain-Jones, Luis Escobar, Lauren White, Amy Kinsley, Matt Michalska-Smith, Katie Worsley-Tonks, and Janine Mistrick. I owe a huge debt of gratitude to my Ecosystem Health family, especially Tyler Garwood, Marissa Milstein, and Janine Mistrick, for providing me with inspiration and community. Massive thanks to my outstanding collaborators in the Felidae project and at the Florida Fish and Wildlife Conservation Commission: none of this research would be possible without your support and enthusiasm. I must also thank the College of Veterinary Medicine's Summer Scholars program and DVM/PhD dual-degree program (in particular, Mark Rutherford) for foundational experiences and support.

None of this work would have been possible without generous financial support. Thank you to the Comparative Medicine and Pathology NIH T32 program (especially Molly McCue, Cathy Carlson, Dave Brown, and Yorie Smart). Thank you also to the University of Minnesota Informatics Institute, and their MnDRIVE graduate assistantship. I am also honored to have received generous support from the Van Sloun Foundation, which has shown an inspiring commitment to Ecosystem Health and the conservation of the Florida panther. I am grateful for the support of numerous scholarships and travel grants which have supported my research and travel to share my work: the PEO Scholar

Award program, the Institute on the Environment mini grant program, the College of Veterinary Medicine Graduate Student travel award, the University of Minnesota Council of Graduate Students travel award, the Minnesota Veterinary Medical Foundation, and the American Association of Wildlife Veterinarians.

Finally, unending thanks to my husband, Matt, who is my partner in all things and whose kindness, thoughtfulness, and support (both emotional and technical) has made the last 7 ½ year journey through veterinary and graduate school possible. I cannot thank my parents enough for their love and unfailing faith in me. Thank you to Kristine for your endless understanding and humor. Thanks to Becky, Desiree, and Kaitlin for preserving my sanity. And thank you to Boe, Brandy, Alice, and Astrid: you make it all worth it.

Dedication

This thesis is dedicated to: my parents, Debbie and Gary, who always believe in me; to my husband Matt, who is the perfect partner (and makes sure the house and I don't fall apart); to Kristine, who laughs with me until we cry; and to Boe, Brandy, Alice, and Astrid who love me as I am.

Abstract

Outbreaks of infectious diseases can have major consequences for public health, food security, and conservation, yet drivers of pathogen transmission are often poorly understood. In transmission models, transmission is generally considered to be a product of contact rates and the probability of transmission, given contact (i.e., transmissibility). Contact rates and patterns are generally easier to empirically observe, particularly with expansions in remote sensing technologies. However, these technologies are imperfect and sampling guidelines for their use in disease studies are generally lacking. Further, both contact rates and transmissibility can be modified by heterogeneities important for subsequent pathogen transmission. For example, transmissibility can be affected by heterogeneities in host defenses, body condition, etc., and many of these factors may even covary with heterogeneities in contact rates. This complexity can thereby muddy our understanding of the ultimate drivers of transmission processes. A holistic approach is therefore needed to move beyond these limitations and help explain *why* individuals interact and transmit. In this dissertation, I fill this gap by reviewing (Chapter 1) and testing (Chapters 2 and 3) innovative methods for determining drivers of transmission in natural systems, predominantly focusing on a naturally occurring model (representative) system: feline retroviruses in the Florida panther (*Puma concolor coryi*). I then implement this new knowledge in an applied context: optimizing pathogen control in endangered panthers (Chapter 4). The findings reported here can improve the ability to identify drivers of transmission across a range of host-pathogen systems, and represent important progress for improving outbreak prevention and management for the benefit of human, animal, and ecosystem health.

Table of Contents

Acknowledgements i

Dedication iii

Abstract iv

Table of Contents v

List of Tables viii

List of Figures ix

Introduction 1

Chapter 1. Incorporating genomic methods into contact networks to reveal new insights into animal behavior and infectious disease dynamics 7

1.1 Synopsis 7

1.2 Interplay between behavior and infectious disease 8

1.3 Contact networks elucidate dynamics of behavior and disease ecology 11

1.4 Genomic tools to understand behavior and pathogen transmission 16

1.5 Integrating contact networks and genomics: current efforts and future directions 22

1.6 Conclusions 33

1.7 Acknowledgements 33

1.8 Figures 35

1.9 Tables 39

Chapter 2. Trade-offs with telemetry-derived contact networks for infectious disease studies in wildlife 42

2.1 Synopsis 42
2.2 Introduction 43
2.3 Materials and Methods 47
2.4 Results 53
2.5 Discussion 56
2.6 Acknowledgements 65
2.7 Data Availability 66
2.8 Figures 67

Chapter 3. Transmission of one predicts another: Apathogenic proxies for transmission dynamics of a fatal virus 71

3.1 Synopsis 71
3.2 Introduction 72
3.3 Materials and Methods 75
3.4 Results 84
3.5 Discussion 85
3.6 Acknowledgements 93
3.7 Tables 94
3.8 Figures 96

Chapter 4. Paradoxes and synergies: optimizing management of a deadly virus in an endangered carnivore 100

4.1 Synopsis 100
4.2 Introduction 101
4.3 Materials and Methods 104
4.4 Results 110

4.5 Discussion 114
4.6 Acknowledgements 121
4.7 Figures 122
Conclusion 127
Bibliography 129
Appendix A. Supporting information for Chapter 2 162
A.1 Materials and Methods 162
A.2 Results 165
A.3 Figures 170
A.4. Tables 200
Appendix B. Supporting information for Chapter 3 202
B.1 Materials and Methods 202
B.2 Results 208
B.3 Tables 212
B.4 Figures 214
Appendix C. Supplementary information for Chapter 4 226
C.1 Materials and Methods 226
C.2 Results 229
C.3 Discussion 233
C.4 Tables 235
C.5 Figures 237

List of Tables

Table 1.1. Advantages and disadvantages of currently used and proposed methods for integrating network and genomic approaches. 39

Table 3.1: Network and transmission simulation parameters 94

Table 3.2: Main FIV transmission network exponential random graph model results..... 95

Table A.1. Parameter values used for the BCRW movement model and corresponding average home range sizes. 200

Table A.2. Contact rates in complete networks by spatial configuration and home range size when using a space-time contact definition with 100m distance threshold. 200

Table B.1: Best ERGM results for each FIV transmission network..... 212

Table B.2: Fixed effects results from model-type performance GLMM..... 213

Table C.1: Network and transmission simulation parameters..... 235

Table C.2: Post-Hoc Sensitivity Analysis Parameter Sets 236

List of Figures

Figure 1.1. Conceptual flow from a contact network to a transmission network. 35

Figure 1.2. Conceptual flow from (a.) a transmission tree to (b.) a transmission network; (c.) depicts a transmission network in the context of the rest of the population, where gray nodes represent uninfected individuals. 36

Figure 1.3. Two examples of proposed integrations of contact networks and genomic tools, in this case focusing on utilizing transmission networks derived from inference of transmission trees. 37

Figure 2.1: Workflow for simulations, sampling, and network estimation. 67

Figure 2.2: Heat maps demonstrating the lower limit of the 95% confidence interval for mean correlation coefficient results for betweenness and strength, the best performing global and local node-level network metrics, respectively. 68

Figure 2.3: Heat maps of network strength (lower limit of the 95% confidence interval for mean correlation coefficient), which showed the strongest improvement with a spatial overlap contact definition, for medium home range simulations at the coarsest frequencies of sampling. 69

Figure 2.4: Heat maps of the lower limit of the 95% confidence interval of mean correlation coefficient for degree. 70

Figure 3.1: Conceptual workflow across all analysis steps. 96

Figure 3.2: Diagram of flows between compartments in transmission model. 97

Figure 3.3: FIV-based networks perform at least as well as other models in predicting number of progressive infections, as seen in violin plots of median total number of progressive infections from parameter sets classified as “feasible.” 98

Figure 3.4: SaTScan cluster analysis for feasible FIV-based and overlap-based network simulations show stronger agreement between empirical observations (red horizontal lines) and FIV-based predictions for (A) predicted FeLV cluster size and (B) Observed/Expected FeLV cases associated with the top detected cluster. 99

Figure 4.1: Steps of the simulation process across baseline and management scenarios. 122

Figure 4.2: Histograms of FeLV mortalities in simulated epidemics with proactive vaccination alone. 123

Figure 4.3: Histograms of FeLV mortalities in simulated epidemics with both reactive and proactive vaccination. 124

Figure 4.4: Histograms of FeLV mortalities in simulated epidemics with both reactive test-and-removal and proactive vaccination. 125

Figure 4.5: Epidemic curves showing the proportion of the population progressively infected over time when 60% of the population was proactively vaccinated. 126

Figure A.1: Examples of different spatial configurations used for simulations. 170

Figure A.2: Expanded flow chart of simulation methods and network comparisons (also see Figure 2.1 in main text). 172

Figure A.3: Betweenness: Results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). 173

Figure A.4: Betweenness: Results are shown across three movement model parameter sets (small home ranges II-IV; SHR II-SHR IV, respectively). 176

Figure A.5: Transitivity: Results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). 177

Figure A.6: Transitivity: Results are shown across three movement model parameter sets (small home ranges II-IV; SHR II-SHR IV, respectively). 179

Figure A.7: Degree: Results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). 180

Figure A.8: Degree: Results are shown across three movement model parameter sets (small home ranges II-IV; SHR II-SHR IV, respectively). 182

Figure A.9: Strength: Results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). 184

Figure A.10: Strength: Results are shown across three movement model parameter sets (small home ranges II-IV; SHR II-SHR IV, respectively). 186

Figure A.11: Density: Results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). 188

Figure A.12: Density: Results are shown across three movement model parameter sets (small home ranges II-IV; SHR II-SHR IV, respectively). 190

Figure A.13: Proportion isolates (evaluated as proportion connected): Results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). 192

Figure A.14: Proportion isolates (evaluated as proportion connected): Results are shown across three movement model parameter sets (small home ranges II-IV; SHR II-SHR IV, respectively). 194

Figure A.15: Modularity from unweighted walktrap community-finding algorithm: Results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). 196

Figure A.16: Modularity from weighted walktrap community-finding algorithm: As with figure A.13, results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). 198

Figure B.1: Phyloscanner-derived main FIV transmission network. 214

Figure B.2: Best ERGM for main FIV network showed reasonable goodness of fit across standard goodness of fit metrics. 216

Figure B.3: Fitting the predictors from the best ERGMs to random networks (based on the main FIV network) shows that all three best models give largely consistent coefficient estimates. 217

Figure B.4: Map of observed panther centroid locations and FeLV status (red = qPCR positive, blue = qPCR negative). 218

Figure B.5: Boxplots of total number of progressive infections from parameter sets classified as “feasible.” 219

Figure B.6: Boxplots of duration of simulated epidemics from parameter sets classified as “feasible.” 220

Figure B.7: Observed/Expected test statistics (T_k) for Cuzick-Edwards tests performed with FIV and overlap-based predictions of FeLV transmission. 221

Figure B.8: Variable importance plots for the FIV-based model. 222

Figure B.9: Partial dependence plots for the FIV-based model, ordered based on variable importance observable in Figure B.8 224

Figure B.10: Partial dependence plots for the C parameter, which was a constant multiplier for probability of transmission from regressives, given effective contact. 225

Figure C.1: Compartmental model diagram of FeLV transmission model 237

Figure C.2: Histograms of FeLV mortalities in simulated epidemics with proactive vaccination alone. 239

Figure C.3: Histograms of the duration of simulated FeLV epidemics (in weeks) with proactive vaccination alone. 241

Figure C.4: Scatterplots of simulated FeLV epidemic durations against mortalities with proactive vaccination alone. 242

Figure C.5: Histograms of FeLV mortalities in simulated epidemics with proactive vaccination and year-round reactive vaccination. 245

Figures C.6: Histograms of FeLV mortalities in simulated epidemics with proactive vaccination and 6-months per year reactive vaccination. 247

Figure C.7: Histograms of the duration of simulated FeLV epidemics with proactive vaccination and year-round reactive vaccination. 249

Figure C.8: Histograms of the duration of simulated FeLV epidemics with proactive vaccination and 6-months per year reactive vaccination. 251

Figure C.9: Histograms of FeLV mortalities in simulated epidemics with both reactive test-and-removal and proactive vaccination. 253

Figure C.10: Histograms of the duration of simulated FeLV epidemics with both reactive test-and-removal and proactive vaccination. 255

Figure C.11: Histograms of FeLV mortalities in simulated epidemics with both reactive underpass closures and proactive vaccination. 257

Figure C.12: Histograms of the duration of simulated FeLV epidemics with both reactive underpass closures and proactive vaccination. 259

Figure C.13: Heat maps of the proportion of simulated FeLV epidemics that failed (fewer than 5 progressive or regressive infections) per 100 successful epidemics. 261

Figure C.14: Histogram of median simulated FeLV mortalities for each of 50 sensitivity analysis parameter sets evaluated under a no-intervention scenario. 262

Figure C.15: Partial Rank Correlation Coefficient (PRCC) estimates for simulation parameters in no-intervention scenario sensitivity analysis. 263

Figure C.16: Partial Rank Correlation Coefficient (PRCC) estimates for simulation parameters in proactive vaccination scenario sensitivity analysis. 264

Figure C.17: Scatterplot of difference in median mortalities with and without proactive vaccination (“mortalities difference” per parameter set) against network density parameterization. 265

Figure C.18: Plot of proactive vaccination sensitivity analysis parameter set classifications across all pairs of parameters. 267

Figure C.19: Scatterplot of post-hoc sensitivity analysis evaluation of influence of proportion adults (Adult_prop) and network density (Net_dens) parameters.... 268

Figure C.20: Scatterplot of *post-hoc* sensitivity analysis results, here showing the proportion adults (Adult_prop) against the weekly probability of contact (Omega/ω) parameters 269

Introduction

Outbreaks of infectious diseases pose major threats to public health, economics, food security, and conservation (e.g., Cleaveland, 2009; Dobson et al., 2020; Jones et al., 2008; Knight-Jones & Rushton, 2013; Wiethoelter et al., 2015). Preventing and interrupting these outbreaks is therefore critical for human, animal, and ecosystem wellbeing. A critical gap in disease control, however, is our generally poor understanding of the underlying drivers of pathogen transmission. Mathematical models provide a useful framework for considering the complexities affecting observed transmission dynamics and help reveal why identifying drivers of transmission is such a formidable challenge. These models consider transmission to be a product of contact rates and the probability of transmission, given contact (hereafter, transmissibility; Anderson & May, 1991). While traditional disease models have assumed well-mixed populations (i.e., homogeneous mixing), these assumptions do not hold for many host-pathogen systems (Keeling & Eames, 2005; Lloyd-Smith et al., 2005). Rather, both contact rates and transmissibility can be affected by heterogeneities that have significant consequences on transmission dynamics (VanderWaal & Ezenwa, 2016).

Contact rates and patterns are generally easier to empirically observe than individual transmissibility. Heterogeneities in host behavior, personality, space use, seasonality, etc. can all affect host contact rates and subsequent pathogen transmission (VanderWaal & Ezenwa, 2016). Network approaches, which are able to incorporate such contact heterogeneities, have therefore been a critical development in disease modeling (Craft, 2015; Craft & Caillaud, 2011; Keeling & Eames, 2005; White et al., 2017). In addition, ongoing advances in remote sensing technology (e.g., GPS telemetry devices) show great promise for effectively documenting heterogeneous contact behaviors in

otherwise difficult-to-observe hosts (Krause et al., 2013). However, gaps remain regarding best practices for the use of such technologies in disease transmission studies, particularly when using network approaches which may be sensitive to missing data.

Heterogeneities in transmissibility are often more difficult to elucidate, but can result from variations in factors such as host defenses, body condition, parasite burden, pathogen virulence, etc. (VanderWaal & Ezenwa, 2016). Further complicating matters, covariation between heterogeneities in contact rates and transmissibility can also arise (White et al., 2018a). For example, space use can produce variation in contact rates (Robert et al., 2012), but also in access to resources that may promote improved host body condition and immune defenses (Becker et al., 2015). These complex dynamics highlight that a holistic approach is needed to move beyond describing individual heterogeneities in contact rates and transmissibility and instead explain why individuals interact and transmit pathogens.

An emerging potential solution to this gap is the exploration of apathogenic infectious agents (hereafter *apathogenic agents*; Archie & Tung, 2015). Several wildlife systems have demonstrated that apathogenic agents within the healthy host microbiome can serve as markers of close social contact (Blasse et al., 2013; Bull et al., 2012; Springer et al., 2016; VanderWaal et al., 2014a), though this has not been true in all cases (e.g., Chiyo et al., 2014). However, these agents have not yet been tested as proxies of direct transmission processes. To evaluate apathogenic agents in this role requires the use of a model (representative) system, for which the endangered Florida panther (*Puma concolor coryi*) is an excellent candidate. Panthers are affected by several directly transmitted feline retroviruses, including apathogenic feline

immunodeficiency virus (FIV) and pathogenic feline leukemia virus (FeLV). FIV, in particular, is expected to have a relatively high intrahost mutation rate, making this virus more amenable to transmission inference between hosts (Grenfell et al., 2004; Hall et al., 2016; Krakoff et al., 2019). In this system, FIV may be able to act as a marker of direct transmission events relevant to analogously transmitted FeLV, and could thereby prospectively predict FeLV transmission dynamics. Further, because panthers experienced an outbreak of FeLV in 2002-2004 (Cunningham et al., 2008), empirical data exists against which to compare FIV-based predictions of FeLV transmission. In this way, transmission dynamics of feline retroviruses in the Florida panther can provide important insight into the utility of apathogenic agents as a holistic tool for understanding drivers of pathogen transmission. Consequently, the overall objective of this dissertation is to test innovative methods for determining drivers of transmission in natural systems. Here, I have primarily focused on the naturally occurring model system of the Florida panther to explore approaches that can translate more generally to other wildlife, domestic animal, and human systems.

In Chapter 1, I reviewed opportunities for integrating network and genomic approaches to better understand animal behavior and pathogen transmission dynamics. Network approaches spanning social network analysis, network statistics, and network modeling are all key tools to link with genomic data in order to elucidate principles of transmission. Recent advances in phylodynamics and transmission inference are particularly powerful genomic approaches and may be highly amenable to linking with network tools. In particular, I highlighted potential advances to be gained by linking high resolution pathogen genetic data with network approaches to (1) determine which

behaviors are most important for pathogen transmission, and (2) determine drivers of transmission-relevant contact.

A key component for successful network modeling—and for disease modeling more generally—is understanding transmission-relevant contact in a population of interest. Such determinations, however, are particularly challenging in difficult-to-observe wildlife species. As discussed above, a suite of remote wildlife monitoring tools (e.g. telemetry) are increasingly used to derive contact rates and patterns from wildlife, but with ongoing uncertainty regarding how remote sampling effort affects subsequent analysis of contact patterns.

In Chapter 2, I addressed this gap in knowledge by determining sampling guidelines for remote telemetry detection of contact and subsequent network generation in free-ranging wildlife. Here, I simulated movement of populations of individuals, sampled from these movement trajectories, and used this movement data to generate contact networks. I compared social network analysis metrics relevant to disease transmission between networks derived from sampled and unsampled movement data to determine consistency of network structures under different regimes of sampling and contact definitions. Results from this chapter demonstrate that (1) local network metrics are more robust to telemetry sampling than global network metrics; (2) populations with infrequent interactions may require the most intensive sampling; and (3) defining contact in terms of spatial overlap may sometimes be able to compensate for coarse telemetry sampling.

Even when scientists can accurately capture contact patterns in wildlife—remotely or otherwise—this may not resolve complexities in understanding why individuals transmit pathogens to each other, as this strategy focuses on only one

component of the transmission process. In Chapter 3, I delved deeper into this issue by using a holistic approach focused on identifying underlying drivers of transmission processes. Here, I linked genomic and network approaches to test an apathogenic viral agent as a proxy for direct contact transmission in my afore-described naturally occurring system, the Florida panther. I first determined drivers of transmission of the apathogenic agent, FIV, and then used these drivers to predict likely transmission pathways for analogously transmitted FeLV in panthers. I simulated FeLV transmission along these pathways and compared predicted transmission to the empirical FeLV outbreak in panthers. Results showed that (1) FIV transmission is driven primarily by geographic distances between panther home range centroids and age classes, and (2) that FIV-based FeLV modeling performed at least as well as simpler, typically retrospective modeling approaches. Apathogenic agents like FIV therefore show promise for prospectively determining drivers of transmission of important pathogens, which would allow for improved proactive disease control planning across a wide range of host-pathogen systems.

Endangered Florida panthers continue to be at risk of FeLV transmission (Chiu et al., 2019), and wildlife managers are hampered by uncertainty regarding best practices for FeLV prevention and response. Using my new knowledge about drivers of retrovirus transmission in panthers, I addressed this gap in Chapter 4, in which I tested several FeLV prevention and response strategies *in silico* to provide practical, evidence-based guidance to panther managers. Here, I simulated FeLV transmission in the contemporary panther population in the absence of interventions, and with a range of proactive and reactive FeLV interventions. These interventions included proactive vaccination, reactive vaccination, reactive test-and-removal, and reactive closure of

wildlife highway underpasses. Importantly, I also considered the effect of partial immunity induced by single and boosted FeLV vaccination. Results found (1) counterintuitive increases in FeLV mortality with inadequate proactive vaccination, likely a result of partial vaccine immunity; (2) synergistic effects of combining proactive vaccination with reactive vaccination or test-and-removal; and (3) limited success of underpass closures for mitigating FeLV impacts under realistic conditions. These findings aid in the conservation and management of an endangered carnivore, and highlight the benefits of linking modeling and management considerations to identify key costs and benefits to different disease management strategies.

Chapter 1. Incorporating genomic methods into contact networks to reveal new insights into animal behavior and infectious disease dynamics

Gilbertson, M. L., Fountain-Jones, N. M., & Craft, M. E. (2018). Incorporating genomic methods into contact networks to reveal new insights into animal behaviour and infectious disease dynamics. *Behaviour*, 155(7-9), 759-791. doi: 10.1163/1568539X-00003471

1.1 Synopsis

Utilization of contact networks has provided opportunities for assessing the dynamic interplay between pathogen transmission and host behavior. Genomic techniques have, in their own right, provided new insight into complex questions in disease ecology, and the increasing accessibility of genomic approaches means more researchers may seek out these tools. The integration of network and genomic approaches provides opportunities to examine the interaction between behavior and pathogen transmission in new ways and with greater resolution. While a number of studies have begun to incorporate both contact network and genomic approaches, a great deal of work has yet to be done to better integrate these techniques. In this review, we give a broad overview of how network and genomic approaches have each been used to address questions regarding the interaction of social behavior and infectious disease, and then discuss current work and future horizons for the merging of these techniques.

1.2 Interplay between behavior and infectious disease

The study and utilization of mathematical models has been revolutionary to the field of disease ecology. Mathematical models can be used to examine mechanisms of pathogen transmission and maintenance, and to make predictions about pathogen spread and management (Keeling & Rohani, 2011; Lloyd-Smith et al., 2009). Initially, models of infectious disease transmission assumed no contact heterogeneity among hosts (Anderson & May, 1991; Keeling & Rohani, 2011). This simplicity can be useful for creating transparent, generalizable models, but may not be appropriate for all systems, as host contact heterogeneity can have significant impacts on disease dynamics (Keeling & Eames, 2005; Meyers, 2007; White et al., 2017). For example, "superspreaders" can have important effects on epidemic outcomes, with their presence resulting in disease outbreaks that are less frequent but more severe than predicted by homogeneous mixing of populations (Lloyd-Smith et al., 2009; Lloyd-Smith et al., 2005). By incorporating contact heterogeneity into model assumptions, contact or "social" networks, applied to disease models, can be important for conveying a more informative picture of pathogen transmission dynamics (Craft, 2015; Craft & Caillaud, 2011; Godfrey, 2013; White et al., 2017).

Contact networks represent connections between individuals or groups of individuals based on a variety of definitions of "contact," and these contacts are frequently influenced by social behavior. Network models can therefore require more intensive data collection compared to other techniques to model disease spread, such as compartmental models, but can also provide unique insight into disease dynamics thanks to their incorporation of contact heterogeneity. While social network models have been much used in studies of infectious disease in human and livestock systems, they

have experienced delayed but growing utilization in wildlife systems (Godfrey, 2013; Martínez-López et al., 2009).

Animal behavior influences contact heterogeneity, more specifically through heterogeneity in rates and patterns of social interaction (VanderWaal & Ezenwa, 2016). Some of the behavioral or social processes that impact contact heterogeneity include individual behavioral phenotypes (Dizney & Dearing, 2013; Natoli et al., 2005), the social structure of a population (Nunn et al., 2015; Sah et al., 2017), seasonal or temporal behavior changes (Chen et al., 2014), and behavioral responses to anthropogenic influences (Becker et al., 2015; Gottdenker et al., 2014). Behavior is therefore important for pathogen transmission, and particularly so for directly transmitted pathogens, where environmental contamination or vectors do not dilute the effect of animal interactions (Godfrey, 2013; VanderWaal & Ezenwa, 2016; White et al., 2017).

While network approaches have made much progress in shaping our understanding of the complicated, dynamic interplay between behavior and infectious disease, they also reveal new, more nuanced questions, demonstrating the need for additional data and methods to better illuminate the processes at work within disease ecology. Genomic methods, applied to host or pathogen, are becoming increasingly recognized for their power to provide new insights into disease ecology (Archie et al., 2009). Genetic and genomic tools can allow us to study and infer pathogen transmission (Kao et al., 2014; Metzker et al., 2002), reconstruct epidemics (Biek et al., 2007; Bird et al., 2007; Sharp & Hahn, 2010), and reveal landscape and environmental factors important to transmission (Archie et al., 2009; Blanchong et al., 2007). Genomic methods are also becoming more available and affordable, putting them within the reach of more researchers. Within the scope of wildlife studies, genomic tools may also

provide important opportunities to maximize the use of non-invasive sampling techniques. Non-invasive techniques are expanding in their range and utility, and may be particularly important for studying populations of endangered or at-risk species, where capture and handling of individuals carries increased risk to overall population health (de Carvalho Ferreira et al., 2014; Hoffmann et al., 2016; Smiley Evans et al., 2016).

To improve predictions and accuracy, network models should be informed by as much of the available data as possible (Welch et al., 2011). While genetic characterization of pathogens can suggest that cases of infection are closely related, many techniques (e.g. strain typing, microsatellite markers, etc.) provide only coarse resolution for inferring transmission (Kao et al., 2014). Whole genome sequencing (WGS), on the other hand, provides significantly more refined data for resolving transmission relationships (Kao et al., 2014). In the context of investigating the dynamics between pathogen transmission and host behavior, it may seem counterintuitive to study pathogens in order to investigate host social structure, but the idea is not necessarily a new one (Welch et al., 2011), and doing so could provide important insights into behavior and disease ecology when integrated with contact networks. Given the increasing availability of genomic sequencing, as well as the expanding utility of non-invasive sampling in wildlife systems, there is an opportunity to incorporate pathogen genomic data into contact networks.

In this review, we discuss how contact networks and genomic methods provide insights into animal behavior and infectious disease transmission, and horizons for the integration of these methods in the future. More specifically, we discuss historical uses of social network analysis and network modeling, and how they are able to highlight the

interplay between social behavior and infectious disease. We then describe how pathogen genetic and genomic techniques have been utilized to make inferences about host behavior and pathogen transmission. Finally, we explore the merging of contact network and genomic methods, and illustrate future directions of network modeling in light of the increasing availability of sequencing tools. Given that much can be gained by utilizing methods from other systems, this review will not limit examples to those from the animal or wildlife literature, but will draw from the human infectious disease literature as well.

1.3 Contact networks elucidate dynamics of behavior and disease ecology

The basics of contact networks

Within network approaches, “nodes” can represent individual humans or animals, or can represent groups of individuals such as herds or even farms. Connections or contacts between nodes are called “edges,” and are defined depending on the research question; for example, in a network where nodes represent individual wild animals, and the pathogen in question is transmitted by sexual contacts, an edge might be drawn between two individuals known to have had sexual contact. Alternatively, in a network of farms, with a pathogen transmitted by direct or indirect contact, an edge might be drawn between two farms if they have shipped animals from one farm to the other. This highlights the importance of carefully defining a “contact” when using network approaches, depending on the system and research question (Perkins et al., 2009).

While social network analysis approaches are often descriptive or observational, network modeling approaches typically involve simulating an outbreak of a pathogen in a population (but see following section for caveats about this dichotomy). Simulating an

outbreak of a pathogen on a contact network, in light of the transmission probability (e.g. properties of the infectious host, susceptible host, and infectivity of the pathogen), can result in the output of a transmission network which can graphically represent who infected whom (Figure 1.1). This transmission network is therefore different from the observed contact network, and could be viewed as a subset of the contact network. Network approaches are, however, dependent on extensive behavioral or observational data to create the initial contact network, and can be highly influenced by how contacts are defined and detected in the population (Eames et al., 2015; Perkins et al., 2009). In some cases, the entire population being studied is observed to build the "complete" contact network; in other cases, this is not feasible or practical, and only a portion of the population of interest will be observed. In these cases, the observed population can be used to create a set of "rules" for contact in the population, and then extrapolated to the full population (e.g. with ERGMs, as in Reynolds et al., 2015; Silk et al., 2017). Thus, even networks built on a sample of empirical observations may be generalized to a larger population.

Social network analysis and network modeling

The use of contact networks could be thought of as being composed of two main "branches," social network analysis (SNA) and network modeling, though these methods are by no means mutually exclusive. Broadly, both SNA and network modeling involve building a social network based on behavioral or observational data. SNA then generally assesses network and node-level metrics that describe connectivity and modularity (Croft et al., 2008; Perkins et al., 2009), while network modeling often describes or simulates an outbreak of an infectious disease on the network. Network modeling is

similar in concept to compartmental “Susceptible-Infectious-Recovered”-type (SIR) dynamic models, but instead of assuming homogeneous mixing, network modeling allows for the incorporation of heterogeneous contact patterns. Of course, our framework for describing contact network approaches is a somewhat artificial simplification, and many studies will not obviously fall into one category or the other. Network approaches could also be viewed as a spectrum ranging from describing networks for making observations and formulating hypotheses, to using statistical models to begin testing these hypotheses, through to simulation models of epidemics on networks to make predictions and test understanding. For the purposes of this review, we will discuss network approaches in the context of observational or descriptive approaches (SNA) and simulation-based approaches (network modeling), but with the acknowledgement that network approaches do not always fall neatly into such a dichotomy, and often involve a continuum of approaches.

In the context of infectious disease research, SNA often compares a node’s (individual or group of individuals) position in the network and node-level characteristics to the disease status of the node (Drewe, 2010; Godfrey, 2013; MacIntosh et al., 2012; VanderWaal et al., 2014a). This approach can provide insight into how social behavior impacts pathogen transmission or prevalence, identify the importance of different behaviors for pathogen transmission, and elucidate the impact of disease prevention and management decisions. For example, in studying how social behavior impacts pathogen prevalence, MacIntosh et al. (2012) found that high social rank in Japanese macaques (*Macaca fuscata yakui*) was associated with greater nematode parasite species richness, as well as a higher probability and intensity of infection with a potentially pathogenic parasite. SNA can also be used to prioritize individuals to target for disease

management interventions, or identify interventions that are less likely to be effective for disease control, based on population social structure (Porphyre et al., 2008; Rushmore et al., 2013).

Similarly, simulating epidemics of pathogens on social networks can accomplish a variety of objectives related to animal behavior and infectious disease. For example, it may not be feasible, or even ethical, to test pathogen spread and control hypotheses with real-world experiments, but modeling these experiments can provide important insights (Craft, 2015; Lloyd-Smith et al., 2009; White et al., 2017). Simulating epidemics on networks can thereby shed light on topics such as how disease is or is not maintained in populations (Craft et al., 2009), how population social structure impacts pathogen transmission (Craft et al., 2011), and what techniques or individuals to target for preventive or epidemic intervention measures (Pellis et al., 2015; Rushmore et al., 2014). For example, network models of pathogen transmission in chimpanzees found that targeting well-connected individuals allows for significant reductions in vaccinations required to prevent disease outbreaks (Rushmore et al., 2014).

Applications of contact networks are not without limitations, however. Observational data used to develop a social network may be limited (e.g. only able to collect behavioral observations during daytime hours, or only able to place GPS collars on a subset of the population), which can significantly impact conclusions (Croft et al., 2008). In addition, proxies of contact such as shared space use—as determined by telemetry data—may not always represent actual contacts; for example, avoidance behaviors could theoretically allow individuals in close proximity to avoid contacts relevant to pathogen transmission. While degree of home range overlap may correlate with increased contact rates between individuals, these assumptions are rarely tested

and may vary seasonally or on an individual basis (Robert et al., 2012). There is therefore a need to improve or refine our detection or definition of transmission-relevant contacts, particularly when considering host species and/or pathogens in which timing of transmission events is not readily observed. In such cases, the increased resolution provided by molecular approaches applied to pathogens may be able to improve our understanding of disease transmission at a fine scale (Kao et al., 2014), and provide clues about microbial transmission in the context of host social structure (Blasse et al., 2013; Blyton et al., 2014; Bull et al., 2012).

Another challenge within the use of contact networks is the fact that social networks are generally dynamic and can vary through time (Craft & Caillaud, 2011; Eames et al., 2015; Rushmore et al., 2013; White et al., 2017). This variability highlights the importance of investigating social networks and infectious disease within appropriate temporal scales, according to the research question (Perkins et al., 2009; White et al., 2017). If the pathogen in question is infectious on the order of days to weeks, contact networks aggregated from year-long observations may not be appropriate and may overestimate pathogen spread. Over the course of a year, a social network is likely far more well-connected than it would be on the days to weeks-long scale on which the chosen pathogen may be operating. However, if the objective is to predict a “worst case scenario” of transmission, an aggregated network may be appropriate. Understanding the different temporal dynamics of host and pathogen within the context of a specific research question is necessary for appropriate study design and interpretation of model simulations. Because pathogen evolution may operate on different time scales from dynamic social networks, when pathogen phylogenies are assessed together with contact networks, they may shed light on the dynamics of transmission within

populations (Vasylyeva et al., 2016). Thus, pathogen phylogenetics may serve a complementary function in balancing the complicated temporal dynamics of many social networks.

In summary, contact network approaches have proven useful for illuminating aspects of the relationship between animal behavior and pathogen transmission by describing how social structure, individual behaviors, and intra- and interspecies interactions contribute to disease transmission and maintenance in populations. Network approaches have also assisted in improving understanding of pathogen dynamics and management in populations, highlighting the importance of contact networks as a tool with real-world applications. An ongoing challenge for the utilization of contact networks is the detection of transmission-relevant contacts, for which genomic tools may provide new avenues for resolving these interactions.

1.4 Genomic tools to understand behavior and pathogen transmission

In this section, we will be reviewing how genetic and genomic-based methods have been used to provide insight into host social behavior and its impact on infectious disease dynamics. While genetics is, broadly, the study of individual genes, and genomics the study of whole genomes, for simplicity and consistency, we will refer to genetic and genomic-based techniques as “genomics” throughout the rest of this review, unless individual examples call for greater specification. We will use “phylogenetics,” the study of evolutionary relationships based on genetics, to encompass both phylogenetic and phylogenomic techniques.

Population genomic tools are well-recognized for their ability to provide important information about host behavior and the consequences of animal behavior on pathogen

transmission. For example, Pope et al. (2007) used population genomic tools to show increased badger dispersal after culling, and inferred important consequences of this increased movement for the spread of bovine tuberculosis. Population genomics have also been used to assess landscape barriers to host dispersal, and the importance of these dispersal behaviors for disease transmission: for example, identifying rivers as semi-permeable barriers to host gene flow, and inferring the impact of rivers on subsequent animal movement behaviors and interactions in the context of pathogen transmission (Cullingham et al., 2009).

Pathogen phylogenetics can also provide insights into animal behavior and disease dynamics, and in some cases give greater clarity than host population genomics alone (Lee et al., 2012). With pathogen phylogenetics, researchers can better understand pathogen transmission dynamics and suggest patterns of host behavior and interactions (Biek et al., 2006; Fountain-Jones, Packer, et al., 2017; Lee et al., 2012; Lembo et al., 2008; Streicker et al., 2016; Wheeler et al., 2010). For example, Lembo et al. (2008) used pathogen phylogenetics to study cross-species transmission and reservoir dynamics of rabies in the Serengeti ecosystem. Their study found more within-species and less between-species transmission than would be expected from random mixing; these findings could be due to the spatial distribution of hosts, or from increased or preferential intraspecific contacts (Lembo et al., 2008). [These hypotheses could be further tested via a network modeling approach, thereby linking the phylogenetic and contact network approaches in the context of non-random contact patterns within and between species.] Pathogen phylogenetics can therefore provide valuable information about transmission dynamics and host contact patterns, including insights into

understanding multi-host pathogen transmission, but these approaches could be enhanced even further by incorporating contact network techniques.

Importantly, investigation of pathogen phylogenetics can provide more detailed information that would be lost or otherwise inapparent by assessing host population genomics alone, particularly for directly transmitted pathogens. For example, phylogenetics of feline immunodeficiency virus in mountain lions revealed recent host demographic history that was not detectable with host population genomics (Biek et al., 2006). In addition, Bayesian phylogeographic approaches can yield important information about behavior and transmission dynamics by examining transmission over space and time (Biek et al., 2007; De Maio et al., 2015; Lemey et al., 2009, 2010), and between groups of hosts (Grad et al., 2014). Biek et al. (2007) used a phylogeographic approach to investigate spatial spread of rabies virus in raccoons in the Northeastern United States, finding, for example, that mountain ranges had a significant impact on rabies spread. These mountain ranges likely slowed spatial expansion due to the poor quality of raccoon habitat and reduced dispersal through these areas (Biek et al., 2007), demonstrating the impact of animal interactions with landscape features on the transmission of pathogens.

When available, the assessment of both pathogen and host genomics provides perhaps the most comprehensive information. For example, coupling host population genomic and pathogen phylogenetic methods have provided information about how solitary carnivores move and respond to landscape features like roads; where host phylogenies may suggest little to no movement across roads, pathogen phylogenies suggest these movements happen, but may be temporary or otherwise fail to result in host reproduction (Fountain-Jones, Craft, et al., 2017; Lee et al., 2012; Wheeler et al.,

2010). These findings show that movement may be restricted across major landscape features, but can still occur and be adequate for pathogen transmission. Host and pathogen evolution, especially in cases of rapidly evolving viruses, often operate on different time scales, so assessing both time scales together can provide information that would be lost by assessing either in isolation (Wheeler et al., 2010). Streicker et al. (2016) used the combined assessment of host population structure and pathogen phylogenies to determine that male vampire bats appear to disperse, while females show greater site fidelity, with important consequences for rabies transmission and expansion. This provides another important example of using host and pathogen phylogenetic analyses together to draw conclusions about host behavior, and then applying those findings to predict pathogen transmission dynamics.

While many phylogenetic studies focus on pathogen relatedness and genetic distance to make inferences, a developing tool for investigating disease dynamics is that of “phylodynamics.” The term was first defined by Grenfell et al. (2004) as the “melding of immunodynamics, epidemiology, and evolutionary biology,” and more specifically, phylodynamics seeks to understand the molecular footprint of epidemiological processes that are difficult to observe (Baele et al., 2017). Phylodynamic approaches have been used to investigate complicated questions such as short-term epidemic dynamics of human immunodeficiency virus in men who have sex with men (Lewis et al., 2008), or the impacts of urbanization on pathogen transmission and evolution in a wildlife system (Fountain-Jones, Craft, et al., 2017). Bayesian methods have been particularly useful within phylodynamics due to their ability to efficiently incorporate complex evolutionary models and uncertainty in parameter estimates (Drummond & Rambaut, 2007).

A developing branch within the broad field of phylodynamics is in the reconstruction of epidemics through the inference of transmission trees. These transmission trees are different from traditional phylogenetic trees, for several important reasons (Figure 1.2) (Didelot et al., 2017). First, in the context of epidemics of infectious disease, phylogenies consider each pathogen sample to be a tip (also known as a leaf) on the phylogeny, removing any possibility of a sampled pathogen being the ancestor of another sample (Picard et al., 2017). This means that phylogenies are unable to describe who infected whom. Transmission trees, however, allow sampled pathogens to be ancestors of other samples, and thereby may be able to reconstruct an epidemic by inferring who infected whom (Picard et al., 2017). The second major difference between transmission trees and phylogenetic tree is in the timing of coalescence events. Where timing of coalescent events in phylogenetic trees reflects branching events, timing in transmission trees reflects actual transmission events, which may occur at different time points from evolutionary branching events (Sintchenko & Holmes, 2015). It is important to understand these differences between phylogenetic and transmission trees, as the lines between the two are sometimes blurred and terminology used inappropriately, leading to confusion about the information provided by these two different methodologies.

Inferring transmission trees from epidemic data is a developing technique, with a variety of proposed methods (Cottam et al., 2008; De Maio et al., 2016; Didelot et al., 2014; Hall et al., 2015; Jombart et al., 2011; Klinkenberg et al., 2017; Numminen et al., 2014; Ypma et al., 2012; Ypma, van Ballegooijen, et al., 2013). While describing the methodological differences between these approaches is outside the scope of this review (but see Hall et al., 2016), the general process involves inferring transmission

trees from pathogen sequence data, while assuming one of the following: no within-host diversity or mutation, no within-host diversity but with mutation, or both within-host diversity and mutation (Hall et al., 2016). Pathogens with rapid mutation rates, such as RNA viruses, are generally more amenable to this type of epidemic reconstruction (Welch et al., 2011; Worby et al., 2014). Even with RNA viruses, however, it should not be assumed that an epidemic can be perfectly reconstructed (Hall et al., 2016), as all techniques will result in some uncertainty, particularly epidemics with incomplete sampling or slower pathogen evolution (Vasylyeva et al., 2016; Worby et al., 2014).

As transmission tree techniques have been further developed, however, they are better able to accommodate previous limitations; new techniques have now been used to study endemic (rather than epidemic) pathogens, ongoing epidemics, epidemics with incomplete sampling, and multiple disease introductions (Didelot et al., 2017; Mollentze et al., 2014). Judicious use of transmission tree methods can thus provide important information about transmission dynamics (Numminen et al., 2014; Ypma, Jonges, et al., 2013). For example, transmission trees have been used to assess the impact of wind direction on transmission of avian influenza in the Netherlands (Ypma, Jonges, et al., 2013). Transmission tree techniques can also shed light on the impact of host behavior and movements on pathogen transmission: Mollentze et al. (2014) used transmission tree reconstruction to show that rabies spread in South Africa appears to include the anthropogenic movement of dogs (i.e. in cars). The use of transmission trees is therefore a technique with great potential for understanding disease transmission dynamics in populations.

Epidemic reconstruction with only genetic sequence data, however, is limited in the resolution it can provide (Hall et al., 2016), and researchers should not expect fully

resolved transmission events for an entire epidemic. The incorporation of other epidemiological data may be able to clarify some of the uncertainty inherent in inference of transmission trees (Hall et al., 2016), for which contact network approaches, in identifying likely avenues of transmission, may provide a fruitful path forward.

In summary, genomic analyses of hosts, pathogens, and both together have demonstrated their utility in investigating the dynamic relationship between social behavior and pathogen transmission. The emerging field of phylodynamics shows great promise for expanding upon and refining our understanding of transmission dynamics in populations. In particular, the developing techniques for transmission tree reconstruction provide opportunities to investigate pathogen transmission at greater resolution. These techniques, however, could be more powerful if integrated with other methodologies, such as epidemiologic and network approaches, for drawing conclusions about patterns of behavior and pathogen transmission.

1.5 Integrating contact networks and genomics: current efforts and future directions

A growing number of studies have begun to integrate pathogen genetic data with contact networks. These studies have provided new insights on identifying risk factors for disease acquisition, studying the social structure of a population, and learning how social structure may predict infection. Previous work has often used molecular techniques that focus on pathogen strain-sharing or genetic distance between pathogen isolates across populations or individuals, and have largely utilized SNA in incorporating contact network techniques. We will discuss some examples in greater detail to highlight

the range of approaches used to integrate contact network and genomic tools for a variety of research questions.

Current research questions

Do node characteristics predict pathogen genetic similarity?

Perhaps the most straightforward integration of social networks and genomics tools is to use SNA to test if node and network-level characteristics predict pathogen strain sharing or genetic similarity of isolates. Villaseñor-Sierra et al. (2007), for example, use this technique to compare node-level characteristics between children with and without group A *Streptococcus*. "Strain sharing" of group A *Streptococcus* was inferred by comparing restriction fragment length polymorphism patterns across isolates, and this study found that spread of group A *Streptococcus* clones was mediated by high connectedness among children, as detected by SNA. This straightforward integration of contact networks and genetic tools highlights the utility of this technique for establishing the importance of social networks for disease spread. This approach may also be useful for initial assessment of pathogen transmission in populations; if social structure is established as an important mediator of pathogen transmission in a given study system, this could justify further, more in-depth analysis, and may subsequently aid in developing targeted control measures.

Which social networks best predict pathogen networks?

A "next step" in advancing the integration of SNA and genomic tools has been to more explicitly compare a social network to pathogen relatedness across a population. This can be done by comparing a population's social network to a "strain-sharing

network" based on microbial relatedness, and testing if the social network predicts the microbial network (Blyton et al., 2014; Bull et al., 2012; Fountain-Jones, Packer, et al., 2017; Marquetoux et al., 2016; Springer et al., 2016; VanderWaal et al., 2014a). Assessing social networks in the context of a pathogen-sharing network can identify social network positions that are important for transmission (VanderWaal et al., 2014a), test if the type of contact used to determine the social network (e.g. movement between farms) can explain pathogen transmission (Marquetoux et al., 2016), or even examine the varying importance of different modes of transmission (Fountain-Jones, Packer, et al., 2017). These strain-sharing networks, however, should not be confused with transmission trees; the former does not explicitly infer who infected whom, while the latter does. In addition, within strain-sharing networks, the molecular representation of strains can vary in resolution, ultimately affecting their power for inferring transmission. Advancements in genomic tools for inferring who infected whom may allow for increased resolution in these types of studies, and thereby potentially shed light on the variable impacts of social or behavioral factors on pathogen transmission.

What behaviors or locations are high risk for transmission?

Integrating contact networks and genomic tools can also be used to identify high risk behaviors or locations for pathogen transmission. For example, Gardy et al. (2011) utilized pathogen WGS and epidemiological data, in conjunction with social network analysis, to determine that a behavioral risk factor (likely crack-cocaine use) was a probable contributing factor in an outbreak of *Mycobacterium tuberculosis* in people in Canada. In addition, Chamie et al. (2015) assessed genotype sharing of *Mycobacterium tuberculosis* in people in Uganda in the context of SNA; they investigated spatial overlap

among genotype-clustered cases to identify likely transmission sites of social importance. Romano et al. (2010), using both social networks and hepatitis C virus gene sequences, assessed factors such as age, risk behaviors, and sexual networks to better understand risk factors even within different subtypes of the same pathogen. All of these studies demonstrate the utility of assessing pathogen genomics in the context of social networks for identifying important behavioral risk factors for transmission, including in instances where genotyping and contact tracing alone are not adequate for establishing pathogen dynamics (Gardy et al., 2011). All of the above example are studies in humans; in wildlife, such work would likely be more difficult as some methods for acquiring specific behavioral data in humans may not translate well to free-ranging animals (e.g. questionnaires are useful for acquiring behavioral data in humans but not applicable for wildlife).

Phylodynamics and future directions

The examples of previous work integrating contact networks and genomic tools have largely focused on using SNA, with limited utilization of higher-resolution molecular data from pathogens (but see Fountain-Jones, Packer, et al., 2017). The insights provided by phylodynamic approaches, however, suggest that integrating contact networks with phylodynamics may be a fruitful path forward. Some phylodynamics studies have begun to incorporate contact networks, including investigating what phylogenies, through phylodynamic analysis, can reveal about underlying host social structure (Colijn & Gardy, 2014; Leventhal et al., 2012; McCloskey et al., 2016; Robinson et al., 2013; Vasylyeva et al., 2016). Several of these studies have concluded that phylogenetic tree structure can provide some information about network structure

and resulting transmission dynamics (Colijn & Gardy, 2014; Leventhal et al., 2012; McCloskey et al., 2016; Vasylyeva et al., 2016). For example, Leventhal et al. (2012) used this approach to conclude that random mixing among hosts was unlikely to result in an observed epidemic's phylogenetic tree. Colijn and Gardy (2014) used this approach to conclude that phylogenetic tree structure can help differentiate between transmission patterns (e.g. super-spreader versus homogenous mixing). However, the conclusions that can be drawn from pathogen phylogenies can be limited, and may not yield particularly novel conclusions for researchers debating the value of sequencing pathogen samples. This may be especially true for hosts with dynamic interactions, as particular caution must be used when inferring a population's underlying social structure for dynamic contact networks (Robinson et al., 2013; Vasylyeva et al., 2016). While these uses of phylogenies to describe underlying host contact structure may be helpful for parameterizing a theoretical network on which to simulate outbreaks of disease, other applications of phylodynamics methods may be better suited to integration with contact networks.

Inference of transmission trees may hold promise to fill this gap for further integrating phylodynamic approaches with contact networks. Like the phylodynamic studies which have attempted to infer host population structure from pathogen phylogenies, some studies have investigated what information about the underlying social network can be gleaned from transmission trees alone (Carnegie, 2017). For example, assessment of HIV transmission trees has been used to infer preferential attachment in the underlying host social network, and this information was used to draw conclusions on the most appropriate management interventions (Leigh Brown et al., 2011). As with the phylogenetic tree approach, using transmission trees to make

inferences about underlying population structure has some limitations. For instance, transmission trees have not demonstrated an ability to detect clustering in the underlying contact network (Welch, 2011). However, the potential utility of pathogen transmission trees for improving understanding of the dynamic interaction between host behavior and pathogen transmission has yet to be fully explored and remains a major gap in the current literature. Future work to explore avenues for integrating genomic and network approaches may specifically benefit from examining how contact networks can inform priors in Bayesian transmission tree reconstruction to resolve some of the inherent uncertainty in this method. In addition, transmission trees, being a higher-resolution representation of the transmission network, should be particularly amenable to statistical approaches that can examine relationships between networks, such as generalized dissimilarity modeling (GDM; Ferrier et al., 2007). Alternatively, a phylogeographic generalized linear model approach could be used, as in Lemey et al. (2014). Ultimately, integrating genomic data—including transmission trees—with network data should provide useful information about the importance of different modes of transmission and provide better predictions about pathogen transmission (Eames et al., 2015), further justifying the effort involved in incorporating these additional layers of data.

While the full scope of the applications of transmission trees to contact network studies has yet to be resolved and will vary depending on research question, some general guidelines can be identified. First, arguably the best-suited study systems for these approaches would be those with fast-evolving pathogens like RNA viruses (Didelot et al., 2017). The rapid evolution of these pathogens may allow for improved inference of transmission events, and could shed light on transmission dynamics at the shorter temporal scale on which many host social networks exist. In addition, if using strain-

sharing or transmission networks to better understand population transmission dynamics, Blyton et al. (2014) point out that a study system should ideally have high strain alpha-diversity (local diversity) and high strain turnover in order to increase the resolution of transmission observations and to better represent contemporary transmission processes, respectively.

While ascertaining the proportion of the population that must be sampled to utilize a transmission tree approach has not been well established to date, pathogens which can be sampled across a greater proportion of the population, perhaps through non-invasive sampling techniques, would be more well-suited to transmission tree approaches. As methods for reconstruction of transmission trees advance, they are becoming more capable of accommodating unsampled individuals (Didelot et al., 2017; Mollentze et al., 2014), so outlining a specific proportion of a population to sample in order to use these methods is currently impractical. Ultimately, such determinations will be dependent on the current state of rapid methodological advances, and will be affected by specific research questions. For example, while low pathogen diversity and/or sampling effort will affect uncertainty to the degree that determining specific transmission events likely happened (rule-in) may be unrealistic, if the objective is instead to rule out certain transmission events, these systems may be perfectly appropriate for transmission tree approaches (Didelot et al., 2017).

In addition to contact networks, incorporation of host data such as relatedness, space use, or varying definitions of contact (e.g. grooming versus fighting) may also be useful for integrating with transmission tree approaches, depending upon the research question (for two examples, see Figure 1.3). Furthermore, linking established phylogeographic models (Lemey et al., 2009, 2010) to transmission tree inference will

enable more spatially explicit estimates of potential transmission events that can be compared directly to individual space use data. While specific examples of the integration of transmission trees to phylogeographic models and contact networks are, to the best of our knowledge, currently lacking in the primary literature, here we present some proposed pathways to integration, with the acknowledgement that, as phylodynamic approaches continue to advance, the opportunities to incorporate these tools into network studies will certainly evolve as well. In addition, we present advantages and disadvantages of current and proposed future methods for integrating network and genomic approaches in Table 1.1.

Which behaviors are most important for pathogen transmission?

The previously described examples of integrating contact networks and genomic techniques are limited in their directionality when describing pathogen transmission. In other words, they do not explicitly investigate who infected whom. In some cases, for example the movement of animals between farms, where movements or shipments are considered a “contact,” directionality is more obvious in the contact network (Marquetoux et al., 2016). In addition, detailed behavioral observations can suggest directionality of pathogen transmission between individual animals by investigating, for example, differences in infection incidence between grooming animals and the individuals being groomed (Drewe, 2010). But these are somewhat implied or correlational indicators of directionality, and would be strengthened by more refined inferences of transmission direction. In circumstances where direct observation is difficult and/or detailed epidemiological information is unavailable, phylodynamics-based methods, such as transmission tree inference, may be more capable of determining who infected whom.

By incorporating higher resolution transmission data gleaned from these types of approaches, contact network techniques can provide clearer information about the impact of behavioral risk factors on pathogen transmission. For example, comparing a transmission network to various behaviorally-derived networks could provide insight into the types of behaviors that are most important for pathogen transmission (Figure 1.3b). As discussed above, GDM may be an important tool for this type of analysis, as it can be useful for scrutinizing complex network structures. Additional questions to investigate by comparing socially-derived networks to high-resolution transmission networks could include (but are not limited to): the importance of turnover of social ties, the relative importance of host genetic versus social distance, and the impact of host social structure or movement.

Can we better understand transmission-relevant contact?

The examples given above have all focused on investigating social networks together with assessments of pathogen "networks." But this is, of course, only applicable in instances in which behavioral or observational data is adequate for describing the social network. Pluciński et al. (2011) discuss the idea that pathogen relatedness could be used to infer information about the underlying contact network. In support of this idea, genetic relatedness between human commensal oral bacteria has been shown to correlate with social network distance, as determined by questionnaire surveys (Francis et al., 2016). Importantly, these studies highlight the potential for commensal or apathogenic organisms to be used as proxies of contact between individuals or even groups of individuals. For example, can microbiome composition or infection with apathogenic organisms be used to establish risk factors or "rules" to describe a mode of

pathogen transmission in a population, and then can those rules predict subsequent transmission with other similar pathogens? Some commensal organisms will, of course, be unsuitable for this kind of transmission analysis. For example, some aspects of the host microbial community can be affected by diet, individual host factors, population of origin, or environmental transmission (Blyton et al., 2013; Chiyo et al., 2014; Degnan et al., 2012). In addition, organisms that transmit among and between multiple hosts would make discerning interactions within a single host species far more challenging; instead, single-host organisms with well-characterized modes of transmission may be most effective at elucidating transmission-relevant contact in a single host species. Rapidly evolving, apathogenic viruses may hold the most promise as proxies of contact, as their diversity and high mutation rates should allow for greater resolution in illuminating transmission events. In addition, identification of such viruses is becoming increasingly achievable due to advances in viromic technologies and non-invasive sampling (Minot et al., 2013; Rasmussen, 2015). The use of such apathogenic infections for detecting transmission-relevant contact may be particularly important for disease systems where the pathogen of interest cannot be easily sampled and sequenced due to, for example, very short shedding periods (e.g. about 1-2 weeks for canine distemper virus, Greene, 2012). Instead, commensal or apathogenic infection information could be utilized to determine a virtual contact network on which to simulate disease outbreaks, potentially without the use of resource intensive observational data typically required to estimate contact rates across populations (Figure 1.3c). This type of approach could potentially reveal and predict, for example, how animal movement behaviors and natural or anthropogenic landscape features affect disease transmission across populations.

Establishing the utility of this type of approach is a clear gap in the literature, and would benefit from future research.

Limitations

As previously noted, transmission tree approaches cannot be expected to reconstruct epidemics with perfect certainty, particularly for slow-evolving pathogens. Incorporating additional bounds and levels of information can, however, improve inferences from phylodynamics and integrative approaches (Ray et al., 2016; Welch et al., 2011). This multimodal approach, which may include data such as host genomics, landscape factors, and epidemic parameters, may be important in studies of pathogens with slow rates of mutation or otherwise limited genetic diversity (Wylie et al., 2005). The current literature on how to utilize a multimodal research framework is limited at this time (Ray et al., 2016), but carries potential for future expansion and may help resolve some areas of uncertainty within phylodynamic approaches. Future work to explore the impacts of incorporating different layers of data—including pathogen genomics—into contact networks should shed light on how these various factors are able to predict and describe the interaction between animal behavior and transmission dynamics.

Of course, within mathematical modeling, there is a perpetual trade-off between precision, generality, and realism (Levins, 1966). In this review, we have discussed the advantages gained by adding complexity to disease models through the use of contact networks and genomics tools, but it must be pointed out that these additions may come at the cost of computational efficiency and generality. While models that are very realistic or precise for a particular study system can provide useful information for managing a particular disease within a specific population, increased complexity within

models is not always the best approach (Buhnerkempe et al., 2015). Ultimately, a key challenge for disease modeling, broadly, is in understanding when simplicity or complexity is more appropriate (Buhnerkempe et al., 2015), and this dichotomy should not be forgotten amongst the advancements offered by more precise or realistic data and modeling.

1.6 Conclusions

This review has focused on describing how contact networks and genomic tools each shed light on the dynamic interaction between animal social behavior and infectious disease dynamics, and how these tools can be integrated for improved and new insights. These techniques may be particularly useful for wildlife researchers looking to make the most of hard-won pathogen data, and to advance their understanding of pathogen dynamics and social behavior. As genomics approaches become more accessible to more researchers, and the statistical and computational tools used to analyze genomic outputs continue to advance, the opportunities to utilize multimodal approaches in disease modeling will expand. Staying up-to-date with this progress will ultimately allow researchers to use novel techniques to answer complicated questions about animal behavior and its impact on pathogen transmission, and consequently make predictions important to the surveillance, prevention, and management of infectious diseases in populations.

1.7 Acknowledgements

M.L.J.G. was supported by the Office of the Director, National Institutes of Health under award number NIH T32OD010993. The content is solely the responsibility of the authors

and does not necessarily represent the official views of the National Institutes of Health. N.M. F-J was funded by National Science Foundation (DEB-1413925) and M.E.C. was funded by National Science Foundation (DEB-1654609), the University of Minnesota's Office of the Vice President for Research and Academic Health Center Seed Grant, and the Cooperative State Research Service, U.S. Department of Agriculture, under Project No. MIN-62-098. Special thanks to L. White, K. Worsley-Tonks, Y. Wang, and two anonymous reviewers for their invaluable input and suggestions.

1.8 Figures

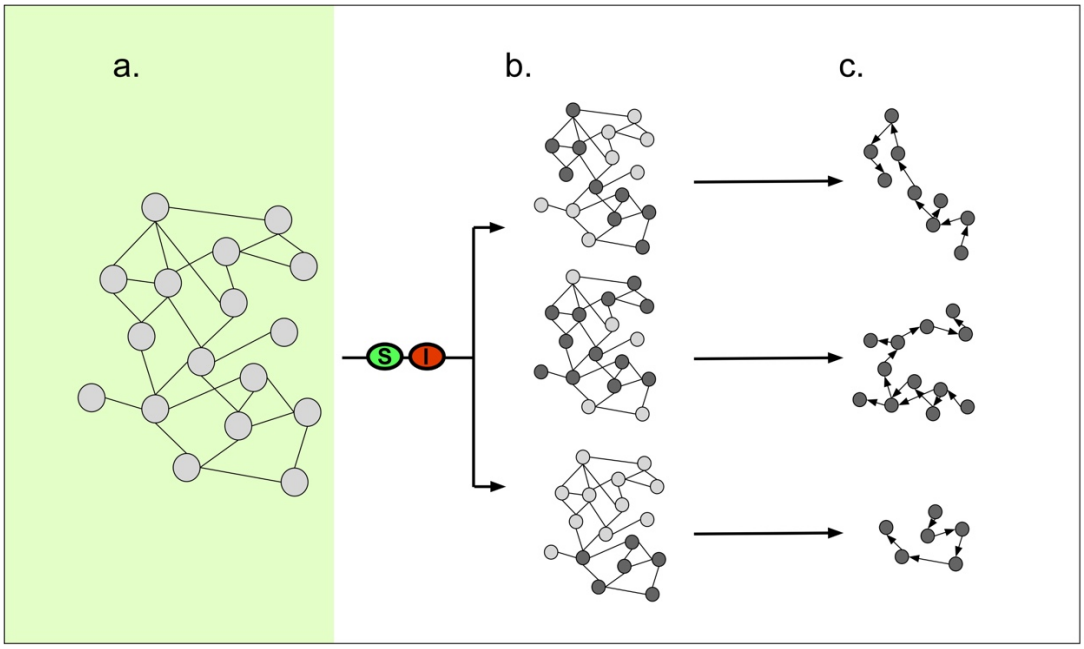


Figure 1.1. Conceptual flow from a contact network to a transmission network. Areas with a green background represent components informed by empirical data; areas with a white background represent components informed by simulations. Network (a.) shows a contact network that would be defined by, for example, observational data of direct contacts. A disease process, such as an SI (susceptible-infectious) model, (b.) could be applied to the contact network many times (pictured three times here, but realistically, a simulation would be run on the order of 1000 times). In this example, the index case was randomly seeded, and darkened nodes represent those nodes that were infected in the course of a simulation. The who-infected-whom, transmission networks (c.) could be a final output of network model simulations.

