COMPLEXITY AND THE INTERSECTION OF SOCIAL AND SEXUAL STRUCTURE, ECOLOGICAL NICHES AND THE EPIDEMIC POTENTIAL OF SEXUALLY TRANSMITTED AND BLOODBORNE INFECTIONS: EMPIRICAL AND THEORETICAL OBSERVATIONS

by

Souradet Yuh-Nan Shaw

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Community Health Sciences
University of Manitoba
Winnipeg

Copyright © 2017 by Souradet Yuh-Nan Shaw
PhD Committee Members

Dr. James Blanchard

Professor, Community Health Sciences & Medical Microbiology, University of Manitoba

Dr. John Wylie

Associate Professor, Medical Microbiology, University of Manitoba
Assistant Professor, Community Health Sciences, University of Manitoba

Dr. Abba Gumel

Professor, Department of Mathematics, University of Manitoba
ABSTRACT

Introduction

Incomplete understanding of the role that context plays in explaining heterogeneity in the transmission dynamics of sexually transmitted and bloodborne infections (STBBIs) has led to deficiencies in prevention and control activities. Like place-based analyses, social network analysis has held much promise for incorporating context to the study of STBBIs. The costs and complexity associated with empirical network data have limited its full potential. Recent advances in the use of exponential random graph models (ERGM) and molecular epidemiology have re-invigorated network-based research. Moreover, ERGM theory focuses on local processes that create global network structure, embodying a generative approach to network formation; this approach contends that networks unfold and evolve predictably, and thus, epidemics should also be similarly predictable. This dissertation aims to combine traditional surveillance methods with advances in network methodologies and orient their use to an applied public health context.

Methods

Using public health surveillance data sources, and focusing on the epidemiology of STBBIs in Winnipeg, Canada, the three studies employ a context-based perspective in understanding the underlying processes that create observable empirical data. Using the Gini coefficient and population attributable fractions, the inequality in the distribution of STBBIs is examined in the context of still-evolving STBBI epidemics. Networks created through molecular genotyping and those created through traditional case-and-contact investigations are compared and contrasted using descriptive statistics and univariate
network metrics. Stochastic simulation modelling, based on the ERGM framework, is used to examine the interaction between pathogen characteristics, mixing patterns and network topology.

**Results**

Each STBBI epidemic examined was revealed to have its own ecological niche, although niches were malleable over time. Examination of geographically-based inequality in the distribution of gonorrhea demonstrated that inequality was decreasing in the context of a growth phase, while also occupying similar geographic space as chlamydia rates increased. Molecular epidemiology is a complementary tool in the construction of sexual networks, revealing potentially hidden links between gonorrhea cases in Winnipeg. The gonorrhea subtype that was most successfully transmitted was associated with chlamydia co-infection. Simulation modelling revealed an intricate relationship between assortative mixing and the infection duration of the pathogen; high levels of assortative mixing muted the trajectory of the modeled STBBI epidemic, with the most drastic effect on infections with shorter duration of infectivity.

**Conclusion**

The three studies cohesively address current challenges in applying context to public health analyses, while expanding our understanding of the mechanisms needed to alter the trajectory of STBBI epidemics. Insights gained from the included analyses form the basis of a proposed context-based surveillance framework.
ACKNOWLEDGEMENTS

Just as nature abhors a vacuum, a PhD dissertation is not completed without a network of support. My thanks first to my advisor, Dr. James Blanchard, whose mentorship, generosity and encouragement has shaped my views on epidemiology, science and public health. To Dr. John Wylie, who first introduced me to social networks, our frequent conversations and your trust in my skills have guided my academic development. Drs. Gumel and Jennings formed my examining committee, and their input and enthusiasm for my dissertation was invaluable.