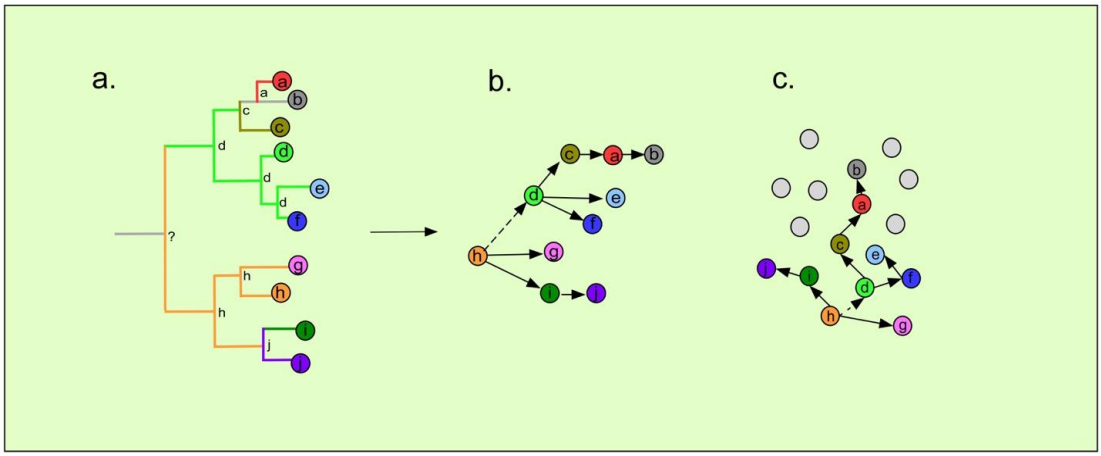


Figure 1.2. Conceptual flow from (a.) a transmission tree to (b.) a transmission network; (c.) depicts a transmission network in the context of the rest of the population, where gray nodes represent uninfected individuals. The green background highlights that this approach is based on empirical data, rather than simulations. Colored circles represent sequenced samples from infected individuals. In (a.), branching events in the transmission tree indicate transmission events, with color changes occurring at these events; lettered labels at internal nodes represent infecting individuals. The colored lines and labels at branching events in the transmission tree highlight a primary difference between phylogenetic trees and transmission trees: pathogens sampled at the “tips” of the transmission tree are allowed to be ancestors of other samples. This then allows for the inference of who infected whom, as demonstrated by the directed networks in (b.) and (c.), but with some uncertainty that cannot be fully resolved (represented here by uncertain transmission from individual “h” to individual “d”).

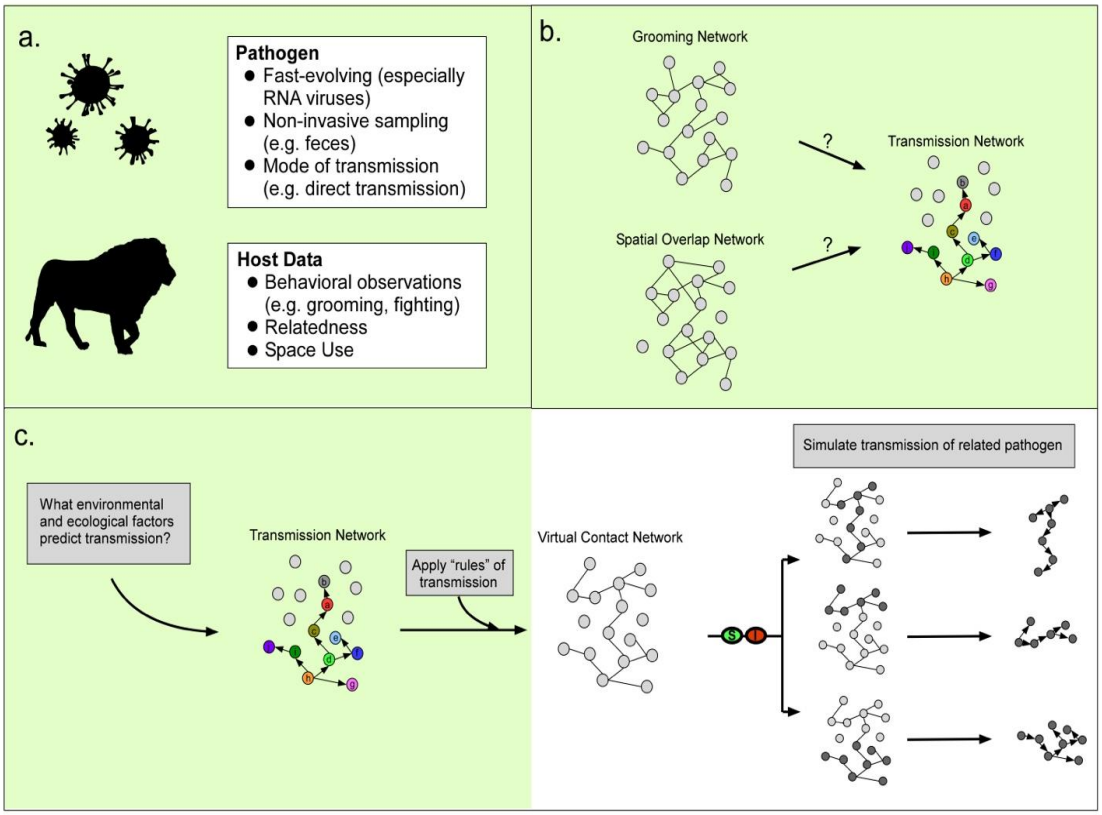


Figure 1.3. Two examples of proposed integrations of contact networks and genomic tools, in this case focusing on utilizing transmission networks derived from inference of transmission trees. Areas with a green background represent components informed by empirical data; areas with a white background represent components informed by simulations. Panel (a.) highlights some characteristics of pathogens and types of host data that may be well-suited to transmission tree approaches, but these are by no means all-inclusive. Panel (b.) demonstrates a SNA application, in which observed networks such as grooming or spatial overlap networks, are compared to high-resolution transmission networks to determine what social or behavioral factors have the greatest impact on pathogen transmission. Panel (c.) depicts a network modeling application, in which a transmission network derived from an apathogenic or commensal organism is used to determine environmental and ecological factors that best predict transmission in

a population. These “rules” are then used to create a “virtual contact network” on which epidemics with a related (pathogenic) organism could be modeled. This approach could help determine the best preventive or intervention measures to be applied prior to an outbreak of a pathogen of concern.

1.9 Tables

Table 1.1. Advantages and disadvantages of currently used and proposed methods for integrating network and genomic approaches. Methods are listed in the order in which they appear in the text. A * indicates that the given method utilizes Bayesian approaches, which have the advantage of being able to incorporate uncertainty.

Method	Advantages	Disadvantages	References
SNA and pathogen strain-sharing network	Requires less intense pathogen sequencing effort. Able to identify if social structure is important for pathogen transmission, and/or the relative importance of different behaviors or locations. In future, may incorporate whole genome sequencing of pathogens, especially in the context of transmission tree reconstruction for higher resolution transmission networks.	Uses low resolution representation of pathogen relatedness with limited ability to represent direction of transmission. May have reduced ability to refine conclusions about impact of specific behaviors, locations, etc. on transmission.	(Fountain-Jones, Packer, et al., 2017; Marquetoux et al., 2016; VanderWaal et al., 2014a; Villaseñor-Sierra et al., 2007)
Phylogenetic tree structure to describe network structure	Able to identify non-random mixing and differentiate some transmission patterns in underlying social network. In future, may be able to be used to parameterize theoretical networks for simulations and hypothesis testing.	Limited in conclusions that can be drawn. Identifying non-random mixing may not be novel in all systems. Poorly suited to dynamic networks.	(Colijn & Gardy, 2014; Leventhal et al., 2012; McCloskey et al., 2016; Robinson et al., 2013; Vasylyeva et al., 2016)

Transmission tree structure* to describe network structure	Able to identify non-random mixing in underlying social network, and potentially draw conclusions about management interventions. In future, may be able to be used to parameterize theoretical networks for simulations and hypothesis testing.	Limited in conclusions that can be drawn. Unable to reliably detect clustering in underlying network. High pathogen sequencing effort required. Unlikely to be the most effective integration of genomic approaches and contact networks.	(Carnegie, 2017; Leigh Brown et al., 2011; Welch, 2011)
Transmission tree reconstruction* with contact networks and/or phylogeographic models informing priors	Contact networks and/or phylogeographic models may help resolve some of the uncertainty involved in transmission tree reconstruction. Transmission trees then provide higher resolution data about transmission events.	Higher effort for sequencing, computational effort for reconstructing transmission trees. Generally requires pathogen with high mutation rate and high-intensity sampling effort.	Theoretical, but see (Hall et al., 2016)
SNA and pathogen transmission trees*	Transmission trees may better represent directionality of transmission than strain-sharing networks. May be able to identify social structure, behaviors, locations, etc. important for transmission. Higher resolution representation of pathogen relatedness than strain-sharing networks.	Higher effort for sequencing, computational effort for reconstructing transmission trees. Generally requires pathogen with high mutation rate and high-intensity sampling effort.	Theoretical

Transmission tree* of apathogenic infections to describe social network	Apathogenic infections may reveal transmission relevant contact prior to an epidemic with a pathogenic infection. May be useful for describing relevant intervention measures in at-risk wildlife populations. May be able to capitalize on non-invasive sampling.	Higher effort for sequencing, computational effort for reconstructing transmission trees. Generally requires agent with high mutation rate and high-intensity sampling effort. Requires agent with well-characterized mode of transmission. Most appropriate for directly transmitted, single-host pathogens. This approach is currently untested; transmission of pathogenic and apathogenic organisms may be too different to be able to translate mechanisms/risk factors between agents.	Theoretical, but see (Blasse et al., 2013; Blyton et al., 2014; Bull et al., 2012; Francis et al., 2016; Pluciński et al., 2011; Springer et al., 2016)

Chapter 2. Trade-offs with telemetry-derived contact networks for infectious disease studies in wildlife

Gilbertson, M. L., White, L. A., & Craft, M. E. (2020). Trade-offs with telemetry-derived contact networks for infectious disease studies in wildlife. *Methods in Ecology and Evolution*. doi: 10.1111/2041-210X.13355

2.1 Synopsis

Network analysis of infectious disease in wildlife can reveal traits or individuals critical to pathogen transmission and help inform disease management strategies. However, estimates of contact between animals are notoriously difficult to acquire. Researchers commonly use telemetry technologies to identify animal associations; but such data may have different sampling intervals and often captures a small subset of the population. The objectives of this study were to outline best practices for telemetry sampling in network studies of infectious disease by determining (1) the consequences of telemetry sampling on our ability to estimate network structure, (2) whether contact networks can be approximated using purely spatial contact definitions, and (3) how wildlife spatial configurations may influence telemetry sampling requirements.

We simulated individual movement trajectories for wildlife populations using a home range-like movement model, creating full location datasets and corresponding "complete" networks. To mimic telemetry data, we created "sample" networks by subsampling the population (10-100% of individuals) with a range of sampling intervals (every minute to every three days). We varied the definition of contact for sample networks, using either spatiotemporal or spatial overlap, and varied the spatial

configuration of populations (random, lattice, or clustered). To compare complete and sample networks, we calculated seven network metrics important for disease transmission and assessed mean ranked correlation coefficients and percent error between complete and sample network metrics.

Telemetry sampling severely reduced our ability to calculate global node-level network metrics, but had less impact on local and network-level metrics. Even so, in populations with infrequent associations, high intensity telemetry sampling may still be necessary. Defining contact in terms of spatial overlap generally resulted in overly connected networks, but in some instances, could compensate for otherwise coarse telemetry data.

By synthesizing movement and disease ecology with computational approaches, we characterized trade-offs important for using wildlife telemetry data beyond ecological studies of individual movement, and found that careful use of telemetry data has the potential to inform network models. Thus, with informed application of telemetry data, we can make significant advances in leveraging its use for a better understanding and management of wildlife infectious disease.

2.2 Introduction

Outbreaks of infectious disease in wildlife can have significant impacts on population health and may also have detrimental effects on domestic animals and humans as a result of cross-species transmission (Jones et al., 2008; Plowright et al., 2017). Disease modeling has proven an illuminating tool in understanding transmission dynamics within and between wildlife populations (Lloyd-Smith et al., 2009). However, disease models rely upon estimates of transmission, which require estimates of contact

between individuals, particularly for directly or sexually transmitted pathogens (Anderson & May, 1991; Begon et al., 2002). Further, contact heterogeneity within a population has significant impacts on epidemic outcomes (Keeling & Eames, 2005; Lloyd-Smith et al., 2005), which may alter the approach needed to appropriately model a pathogen system. Network analysis (including social network analysis and network modeling) incorporates contact heterogeneity which can help address this challenge, but network tools are data hungry and further dependent on quality estimates of contact rates (Keeling & Eames, 2005).

The cost and effort required to monitor free-ranging populations, the size or secretive nature of a species, or inhospitable habitats can all make contacts in wildlife challenging to observe (Krause et al., 2013). Estimates of contact in wildlife populations, therefore, increasingly utilize remote or automated detection of animal associations via animal tracking devices (Cross et al., 2012; Krause et al., 2013). Among such approaches, arguably the most direct approximation of contact is to identify co-location in both space and time (*spatiotemporal overlap*) via proximity loggers (e.g., Hamede et al., 2009). However, proximity loggers are not a perfect solution; a short battery life can limit the duration of observation, and the larger size of these loggers is inappropriate for some species (Krause et al., 2013). Alternatively, GPS or VHF telemetry technologies have also been used to detect spatiotemporal overlap as a proxy for contact (e.g., Godfrey et al., 2014; Perkins et al., 2009; Schauber et al., 2015), but the localization error associated with these tools may affect their contact detection accuracy (Hulbert & French, 2001).

Another proxy mechanism for estimating contacts in wildlife is through the use of *spatial overlap* contact definitions, which assumes that the extent of spatial overlap

between any two individuals is representative of the probability of contact. A number of studies have used this approach to estimate networks and contact rates from observed or telemetry-recorded locations (e.g., Godfrey et al., 2010; Lewis et al., 2017; Schauber et al., 2007). Further, home range and subsequent spatial overlap estimation can be accomplished with lower frequency of sampling than would be required for spatiotemporal definitions of contact. However, while some studies have supported the underlying assumption that increased dyadic spatial overlap is associated with increased dyadic contact rate (Robert et al., 2012; Vander Wal et al., 2014), this assumption has not been thoroughly tested across systems (Schauber et al., 2015) and may be more likely to reflect the likelihood of shared space use rather than actual interaction frequency (Wanelik & Farine, 2019). Further, studies examining animal social networks often correct for spatial overlap (Whitehead & James, 2015), indicating that spatial overlap is, as yet, an unproven proxy for animal associations.

Perhaps the most pervasive challenge for any of these contact estimation approaches or any telemetry study more broadly is the issue of appropriate sampling effort — both the proportion of the population sampled and the frequency with which locations are recorded for each monitored individual. While some work has attempted to establish the proportion of a population that must be sampled (i.e. node sampling) for appropriate estimates of contact networks (e.g., Silk et al., 2015; Smith & Moody, 2013; Wey et al., 2008), less effort has been focused on the frequency with which individuals should be monitored (i.e. edge sampling, but see Davis et al., 2018). Further, we lack realistic sampling recommendations for deriving contact networks for wildlife (Cross et al., 2012, but see Costenbader & Valente, 2003), and, to the best of the authors'

knowledge, no research has yet explored the impact of the frequency of telemetry locations on subsequent network estimation.

Despite these challenges, telemetry estimation of wildlife contacts holds great promise for informing network analysis and improving understanding of pathogen transmission in wildlife. Researchers therefore need best practices recommendations for using existing and future telemetry datasets for wildlife network estimation, particularly in light of the lack of information about necessary telemetry sampling effort.

Because sampling or directly observing an entire wildlife population and all of its resulting contacts is currently unfeasible, we employed a simulation approach to investigate the impact of telemetry sampling on contact detection and network estimation. The first objective of this study was to simulate the movement trajectories of a theoretical wildlife population to produce *complete networks* of known contacts and then to simulate different sampling effort regimes by “collaring” or “tagging” different proportions of the population and also varying the frequency of telemetry sampling. We hypothesized that the impact of sampling would vary based on the metrics used to describe network structure, with metrics determined by local network topology being the most resilient to reduced sampling effort (Cross et al., 2012; Davis et al., 2018).

In addition, because different approaches to detecting and defining contact, even in the same population, can result in different contact network structures (Perkins et al., 2009), our second objective was to investigate how our methodological definition of a contact (spatiotemporal overlap versus spatial overlap) affected network structure and resulting network metrics. In particular, we investigated if more sensitive, but less specific definitions of contact could compensate for reduced sampling effort.

Lastly, social systems of wildlife are highly variable, spanning from solitary to highly gregarious, and from non-territorial to territorial (Sah et al., 2018). Our third objective was therefore to determine if the underlying social system of the population has consequences for the effect of sampling on network estimation. We tested variations in the *spatial configuration* of simulated populations to represent a range of realistic social systems, including highly territorial populations with infrequent associations between individuals (e.g. solitary carnivores) and highly aggregated populations with frequent associations between individuals (e.g. herd species or animals aggregating around limited resources). We hypothesized that territorial populations with infrequent associations would be more sensitive to sampling, as their rare associations with limited partners are more likely to be missed with telemetry sampling.

2.3 Materials and Methods

Simulations

In order to examine the effects of telemetry sampling on subsequent network estimation, we simulated movements of individuals in wildlife populations and sampled from these movement trajectories to mimic data collection from telemetry devices (Figure 2.1 and Appendix A, Figure A.2). Hereafter, a single *simulation* refers to movement trajectories for a population of 100 animals. Movement trajectories were created using a simple biased correlated random walk (BCRW; Long et al., 2014), which produces home range-like movement when the bias is directed toward the starting point of the trajectory (Van Moorter et al., 2009). A home range movement model was necessary in order to estimate home ranges and thereby test the robustness of spatial overlap as an indicator of animal associations (see “Sample Networks” below). The

movement trajectories were generated on a per-minute basis, for a duration of 90 days. We considered 90 days an appropriate time frame for contact network estimation because animal social dynamics and pathogen transmission are known to vary seasonally (Reynolds et al., 2015).

Movement trajectories were generated using six parameter sets. Three of these varied home range sizes by altering the step length distribution scaling parameter, while leaving all other parameters constant; this produced small, medium, and large home range parameter sets (Appendix A, Table A.1). Because the BCRW is a simple, tractable movement model, step length distributions based on empirical movement data do not necessarily produce appropriate home range sizes for a chosen species. Rather, the BCRW allowed us to test the sensitivity of our findings to changes in movement model parameters without the decreased transparency associated with more complex movement models. With these benefits in mind, we tested three additional parameter sets which kept home range size roughly constant with the small home range model while varying other BCRW parameters (Appendix A, Table A.1).

To mimic different spatial configurations, with a given parameter set, we varied the starting locations (and thus the home range centers) of individuals, testing random, lattice, and clustered spatial configurations (Appendix A, Figure A.1). Random configurations were a “null” spatial layout. Lattice configurations maintained approximately the same degree of home range overlap, independent of home range size, and represented highly territorial, solitary species such as many large carnivores (e.g. jaguars, Ethiopian wolves; de Azevedo & Murray, 2007; Zubiri & Gottelli, 1995). Clustered configurations placed individuals in one, five, or ten equally sized clusters, thereby representing highly social species (e.g. social ungulates such as white-tailed

deer or American bison; Schauber et al., 2015) or animals aggregated around heterogeneous resources (e.g. urban racoons; Hirsch et al., 2013). Further details on simulation methods can be found in Apendix A. We performed all simulations in R, version 3.5.0 (R Core Team, 2018).

Complete Networks

We constructed complete contact networks by detecting contact events between all 100 individual movement trajectories in a given simulated population. For complete networks, contact was always determined by spatiotemporal overlap, with contacts defined as simultaneous locations within a given distance threshold. We varied the distance threshold used to construct complete networks; because movement simulations were scale-free, we refer to the thresholds as small, medium, or large thresholds (hereafter, *small thresholds*, etc.), with these resulting in three complete networks for any given simulation. The distance thresholds should be considered relative to the step length distributions used in the movement model. For example, even with the smallest step length distribution, an individual's cumulative steps would surpass the largest contact threshold within about 8-10 steps (i.e. minutes), meaning that even our large contact threshold is fairly strict (see Appendix A for further details).

Sample Networks

To construct *sample networks*, we first mimicked sampling a subset of the population by randomly selecting a portion of the simulated individuals to "collar." We varied the sampling effort from 10-100% of the population at 10% intervals (hereafter, *proportion of the population sampled*). In reality, the proportion of a wildlife population

sampled is often unknown, but is likely to be on the lower end of this range (i.e. <50%; Cross et al., 2012). With those monitored individuals, we then varied the frequency of location sampling by recording individual locations at an interval of every 1 minute, 15 minutes, 60 minutes, 3 hours, 12 hours, 24 hours, or 72 hours (hereafter, *frequency of sampling*). This range spans functionally continuous sampling (every 1 minute), which is rare in wildlife studies, to much more common frequencies of GPS sampling (every 12-72 hours). The composite of the proportion of the population sampled and frequency of sampling is hereafter referred to collectively as *sampling effort*.

We then estimated spatiotemporal overlap using simultaneous locations within large, medium, or small distance thresholds for each sample dataset. For sample networks, we also detected contacts as determined by spatial overlap, which we calculated using the utilization distribution overlap index (UDOI) of the 95% bivariate normal kernel density estimate (KDE) home range (Fieberg et al., 2005). Spatial overlap contacts were defined as UDOI greater than zero (using both binary and weighted edges). Because KDE approaches assume independent locations (Worton, 1989), we only calculated spatial overlap contacts for telemetry frequencies of every 24 or 72 hours. Spatial overlap was calculated using the *adehabitatHR* package in R (Calenge, 2006).

Network Comparisons

Complete and sample networks were compared based on contact definitions (Appendix A, Figure A.2). The complete networks with a large contact threshold were therefore compared to sample networks with the same large threshold, and so on. In addition, we compared complete networks with a medium threshold to corresponding

sample networks with a large threshold to assess the effect of using a more sensitive but less specific spatiotemporal contact definition in the sample network. Lastly, we compared the three types of complete networks (large, medium, or small threshold) to sample networks derived from the spatial overlap contact definition. Hereafter, sample networks with the same contact definition as their corresponding complete network are referred to as *strict* contact definitions, while sample networks with more sensitive/less specific spatiotemporal or spatial overlap contact definitions are referred to as *less precise* contact definitions.

For all complete and sample networks, we calculated seven structural network metrics which can be important for pathogen transmission. We followed Silk *et al.* (2015) and Davis *et al.* (2018) in selecting the node-level metrics of degree, strength, betweenness, and transitivity; we also calculated network-level metrics of density, proportion isolates, and modularity (see Appendix A for modularity details). As in Davis *et al*. (2018), we characterized our node-level metrics as *local* or *global*. Degree and strength are determined by local connectivity, and individuals with high degree or strength have more or stronger connections. Such well-connected individuals may therefore be candidates as "superspreaders" in studies of pathogen transmission (Lloyd-Smith et al., 2005). Betweenness and transitivity are determined by global network topology; individuals with high betweenness may be important "firebreak" individuals for interrupting pathogen transmission (VanderWaal et al., 2014b), and transitivity is a measure of clustering in a network (Farine & Whitehead, 2015) which can have significant impacts on epidemic outcomes (Keeling & Eames, 2005). Network analysis was performed with the R package *igraph* (Csardi & Nepusz, 2006).

To compare node-level metrics between complete and sample networks, we followed Davis *et al*. (2018) in calculating the ranked correlation coefficient between metrics from individuals in the sample network to corresponding individuals from the complete network. We calculated mean and 95% confidence intervals for ranked correlation coefficients as averaged across variations in movement model parameterization and sampling effort. Sample networks often became very poorly connected with low sampling effort, meaning that fewer networks had measurable node-level metrics, with subsequently greater variation in mean ranked correlation coefficients. We therefore conservatively used the lower limit of the 95% confidence interval, rather than the mean, to interpret results. Target correlations were between 0.80 and 1.00, following Smith and Moody (2013). When networks were completely unconnected at a given level of sampling effort, we classified them as “disconnected.”

To compare agreement between sample and complete networks for the network-level metrics (density, proportion isolates, modularity), we calculated the percent error of mean metric values between complete and sample networks across given simulation variations. Here percent error is the difference between the network metric of the sample and complete networks, divided by the metric of the complete network, and multiplied by 100%. For proportion isolates, we assessed the percent error of the complement of proportion isolates (i.e., proportion connected or effective number of nodes; Sah et al., 2018) as connected individuals are of primary interest in pathogen transmission studies. Percent error of mean density and proportion isolates describes agreement in overall network connectivity; a positive percent error equates to more connections in the sample network than the complete network. For modularity, a positive percent error indicates higher modularity in the sample networks, suggesting stronger community structure.

2.4 Results

We completed a total of 11,994 simulations, each with three definitions of contact (large, medium, or small threshold) used to construct complete networks, for a total of 35,982 complete networks. Simulation variations did not yield substantial changes in the patterns reported here, but more extensive results are reported in Appendix A (Figures A.1-A.12). Here we highlight key results for medium home range simulations which are illustrative of our overall findings.

Local metrics outperform global metrics

When using spatiotemporal definitions of contact, all metrics performed best (highest correlation coefficients and smallest percent error) for comparisons of complete and sample networks with large threshold contact definitions, and all metrics performed poorly with small thresholds. This finding was consistent across all simulation variations, and was likely driven by fewer associations with the small thresholds. In general, the local metrics (strength, degree) consistently outperformed the global metrics (betweenness, transitivity; Figure 2.2 and Appendix A, Figures A.1-A.8). Strength showed the highest resilience to sampling of all metrics with the best performance in clustered populations. However, even among the local metrics, high sampling effort was still required to achieve high correlation scores in some instances (e.g. for lattice spatial configurations, as in Figure 2.2).

Density, proportion isolates, and modularity performed intermediately as compared to the other metrics. Density, especially, required higher frequency of sampling to achieve a smaller percent error. As with the node-level metrics, both density

and proportion isolates performed best (smallest percent errors) in clustered populations, with random and lattice spatial configurations more sensitive to sampling (Appendix A, Figures A.9-A.12). Modularity, on the other hand, showed strong positive percent error in clustered configurations, and greater variability in results overall (Appendix A, Figures A.13-A.14).

Frequency of sampling has major consequences for network estimation

As expected, reducing the proportion of the population sampled resulted in reduced metric performance for most metrics, but with the notable exception of density (Appendix A, Figures A.9-A.10). Density was largely unaffected by the proportion of the population sampled, regardless of simulation variations, though this was likely a result of the random sampling procedure used. Surprisingly, frequency of telemetry sampling had at least as much impact on metric performance as the proportion of the population sampled (Figure 2.2). This was especially notable for global metrics, density, lattice spatial configurations, and more restrictive contact definitions (medium or small thresholds), where seemingly minor reductions in the frequency of sampling resulted in rapid decreases in correlation scores and more negative percent error.

Less precise contact definitions produce overly connected networks

We also examined metric performance when less precise contact definitions were used in the sample networks in order to determine if such approaches could compensate for reduced sampling effort. We found mixed results when using less precise contact definitions. As with the spatiotemporal contact definitions, global metrics performed poorly across all simulation variations. The local metrics, however,

occasionally recovered surprising levels of correlation performance with our less precise contact definitions (Figure 2.3 and Appendix A, Figures A.5-A.8), especially when compared to the equivalent sampling effort with a standard spatiotemporal contact definition (Figure 2.3). The correlation score improvements were the most pronounced for spatial overlap definitions of contact when the complete networks were defined by a large threshold. Even with the local metrics, improvements in correlation scores varied across types of spatial configurations: clustered configurations showed the most improvement with less precise contact definitions, but lattice layouts displayed no major increase in correlation scores. The network-level metrics, density and proportion isolates (evaluated as proportion connected), showed increased positive percent error, demonstrating that these less precise definitions of contact produced overly connected networks (Appendix A, Figures A.9-A.12).

Spatial configuration has significant impacts on network estimation

The performance for all social network metrics varied with spatial configuration (degree shown for illustrative purposes in Figure 2.4; see also Appendix A, Figures A.1-A.12). Lattice layouts, in particular, consistently showed poor metric performance, while dense, frequently interacting populations (i.e. one cluster) had the best metric performance. In general, the lattice and random layout populations had fewer associations than clustered populations (Appendix A, Table A.2; from 1-20% of the mean number of contacts for single cluster configurations), supporting our hypothesis that infrequently interacting populations would be more sensitive to telemetry sampling. However, even in instances in which lattice populations had greater than or equivalent mean numbers of contacts per dyad (Appendix A, Table A.2), correlation scores in lattice

populations remained low compared to other spatial configurations, suggesting that sampling sensitivity is not entirely mediated by the strength of an association. The global metrics showed less variation in performance with spatial configuration, although these metrics generally performed poorly across all variations.

Simulated home range size varied from 5×10^6 (small) to 4×10^7 (large) square units (equivalent to 5 to 40 square “kilo-units”; a mean area increase of 800%; Appendix A, Table A.1). Home range size also impacted metric performance, with a general trend toward reduced performance as home range size increased. For density, this trend resulted in large home range simulations having larger positive percent error when using a spatial overlap definition of contact in sample networks. This is notable for lattice simulations, which were designed to keep roughly equivalent amounts of spatial overlap between individuals across simulations (see Appendix A for details on simulations), meaning it is unlikely that large home range simulations had proportionally more overlap than small home range simulations. Rather, the higher positive percent error in the lattice scenarios suggests that large home range simulations generally had fewer associations between individuals, making these simulations more sensitive to sampling across metrics and simulation variations (Appendix A, Table A.2).

2.5 Discussion

Wildlife telemetry data can be a powerful tool for remotely detecting animal associations for use in network estimation, social network analysis, and pathogen transmission modeling (Robitaille et al., 2019). However, regardless of study objective, variations in telemetry sampling effort can have a significant impact on the structure of subsequent contact networks. Our results suggest that the impact of telemetry sampling

effort varies with contact definition, spatial configuration, and the chosen network metrics, and we provide best practices recommendations for telemetry sampling design for network studies of wildlife infectious disease below.

Prioritize local network metrics and consider maximizing sampling frequency

Across all simulation variations, we found that local network metrics consistently outperformed global metrics. However, the local metrics still required intensive sampling under some conditions, especially lattice layout populations. Further, the frequency of telemetry sampling was often at least as important as the proportion of the population sampled in order to achieve high metric correlation values. In the context of networks, the proportion of the population monitored can be thought of as "node sampling," while the frequency of telemetry sampling relates to the "edges" of the network. Given the profound impact of undersampling the edges of the network which we observed here, we suggest that researchers may choose to maximize frequency of telemetry locations over the proportion of the population sampled, especially for territorial species or those with infrequent interactions (e.g. solitary species such as puma; Elbroch & Quigley, 2016). This recommendation expands upon work by Franks *et al.* (2010), who argued for emphasizing edge sampling over node sampling when observing gambit of the group associations in resource-limited circumstances. In infectious disease studies, estimates of local metrics may be important for identifying highly connected individuals (i.e. "superspreaders"), but also for determining if phenotypic traits are associated with connectedness. For example, traits like sex, age, or social rank may be associated with high node degree, allowing for more efficient vaccination protocols, interventions, or management strategies (Rushmore et al., 2014). The population sample sizes

necessary to achieve the statistical power to identify such associations may therefore need to be balanced against the frequency of locations necessary to accurately estimate local network metrics, and our simulation results suggest this conflict will be most profound for territorial or infrequently interacting species.

In more extreme situations, researchers may also consider using a smaller number of intensively monitored individuals to develop movement and association models in order to simulate contacts in populations. Such an approach may be particularly useful if global network metrics are critical to the research objective (e.g. using high betweenness to identify important “firebreak” individuals for vaccination or quarantine), as these metrics tended only to perform well with exceptionally high sampling effort which is generally not feasible for wildlife populations. Alternatively, many network metrics tend to be highly correlated with each other, including betweenness and degree (Farine & Whitehead, 2015; VanderWaal et al., 2014b), which could further depend on spatial configuration. Researchers may therefore consider if a local network metric is able to accomplish their research objectives, rather than attempting to manage the sampling challenges needed to achieve reliable global metrics.

Spatial overlap is a viable contact definition in clustered spatial configurations

We also examined the effect of using less precise contact definitions on network estimation. First, we examined the effects of using a large spatiotemporal threshold in sample networks, as compared to a medium threshold in the complete network. While simulating location error was beyond the scope of the present study, a less precise spatiotemporal contact definition approximates some of the precision lost with telemetry location error. The less precise spatiotemporal contact definitions showed modest

improvement in local metric correlation scores suggesting location error may have limited effects on network estimation. However, further research more specifically into the effects of location errors are necessary to refine this conclusion.

Importantly, in some instances, less precise spatial overlap contact definitions were able to compensate for coarse sampling by still identifying the relative connectedness of individual nodes in our simulations. Thus, if the research objective is to identify the most connected individuals in a population (Craft, 2015; Farine & Whitehead, 2015; White et al., 2017), less precise contact definitions like spatial overlap may be a viable strategy. However, because these contact definitions simultaneously produced overly dense networks, caution should be applied if using these contact definitions for metrics at the network level or in the context of pathogen transmission simulations. It may be possible to correct for overly dense networks when using a less precise contact definition by "thresholding" at the network level — for example, only counting as contacts those pairs that reach a specified level of spatial or spatiotemporal overlap — however excluding edges below a given weight is generally not recommended in network studies (Farine & Whitehead, 2015), and determining the exact threshold to use may be ambiguous or lack biological motivation.

A major caveat of using less precise contact definitions to compensate for coarse sampling, however, is that the performance of these contact definitions varied across spatial configurations. For example, lattice layout populations surprisingly showed no major improvement in correlation scores with less precise contact definitions. Thus, we recommend using empirical data or system-specific movement simulations to justify using less precise contact definitions as proxies of direct contact (as in Brandell et al., 2020), particularly in highly territorial populations.

Our results are largely consistent with prior work which has found that spatial overlap may be used as a proxy for direct contact (Robert et al., 2012). However, such findings have not been consistent across all studies (e.g., Schauber et al., 2007) and spatial overlap is often corrected for in social network studies (Whitehead & James, 2015), suggesting that the utility of spatial overlap as a proxy of direct contact may be system specific. Further, in our simulations, spatial overlap was more representative of less restrictive contact definitions (large threshold), and may therefore be more appropriate when studying host-pathogen systems in which the pathogen is transmitted over larger distances or longer time periods (e.g. long-distance aerosol, vector, or persistent environmental transmission; Burgin et al., 2013; Tissot-Dupont et al., 2004). Because less precise contact definitions may be able to make the most of otherwise low frequency sampling, this approach may allow researchers to sample edges with reduced effort. While any remote contact detection approaches may struggle to determine the exact nature of an association (e.g. aggressive vs. passive; Krause et al., 2013), the less precise contact definitions are even more limited. Spatial overlap, for example, cannot directly characterize the duration or frequency of associations, both of which may be important for pathogen transmission (Sah et al., 2018). Thus, less precise contact definitions may augment but should not replace stricter, more specific definitions.

Spatial configurations influence telemetry sampling requirements

Our final objective was to determine if the effect of telemetry sampling varied with population spatial configuration. We found that random and lattice configured populations, which had lower density networks and less frequent associations, required high sampling effort to achieve high correlations between complete and sample network

metrics, especially compared to clustered populations. Lattice layouts, which are most representative of highly territorial populations with limited associations such as many large carnivore species (e.g. jaguars; de Azevedo & Murray, 2007) or between groups of some social species (e.g. Ethiopian wolves; Zubiri & Gottelli, 1995), were generally the most sensitive to the effects of telemetry sampling and may therefore require especially high sampling effort to accurately estimate network metrics. Further, relatively solitary species tend to have higher variation in their number of contacts (Sah et al., 2018), which may make these populations more sensitive to undersampling, particularly if the distribution of contacts is highly skewed (Perkins et al., 2009; Wilson et al., 2002).

This increased effort may be especially important as the rare or infrequent associations between solitary animals may be pivotal for pathogen transmission, and would be easily missed with sampling effort deemed adequate for other species (social species or species aggregating over heterogeneous resources, e.g. deer or urban racoons). For instance, among our lattice simulations, even local metrics required a frequency of sampling of every 1 to 15 minutes to achieve moderate correlation scores; this level of sampling effort is quite high, even for GPS collars, and is likely to be beyond the level of effort used in many wildlife monitoring studies. VHF telemetry, in particular, tends to capture animal locations far less frequently, such that our locations every 72 hours represent a high level of effort for VHF telemetry (Krause et al., 2013). When using spatiotemporal contact definitions, VHF telemetry is therefore unlikely to adequately capture network structure for territorial, solitary species. Further, GPS telemetry for these species is likely to require higher than average frequency of sampling when the study objective is to estimate contact networks for disease transmission. Our simulations are limited in their system specificity, however, so actual sampling effort

should be determined by empirical data or species-specific simulations. Importantly, spatial overlap definitions of contact performed poorly for these infrequently interacting species, meaning that spatial overlap is unlikely to be able to compensate for their higher sampling needs. Spatial overlap contact definitions may be more appropriate in heterogeneous landscapes or instances of resource provisioning where animals are expected to congregate (Becker & Hall, 2014) — a scenario approximated by our clustered simulations.

Challenges and future opportunities

By examining aggregated populations of individuals, our clustered simulations capture some impact of landscape heterogeneity and social biology on telemetry sampling. However, more explicit inclusion of landscape heterogeneity (e.g. through simulations based on resource selection functions, RSF; Dougherty et al., 2018; White et al., 2018b) would be particularly useful in future work to further ascertain the effect of landscape heterogeneity on telemetry sampling and subsequent network estimation. We chose a relatively simple movement model, the BCRW, as this improved model transparency and better allowed us to determine the robustness of our results to variations in model parameterization. RSF models would be an appropriate future direction to add an additional layer of complexity and further biological realism to our findings.

Further, we utilized the same movement model across all individuals in a given simulation. In reality, animal movement varies across time of day, and between different individuals. For example, in species where males maintain larger home ranges, these individuals may require higher sampling effort, given that larger home range simulations

in our study demonstrated higher sensitivity to telemetry sampling. For determining system-specific telemetry sampling needs with a simulation approach, increased biological realism should incorporate both landscape and these individual-level heterogeneities.

Contact networks used for transmission modeling should be aggregated over time frames representative of the infectious period for the pathogen of interest (White et al., 2017). Our 90-day simulation durations, therefore, make our simulations more representative of pathogens with short to moderate infectious periods (Craft et al., 2011); for longer infectious periods or populations with more rapid changes to social organization, dynamic networks might be used (Bansal et al., 2010; Volz & Meyers, 2007). Future work examining the effect of telemetry sampling over dynamic networks would help refine sampling needs for populations with fission-fusion social dynamics or seasonal changes in network structure, where such temporal changes may result in significantly different epidemic outcomes (VanderWaal et al., 2017).

Another key assumption of our simulations was that individuals behave independently of one another. Attraction behaviors would make animals more likely to interact or maintain associations for longer durations than is represented by our simulations. Such associations would be expected to make contact networks less sensitive to reduced telemetry sampling. In contrast, for avoidance behaviors, we would expect increased sensitivity to telemetry sampling. Our current simulation model therefore represents a “null” case for the impact of telemetry sampling on network estimation. Our results are, therefore, broadly generalizable across systems, and set an expectation for populations or research questions that may require more intensive

sampling. More species-specific work in the future could clarify how association behaviors impact telemetry needs.

As with any simulation study, some events have been simplified, particularly in comparison to field conditions. Our sampling approach was truly random, but even when the goal is random sampling, this condition is often unmet in the field. For example, certain geographic study areas of populations may be more thoroughly sampled, rather than a random sample of the whole population (Craft et al., 2009; Reynolds et al., 2015). While our network density results appear to be insensitive to the proportion of the population sampled, this is likely due at least in part to our random sampling, especially given that network-level metrics are generally expected to be negatively impacted by low population sampling (Farine & Whitehead, 2015). Under field conditions, where random sampling is less likely, network density may be more sensitive to the proportion of the population sampled. Future research may examine the effect of non-random or geographically biased telemetry sampling on network metric estimation.

In addition, while we varied the distance threshold for contacts, using time lags when estimating associations was beyond the scope of this study. Given the improvement in some network metrics when using less precise contact definitions, we expect that time lags also have the potential to further maximize the utility of telemetry data for estimating wildlife contact networks. Future studies characterizing the bounds of this approach would therefore be of great practical use.

Conclusions

While some network metrics and spatial configurations are more sensitive to telemetry sampling effort, remote detection of animal contacts through telemetry

technology appears to be a viable approach for estimating some network metrics. In particular, local metrics were well-approximated, particularly in clustered populations which may represent social species or animals navigating highly heterogeneous landscapes. While our results should not be used as a system-specific guide for designing telemetry protocols, they do outline several important sampling guidelines for estimating network structure in infectious disease studies. In particular, we recommend (1) using local network metrics over global metrics, (2) prioritizing frequency of sampling for territorial or infrequently interacting species, and (3) considering a spatial overlap contact definition if sampling is coarse, but only for frequently interacting, highly aggregated species (e.g. herd species). Our findings are broadly generalizable across species, but also demonstrate that future system-specific sampling efforts can be designed following our simulation approach, incorporating additional heterogeneity to inform reliable telemetry approaches for network studies of infectious disease.

2.6 Acknowledgements

Thanks to K. VanderWaal for meaningful feedback. MLJG was supported by the Office of the Director, National Institutes of Health under award number NIH T32OD010993. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. LAW was supported by the National Socio-Environmental Synthesis Center (SESYNC) under funding received from the National Science Foundation DBI-1639145. MEC was funded by the National Science Foundation (DEB-1413925 and 1654609) and CVM Research Office UMN Ag Experiment Station General Ag Research Funds. The authors acknowledge the

Minnesota Supercomputing Institute at the University of Minnesota for providing resources that contributed to the research results reported within this paper.

2.7 Data Availability

Full R code for simulations is available on GitHub (https://github.com/mjones029/Telemetry_Network_Simulations) and archived at Zenodo (https://doi.org/10.5281/zenodo.3610569).

2.8 Figures

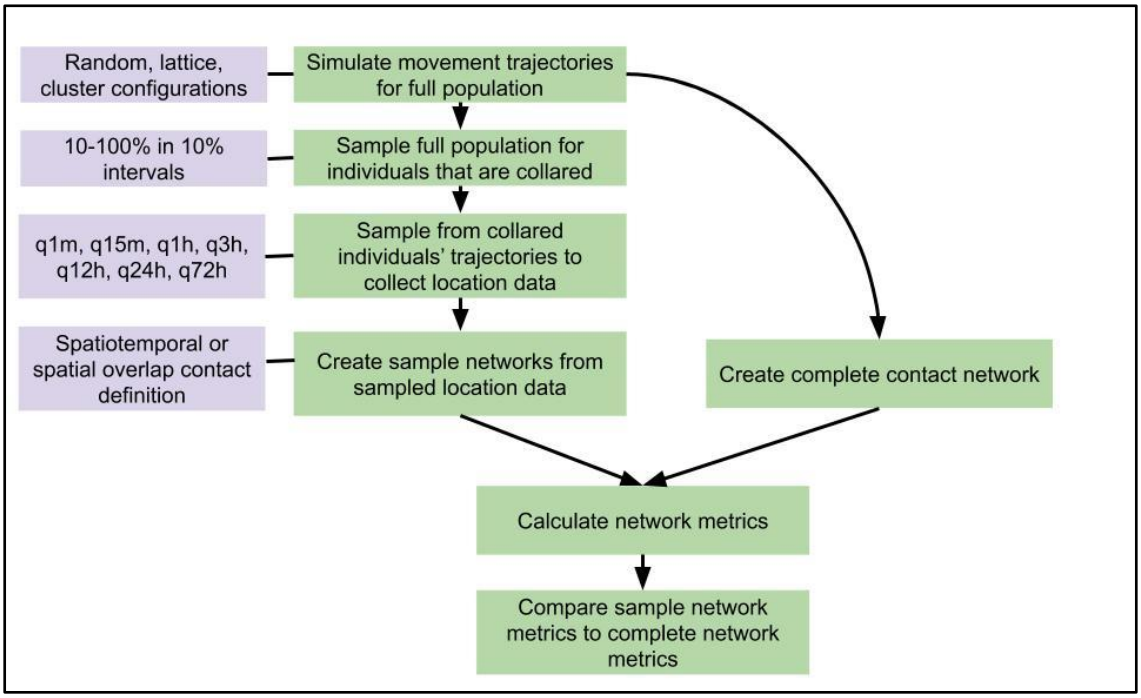


Figure 2.1: Workflow for simulations, sampling, and network estimation. Green boxes indicate individual steps in the workflow, and purple boxes indicate treatments introduced to simulations, sampling, or contact detection protocols. “q1m” means “every 1 minute,” and so on.

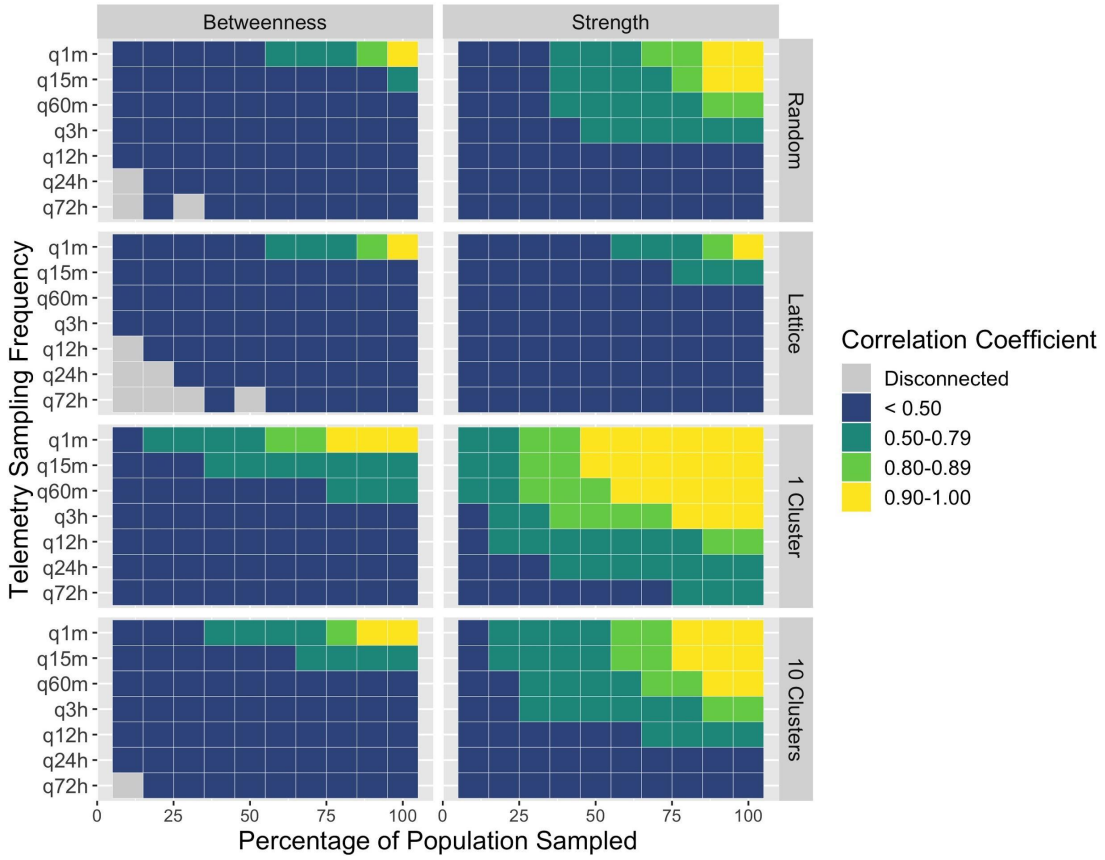


Figure 2.2: Heat maps demonstrating the lower limit of the 95% confidence interval for mean correlation coefficient results for betweenness and strength, the best performing global and local node-level network metrics, respectively. Yellow and light green cells indicate the highest correlation categories. Results here represent a comparison of complete networks and sample networks generated with large threshold contact definitions. Results shown are for medium home range movement simulations across four variations in spatial configurations (random, lattice, one cluster, and ten clusters). For telemetry sampling frequencies, “q1m” means “every 1 min,” and so on.

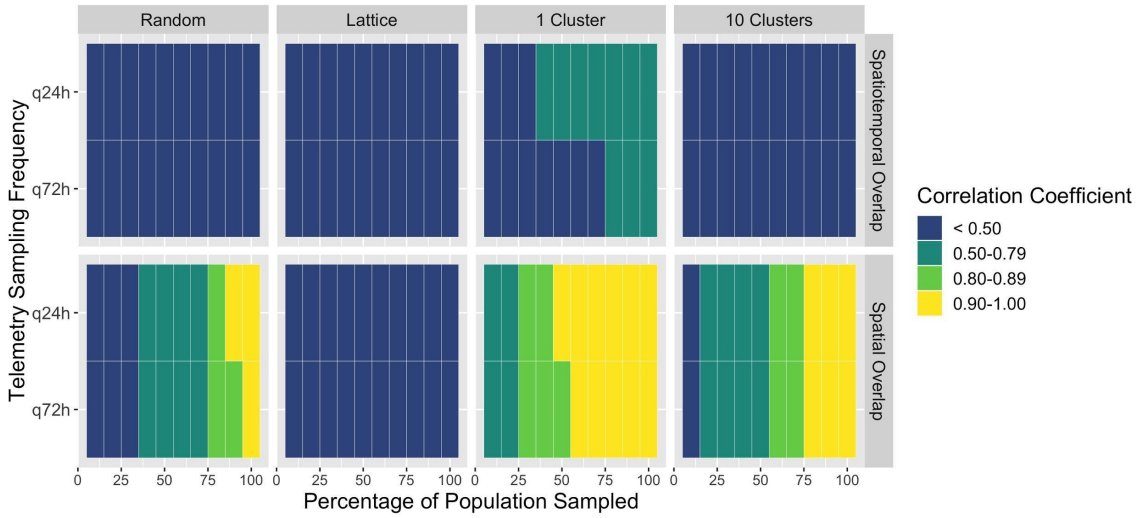


Figure 2.3: Heat maps of network strength (lower limit of the 95% confidence interval for mean correlation coefficient), which showed the strongest improvement with a spatial overlap contact definition, for medium home range simulations at the coarsest frequencies of sampling. The top panel represents the comparison between complete and sample networks each generated with large threshold spatiotemporal contact definitions (Spatiotemporal Overlap). The lower panel shows results for the comparison between complete networks with a large threshold spatiotemporal contact definition and sample networks with a spatial overlap contact definition (Spatial Overlap).

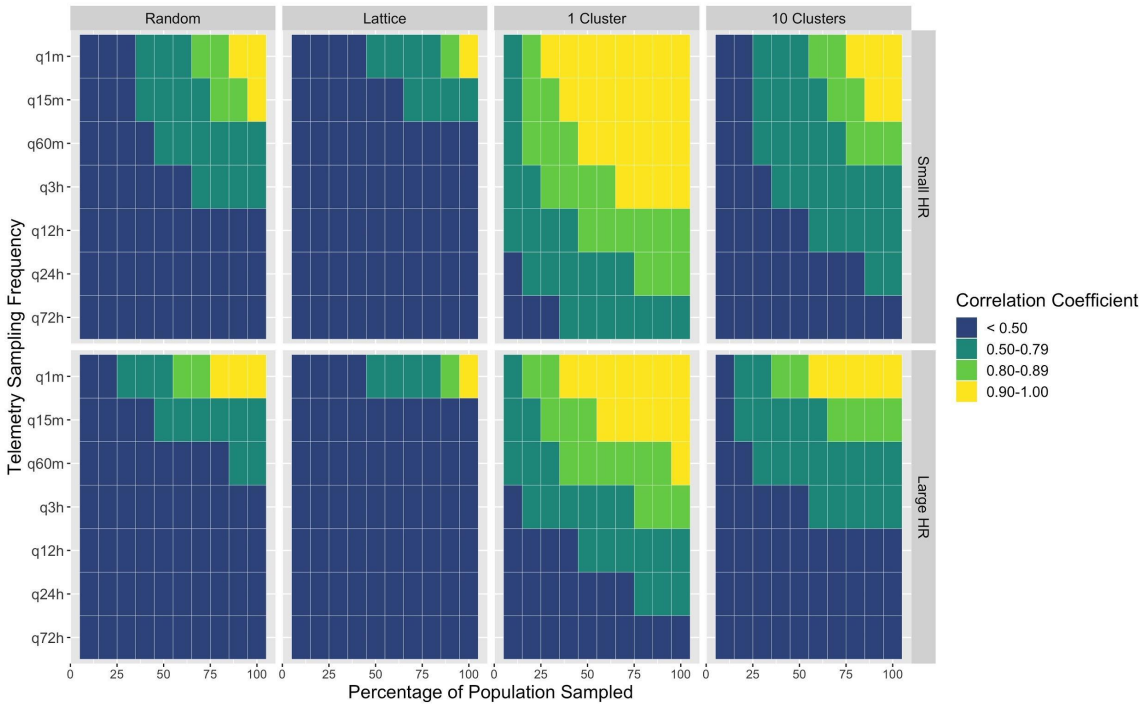


Figure 2.4: Heat maps of the lower limit of the 95% confidence interval of mean correlation coefficient for degree. Yellow and green represent the highest correlation values; results are shown across two movement model variations (small and large home ranges; Small HR and Large HR, respectively) and four spatial configurations (random, lattice, one cluster, and ten clusters), with large threshold contact definitions.

Chapter 3. Transmission of one predicts another: Apathogenic proxies for transmission dynamics of a fatal virus

Marie L.J. Gilbertson, Nicholas M. Fountain-Jones, Jennifer L. Malmberg, Roderick B. Gagne, Justin S. Lee, Simona Kraberger, Sarah Kechejian, Raegan Petch, Elliott Chiu, Dave Onorato, Mark W. Cunningham, Kevin R. Crooks, W. Chris Funk, Scott Carver, Sue VandeWoude, Kimberly VanderWaal, Meggan E. Craft

3.1 Synopsis

Identifying drivers of transmission prior to an epidemic—especially of an emerging pathogen—is a formidable challenge for proactive disease management efforts. To overcome this gap, we tested a novel approach hypothesizing that an apathogenic virus could elucidate drivers of direct contact transmission processes, and thereby predict transmission dynamics of an analogously transmitted virulent pathogen. We tested our hypothesis in a model system, the Florida panther (*Puma concolor coryi*), which is affected by several feline retroviruses, including apathogenic feline immunodeficiency virus (FIV) and pathogenic feline leukemia virus (FeLV). We derived a transmission network using FIV whole genome sequences, and used exponential random graph models to determine drivers structuring this network. We used the identified factors to predict transmission pathways among panthers, simulated FeLV transmission using these pathways and three alternate modeling approaches, and compared predictions against empirical observations from a historical FeLV outbreak in panthers. FIV transmission was primarily driven by panther age class and pairwise geographic distances. Prospective FIV-based predictions of FeLV transmission performed at least

as well as simpler, often retrospective approaches, with evidence that FIV-based predictions could capture the spatial structuring of the observed FeLV outbreak. Our finding that an apathogenic agent can predict transmission of an analogously transmitted pathogen is an innovative approach that warrants testing in other host-pathogen systems to determine its generalizability. Use of such apathogenic agents holds promise for improving predictions of pathogen transmission in novel host populations, and can thereby revolutionize proactive pathogen management in human and animal systems.

3.2 Introduction

Infectious disease outbreaks can have profound impacts on conservation, food security, and global health and economics. Mathematical models have proven a vital tool for understanding transmission dynamics of pathogens (Anderson & May, 1991; Antonovics, 2017), but may struggle to predict the dynamics of novel or emerging agents (Metcalf & Lessler, 2017; Plowright et al., 2017). This is at least partially due to the challenges associated with characterizing contacts relevant to transmission processes. Common modeling assumptions that all hosts interact and transmit infections to the same degree may ignore key drivers of transmission such as specific transmission-relevant behaviors (e.g., grooming or fighting in animals: Drewe, 2010; concurrent sexual partnerships in humans: Morris & Kretzschmar, 1995) or assortative mixing (Cauchemez et al., 2011), resulting in flawed epidemic predictions (Craft & Caillaud, 2011; Keeling & Eames, 2005). Further, identifying drivers of transmission and consequent control strategies for any given pathogen is typically done reactively or retrospectively in an effort to stop or prevent further outbreaks or spatial spread (e.g. Keeling et al., 2003; Lloyd-Smith et al., 2009; Smith et al., 2002). These constraints limit

the ability to perform prospective disease management planning tailored to a given target population, increasing the risk of potentially catastrophic pathogen outbreaks, as observed in humans (Dobson et al., 2020; Keogh-Brown & Smith, 2008; Pike et al., 2014), domestic animals (Blake et al., 2003; Knight-Jones & Rushton, 2013) and species of conservation concern including Ethiopian wolves (Sillero-Zubiri et al., 1996), African lions (Roelke-Parker et al., 1996), black-footed ferrets (Williams et al., 1988), and Florida panthers (Cunningham et al., 2008).

Some efforts have been made to determine if common infectious agents present in the healthy animal microbiome or virome can capture contact patterns, which may translate to interactions relevant to transmission. Such an approach could circumvent some of the uncertainties associated with more traditional approaches to contact detection by using genetic evidence from the transmissible agent itself to define between-individual interactions for which contact was sufficient for transmission to occur. Prior studies, however, have found mixed results thus far (Blasse et al., 2013; Blyton et al., 2013; Bull et al., 2012; Chiyo et al., 2014; Springer et al., 2016; VanderWaal et al., 2014a). For example, in humans, members of the same household have been found to share microbiota (Lax et al., 2014; Song et al., 2013), but disentangling social mechanisms of this sharing is complicated by shared diets, environments, behaviors, etc. (Archie & Tung, 2015). In animals, studies of *Escherichia coli* in Verreaux's sifaka and giraffe have found strain sharing relationships to be tied to social interactions (Springer et al., 2016; VanderWaal et al., 2014a), but the same was not found in a similar study of elephants (Chiyo et al., 2014).

These studies, however, reveal some characteristics of ideal non-disease-inducing infectious agents (hereafter, *apathogenic agents*) for use as potential markers

of transmission-relevant interactions. Such apathogenic agents should have rapid mutation rates to facilitate discernment of transmission relationships between individuals over time (Gilbertson et al., 2018; Hall et al., 2016). Furthermore, these agents should be relatively common and well-sampled in a target population, have a well-characterized mode of transmission, and feature high strain alpha-diversity (local diversity) and high strain turnover (Blyton et al., 2014; Gilbertson et al., 2018). Many RNA viruses may best align with these characteristics (Archie et al., 2009; Grenfell et al., 2004). Some apathogenic RNA viruses could thereby act as "proxies" of specific modes of transmission (i.e., direct transmission) and determine which factors drive these transmission processes. Such factors may subsequently be able to predict transmission dynamics of pathogenic agents operating under the same mode of transmission (Gilbertson et al., 2018). To the best of the authors' knowledge, however, RNA viruses have not been tested as models for direct transmission processes in this way.

Here, we use a naturally occurring host-pathogen system to test if an apathogenic RNA virus can act as a model for direct transmission processes and subsequently predict transmission of a related pathogenic agent. Florida panthers (*Puma concolor coryi*) are an endangered subspecies of puma found only in southern Florida and have been extensively studied and monitored for almost four decades. Panthers are also infected by several feline retroviruses relevant to our study questions. Feline immunodeficiency virus (FIVpco; hereafter, FIV) is a relatively common lentivirus in panthers, and does not appear to cause significant clinical disease. FIV is transmitted by close contact (i.e., fighting and biting), generally has a rapid mutation rate (intra-individual evolution rate of 0.00129 substitutions/site/year; Krakoff et al., 2019), and, as a chronic retrovirus infection, can be persistently detected after the time of infection. In

contrast, panthers are also affected by feline leukemia virus (FeLV), a related retrovirus, which spills over into their population owing to predation of infected domestic cats (Brown et al., 2008). Once spillover has occurred, FeLV is also transmitted by close contact, and causes significant clinical disease in panthers, with a major outbreak documented among panthers in 2002-2004 (Cunningham et al., 2008).

We hypothesized that FIV can act as a proxy for direct transmission in panthers, revealing which measurable factors drive direct transmission events. We then tested the accuracy and utility of FIV in this role by predicting transmission dynamics of FeLV based on these factors. A key advantage of this naturally-occurring system is the well-observed historical FeLV outbreak in panthers, which allowed us to compare our predictions against empirical observations. The objectives of this study were therefore: (1) to determine which factors shape FIV transmission in Florida panthers, and (2) test if these factors can predict transmission dynamics of analogously-transmitted FeLV in panthers. Success of this approach in our model system would pave the way for testing similar apathogenic agents in other host-pathogen systems, thereby improving our ability to predict transmission dynamics of novel agents in human and animal populations.

3.3 Materials and Methods

Dataset assembly

We assembled an extensive dataset covering more than 30 years of Florida panther research. Ongoing panther management has documented age and sex of monitored panthers. In addition, a subset of the population is monitored using very high frequency (VHF) telemetry collars, with relocations determined via aircraft typically three times per week. Previous panther research has generated a microsatellite dataset for

monitored panthers (Van De Kerk et al., 2019), and a dataset of 60 full FIV genomes (proviral DNA sequenced within a tiled amplicon framework in Malmberg et al., 2019). In addition, the historical FeLV outbreak in panthers was well documented (Brown et al., 2008; Cunningham et al., 2008), providing key observations regarding FeLV dynamics in free-ranging panthers. To augment these observations, we also used an FeLV database which documents FeLV status (positive and negative) for 31 panthers as determined by qPCR from 2002-04.

FIV transmission inference

To determine predictors of FIV transmission, we first generated a "who transmitted to whom" transmission network using panther FIV genomes (see Figure 3.1 for workflow across all analyses). We used the program Phyloscanner (Wymant et al., 2018), which maximizes the information gleaned from next generation sequencing reads to infer transmission relationships. Phyloscanner operates in a two-step process, first inferring within and between host phylogenies in windows along the FIV genome, and then analysing those phylogenies to produce transmission trees or networks. For step one, we used 150bp windows, allowing 25bp overlap between windows. To test sensitivity to this choice, we separately ran a full Phyloscanner analysis with 150bp windows, but without overlap between windows (Appendix B). For step two, we held k, which penalizes within-host diversity, equal to 0. We used a patristic distance threshold of 0.05 and allowed missing and more complex transmission relationships. Because we had uneven read depth across FIV genomes, we downsampled to a maximum of 200 reads per host. The output of the full Phyloscanner analysis was a single transmission network (hereafter, *main FIV network*), but see Appendix B for details regarding analysis

of the sensitivity of our results to variations in and summary across multiple transmission networks.

Statistical analysis of FIV transmission networks

We performed statistical analysis of unweighted, binary FIV transmission networks using exponential random graph models (ERGMs), which account for non-independence in network structure (Silk & Fisher, 2017). The network structural terms we considered included an intercept-like edges term (Silk & Fisher, 2017), geometrically weighted edgewise shared partner distribution (*gwesp*; representation of network triangles), alternating k-stars (*altkstar*; representation of star structures), and 2-paths (2 step paths from *i* to *k* via *j*; (Morris et al., 2008).

The dyad-independent variables included panther sex (both as a node factor and node mixing variable; see Appendix B for terminology and additional variable details). We assessed panther age as a categorical variable (both as a node factor and node mixing variable), with subadults being individuals between the ages of 6 months and two years, and adults being individuals over two years of age. We included pairwise genetic relatedness from panther microsatellite data as an edge covariate. Spatial variables included a node-matching variable for the location of panthers' minimum convex polygon (MCP) home range centroid or capture location (hereafter *centroid*; see Appendix B) north versus south of the major I-75 freeway. In addition, we included a node covariate term for the distance from the centroid to the nearest urban area (in km; USA Urban Areas layer, ArcGIS; Esri, National Atlas of the United States, United States Geological Survey, Department of Commerce, Census Bureau-Geography Division). Pairwise geographic distances between panthers were calculated using distances between

centroids (in km), and log-transformed for use as an edge covariate. Lastly, we included a spatial overlap edge covariate based on the pairwise utilization distribution overlap indices of 95% home range kernels (Fieberg et al., 2005), using the *adehabitat* package in R (Calenge, 2006).

Because ERGMs are prone to degeneracy with increasing complexity (Silk & Fisher, 2017), we first performed forward selection for network structural variables, followed by forward selection of dyad-independent variables, while controlling for network structure. Model selection was based on AIC and goodness of fit, and MCMC diagnostics were assessed for the final model (Appendix B).

Panther population simulations

To test if predictors of FIV transmission identified in the ERGM analysis can predict FeLV transmission, we next simulated FeLV transmission through a network which was based on these FIV predictors among populations representing panthers during the historical FeLV outbreak (2002-2004). Hereafter, a *full-simulation* includes both simulation of the panther population and network during the historical outbreak period and simulation of FeLV transmission within that population. Below, we describe the process for a single simulation, but these procedures were repeated for each full simulation.

To simulate the historic panther population, we first based the simulated population size on the range of empirical estimates from 2002-2004 (McClintock et al., 2015). Additional characteristics of the simulated population included those identified as significant predictors in the ERGM analysis: age category and pairwise geographic distances between panther home range centroids. We randomly assigned age

categories to the simulated population based on proportion adult versus subadult. Age proportions were based on age distributions in the western United States (Logan & Sweanor, 2001) which qualitatively align with the historically elevated mean age of the panther population (Johnson et al., 2010), as these distributions are not well established in Florida panthers. Pairwise geographic distances for the simulated population were generated by randomly assigning simulated home range centroids based on the distribution of observed centroids on the landscape (Appendix B).

With the simulated panther population, we then used ERGM coefficients to generate network edges in the population representing potential transmission pathways between panthers. The original FIV transmission network spanned 15 years of observations and represents a subset of the actual contact network, as it includes only those interactions that resulted in successful transmission (Craft, 2015). We therefore had a high degree of uncertainty regarding the appropriate network density for our simulations. To manage this uncertainty, we constrained density in our network simulations across a range of parameter space (net_dens, Table 3.1).

Simulation of FeLV transmission on FIV-based networks

The next step in each full simulation was to model FeLV transmission on the network generated from FIV predictors of transmission. FeLV transmission was based on a chain binomial process on the simulated networks, following a modified SIR compartmental model progression (Figure 3.2). Time steps were in weeks, and transmission simulations lasted until no infectious individuals remained or until 2.5 years, whichever came first. Simulations were initiated with one randomly selected infectious individual.

Transmission (Figure 3.2) was dependent on the following: (1) existence of an edge between two individuals, (2) the dyad in question involving a susceptible and infectious individual, and (3) a random binomial draw based on the probability of transmission given contact (β, Table 3.1). In addition, puma generally have low expected weekly contact rates (Elbroch & Quigley, 2016); we therefore included an additional weekly contact probability, represented as a random binomial draw for contact in a given week (ω, Table 3.1).

Upon successful transmission, infectious individuals were randomly assigned to one of three outcomes of FeLV infection (Cunningham et al., 2008). *Progressive* infections (probability P, Table 3.1) are infectious, develop clinical disease and die due to infection at a mortality rate, μ. *Regressive* infections (probability P) recover at a rate based on a constant, K, multiplied by the mortality rate of progressives (Table 3.1). Anecdotal evidence suggests regressive individuals are not infectious (Cunningham et al., 2008), but given ongoing uncertainty, we allowed regressives to be infectious by multiplying the probability of transmission for progressives (β) by a constant, C (Table 3.1). *Abortive* infections (probability 1-2P) are never infectious, clearing infection and joining the recovered class.

A vaccination process was included in simulations as panthers were vaccinated against FeLV during the historical FeLV outbreak. As in the observed outbreak, simulated vaccination only occurred after an outbreak's first year. Vaccination occurred at a rate, τ, applied to the whole population, as wildlife managers are unlikely to know if a panther is susceptible at the time of capture or darting. However, only susceptible individuals transitioned to the vaccinated class (i.e. vaccination failed in non-susceptibles). Because panthers were vaccinated in the empirical outbreak with a

domestic cat vaccine with unknown efficacy in panthers, we allowed vaccinated individuals to become infected in transmission simulations by including a binomial probability for vaccine failure (1-vaccine efficacy, ve, Table 3.1).

The panther population size remained roughly static through the course of the FeLV outbreak (McBride et al., 2008; McClintock et al., 2015). We therefore elected not to include background mortality, but did include infection-induced mortality. To maintain a consistent population size, we therefore included a birth/recruitment process. Because FIV-based simulated networks drew edges based on population characteristics, we treated births as a "respawning" process, in which territories vacated due to mortality were reoccupied by a new susceptible at rate, ν. This approach allowed us to maintain the ERGM-based network structure and is biologically reasonable, as vacated panther territories are unlikely to remain unoccupied for long.

While FeLV is well-documented in domestic cats, infection is uncommon in puma as a species, resulting in a high degree of uncertainty regarding differences in within-individual infection dynamics in panthers. Given these uncertainties, we performed all simulations across a range of parameter space. To more efficiently cover this parameter space, we generated parameter sets using a Latin hypercube design (LHS), using the *lhs* package in R (Carnell, 2012). We generated 150 parameter sets, conducting 50 full simulations per parameter set.

Comparison of simulation predictions to observed FeLV outbreak

We compared FIV-based simulation predictions to the observed FeLV outbreak, but also compared predictions from three simpler types of models: random networks,

home range overlap-based networks, and a well-mixed model. All models used the parameterizations from our LHS parameter sets, as relevant.

For our random networks model, we generated Erdos-Renyi random networks, with the simulated network densities from our LHS parameter sets (Table 3.1). Overlap-based networks were generated using the degree distributions of panther home range overlap networks from 2002-2004 and simulated annealing with the R package *statnet* (Handcock et al., 2008; Reynolds et al., 2015; Appendix B). These overlap-based networks were not spatially explicit, as they were based only on the degree distributions from real spatial overlap networks. For both random and overlap-based networks, FeLV transmission was simulated as in the FIV-based simulations. The well-mixed model was a Gillespie algorithm (stochastic, continuous time compartmental model), with rate functions aligning with the chain binomial FeLV transmission probabilities (Appendix B).

Target ranges for predicted outcomes were based on observed FeLV dynamics (Cunningham et al., 2008), with ranges to account for uncertainty in observations and population size in this cryptic species (Appendix B). The primary outcomes of interest were (1) duration of outbreak: 78-117 weeks, (2) total number of progressive infections: 5-20, and (3) presence of spatial clustering (see below). While our primary focus was progressive infections, we also included an expectation that at least 5 individuals were abortive infections. Empirically, these individuals were the most numerous, but as they were not clinically ill, abortive infections were less likely to be detected in normal panther management; we therefore did not include an upper bound for this target.

Using our database of qPCR results for FeLV in panthers (positive and negative tests), we performed a local spatial clustering analysis of FeLV cases and controls using SaTScan (50% maximum, circular window; Kulldorff, 1997), and a global cluster analysis

with Cuzick and Edward's test in the R package *smacpod* (1, 3, 5, 7, 9, and 11 nearest neighbors; 999 iterations; Cuzick & Edwards, 1990; French, 2020). These analyses found evidence of local (weak) and global clustering (at 3, 5, and 7 neighborhood levels) among progressive and regressive cases (see results, Appendix B). In simulations, we therefore included spatial clustering of progressive and regressive cases as a target outcome.

For the duration of outbreaks and total number of progressives, we calculated median values of both outcomes for each parameter set (i.e., 50 simulations) within each model type. If a parameter set's medians were within the target ranges for both of these outcomes, it was considered *feasible*. To quantify differences in model prediction performance, we fit a binomial generalized linear mixed model (GLMM), assuming a logistic regression with "feasible" as the outcome, model type as a predictor variable, and a random intercept for LHS parameter set.

To determine if simulated results demonstrated spatial clustering, we performed SaTScan spatial cluster analysis (50% maximum, circular window) and Cuzick and Edward's tests (at 3, 5, and 7 nearest neighbors) on simulation results. Only FIV-based simulations were spatially explicit, and we performed spatial analyses only on those parameter sets that were classified as feasible. To determine if any detected clustering in FIV simulations was simply based on our respawning protocol, however, we also performed both spatial analyses with feasible overlap-based simulation results. For these, we assigned the same geographic locations to nodes in the overlap-based networks from the corresponding FIV-based networks (i.e., matching simulation number from matching parameter set).

3.4 Results

FIV transmission network analysis

The main FIV network included 19 nodes (individuals) with 42 edges (representing potential transmission events; network density = 0.25) after removing 9 edges that were between individuals known not to be alive at the same time (Appendix B, Figure B.1). ERGM results for the main FIV network identified geometrically weighted edgewise shared partner distribution (*gwesp*) and alternating k-stars (*altkstar*) as key structural variables, and age category (as a node-level factor) and log transformed pairwise geographic distance as key dyad-independent variables (Table 3.2). Though altkstar was not statistically significant, inclusion of this variable contributed to improved AIC and goodness of fit outcomes. Adults were more likely to be involved in transmission events (but see discussion of sample size limitations) and inferred transmission events were more likely between individuals which were geographically closer to each other. The fitted model showed reasonable goodness of fit (Appendix B, Figure B.2). ERGM results were largely consistent across replicate analyses with alternative transmission networks formed by summarizing across four single Phyloscanner outputs (see Appendix B for further details).

FeLV simulations

SaTScan analysis of observed FeLV status found weak evidence of local spatial clustering (two clusters detected, but not statistically significant with p=0.165 and 0.997, respectively; Appendix B, Figure B.4). Cuzick and Edward's tests found evidence of global clustering at 3, 5, and 7 nearest neighbor levels (test statistic T_k where k is number of nearest neighbors considered: $T_3 = 20$, $p = 0.049$; $T_5 = 32$, $p = 0.028$; $T_7 = 43$,

p = 0.023). Both sets of spatial analysis results were then compared against FeLV predictions from FIV and overlap-based models.

About 9% of parameter sets across all model types were classified as feasible (Figure 3.3; Appendix B, Figures B.5-B.6). The GLMM for model type performance (i.e., FIV-based, random, overlap-based, or well-mixed) did not find statistically significant differences between odds of generating feasible simulation outcomes, though the FIV-based model had the highest odds of feasibility (exponentiated estimate = 1.55, though p = 0.30; Appendix B, Table B.2). Feasible parameter sets from both the FIV-based and overlap-based models produced some evidence of local and global spatial clustering of simulated FeLV cases (Figure 3.4; Appendix B, Figure B.7). However, the FIV-based model was better able to capture the size and strength (observed/expected FeLV cases) of predicted local clusters (Figure 3.4) and was moderately better at capturing global spatial patterns (Appendix B, Figure B.7).

In order to determine if certain transmission parameters were particularly important for feasible performance, we performed *post hoc* random forest analyses using the R package *randomForest* (Liaw & Wiener, 2002; White et al., 2020) for each of the four model types (see Appendix B). While random forests typically showed poor balanced accuracy and area under the curve (AUC) results, the parameter shaping transmission from regressively infected individuals (C), showed support for weak to moderate transmission from regressives (i.e., C = 0.1 or 0.5; Appendix B, Figure B.10).

3.5 Discussion

In this study we develop a new approach whereby we leveraged knowledge of transmission dynamics of a common apathogenic agent to prospectively predict

dynamics of an uncommon and virulent pathogen. Our approach was distinctly different from simpler models we tested, as the apathogenic (FIV)--based approach could be used to prospectively identify predictors of transmission and develop disease control plans prior to an outbreak of a virulent pathogen (FeLV), while other approaches either make broad assumptions about transmission-relevant contacts (e.g., homogeneous mixing), or rely on retrospective or reactive modeling. We found that FIV transmission in panthers is primarily driven by distance between home range centroids and age class, and that our prospective FIV-based approach predicted FeLV transmission dynamics at least as well as simpler or more reactive approaches. While we do not propose that this apathogenic agent approach could accurately predict exactly when, where, and to whom transmission might occur, our results support the role of apathogenic agents as novel tools for prospectively identifying relevant drivers of transmission and consequently improving proactive disease management.

Pairwise geographic distances and panther age class predict FIV transmission

Combining genomic and network approaches, we determined that pairwise geographic distances and age category structure FIV transmission in the Florida panther. Because FIV is a persistent infection, we would expect cumulative risk of transmission to increase over an individual's lifetime and adults would consequently be involved in more transmission events. The low number of subadult individuals in our dataset, however, means that this finding must be interpreted with caution.

Panthers are wide-ranging animals but maintain home ranges, and this appears to translate to increased transmission between individuals that are close geographically. This finding is further supported by the tendency for FIV phylogenies to show distinct

geographic clustering (Franklin et al., 2007; Lee et al., 2014), but is in contrast to other infectious agents of puma. A related feline retrovirus, feline foamy virus (FFV), does not show distinct geographic clustering but is commonly transmitted between domestic cats and puma (Kraberger et al., 2020). A prior study of several pathogens in puma across the United States rarely identified spatial autocorrelation in pathogen exposures, but notably found that FIV infection status approached statistical significance specifically in Florida panthers (Gilbertson et al., 2016). The wide-ranging nature of puma appears to limit geographic clustering of many infectious agents, with FIV a notable exception to this pattern. Multi-host agents such as FFV are presumably more able to escape geographical limitations, and thereby lack spatial structuring. More generally, the importance of distance rather than panther relatedness in structuring transmission may support the resource dispersion or land tenure hypotheses as drivers of spatial and social structuring in panthers, rather than kinship. The high inbreeding among panthers (Johnson et al., 2010) may limit our power for identifying such a relationship between relatedness and FIV transmission, but support for resource dispersion or land tenure would be in agreement with findings in other puma systems (Elbroch et al., 2016) and even other territorial carnivores (Brandell et al., 2020).

Surprisingly, sex was not a significant predictor of FIV transmission. FIV force of infection is generally higher in male panthers, likely due to their increased fighting behaviors (Reynolds et al., 2019). However, studies in other felid species have found mixed importance of sex for FIV transmission, ranging from little to no importance (puma in the western United States: Fountain-Jones et al., 2019; bobcat: Fountain-Jones, Craft, et al., 2017) to importance only among certain FIV subtypes (African lions: Fountain-Jones, Packer, et al., 2017). Our results add to this body of research to suggest that the

relationship between host sex and FIV transmission is more complex than can be explained by sex-specific behaviors or susceptibility alone.

An FIV-based model captures FeLV transmission dynamics

In our study, a network model based on principles of FIV transmission produced FeLV outbreak predictions consistent with the observed FeLV outbreak. The FIV-based approach performed at least as well as simpler models, with evidence that FIV better predicted the observed spatial dynamics for FeLV transmission. A key difference between the FIV-based approach and other spatially-explicit methods is that FIV allowed us to determine the importance of spatial dynamics prospectively and then translate to predictions of FeLV transmission, rather than relying on retrospective FeLV spatial analyses. Furthermore, while more complex potential drivers of transmission (e.g., host relatedness or assortative mixing by age or sex) were not found to be important for FIV transmission, these may yet be key factors structuring transmission in other systems. Simpler model types like random networks or metapopulation models may struggle to make transmission predictions that incorporate these factors as drivers of transmission-relevant contact. The predictive power we observed here using an apathogenic virus could thus significantly shape proactive epidemic management strategies for pathogens such as FeLV.

FIV and FeLV have different epidemiologies (e.g., progressive versus regressive infections, duration of infectiousness); despite this, FIV-based predictors of transmission were able to capture dynamics of FeLV transmission. Here, FIV determined the key drivers of close, direct contact transmission in panthers, fundamentally acting as a proxy for this type of contact. FeLV simulations were then able to independently account for

differences in epidemiology to produce appropriate predictions for a different but analogously transmitted virus. Key components of this success are likely that (1) FIV is a largely species-specific virus with transmission pathways closely matching intra-species transmission of FeLV, and (2) both FIV and FeLV, perhaps unusually for infectious agents of puma, display spatial clustering of infection. If, for example, FIV also exhibited strong vertical or environmental transmission, we would no longer expect the predictive success for FeLV we observed here.

Notably, FeLV is not species-specific, having originated in panthers through spillover from domestic cats (Brown et al., 2008). Recent research suggests that FeLV may continue to percolate within the panther population but is also subject to repeated spillover events from domestic cats (Chiu et al., 2019). Determining the predictors and frequency of these spillover events would be less feasible with FIV, which is largely species-specific (Lagana et al., 2013; VandeWoude et al., 2010; VandeWoude & Apetrei, 2006). Rather, targeted investigation of spillover dynamics between these species would be necessary, and could use other apathogenic viral candidates that are frequently transmitted from domestic cats to puma, such as FFV (Kraberger et al., 2020). Using such agents to identify drivers of spillover events could be key for better understanding the dynamics of “pathogen release from reservoir hosts” (Plowright et al., 2017), which is of profound relevance across wildlife, domestic animal, and human systems.

While few parameter sets were classified as feasible across all model types, this appears to be predominantly the result of the wide range of parameter space explored through our LHS sampling design. This limitation was fundamentally due to uncertainties in FeLV transmission parameters, and is representative of the uncertainties experienced

in predicting transmission of emerging or understudied pathogens. Key here were the interacting uncertainties regarding infectiousness of regressives, the number of introductions of FeLV to the panther population, and the duration of FeLV infection in panthers. All three of these dynamics can have significant impacts on the duration of a simulated epidemic, allowing an epidemic to continue to "stutter" along at low levels (Blumberg & Lloyd-Smith, 2013), much as was observed in the empirical FeLV outbreak. Our *post hoc* random forest analysis provided some evidence of weak transmission from regressive individuals, but this finding would need to be validated with additional research, as it is in stark contrast to FeLV dynamics in domestic cats. Reducing uncertainties in these three key dynamics would significantly narrow the range of our predictions, and even assist in ongoing management efforts for FeLV in endangered panthers.

Furthermore, the effect of transmission parameter uncertainties underscores the importance of linking laboratory and model-based research to generate more accurate transmission forecasts (Plowright et al., 2008). Experimental research could help to narrow the range of parameter space for FeLV—or other emerging pathogens—to produce more consistent and accurate model predictions. This necessity is all the more apparent during the current COVID-19 pandemic, in which mathematical models have benefited from rapid laboratory and epidemiological research to reduce uncertainty in model parameters.

Limitations and future directions

This study found evidence for the utility of an apathogenic agent to predict transmission of a related pathogenic agent, but this approach must now be tested in

additional host-pathogen systems. The mixed results when using commensal agents to identify close social relationships in other systems (Blasse et al., 2013; Blyton et al., 2013; Bull et al., 2012; Chiyo et al., 2014; Springer et al., 2016; VanderWaal et al., 2014a) highlights that some host-apathogenic agent combinations will work better than others for determining drivers of transmission. More research is therefore necessary to determine which apathogenic agents may be most suitable as markers of transmission, and how divergent an apathogenic agent may be from a pathogen of interest while still predicting transmission dynamics.

The suite of tools for inferring transmission networks from infectious agent genomes is rapidly expanding (Firestone et al., 2019; Hall et al., 2016). In this study, we used the program Phyloscanner as it maximized the information from our next generation sequencing viral data. However, our FIV sequences were generated within a tiled amplicon framework (Malmberg et al., 2019; Quick et al., 2017), which biases intrahost diversity and likely limits viral haplotypes (Grubaugh et al., 2019). Phyloscanner was originally designed to analyze RNA from virions and not proviral DNA, as we have done here. We have attempted to mitigate the effects of these limitations by analyzing several different Phyloscanner outputs to confirm consistency in our results, and by using only binary networks to avoid putting undue emphasis on transmission network edge probabilities, as these are likely highly uncertain. Further, our primary conclusions from the transmission networks—that age and pairwise distance are important for transmission—are biologically plausible and supported by other literature, as discussed above. Nevertheless, future work should evaluate additional or alternative transmission network inference platforms.

In addition, ERGMs assume the presence of the “full network” and it is as yet unclear how missing data may affect transmission inferences (Silk & Fisher, 2017). ERGMs are also prone to degeneracy with increased complexity and do not easily capture uncertainty in transmission events, as most weighted network ERGM (or generalized ERGM) approaches have been tailored for count data (e.g., Krivitsky, 2012). ERGMs may therefore not be the ideal solution for identifying drivers of transmission networks in all systems. Alternatives may include advancing dyad-based modeling strategies (Wilber et al., 2019), which may more easily manage weighted networks and instances of missing data.

Our FIV-based approach required extensive field sampling, and many disciplines from viral genomics through simulation modeling. However, with increasing availability of virome data and even field-based sequencing technology, our approach may become more accessible with time. Further, the predictive benefits seen here, while needing further testing and validation, could become a key strategy for proactive pathogen management in species of conservation concern, populations of high economic value (e.g. production animals), or populations at high risk of spillover, all of which may most benefit from rapid, efficient epidemic responses.

Conclusions

Here, we integrated genomic and network approaches to identify drivers of FIV transmission in the Florida panther. This apathogenic agent acted as a marker of close, direct contact transmission, and was subsequently successful in predicting the observed transmission dynamics of the related pathogen, FeLV. Further testing of apathogenic agents as markers of transmission and their ability to predict transmission of related

pathogenic agents is needed, but holds great promise for revolutionizing proactive epidemic management across host-pathogen systems.

3.6 Acknowledgements

Thanks to M. Michalska-Smith, K. Worsley-Tonks, J. Mistrick, and S. N. Hart for key feedback. This research was supported by the National Science Foundation (DEB-1413925). MLJG was supported by the Office of the Director, National Institutes of Health (NIH T32OD010993), the University of Minnesota Informatics Institute MnDRIVE program, and the Van Sloun Foundation. JLM was supported by the ACVP/STP Coalition for Veterinary Pathology Fellows and the Linda Munson Fellowship for Wildlife Pathology Research. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

3.7 Tables

Table 3.1: Network and transmission simulation parameters

Parameter	Definition	Range	Citations
Pop_size	Population size	80-120	(McClintock et al., 2015)
Adult_prop	Proportion adults versus subadults	0.82-0.99	(Logan & Sweanor, 2001)
Net_dens	Simulated network density	0.05-0.15	*NA*
β	Probability of transmission from progressives, given effective contact	0.17-0.29	(Fromont et al., 1997)
C	Constant multiplier for probability of transmission from regressives, given effective contact	0, 0.1, 0.5, 1	*NA*
ω	Weekly probability of contact	0.1-0.4	(Elbroch & Quigley, 2016)
μ	Weekly probability of death from progressive infection	1/18, 1/26*	(Cunningham et al., 2008)
K	Constant multiplier for weekly probability of recovery from regressive infection	0.5, 1	*NA*
ν	Weekly probability of territory repopulation ("respawn rate")	1/12-1/4	*NA*
τ	Weekly probability of vaccination	0.5-1	*NA*
ve	Probability of vaccine efficacy	0.4-1	(Cunningham et al., 2008)
P	Proportion randomly assigned to progressive, regressive	0.25	(Cunningham et al., 2008)

**We tested a lower death rate (prolonged duration of infection) due to the low number of observed panther cases and the generally longer infection duration in domestic cats (Hartmann, 2012).*

Table 3.2: Main FIV transmission network exponential random graph model results

Variable	**Estimate**	**SE**	***p*-value**
Edges (intercept)	-2.56	1.33	0.055
gwesp	0.98	0.26	<0.001
altkstar	-0.70	0.96	0.47
Age (Adult)	0.93	0.44	0.03
Log pairwise distance	-0.45	0.21	0.03

Note: "gwesp" is geometrically weighted edgewise shared partner distribution and "altkstar" is alternating k-stars. Age classes were subadult and adult; pairwise distances were between home range centroids. Only those variables from the final model are shown. Estimates shown are untransformed; SE represents standard error; p-*values less than 0.05 were considered statistically significant.*

3.8 Figures

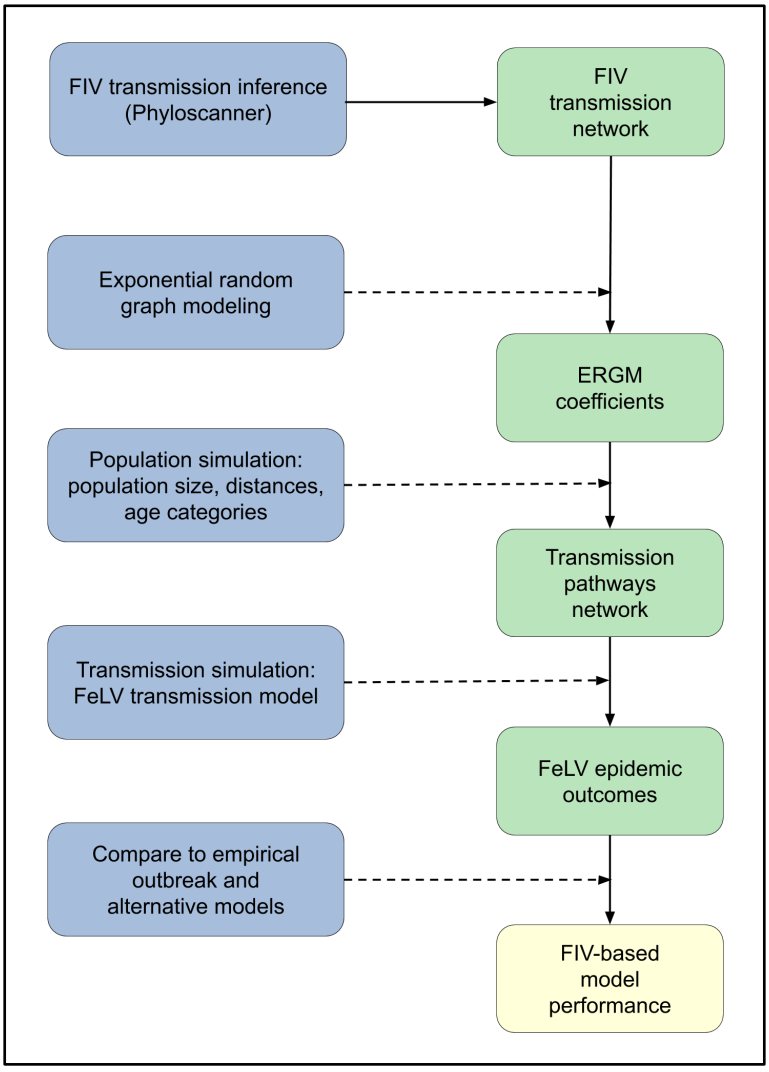


Figure 3.1: Conceptual workflow across all analysis steps. Processes are shown on the left in blue; specific outcomes are shown on the right in green; the final analysis outcome is in yellow at the bottom right. Solid lines show direct flows or outcomes. Dashed lines show processes acting on or in concert with prior outcomes: for example, exponential random graph modeling was performed using the FIV transmission network, and the combination of the two produced the ERGM coefficients outcome.

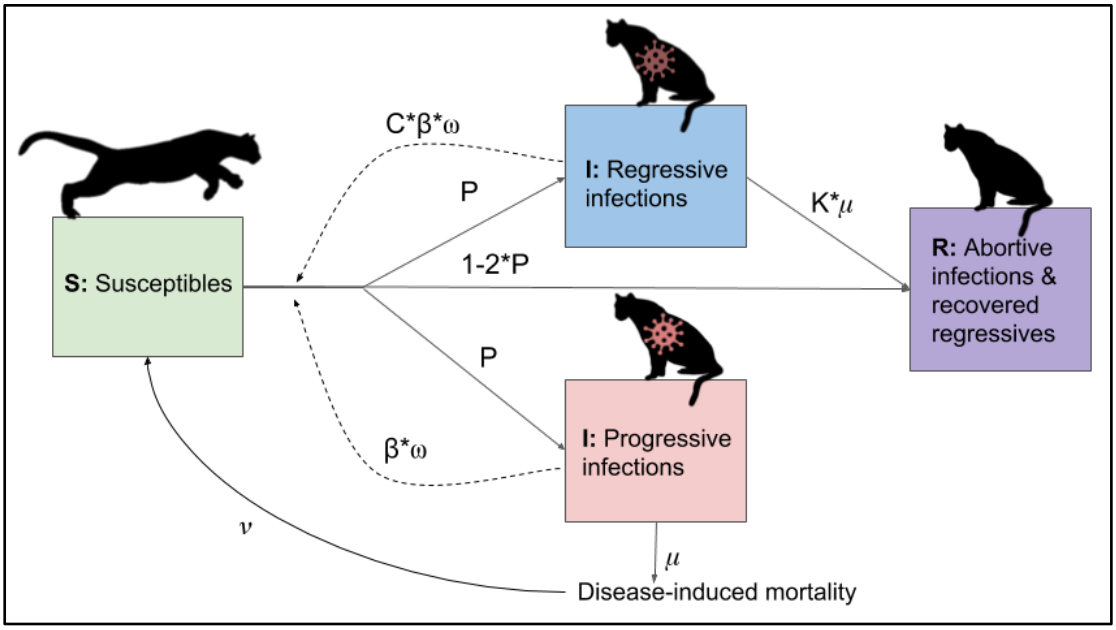


Figure 3.2: Diagram of flows between compartments in transmission model. Virus icons indicate infectious states, with the regressive infection icon darkened to represent reduced or uncertain infectiousness of this class. Note: a vaccination process was also included in the transmission model, but is not shown for simplicity. Susceptibles could be vaccinated, and vaccinated individuals subsequently infected analogously to susceptibles, but with an additional probability of (1-ve). See Table 3.1 for definitions of parameters.

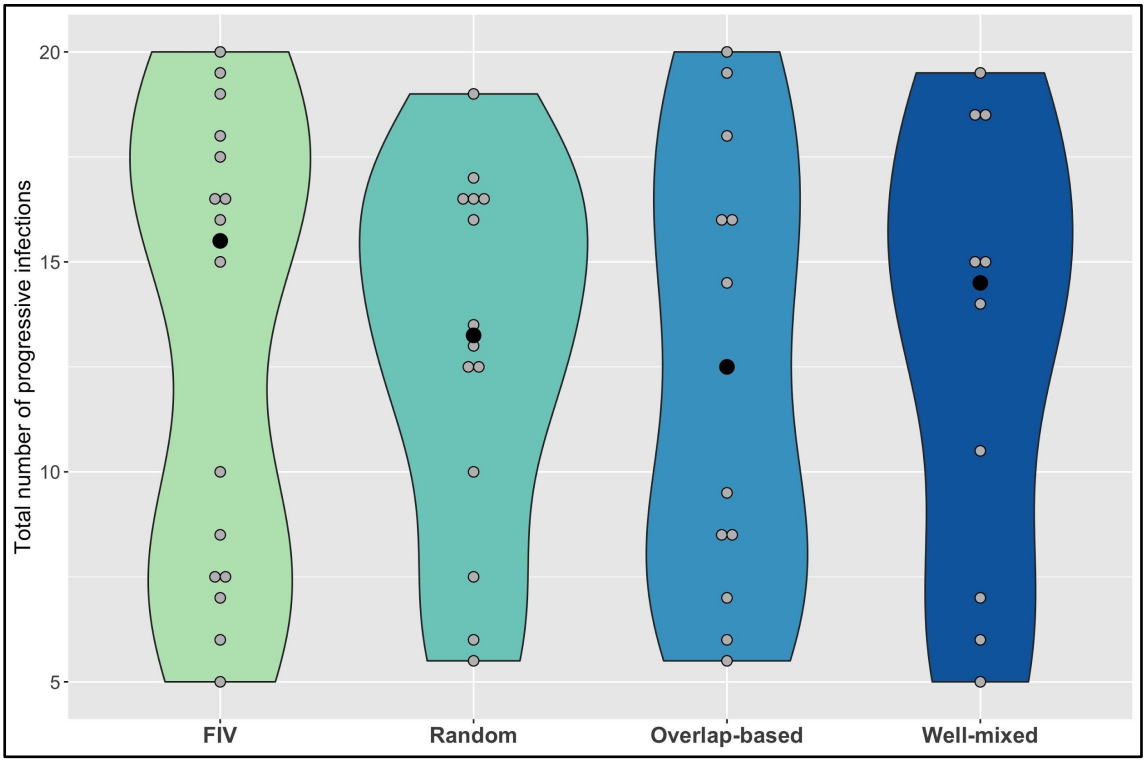


Figure 3.3: FIV-based networks perform at least as well as other models in predicting number of progressive infections, as seen in violin plots of median total number of progressive infections from parameter sets classified as “feasible.” To be feasible, medians needed to fall between 5 and 20 progressive infections, while also having median epidemic duration between 78-117 weeks, and a median of at least 5 abortive infections. Gray points show median values from each feasible parameter set; black points are the median value within each violin plot. Model types are given on the x-axis.

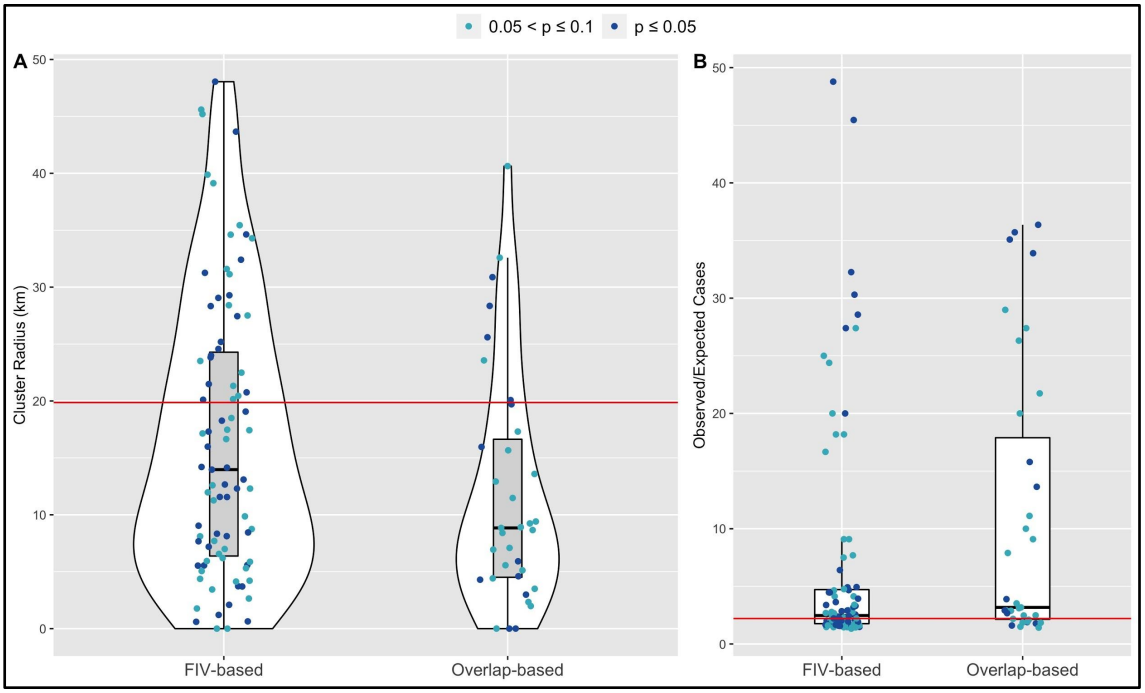


Figure 3.4: SaTScan cluster analysis for feasible FIV-based and overlap-based network simulations show stronger agreement between empirical observations (red horizontal lines) and FIV-based predictions for (A) predicted FeLV cluster size and (B) Observed/Expected FeLV cases associated with the top detected cluster. Shown are feasible simulation results in which at least one cluster was detected with p-value less than or equal to 0.1; further, only the results from the top cluster are shown.

Chapter 4. Paradoxes and synergies: optimizing management of a deadly virus in an endangered carnivore

Marie L.J. Gilbertson, Dave Onorato, Mark Cunningham, Sue VandeWoude, Meggan E. Craft

4.1 Synopsis

Outbreaks of infectious disease can have serious consequences for wildlife population health, especially species of conservation concern. The endangered Florida panther, for example, experienced an outbreak of feline leukemia virus (FeLV) in 2002-2004, and continues to be affected by this deadly virus. Ongoing management efforts aim to mitigate the effects of FeLV on panthers, but with limited information about which strategies may be most effective and efficient. Examples such as this, however, can be excellent opportunities to link management with simulation models in order to optimize conservation efforts for populations of concern.

We used a simulation-based approach to determine optimal FeLV management strategies. We simulated use of proactive FeLV management strategies (i.e., proactive vaccination) and several reactive strategies. Reactive strategies included reactive vaccination (with and without proactive vaccination; random or spatially targeted); test and removal protocols; and temporary spatial segregation of the panther population by blocking freeway wildlife underpasses. Vaccination strategies accounted for partial vaccine immunity, including with single or boosted inoculations. We compared the effectiveness of these different strategies in mitigating FeLV mortalities and the duration of outbreaks.

Results showed that inadequate proactive vaccination can paradoxically increase the number of disease-induced mortalities in FeLV outbreaks, most likely due to effects of partial vaccine immunity. Combinations of proactive vaccination with reactive test-and-removal or vaccination had a synergistic effect in reducing impacts of FeLV outbreaks. Temporary spatial restrictions, however, were unlikely to be effective under realistic conditions.

Synthesis and applications: Management-informed disease simulations demonstrated unexpected negative consequences and synergies with active management strategies for a deadly virus in Florida panthers. We recommend a combination of proactive and reactive management approaches, and suggest prioritizing boosted vaccination over broad distribution of unboosted inoculations. Our results highlight the importance of integrating management and modeling approaches to aid in conservation of at-risk species.

4.2 Introduction

Outbreaks of infectious disease can have significant impacts on the population health of free-ranging wildlife, and are all the more serious in species of conservation concern (Breed et al., 2009). The feline retrovirus, feline leukemia virus (FeLV), for example, has been the source of significant outbreaks in two endangered felids: Iberian lynx (*Lynx pardinus*) and Florida panthers (*Puma concolor coryi*). In the case of panthers, FeLV caused an outbreak in 2002-2004 (Cunningham et al., 2008), spilling over from domestic cats, with subsequent direct transmission among panthers (Brown et al., 2008). In addition, there is recent evidence of ongoing FeLV spillover to and transmission among panthers (Chiu et al., 2019), necessitating continued management

of this deadly pathogen. FeLV inoculation with a domestic cat vaccine has been used previously, but with unknown efficacy in panthers (Cunningham et al., 2008). Further, the proportion of the population that must be vaccinated to most efficiently prevent future FeLV outbreaks is unknown, as is how proactive vaccination might interact with other reactive interventions to interrupt an FeLV outbreak. Such uncertainties are common among free-ranging wildlife systems threatened by infectious disease and hamper efforts to effectively control pathogen transmission.

Mathematical models of infectious disease transmission, however, are a powerful tool for filling such disease management knowledge gaps (McCallum, 2016). Models allow the ethical testing of a wide range of different management approaches and can reveal unexpected consequences of disease control interventions (Lloyd-Smith et al., 2009). Models have been used to optimize disease management protocols in a variety of free-ranging wildlife species of conservation concern, including Ethiopian wolves (Haydon et al., 2006), chimpanzees (Rushmore et al., 2014), and Amur tigers (Gilbert et al., 2020). Further, models can serve the important function of balancing reality and ideal disease control protocols to provide practical, effective guidance for wildlife managers (e.g., Baker et al., 2019; Robinson Stacie J. et al., 2018). Here, we use mathematical models of FeLV transmission in Florida panthers to determine optimal disease prevention and control strategies in this iconic carnivore.

Pathogen management strategies can be preventive (hereafter, *proactive*) or reactive. Among proactive strategies, vaccination is a cornerstone in veterinary medicine (Cleaveland, 2009), and is available for control of FeLV in panthers, but with uncertainties regarding optimal distribution, individual efficacy, and level of population protection needed (Barnett & Civitello, 2020). Reactive strategies include vaccination

after an outbreak has been detected, as in the historical FeLV outbreak in panthers. In addition, FeLV management in domestic cats has relied upon test-and-removal or isolation of infected individuals (Little et al., 2020). This strategy was used to control FeLV in Iberian lynx (López et al., 2009) and is part of future FeLV mitigation plans in panthers. However, the predicted effectiveness of these reactive strategies is unknown, or how they might interact with proactive vaccination efforts to more effectively reduce FeLV impacts.

Larger-scale isolation or quarantine measures may also be used for disease management in free-ranging wildlife, in the form of physical or behavioral barriers between affected and unaffected subsets of a population. For example, fencing has been used to prevent transmission of foot-and-mouth disease between wildlife and cattle in South Africa (Mysterud & Rolandsen, 2019). For panthers, while physical barriers to prevent spillover from domestic cats are impractical, it may be possible to use temporary spatial barriers to reduce the spatial spread of FeLV among panthers after a spillover event. Specifically, the major I-75 freeway is fenced throughout Florida panther habitat to reduce vehicle strikes, with regular wildlife underpasses the main means for wildlife to traverse this barrier. While never before attempted, it may be possible to physically block these underpasses under emergency conditions to prevent the spread of FeLV from the northern to southern subsets of the panther population, or vice versa.

Given the ongoing risks of FeLV to panther population health and conservation and the uncertainties regarding best application of both proactive and reactive FeLV management strategies, the objectives of this study were to test the effectiveness of: (1) proactive vaccination, (2) reactive vaccination, (3) reactive test-and-removal, and (4) reactive temporary spatial restrictions singly, and in combination, for reducing the

population level impacts of FeLV in Florida panthers. We used a mathematical modeling approach to address these objectives in order to efficiently and ethically test a wide range of intervention protocols for the protection of this endangered carnivore.

4.3 Materials and Methods

Simulation pipeline

To examine the effect of different disease management regimes on FeLV control, we used a spatially-explicit, network simulation approach adapted from our models determining drivers of retrovirus transmission in panthers (Chapter 3). This pipeline involves two steps: (1) simulation of a contact network among panthers, and (2) simulation of FeLV transmission on this network. In brief, we simulated panther populations of 150 individuals (McClintock et al., 2015) and used our previously described exponential random graph model for retrovirus transmission in panthers (Chapter 3) to simulate contact networks constrained by density among these populations (see Appendix C methods and Table C.1 for further details).

For modeling FeLV transmission on simulated contact networks, we used the transmission model from Chapter 3. Briefly, this model used a susceptible-infectious-recovered compartmental framework, where individuals could be progressively, regressively, or abortively infected (Cunningham et al., 2008). Progressive infections always resulted in death, while regressive infections eventually recovered with immunity, and abortive infections were always considered immune. Importantly, based on our previous work, we allowed both progressives and regressives to be infectious, though regressives were less likely to transmit (Appendix C, Table C.1). As in Chapter 3, we included only disease-induced mortality, and, in order to preserve key network structure,

allowed territories vacated by deaths to be reoccupied by new susceptible individuals (hereafter, *respawning*). Outbreaks were initiated by a single, randomly selected non-isolate individual in the population (see Appendix C discussion), and proceeded in weekly time steps for up to five years.

Our primary objective was to examine the effect of different FeLV management regimes on epidemic outcomes, so in our primary simulations, we held network generation and transmission parameters constant at previously supported values (see Appendix C methods and Table C.1 for further details). We evaluated the consistency of our results to the choice of parameter values, however, with a sensitivity analysis (see below). Hereafter, a *parameter set* represents the unique set of network, transmission, and management parameters for any given set of simulations. A *full simulation* includes simulation of a single contact network and FeLV transmission on that network (with or without management interventions). For each parameter set, we performed 100 full simulations.

For a *baseline scenario*, we recorded key epidemic outcomes in the absence of interventions; *management scenarios* recorded key outcomes in the presence of interventions (Figure 4.1). These key outcomes were (1) the number of mortalities, (2) the duration of an epidemic, and (3) the proportion of epidemics that failed per 100 successful epidemics. A failed epidemic was one in which fewer than 5 individuals acquired progressive or regressive infections. The outcomes of mortalities and epidemic durations were summarized as median values per parameter set, as results were often skewed; all outcomes were compared between baseline and management scenarios. All simulations were performed in R version 3.6.3 (R Core Team, 2018).

Proactive vaccination

In the first management scenario, we examined the effect of different levels of population proactive vaccination (proportion of the population vaccinated prior to an outbreak), and different ratios of single versus boosted vaccination. We simulated from 10-80% (in 10% increments) of the population having some degree of vaccine-induced immunity to FeLV at the onset of an outbreak. These vaccinations were distributed randomly in the population. Among the vaccinated individuals, 0, 50, or 100% received a second, boosting inoculation. Actual vaccine efficacies are unknown for panthers, but based on efficacy studies in domestic cats (Sparkes, 1997; Torres et al., 2005), we conservatively assumed that boosted vaccination would prevent 80% of infections and single vaccination would prevent 40% of infections. We modeled this efficacy as a binomial probability, given an effective contact between an infectious individual and a vaccinated individual (i.e. vaccination induces partial immunity; Barnett & Civitello, 2020). We conducted proactive vaccination scenarios in a full factorial design, for a total of 24 proactive vaccination parameter sets, and 2,400 full simulations (100 full simulations per parameter set).

Reactive vaccination

During simulated periods of reactive vaccination administration, panthers were selected at a rate of one panther per week for vaccination. We assumed that managers would not know the disease status of an individual selected for vaccination and that vaccination would be ineffective in infectious or recovered individuals. In the case of previously vaccinated individuals, a re-vaccination changed the vaccine efficacy of singly vaccinated individuals (efficacy of 40%) to the efficacy for boosted individuals (efficacy of

80%). We further assumed that managers would know which individuals had received two vaccinations, so boosted individuals were not selected for additional vaccination attempts.

We varied the timing of the onset of reactive vaccination after the initiation of an FeLV outbreak to reflect the difficulty of epidemic detection in this elusive carnivore. We therefore began reactive vaccination at an optimistic, but difficult-to-attain time point of 26 weeks, and a more realistic time point of 52 weeks. In addition, we varied the distribution of reactive vaccination. While proactive vaccines were always distributed randomly, reactive vaccination was either randomly distributed or spatially distributed in an attempted vaccine barrier along the I-75 freeway (see Appendix C for further details).

Because vaccination is resource and time intensive, we evaluated the effect of reactive vaccination for 6 months per year versus year-round. Among reactive vaccination scenarios, we also included a scenario with proactive vaccination, specifically where 0-60% of the population was proactively vaccinated (in 20% increments) at the initiation of an outbreak. Because it is highly unlikely that 100% of the proactively vaccinated population would have received boosted vaccination, we used a more conservative—yet still challenging to attain—ratio of 50% of the proactively vaccinated population being boosted (i.e. efficacy of 80%). We held this ratio of efficacy constant for proactive vaccination in all reactive scenarios. We again used a factorial design across all variations, resulting in 32 parameter sets and a total of 3,200 full simulations under reactive vaccination scenarios.

Reactive test-and-removal

Test-and-removal scenarios were built around a protocol in which panthers infectious at capture are removed from the population through humane euthanasia or temporary removal until recovery. For simplicity, we assumed that all progressively infected individuals would show clinical signs and be humanely euthanized at capture, while infectious regressive individuals were temporarily removed from the population until their recovery and re-release into any open territory. A maximum of five individuals were allowed to be temporarily removed in this way at one time.

We expected that managers would be able to capture and test one panther per week at most, and that captures occurred during a 17 week (about 4 month) capture season, in accordance with current panther capture protocols. We assumed that managers would not know the disease status of a target individual until the capture occurred, so captures were not targeted by infection state. We varied the onset of test-and-removal, such that the intervention began 26 or 52 weeks after the initiation of an epidemic. Captures were also random or spatially targeted. If spatially targeted, captures (and consequent removals) only occurred on the same side of the I-75 freeway as the initial FeLV infection.

As in the reactive vaccination scenarios, we again also included varying degrees of proactive vaccination (0-60% in 20% increments). Reactive test-and-removal scenarios were simulated in a full factorial design across variations, for a total of 16 parameter sets and 1,600 full simulations.

Reactive underpass closures

To determine the potential utility of closing I-75 wildlife underpasses in a FeLV outbreak, we considered a “best case” scenario in which underpasses were closed (and

later reopened) instantaneously, and where closures were completely effective at preventing transmission across the freeway. Underpasses were closed either 26 or 52 weeks after the initiation of an FeLV outbreak, and remained closed for 4, 13, 26, or 52 weeks. We again included variations in proactive vaccination (0-60% in 20% increments), as in other reactive management scenarios. The factorial design here resulted in 32 parameter sets for underpass closure scenarios, for a total of 3,200 full simulations.

Sensitivity analysis

We used a latin hypercube sampling (LHS) approach to generate 50 sensitivity analysis parameter sets across our 8 network and transmission parameters using the *lhs* package in R (Carnell, 2012). We repeated our baseline scenario simulations across these 50 parameter sets; to manage computational effort, we completed 50 simulations per parameter set for all sensitivity analyses.

Due to the high computational effort required to perform sensitivity simulations across all management scenarios, we focused only on the proactive vaccination scenarios. Specifically, we evaluated a subset of 12 LHS parameter sets across a subset of the proactive vaccination conditions in a factorial design (see Appendix C methods for further details), resulting in 108 parameter sets with 50 full simulations per parameter set (5,400 full simulations). This approach allowed us to examine sensitivity of our proactive vaccination results across different outbreak sizes and network and transmission parameters, while mitigating computational effort associated with exploring such a wide range of parameters and scenario variations. Sensitivity analysis simulation results were evaluated using scatterplots and Partial Rank Correlation Coefficients

(PRCC; Marino et al., 2008; Wu et al., 2013); proactive vaccination scenarios were evaluated for alignment with our qualitative result that low proactive vaccination can increase disease-induced mortalities.

4.4 Results

Progressive infections are expected to result in death in panthers and are therefore of key concern for management efforts. For simplicity, hereafter, we refer to the number of progressive infections in simulations as the number of mortalities. In the baseline, no-intervention scenario, there was a median of 34 mortalities (range: 1-54); median duration of epidemics was 119.5 weeks, and 34 epidemics failed (fewer than 5 progressive or regressive infections) per 100 successful epidemics.

Proactive vaccination alone

Proactive vaccination paradoxically increased the number of mortalities across a range of conditions, especially without vaccine boosting (Figure 4.2; Appendix C, Figure C.2). Even with 50% of vaccinates receiving a booster, proactive vaccination only reduced mortalities from the baseline scenario at high levels of population vaccination (i.e., 60-80%). With 100% boosting, proactive vaccination increased mortalities at low levels of population vaccination (10-20%), had marginal effects at 30-40% population vaccination, and was strongly effective at about 50% population vaccination levels and higher (at 80% population vaccination, median 17.5 mortalities versus 34 with no interventions). Proactive vaccination consistently lengthened the duration of epidemics, relative to the baseline scenario (up to a median duration of 143.5 weeks; Appendix C, Figures C.3-4). When all vaccinated individuals received a booster, vaccination reduced

the probability of a successful outbreak even at 40% population vaccination (52 failures versus 34 failures per 100 successful epidemics; Appendix C, Figure C.13). In contrast, when no vaccinated individuals received a booster, proactive vaccination was largely only effective at reducing the probability of epidemics at very high levels of population vaccination.

Proactive and reactive vaccination

Reactive vaccination alone did not reduce mortalities (Figure 4.3). However, reactive vaccination appeared to work synergistically with proactive vaccination, particularly at moderate to high levels of proactive vaccination (i.e., at least 40-60% of the population proactively vaccinated). This largely held true regardless of the timing of intervention onset and the strategy for reactive vaccination distribution (i.e., random versus spatial; Appendix C, Figures C.5-6). The mortality-reducing effects of reactive vaccination were, however, slightly reduced if reactive vaccination occurred for only 6 months out of the year (Figure 4.3). A ratio of greater than 1.5 inoculations per vaccinated individual appeared to promote the largest reductions in mortalities (e.g., as few as a median of 25 mortalties with year-round reactive vaccination and 60% proactive vaccination versus 34 mortalities with no interventions; Figure 4.3). Adding reactive vaccination largely did not affect the durations of simulated epidemics (Appendix C, Figures C.7-8), and had little impact on the probability of epidemics failing, particularly in comparison to proactive vaccination alone (Appendix C, Figure C.13).

Proactive vaccination with test-and-removal

Test-and-removal alone did not reduce mortalities (Figure 4.4). Like reactive vaccination, however, test-and-removal appeared to work synergistically with proactive vaccination, especially at moderate to high levels of proactive vaccination (i.e., at least 40-60% of the population proactively vaccinated; as few as a median of 26 mortalities). Again, this largely held true regardless of the timing of the onset of the intervention or the targeting of captures (i.e., random versus spatial). Notably, simulated captures were only conducted for about 4 months per simulation year, in contrast to at least 6 months of reactive vaccination per year. Captures were marginally more likely to successfully identify actively infectious individuals when initiated earlier in an outbreak (Appendix C, Figure C.9). The addition of test-and-removal largely did not affect the durations of epidemics (Appendix C, Figure C.10). When coupled with proactive vaccination, test-and-removal had a modest effect in reducing the probability of a successful epidemic (e.g., maximum of 50 failed epidemics per 100 successful versus 34 failed epidemics per 100 successes with no interventions; Appendix C, Figure C.13).

Reactive underpass closures

Reactive underpass closures were ineffective in the absence of proactive vaccination (Appendix C, Figure C.11). The most notable impact of underpass closures on reducing FeLV mortalities occurred when onset of closure was early (26 weeks after epidemic initiation), lasted for at least 13 weeks, and occurred in conjunction with at least 40-60% of the population being proactively vaccinated (though the clearest effects occurred when at least 60% were proactively vaccinated). Under these conditions, underpass closures synergistically reduced mortalities from baseline scenarios (as few as a median of 22 mortalities with underpass closures) and increased the probability that

an epidemic would fail (maximum of 75 failed epidemics per 100 successes with underpass closures; Appendix C, Figure C.13). In some cases, simulations showed a marked decrease in transmission during the period of underpass closures, but reopening often resulted in a subsequent resurgence of infections (Figure 4.5). Underpass closures had no clear impact on the duration of outbreaks (Appendix C, Figure C.12).

Sensitivity analyses

Simulated epidemic sizes were variable across the full 50 sensitivity analysis parameter sets in the absence of FeLV interventions (range: median 7.5 mortalities-median 47.5 mortalities; Appendix C, Figure C.14). Based on PRCC, the parameters for network density, transmission potential from regressives, weekly contact rates, and baseline transmission potential were positively associated with median mortalities in the absence of interventions; the parameter for the infection-induced mortality rate was negatively associated with median mortalities (see Appendix C results, Figure C.15).

When focusing on a subset of parameter sets for proactive vaccination sensitivity analysis, low levels of proactive vaccination (e.g., 20% population proactive vaccination) were sometimes effective in reducing the number of mortalities, in contrast to our primary results. PRCC results from proactive scenario sensitivity analysis suggested that the parameters for network density, transmission potential from regressives, and weekly contact rates were positively correlated with increased mortalities at low levels of vaccination (Appendix C, Figure C.16). However, network density did not have a clearly monotonic relationship with the difference between mortalities with and without proactive vaccination (Appendix C, Figure C.17). We therefore performed additional *post-hoc* sensitivity analyses to further interrogate the relationship between network density and

our qualitative outcome of increased mortalities at low levels of proactive vaccination (see Appendix C results for detailed discussion of *post-hoc* analysis). These additional analyses found some evidence that low levels of proactive vaccination were least effective at reducing mortalities at intermediate values of parameters governing network connectivity (network density and proportion adults), especially when coupled with increased transmission potential (e.g. higher infectiousness of regressive individuals and/or increased weekly contact rates; Appendix C, Figures C.19-20).

4.5 Discussion

In this study we found unexpected consequences and impacts of several epidemic management strategies for disease control in small populations of conservation concern. Although our simulation results provide guidelines for FeLV management in Florida panthers, they also demonstrate the power of partnering modeling approaches and population management questions to test and optimize disease control strategies in free-ranging wildlife (Joseph et al., 2013). Furthermore, the principles of transmission and available methods of disease control underlying our findings provide insights for pathogen control in other host-pathogen systems.

Proactive vaccination alone may worsen epidemic outcomes under some conditions

Our simulation results showed a paradoxical increase in FeLV mortalities with low levels of proactive vaccination. This counterintuitive finding is likely due, at least in part, to partial vaccine immunity, a type of vaccine imperfection often overlooked in studies of wildlife disease (Barnett & Civitello, 2020). Under partial vaccine immunity, vaccinates can act as a semi-protected susceptible pool, contracting infection later in the

course of an epidemic, ultimately prolonging epidemics and increasing the total number of mortalities. Only with adequate herd immunity can these effects be avoided. This counterintuitive result is consistent with findings by Rees et al. (2013), who found reduced vaccine effectiveness with heterogeneity in the spatial distribution of hosts. Importantly, our results suggest that studies assuming full protection from vaccination may significantly underestimate the level of vaccination needed for population protection—underestimates which may even make epidemics worse if only partial vaccine immunity can be achieved.