I am indebted to my numerous colleagues, advisors and friends that I have met along this journey, and whose encouragement and kind words have been instrumental in my academic career. To name but a few: my colleagues at Teen Talk, the Centre for Global Public Health, the National Collaborating Centre for Infectious Diseases and the Winnipeg Regional Health Authority – you have put up with my antics and endured many a late night; Dr. Lisa Lix, Dr. Lawrence Elliott, and Dr. Carole Beaudoin – who all had a hand in honing my quantitative skills and in asking the right questions. To Dr. Margaret Fast, the first epidemiologist I ever met (even before I knew what that term meant) – and whose career and determination was an early inspiration - it is a wondrous thing to enter my career knowing that your colleagues affirm your humility, expertise, and humanity.

All my love to my mother Vene, and my sisters Vichit, Jindara, Thidara, Soudara and Dara – you have been resilient in the face of life’s triumphs and tragedies, and I am humbled by your continual courage - it has been a long journey since Laos. My love as well, to my mother-in-law, Monique Roy. And finally, to Shelley and Nathan, Adeline
and Isaak. Our interactions deeply inspired these pages and are woven throughout the narrative, as I’ve ruminated daily on the impossible web of connections that resulted in all of us being together.
DEDICATION

To my mother, Vene; my wife, Shelley; my sisters Vichit, Jindara, Thidara, Soudara and Dara; to Nathan, Adeline and Isaak. My stars, my stars in the infinite sky.
PREFACE

This statement is to certify that the work presented in this thesis was conceived, conducted, written, and disseminated by Souradet Yuh-Nan Shaw (SYS).

With advice from supervisor Dr. James Blanchard (JFB) and committee members Dr. John Wylie (JLW) and Dr. Abba Gumel (AG), SYS designed the studies, performed the statistical analyses and programmed the simulation work in Chapters 3-5. SYS conducted the review described in Chapter 2.

Deborah L. Nowicki (DLN), Dr. Christopher Green (CG), Craig P. Ross (CPR), Dr. Carla Loeppky (CL), Dr. Joss Reimer (JR), Dr. Pierre J. Plourde (PJP) Dr. Lawrence J. Elliott (LJE), JLW and JFB provided important critical and intellectual feedback on Chapter 3. Data for Chapter 3 were provided by Manitoba Health, Seniors and Active Living (MHSAL).

CPR, JR, PJP, JLW and Dr. Paul van Caeseele (PVC) provided interpretation and intellectual feedback on Chapter 4. Irene Martin (IM) provided technical expertise and interpretation of the NG-MAST methods in Chapter 4. Adriana Mudryj (AM), Nicole Herpai (NH) and Nicole Choptain (NC) provided assistance in analyses and data cleaning.

JLW, AG and JFB provided important scientific and methodological insight for Chapter 5.
# TABLE OF CONTENTS

ABSTRACT ............................................................................................................................ iii

ACKNOWLEDGEMENTS........................................................................................................ v

DEDICATION .......................................................................................................................... vii

PREFACE ............................................................................................................................... viii

TABLE OF CONTENTS .......................................................................................................... ix

LIST OF TABLES .................................................................................................................... xii

LIST OF FIGURES ................................................................................................................ xiii

LIST OF ABBREVIATIONS ..................................................................................................... xvi

CHAPTER 1: BACKGROUND, RATIONALE AND OBJECTIVES ........................................... 1

1.1. INTRODUCTION ............................................................................................................. 1

1.2. THEORETICAL ORIENTATION ................................................................................... 4

1.3. RATIONALE FOR CURRENT STUDY .......................................................................... 4

1.4. STUDY OBJECTIVES & DISSERTATION OVERVIEW ............................................. 7

CHAPTER 2: LITERATURE REVIEW .................................................................................... 9

2.1. INTRODUCTION ............................................................................................................. 9

2.2. HISTORY OF SOCIAL NETWORK ANALYSIS ........................................................... 11

2.3. SOCIAL NETWORK BASICS ........................................................................................ 12

2.4. BASIC SNA METRICS .................................................................................................. 14

2.5. THE IMPACT OF NETWORK STRUCTURE ON TRANSMISSION ............................. 19

2.6. PERSISTENCE OF FORM AND UNDERLYING GENERATIVE PROCESSES .............. 22
PREFACE TO CHAPTER 2

2.7. ANALYTICAL FRAMEWORK ........................................................................................................ 23
2.8. EXPONENTIAL RANDOM GRAPH MODELS .................................................................................. 24
2.9. EXPONENTIAL RANDOM GRAPH MODELS IN THE LITERATURE ........................................ 26

CHAPTER 3: THE EPIDEMIOLOGY OF SEXUALLY TRANSMITTED AND BLOODBORNE INFECTIONS IN WINNIPEG, CANADA: 2007-2016 ................................................................. 32

3.1. INTRODUCTION .......................................................................................................................... 32
3.2. METHODS .................................................................................................................................. 35
3.3. RESULTS ...................................................................................................................................... 40
3.4. DISCUSSION ................................................................................................................................. 52

PREFACE TO CHAPTER 4 .................................................................................................................. 65

CHAPTER 4: SEXUAL NETWORK AND GENOTYPIC ANALYSIS OF AN OUTBREAK OF GONORRHEA IN WINNIPEG, CANADA ......................................................................................... 66

4.1. INTRODUCTION .......................................................................................................................... 66
4.2. METHODS .................................................................................................................................. 69
4.3. RESULTS ...................................................................................................................................... 73
4.4. DISCUSSION ................................................................................................................................. 85

PREFACE TO CHAPTER 5 .................................................................................................................. 94

CHAPTER 5: THE ROLE OF ASSORTATIVE MIXING AND DURATION OF INFECTION ON EPIDEMIC POTENTIAL OF SEXUALLY TRANSMITTED INFECTIONS: A SIMULATION STUDY OF NETWORK DYNAMICS ................................................................. 95

5.1. INTRODUCTION .......................................................................................................................... 95
5.2. METHODS .................................................................................................................................. 98
CHAPTER 6: DISCUSSION, IMPLICATIONS, & CONCLUSIONS ............. 123

6.1. SUMMARY OF RESEARCH FINDINGS .................................................. 123

6.2. STUDY STRENGTHS AND UNIQUE CONTRIBUTIONS TO THE LITERATURE ........................................................................................................................... 125

6.3. IMPLICATIONS .......................................................................................... 127

6.4. CONCLUSION ............................................................................................ 138

BIBLIOGRAPHY ............................................................................................... 139

APPENDIX 1: SOCIAL NETWORK BASIC METRICS CALCULATION .... 174

APPENDIX 2: DERIVATION OF GINI INDEX .................................................. 176

APPENDIX 3: TRIAD CENSUS ...................................................................... 177
LIST OF TABLES

Table 3.1: Frequency, Crude and Age-Standardized Rates (per 100,000), Sexually Transmitted and Bloodborne Infections, Winnipeg Health Region (2007-2016) .... 43

Table 3.2: Results from Poisson Regression models, Sexually Transmitted and Bloodborne Infections, Winnipeg Health Region .................................................. 44

Table 3.3: Indices of Geographic Concentration, Sexually Transmitted and Bloodborne Infections in the Winnipeg Health Region, by Time Period .................................... 47

Table 3.4: Spearman’s Rank Correlation, Comparing Age-Adjusted Rates in Gonorrhea Outbreak Years, by Neighbourhood Cluster ................................................ 51

Table 4.1: Gonorrhea Subtypes by Socio-Demographic and Testing Characteristics, Winnipeg Health Region (2014) ........................................................................ 76

Table 4.2: Subtype-3671 vs. all Other Gonorrhea Subtypes, Winnipeg Health Region (2014) ........................................................................................................ 78

Table 4.3: Adjusted Odds Ratios (OR) and 95% Confidence Intervals (95%CI) from Logistic Regression Models, Socio-demographic and Testing Characteristics Associated with Subtype-3671 (>15 years of age) ............................................................. 79