Our sensitivity analysis suggests that populations with intermediate levels of network connectivity and/or high transmission potential may be most vulnerable to these paradoxical effects. As in spatially structured populations (McCallum, 2008), high or low connectivity causes infection to fade out quickly; with intermediate connectivity, vaccine failures provide a steady supply of new susceptibles. Alternatively, in cases of high transmission potential, vaccination may shift a rapid fade-out epidemic to a sustained epidemic scenario, as in Rees et al. (2013). While our sensitivity analysis was limited by computational complexity and use of some discrete parameters (which may affect PRCC inference; Marino et al., 2008), our findings are consistent with this broader body of literature.

FeLV vaccine efficacy may operate differently in reality from our simulation structure: for example, some vaccinates may have zero vaccine-induced immunity, while others have 100% protection (binary immunity). Alternatively, vaccination may not protect from infection but could reduce viral shedding or increase survival of infected individuals (Barnett & Civitello, 2020). In the case of binary immunity, vaccine efficacy would be unlikely to prolong and worsen epidemics as we saw here. In contrast,

increased survival of infected individuals without changes to shedding potential could extend or worsen outbreaks, and even favor the evolution of virulence (Barnett & Civitello, 2020).

Based on our results, we argue that wildlife managers should continue vaccinating available panthers and prioritize boosting vaccinated individuals to develop a core population of high-immunity individuals, rather than a broadly distributed low-immunity population. This recommendation should be most effective at increasing the probability of epidemic failure, and aligns with other wildlife studies which have emphasized vaccination of core populations (Vial et al., 2006) or risk-based sub-populations (Beyer et al., 2012).

Temporary spatial restrictions are unlikely to be effective under realistic scenarios

Here, we examined the effect of wildlife underpass closures as a novel method to restrict connectivity of the panther population under emergency disease control conditions. Such temporary spatial restrictions could increase other types of panther mortality (e.g. vehicle strikes or intraspecific conflict), so such an intervention would need to reduce FeLV mortalities at a rate greater than these other types of panther mortality in order to be a viable disease control strategy. Unfortunately, our simulations found that temporary underpass closures were generally no more effective at reducing FeLV mortalities than less risky reactive interventions. Further, underpass closures were less effective when occurring after the peak of simulated epidemics, such that spatial restrictions would need to occur early in an epidemic to be effective. However, it should be noted that spatial restrictions may be more effective if used in combination with other reactive FeLV management strategies (e.g., reactive vaccination in concert with

temporary spatial restrictions to prevent transmission resurgence after restrictions are removed).

Extensive research has considered how landscape barriers and fragmentation may affect pathogen transmission and control in wildlife (e.g., Rees et al., 2013; Smith et al., 2002; Tracey et al., 2014; White et al., 2018b). By closing underpasses, we artificially fragmented panther habitat, assuming complete efficacy of closures in preventing transmission across the I-75 freeway. In reality, some individuals would likely successfully traverse this barrier. Because habitat fragmentation can promote pathogen outbreaks and persistence (White et al., 2018b), a "semipermeable" freeway barrier may, in fact, ultimately worsen epidemic dynamics. Alternatively, landscape restrictions to animal movement can facilitate more efficient disease control, as in Haydon *et al* (2006), where vaccinating to control pathogen spread along habitat corridors reduced the level of vaccine coverage needed in endangered Ethiopian wolves. While panthers are generally well-connected spatially, expansion of the population north of its current range may favor similar metapopulation-oriented disease control strategies. However, given uncertainties in the effect of underpass closures on other sources of mortalities and their limited effectiveness in our simulations, we suggest temporary underpass closures should currently be considered a low priority for FeLV management in panthers.

Reactive and proactive strategies can work synergistically to reduce epidemic impacts

Our simulations showed that both reactive vaccination and test-and-removal strategies reduced FeLV mortalities in panthers when used in combination with moderate levels of proactive vaccination. Test-and-removal had more consistent effects with arguably less effort than reactive vaccination (4 months of captures compared to

year-round reactive vaccination). We therefore suggest that test-and-removal be prioritized over reactive vaccination, especially if identification of actively infectious individuals can be improved.

In our simulations, captures that were most aligned with the initial wave of infectious individuals (i.e., with earlier onset) were more likely to successfully identify actively infectious individuals for removal. This finding highlights the importance of targeting captures to individuals likely to be actively infectious. Determining infection status in cryptic wildlife is difficult, however, and consequently supports the increased use of remote tracking technologies that may be able to (1) identify behavior changes associated with sickness, and (2) detect the onset of an epidemic more quickly. This conclusion is consistent with similar findings in Channel Island foxes, where increasing the number and frequency of tracking of sentinel individuals was important for early identification of epidemics (Sanchez & Hudgens, 2020).

A key component of the success of test-and-removal here is the selective removal we simulated, which avoids removing immune individuals that contribute to overall herd immunity (Miguel et al., 2020; Potapov et al., 2012). However, we have simplified the field testing process in our simulations. The common field-available FeLV diagnostic test identifies antigenemia, which is key for identifying actively infectious individuals. However, the duration of antigenemia—and even degree of infectiousness—in regressively infected individuals is unclear in panthers. We may therefore overestimate the effect of removing regressive individuals, but given their reduced infectiousness in our simulations, we still expect test-and-removal to be a key strategy for mitigating FeLV impacts in panthers.

Notably, reactive vaccination, in concert with proactive vaccination, is also a viable alternative strategy to test-and-removal. Our simulations showed reactive vaccination to have the strongest effects for reducing FeLV mortalities when at least 50% of the population was vaccinated, and with a ratio of about 1.5 vaccines per vaccinated individual (i.e., 50% of vaccinated individuals received a booster). We therefore suggest that managers should prioritize boosting at least half of vaccinated individuals in a reactive vaccination response scenario. We did not see a strong effect of attempting a vaccine barrier, in contrast to Sanchez *et al* (2020), where a simulated vaccine barrier could effectively halt spread of a pathogen in Channel Island foxes. This difference is likely due to the differences in home range size and movement capacity between the two species, with foxes ranging far less widely than panthers. In reality, if an empirical outbreak of FeLV in panthers exhibited a stronger spatial signal than was featured in our simulations, spatially targeted reactive vaccination may yet be a worthwhile intervention strategy.

Importantly—particularly if levels of population protection from proactive vaccination are unknown—both reactive vaccination and test-and-removal strategies mitigated the negative effects of low levels of population protection seen with inadequate proactive vaccination. It is therefore vital for managers to incorporate these reactive strategies in the event of future FeLV outbreaks.

Limitations and future directions

In this study, we considered the effects of partial vaccine immunity, but we made the simplifying assumption of no waning vaccine or infection-induced immunity over time. However, our findings with regard to imperfect efficacy are at least partially

representative of the likely consequences of waning immunity, in that the loss of immunity supplies new susceptibles to the population. Should immunity not outlast the course of an FeLV outbreak, this process could prolong outbreaks and result in increased mortalities. Future research should therefore examine the effects of waning vaccine immunity, particularly considering the value of revaccinating individuals which may be experiencing loss of vaccine protection.

Our simulation results found that relatively high levels of vaccination were required to reduce impacts of FeLV in panthers, in contrast to studies in other endangered species (Gilbert et al., 2020; Haydon et al., 2006). However, here we are investigating a pathogen with a long duration of infectiousness and lack the advantages of distinct corridors between panther sub-populations for reducing required levels of vaccination. It is therefore unsurprising that panthers would require higher levels of FeLV population vaccination than was found, for example, for rabies vaccination in Ethiopian wolves (Haydon et al., 2006) or canine distemper virus vaccination in Amur tigers (Gilbert et al., 2020). However, we also did not assume the presence of preexisting population immunity prior to proactive vaccination, and some degree of population immunity likely already exists in panthers, given ongoing exposures (Chiu et al., 2019). This would reduce necessary vaccination levels in panthers, as would higher vaccine efficacy than we conservatively assumed here (Vial et al., 2006). Future research could prioritize better understanding the realized individual immunity after single and boosted FeLV vaccination in panthers to determine if inoculations are more protective than we conservatively assumed here.

Conclusions

Our simulation results highlight the risks of inadequate proactive vaccination, particularly with partial vaccine immunity. We recommend prioritizing boosted vaccination in panthers, and joint use of proactive vaccination and reactive strategies to mitigate the risks of imperfect vaccination and most effectively reduce the impacts of FeLV in this iconic carnivore. This research highlights the value of linking modeling and management priorities to identify unexpected consequences of interventions and determine optimal pathogen management strategies in free-ranging wildlife.

4.6 Acknowledgements

This research was supported by the National Science Foundation (DEB-1413925, 1654609, and 2030509). MLJG was supported by the Office of the Director, National Institutes of Health (NIH T32OD010993), the University of Minnesota Informatics Institute MnDRIVE program, and the Van Sloun Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Florida panther data collected by the Florida Fish and Wildlife Conservation Commission is fully supported by donations to the Florida Panther Research and Management Trust Fund via the registration of “Protect the Panther” license plates. We acknowledge the efforts of National Park Service staff in the collection of Florida panther data utilized in this study.

4.7 Figures

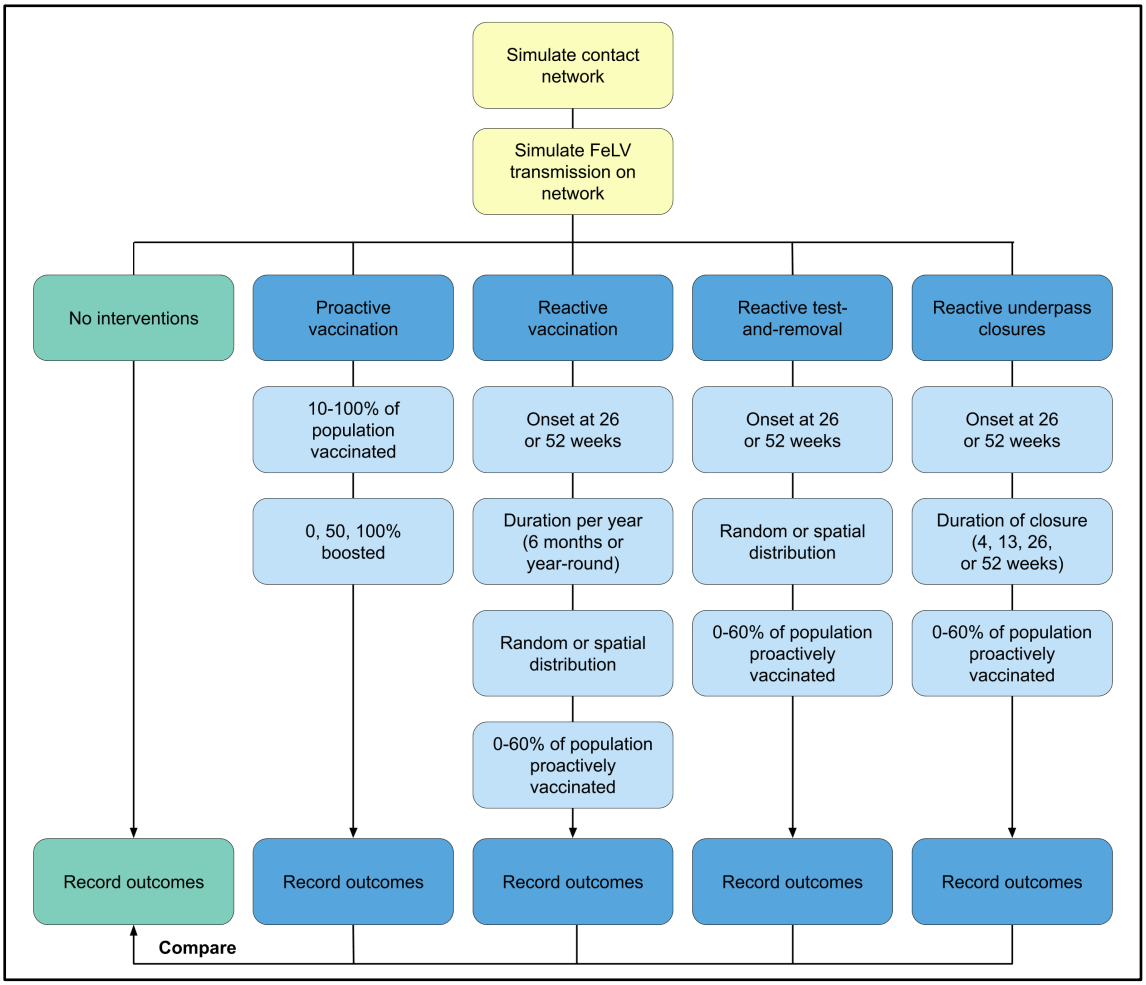


Figure 4.1: Steps of the simulation process across baseline and management scenarios. The basic network and FeLV transmission simulation steps are shown in yellow, the baseline (no intervention) scenario in green, and the four different overarching management scenarios in blue. Within each management scenario, we investigated several variations for the given approach (shown in light blue) in a factorial design.

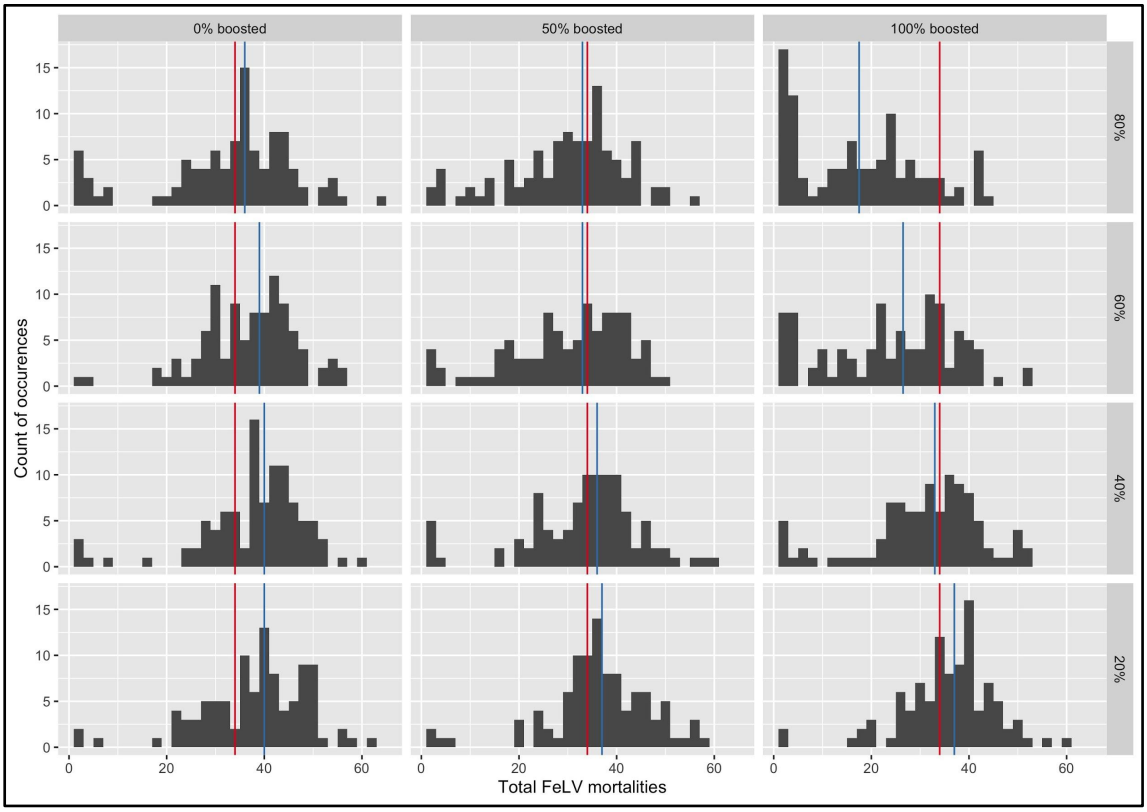


Figure 4.2: Histograms of FeLV mortalities in simulated epidemics with proactive vaccination alone. Red vertical lines indicate the median number of mortalities from the baseline scenario without interventions; blue lines, scenarios with proactive vaccination. Panel rows represent the proportion of the population proactively vaccinated; columns represent ratios of vaccine efficacy among vaccinates. Each histogram plot represents the results of 100 simulations. An expanded version of this figure is available in Appendix C (Figure C.2).

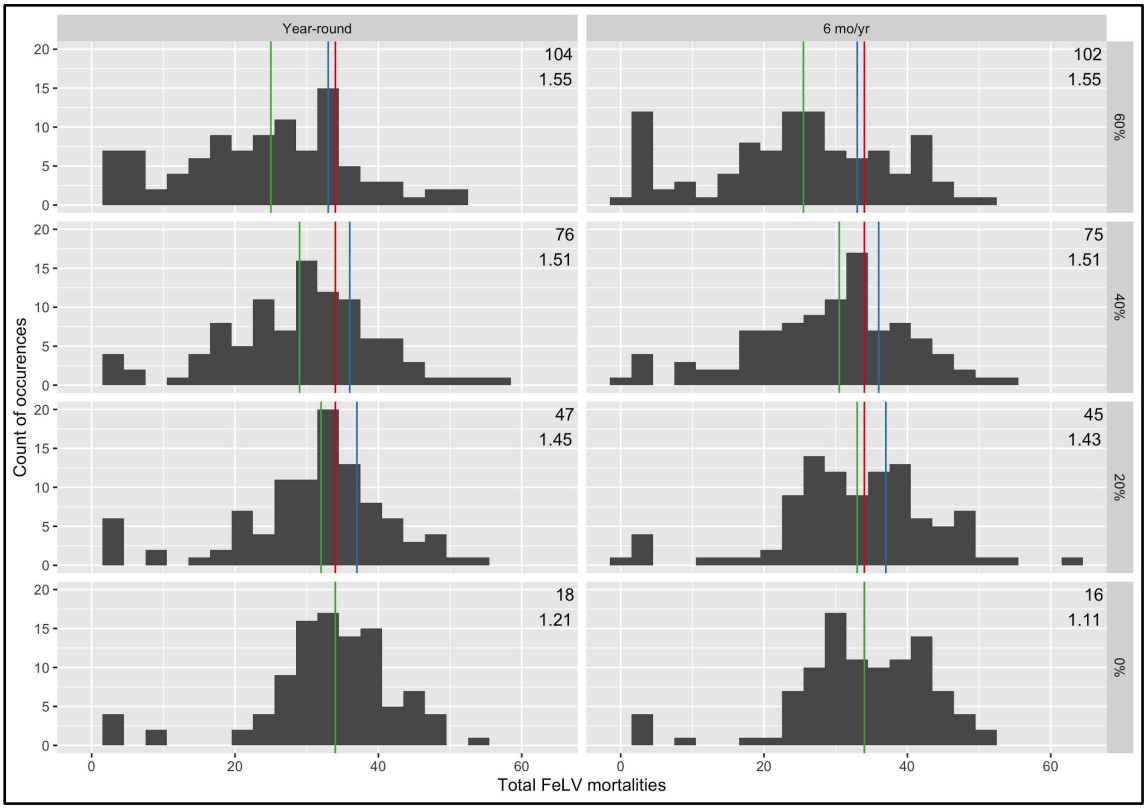


Figure 4.3: Histograms of FeLV mortalities in simulated epidemics with both reactive and proactive vaccination. Results shown are for randomly distributed reactive vaccination starting 52 weeks after epidemic onset. Panel rows represent the proportion of the population proactively vaccinated (with 50% boosted vaccination); columns represent duration of reactive vaccination per year. Red vertical lines indicate the median number of mortalities from simulations without interventions; blue lines for proactive vaccination alone; green lines for the given combination of reactive and proactive vaccination. Numbers in the upper right of each histogram represent (above) median total number of vaccinated individuals (both proactive and reactive), and (below) median value for the total vaccinations used per total vaccinated individuals.

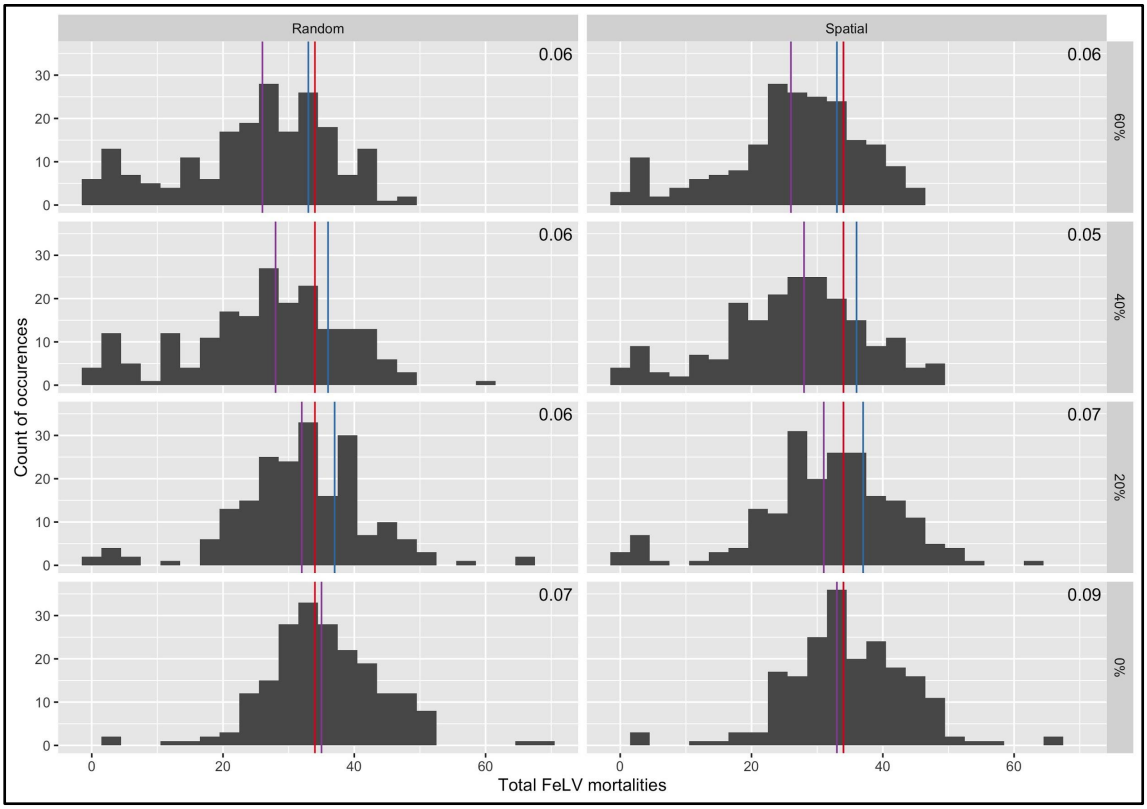


Figure 4.4: Histograms of FeLV mortalities in simulated epidemics with both reactive test-and-removal and proactive vaccination. Panel rows represent the proportion of the population proactively vaccinated (with 50% boosted vaccination); columns represent the capture and testing targeting strategy (random versus spatially targeted). Vertical red lines indicate the median number of mortalities from simulations without interventions; blue lines for proactive vaccination alone; purple lines for the given combination of test and removal and proactive vaccination. Numbers in the upper right of each histogram represent the median value for the proportion of capture events that resulted in a removal or humane euthanasia.

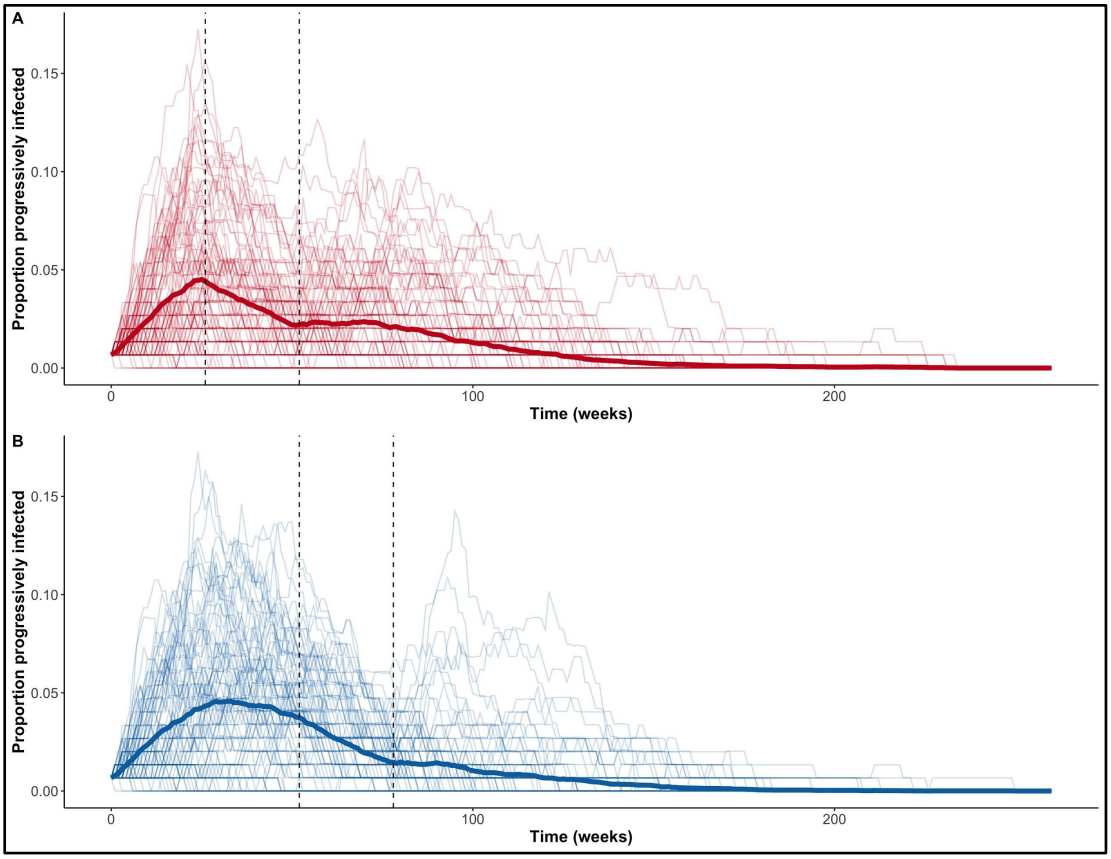


Figure 4.5: Epidemic curves showing the proportion of the population progressively infected over time when 60% of the population was proactively vaccinated. Lighter lines show individual simulation results; dark lines show mean values across all simulations. Vertical black dashed lines show when underpasses were closed (left) and reopened (right). Panel A shows underpasses closed for 26 weeks starting 26 weeks after an epidemic started; panel B shows underpasses closed for 26 weeks starting at 52 weeks.

Conclusion

In this dissertation, I have tested novel and developing approaches for understanding the drivers of pathogen transmission, predominantly focused on the naturally occurring model system of feline retroviruses in the Florida panther. In Chapter 1, I reviewed key opportunities for integrating network and genomic approaches for improved understanding of animal behavior and the underlying dynamics of pathogen transmission. In Chapter 2, I identified best practices for using telemetry data in network studies of wildlife infectious disease, which can provide new opportunities and insights for remote study of transmission processes in difficult-to-observe species. Chapter 3 first determined drivers of transmission of an apathogenic virus in the Florida panther, then determined that those drivers could predict transmission of an analogously transmitted pathogen. These results have major implications for the proactive development of infectious disease prevention and control plans, and may be translatable across wildlife, domestic animal, and human systems. Finally, in Chapter 4, I determined optimal pathogen management strategies in the endangered Florida panther, including identifying surprising paradoxes and synergies across different management interventions. The findings of this chapter will have significant consequences for ongoing conservation efforts of the Florida panther, but also highlight key gaps and opportunities in model-based testing of pathogen control interventions.

Together, these individual chapters have resulted in new knowledge about and methods for identifying underlying drivers of pathogen transmission, a key step for improving transmission forecasting and management. Further, my work has demonstrated strategies for linking diverse data streams and disciplines to expand

understanding of disease ecology across systems. This research therefore represents a valuable contribution to the literature that can reduce impacts of infectious disease in humans and animals, for the benefit of public, animal, and ecosystem health.

Bibliography

Anderson, R. M., & May, R. M. (1991). *Infectious Diseases of Humans: Dynamics and Control*. OUP Oxford.

Antonovics, J. (2017). Transmission dynamics: critical questions and challenges. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1719). https://doi.org/10.1098/rstb.2016.0087

Archie, E. A., Luikart, G., & Ezenwa, V. O. (2009). Infecting epidemiology with genetics: a new frontier in disease ecology. *Trends in Ecology & Evolution*, *24*(1), 21–30.

Archie, E. A., & Tung, J. (2015). Social behavior and the microbiome. *Current Opinion in Behavioral Sciences*, *6*, 28–34.

Baele, G., Suchard, M. A., Rambaut, A., & Lemey, P. (2017). Emerging Concepts of Data Integration in Pathogen Phylodynamics. *Systematic Biology*, *66*(1), e47–e65.

Baker, L., Matthiopoulos, J., Müller, T., Freuling, C., & Hampson, K. (2019). Optimizing spatial and seasonal deployment of vaccination campaigns to eliminate wildlife rabies. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *374*(1776), 20180280.

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*(5), 478–489.

Barnett, K. M., & Civitello, D. J. (2020). Ecological and Evolutionary Challenges for Wildlife Vaccination. *Trends in Parasitology*, *36*(12), 970–978.

Barton, K. A., Phillips, B. L., Morales, J. M., & Travis, J. M. J. (2009). The evolution of an “intelligent” dispersal strategy: biased, correlated random walks in patchy

landscapes. *Oikos*, *118*(2), 309–319.

Becker, D. J., & Hall, R. J. (2014). Too much of a good thing: resource provisioning alters infectious disease dynamics in wildlife. *Biology Letters*, *10*(7). https://doi.org/10.1098/rsbl.2014.0309

Becker, D. J., Streicker, D. G., & Altizer, S. (2015). Linking anthropogenic resources to wildlife-pathogen dynamics: a review and meta-analysis. *Ecology Letters*, *18*(5), 483–495.

Begon, M., Bennett, M., Bowers, R. G., French, N. P., Hazel, S. M., & Turner, J. (2002). A clarification of transmission terms in host-microparasite models: numbers, densities and areas. *Epidemiology and Infection*, *129*(1), 147–153.

Beyer, H. L., Hampson, K., Lembo, T., Cleaveland, S., Kaare, M., & Haydon, D. T. (2012). The implications of metapopulation dynamics on the design of vaccination campaigns. *Vaccine*, *30*(6), 1014–1022.

Biek, R., Drummond, A. J., & Poss, M. (2006). A virus reveals population structure and recent demographic history of its carnivore host. *Science*, *311*(5760), 538–541.

Biek, R., Henderson, J. C., Waller, L. A., Rupprecht, C. E., & Real, L. A. (2007). A high-resolution genetic signature of demographic and spatial expansion in epizootic rabies virus. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(19), 7993–7998.

Bird, B. H., Khristova, M. L., Rollin, P. E., Ksiazek, T. G., & Nichol, S. T. (2007). Complete genome analysis of 33 ecologically and biologically diverse Rift Valley fever virus strains reveals widespread virus movement and low genetic diversity due to recent common ancestry. *Journal of Virology*, *81*(6), 2805–2816.

Bivand, R., & Rundel, C. (2018). *rgeos: Interface to Geometry Engine - Open Source*

('GEOS'). https://CRAN.R-project.org/package=rgeos

Blake, A., Sinclair, M. T., & Sugiyarto, G. (2003). Quantifying the Impact of Foot and Mouth Disease on Tourism and the UK Economy. *Tourism Economics*, *9*(4), 449–465.

Blanchong, J. A., Samuel, M. D., Scribner, K. T., Weckworth, B. V., Langenberg, J. A., & Filcek, K. B. (2007). Landscape genetics and the spatial distribution of chronic wasting disease. *Biology Letters*, *4*(1), 130–133.

Blasse, A., Calvignac-Spencer, S., Merkel, K., Goffe, A. S., Boesch, C., Mundry, R., & Leendertz, F. H. (2013). Mother-offspring transmission and age-dependent accumulation of simian foamy virus in wild chimpanzees. *Journal of Virology*, *87*(9), 5193–5204.

Blumberg, S., & Lloyd-Smith, J. O. (2013). Inference of R(0) and transmission heterogeneity from the size distribution of stuttering chains. *PLoS Computational Biology*, *9*(5), e1002993.

Blyton, M. D. J., Banks, S. C., Peakall, R., & Gordon, D. M. (2013). High temporal variability in commensal *Escherichia coli* strain communities of a herbivorous marsupial. *Environmental Microbiology*, *15*(8), 2162–2172.

Blyton, M. D. J., Banks, S. C., Peakall, R., Lindenmayer, D. B., & Gordon, D. M. (2014). Not all types of host contacts are equal when it comes to *E. coli* transmission. *Ecology Letters*, *17*(8), 970–978.

Brandell, E. E., Fountain-Jones, N. M., Gilbertson, M. L. J., Cross, P. C., Hudson, P. J., Smith, D. W., Stahler, D. R., Packer, C., & Craft, M. E. (2020). Group density, disease, and season shape territory size and overlap of social carnivores. *The Journal of Animal Ecology*. https://doi.org/10.1111/1365-2656.13294

Breed, A. C., Plowright, R. K., Hayman, D. T. S., Knobel, D. L., Molenaar, F. M., Gardner–Roberts, D., Cleaveland, S., Haydon, D. T., Kock, R. A., Cunningham, A. A., Sainsbury, A. W., & Delahay, R. J. (2009). Disease Management in Endangered Mammals. In R. J. Delahay, G. C. Smith, & M. R. Hutchings (Eds.), *Management of Disease in Wild Mammals* (pp. 215–239). Springer Japan.

Brown, M. A., Cunningham, M. W., Roca, A. L., Troyer, J. L., Johnson, W. E., & O'Brien, S. J. (2008). Genetic characterization of feline leukemia virus from Florida panthers. *Emerging Infectious Diseases*, *14*(2), 252–259.

Buhnerkempe, M. G., Roberts, M. G., Dobson, A. P., Heesterbeek, H., Hudson, P. J., & Lloyd-Smith, J. O. (2015). Eight challenges in modelling disease ecology in multi-host, multi-agent systems. *Epidemics*, *10*, 26–30.

Bull, C. M., Godfrey, S. S., & Gordon, D. M. (2012). Social networks and the spread of *Salmonella* in a sleepy lizard population. *Molecular Ecology*, *21*(17), 4386–4392.

Burgin, L. E., Gloster, J., Sanders, C., Mellor, P. S., Gubbins, S., & Carpenter, S. (2013). Investigating incursions of bluetongue virus using a model of long-distance *Culicoides* biting midge dispersal. *Transboundary and Emerging Diseases*, *60*(3), 263–272.

Calenge, C. (2006). The package adehabitat for the R software: tool for the analysis of space and habitat use by animals. In *Ecological Modelling* (Vol. 197, p. 1035).

Carnegie, N. B. (2017). Effects of contact network structure on epidemic transmission trees: implications for data required to estimate network structure. *Statistics in Medicine*. https://doi.org/10.1002/sim.7259

Carnell, R. (2012). lhs: Latin hypercube samples. *R Package Version 0. 10, URL http://CRAN. R-Project. Org/package= Lhs*.

Cauchemez, S., Bhattarai, A., Marchbanks, T. L., Fagan, R. P., Ostroff, S., Ferguson, N. M., Swerdlow, D., & Pennsylvania H1N1 working group. (2011). Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 2825–2830.

Chamie, G., Wandera, B., Marquez, C., Kato-Maeda, M., Kamya, M. R., Havlir, D. V., & Charlebois, E. D. (2015). Identifying locations of recent TB transmission in rural Uganda: a multidisciplinary approach. *Tropical Medicine & International Health: TM & IH*, *20*(4), 537–545.

Chen, S., White, B. J., Sanderson, M. W., Amrine, D. E., Ilany, A., & Lanzas, C. (2014). Highly dynamic animal contact network and implications on disease transmission. *Scientific Reports*, *4*, 4472.

Chiu, E. S., Kraberger, S., Cunningham, M., Cusack, L., Roelke, M., & VandeWoude, S. (2019). Multiple Introductions of Domestic Cat Feline Leukemia Virus in Endangered Florida Panthers. *Emerging Infectious Diseases*, *25*(1), 92–101.

Chiyo, P. I., Grieneisen, L. E., Wittemyer, G., Moss, C. J., Lee, P. C., Douglas-Hamilton, I., & Archie, E. A. (2014). The influence of social structure, habitat, and host traits on the transmission of *Escherichia coli* in wild elephants. *PloS One*, *9*(4), e93408.

Cleaveland, S. (2009). Viral threats and vaccination: disease management of endangered species. *Animal Conservation*, *12*(3), 187–189.

Colijn, C., & Gardy, J. (2014). Phylogenetic tree shapes resolve disease transmission patterns. *Evolution, Medicine, and Public Health*, *2014*(1), 96–108.

Costenbader, E., & Valente, T. W. (2003). The stability of centrality measures when

networks are sampled. *Social Networks*, *25*(4), 283–307.

Cottam, E. M., Thébaud, G., Wadsworth, J., Gloster, J., Mansley, L., Paton, D. J., King, D. P., & Haydon, D. T. (2008). Integrating genetic and epidemiological data to determine transmission pathways of foot-and-mouth disease virus. *Proceedings. Biological Sciences / The Royal Society*, *275*(1637), 887–895.

Craft, M. E. (2015). Infectious disease transmission and contact networks in wildlife and livestock. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1669). https://doi.org/10.1098/rstb.2014.0107

Craft, M. E., Volz, E., Packer, C., & Meyers, L. A. (2009). Distinguishing epidemic waves from disease spillover in a wildlife population. *Proceedings. Biological Sciences / The Royal Society*, *276*(1663), 1777–1785.

Craft, M. E., & Caillaud, D. (2011). Network models: an underutilized tool in wildlife epidemiology? *Interdisciplinary Perspectives on Infectious Diseases*, *2011*, 676949.

Craft, M. E., Volz, E., Packer, C., & Meyers, L. A. (2011). Disease transmission in territorial populations: the small-world network of Serengeti lions. *Journal of the Royal Society, Interface / the Royal Society*, *8*(59), 776–786.

Croft, D. P., James, R., & Krause, J. (2008). *Exploring Animal Social Networks*. Princeton University Press.

Cross, P. C., Creech, T. G., Ebinger, M. R., Heisey, D. M., Irvine, K. M., & Creel, S. (2012). Wildlife contact analysis: emerging methods, questions, and challenges. *Behavioral Ecology and Sociobiology*, *66*(10), 1437–1447.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. In *InterJournal: Vol. Complex Systems* (p. 1695). http://igraph.org

Cullingham, C. I., Kyle, C. J., Pond, B. A., Rees, E. E., & White, B. N. (2009). Differential permeability of rivers to raccoon gene flow corresponds to rabies incidence in Ontario, Canada. *Molecular Ecology*, *18*(1), 43–53.

Cunningham, M. W., Brown, M. A., Shindle, D. B., Terrell, S. P., Hayes, K. A., Ferree, B. C., McBride, R. T., Blankenship, E. L., Jansen, D., Citino, S. B., Roelke, M. E., Kiltie, R. A., Troyer, J. L., & O'Brien, S. J. (2008). Epizootiology and management of feline leukemia virus in the Florida puma. *Journal of Wildlife Diseases*, *44*(3), 537–552.

Cuzick, J., & Edwards, R. (1990). Spatial clustering for inhomogeneous populations. *Journal of the Royal Statistical Society*, *52*(1), 73–96.

Davis, G. H., Crofoot, M. C., & Farine, D. R. (2018). Estimating the robustness and uncertainty of animal social networks using different observational methods. *Animal Behaviour*, *141*, 29–44.

de Azevedo, F. C. C., & Murray, D. L. (2007). Spatial organization and food habits of jaguars (*Panthera onca*) in a floodplain forest. *Biological Conservation*, *137*(3), 391–402.

de Carvalho Ferreira, H. C., Weesendorp, E., Quak, S., Stegeman, J. A., & Loeffen, W. L. A. (2014). Suitability of faeces and tissue samples as a basis for non-invasive sampling for African swine fever in wild boar. *Veterinary Microbiology*, *172*(3-4), 449–454.

Degnan, P. H., Pusey, A. E., Lonsdorf, E. V., Goodall, J., Wroblewski, E. E., Wilson, M. L., Rudicell, R. S., Hahn, B. H., & Ochman, H. (2012). Factors associated with the diversification of the gut microbial communities within chimpanzees from Gombe National Park. *Proceedings of the National Academy of Sciences of the*

United States of America, *109*(32), 13034–13039.

Delignette-Muller, M. L., Dutang, C., & Others. (2015). fitdistrplus: An R package for fitting distributions. *Journal of Statistical Software*, *64*(4), 1–34.

De Maio, N., Wu, C.-H., O'Reilly, K. M., & Wilson, D. (2015). New Routes to Phylogeography: A Bayesian Structured Coalescent Approximation. *PLoS Genetics*, *11*(8), e1005421.

De Maio, N., Wu, C.-H., & Wilson, D. J. (2016). SCOTTI: Efficient Reconstruction of Transmission within Outbreaks with the Structured Coalescent. *PLoS Computational Biology*, *12*(9), e1005130.

Didelot, X., Fraser, C., Gardy, J., & Colijn, C. (2017). Genomic Infectious Disease Epidemiology in Partially Sampled and Ongoing Outbreaks. *Molecular Biology and Evolution*, *34*(4), 997–1007.

Didelot, X., Gardy, J., & Colijn, C. (2014). Bayesian inference of infectious disease transmission from whole-genome sequence data. *Molecular Biology and Evolution*, *31*(7), 1869–1879.

Dizney, L., & Dearing, M. D. (2013). The role of behavioural heterogeneity on infection patterns: implications for pathogen transmission. *Animal Behaviour*, *86*(5).

Dobson, A. P., Pimm, S. L., Hannah, L., Kaufman, L., Ahumada, J. A., Ando, A. W., Bernstein, A., Busch, J., Daszak, P., Engelmann, J., Kinnaird, M. F., Li, B. V., Loch-Temzelides, T., Lovejoy, T., Nowak, K., Roehrdanz, P. R., & Vale, M. M. (2020). Ecology and economics for pandemic prevention. *Science*, *369*(6502), 379–381.

Dougherty, E. R., Seidel, D. P., Carlson, C. J., Spiegel, O., & Getz, W. M. (2018). Going through the motions: incorporating movement analyses into disease research.

Ecology Letters, *21*(4), 588–604.

Drewe, J. A. (2010). Who infects whom? Social networks and tuberculosis transmission in wild meerkats. *Proceedings. Biological Sciences / The Royal Society*, *277*(1681), 633–642.

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*, 214.

Eames, K., Bansal, S., Frost, S., & Riley, S. (2015). Six challenges in measuring contact networks for use in modelling. *Epidemics*, *10*, 72–77.

Elbroch, L. M., Lendrum, P. E., Quigley, H., & Caragiulo, A. (2016). Spatial overlap in a solitary carnivore: support for the land tenure, kinship or resource dispersion hypotheses? *The Journal of Animal Ecology*, *85*(2), 487–496.

Elbroch, L. M., & Quigley, H. (2016). Social interactions in a solitary carnivore. *Current Zoology*, *63*(4), 357–362.

Esri, National Atlas of the United States, United States Geological Survey, Department of Commerce, Census Bureau-Geography Division. (n.d.). *USA Urban Areas (FeatureServer)*.

Farine, D. R., & Whitehead, H. (2015). Constructing, conducting and interpreting animal social network analysis. *The Journal of Animal Ecology*, *84*(5), 1144–1163.

Ferrier, S., Manion, G., Elith, J., & Richardson, K. (2007). Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions*, *13*(3), 252–264.

Fieberg, J., Kochanny, C. O., & Lanham. (2005). Quantifying home-range overlap: the importance of the utilization distribution. *The Journal of Wildlife Management*, *69*(4), 1346–1359.

Firestone, S. M., Hayama, Y., Bradhurst, R., Yamamoto, T., Tsutsui, T., & Stevenson, M. A. (2019). Reconstructing foot-and-mouth disease outbreaks: a methods comparison of transmission network models. *Scientific Reports*, *9*(1), 4809.

Fountain-Jones, N. M., Craft, M. E., Funk, W. C., Kozakiewicz, C., Trumbo, D., Boydston, E. E., Lyren, L. M., Crooks, K., Lee, J. S., VandeWoude, S., & Carver, S. (2017). Urban landscapes can change virus gene flow and evolution in a fragmentation-sensitive carnivore. *Molecular Ecology*. https://doi.org/10.1111/mec.14375

Fountain-Jones, N. M., Kraberger, S., Gagne, R., Trumbo, D. R., Salerno, P., Chris Funk, W., Crooks, K., Biek, R., Alldredge, M., Logan, K., Baele, G., Dellicour, S., Ernest, H. B., VandeWoude, S., Carver, S., & Craft, M. E. (2019). Host relatedness and landscape connectivity shape pathogen spread in a large secretive carnivore. *bioRxiv*, 816009.

Fountain-Jones, N. M., Packer, C., Troyer, J. L., VanderWaal, K., Robinson, S., Jacquot, M., & Craft, M. E. (2017). Linking social and spatial networks to viral community phylogenetics reveals subtype-specific transmission dynamics in African lions. *The Journal of Animal Ecology*. https://doi.org/10.1111/1365-2656.12751

Francis, S. S., Plucinski, M. M., Wallace, A. D., & Riley, L. W. (2016). Genotyping Oral Commensal Bacteria to Predict Social Contact and Structure. *PloS One*, *11*(9), e0160201.

Franklin, S. P., Troyer, J. L., Terwee, J. A., Lyren, L. M., Boyce, W. M., Riley, S. P. D., Roelke, M. E., Crooks, K. R., & Vandewoude, S. (2007). Frequent transmission of immunodeficiency viruses among bobcats and pumas. *Journal of Virology*, *81*(20), 10961–10969.

Franks, D. W., Ruxton, G. D., & James, R. (2010). Sampling animal association networks with the gambit of the group. *Behavioral Ecology and Sociobiology*, *64*(3), 493–503.

French, J. (2020). *smacpod: Statistical Methods for the Analysis of Case-Control Point Data*. https://CRAN.R-project.org/package=smacpod

Fromont, E., Artois, M., Langlais, M., Courchamp, F., & Pontier, D. (1997). Modelling the feline leukemia virus (FeLV) in natural populations of cats (*Felis catus*). *Theoretical Population Biology*, *52*(1), 60–70.

Fronhofer, E. A., Hovestadt, T., & Poethke, H.-J. (2013). From random walks to informed movement. *Oikos* , *122*(6), 857–866.

Gardy, J. L., Johnston, J. C., Ho Sui, S. J., Cook, V. J., Shah, L., Brodkin, E., Rempel, S., Moore, R., Zhao, Y., Holt, R., Varhol, R., Birol, I., Lem, M., Sharma, M. K., Elwood, K., Jones, S. J. M., Brinkman, F. S. L., Brunham, R. C., & Tang, P. (2011). Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *The New England Journal of Medicine*, *364*(8), 730–739.

Geret, C. P., Cattori, V., Meli, M. L., Riond, B., Martinez, F., Lopez, G., ... & Lutz, H. (2011). Feline leukemia virus outbreak in the critically endangered Iberian lynx (*Lynx pardinus*): high-throughput sequencing of envelope variable region A and experimental transmission. *Archives of virology*, 156(5), 839-854.

Gilbert, M., Sulikhan, N., Uphyrkina, O., Goncharuk, M., Kerley, L., Hernandez Castro, E., Reeve, R., Seimon, T., McAloose, D., Seryodkin, I. V., Naidenko, S. V., Davis, C. A., Wilkie, G. S., Vattipally, S. B., Adamson, W. E., Hinds, C., Thomson, E. C., Willett, B. J., Hosie, M. J., … Cleaveland, S. (2020). Distemper, extinction, and vaccination of the Amur tiger. *Proceedings of the National*

Academy of Sciences of the United States of America. https://doi.org/10.1073/pnas.2000153117

Gilbertson, M. L., Carver, S., VandeWoude, S., Crooks, K. R., Lappin, M. R., & Craft, M. E. (2016). Is pathogen exposure spatially autocorrelated? Patterns of pathogens in puma (*Puma concolor*) and bobcat (*Lynx rufus*). *Ecosphere*, 7(11), e01558.

Gilbertson, M. L., Fountain-Jones, N. M., & Craft, M. E. (2018). Incorporating genomic methods into contact networks to reveal new insights into animal behaviour and infectious disease dynamics. *Behaviour*, 155(7-9), 759-791.

Godfrey, S. S., Moore, J. A., Nelson, N. J., & Bull, C. M. (2010). Social network structure and parasite infection patterns in a territorial reptile, the tuatara (*Sphenodon punctatus*). *International Journal for Parasitology*, *40*(13), 1575–1585.

Godfrey, S. S. (2013). Networks and the ecology of parasite transmission: A framework for wildlife parasitology. *International Journal for Parasitology. Parasites and Wildlife*, *2*, 235–245.

Godfrey, S. S., Ansari, T. H., Gardner, M. G., Farine, D. R., & Bull, C. M. (2014). A contact-based social network of lizards is defined by low genetic relatedness among strongly connected individuals. *Animal Behaviour*, *97*, 35–43.

Gottdenker, N. L., Streicker, D. G., Faust, C. L., & Carroll, C. R. (2014). Anthropogenic land use change and infectious diseases: a review of the evidence. *EcoHealth*, *11*(4), 619–632.

Grad, Y. H., Kirkcaldy, R. D., Trees, D., Dordel, J., Harris, S. R., Goldstein, E., Weinstock, H., Parkhill, J., Hanage, W. P., Bentley, S., & Lipsitch, M. (2014). Genomic epidemiology of *Neisseria gonorrhoeae* with reduced susceptibility to cefixime in the USA: a retrospective observational study. *The Lancet Infectious*

Diseases, *14*(3), 220–226.

Greene, C. E. (2012). *Infectious diseases of the dog and cat* (4th ed..). Elsevier/Saunders.

Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L. N., Daly, J. M., Mumford, J. A., & Holmes, E. C. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *Science*, *303*(5656), 327–332.

Grubaugh, N. D., Gangavarapu, K., Quick, J., Matteson, N. L., De Jesus, J. G., Main, B. J., Tan, A. L., Paul, L. M., Brackney, D. E., Grewal, S., Gurfield, N., Van Rompay, K. K. A., Isern, S., Michael, S. F., Coffey, L. L., Loman, N. J., & Andersen, K. G. (2019). An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biology*, *20*(1), 8.

Hall, M. D., Woolhouse, M. E. J., & Rambaut, A. (2016). Using genomics data to reconstruct transmission trees during disease outbreaks. *Revue Scientifique et Technique* , *35*(1), 287–296.

Hall, M., Woolhouse, M., & Rambaut, A. (2015). Epidemic Reconstruction in a Phylogenetics Framework: Transmission Trees as Partitions of the Node Set. *PLoS Computational Biology*, *11*(12), e1004613.

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*(11), 1147–1157.

Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). statnet: Software tools for the representation, visualization, analysis and simulation of network data. *Journal of Statistical Software*, *24*(1), 1548.

Hartmann, K. (2012). Clinical aspects of feline retroviruses: a review. *Viruses*, *4*(11), 2684–2710.

Haydon, D. T., Randall, D. A., Matthews, L., Knobel, D. L., Tallents, L. A., Gravenor, M. B., Williams, S. D., Pollinger, J. P., Cleaveland, S., Woolhouse, M. E. J., Sillero-Zubiri, C., Marino, J., Macdonald, D. W., & Laurenson, M. K. (2006). Low-coverage vaccination strategies for the conservation of endangered species. *Nature*, *443*(7112), 692–695.

Hirsch, B. T., Prange, S., Hauver, S. A., & Gehrt, S. D. (2013). Raccoon social networks and the potential for disease transmission. *PloS One*, *8*(10), e75830.

Hoffmann, C., Stockhausen, M., Merkel, K., Calvignac-Spencer, S., & Leendertz, F. H. (2016). Assessing the feasibility of fly based surveillance of wildlife infectious diseases. *Scientific Reports*, *6*, 37952.

Hulbert, I. A. R., & French, J. (2001). The accuracy of GPS for wildlife telemetry and habitat mapping. *The Journal of Applied Ecology*, *38*(4), 869–878.

Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks. *Journal of Statistical Software*, *24*(3), nihpa54860.

Johnson, W. E., Onorato, D. P., Roelke, M. E., Land, E. D., Cunningham, M., Belden, R. C., McBride, R., Jansen, D., Lotz, M., Shindle, D., Howard, J., Wildt, D. E., Penfold, L. M., Hostetler, J. A., Oli, M. K., & O'Brien, S. J. (2010). Genetic restoration of the Florida panther. *Science*, *329*(5999), 1641–1645.

Jombart, T., Eggo, R. M., Dodd, P. J., & Balloux, F. (2011). Reconstructing disease outbreaks from genetic data: a graph approach. *Heredity*, *106*(2), 383–390.

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., &

Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, *451*(7181), 990–993.

Joseph, M. B., Mihaljevic, J. R., Arellano, A. L., Kueneman, J. G., Preston, D. L., Cross, P. C., & Johnson, P. T. J. (2013). Taming wildlife disease: bridging the gap between science and management. *The Journal of Applied Ecology*, *50*(3), 702–712.

Kao, R. R., Haydon, D. T., Lycett, S. J., & Murcia, P. R. (2014). Supersize me: how whole-genome sequencing and big data are transforming epidemiology. *Trends in Microbiology*, *22*(5), 282–291.

Keeling, M. J., & Eames, K. T. D. (2005). Networks and epidemic models. *Journal of the Royal Society, Interface / the Royal Society*, *2*(4), 295–307.

Keeling, M. J., & Rohani, P. (2011). *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press.

Keeling, M. J., Woolhouse, M. E. J., May, R. M., Davies, G., & Grenfell, B. T. (2003). Modelling vaccination strategies against foot-and-mouth disease. *Nature*, *421*(6919), 136–142.

Keogh-Brown, M. R., & Smith, R. D. (2008). The economic impact of SARS: How does the reality match the predictions? *Health Policy* , *88*(1), 110–120.

Klinkenberg, D., Backer, J. A., Didelot, X., Colijn, C., & Wallinga, J. (2017). Simultaneous inference of phylogenetic and transmission trees in infectious disease outbreaks. *PLoS Computational Biology*, *13*(5), e1005495.

Knight-Jones, T. J. D., & Rushton, J. (2013). The economic impacts of foot and mouth disease--What are they, how big are they and where do they occur? *Preventive Veterinary Medicine*, *112*(3-4), 161–173.

Kraberger, S., Fountain-Jones, N. M., Gagne, R. B., Malmberg, J., Dannemiller, N. G., Logan, K., Alldredge, M., Varsani, A., Crooks, K. R., Craft, M., Carver, S., & VandeWoude, S. (2020). Frequent cross-species transmissions of foamy virus between domestic and wild felids. *Virus Evolution*, *6*(1), vez058.

Krakoff, E., Gagne, R. B., VandeWoude, S., & Carver, S. (2019). Variation in Intra-individual Lentiviral Evolution Rates: a Systematic Review of Human, Nonhuman Primate, and Felid Species. *Journal of Virology*, *93*(16).

Krause, J., Krause, S., Arlinghaus, R., Psorakis, I., Roberts, S., & Rutz, C. (2013). Reality mining of animal social systems. *Trends in Ecology & Evolution*, *28*(9), 541–551.

Krivitsky, P. N. (2012). Exponential-family random graph models for valued networks. *Electronic Journal of Statistics*, *6*, 1100–1128.

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics - Theory and Methods*, *26*(6), 1481–1496.

Lagana, D. M., Lee, J. S., Lewis, J. S., Bevins, S. N., Carver, S., Sweanor, L. L., McBride, R., McBride, C., Crooks, K. R., & VandeWoude, S. (2013). Characterization of regionally associated feline immunodeficiency virus (FIV) in bobcats (*Lynx rufus*). *Journal of Wildlife Diseases*, *49*(3), 718–722.

Lax, S., Smith, D. P., Hampton-Marcell, J., Owens, S. M., Handley, K. M., Scott, N. M., Gibbons, S. M., Larsen, P., Shogan, B. D., Weiss, S., Metcalf, J. L., Ursell, L. K., Vázquez-Baeza, Y., Van Treuren, W., Hasan, N. A., Gibson, M. K., Colwell, R., Dantas, G., Knight, R., & Gilbert, J. A. (2014). Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science*, *345*(6200), 1048–1052.

Lee, J. S., Bevins, S. N., Serieys, L. E. K., Vickers, W., Logan, K. A., Aldredge, M., Boydston, E. E., Lyren, L. M., McBride, R., Roelke-Parker, M., Pecon-Slattery, J., Troyer, J. L., Riley, S. P., Boyce, W. M., Crooks, K. R., & VandeWoude, S. (2014). Evolution of puma lentivirus in bobcats (*Lynx rufus*) and mountain lions (*Puma concolor*) in North America. *Journal of Virology*, *88*(14), 7727–7737.

Lee, J. S., Ruell, E. W., Boydston, E. E., Lyren, L. M., Alonso, R. S., Troyer, J. L., Crooks, K. R., & Vandewoude, S. (2012). Gene flow and pathogen transmission among bobcats (*Lynx rufus*) in a fragmented urban landscape. *Molecular Ecology*, *21*(7), 1617–1631.

Leigh Brown, A. J., Lycett, S. J., Weinert, L., Hughes, G. J., Fearnhill, E., Dunn, D. T., & UK HIV Drug Resistance Collaboration. (2011). Transmission network parameters estimated from HIV sequences for a nationwide epidemic. *The Journal of Infectious Diseases*, *204*(9), 1463–1469.

Lembo, T., Hampson, K., Haydon, D. T., Craft, M., Dobson, A., Dushoff, J., Ernest, E., Hoare, R., Kaare, M., Mlengeya, T., Mentzel, C., & Cleaveland, S. (2008). Exploring reservoir dynamics: a case study of rabies in the Serengeti ecosystem. *The Journal of Applied Ecology*, *45*(4), 1246–1257.

Lemey, P., Rambaut, A., Bedford, T., Faria, N., Bielejec, F., Baele, G., Russell, C. A., Smith, D. J., Pybus, O. G., Brockmann, D., & Suchard, M. A. (2014). Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. *PLoS Pathogens*, *10*(2), e1003932.

Lemey, P., Rambaut, A., Drummond, A. J., & Suchard, M. A. (2009). Bayesian phylogeography finds its roots. *PLoS Computational Biology*, *5*(9), e1000520.

Lemey, P., Rambaut, A., Welch, J. J., & Suchard, M. A. (2010). Phylogeography takes a

relaxed random walk in continuous space and time. *Molecular Biology and Evolution*, *27*(8), 1877–1885.

Leventhal, G. E., Kouyos, R., Stadler, T., Wyl, V. von, Yerly, S., Böni, J., Cellerai, C., Klimkait, T., Günthard, H. F., & Bonhoeffer, S. (2012). Inferring epidemic contact structure from phylogenetic trees. *PLoS Computational Biology*, *8*(3), e1002413.

Levins, R. (1966). The strategy of model building in population biology. *American Scientist*, *54*(4), 421–431.

Lewis, F., Hughes, G. J., Rambaut, A., Pozniak, A., & Leigh Brown, A. J. (2008). Episodic sexual transmission of HIV revealed by molecular phylodynamics. *PLoS Medicine*, *5*(3), e50.

Lewis, J. S., Logan, K. A., Alldredge, M. W., Theobald, D. M., VandeWoude, S., & Crooks, K. R. (2017). Contact networks reveal potential for interspecific interactions of sympatric wild felids driven by space use. *Ecosphere*, *8*(3). https://doi.org/10.1002/ecs2.1707

Liaw, A., Wiener, M., & Others. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22.

Little, S., Levy, J., Hartmann, K., Hofmann-Lehmann, R., Hosie, M., Olah, G., & Denis, K. S. (2020). 2020 AAFP Feline Retrovirus Testing and Management Guidelines. *Journal of Feline Medicine and Surgery*, *22*(1), 5–30.

Lloyd-Smith, J. O., George, D., Pepin, K. M., Pitzer, V. E., Pulliam, J. R. C., Dobson, A. P., Hudson, P. J., & Grenfell, B. T. (2009). Epidemic dynamics at the human-animal interface. *Science*, *326*(5958), 1362–1367.

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*(7066),

355–359.

Logan, K. A., & Sweanor, L. L. (2001). *Desert Puma: Evolutionary Ecology And Conservation Of An Enduring Carnivore*. Island Press.

Long, J. A., Nelson, T. A., Webb, S. L., & Gee, K. L. (2014). A critical examination of indices of dynamic interaction for wildlife telemetry studies. *The Journal of Animal Ecology*, *83*(5), 1216–1233.

López, G., López-Parra, M., Fernández, L., Martínez-Granados, C., Martínez, F., Meli, M. L., Gil-Sánchez, J. M., Viqueira, N., Díaz-Portero, M. A., Cadenas, R., Lutz, H., Vargas, A., & Simón, M. A. (2009). Management measures to control a feline leukemia virus outbreak in the endangered Iberian lynx. *Animal Conservation*, *12*(3), 173–182.