Table 4.4: Component Sizes, Case and Contact Investigations, and from Gonorrhea Genotyping, Winnipeg Health Region (2014) .................................................... 83

Table 4.5: Univariate Statistics, Sexual Network Constructed from Case & Contact Investigations, Winnipeg Health Region (2014) ..................................................... 84

Table 5.1: Model Parameters for Network Models .............................................................................. 102

Table 5.2: Prevalence Rate (per 100 population) and Rate Ratio, by Risk Group and Assortativity Level .............................................................................................. 106
LIST OF FIGURES

Figure 3.1: Incidence of Select Sexually Transmitted and Bloodborne Infections, Winnipeg Health Region 2007-2016 (logarithmic scale) ................................................. 42

Figure 3.2: Age-Adjusted and Relative Rates of Gonorrhea by Neighbourhood Clusters, Outbreak Years (2012 & 2016) vs. Endemic Rate (2009-2011 & 2013-2015) ........ 49

Figure 3.3: Age-Adjusted Rates of Gonorrhea & Chlamydia by Neighbourhood Clusters, Gonorrhea Outbreak Years (2012 & 2016) vs. Endemic Rate (2009-2011 & 2013-2015) .................................................................................................................................................. 51

Figure 3.4: Map of Sexually Transmitted and Bloodborne Infection Cases in the Winnipeg Health Region, 2007-2009 ................................................................. 61

Figure 3.5: Map of Sexually Transmitted and Bloodborne Infection Cases in the Winnipeg Health Region, 2010-2012 ........................................................................ 62

Figure 3.6: Map of Sexually Transmitted and Bloodborne Infection Cases in the Winnipeg Health Region, 2013-2015 ........................................................................ 63

Figure 3.7: Map of Sexually Transmitted and Bloodborne Infection Cases in the Winnipeg Health Region, 2016 .............................................................................. 64

Figure 4.1: Distribution of Gonorrhea Subtypes, Winnipeg Health Region (2014, N=126) .............................................................................................................. 75

Figure 4.2: Sexual Contact Network from Case and Contact Investigations and Gonorrhea Subtypes, Winnipeg Health Region (2014) .................................................... 81

Figure 4.3: Results from Conditional Uniform Graph Simulations of In-degree and Out-degree Centralization Measures .............................................................................. 85

Figure 5.1: Prevalence of Long Duration Pathogen From Estimated Networks as Assortativity Between High- and Low-Activity Groups is Modified ....................... 103

Figure 5.2: Prevalence of Short Duration Pathogen From Estimated Networks as Assortativity Between High- and Low-Activity Groups is Modified ....................... 104

Figure 5.3: Average Cumulative Incidence of Simulated Epidemics of Long- and Short-Duration Pathogens Within Networks Estimated from Separable Temporal Exponential Random Graph Models as Assortativity Between High- and Low-Activity Groups is Modified ........................................................................... 105
Figure 5.4: Prevalence Rate of Simulated Epidemics of Short-Duration Pathogens Within Networks Estimated from Separable Temporal Exponential Random Graph Models, by Risk Group and Assortativity Level ................................................................. 108

Figure 5.5: Prevalence Rate of Simulated Epidemics of Long-Duration Pathogens Within Networks Estimated from Separable Temporal Exponential Random Graph Models, by Risk Group and Assortativity Level ................................................................. 109

Figure 5.6: Network Visualization Showing Infected and Non-Infected Nodes, Simulated Epidemic of Short Duration Pathogen Within Network Estimated from Separable Temporal Exponential Random Graph Models, at 50% Assortative Mixing ............ 110

Figure 5.7: Network Visualization Showing Infected and Non-Infected Nodes, Simulated Epidemic of Short Duration Pathogen Within Network Estimated from Separable Temporal Exponential Random Graph Models, at 99% Assortative Mixing ............ 111

Figure 5.8: Largest Component in Network Estimated from Separable Temporal Exponential Random Graph Models at 2 Assortative Mixing Levels ............... 112