MacIntosh, A. J. J., Jacobs, A., Garcia, C., Shimizu, K., Mouri, K., Huffman, M. A., & Hernandez, A. D. (2012). Monkeys in the middle: parasite transmission through the social network of a wild primate. *PloS One*, *7*(12), e51144.

Malmberg, J. L., Lee, J. S., Gagne, R. B., Kraberger, S., Kechejian, S., Roelke, M., McBride, R., Onorato, D., Cunningham, M., Crooks, K. R., & VandeWoude, S. (2019). Altered lentiviral infection dynamics follow genetic rescue of the Florida panther. *Proceedings. Biological Sciences / The Royal Society*, *286*(1913), 20191689.

Marino, S., Hogue, I. B., Ray, C. J., & Kirschner, D. E. (2008). A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology*, *254*(1), 178–196.

Marquetoux, N., Heuer, C., Wilson, P., Ridler, A., & Stevenson, M. (2016). Merging DNA typing and network analysis to assess the transmission of paratuberculosis

between farms. *Preventive Veterinary Medicine*, *134*, 113–121.

Martínez-López, B., Perez, A. M., & Sánchez-Vizcaíno, J. M. (2009). Social network analysis. Review of general concepts and use in preventive veterinary medicine. *Transboundary and Emerging Diseases*, *56*(4), 109–120.

McBride, R. T., McBride, R. T., McBride, R. M., & McBride, C. E. (2008). Counting Pumas by Categorizing Physical Evidence. *Southeastern Naturalist*, *7*(3), 381–400.

McCallum, H. (2008). Landscape structure, disturbance, and disease dynamics. *Infectious Disease Ecology: Effects of Ecosystems on Disease and of Disease on Ecosystems*, 100–122.

McCallum, H. (2016). Models for managing wildlife disease. *Parasitology*, *143*(7), 805–820.

McClintock, B. T., Onorato, D. P., & Martin, J. (2015). Endangered Florida panther population size determined from public reports of motor vehicle collision mortalities. *The Journal of Applied Ecology*, *52*(4), 893–901.

McCloskey, R. M., Liang, R. H., & Poon, A. F. Y. (2016). Reconstructing contact network parameters from viral phylogenies. *Virus Evolution*, *2*(2), vew029.

Metcalf, C. J. E., & Lessler, J. (2017). Opportunities and challenges in modeling emerging infectious diseases. *Science*, *357*(6347), 149–152.

Metzker, M. L., Mindell, D. P., Liu, X.-M., Ptak, R. G., Gibbs, R. A., & Hillis, D. M. (2002). Molecular evidence of HIV-1 transmission in a criminal case. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(22), 14292–14297.

Meyers, L. A. (2007). Contact network epidemiology: Bond percolation applied to

infectious disease prediction and control. *Bulletin of the American Mathematical Society*, *44*(01), 63–87.

Miguel, E., Grosbois, V., Caron, A., Pople, D., Roche, B., & Donnelly, C. A. (2020). A systemic approach to assess the potential and risks of wildlife culling for infectious disease control. *Communications Biology*, *3*(1), 353.

Minot, S., Bryson, A., Chehoud, C., Wu, G. D., Lewis, J. D., & Bushman, F. D. (2013). Rapid evolution of the human gut virome. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(30), 12450–12455.

Mollentze, N., Nel, L. H., Townsend, S., le Roux, K., Hampson, K., Haydon, D. T., & Soubeyrand, S. (2014). A Bayesian approach for inferring the dynamics of partially observed endemic infectious diseases from space-time-genetic data. *Proceedings. Biological Sciences / The Royal Society*, *281*(1782), 20133251.

Morris, M., Handcock, M. S., & Hunter, D. R. (2008). Specification of Exponential-Family Random Graph Models: Terms and Computational Aspects. *Journal of Statistical Software*, *24*(4), 1548–7660.

Morris, M., & Kretzschmar, M. (1995). Concurrent partnerships and transmission dynamics in networks. *Social Networks*, *17*(3), 299–318.

Mysterud, A., & Rolandsen, C. M. (2019). Fencing for wildlife disease control. *The Journal of Applied Ecology*, *56*(3), 519–525.

Natoli, E., Say, L., Cafazzo, S., Bonanni, R., Schmid, M., & Pontier, D. (2005). Bold attitude makes male urban feral domestic cats more vulnerable to Feline Immunodeficiency Virus. *Neuroscience and Biobehavioral Reviews*, *29*(1), 151–157.

Numminen, E., Chewapreecha, C., Sirén, J., Turner, C., Turner, P., Bentley, S. D., &

Corander, J. (2014). Two-phase importance sampling for inference about transmission trees. *Proceedings. Biological Sciences / The Royal Society*, *281*(1794), 20141324.

Nunn, C. L., Jordán, F., McCabe, C. M., Verdolin, J. L., & Fewell, J. H. (2015). Infectious disease and group size: more than just a numbers game. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1669). https://doi.org/10.1098/rstb.2014.0111

O'Brien, J. M., O'Brien, C. S., MCcarthy, C., & Carpenter, T. E. (2014). Incorporating foray behavior into models estimating contact risk between bighorn sheep and areas occupied by domestic sheep. *Wildlife Society Bulletin*, *38*(2), 321–331.

Pellis, L., Ball, F., Bansal, S., Eames, K., House, T., Isham, V., & Trapman, P. (2015). Eight challenges for network epidemic models. *Epidemics*, *10*, 58–62.

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. *The Journal of Animal Ecology*, *78*(5), 1015–1022.

Picard, C., Dallot, S., Brunker, K., Berthier, K., Roumagnac, P., Soubeyrand, S., Jacquot, E., & Thébaud, G. (2017). Exploiting Genetic Information to Trace Plant Virus Dispersal in Landscapes. *Annual Review of Phytopathology*, *55*, 139–160.

Pike, J., Bogich, T., Elwood, S., Finnoff, D. C., & Daszak, P. (2014). Economic optimization of a global strategy to address the pandemic threat. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(52), 18519–18523.

Plowright, R. K., Parrish, C. R., McCallum, H., Hudson, P. J., Ko, A. I., Graham, A. L., &

Lloyd-Smith, J. O. (2017). Pathways to zoonotic spillover. *Nature Reviews. Microbiology*, *15*(8), 502–510.

Plowright, R. K., Sokolow, S. H., Gorman, M. E., Daszak, P., & Foley, J. E. (2008). Causal inference in disease ecology: investigating ecological drivers of disease emergence. *Frontiers in Ecology and the Environment*, *6*(8), 420–429.

Pluciński, M. M., Starfield, R., & Almeida, R. P. P. (2011). Inferring social network structure from bacterial sequence data. *PloS One*, *6*(8), e22685.

Pope, L. C., Butlin, R. K., Wilson, G. J., Woodroffe, R., Erven, K., Conyers, C. M., Franklin, T., Delahay, R. J., Cheeseman, C. L., & Burke, T. (2007). Genetic evidence that culling increases badger movement: implications for the spread of bovine tuberculosis. *Molecular Ecology*, *16*(23), 4919–4929.

Porphyre, T., Stevenson, M., Jackson, R., & McKenzie, J. (2008). Influence of contact heterogeneity on TB reproduction ratio R0 in a free-living brushtail possum *Trichosurus vulpecula* population. *Veterinary Research*, *39*(3), 31.

Potapov, A., Merrill, E., & Lewis, M. A. (2012). Wildlife disease elimination and density dependence. *Proceedings. Biological Sciences / The Royal Society*, *279*(1741), 3139–3145.

Quick, J., Grubaugh, N. D., Pullan, S. T., Claro, I. M., Smith, A. D., Gangavarapu, K., Oliveira, G., Robles-Sikisaka, R., Rogers, T. F., Beutler, N. A., Burton, D. R., Lewis-Ximenez, L. L., de Jesus, J. G., Giovanetti, M., Hill, S. C., Black, A., Bedford, T., Carroll, M. W., Nunes, M., … Loman, N. J. (2017). Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nature Protocols*, *12*(6), 1261–1276.

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. R

Foundation for Statistical Computing. https://www.R-project.org/

Rasmussen, A. L. (2015). Probing the Viromic Frontiers. *mBio*, *6*(6), e01767–15.

Ray, B., Ghedin, E., & Chunara, R. (2016). Network inference from multimodal data: A review of approaches from infectious disease transmission. *Journal of Biomedical Informatics*, *64*, 44–54.

Rees, E. E., Pond, B. A., Tinline, R. R., & Bélanger, D. (2013). Modelling the effect of landscape heterogeneity on the efficacy of vaccination for wildlife infectious disease control. *The Journal of Applied Ecology*, *50*(4), 881–891.

Reynolds, J. J. H., Carver, S., Cunningham, M. W., Logan, K. A., Vickers, W., Crooks, K. R., VandeWoude, S., & Craft, M. E. (2019). Feline immunodeficiency virus in puma: Estimation of force of infection reveals insights into transmission. *Ecology and Evolution*, *9*(19), 11010–11024.

Reynolds, J. J. H., Hirsch, B. T., Gehrt, S. D., & Craft, M. E. (2015). Raccoon contact networks predict seasonal susceptibility to rabies outbreaks and limitations of vaccination. *The Journal of Animal Ecology*, *84*(6), 1720–1731.

Robert, K., Garant, D., & Pelletier, F. (2012). Keep in touch: Does spatial overlap correlate with contact rate frequency? *The Journal of Wildlife Management*, *76*(8), 1670–1675.

Robinson, K., Fyson, N., Cohen, T., Fraser, C., & Colijn, C. (2013). How the dynamics and structure of sexual contact networks shape pathogen phylogenies. *PLoS Computational Biology*, *9*(6), e1003105.

Robinson S. J., Barbieri Michelle M., Murphy Samantha, Baker Jason D., Harting Albert L., Craft Meggan E., & Littnan Charles L. (2018). Model recommendations meet management reality: implementation and evaluation of a network-informed

vaccination effort for endangered Hawaiian monk seals. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1870), 20171899.

Robitaille, A. L., Webber, Q. M. R., & Vander Wal, E. (2019). Conducting social network analysis with animal telemetry data: Applications and methods using spatsoc. *Methods in Ecology and Evolution / British Ecological Society*, *0*(0). https://doi.org/10.1111/2041-210X.13215

Roelke, M. E., Forrester, D. J., Jacobson, E. R., Kollias, G. V., Scott, F. W., Barr, M. C., Evermann, J. F., & Pirtle, E. C. (1993). Seroprevalence of infectious disease agents in free-ranging Florida panthers (*Felis concolor coryi*). *Journal of Wildlife Diseases*, *29*(1), 36–49.

Roelke-Parker, M. E., Munson, L., Packer, C., Kock, R., Cleaveland, S., Carpenter, M., O'Brien, S. J., Pospischil, A., Hofmann-Lehmann, R., Lutz, H., Mwamengele, G. L., Mgasa, M. N., Machange, G. A., Summers, B. A., & Appel, M. J. (1996). A canine distemper virus epidemic in Serengeti lions (*Panthera leo*). *Nature*, *379*(6564), 441–445.

Romano, C. M., de Carvalho-Mello, I. M. V. G., Jamal, L. F., de Melo, F. L., Iamarino, A., Motoki, M., Pinho, J. R. R., Holmes, E. C., de Andrade Zanotto, P. M., & VGDN Consortium. (2010). Social networks shape the transmission dynamics of hepatitis C virus. *PloS One*, *5*(6), e11170.

Rushmore, J., Caillaud, D., Hall, R. J., Stumpf, R. M., Meyers, L. A., & Altizer, S. (2014). Network-based vaccination improves prospects for disease control in wild chimpanzees. *Journal of the Royal Society, Interface / the Royal Society*, *11*(97), 20140349.

Rushmore, J., Caillaud, D., Matamba, L., Stumpf, R. M., Borgatti, S. P., & Altizer, S.

(2013). Social network analysis of wild chimpanzees provides insights for predicting infectious disease risk. *The Journal of Animal Ecology*, *82*(5), 976–986.

Sah, P., Leu, S. T., Cross, P. C., Hudson, P. J., & Bansal, S. (2017). Unraveling the disease consequences and mechanisms of modular structure in animal social networks. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(16), 4165–4170.

Sah, P., Mann, J., & Bansal, S. (2018). Disease implications of animal social network structure: A synthesis across social systems. *The Journal of Animal Ecology*, *87*(3), 546–558.

Sah, P., Méndez, J. D., & Bansal, S. (2019). A multi-species repository of social networks. *Scientific Data*, *6*(1), 44.

Sanchez, J. N., & Hudgens, B. R. (2020). Vaccination and monitoring strategies for epidemic prevention and detection in the Channel Island fox (*Urocyon littoralis*). *PloS One*, *15*(5), e0232705.

Sharp, P. M., & Hahn, B. H. (2010). The evolution of HIV-1 and the origin of AIDS. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *365*(1552), 2487–2494.

Silk, M. J., Croft, D. P., Delahay, R. J., Hodgson, D. J., Weber, N., Boots, M., & McDonald, R. A. (2017). The application of statistical network models in disease research. *Methods in Ecology and Evolution / British Ecological Society*, *8*(9), 1026–1041.

Sintchenko, V., & Holmes, E. C. (2015). The role of pathogen genomics in assessing disease transmission. *BMJ*, *350*, h1314.

Smiley Evans, T., Gilardi, K. V. K., Barry, P. A., Ssebide, B. J., Kinani, J. F., Nizeyimana, F., Noheri, J. B., Byarugaba, D. K., Mudakikwa, A., Cranfield, M. R., Mazet, J. A. K., & Johnson, C. K. (2016). Detection of viruses using discarded plants from wild mountain gorillas and golden monkeys. *American Journal of Primatology*, *78*(11), 1222–1234.

Springer, A., Mellmann, A., Fichtel, C., & Kappeler, P. M. (2016). Social structure and *Escherichia coli* sharing in a group-living wild primate, Verreaux's sifaka. *BMC Ecology*, *16*, 6.

Schauber, E. M., Nielsen, C. K., Kjær, L. J., Anderson, C. W., & Storm, D. J. (2015). Social affiliation and contact patterns among white-tailed deer in disparate landscapes: implications for disease transmission. *Journal of Mammalogy*, *96*(1), 16–28.

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. *The Journal of Wildlife Management*, *71*(1), 155–163.

Silk, M. J., & Fisher, D. N. (2017). Understanding animal social structure: exponential random graph models in animal behaviour research. *Animal Behaviour*, *132*, 137–146.

Silk, M. J., Jackson, A. L., Croft, D. P., Colhoun, K., & Bearhop, S. (2015). The consequences of unidentifiable individuals for the analysis of an animal social network. *Animal Behaviour*, *104*, 1–11.

Sillero-Zubiri, C., King, A. A., & Macdonald, D. W. (1996). Rabies and mortality in Ethiopian wolves (*Canis simensis*). *Journal of Wildlife Diseases*, *32*(1), 80–86.

Smith, D. L., Lucey, B., Waller, L. A., Childs, J. E., & Real, L. A. (2002). Predicting the

spatial dynamics of rabies epidemics on heterogeneous landscapes. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(6), 3668–3672.

Smith, J. A., & Moody, J. (2013). Structural Effects of Network Sampling Coverage I: Nodes Missing at Random. *Social Networks*, *35*(4). https://doi.org/10.1016/j.socnet.2013.09.003

Song, S. J., Lauber, C., Costello, E. K., Lozupone, C. A., Humphrey, G., Berg-Lyons, D., Caporaso, J. G., Knights, D., Clemente, J. C., Nakielny, S., Gordon, J. I., Fierer, N., & Knight, R. (2013). Cohabiting family members share microbiota with one another and with their dogs. *eLife*, *2*, e00458.

Sparkes, A. H. (1997). Feline leukaemia virus: a review of immunity and vaccination. *The Journal of Small Animal Practice*, *38*(5), 187–194.

Streicker, D. G., Winternitz, J. C., Satterfield, D. A., Condori-Condori, R. E., Broos, A., Tello, C., Recuenco, S., Velasco-Villa, A., Altizer, S., & Valderrama, W. (2016). Host-pathogen evolutionary signatures reveal dynamics and future invasions of vampire bat rabies. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(39), 10926–10931.

Tissot-Dupont, H., Amadei, M.-A., Nezri, M., & Raoult, D. (2004). Wind in November, Q fever in December. *Emerging Infectious Diseases*, *10*(7), 1264–1269.

Torres, A. N., Mathiason, C. K., & Hoover, E. A. (2005). Re-examination of feline leukemia virus: host relationships using real-time PCR. *Virology*, *332*(1), 272–283.

Tracey, J. A., Bevins, S. N., VandeWoude, S., & Crooks, K. R. (2014). An agent-based movement model to assess the impact of landscape fragmentation on disease

transmission. *Ecosphere* , *5*(9), 1–24.

Van de Kerk, M., Onorato, D. P., Hostetler, J. A., Bolker, B. M., & Oli, M. K. (2019). Dynamics, persistence, and genetic management of the endangered florida panther population. *Wildlife Monographs*, *203*(1), 3–35.

Van Moorter, B., Visscher, D., Benhamou, S., Börger, L., Boyce, M. S., & Gaillard, J.-M. (2009). Memory keeps you at home: a mechanistic model for home range emergence. *Oikos* , *118*(5), 641–652.

VanderWaal, K. L., Atwill, E. R., Isbell, L. A., & McCowan, B. (2014a). Linking social and pathogen transmission networks using microbial genetics in giraffe (*Giraffa camelopardalis*). *The Journal of Animal Ecology*, *83*(2), 406–414.

VanderWaal, K. L., Atwill, E. R., Isbell, L. A., & McCowan, B. (2014b). Quantifying microbe transmission networks for wild and domestic ungulates in Kenya. *Biological Conservation*, *169*, 136–146.

VanderWaal, K. L., & Ezenwa, V. O. (2016). Heterogeneity in pathogen transmission: mechanisms and methodology. *Functional Ecology*, *30*(10), 1606–1622.

VanderWaal, K., Gilbertson, M., Okanga, S., Allan, B. F., & Craft, M. E. (2017). Seasonality and pathogen transmission in pastoral cattle contact networks. *Royal Society Open Science*, *4*(12), 170808.

VandeWoude, S., & Apetrei, C. (2006). Going wild: lessons from naturally occurring T-lymphotropic lentiviruses. *Clinical Microbiology Reviews*, *19*(4), 728–762.

VandeWoude, S., Troyer, J., & Poss, M. (2010). Restrictions to cross-species transmission of lentiviral infection gleaned from studies of FIV. *Veterinary Immunology and Immunopathology*, *134*(1-2), 25–32.

Vander Wal, E., Laforge, M. P., & McLoughlin, P. D. (2014). Density dependence in

social behaviour: home range overlap and density interacts to affect conspecific encounter rates in a gregarious ungulate. *Behavioral Ecology and Sociobiology*, *68*(3), 383–390.

Vasylyeva, T. I., Friedman, S. R., Paraskevis, D., & Magiorkinis, G. (2016). Integrating molecular epidemiology and social network analysis to study infectious diseases: Towards a socio-molecular era for public health. *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, *46*, 248–255.

Vial, F., Cleaveland, S., Rasmussen, G., & Haydon, D. T. (2006). Development of vaccination strategies for the management of rabies in African wild dogs. *Biological Conservation*, *131*(2), 180–192.

Villaseñor-Sierra, A., Quiñonez-Alvarado, M. G., & Caballero-Hoyos, J. R. (2007). Interpersonal relationships and group A *streptococcus* spread in a Mexican day-care center. *Salud Publica de Mexico*, *49*(5), 323–329.

Volz, E., & Meyers, L. A. (2007). Susceptible--infected--recovered epidemics in dynamic contact networks. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1628), 2925–2934.

Wanelik, K. M., & Farine, D. R. (2019). How to characterise shared space use networks. *bioRxiv*. https://doi.org/10.1101/839530

Welch, D. (2011). Is network clustering detectable in transmission trees? *Viruses*, *3*(6), 659–676.

Welch, D., Bansal, S., & Hunter, D. R. (2011). Statistical inference to advance network models in epidemiology. *Epidemics*, *3*(1), 38–45.

Wey, T., Blumstein, D. T., Shen, W., & Jordán, F. (2008). Social network analysis of

animal behaviour: a promising tool for the study of sociality. *Animal Behaviour*, *75*(2), 333–344.

Wheeler, D. C., Waller, L. A., & Biek, R. (2010). Spatial analysis of feline immunodeficiency virus infection in cougars. *Spatial and Spatio-Temporal Epidemiology*, *1*(2-3), 151–161.

Whitehead, H., & James, R. (2015). Generalized affiliation indices extract affiliations from social network data. *Methods in Ecology and Evolution / British Ecological Society*, *6*(7), 836–844.

White, L. A., Forester, J. D., & Craft, M. E. (2017). Using contact networks to explore mechanisms of parasite transmission in wildlife. *Biological Reviews of the Cambridge Philosophical Society*, *92*(1), 389–409.

White, L. A., Forester, J. D., & Craft, M. E. (2018a). Covariation between the physiological and behavioral components of pathogen transmission: Host heterogeneity determines epidemic outcomes. *Oikos,* 127(4), 538-552.

White, L. A., Forester, J. D., & Craft, M. E. (2018b). Disease outbreak thresholds emerge from interactions between movement behavior, landscape structure, and epidemiology. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(28), 7374–7379.

White, L. A., VandeWoude, S., & Craft, M. E. (2020). A mechanistic, stigmergy model of territory formation in solitary animals: Territorial behavior can dampen disease prevalence but increase persistence. *PLoS Computational Biology*, *16*(6), e1007457.

Wiethoelter, A. K., Beltrán-Alcrudo, D., Kock, R., & Mor, S. M. (2015). Global trends in infectious diseases at the wildlife-livestock interface. *Proceedings of the National*

Academy of Sciences of the United States of America, *112*(31), 9662–9667.

Wilber, M. Q., Pepin, K. M., Campa, H., III, Hygnstrom, S. E., Lavelle, M. J., Xifara, T., VerCauteren, K. C., & Webb, C. T. (2019). Modelling multi-species and multi-mode contact networks: Implications for persistence of bovine tuberculosis at the wildlife--livestock interface. *The Journal of Applied Ecology*, *56*(6), 1471–1481.

Williams, E. S., Thorne, E. T., Appel, M. J., & Belitsky, D. W. (1988). Canine distemper in black-footed ferrets (*Mustela nigripes*) from Wyoming. *Journal of Wildlife Diseases*, *24*(3), 385–398.

Wilson, K., Bjørnstad, O. N., Dobson, A. P., Merler, S., Poglayen, G., Randolph, S. E., Read, A. F., & Skorping, A. (2002). Heterogeneities in macroparasite infections: patterns and processes. *The Ecology of Wildlife Diseases*, *44*, 6–44.

Worby, C. J., Lipsitch, M., & Hanage, W. P. (2014). Within-host bacterial diversity hinders accurate reconstruction of transmission networks from genomic distance data. *PLoS Computational Biology*, *10*(3), e1003549.

Worton, B. J. (1989). Kernel methods for estimating the utilization distribution in home-range studies. *Ecology*, 70(1), 164-168.

Wu, J., Dhingra, R., Gambhir, M., & Remais, J. V. (2013). Sensitivity analysis of infectious disease models: methods, advances and their application. *Journal of the Royal Society, Interface / the Royal Society*, *10*(86), 20121018.

Wylie, J. L., Cabral, T., & Jolly, A. M. (2005). Identification of networks of sexually transmitted infection: a molecular, geographic, and social network analysis. *The Journal of Infectious Diseases*, *191*(6), 899–906.

Wymant, C., Hall, M., Ratmann, O., Bonsall, D., Golubchik, T., de Cesare, M., Gall, A., Cornelissen, M., Fraser, C., & STOP-HCV Consortium, The Maela

Pneumococcal Collaboration, and The BEEHIVE Collaboration. (2018). PHYLOSCANNER: Inferring Transmission from Within- and Between-Host Pathogen Genetic Diversity. *Molecular Biology and Evolution*, *35*(3), 719–733.

Ypma, R. J. F., Bataille, A. M. A., Stegeman, A., Koch, G., Wallinga, J., & van Ballegooijen, W. M. (2012). Unravelling transmission trees of infectious diseases by combining genetic and epidemiological data. *Proceedings. Biological Sciences / The Royal Society*, *279*(1728), 444–450.

Ypma, R. J. F., Jonges, M., Bataille, A., Stegeman, A., Koch, G., van Boven, M., Koopmans, M., van Ballegooijen, W. M., & Wallinga, J. (2013). Genetic data provide evidence for wind-mediated transmission of highly pathogenic avian influenza. *The Journal of Infectious Diseases*, *207*(5), 730–735.

Ypma, R. J. F., van Ballegooijen, W. M., & Wallinga, J. (2013). Relating phylogenetic trees to transmission trees of infectious disease outbreaks. *Genetics*, *195*(3), 1055–1062.

Zubiri, C. S., & Gottelli, D. (1995). Spatial organization in the Ethiopian wolf *Canis simensis*: large packs and small stable home ranges. *Journal of Zoology*, *237*(1), 65–81.

Appendix A. Supporting information for Chapter 2

A.1 Materials and Methods

Simulation details

Movement simulations were based on a biased, correlated random walk (BCRW; Long et al., 2014). The BCRW is a relatively simple movement model, described by a scaled step length distribution (*h*), correlation of turning angle (ρ), bias strength (*b*), and bias shape (*c*) (Barton et al., 2009; Van Moorter et al., 2009). To isolate the effect of subsampling, all trajectories within a simulation were generated with the same BCRW parameterization, but six different parameter sets were used across simulations to evaluate the robustness of our findings to parameter variations that altered home range size and movement (Table A.1). BCRW parameter sets were based on previously published BCRW models (Barton et al., 2009; Fronhofer et al., 2013; Long et al., 2014), but with reduced bias strength, which produced home range movement with occasional excursions. Excursion behavior can be important for pathogen transmission (O'Brien et al., 2014), but may also be more likely to be missed in instances of subsampling telemetry data.

Three parameter sets varied home range size by altering the step length distribution scaling parameter (*h*), while leaving all other parameters constant; this produced small, medium, and large home range parameter sets (Table A.1). We tested three additional parameter sets which kept home range size roughly constant with the small home range model while varying other BCRW parameters such as ρ and *b* (Table A.1).

To mimic different *spatial configurations*, with a given parameter set, we varied the starting locations (and thus the home range centers) of individuals on a 30,000 by 30,000m “study area”; we tested random, lattice, and clustered spatial configurations (Figure A.1). In randomly distributed populations, starting locations were assigned by drawing from a uniform distribution. In lattice-distributed populations, starting locations were assigned by keeping the same distance in the x and y direction between all starting locations. Within a simulation, this distance was the same for all individuals, but it varied between simulations, and was set from one to two times the radius of an average home range for the given parameter set. Clustered simulations had a further level of variation; we assigned all 100 individuals in a clustered simulation to one, five, or ten equally sized clusters. For each clustered population, the first individual per cluster was assigned a random starting location within the simulated study area, and the rest of the individuals in that same cluster were assigned starting locations relative to this individual. Distances from the initial individual for each subsequent individual in a cluster were drawn from a negative binomial distribution, with the mean set to one- to two-times the radius of an average home range for the given parameter set, and the size parameter ranging from one to five (Wilson et al., 2002). Direction from the initial individual was drawn from a wrapped normal distribution with a mean of zero and a concentration parameter of zero (functionally yielding a wrapped uniform distribution). For random and lattice populations, we ran 500 simulations per parameter set per spatial configuration, and we ran 999 simulations per parameter set for clustered populations (i.e. 333 simulations per cluster size described above; Figure A.2). We performed all simulations in R, version 3.5.0 (R Core Team, 2018). BCRW R code was adapted from Long et al. (2014).

Spatiotemporal Contact Definitions

When detecting contacts in complete and sample networks, spatiotemporal contacts were defined as simultaneous associations within small, medium, or large distance thresholds, corresponding to 1, 10, or 100 distance units, respectively. These contact definitions should be considered relative to the step length distributions used in movement simulations (Table A.1), as simulations were scale free. For example, the smallest step length scaling parameter was h = 15, which was used with a chi distribution in the *adehabitatLT* R package (Calenge, 2006), based on code from Long et al (2014). With a sample of 129,600 steps (i.e. 90 days of simulated movement) drawn from this combination of scaling parameter and distribution, the cumulative straight-line distance of 8 steps exceeded 100 distance units (large contact threshold) in 97.0% of cases; 10 steps exceeded 100 units in 99.9% of cases. While simulated movements were not in a straight line, because even movements from our smallest step length scaling parameter could achieve the distance of our largest contact threshold within about 8-10 steps (i.e. 8-10 minutes)—and while considering only the movement of one individual—we consider our large contact threshold fairly strict and therefore representative of direct contact. Further, broad definitions of contact are used in the literature (e.g. Schauber et al., 2015; VanderWaal et al., 2017), often without reference to biological realism.

Modularity Analysis

Modularity is a network-level metric that describes the strength of within versus between community interactions in a network. We estimated modularity using three different community-finding algorithms: walktrap (with four steps), edge-betweenness,

and fast-greedy. For each algorithm, we calculated modularity using weighted and unweighted networks. Modularity was then calculated based on the results of the community-finding algorithm results. All community and modularity analyses were performed in R, version 3.5.0 (R Core Team, 2018), using the package *igraph* (Csardi & Nepusz, 2006). As with density and proportion isolates, performance of modularity under telemetry sampling was evaluated using percent error of mean modularity scores across simulation variations.

A.2 Results

The six parameter sets did not yield substantial changes in the patterns reported in the main text but more extensive results are reported here by network metric type. In addition, we calculated metrics and correlations for weighted betweenness and transitivity to examine if weighted values altered results; the weighted results were consistent with those for unweighted values.

For clarity, in all supplementary figures for betweenness, transitivity, degree, and strength, heat maps show the lower limit of the 95% confidence interval of mean correlation coefficient for any given network metric. Yellow and green represent the highest correlation values. Results are shown across movement model parameter sets, all five spatial configurations (random start, lattice start, one cluster, five clusters, and ten clusters), and five types of network comparisons (given on the right y-axis). For network comparisons, the first label gives the contact definition for the complete networks, and the second label the contact definition for the sample networks. For example, “10m:SO” indicates comparison of 10m complete networks to spatial overlap

sample networks. For frequency of telemetry sampling, “q1m” means “every 1 minute” and so on.

Global network metrics

Global network metrics, betweenness and transitivity, generally performed poorly across all simulation variations (Figures A.3-A.6). Even when using less precise contact definitions in sample networks, improvements in metric performance were minimal (Figures A.3 and A.5). Results were robust to changes in the movement model when home range size was held constant (Figures A.4 and A.6).

Local network metrics

Local network metrics, degree and strength, outperformed global network metrics, especially in clustered populations (Figures A.7-A.10). Strength, in particular, performed the best out of all network metrics with the most consistently high correlation scores across simulation variations. However, under some conditions, even local network metrics performed poorly. For example, lattice layouts showed poorer local metric performance, as did more restrictive contact definitions (e.g. 1m or 10m). As discussed in the main text, less precise contact definitions were often associated with surprising increases in local metric performance, particularly in clustered populations (Figures A.7 and A.9). The results were robust to changes in the movement model when home range size was held constant (Figures A.8 and A.10).

Network-level metrics

The network-level metrics, density, proportion isolates (evaluated as its complement, proportion connected), and modularity performed intermediately, compared to the node-level metrics (Figures A.11-A.16). Density was apparently robust to the proportion of the population sampled, but this was likely driven by the random sampling protocol we used (Figures A.11-A.12). When using less precise contact definitions, both density and proportion isolates generally demonstrated increased positive percent error, demonstrating that these contact definitions result in overly connected networks compared to the complete network (Figures A.11 and A.13). Results were largely robust to BCRW parameterization when home range was held constant. However, there was some variation in results from small home range parameterizations when comparing 100m complete networks to spatial overlap sample networks, with instances of negative percent error for both density and proportion isolates (Figures A.12 and A.14). This is likely because the 100m contact definition was fairly large relative to the step length distribution for these parameter sets. The variation in results then likely reflects the effect of “thresholding,” where at a certain degree of sampling or less precise contact definition, networks are no longer over-connected relative to the complete network.

Modularity results were much more variable than those for other metrics, though these results were generally robust to the choice of community-finding algorithm or the use of weighted versus unweighted network (illustrative results from unweighted and weighted walktrap algorithms are shown in Figures A.15-A.16). For modularity, a positive percent error indicates higher modularity in the sample networks, suggesting stronger community structure in the sample networks than in the complete networks. When estimated with a weighted walktrap algorithm, modularity estimates tended to have

higher positive percent error than the other modularity estimates, but these differences were negligible (Figure A.16).

Of note, single cluster spatial configurations consistently showed high positive percent error regardless of the community finding algorithm, suggesting these configurations are highly sensitive to overestimates of modularity. For all spatial configurations, exceptionally low sampling effort consistently resulted in strongly negative percent error, demonstrating that the sample networks showed lower modularity and weaker community structure than complete networks. Spatial configuration interacted with home range size in determining the effect of telemetry sampling effort on modularity; for example, large home range simulations showed highest positive percent error for random configurations, but small home range simulations showed highest positive percent error for lattice configurations (Figures A.15-A.16). In addition, there was often a sharp threshold in the transition from positive to negative percent error which was generally dependent on the combination of both proportion of the population sampled and the frequency of sampling, but the sampling effort combination describing this threshold varied across spatial configurations.

When comparing complete networks to spatial overlap sample networks, modularity was generally underestimated except for some cases of lattice configuration (Figures A.15-A.16). These results suggest that the precise effects of telemetry sampling on modularity estimation are highly variable across both spatial configuration and sampling effort. The sensitivity of modularity to sampling effort merits further investigation, perhaps with networks derived from empirical systems (e.g. Sah et al., 2019), in addition to simulated data.

For density, proportion isolates (evaluated as proportion connected), and modularity, all heat maps indicate percent error of mean metrics between complete and sample networks, with blue indicating negative percent error (low connectedness or modularity of sample network relative to the complete network) and red indicating positive percent error (high connectedness or modularity of sample network relative to the complete network). All other abbreviations in the plots below are consistent with previous node-level metric plots.

Emergent contact rates

Contact rates were an emergent property of simulations and spatial configurations. Lattice layout populations, which best represented territorial, infrequently interacting populations had the fewest contacts, and single cluster populations had the most (Table A.2). The mean duration of contacts between any given pair was consistent across spatial configuration, but shorter contact durations were observed for large home range simulations. This was likely a result of the larger step length distribution scaling parameter used in the large home range parameter set.

A.3 Figures

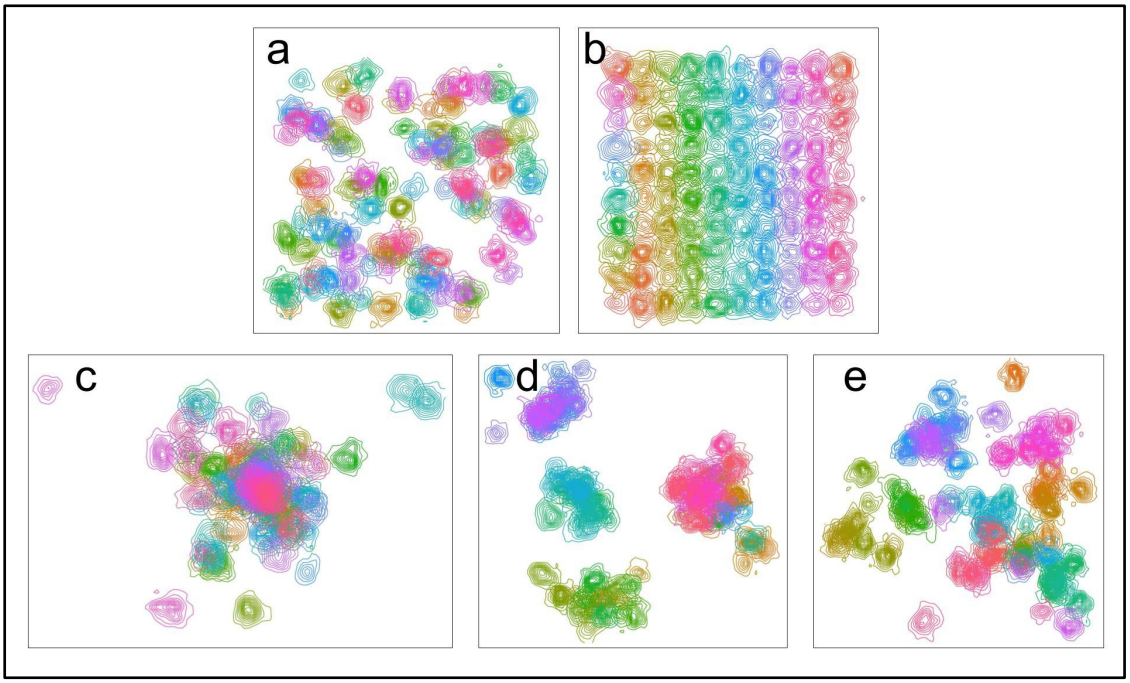


Figure A.1: Examples of different spatial configurations used for simulations. Each panel represents densities of movements for 100 individuals in a single simulation. Panel (a) shows a random spatial configuration; panel (b) shows a lattice configuration; panels (c-d) show single, five, and ten cluster spatial configurations, respectively. Individual groups of home ranges in clustered configurations often overlap, so each individual group or cluster may not be readily identifiable.

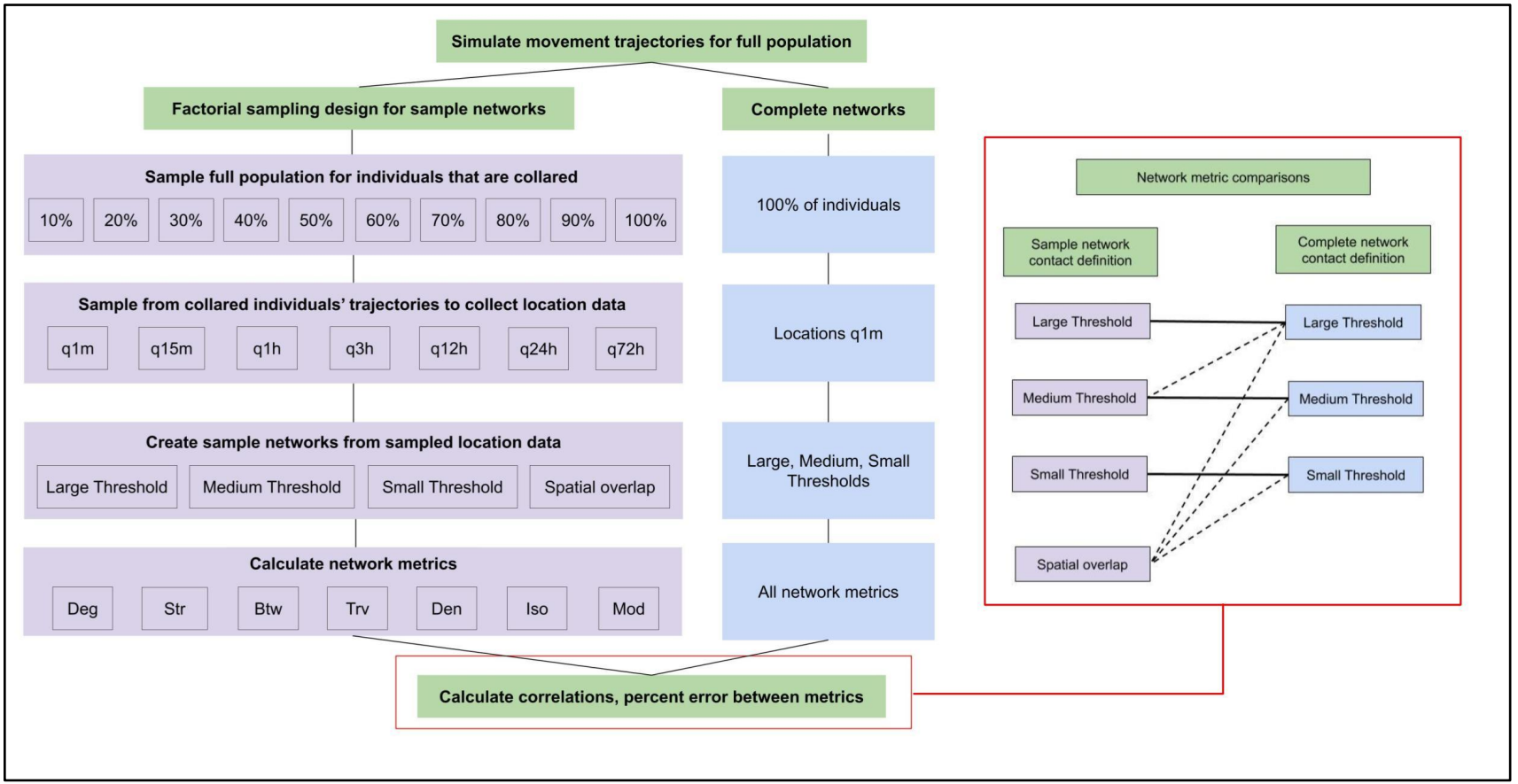

Simulate movement trajectories for full population
Factorial sampling design for sample networks
Complete networks
Sample full population for individuals that are collared
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
100% of individuals
Sample from collared individuals' trajectories to collect location data
q1m
q15m
q1h
q3h
q12h
q24h
q72h
Locations q1m
Create sample networks from sampled location data
Large Threshold
Medium Threshold
Small Threshold
Spatial overlap
Large, Medium, Small Thresholds
Calculate network metrics
Deg
Str
Btw
Trv
Den
Iso
Mod
All network metrics
Calculate correlations, percent error between metrics
Network metric comparisons
Sample network contact definition
Complete network contact definition
Large Threshold
Large Threshold
Medium Threshold
Medium Threshold
Small Threshold
Small Threshold
Spatial overlap

Figure A.2: Expanded flow chart of simulation methods and network comparisons (also see Figure 2.1 in main text). The main flow chart shows variations in sampling and contact definitions, and all network metrics calculated. For sampling locations from trajectories, “q1m” means “every 1 minute” and so on. For network metrics, Deg = Degree, Str = Strength, Btw = Betweenness, Trv = Transitivity, Den = Density, Iso = Proportion Isolates, and Mod = Modularity. Green boxes represent major steps in the methods, purple boxes represent treatments or steps for sample networks, and blue boxes the treatments or steps for complete networks. The inset box expands to show what comparisons were made between sample and complete network metrics. In the inset, solid lines represent strict comparisons, in which the contact definition in the complete and sample networks is the same. Dashed lines represent less precise comparisons, in which the sample network has a less strict contact definition than the complete network.

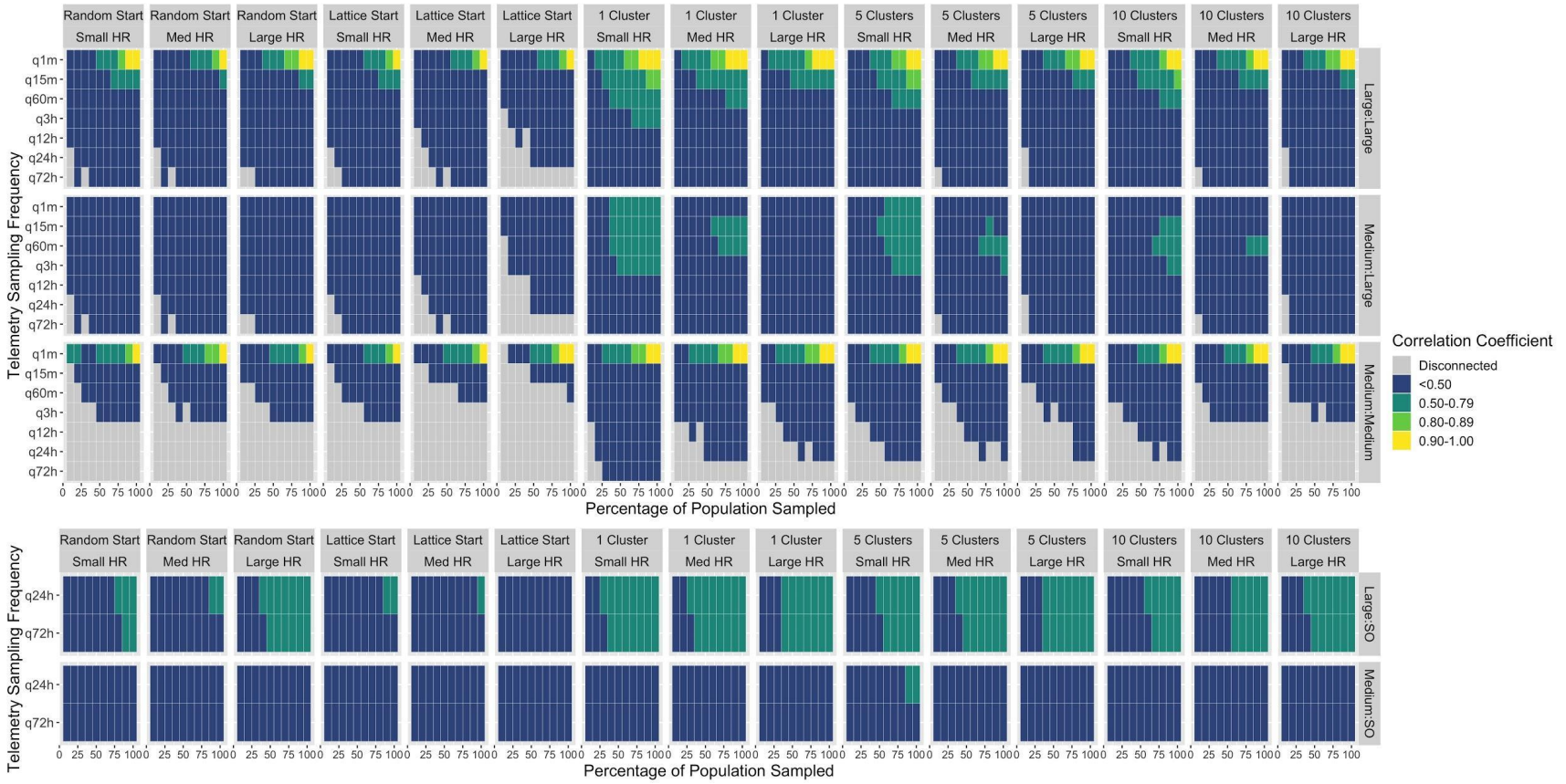


Figure A.3: Betweenness: Results are shown across three movement model parameter sets (small, medium, and large home

ranges; Small HR, Med HR, and Large HR, respectively). The top panel represents network comparisons where both complete and sample networks were generated with space-time contact definitions; the bottom panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

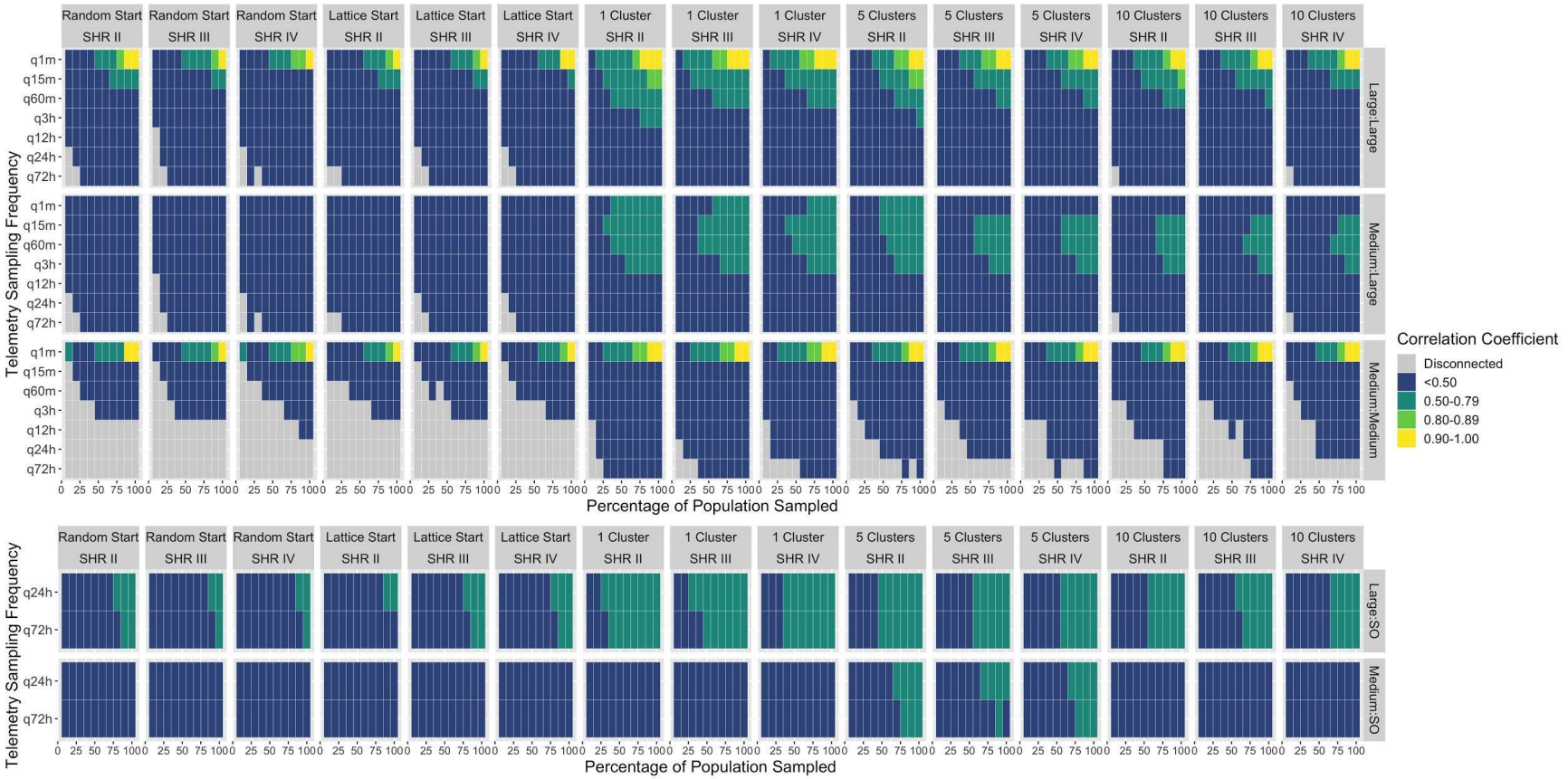

Random Start SHR II
Random Start SHR III
Random Start SHR IV
Lattice Start SHR II
Lattice Start SHR III
Lattice Start SHR IV
1 Cluster SHR II
1 Cluster SHR III
1 Cluster SHR IV
5 Clusters SHR II
5 Clusters SHR III
5 Clusters SHR IV
10 Clusters SHR II
10 Clusters SHR III
10 Clusters SHR IV
Telemetry Sampling Frequency
q1m
q15m
q60m
q3h
q12h
q24h
q72h
Large:Large
Medium:Large
Medium:Medium
Large:SO
Medium:SO
0 25 50 75 100
Percentage of Population Sampled
Correlation Coefficient
Disconnected
<0.50
0.50-0.79
0.80-0.89
0.90-1.00

Figure A.4: Betweenness: Results are shown across three movement model parameter sets (small home ranges II-IV; SHR II-SHR IV, respectively). The top panel represents network comparisons where both complete and sample networks were generated with space-time contact definitions; the bottom panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

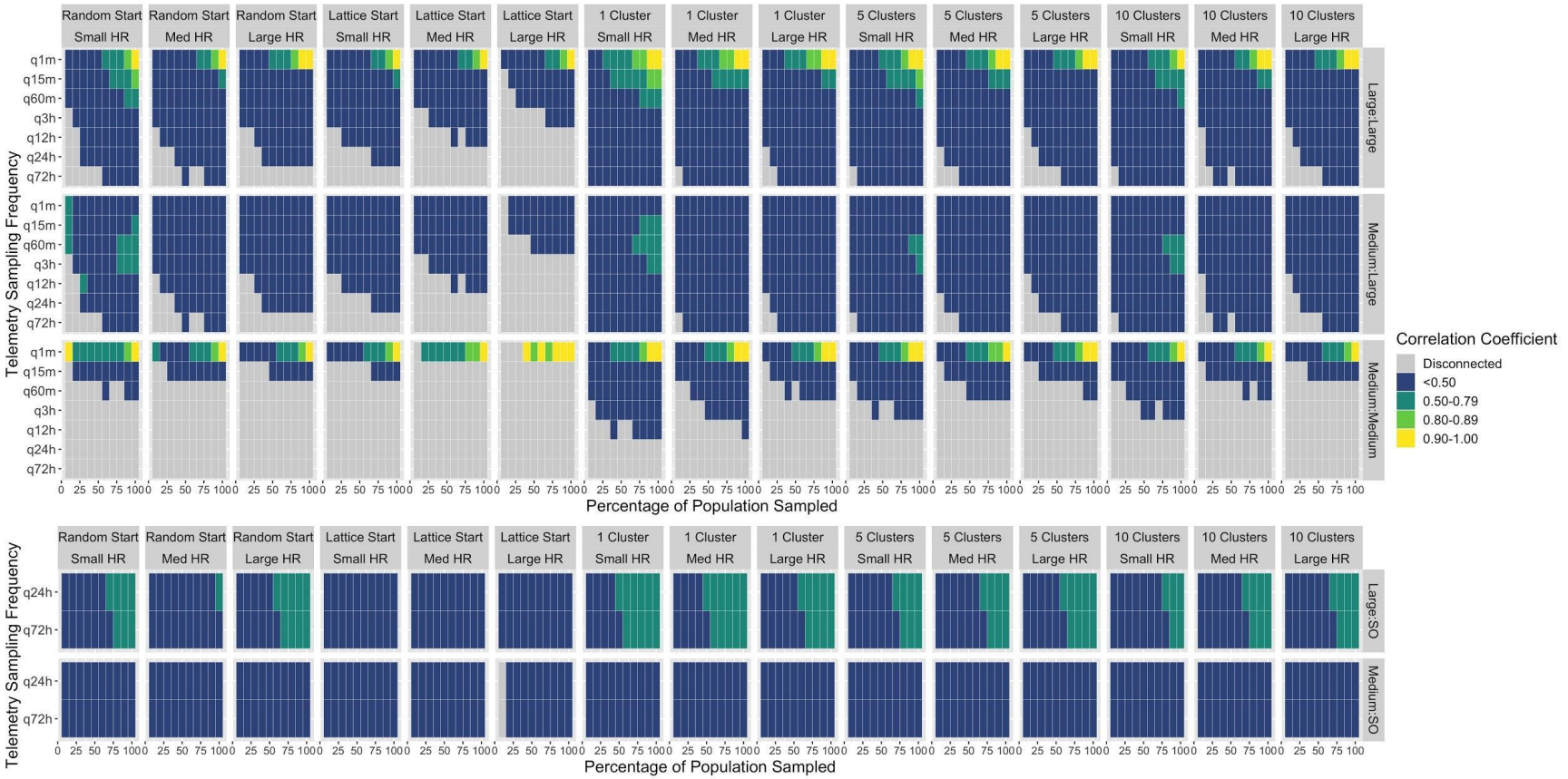


Figure A.5: Transitivity: Results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). The top panel represents network comparisons where both complete and sample

networks were generated with space-time contact definitions; the bottom panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

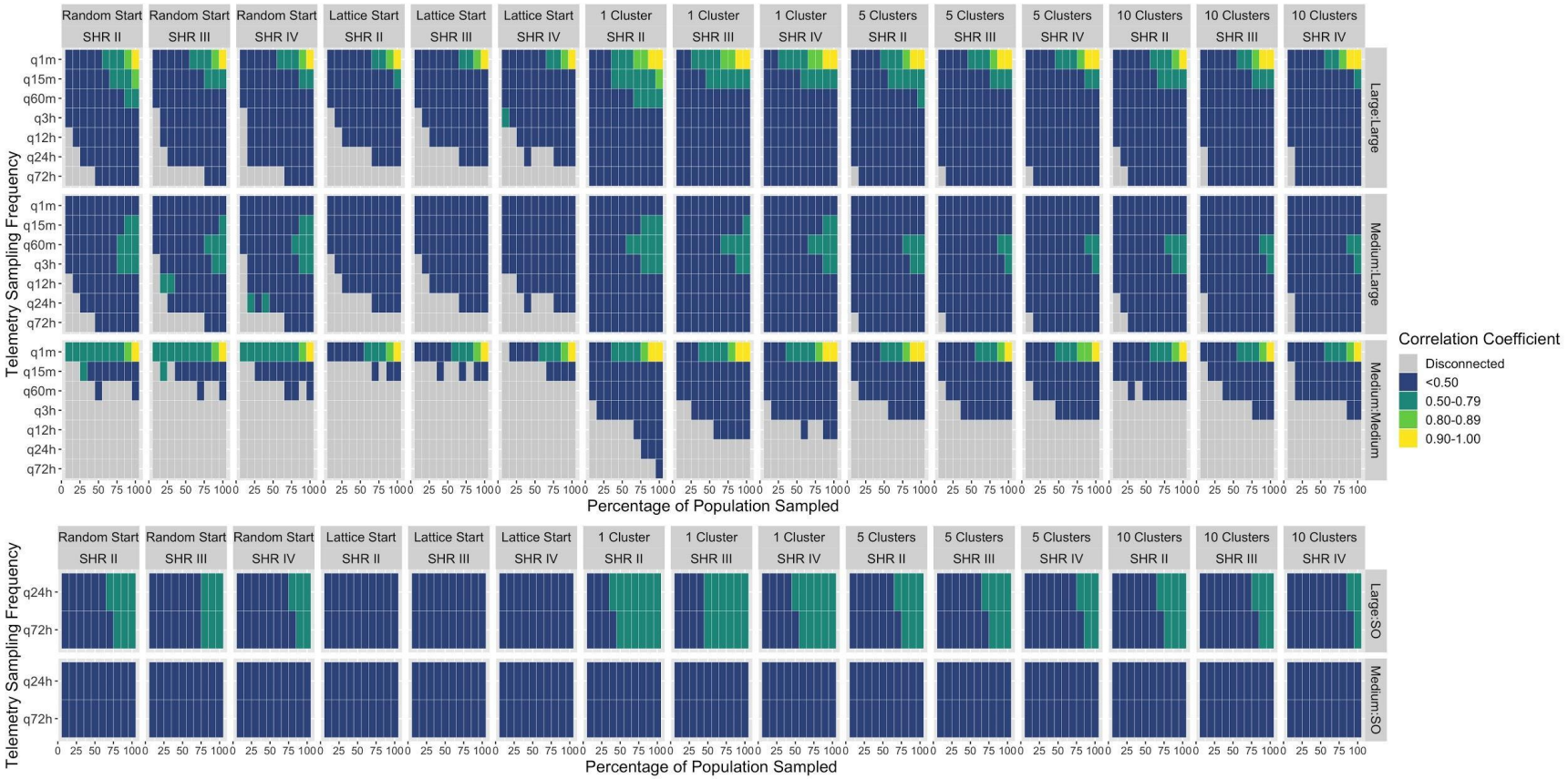


Figure A.6: Transitivity: Results are shown across three movement model parameter sets (small home ranges II-IV; SHR II-SHR IV, respectively).

The top panel represents network comparisons where both complete and sample networks were generated with space-time contact definitions; the bottom panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

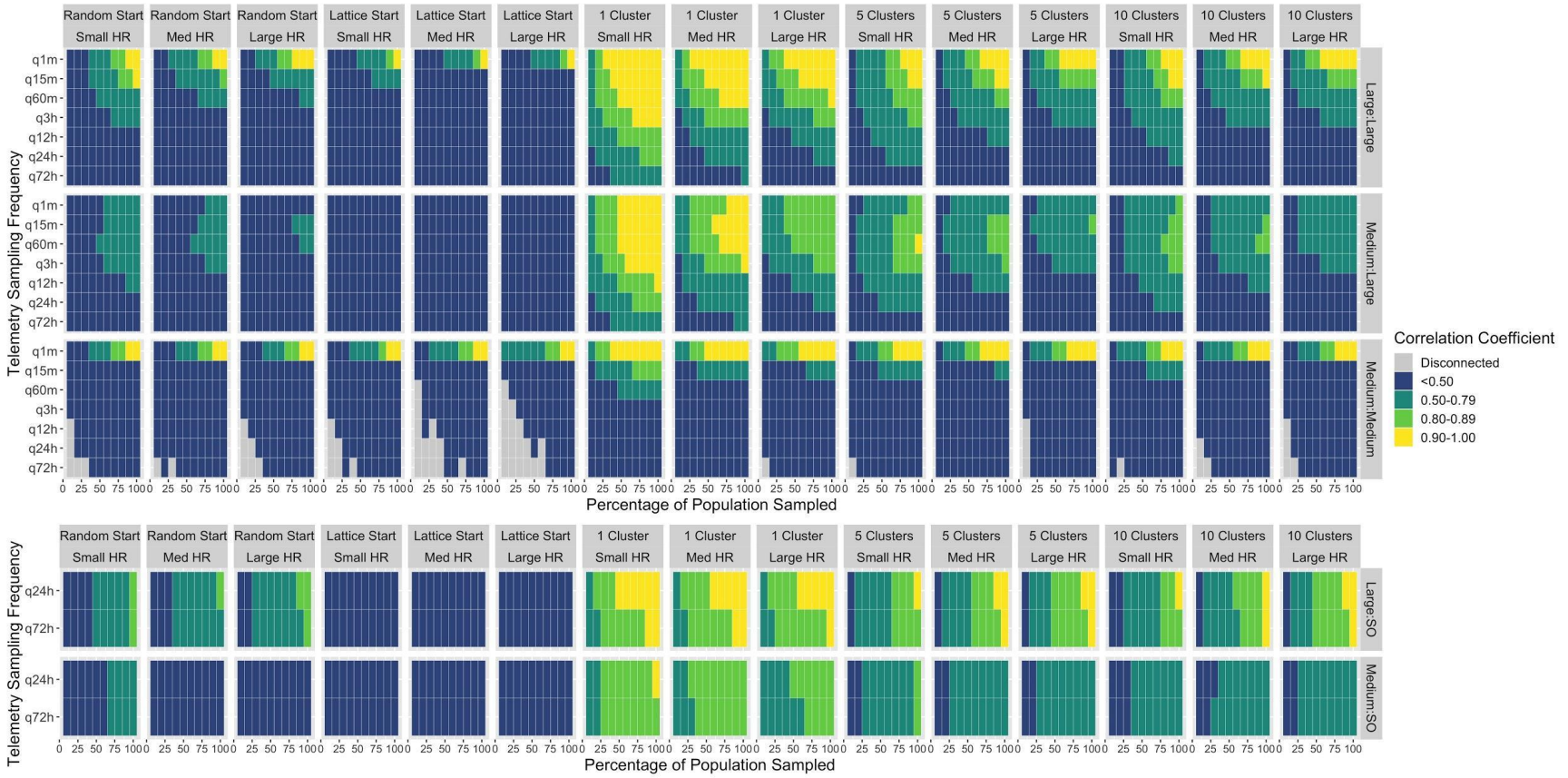


Figure A.7: Degree: Results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). The top panel represents network comparisons where both complete and sample

networks were generated with space-time contact definitions; the bottom panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

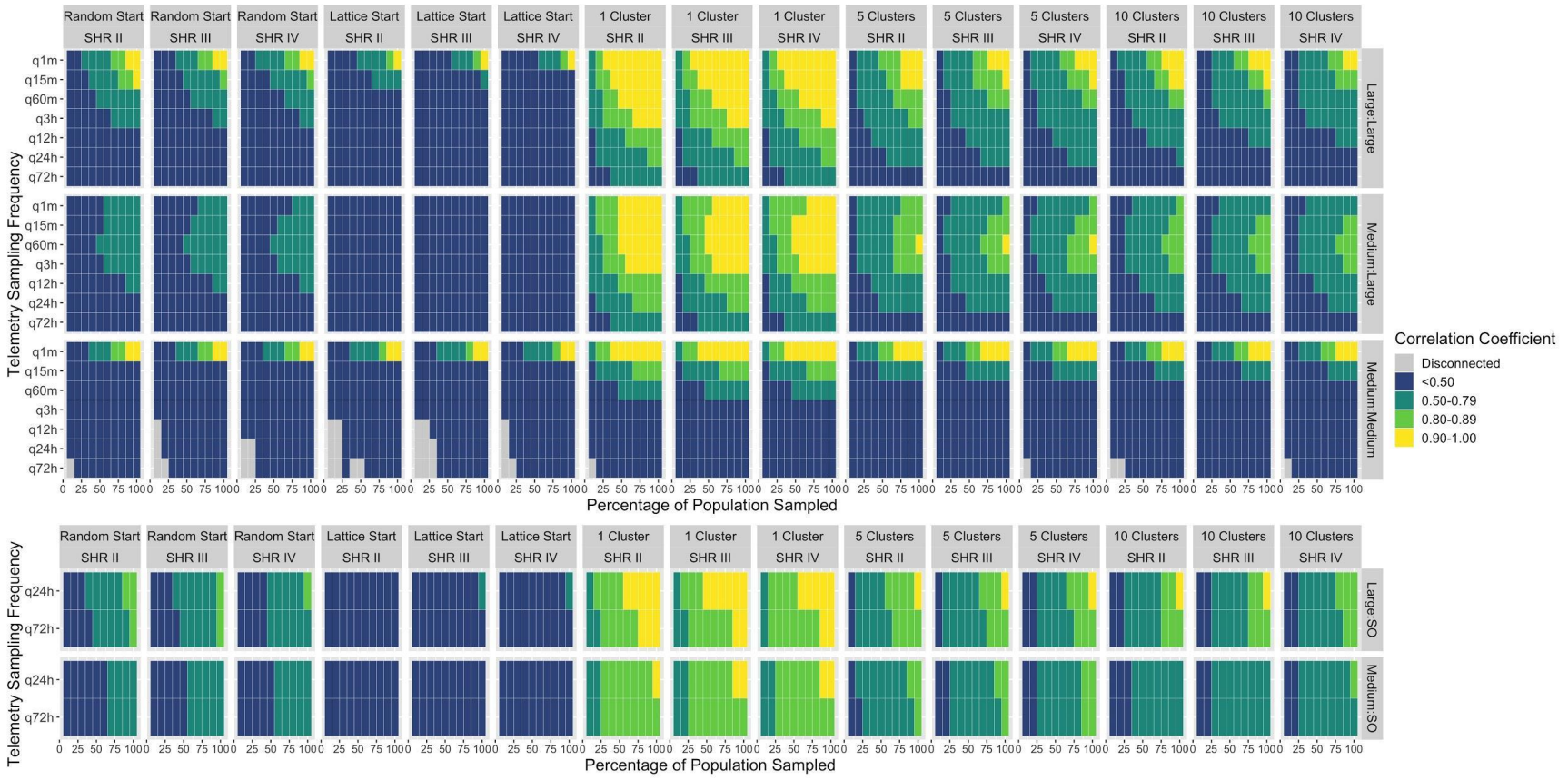


Figure A.8: Degree: Results are shown across three movement model parameter sets (small home ranges II-IV; SHR II-SHR IV, respectively). The top panel represents network comparisons where both complete and sample networks were generated with space-

time contact definitions; the bottom panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

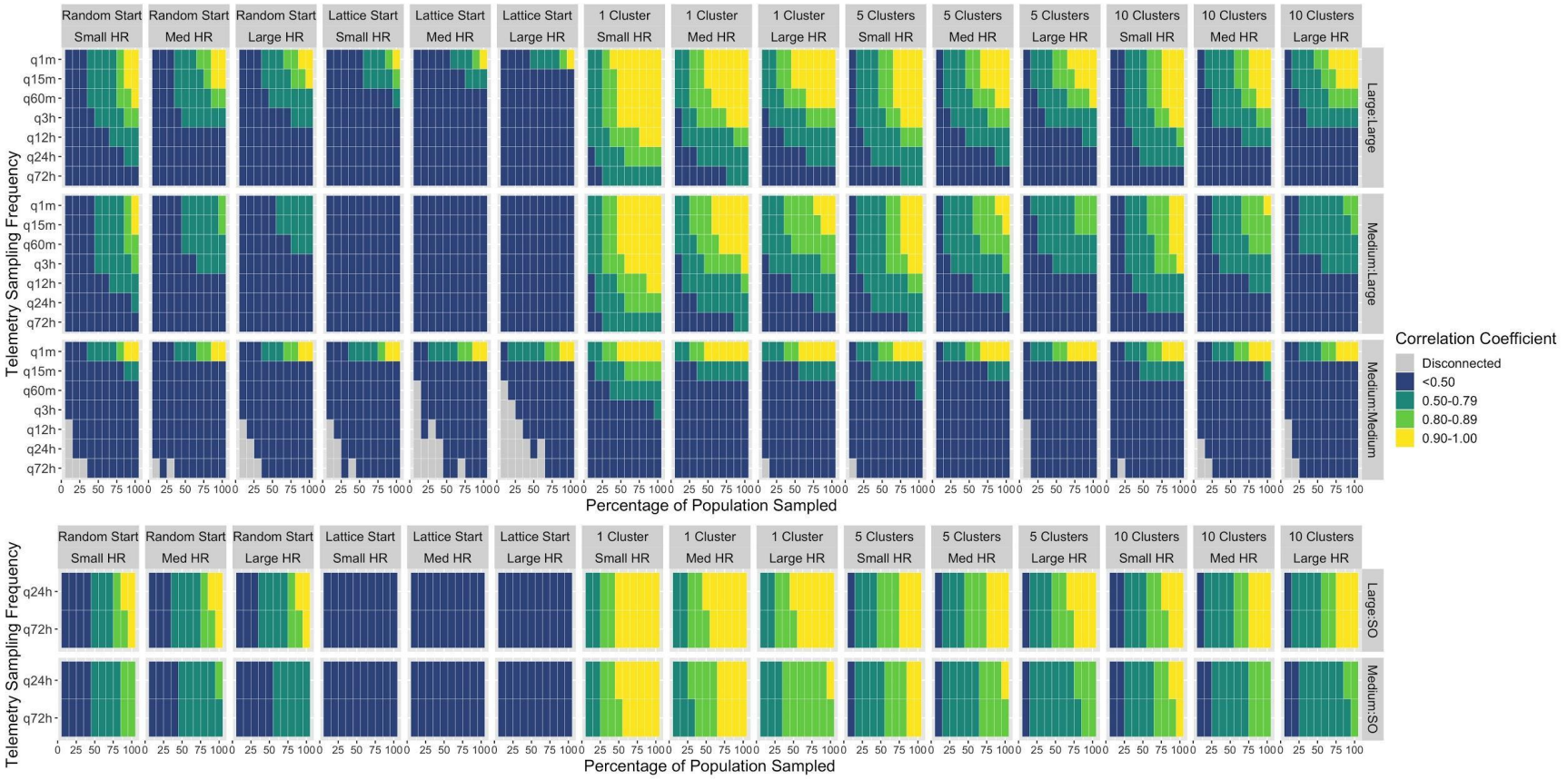


Figure A.9: Strength: Results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). The top panel represents network comparisons where both complete and sample

networks were generated with space-time contact definitions; the bottom panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

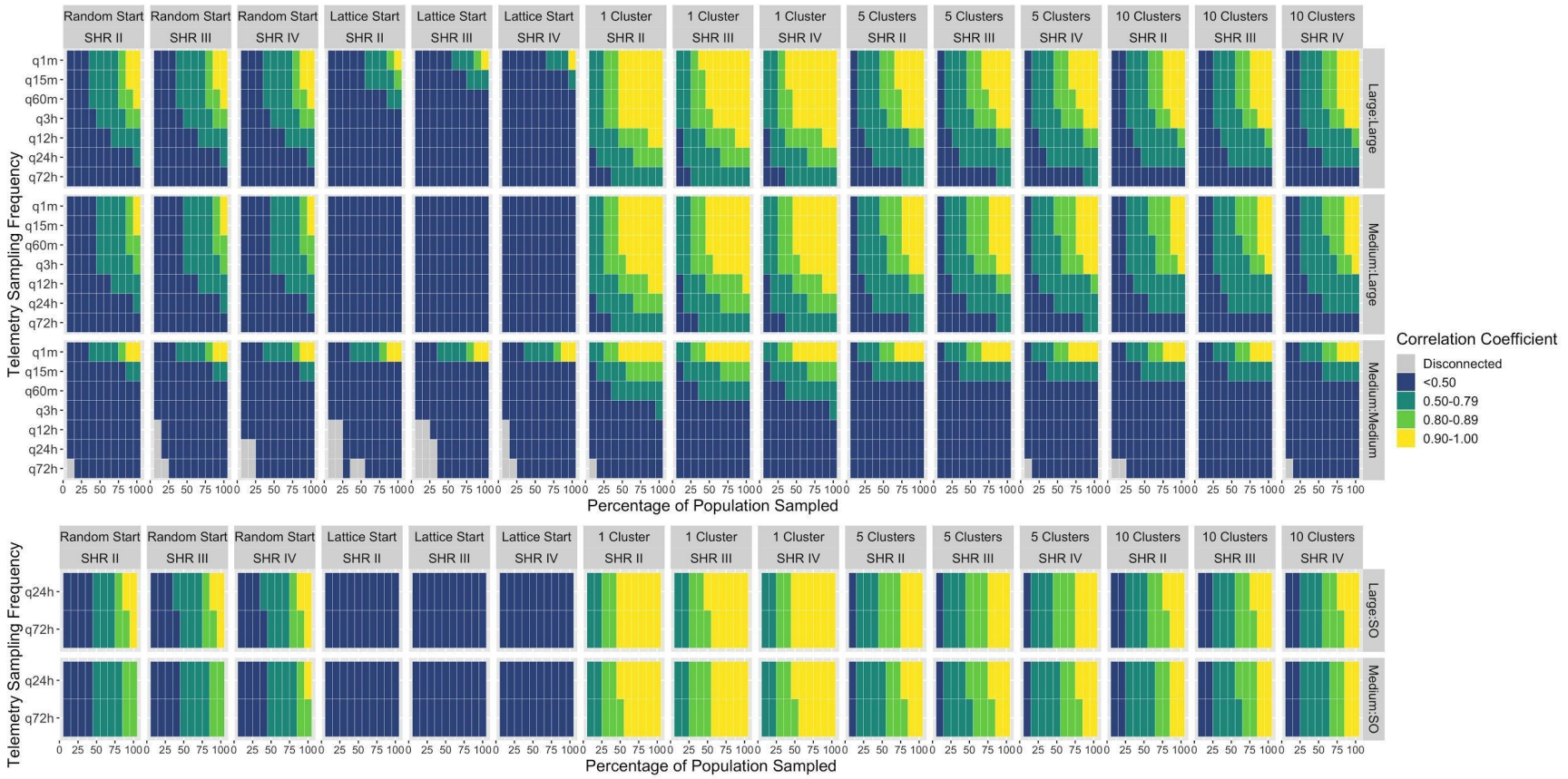


Figure A.10: Strength: Results are shown across three movement model parameter sets (small home ranges II-IV; SHR II-SHR IV, respectively). The top panel represents network comparisons where both complete and sample networks were generated with space-

time contact definitions; the bottom panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

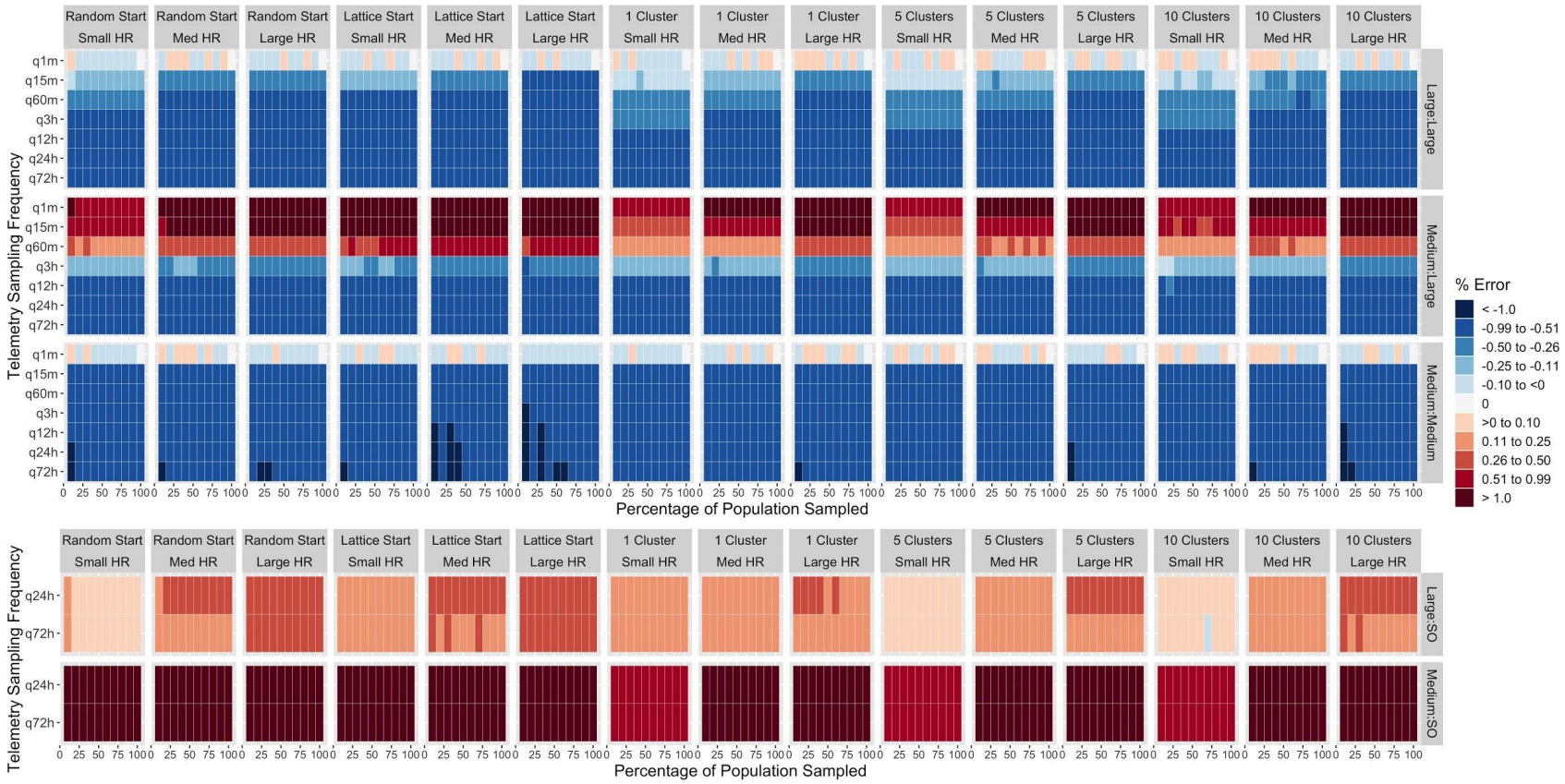


Figure A.11: Density: Results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). The top panel represents network comparisons where both complete and sample

networks were generated with space-time contact definitions; the bottom panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

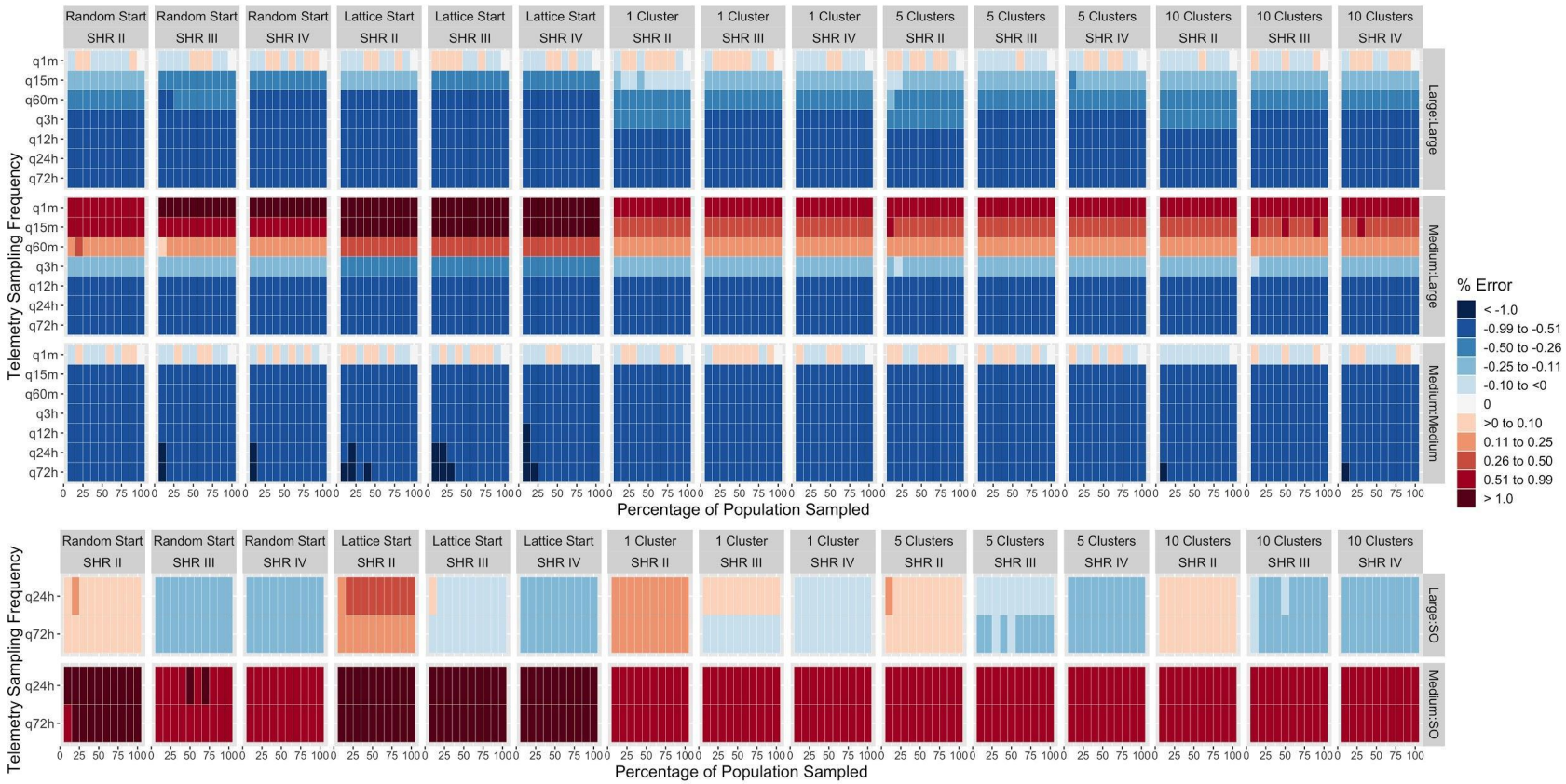


Figure A.12: Density: Results are shown across three movement model parameter sets (small home ranges II-IV; SHR II-SHR IV, respectively). The top panel represents network comparisons where both complete and sample networks were generated with space-

time contact definitions; the bottom panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

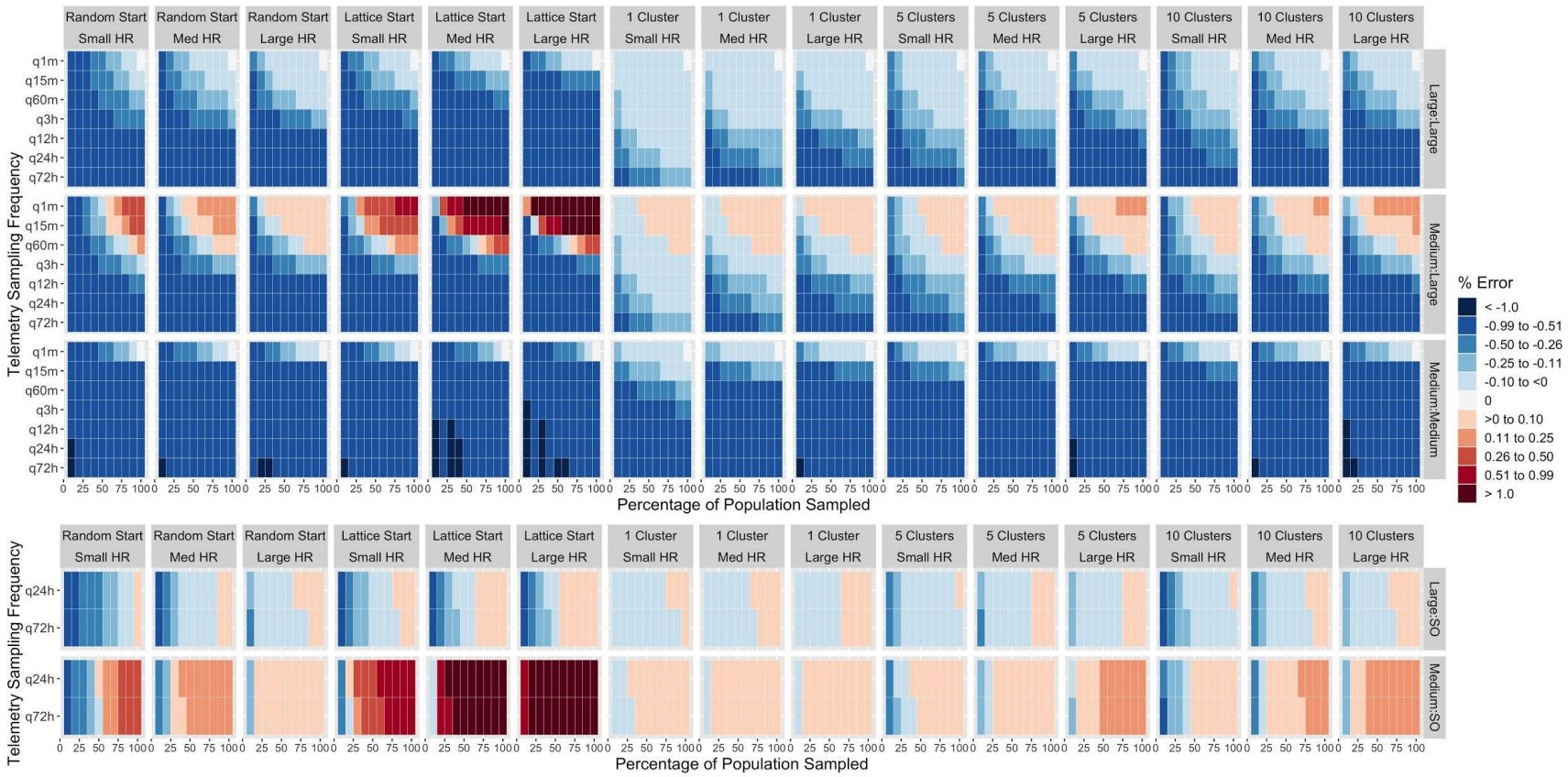


Figure A.13: Proportion isolates (evaluated as proportion connected): Results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). The top panel represents network

comparisons where both complete and sample networks were generated with space-time contact definitions; the bottom panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

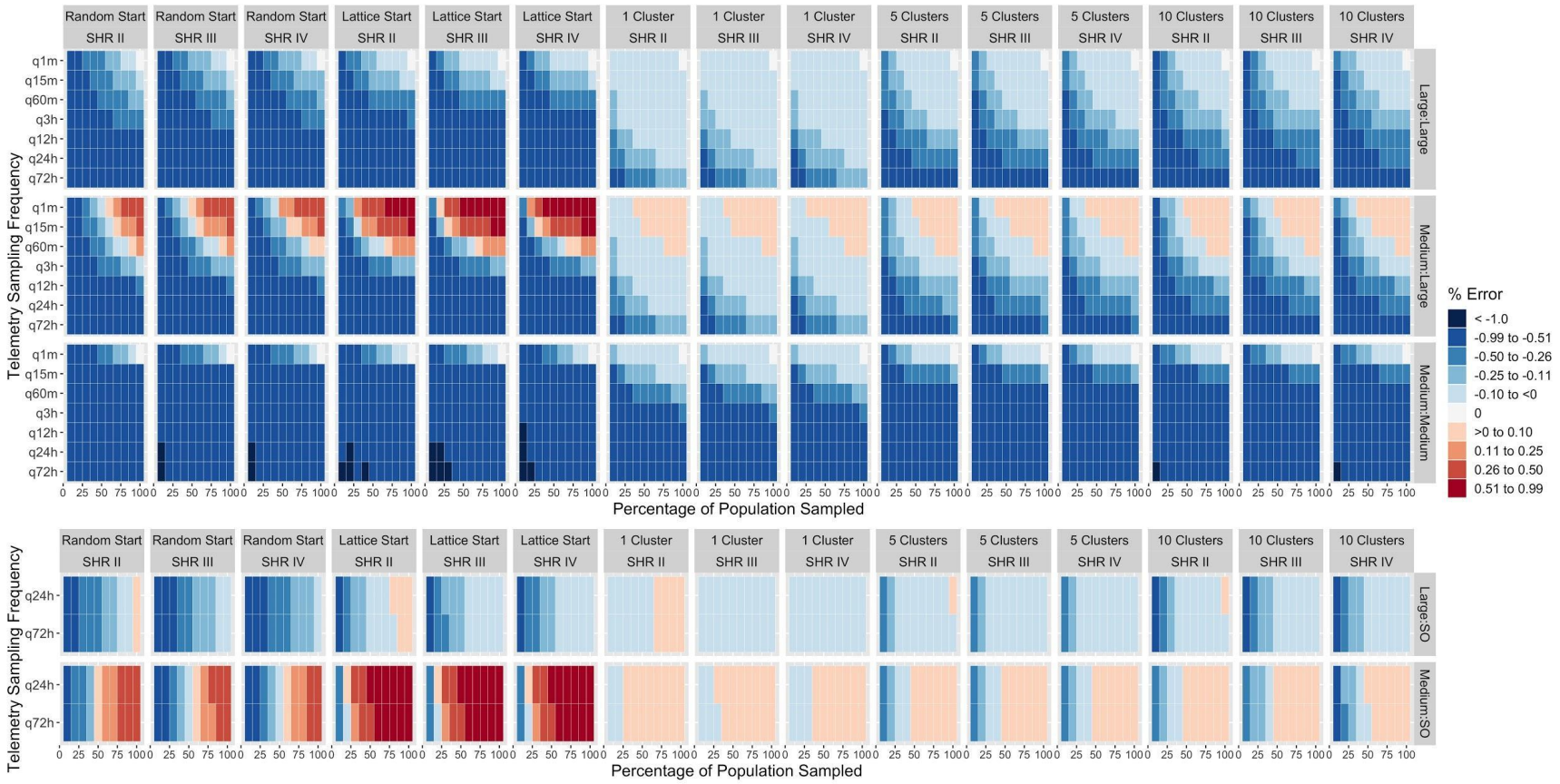


Figure A.14: Proportion isolates (evaluated as proportion connected): Results are shown across three movement model parameter sets (small home ranges II-IV; SHR II-SHR IV, respectively). The top panel represents network comparisons where both complete

and sample networks were generated with space-time contact definitions; the bottom panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

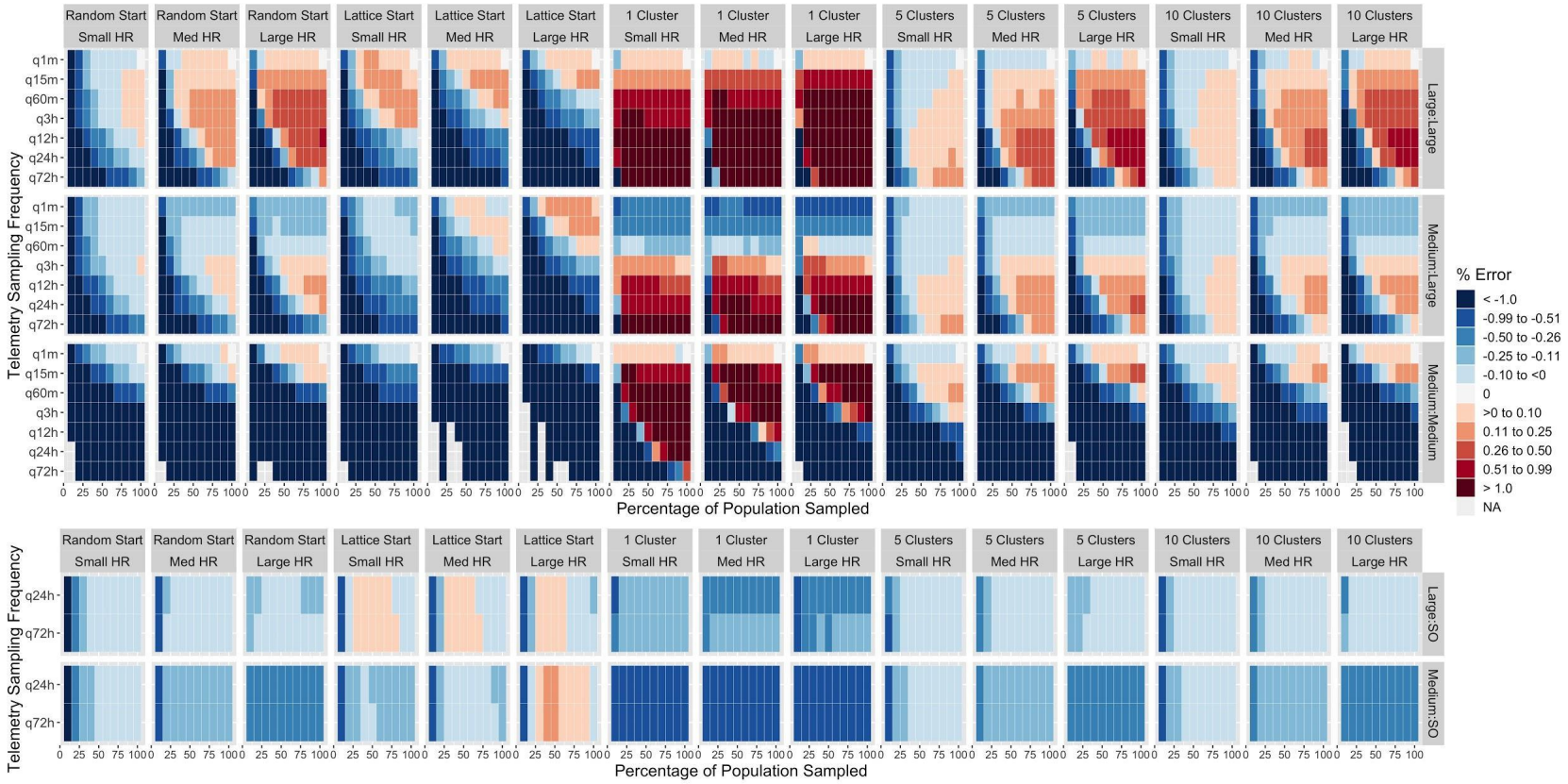


Figure A.15: Modularity from unweighted walktrap community-finding algorithm: Results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). The top panel represents network comparisons where both complete and sample networks were generated with space-time contact definitions; the bottom

panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

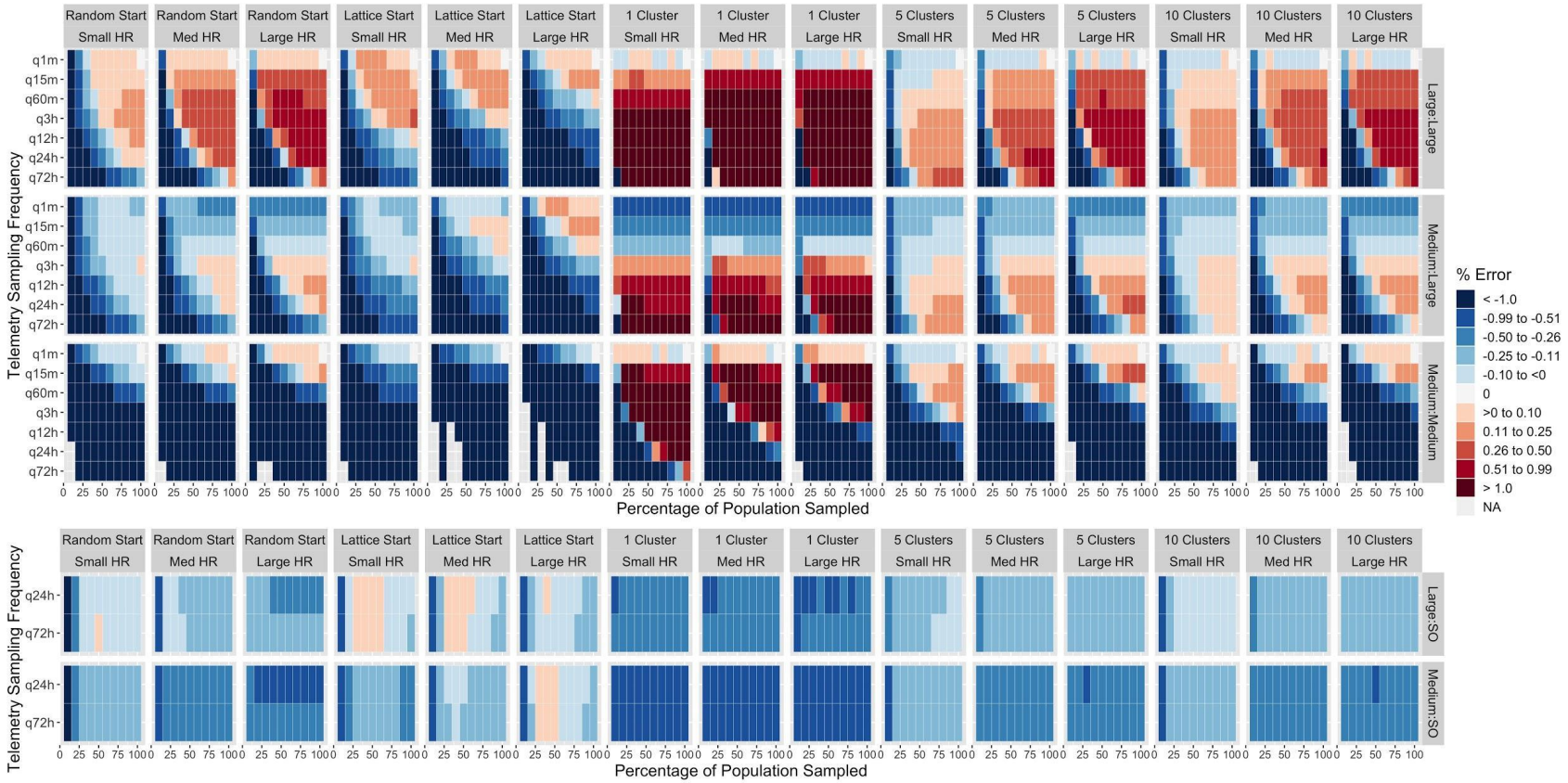


Figure A.16: Modularity from weighted walktrap community-finding algorithm: As with figure A.13, results are shown across three movement model parameter sets (small, medium, and large home ranges; Small HR, Med HR, and Large HR, respectively). The top

panel represents network comparisons where both complete and sample networks were generated with space-time contact definitions; the bottom panel represents results from network comparisons where the sample network was generated with a spatial overlap contact definition.

A.4. Tables

Table A.1. Parameter values used for the BCRW movement model and corresponding average home range sizes. The parameter *h* gives the scaling of the step length distribution (in scale-free units); ρ is the strength of the turning angle correlation; *b* is the bias strength; *c* is the bias direction. Average home range (HR) areas are based on a sample of 50,000 simulated movement trajectories for each parameter set. Average HR areas were determined using the 95% bivariate normal kernel density estimates with a subset of simulated locations per trajectory (one location every 24 hours), and are given in 1000 units^2 (akin to kilometers). Standard deviations for home range areas are given in parentheses. The four "small home range" parameter sets maintain roughly constant average home range areas while varying simulation parameters. Because home range size was an emergent property of the movement simulations, home ranges could not be exactly the same between small home range parameter sets, but they are comparable.

Parameter Set	***h***	ρ	***b***	***c***	**Avg HR Area (sd) (1000 units^2)**
Large Home Ranges	60	0.8	0.01	0.3	40.0 (5.7)
Medium Home Ranges	34	0.8	0.01	0.3	17.0 (2.4)
Small Home Ranges I	15	0.8	0.01	0.3	5.0 (0.7)
Small Home Ranges II	20	0.6	0.0153	0.3	5.2 (0.7)
Small Home Ranges III	34	0.8	0.02	0.3	5.2 (0.7)
Small Home Ranges IV	60	0.8	0.0314	0.3	5.2 (0.7)

Table A.2. Contact rates in complete networks by spatial configuration and home range size when using a space-time contact definition with 100m distance threshold. Mean

contacts records the average number of raw contacts, where sequential contacts are considered separate contact events. SD contacts is the standard deviation for the mean number of contacts. Mean contacts/dyad records the mean number of contacts for all pairs with contacts. Mean duration/dyad records the mean duration (in minutes) of contacts for all pairs with contacts. Duration was determined by the number of continuous time steps over which a contact occurred between a given pair.

Spatial Configuration	HR Size	Mean Contacts	SD Contacts	Mean Contacts/Dyad	Mean Duration/Dyad
Random	Small	21807.2	4382.3	245.9	6.0
	Medium	20943.2	2312.7	77.0	2.7
	Large	20287.7	1608.3	36.1	1.6
Lattice	Small	11970.3	12431.5	43.4	6.0
	Medium	3458.2	3581.2	13.4	2.7
	Large	1472.8	1530.0	6.2	1.6
1 Cluster	Small	879776.2	350159.3	349.8	5.9
	Medium	248177.5	101418.2	104.0	2.7
	Large	107537.5	47064.4	46.6	1.7
5 Clusters	Small	193512.5	78318.7	347.2	6.0
	Medium	65233.2	24954.9	98.4	2.7
	Large	33855.2	12061.4	41.8	1.6
10 Clusters	Small	106230.9	40553.5	339.0	5.9
	Medium	43629.9	15311.2	95.0	2.7
	Large	26098.3	7380.8	39.9	1.6

Appendix B. Supporting information for Chapter 3

B.1 Materials and Methods

Phyloscanner

For feline immunodeficiency virus (FIV) transmission inference, we used Phyloscanner to generate a main transmission network (see Chapter 3 main text). To test sensitivity of our results to variations in Phyloscanner results, we also generated two *summary FIV networks*, varying the degree of window overlap in the first step of Phyloscanner analysis. With Phyloscanner step one set to 25bp overlap, we generated four additional FIV transmission networks, but kept only those edges that were found in at least two of these four networks. We repeated this process with Phyloscanner step one set to 0bp overlap, again keeping only those edges found in at least two of four resulting transmission networks. We then removed any edges between individuals that were known not to be alive at the same time, based on panther monitoring data.This process resulted in our main FIV network, and the two additional summary FIV networks.

Statistical analysis of FIV transmission networks

To determine factors structuring FIV transmission networks, we considered a suite of network structural variables (also considered dyad-dependent variables; see Chapter 3 main text) and our key population predictors of interest (dyad-independent variables). Among dyad-independent variables, an ERGM approach can include both node-level variables (e.g., node age or sex) and edge-level variables (e.g., genetic distance). Node-level variables can be evaluated as continuous or categorical variables (e.g., are males involved in more transmission events?), but also for difference or

matching relationships (e.g. do more transmission events occur between male/female dyads or same-sex dyads?). In keeping with ERGM analysis terminology, categorical node variables are referred to as *node factor*, continuous as *node covariate*, matching relationships as *node mixing*, and continuous edge variables as *edge* covariate (Morris et al., 2008; Silk & Fisher, 2017).

Among the dyad-independent variables we examined in our ERGM analysis (see Chapter 3 main text), we included several spatial variables. Home range centroids were used in the generation of several of these variables and were generated by first estimating 95% minimum convex polygon (MCP) home ranges for telemetry-monitored panthers. To estimate these MCP home ranges, we used only those telemetry data collected in the 12 months after an individual's initial capture, and only for those individuals with at least 30 relocations in that time period. MCPs were generated with the *adehabitat* package in R (Calenge, 2006), and centroids were calculated using the *rgeos* package (Bivand & Rundel, 2018). Our priority was to capture the range of panther-occupied landscape, so we also incorporated point locations for individuals without at least 30 telemetry relocations. For these point locations, we prioritized using the location of an individual at capture; if this information was not available, we instead used the telemetry relocation captured closest to the date of capture. For the main FIV network, this approach resulted in 11 point locations from MCP centroids, 7 were capture locations, and 1 was a relocation closest to capture date. Hereafter, these locations are referred to as *centroids*.

Major roadways have been shown to alter puma movement in North America (Wheeler et al., 2010), so our ERGM analysis included a node-matching variable for location of panthers' centroids north versus south of the major I-75 freeway, defined as

latitude 26.15. We further hypothesized that panthers closer to urban areas would face greater competition for resources and therefore be involved in more transmission events due to increased fighting behaviors. We therefore also examined a node covariate term for distance to nearest urban area (in km). We used the “near table” function in ArcGIS to determine the distance of each centroid to the closest urban area edge, defining urban areas using the USA Urban Areas layer publicly available in ArcGIS (Census 2010 Urbanized Areas and Clusters; Esri, National Atlas of the United States, United States Geological Survey, Department of Commerce, Census Bureau-Geography Division). We included pairwise geographic distances between panthers using distances between centroids (in km), and log-transformed this edge covariate for ERGM analysis. Lastly, we hypothesized that panthers with overlapping home ranges would be more likely to transmit to each other, so we included a spatial overlap edge covariate based on the pairwise utilization distribution overlap indices of 95% home range kernels (Fieberg et al., 2005), using the adehabitat package in R (Calenge, 2006).

For our pairwise relatedness variable (see Chapter 3 main text), we used previously collected microsatellite data (Van De Kerk et al., 2019). One individual in the FIV transmission networks lacked microsatellite data but had known pedigree sibling relationships with other individuals in the transmission networks (Johnson et al., 2010). In order to preserve available data (ERGMs cannot operate with missing data), we interpolated sibling relatedness values for this individual using mean relatedness values from other known sibling pairs. Non-sibling relationships for this individual were conservatively interpolated at population mean relatedness, functionally assuming no relatedness.

Goodness of fit for ERGMs was performed using the *ergm* package in R (Hunter et al., 2008). We evaluated fit for degree distribution, geodesic distance, and triad census ("degree", "distance," and "triadcensus" terms, respectively). Model selection included evaluation of AIC and improvement to goodness of fit, using these terms.

Panther centroid simulation

Because pairwise geographic distances were found to be significant in ERGM analysis of FIV transmission networks (see Chapter 3 main text), in order to simulate potential transmission pathways among panthers, we also had to simulate these geographic pairwise distances. We did so by simulating home range centroids based on the empirical panther population. Simulated centroids were generated by plotting the observed MCP centroids from 2002-2004; the polygon encompassing these centroids was then split into 70 quadrats, and simulated centroids were randomly drawn from these quadrats, according the proportion of the observed population that was found within each quadrat. This functionally kept much of the heterogeneity in distribution of home range centroids across panther habitat. Pairwise distances were then calculated between simulated centroids and log transformed, as was done to calculate pairwise distances for the original ERGM analysis (see Chapter 3 main text).

Overlap-based networks

To compare FeLV transmission predictions from FIV-based networks against simpler model types, we generated spatial overlap-based networks, on which we also simulated FeLV transmission (see Chapter 3 main text). To do so, we first generated networks of utilization distribution overlap index (UDOI) spatial overlap (with 95% kernel)

among collared panthers in each year from 2002-2004 (three total networks; Calenge, 2006), considering an edge to exist if UDOI was greater than 1. We calculated the degree distribution from each of these networks, and fit a negative binomial distribution to the degree distribution for each year using the *fitdistrplus* package in R (Delignette-Muller et al., 2015). We took the means of the parameter values for the resulting three negative binomial distributions to create a single "summary" negative binomial distribution. We simulated new overlap-based networks using this summary negative binomial to draw degree distributions, and then used simulated annealing (Handcock et al., 2008; Reynolds et al., 2015) to generate random networks based on the drawn distributions. Because simulated overlap-based networks were informed by degree distributions, they were not spatially explicit, but represent data typically available in long-term wildlife monitoring studies.

Gillespie algorithm

We also compared FeLV transmission predictions from FIV-based networks to predictions from a homogeneous mixing model: in this case, a Gillespie algorithm (stochastic, time-to-event model). This model was specified in order to align with the chain binomial network model specifications, resulting in the following rate functions:

Susceptibles infection rate = $\omega*\beta*Net\_dens*S(I_p + C*I_r)$

Vaccinates infection rate = $\omega*\beta*(1\text{-}ve)*Net\_dens*V(I_p + C*I_r)$

Progressives mortality rate = μI_p

Recovery rate = $\mu*K*I_r$

Respawn rate = νD

Vaccination rate (after one simulation year) = τ*S/N

In the above rate functions, N is the total population, S is susceptibles, I_p is progressively infected individuals, I_r is regressively infected individuals, V is vaccinated individuals, and D marks unoccupied territories after death of the prior occupant and prior to "respawning" (as in network models). All other parameters are as in Table 3.1 of the main text for Chapter 3. Of special note, Net_dens represents the density of networks from network transmission models, and here functions as a population size-scaled contact rate. This contact rate is further modified by the weekly probability of contact, ω, as was done in network models (see Chapter 3 main text). The vaccination rate is scaled by population size as vaccination was applied to the whole population in both network and Gillespie models, but only susceptibles could transition from susceptible to vaccinated.

FeLV spatial analyses

To test for spatial clustering of FeLV in the historical panther outbreak, we used a dataset of FeLV qPCR results (n = 31), in which 12 individuals tested positive. We used a circular window with a maximum spatial cluster of 50% of the population at risk. In addition, we used the same data to test for global clustering using Cuzick and Edward's test with the *smacpod* package in R (Cuzick & Edwards, 1990; French, 2020). Here, we evaluated nearest neighbor levels (k) of 1, 3, 5, 7, 9 and 11, and used 999 iterations for inference. We used these same parameterizations for SaTScan and Cuzick and Edward's analyses of FeLV simulations, with the exception that we only evaluated k = 3, 5, and 7 for simulated data (see Chapter 3 main text).

FeLV prediction target ranges

When comparing FeLV simulation predictions against observations from the historical outbreak, we used several "target" ranges for outbreak duration and the number of progressive infections. More specifically, the empirical outbreak is considered to have occurred from July 1, 2002 - June 30, 2004 (104 weeks), but due to uncertainty in the precise duration of the historical outbreak, we considered a simulated duration of 78-117 weeks to be "on target." During the observed outbreak, 5 individuals were documented with progressive (or transient) infection. Panthers are cryptic, difficult-to-observe animals, resulting in uncertainty in detection of all progressive infections and full population size at the time. We therefore considered 5-20 progressive infections in simulations to be on target.

B.2 Results

FIV transmission network inference

Phyloscanner FIV network results indicated some dissimilarities between single runs. As reported in the main text for Chapter 3, the main FIV network included 19 nodes with 42 edges (network density = 0.25) after removing 9 edges that were between individuals known not to be alive at the same time (Figure B.1). The summary transmission network allowing scanning window overlap included 20 nodes with 43 edges (network density = 0.23), and the summary network without window overlap included 20 nodes with 35 edges (network density = 0.18; after 8 and 6 edges removed, respectively, due to dates known alive).

Supplementary ERGM results

Results from the best ERGM for each FIV network are given in Table B.1. Of note, the best ERGM for the main FIV network showed reasonable goodness of fit (Figure B.2). ERGM results for the two summary networks were comparable to the main FIV network (Table B.1). The key difference was that the summary network with no window overlap did not find log transformed pairwise geographic distances to be a significant variable, though this fitted model showed evidence of degeneracy. To further confirm consistency of our ERGM findings, we performed a *post hoc* analysis with simulated random networks (see below; Figure B.3).

Post hoc random network ERGM analysis

Because there were some differences between ERGM results from the three FIV transmission networks (Chapter 3 main text; Table B.1), we performed a *post hoc* random network analysis to determine the consistency of our results against "null" random networks. Using the same panthers and descriptive data from the main FIV transmission tree, we rewired this transmission network as an Erdos-Renyi random network of the same density and fit an ERGM with the same variables from our main ERGM. We repeated this procedure 50 times, recording variable coefficients with each iteration. We then compared the distribution coefficients from simulated random networks to those from our three ERGMs, finding strong consistency among our ERGM coefficients relative to those from random networks (Figure B.3).

FeLV spatial analyses

As reported in the main text for Chapter 3, SaTScan analysis of observed FeLV status found weak evidence of spatial clustering (two clusters detected, but not statistically significant with p=0.165 and 0.997, respectively; Figure B.4).

FeLV simulations

Main results of the generalized linear mixed model (GLMM) for FeLV predictive model performance (see Chapter 3 main text) are given in Table B.2 (homogeneous mixing model was reference group; for parameter set random intercepts: variance = 0.90; standard deviation = 0.95). While the FIV-based approach did not show statistically significant improvements in performance, it did trend toward best performance, having the highest number of "feasible" parameter sets (Figures B.5, B.6).

SaTScan results for simulated FeLV cases and controls are given in the main text for Chapter 3. Cuzick and Edward's tests found evidence of global clustering of simulated FeLV cases with both the FIV and overlap-based models. However, for simulations with p-values less than or equal to 0.1, the FIV-based model was moderately more likely to capture the strength of global clustering (observed/expected test statistic, T_k, ratio) from the empirical FeLV data (Figure B.7).

In addition, to better understand the importance of FeLV transmission parameters in generating "feasible" results, we performed a *post hoc* random forest analysis for each model type, with "feasible" as a binary outcome for each parameter set (as in White et al., 2020). Predictors were the FeLV transmission transmission parameters, and data were split into 80% training/20% testing data sets. Because few parameter sets were categorized as feasible, we tested different resampling strategies for balancing the data. These included no resampling, down sampling, up sampling,

up/down sampling, and SMOTE sampling. The sampling protocol that produced the highest balanced accuracy was carried forward for analysis. In addition, we optimized hyper-parameters for the final random forest model. All random forests were performed using the *randomForest* package in R (Liaw et al., 2002). Across all model types, final random forest results tended to show poor balanced accuracy, low area under the curve (AUC; observed as low as AUC of 0.5) results, and were often inconsistent between repetitions (i.e. changes to training/testing data sets). Example random forest output for the FIV-based model is shown in Figures B.8 and B.9 for transparency, but should be interpreted with caution. Of particular note, however, was that C, the modifier shaping potential transmission from regressively infected individuals, had a strong tendency across model types to show best performance at C = 0.1 or 0.5 (Figure B.10); this would support the possibility of low transmissibility of regressively infected individuals. See main text for Chapter 3 for further discussion of this finding.

B.3 Tables

Table B.1: Best ERGM results for each FIV transmission network

Transmission network	Variable	Estimate	SE	*p*-value
Main FIV network	Edges (intercept)	-2.56	1.33	0.055
	gwesp	0.98	0.26	<0.001
	altkstar	-0.70	0.96	0.47
	Age (Adult)	0.93	0.44	0.03
	Log pairwise distance	-0.45	0.21	0.03
Summary network with window overlap	Edges (intercept)	-0.15	1.48	0.92
	gwesp	1.03	0.31	<0.001
	altkstar	-3.51	1.22	0.004
	Age (Adult)	1.36	0.61	0.02
	Log pairwise distance	-0.63	0.22	0.004
Summary network without window overlap	Edges (intercept)	-2.76	1.33	0.038
	gwesp	1.03	0.32	0.001
	altkstar	-2.17	0.99	0.029
	Age (Adult)	1.03	0.57	0.073

Table B.2: Fixed effects results from model-type performance GLMM

Variable	Estimate*	SE	*p*-value
Intercept	0.055	0.40	<0.001
FIV-based network model	1.55	0.42	0.30
Random network model	1.32	0.43	0.52
Overlap-based network model	1.21	0.44	0.66

**Note: estimates provided are exponentiated*

B.4 Figures

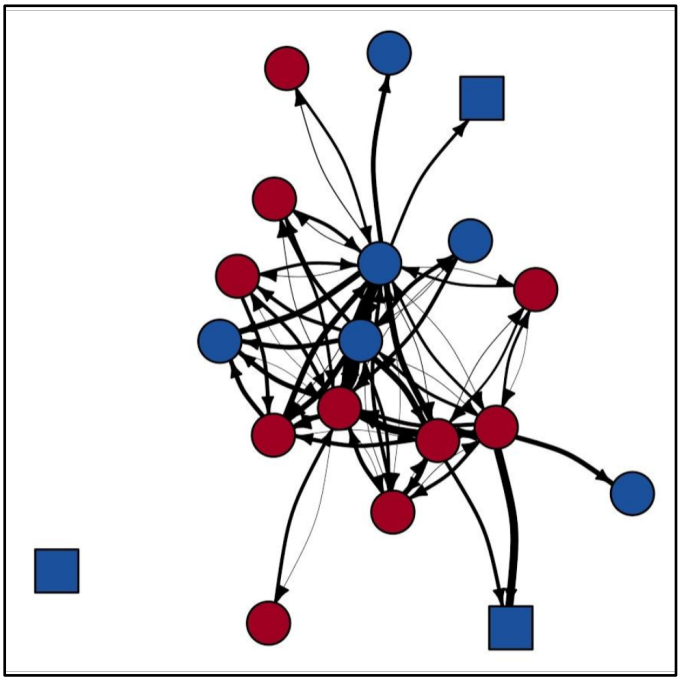


Figure B.1: Phyloscanner-derived main FIV transmission network. Node shape indicates panther age category (square = subadult; circle = adult). Node color indicates panther sex (blue = male; red = female). Edge weight represents Phyloscanner tree support for each edge (thicker edge = increased support); for visualization purposes, edges are displayed as the inverse of the absolute value of the log of these support values. While pictured as a directed and weighted network, statistical analyses used binary, undirected networks.

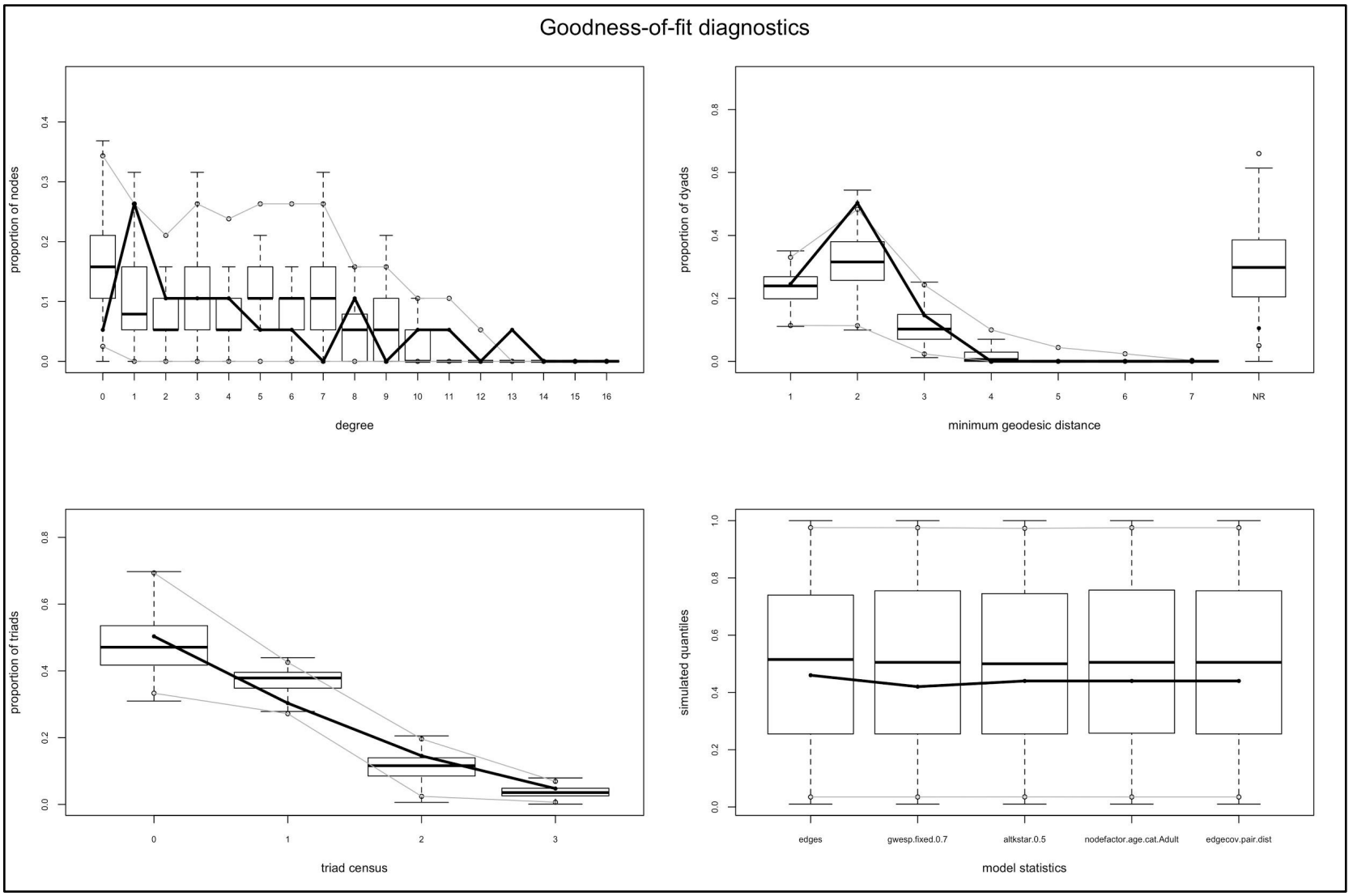

Goodness-of-fit diagnostics
proportion of nodes
degree
proportion of dyads
minimum geodesic distance
proportion of triads
triad census
simulated quantiles
edges
gwesp.fixed.0.7
altkstar.0.5
nodefactor.age.cat.Adult
edgecov.pair.dist
model statistics

Figure B.2: Best ERGM for main FIV network showed reasonable goodness of fit across standard goodness of fit metrics. Clockwise, from top left: degree, minimum geodesic distance, model statistics, and triad census. Boxplots show model predictions; solid black lines observations.