Figure 5.9: Framework for Describing Epidemics ................................................................. 114

Figure 5.10: An Example of the Impact of Assortative Mixing on Epidemic Potential. 122

Figure 6.1: Schematic of Traditional Case-Based Surveillance ......................................................... 128

Figure 6.2: Schematic of Traditional Case-Based Surveillance: Pathogen Comparison 129

Figure 6.3: Schematic of Context-Based Surveillance ................................................................. 129

Figure 6.4: Schematic of Context-Based Surveillance: Incorporating Case-Based and Case-Contact Data .......................................................................................... 130

Figure 6.5: Schematic of Context-Based Surveillance: Incorporating Case-Based, Case-Contact, and Place-Based Data ........................................................................ 131

Figure 6.6: Schematic of Context-Based Surveillance: Incorporating Case-Based, Case-Contact, and Molecular Data ........................................................................ 132

Figure 6.7: Schematic of Context-Based Surveillance: Incorporating All Levels of Data ........................................................................................................ 133

Figure 6.8: Application of the Context-Based Surveillance Framework to Data from Chapter 4........................................................................................................ 134

Figure 6.9: Schematic of Context-Based Surveillance: Co-infection Success .................. 135
Figure 6.10: Schematic of Context-Based Surveillance: Co-infection Failure............. 136

Figure 6.11: Schematic of Context-Based Surveillance: Low Level of Assortative Mixing
............................................................................................................................................. 137

Figure 6.12: Schematic of Context-Based Surveillance: High Level of Assortative Mixing
............................................................................................................................................. 137
LIST OF ABBREVIATIONS

95% CI: 95% Confidence Intervals
AIDS: Acquired Immuno-Deficiency Syndrome
AIRR: Adjusted Incidence Rate Ratio
AOR: Adjusted Odds Ratio
CA: Community Area
CDC: Centers for Disease Control
CPL: Cadham Provincial Laboratory
CT: Chlamydia
CUG: Conditional Uniform Graph
DNA: Deoxyribonucleic Acid
ERGM: Exponential Random Graph Model
HCV: Hepatitis C
HIV: Human Immunodeficiency Virus
MCMC: Markov Chain Monte Carlo
MHSAL: Manitoba Health, Seniors and Active Living
MSM: Men Who Have Sex With Men
MSMW: Men Who Have Sex With Men And Women
NAAT: Nucleic Acid Amplification Test
NC: Neighbourhood Cluster
NG: Gonorrhea
NG-MAST: Neisseria gonorrhea Multi Antigen Sequence Typing
NML: National Microbiology Laboratory
PAF: Population Attributable Fraction
PDBC: Post-Diagnosis Behaviour Change
PHN: Public Health Nurse
PrEP: Pre-Exposure Prophylaxis
PWID: People Who Inject Drugs
RNA: Ribonucleic Acid
SFHR: Social Factors and HIV Risk
SIS: Susceptible-Infectious-Susceptible
SNA: Social Network Analysis
ST: Subtype
STI: Sexually Transmitted Infection
STBBI: Sexually Transmitted and Bloodborne Infection
STERGM: Separable Temporal Exponential Random Graph Model
UIRR: Unadjusted Incidence Rate Ratio
US: United States
WHR: Winnipeg Health Region
WRHA: Winnipeg Regional Health Authority
CHAPTER 1: BACKGROUND, RATIONALE AND OBJECTIVES

1.1. INTRODUCTION

Researchers have suggested epidemiology, as a science, has long been hampered by its inattention to context in understanding heterogeneity in disease transmission dynamics. According to prominent researchers such as Sandro Galea and Mervyn Susser,\(^1\)-\(^4\) this inattention to context is exemplified by the prodigious use of regression models in epidemiology, where the main intent is to isolate the effect of a variable of interest from all other contextual variables. Methods that incorporate context, such as place-based analyses, cluster analysis and social network analysis have been so far underutilized. This inattention to incorporating context has resulted in an incomplete ability to explain why some groups of individuals are at higher risk of disease acquisition and transmission over others. This incomplete understanding, in turn, has led to deficiencies in prevention, intervention and control activities.