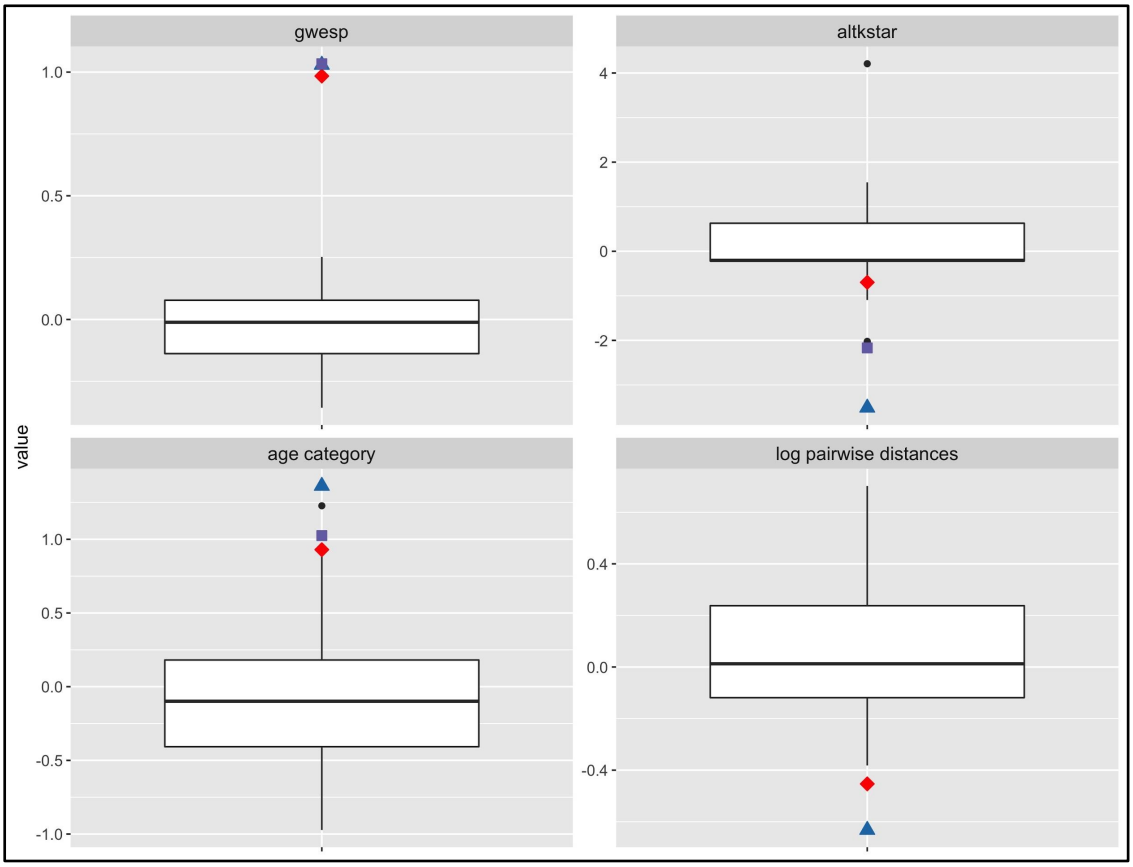


Figure B.3: Fitting the predictors from the best ERGMs to random networks (based on the main FIV network) shows that all three best models give largely consistent coefficient estimates. Boxplots show coefficient estimates from 50 random networks. Red diamonds are estimates from the main FIV network ERGM; blue triangles are estimates from the summary network with window overlap; purple squares are estimates from the summary network without window overlap. Note that the primary inconsistency between models is that the summary network without overlap did not identify log pairwise distances as a significant predictor.

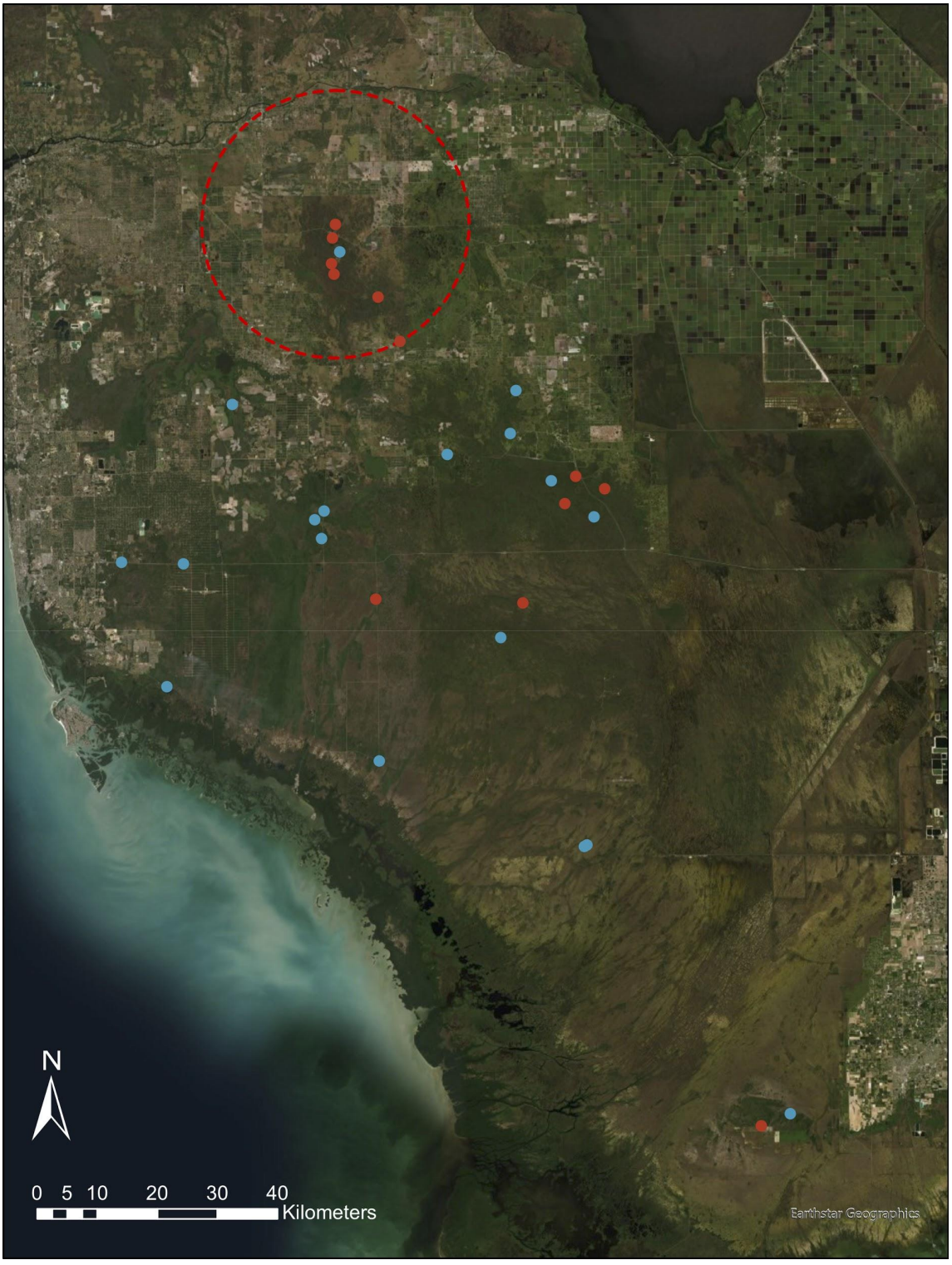


Figure B.4: Map of observed panther centroid locations and FeLV status (red = qPCR positive, blue = qPCR negative). The red dashed circle shows the location and size of

the top SaTScan cluster candidate, though this cluster was not considered statistically significant (p = 0.165).

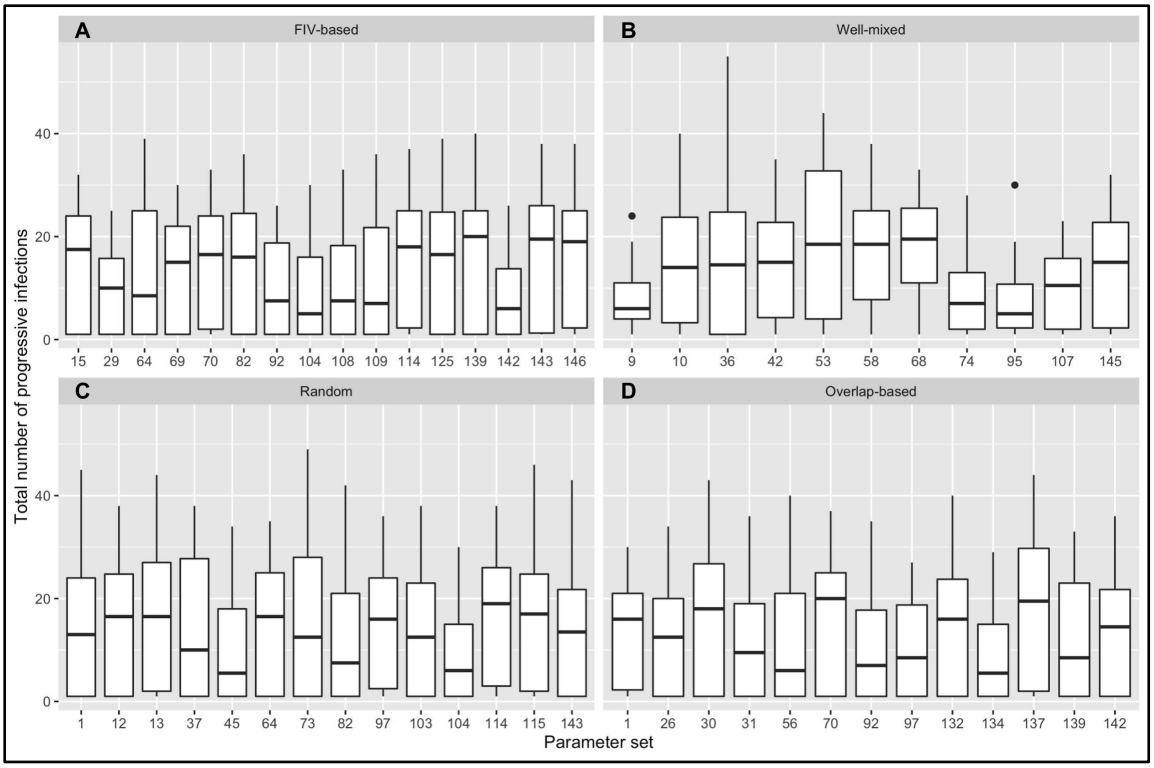


Figure B.5: Boxplots of total number of progressive infections from parameter sets classified as “feasible.” Results are shown for model types (A) FIV-based network, (B) well-mixed compartmental model, (C) random network, and (D) overlap-based network. Parameter set on the x-axis represents the unique parameter set drawn from our LHS sampling design; for example, set 1 for the FIV-based model type is identical to set 1 for random network model type, but feasible sets are not necessarily the same across model types.

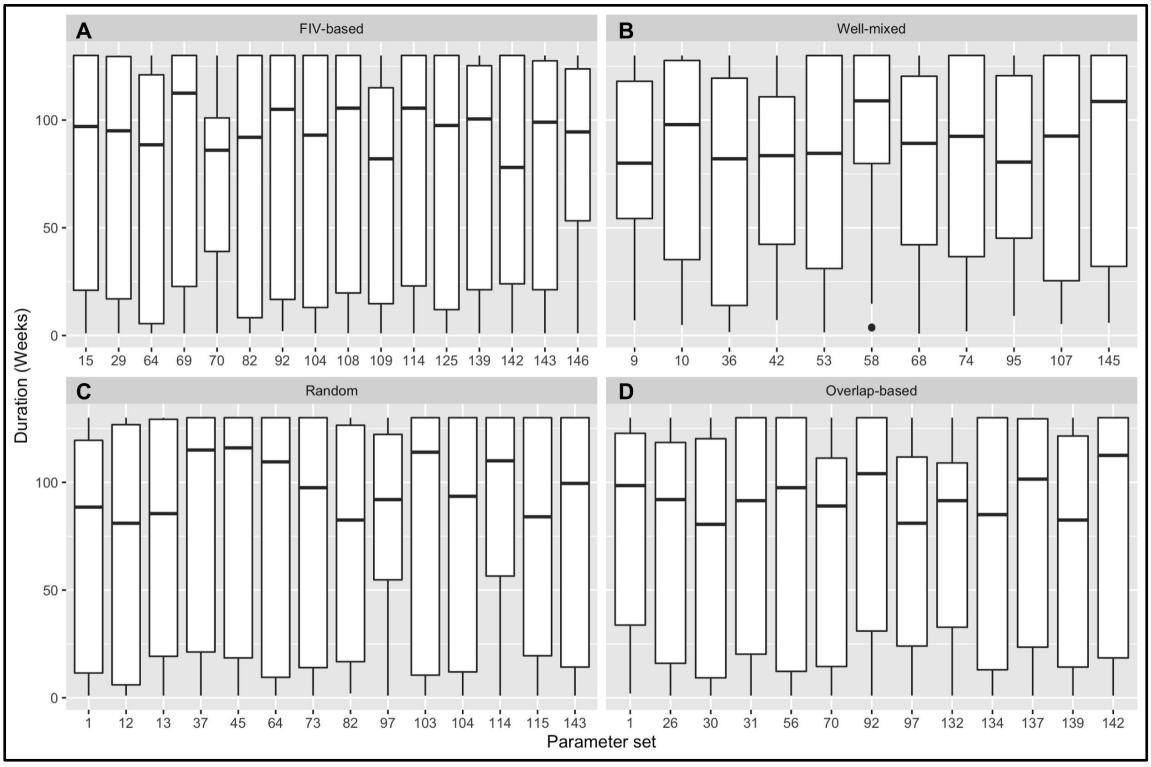


Figure B.6: Boxplots of duration of simulated epidemics from parameter sets classified as “feasible.” Results are shown for model types (A) FIV-based network, (B) well-mixed compartmental model, (C) random network, and (D) overlap-based network. Parameter set on the x-axis represents the unique parameter set drawn from our LHS sampling design; for example, set 1 for the random network model type is identical to set 1 for the overlap-based network model type, but feasible sets are not necessarily the same across model types.

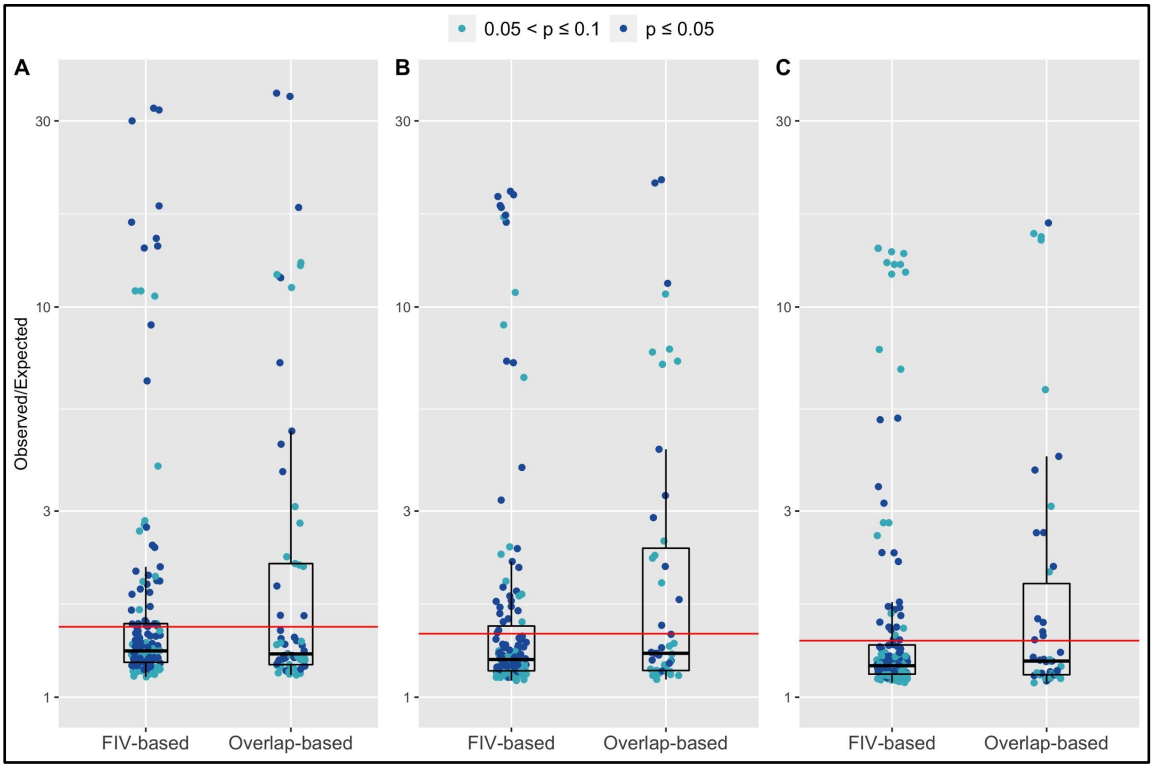


Figure B.7: Observed/Expected test statistics (T_k) for Cuzick-Edwards tests performed with FIV and overlap-based predictions of FeLV transmission. Shown are results from feasible parameter set simulations in which Cuzick-Edwards test results had p-value less than or equal to 0.1. Plots represent the neighbor levels that demonstrated statistically significant clustering for the empirical FeLV data: (A) k = 3; (B) k = 5; (C) k = 7. The red horizontal line in all cases is the Observed/Expected T_k ratio for the empirical FeLV data.

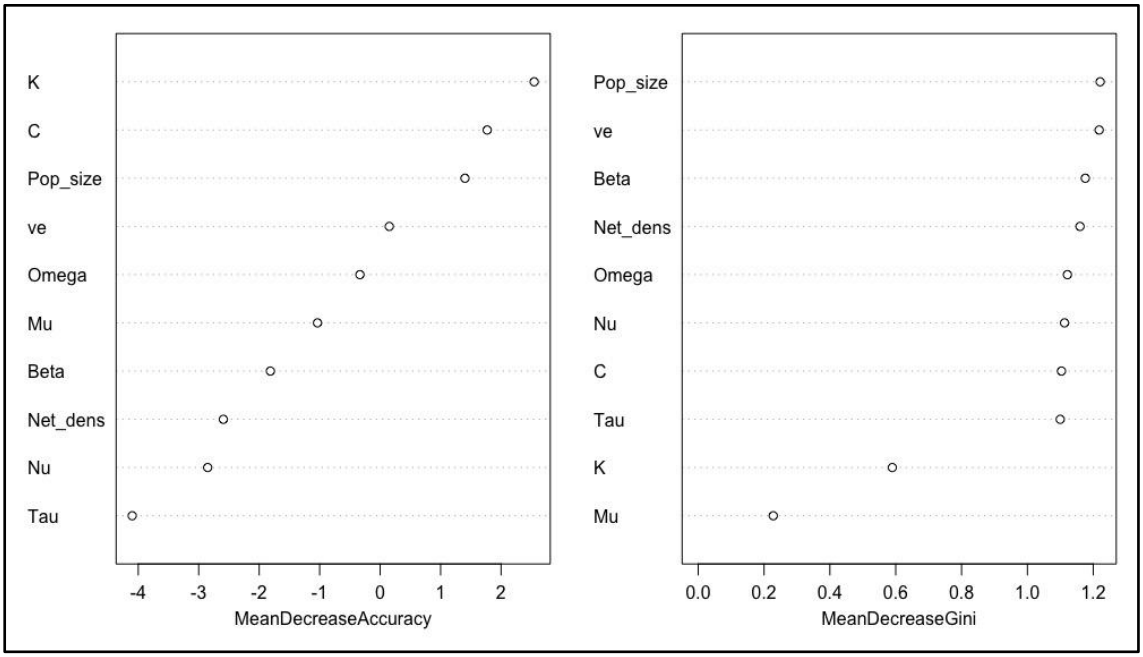


Figure B.8: Variable importance plots for the FIV-based model. While AUC was 0.889 for this random forest analysis, results were inconsistent between random forests and should be interpreted with caution. Variable names are given on the x-axis (see Table 3.1 in Chapter 3 main text). Mean decrease in accuracy scores is given on the x-axis in the left panel; mean decrease in Gini index on the x-axis in the right panel.

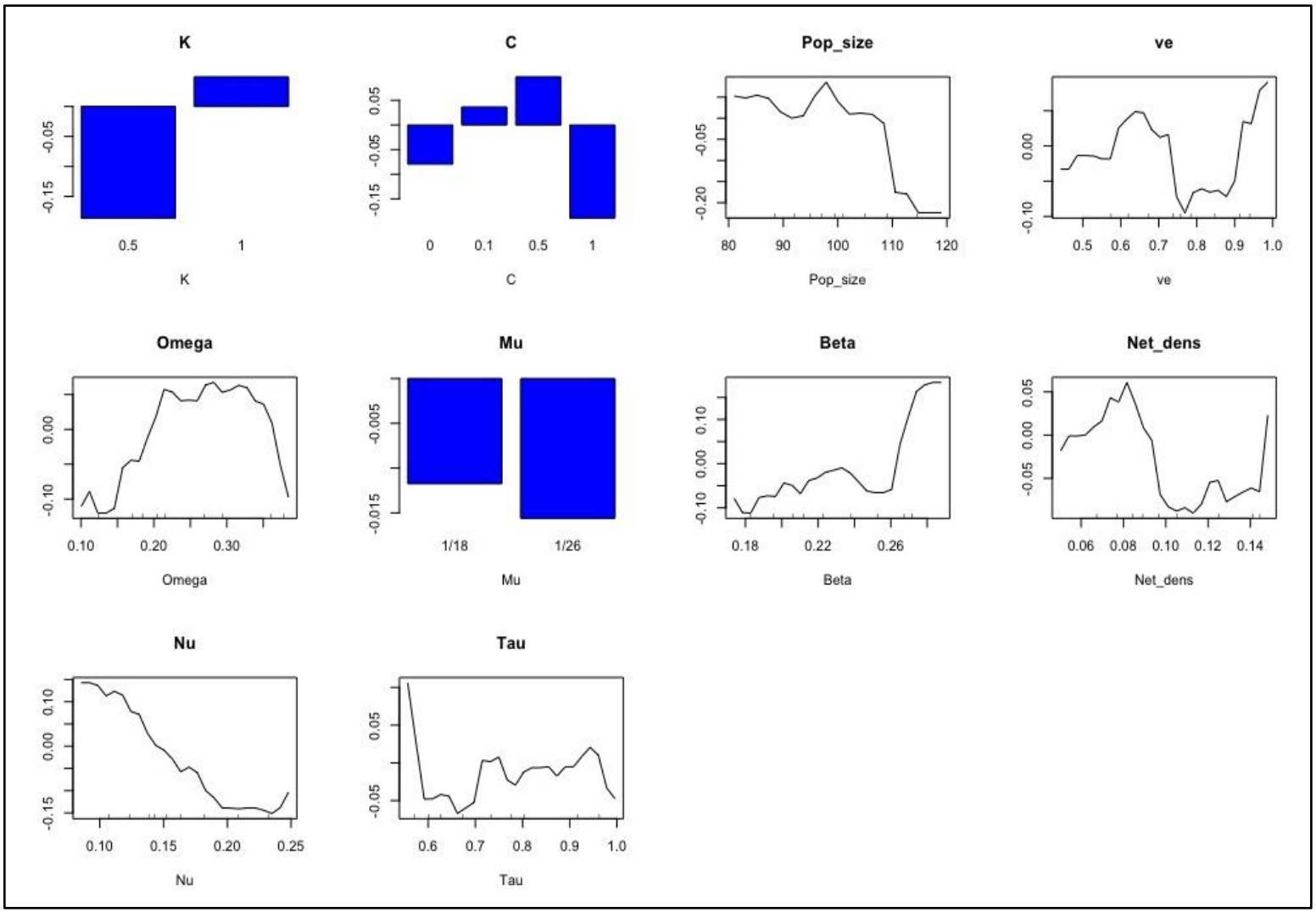

K
-0.05
-0.15
0.5
1
K
C
0.05
-0.05
-0.15
0
0.1
0.5
1
C
Pop_size
-0.05
-0.20
80
90
100
110
120
Pop_size
ve
0.00
-0.10
0.5
0.6
0.7
0.8
0.9
1.0
ve
Omega
0.00
-0.10
0.10
0.20
0.30
Omega
Mu
-0.005
-0.015
1/18
1/26
Mu
Beta
0.10
0.00
-0.10
0.18
0.22
0.26
Beta
Net_dens
0.05
0.00
-0.05
0.06
0.08
0.10
0.12
0.14
Net_dens
Nu
0.10
0.00
-0.15
0.10
0.15
0.20
0.25
Nu
Tau
0.05
-0.05
0.6
0.7
0.8
0.9
1.0
Tau

Figure B.9: Partial dependence plots for the FIV-based model, ordered based on variable importance observable in Figure B.8 (according to mean decrease in accuracy scores; highest importance in top left). While AUC was 0.889 for this random forest analysis, results were inconsistent between random forests and should be interpreted with caution. For example, note that partial dependence results differ quantitatively for the C parameter between this figure and Figure B.10, though they are qualitatively consistent. Variable names are given in plot titles and x-axes (see Table 3.1 in the Chapter 3 main text).

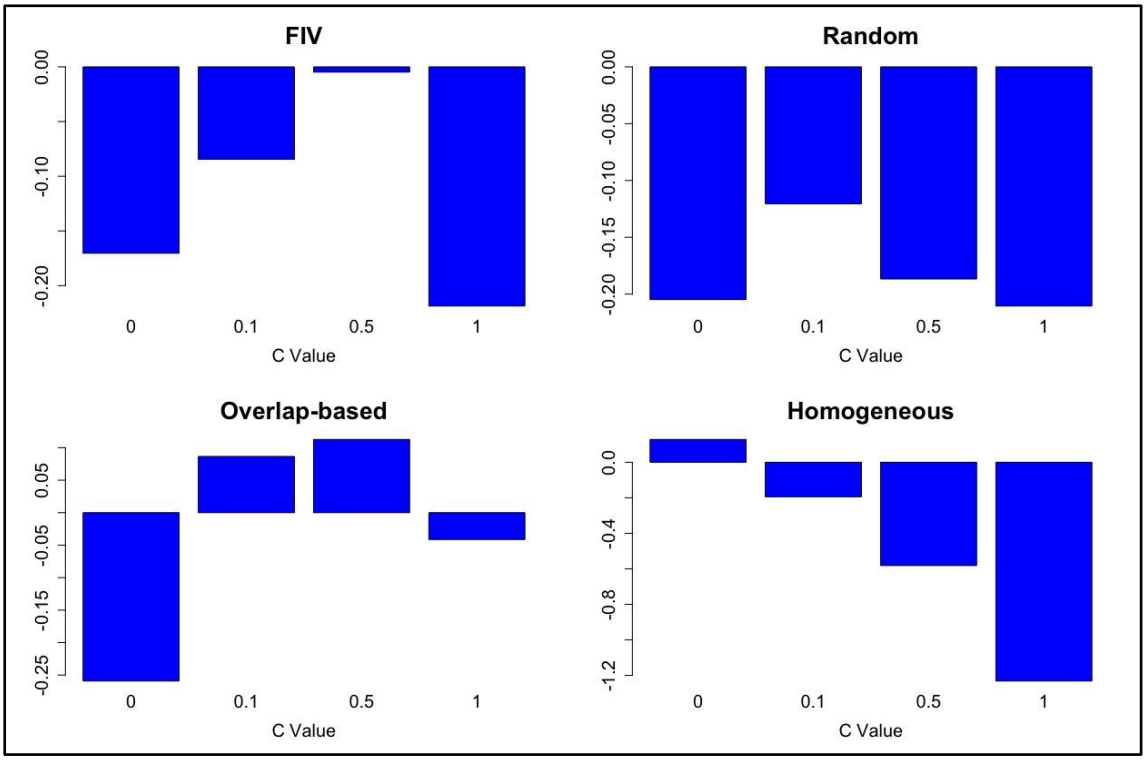


Figure B.10: Partial dependence plots for the C parameter, which was a constant multiplier for probability of transmission from regressives, given effective contact. Because random forest analyses were sensitive to sampling, the plotted results are from random forest models in which the area under the curve was greater than or equal to 0.8. All models but homogeneous mixing showed at least some support for values of C greater than 0 but less than 1.

Appendix C. Supplementary information for Chapter 4

C.1 Materials and Methods

Network and transmission simulations

The following is summarized from Chapter 3 (and Appendix B) and additional details about network and transmission simulations can be found there.

We used a spatially-explicit network simulation approach for our simulations, as this allowed us to incorporate both our previously identified drivers of retrovirus transmission and contact heterogeneity important for shaping epidemic outcomes (Keeling & Eames, 2005; Lloyd-Smith et al., 2005). More specifically, we used our previously defined exponential random graph model (ERGM) for retrovirus transmission pathways in panthers to simulate contact networks. Networks were therefore simulated based on pairwise geographic distances between panther home range centroids, proportion of the population that is adult (versus subadult), and two network structural terms, alternating k-stars and geometrically weighted edgewise shared partner distribution. To generate these networks, we first simulated panther populations with these key characteristics (i.e. age classes and distances), and then used ERGM coefficients from Chapter 3 to generate contact networks representing likely FeLV transmission pathways.

We used empirical telemetry data to simulate pairwise geographic distances among each simulated panther population. Panthers have been monitored using predominantly VHF telemetry for decades, with aerial relocations recorded typically every 3 days. The number of collared individuals has decreased in recent years, but as we were focused on FeLV management in the contemporary population, we focused our

telemetry use on the most recent years with consistently high relative telemetry coverage (2010-2012; 56 total panthers monitored, with an average of 34 monitored per year). This telemetry data was then used to simulate home range centroids and subsequent pairwise distances.

To simulate age distributions, we randomly assigned age categories (adult versus subadult) based on a "proportion adult" parameter (Table C.1), taken from random forest analysis from Chapter 3 (and Appendix B). In simulating networks, we also constrained network density based on target values again identified in random forest analysis (Table C.1). Similarly, for simulating FeLV transmission along networks, we used a susceptible-infectious-recovered framework and parameterization adapted random forest results (Table C.1, Figure C.1).

We held all network and transmission parameters constant for our primary baseline (no-intervention) and management scenarios to evaluate only the effect of interventions (Table C.1). For sensitivity analysis, we considered the same reasonable parameter ranges used Chapter 3 (Table C.1).

Spatial distribution of reactive vaccination

To simulate spatially-targeted vaccination in a *cordon sanitaire* along the I-75 freeway, we selected individuals for attempted vaccination based on their proximity to the freeway. Proximity was determined by the distance between each individual's simulated home range centroid and latitude 26.16 (representing I-75). Proximities were ranked into three categories (i.e., closest, middle, and furthest thirds of the population), and weighted for sampling. The closest third received a weight of 0.7; the middle third, 0.2; and the furthest third, 0.1. Weighting allowed us to preferentially select individuals

with home range centroids close to I-75, but also represented the real-world possibility that panthers range widely and may be identified further from these centroids. Weighting also accounted for the fact that, as individuals close to the freeway were “saturated” with vaccination, other, more distant individuals might become more available candidates for vaccination.

Sensitivity analyses supplementary methods

To mitigate computational complexity for evaluating sensitivity of our proactive vaccination results, we focused on the subset of proactive vaccination conditions in which 20, 40, or 60% of the population was proactively vaccinated, with 0, 50, or 100% of vaccinates receiving a booster. Furthermore, we selected a subset of LHS parameter sets for these proactive vaccination sensitivity tests, choosing these based on the results of the full sensitivity analysis under the baseline scenario (see main text). There, we recorded the median number of progressive infections for each of the 50 parameter sets, then randomly selected sets based on quantiles (3 parameter sets from the lower quantile, 6 from the interquartile range, and 3 from the upper quantile). As reported in the main text, this approach allowed us to examine sensitivity of our proactive vaccination results across different outbreak sizes and network and transmission parameters, while mitigating computational effort associated with exploring such a wide range of parameters and scenario variations. The resulting 12 LHS parameter sets, used in a factorial design across proactive vaccination conditions, resulted in 108 parameter sets with 50 full simulations per set (5,400 full simulations).

To evaluate sensitivity analysis simulation results, we used a combination of scatterplots and Partial Rank Correlation Coefficients (PRCC; Marino et al., 2008; Wu et

al., 2013). For baseline, no intervention scenarios, we tested for parameters of significance for the outcome of median mortalities per parameter set. We used averaged mortalities to both minimize the effect of aleatory uncertainty and increase statistical power of PRCC (Marino et al., 2008).

For proactive vaccination sensitivity analyses, we first determined if simulation results were consistent with our qualitative outcome that low vaccination levels can increase mortalities using histograms and median mortalities as in the main text. Finding that this result was not consistent across all parameter sets, we performed additional *post-hoc* analyses and simulations to determine parameters important for increased mortalities at low vaccination levels (see supplementary results below).

C.2 Results

Management scenarios supplementary results

For transparency, we present the results for all management scenarios here. Full results for proactive vaccination alone are shown in Figures C.2 (mortalities) and C.3 (epidemic durations). In addition, we show the relationship between outbreak duration and number of mortalities in Figure C.4, demonstrating that longer outbreaks are typically associated with higher numbers of mortalities. Full results for reactive vaccination mortalities are shown in Figures C.5-6 and durations in Figure C.7-8. Full results for reactive test-and-removal are shown in Figures C.9 (mortalities) and C.10 (epidemic durations). Full results for reactive underpass closures are shown in Figures C.11 (mortalities) and C.12 (epidemic durations). The proportion of failed epidemics (fewer than 5 progressive or regressive infections) per 100 successful epidemics for all management scenarios are shown in Figure C.13.

Sensitivity analysis of no-intervention scenarios

As reported in the main text, simulated FeLV outbreak sizes were variable across the 50 sensitivity analysis parameter sets examined under a no-intervention scenario (Figure C.14). PRCC results identified network density (Net_dens), infectiousness of regressives (C), weekly contact rates (ω), and the baseline probability of transmission (β) as statistically significant parameters with positive correlations to median mortalities; the weekly probability of mortality (μ) showed a significant negative correlation with median mortalities (Figure C.15). Note, however, that PRCC inference may be affected by discrete parameters (Marino et al., 2008), so these results should be interpreted with some caution.

Post-hoc sensitivity analysis

In our initial sensitivity analysis of proactive vaccination scenarios, we found that our results varied qualitatively from our main simulations, with low levels of proactive vaccination sometimes effective at reducing FeLV outbreak sizes. Focusing on the subset of sensitivity scenarios in which 20% of the population was proactively vaccinated with 100% boosting, we examined scatterplots and PRCC for parameter importance in relation to the outcome of the difference between median mortalities with and without proactive vaccination (hereafter, *mortality difference*; i.e., a difference value greater than 0 means that 20% vaccination with 100% boosting increased mortalities). PRCC analysis identified parameters of network density, infectiousness for regressives, and weekly contact rates (Net_dens, C, and ω, respectively) as statistically significant

parameters positively correlated with mortality difference at 20% proactive vaccination with 100% boosting (Figure C.16). However, PRCC inference may be affected by discrete parameters and may be confounded by non-monotonic data (Marino et al., 2008). We therefore also evaluated scatterplots of our outcome across parameter values, finding that the relationship between mortality difference and network density, in particular, was not obviously monotonic, though data was limited (Figure C.17). Given these limitations, we completed additional simulations to further evaluate the effect of network density on the mortality difference outcome.

To determine parameter space to target for additional simulations—and potential additional non-monotonicities—we first classified the results for each proactive vaccination sensitivity analysis parameter set (n=12) according to if low levels of proactive vaccination qualitatively reduced mortalities, made little to no change, or increased mortalities. We then examined these classifications across the parameter space represented by each pair of these 12 parameter sets (Figure C.18) to determine if any parameters or pairs of parameters showed clustering of classifications in parameter space. We were limited to this qualitative analysis by computational complexity (these 12 parameter sets alone required 5,400 simulations), but found some indications that, in addition to the previously identified target parameter of network density, the parameter for proportion adults also appeared to be associated with classification of simulation results as either “little to no change in mortalities” or “increased mortalities” at low levels of vaccination (Figure C.18).

To further examine this potential relationship, we generated an additional two parameter sets for a *post-hoc* sensitivity analysis of proactive vaccination scenarios (Table C.2). In the first set, we used the same target parameters from our main analyses

(Table C.1) for all parameters except network density and proportion adults. The latter two parameters were assigned based on apparent intermediate parameter space, which we hypothesized should preferentially result in low levels of proactive vaccination being associated with increased mortalities. This resulted in an additional 4 parameter sets across a no-intervention baseline and 9 variations in proactive vaccination (20/40/60% of the population vaccinated with 0/50/100% of individuals receiving boosted inoculations); with 50 simulations for each variation, this totaled an additional 2,000 simulations.

The second additional sensitivity parameter set held network density and proportion adults parameters constant from our main analyses (Table C.1), but drew all other simulation parameters from the seven proactive vaccination sensitivity analysis parameter sets that qualitatively produced results most divergent from our main analyses (Table C.2). Here, we hypothesized that if network density and proportion adults really are important drivers behind our results, these simulations should now produce results more qualitatively aligned with our main analyses. This resulted in an additional 7 parameter sets across a no-intervention baseline and 6 variations in proactive vaccination (20/40/60% of the population vaccinated with 50/100% of individuals receiving boosted inoculations; we did not include 0% boosted in order to reduce computational intensity). With 50 simulations for each variation, this totaled an additional 2,450 simulations.

For both *post-hoc* sensitivity analysis parameter sets, we examined qualitative alignment with our main results to determine if particular parameter space was associated with our main finding that low levels of proactive vaccination could paradoxically increase FeLV outbreak mortalities. These additional analyses found some evidence that intermediate values for network structural parameters (network density

and proportion adults) may be associated with the qualitative result that low proactive vaccination can increase FeLV mortalities; this may be especially true under conditions of high transmission potential (e.g. higher infectiousness of regressive individuals and/or increased weekly contact rates; Figure C.19). While our *post-hoc* sensitivity analysis did not specifically examine individual transmission parameters, we do note that high weekly probabilities of contact are also tentatively associated with our qualitative results (Figure C.20), which is consistent with our initial PRCC analysis with proactive vaccination sensitivity scenarios.

C.3 Discussion

FeLV transmission

While our sensitivity analysis highlighted the impacts of network structure on pathogen management requirements, additional uncertainty in FeLV transmission parameters should be considered. In particular, our main simulations allowed limited transmission from regressively infected individuals, which contrasts with expectations from domestic cats, but was supported by model-based evidence in Chapter 3. In addition, we assumed only a single spillover event from domestic cats, but these events occur with uncertain frequency (Chiu et al., 2019). The virulence of any circulating FeLV strain may also affect model predictions. There was some evidence that the FeLV strain in the 2002-2004 outbreak among panthers was particularly virulent (Brown et al., 2008). However, the same does not appear to have been true for Iberian lynx (Geret et al., 2011). Severity of FeLV-induced disease may therefore reflect the effects of inherent genetic susceptibility, particularly in these highly inbred populations (Geret et al., 2011; Johnson et al., 2010; Roelke et al., 1993). Lower virulence FeLV strains than we have

modeled here and/or increased population genetic diversity may therefore reduce severity of disease and even transmission potential in panthers.

C.4 Tables

Table C.1: Network and transmission simulation parameters

Parameter	Definition	Target Value	Range
Adult_prop	Proportion adults versus subadults	0.89	0.82-0.99
Net_dens	Simulated network density	0.08	0.05-0.15
β	Probability of transmission from progressives, given effective contact	0.27	0.17-0.29
C	Constant multiplier for probability of transmission from regressives, given effective contact	0.1	0, 0.1, 0.5
ω	Weekly probability of contact	0.3	0.1-0.4
μ	Weekly probability of death from progressive infection	1/18	1/18, 1/26
K	Constant multiplier for weekly probability of recovery from regressive infection	1	0.5, 1
ν	Weekly probability of territory repopulation ("respawn rate")	0.11	0.083-0.25
P	Proportion randomly assigned to progressive, regressive	0.25	0.25

Note: Values and ranges adapted from Chapter 3. Target values are those used in no-intervention and management scenarios. The ranges column gives the range of possible values used in the Latin hypercube sampling sensitivity analysis (see below). See Figure C.1 for transmission parameters in the context of the compartmental model.

Table C.2: Post-Hoc Sensitivity Analysis Parameter Sets

	Set Name	***Post-Hoc* Set 1**				***Post-Hoc* Set 2**						
	Set Number	1	2	3	4	1	2	3	4	5	6	7
Parameter Name	Adult_prop	0.8875	0.8875	0.9125	0.9125	0.89	0.89	0.89	0.89	0.89	0.89	0.89
	Net_dens	0.0875	0.1125	0.0875	0.1125	0.08	0.08	0.08	0.08	0.08	0.08	0.08
	β	0.27	0.27	0.27	0.27	0.259	0.243	0.249	0.196	0.251	0.191	0.279
	C	0.1	0.1	0.1	0.1	0	0	0.5	0.5	0	0	0.1
	ω	0.3	0.3	0.3	0.3	0.197	0.102	0.284	0.373	0.190	0.219	0.242
	μ	1/18	1/18	1/18	1/18	1/26	1/18	1/18	1/18	1/26	1/26	1/18
	K	1	1	1	1	1	0.5	0.5	0.5	0.5	0.5	1
	ν	0.11	0.11	0.11	0.11	0.176	0.181	0.167	0.186	0.120	0.197	0.113
	P	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25

Note: Parameter definitions can be found in Table C.1.

C.5 Figures

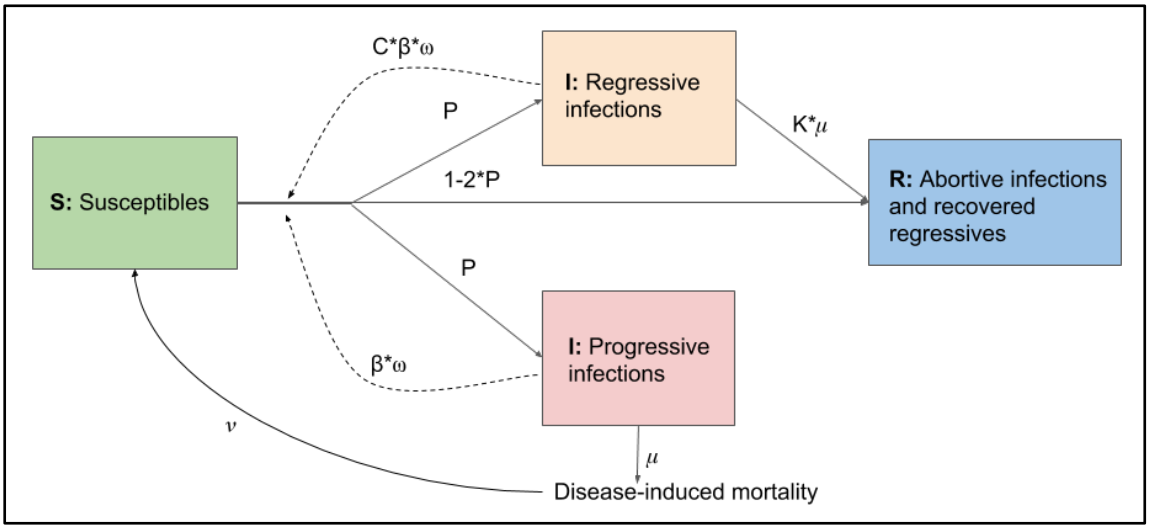


Figure C.1: Compartmental model diagram of FeLV transmission model (see Table C.1 for parameter definitions). Figure adapted from Chapter 3. Note that, for simplicity, the vaccinated class of individuals is not shown.

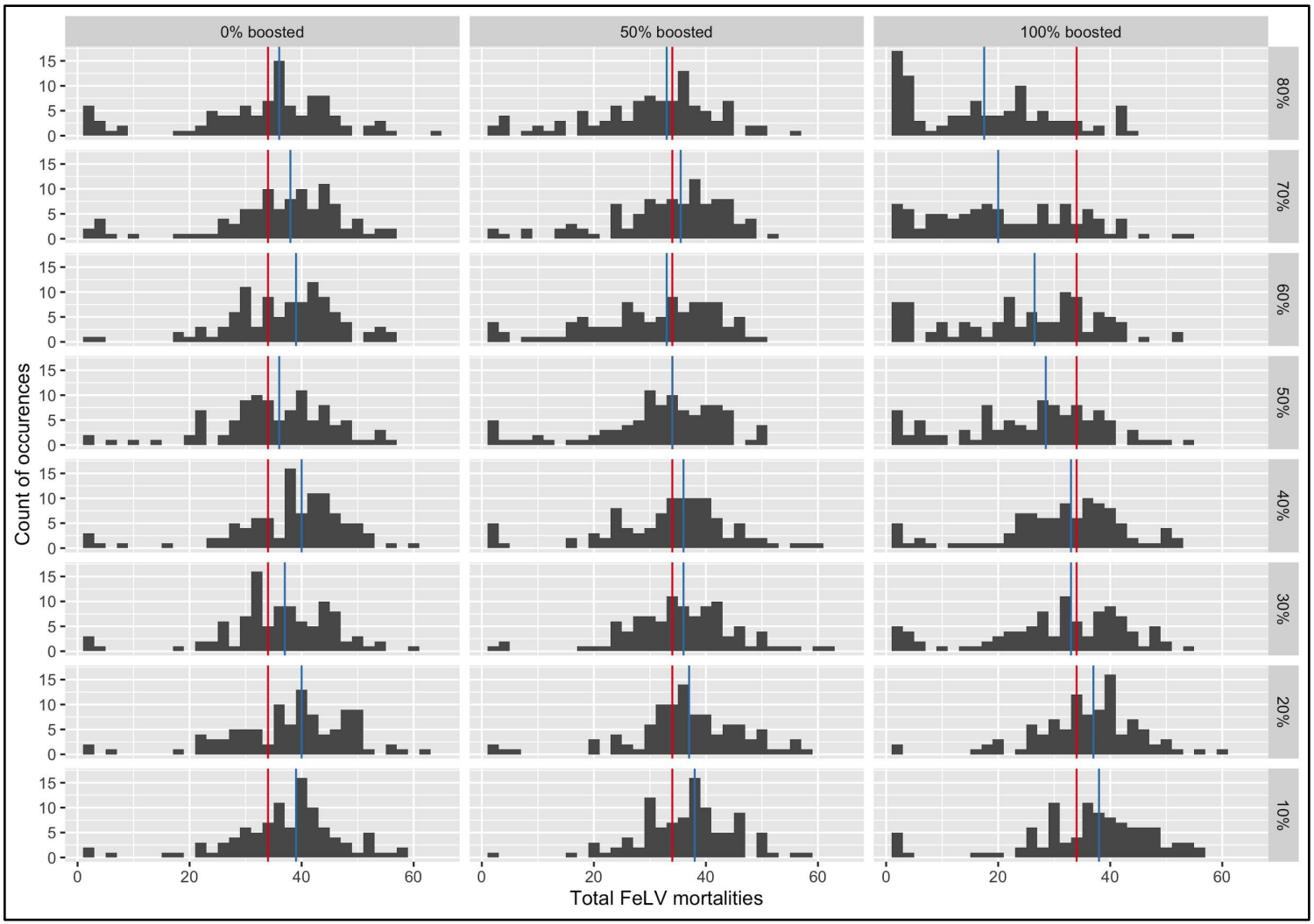

0% boosted
50% boosted
100% boosted
80%
70%
60%
50%
40%
30%
20%
10%
Count of occurences
Total FeLV mortalities
0
5
10
15
20
40
60

Figure C.2: Histograms of FeLV mortalities in simulated epidemics with proactive vaccination alone. Red vertical lines indicate the median number of mortalities from the baseline scenario without interventions; blue lines indicate median number of mortalities with proactive vaccination. Panel rows represent the proportion of the population proactively vaccinated; columns represent ratios of vaccine efficacy among vaccinates. Vaccination without a booster was assumed to have 40% efficacy, versus 80% efficacy with a boosting inoculation. Each histogram plot represents the results of 100 simulations.

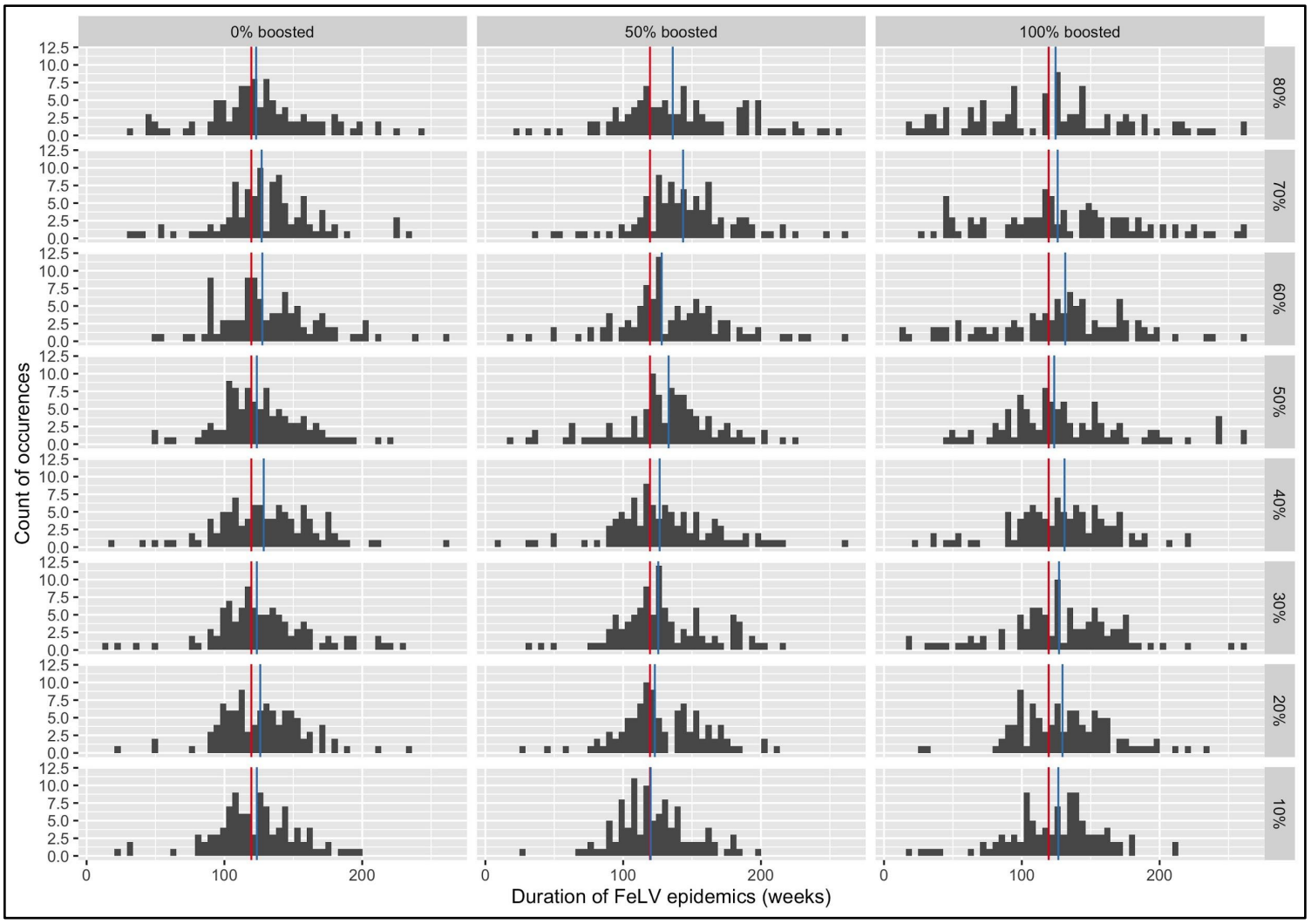

0% boosted
50% boosted
100% boosted
80%
70%
60%
50%
40%
30%
20%
10%
Count of occurences
0.0
2.5
5.0
7.5
10.0
12.5
0
100
200
Duration of FeLV epidemics (weeks)

Figure C.3: Histograms of the duration of simulated FeLV epidemics (in weeks) with proactive vaccination alone. Red vertical lines indicate the median duration from the baseline scenario without interventions; blue lines indicate median duration with proactive vaccination. Panel rows represent the proportion of the population proactively vaccinated; columns represent ratios of vaccine efficacy among vaccinates. Vaccination without a booster was assumed to have 40% efficacy, versus 80% efficacy with a boosting inoculation. Each histogram plot represents the results of 100 simulations.

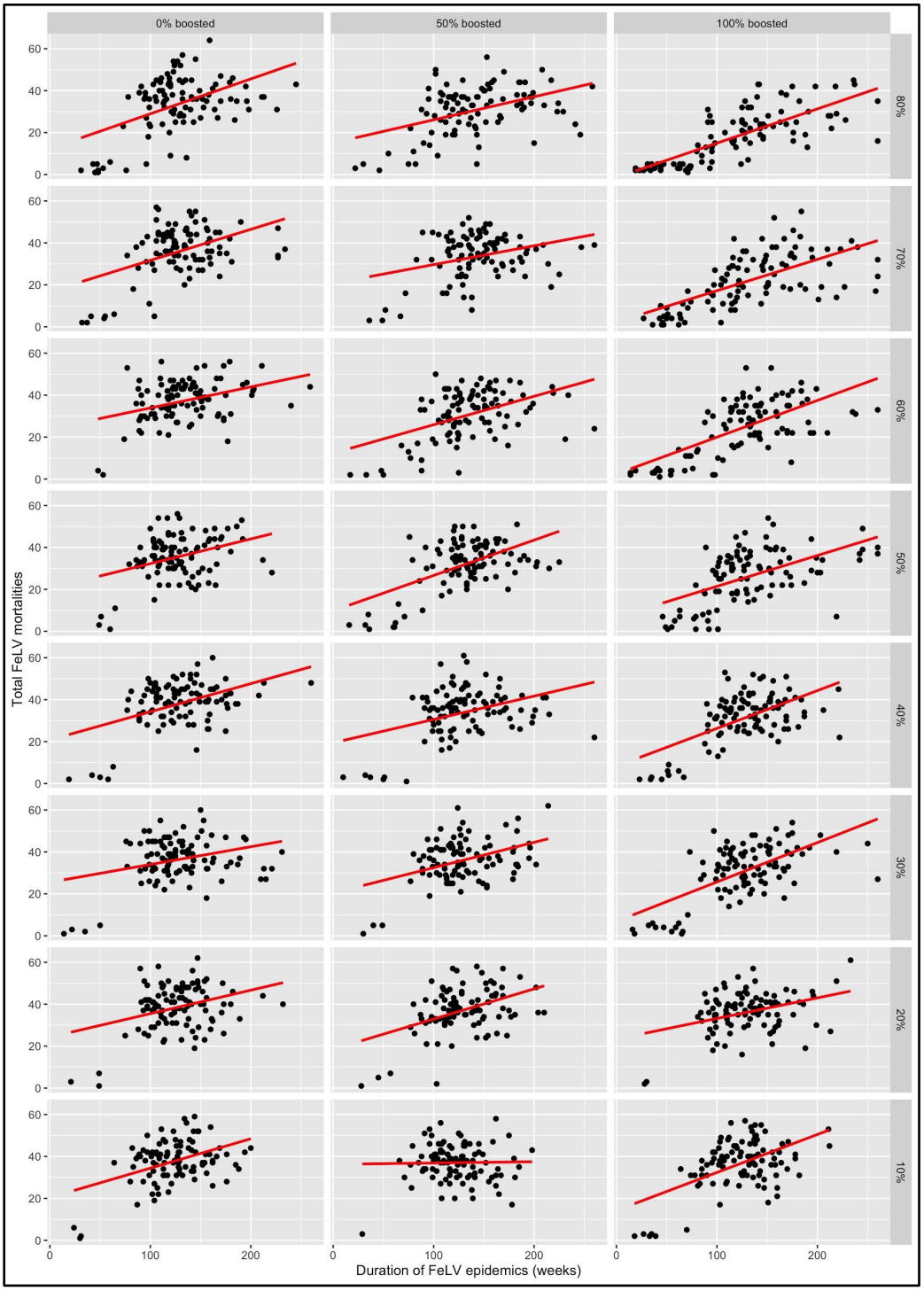


Figure C.4: Scatterplots of simulated FeLV epidemic durations against mortalities with proactive vaccination alone. Red lines are the fitted linear models for mortalities as a function of epidemic duration. Panel rows represent the proportion of the population

proactively vaccinated; columns represent ratios of vaccine efficacy among vaccinates. Vaccination without a booster was assumed to have 40% efficacy, versus 80% efficacy with a boosting inoculation. Each histogram plot represents the results of 100 simulations.

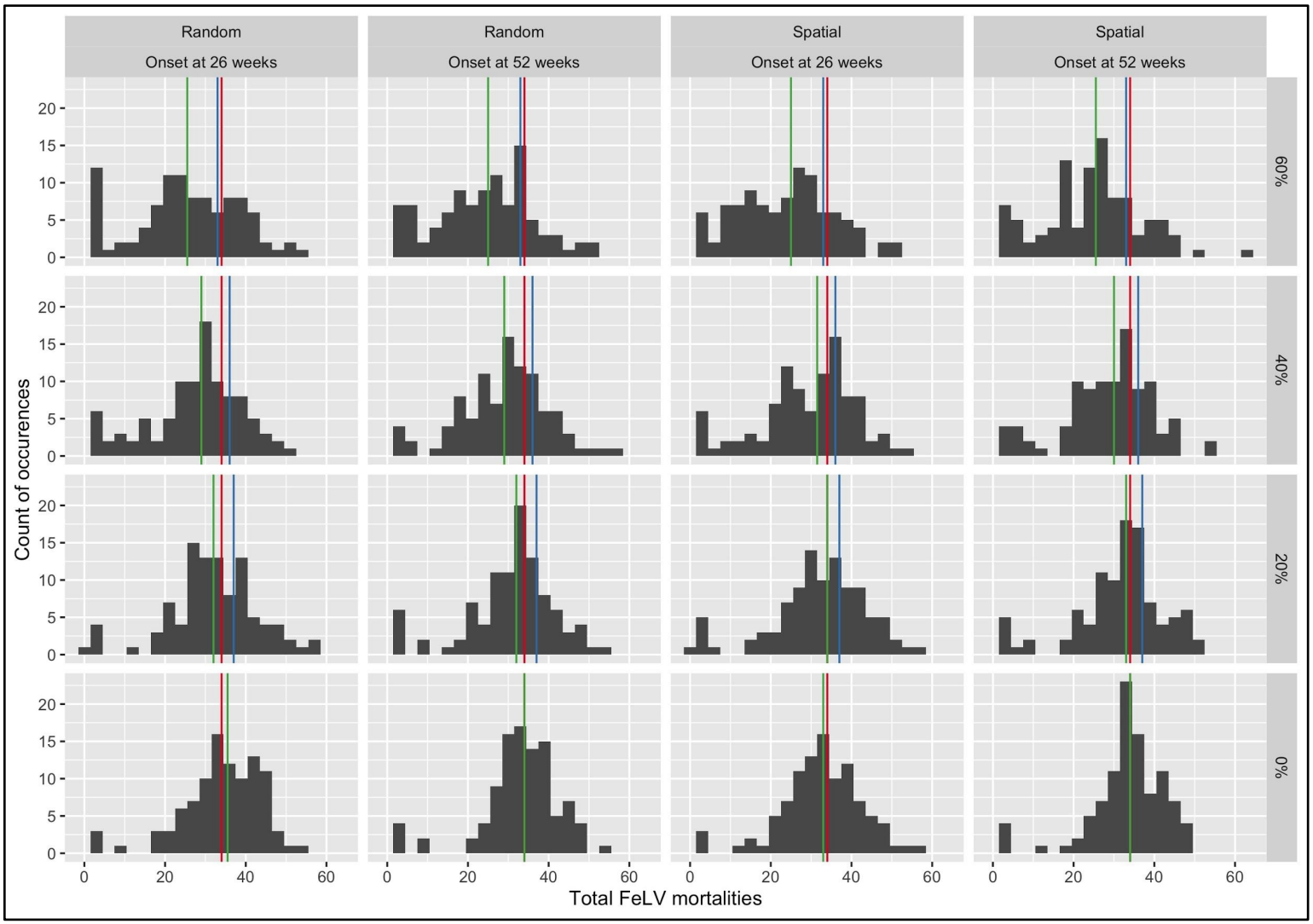
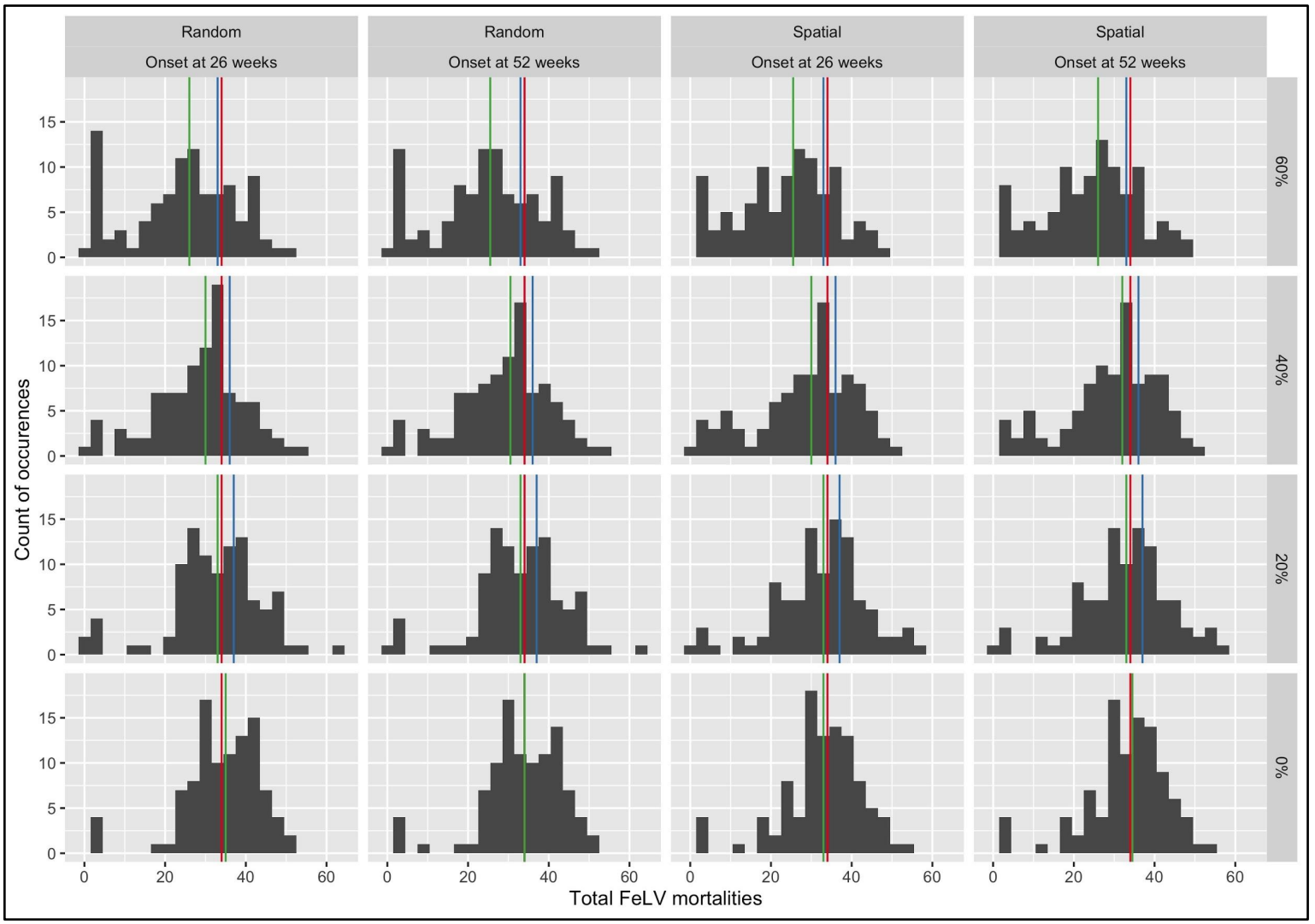

Random
Onset at 26 weeks
Random
Onset at 52 weeks
Spatial
Onset at 26 weeks
Spatial
Onset at 52 weeks
60%
40%
20%
0%
Count of occurences
Total FeLV mortalities
0
5
10
15
20
0
20
40
60

Figure C.5: Histograms of FeLV mortalities in simulated epidemics with proactive vaccination and year-round reactive vaccination. Panel rows represent the proportion of the population proactively vaccinated (with 50% of those vaccinates receiving boosted vaccination); columns represent both distribution strategy (random versus spatial) and timing of onset of reactive vaccination after epidemic initiation (26 or 52 weeks). Red vertical lines indicate the median number of mortalities from simulations without interventions; blue lines for proactive vaccination alone; green lines for the given combination of reactive and proactive vaccination. Each histogram plot represents the results of 100 simulations.


Random
Onset at 26 weeks
Random
Onset at 52 weeks
Spatial
Onset at 26 weeks
Spatial
Onset at 52 weeks
60%
40%
20%
0%
Count of occurences
Total FeLV mortalities
0
5
10
15
0
20
40
60

Figures C.6: Histograms of FeLV mortalities in simulated epidemics with proactive vaccination and 6-months per year reactive vaccination. Panel rows represent the proportion of the population proactively vaccinated (with 50% of those vaccinates receiving boosted vaccination); columns represent both distribution strategy (random versus spatial) and timing of onset of reactive vaccination after epidemic initiation (26 or 52 weeks). Red vertical lines indicate the median number of mortalities from simulations without interventions; blue lines for proactive vaccination alone; green lines for the given combination of reactive and proactive vaccination. Each histogram plot represents the results of 100 simulations.

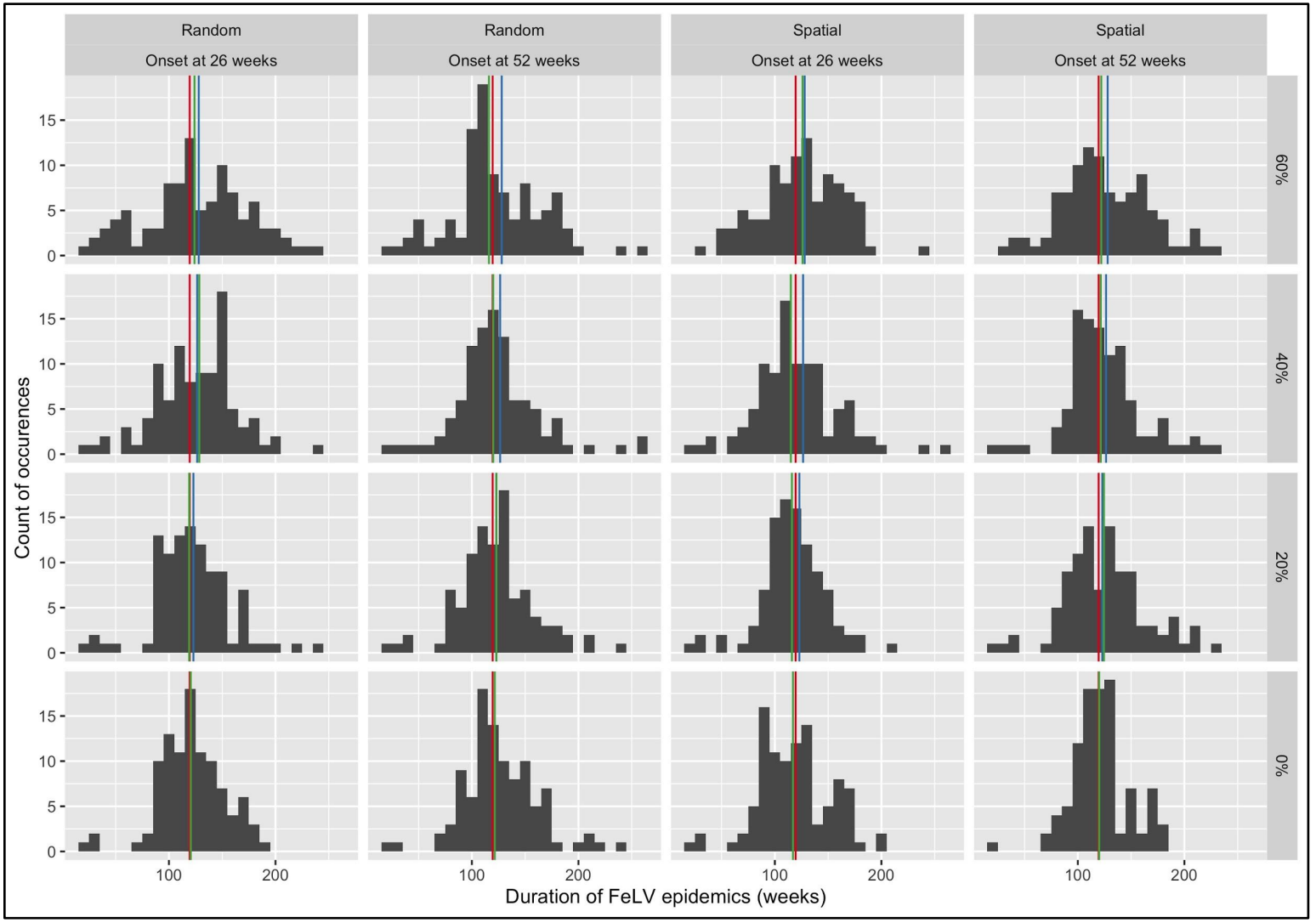

Random
Onset at 26 weeks
Random
Onset at 52 weeks
Spatial
Onset at 26 weeks
Spatial
Onset at 52 weeks
60%
40%
20%
0%
Count of occurences
0
5
10
15
100
200
Duration of FeLV epidemics (weeks)

Figure C.7: Histograms of the duration of simulated FeLV epidemics with proactive vaccination and year-round reactive vaccination. Panel rows represent the proportion of the population proactively vaccinated (with 50% of those vaccinates receiving boosted vaccination); columns represent both distribution strategy (random versus spatial) and timing of onset of reactive vaccination after epidemic initiation (26 or 52 weeks). Red vertical lines indicate the median duration without interventions; blue lines for proactive vaccination alone; green lines for the given combination of reactive and proactive vaccination. Each histogram plot represents the results of 100 simulations.

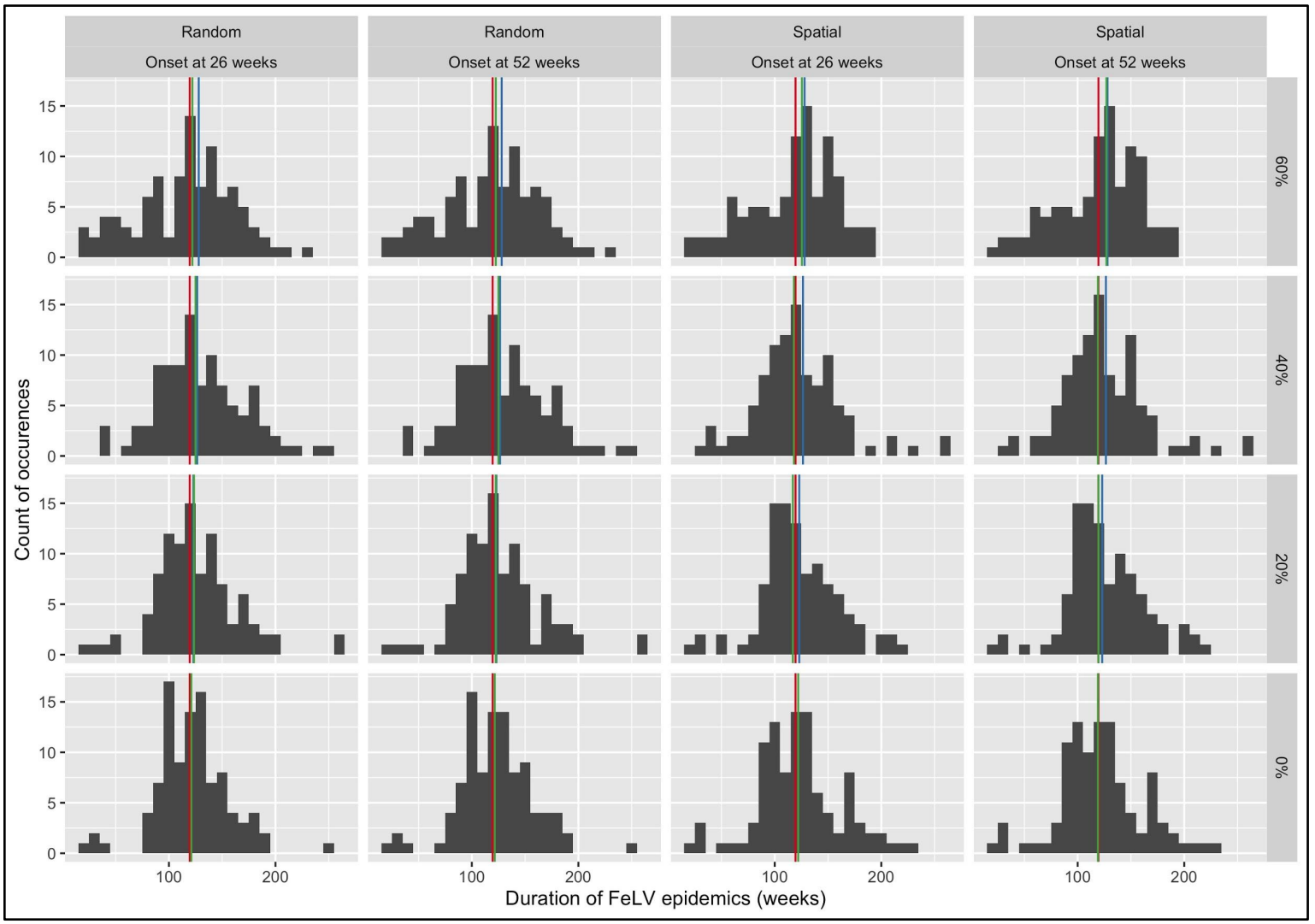

Random
Onset at 26 weeks
Random
Onset at 52 weeks
Spatial
Onset at 26 weeks
Spatial
Onset at 52 weeks
60%
40%
20%
0%
Count of occurences
Duration of FeLV epidemics (weeks)
0
5
10
15
100
200

Figure C.8: Histograms of the duration of simulated FeLV epidemics with proactive vaccination and 6-months per year reactive vaccination. Panel rows represent the proportion of the population proactively vaccinated (with 50% of those vaccinates receiving boosted vaccination); columns represent both distribution strategy (random versus spatial) and timing of onset of reactive vaccination after epidemic initiation (26 or 52 weeks). Red vertical lines indicate the median duration without interventions; blue lines for proactive vaccination alone; green lines for the given combination of reactive and proactive vaccination. Each histogram plot represents the results of 100 simulations.

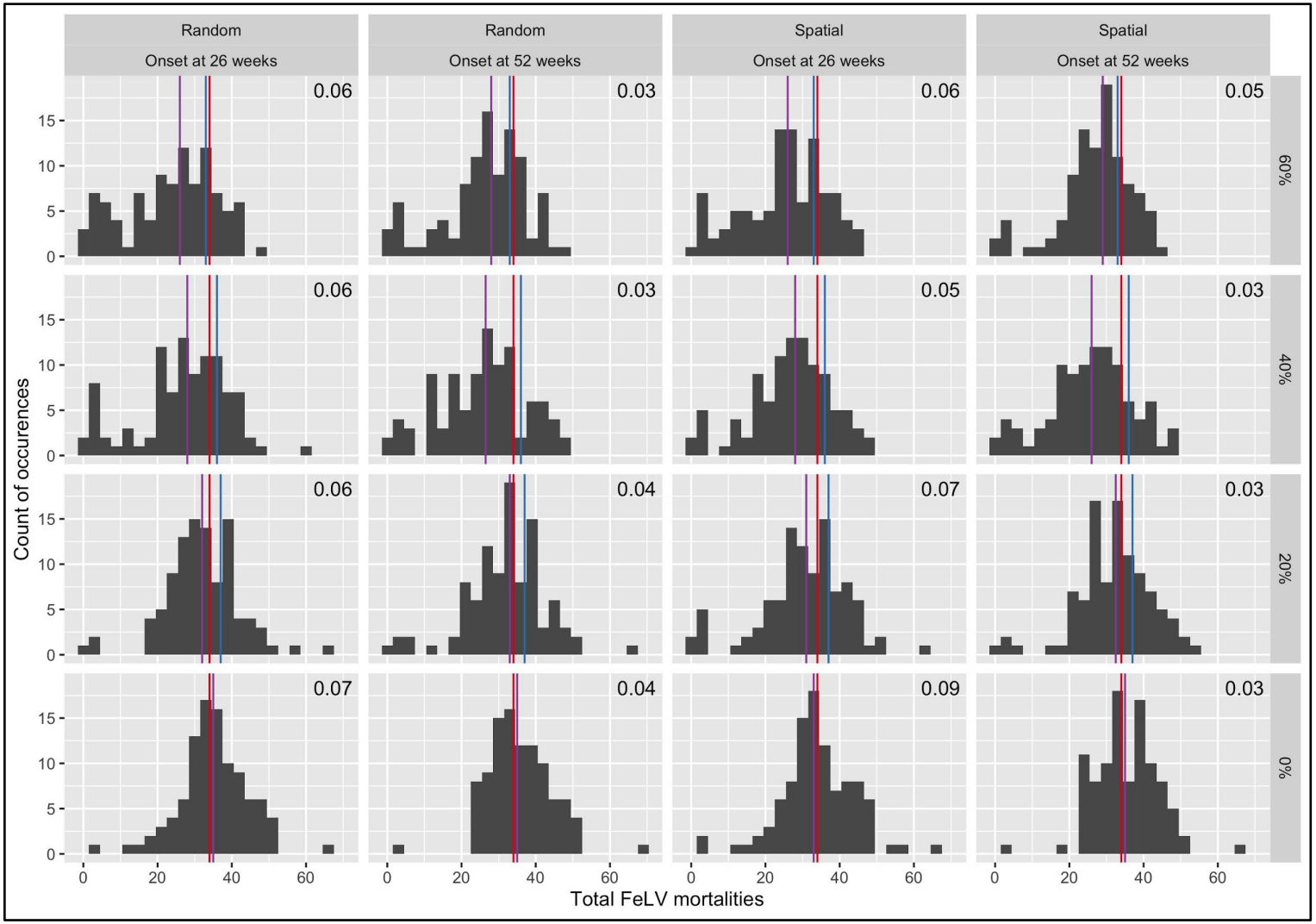

Random
Onset at 26 weeks
Random
Onset at 52 weeks
Spatial
Onset at 26 weeks
Spatial
Onset at 52 weeks
0.06
0.03
0.06
0.05
0.06
0.03
0.05
0.03
0.06
0.04
0.07
0.03
0.07
0.04
0.09
0.03
60%
40%
20%
0%
Count of occurences
Total FeLV mortalities
0
5
10
15
0
20
40
60

Figure C.9: Histograms of FeLV mortalities in simulated epidemics with both reactive test-and-removal and proactive vaccination. Panel rows represent the proportion of the population proactively vaccinated; columns represent both capture/testing strategy (random versus spatial) and timing of onset of reactive test-and-removal after epidemic initiation (26 or 52 weeks). Red vertical lines indicate the median number of mortalities from simulations without interventions; blue lines for proactive vaccination alone; purple lines for the given combination of reactive test-and-removal and proactive vaccination. Numbers in the upper right of each histogram represent the median value for the proportion of capture events that resulted in a removal or humane euthanasia. Each histogram plot represents the results of 100 simulations.

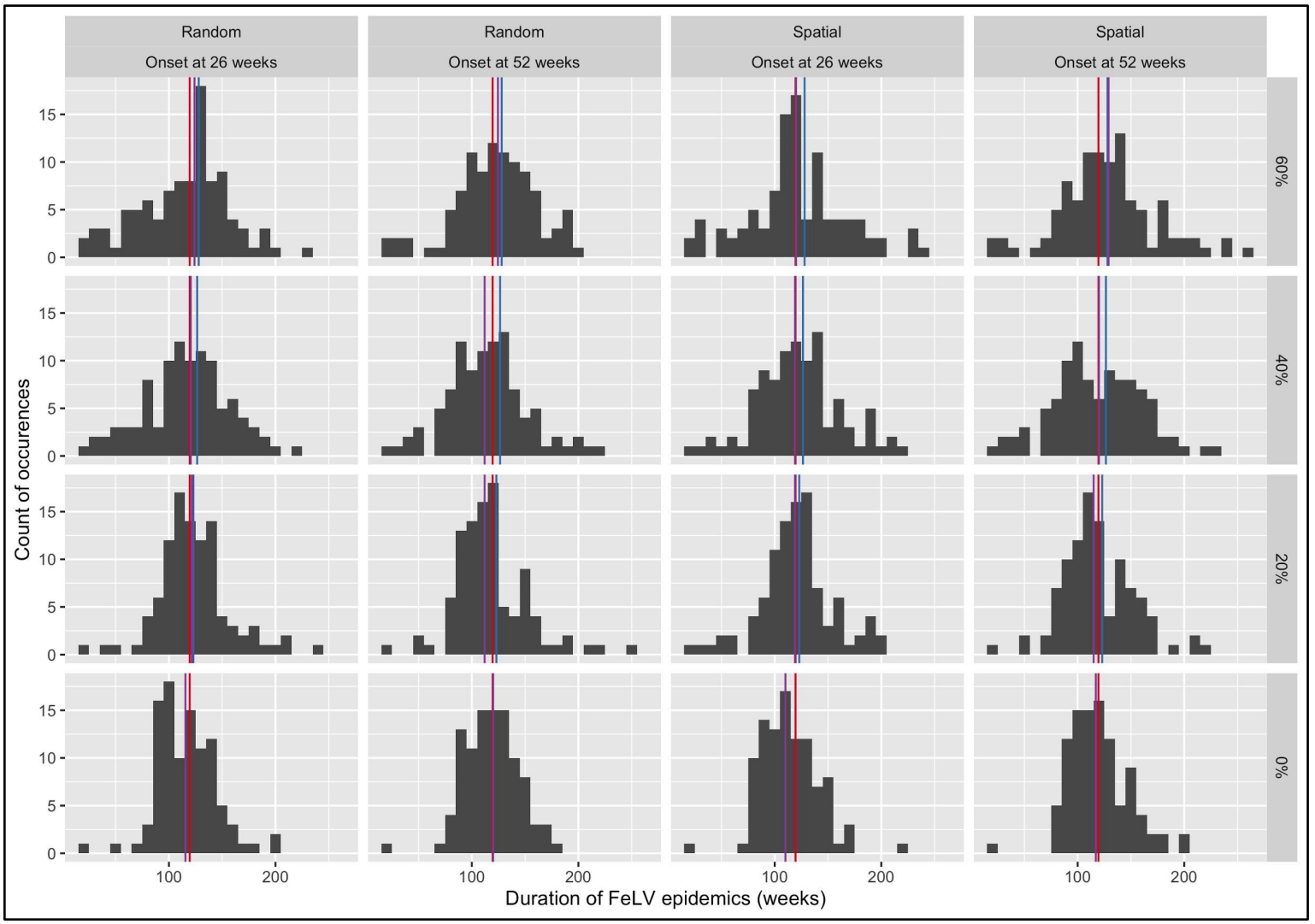

Random
Onset at 26 weeks
Random
Onset at 52 weeks
Spatial
Onset at 26 weeks
Spatial
Onset at 52 weeks
60%
40%
20%
0%
Count of occurences
0
5
10
15
100
200
Duration of FeLV epidemics (weeks)

Figure C.10: Histograms of the duration of simulated FeLV epidemics with both reactive test-and-removal and proactive vaccination. Panel rows represent the proportion of the population proactively vaccinated; columns represent both capture/testing strategy (random versus spatial) and timing of onset of reactive test-and-removal after epidemic initiation (26 or 52 weeks). Red vertical lines indicate the median duration without interventions; blue lines for proactive vaccination alone; purple lines for the given combination of reactive test-and-removal and proactive vaccination. Each histogram plot represents the results of 100 simulations.

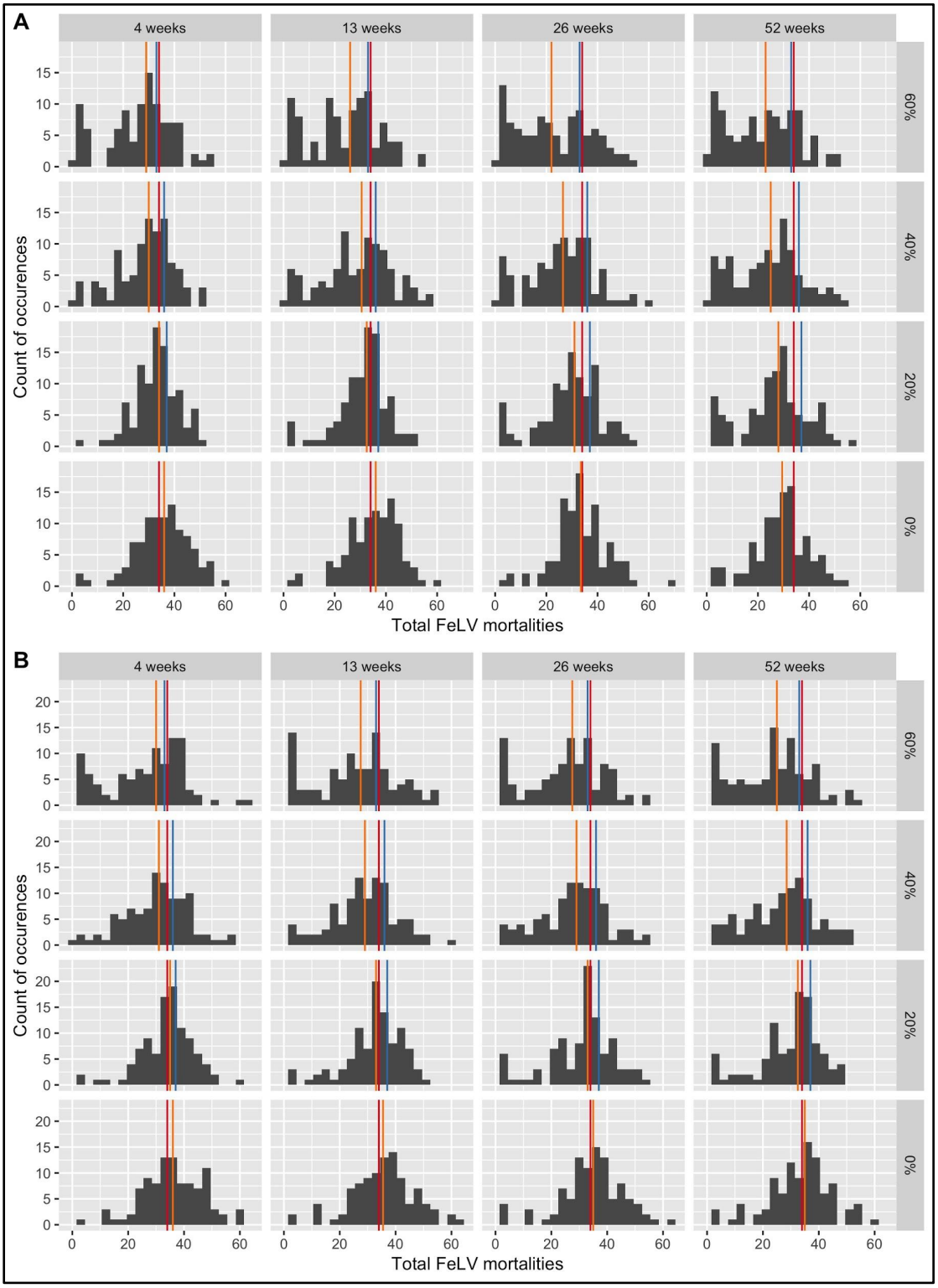

A
4 weeks
13 weeks
26 weeks
52 weeks
60%
40%
20%
0%
Count of occurences
Total FeLV mortalities
B
4 weeks
13 weeks
26 weeks
52 weeks
60%
40%
20%
0%
Count of occurences
Total FeLV mortalities

Figure C.11: Histograms of FeLV mortalities in simulated epidemics with both reactive underpass closures and proactive vaccination. Plot subsections represent (A) onset of underpass closures 26 weeks after epidemic initiation, and (B) onset 52 weeks after initiation. Panel rows represent the proportion of the population proactively vaccinated; columns show the duration of underpass closures. Red vertical lines indicate the median number of mortalities from simulations without interventions; blue lines for proactive vaccination alone; orange lines for the given combination of reactive underpass closure and proactive vaccination. Each histogram plot represents the results of 100 simulations.

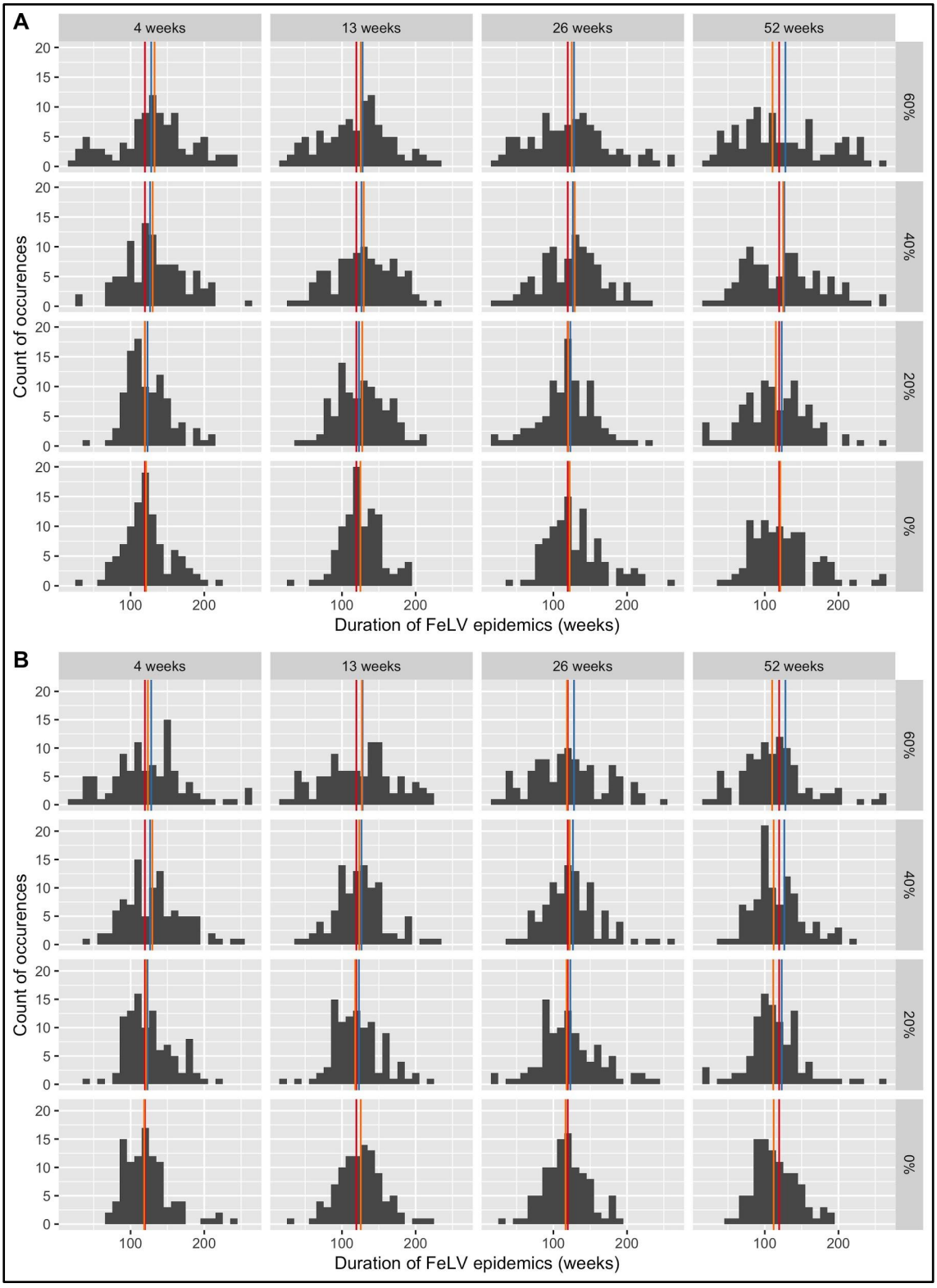

A
4 weeks
13 weeks
26 weeks
52 weeks
60%
40%
20%
0%
Count of occurences
Duration of FeLV epidemics (weeks)
B
4 weeks
13 weeks
26 weeks
52 weeks
60%
40%
20%
0%
Count of occurences
Duration of FeLV epidemics (weeks)

Figure C.12: Histograms of the duration of simulated FeLV epidemics with both reactive underpass closures and proactive vaccination. Plot subsections represent (A) onset of underpass closures 26 weeks after epidemic initiation, and (B) onset 52 weeks after initiation. Panel rows represent the proportion of the population proactively vaccinated; columns show the duration of underpass closures. Red vertical lines indicate the median duration without interventions; blue lines for proactive vaccination alone; orange lines for the given combination of reactive underpass closure and proactive vaccination. Each histogram plot represents the results of 100 simulations.

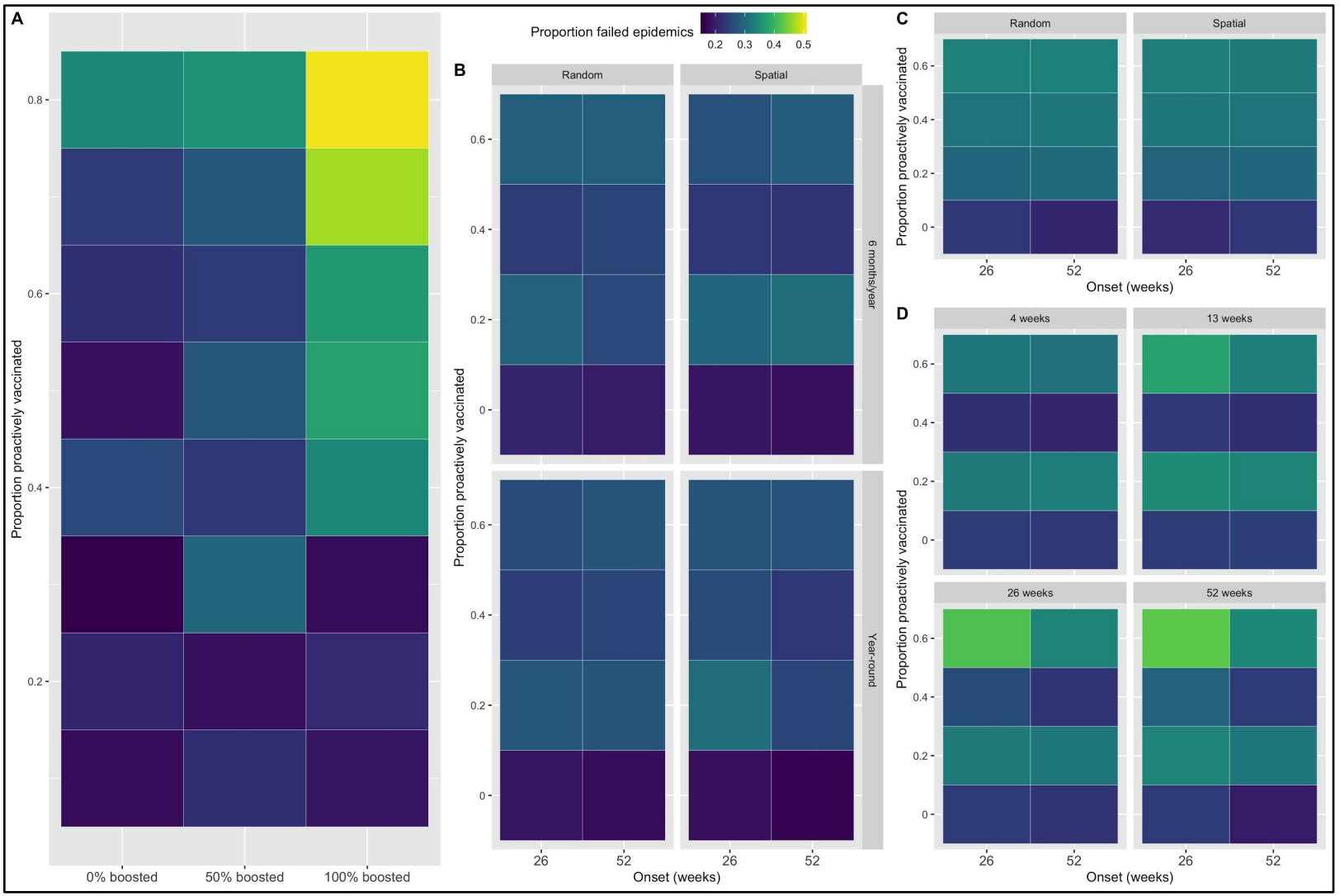

A
Proportion proactively vaccinated
0.8
0.6
0.4
0.2
0% boosted
50% boosted
100% boosted
Proportion failed epidemics
0.2 0.3 0.4 0.5
B
Random
Spatial
6 months/year
Year-round
Proportion proactively vaccinated
0.6
0.4
0.2
0
26
52
Onset (weeks)
C
Random
Spatial
Proportion proactively vaccinated
Onset (weeks)
D
4 weeks
13 weeks
26 weeks
52 weeks
Proportion proactively vaccinated
Onset (weeks)

Figure C.13: Heat maps of the proportion of simulated FeLV epidemics that failed (fewer than 5 progressive or regressive infections) per 100 successful epidemics. Results are shown for (A) proactive vaccination alone; (B) reactive and proactive vaccination; (C) reactive test-and-removal with proactive vaccination; and (D) reactive underpass closures with proactive vaccination. In all plots, the y-axis gives the proportion of the population proactively vaccinated. For (A), the x-axis gives the proportion of individuals receiving boosted vaccination; in all other plots, 50% of all vaccinated individuals received boosted vaccinations. For plots B-D, the x-axis gives the onset of reactive interventions after epidemic initiation (26 or 52 weeks). In B and C, panel columns represent strategy for reactive vaccination and captures/testing, respectively. In addition, panel rows in B show the duration of reactive vaccination per year. Lastly, each panel in D represents the duration of underpass closures. Each colored square represents the results of 100 simulations.

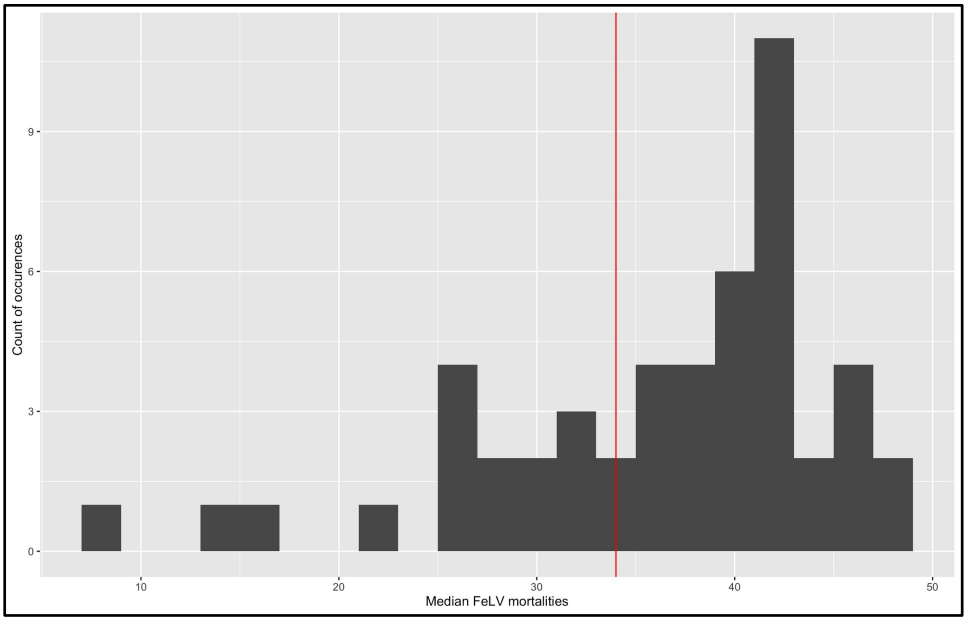


Figure C.14: Histogram of median simulated FeLV mortalities for each of 50 sensitivity analysis parameter sets evaluated under a no-intervention scenario. The vertical red line represents the median mortalities with no interventions from our main analysis simulations.

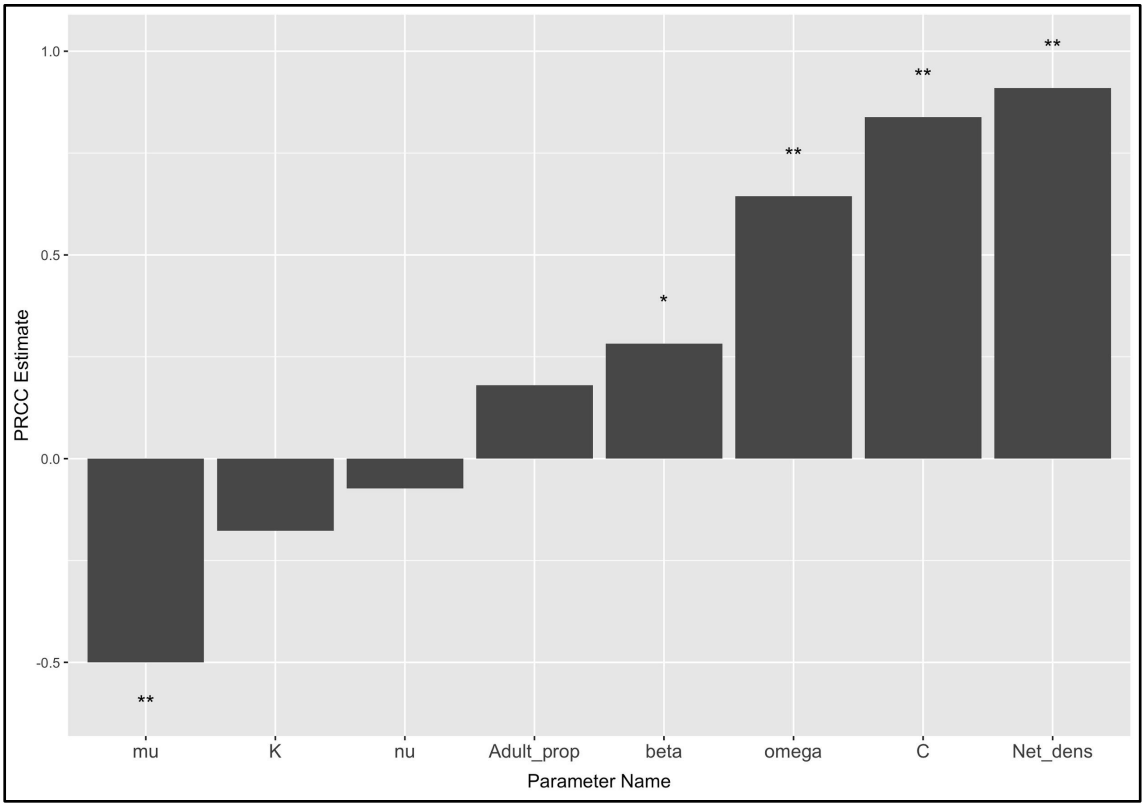


Figure C.15: Partial Rank Correlation Coefficient (PRCC) estimates for simulation parameters in no-intervention scenario sensitivity analysis. The outcome considered was the median mortalities per parameter set. Parameter names and descriptions can be found in Table C.1. Asterisks indicate statistical significance: * = $p \leq 0.05$; ** = $p \leq 0.01$.

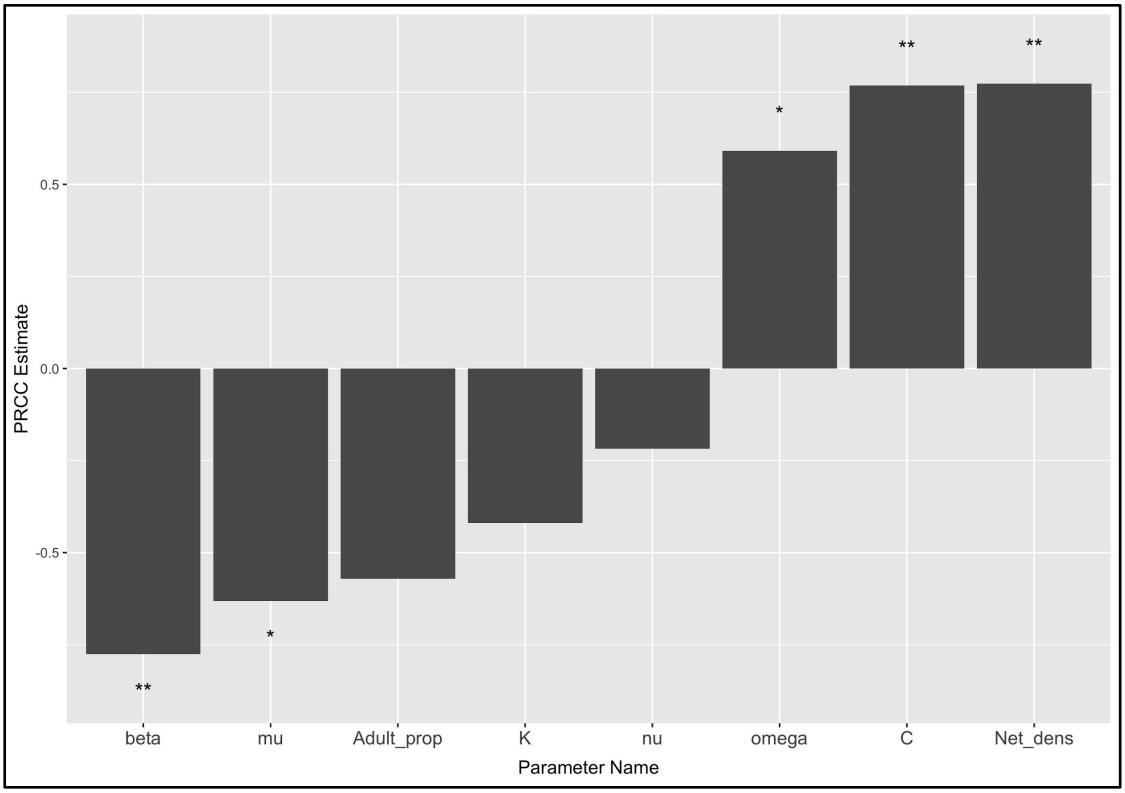


Figure C.16: Partial Rank Correlation Coefficient (PRCC) estimates for simulation parameters in proactive vaccination scenario sensitivity analysis. The outcome considered was the difference in median mortalities with and without proactive vaccination (per parameter set). Shown are results for parameter sets in which 20% of the population was proactively vaccinated with 100% boosting. Parameter names and descriptions can be found in Table C.1. Asterisks indicate statistical significance: * = $p \leq 0.05$; ** = $p \leq 0.01$.

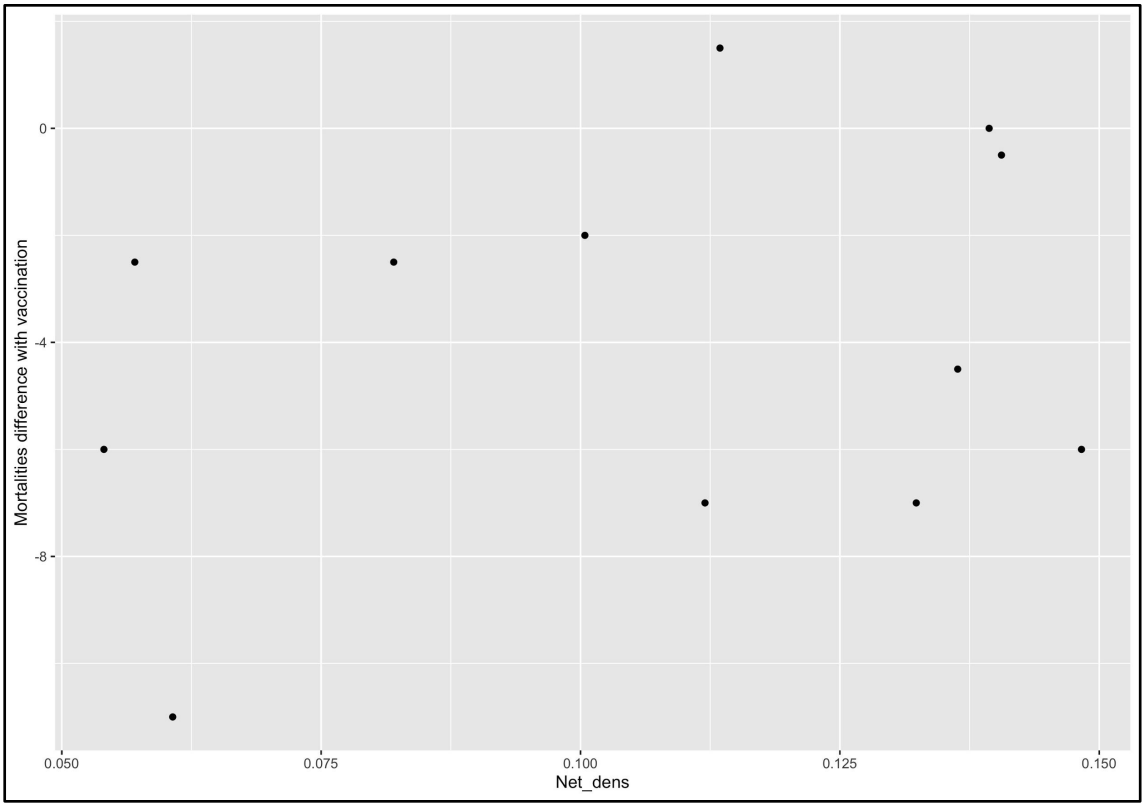


Figure C.17: Scatterplot of difference in median mortalities with and without proactive vaccination (“mortalities difference” per parameter set) against network density parameterization. Shown are results for sensitivity analysis parameter sets in which 20% of the population was proactively vaccinated with 100% boosting.

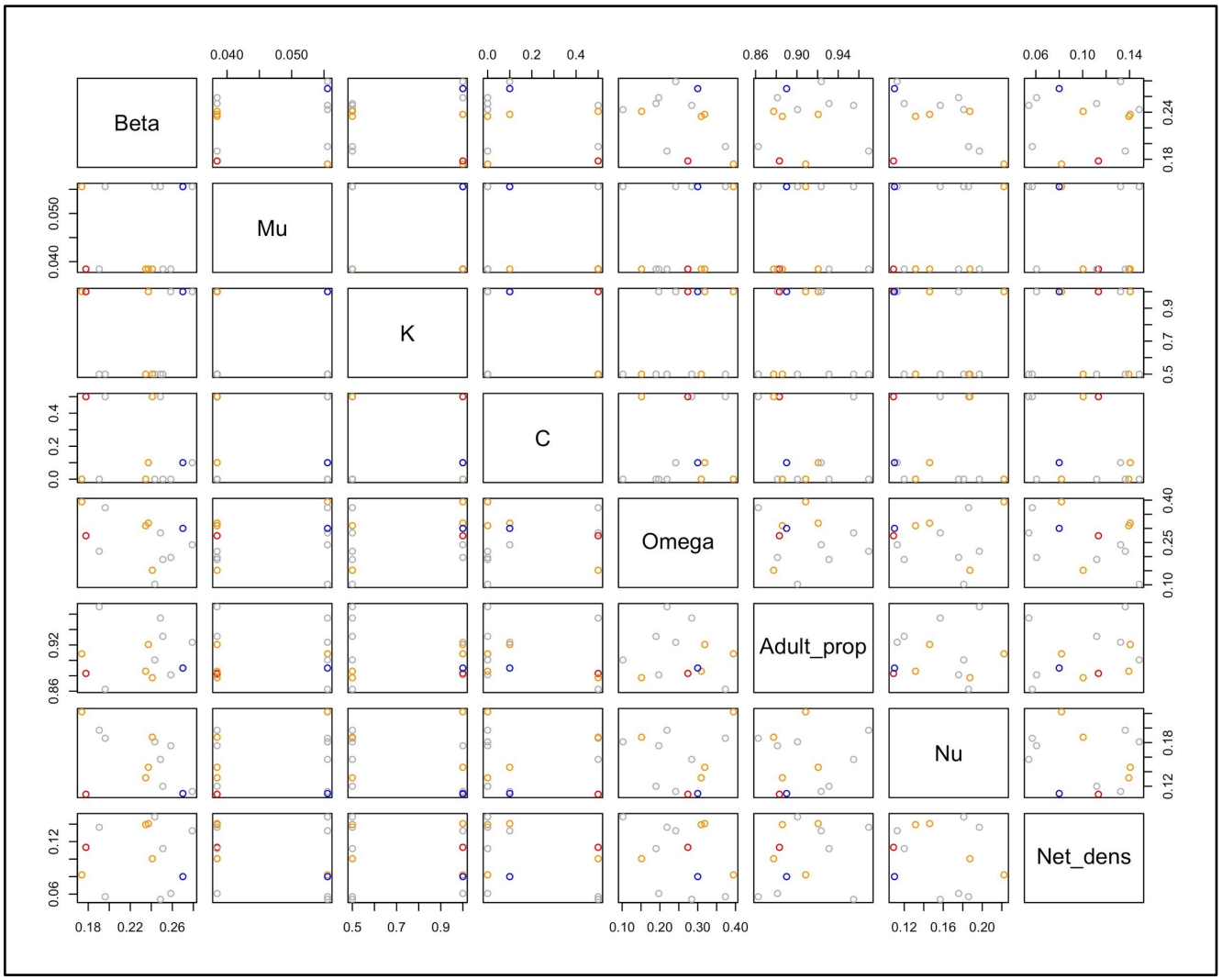

Beta
Mu
K
C
Omega
Adult_prop
Nu
Net_dens
0.040
0.050
0.0
0.2
0.4
0.86
0.90
0.94
0.06
0.10
0.14
0.18
0.24
0.5
0.7
0.9
0.25
0.40
0.92
0.12
0.22
0.26
0.30
0.16
0.20

Figure C.18: Plot of proactive vaccination sensitivity analysis parameter set classifications across all pairs of parameters. Points represent individual parameter sets for the main sensitivity analysis of proactive vaccination (n = 12), and colors represent qualitative classifications of results. The blue point per plot shows the parameter space represented by our main simulations. For all other points, red = increased mortalities with low levels of proactive vaccination; orange = little to no change in mortalities with low levels of proactive vaccination; gray = reduced mortalities with low levels of proactive vaccination.

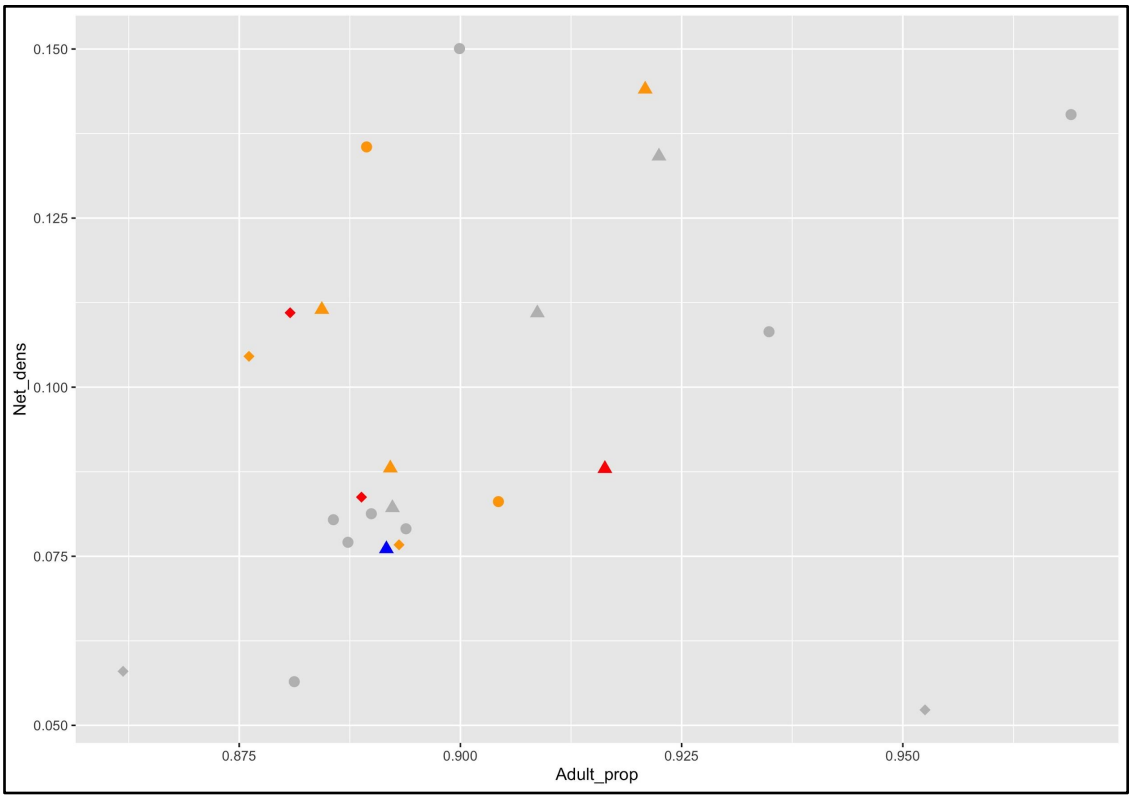


Figure C.19: Scatterplot of post-hoc sensitivity analysis evaluation of influence of proportion adults (Adult_prop) and network density (Net_dens) parameters. Points represent individual parameter sets for both the main sensitivity analysis of proactive vaccination (n=12) and both *post-hoc* parameter sets (n = 4 and 7). Point colors represent qualitative classifications of results. Blue shows the parameter space represented by our main simulations. For all other points, red = increased mortalities with low levels of proactive vaccination; orange = little to no change in mortalities with low levels of proactive vaccination; gray = reduced mortalities with low levels of proactive vaccination. Point shapes give the value for the constant modifying infectiousness of regressives (parameter C): circle = 0, triangle = 0.1, diamond = 0.5. Note: points are jittered for visibility.

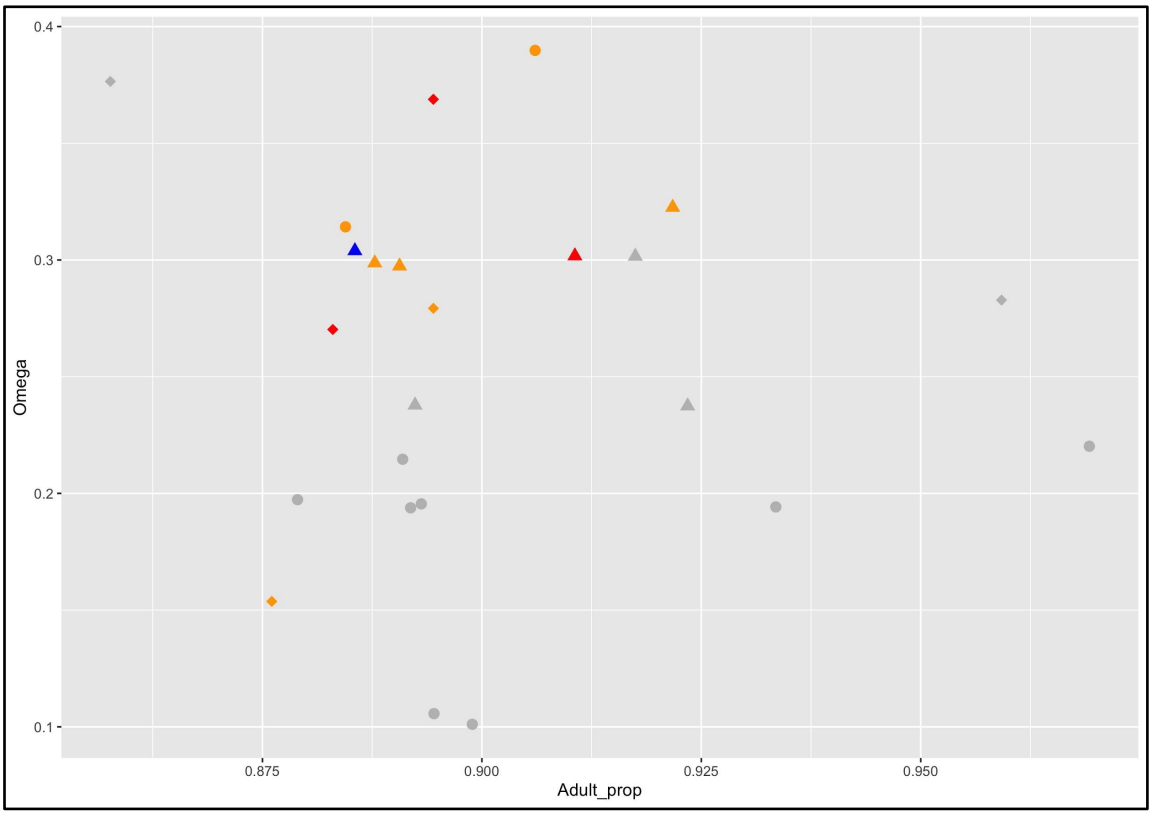


Figure C.20: Scatterplot of *post-hoc* sensitivity analysis results, here showing the proportion adults (Adult_prop) against the weekly probability of contact (Omega/ω) parameters, though ω was not specifically evaluated in *post-hoc* analysis. Points represent individual parameter sets for both the main sensitivity analysis of proactive vaccination (n=12) and both *post-hoc* parameter sets (n = 4 and 7). Point colors represent qualitative classifications of results. Blue shows the parameter space represented by our main simulations. For all other points, red = increased mortalities with low levels of proactive vaccination; orange = little to no change in mortalities with low levels of proactive vaccination; gray = reduced mortalities with low levels of proactive vaccination. Point shapes give the value for the constant modifying infectiousness of regressives

(parameter C): circle = 0, triangle = 0.1, diamond = 0.5. Note: points are jittered for visibility.