Since the late 1990s social network analysis (SNA) has held much promise in incorporating context to the study of infectious disease epidemiology. The use of SNA in infectious disease epidemiology has demonstrated that network structure can influence the transmission of infectious diseases, such as sexually transmitted and bloodborne infections (STBBIs).\(^5\) In the production of infectious diseases, and especially in the context of STBBIs, social relationships matter. SNA explicitly states that networks are the formal structure of these relationships; it is through these structures that infectious pathogens are propagated and transmitted. As Morris states, “...the diffusion of pathogens through a human population traces the structure of social networks. The pattern of
spread is jointly determined by the biology of the pathogen and the social structure that can support it...” (p. 109: 6).

Given that the “pattern of spread” of a disease, or its epidemiology, is determined by the interaction between the biology of the pathogen, and unique local social structures, and considering sexually transmitted infections are propagated along the most exclusive form of social contact, why then, are similar patterns of STBBI epidemiology observed across a wide range of geographies globally? For example, throughout much of the late 20th century, infectious syphilis was thought to have all but been eliminated; however, starting at the end of the first decade of the 21st century, infectious syphilis has now fully re-emerged in most industrialized nations. At the same time that there seems to exist, at a very broad level, global alignment in secular trends of STBBI, there exists persistent and substantial differentials in STBBI prevalence across subpopulations at the local level, whether measured by age, sex, race or socio-economic position. Moreover, not only do STBBI rely on the most intimate type of contact, they also require a specific sequence of events for propagation – the infected case must transmit the pathogen to a susceptible contact or contacts within a limited time period.

The studies within this dissertation attempt to explore these two seemingly oppositional and contradictory forces – that the success of STBBI relies on the most intimate type of contact, and a very specific sequence of events; and secondly, that differentials in STBBI prevalence between subpopulations, and across localities, are persistent and substantial. We theorize understanding this paradox requires moving beyond the individually-centred research paradigm, and more intently exploring how context plays a large role in determining the outcome of STBBI epidemics.
Understanding context, whether geographically- or relationship-based, may help to explain the possible trajectories that sexually transmitted epidemics can take. Although the main focus will be on the use of SNA, other traditional forms of analyses, as well as non-traditional place-based analyses will be incorporated into the studies included in the dissertation.

One of the central questions for SNA is the explicit role that network structure plays, either alone, or in concert with both individual-level characteristics and broader-level determinants, in the production and the sustainability of infectious diseases. Up until recently, and due to important complexities and challenges on both the analytical and research design sides, SNA has mostly been used in a descriptive manner. Networks, by definition, are contingent upon connections between people and/or places, and thus the data they generate are dependent. Although this dependency is at the heart of what makes network analysis insightful, the dependent nature of network data has made the statistical analysis of social network data a challenge for traditional regression-type analytical techniques, as the very structure of network data violates the fundamental assumption of independent and identically distributed random variables.

Over the last decade, however, the use of statistical frameworks for the analysis of social network data has accelerated, leading to advances in hypothesis testing, and a more complete understanding of the types of structures that contribute to infectious disease propagation and sustainability. More specifically, the development of exponential random graph models (ERGMs), and its longitudinal corollary, separable temporal exponential random graph models (STERGMs), has allowed researchers to incorporate the dependent nature of network data into statistical models. The outputs of these models
are similar to those produced from traditional regression-type models, and are thus easily translated by, and accessible to, public health audiences. Moreover, because the statistical frameworks allow comparability between different networks, a more generalizable theory of network formation can be formulated.

1.2. THEORETICAL ORIENTATION

Equally important, the statistical network modelling perspective is consistent with a *generative* view of network formation: that the complex structures of networks arise as an agglomeration of substructures generated from the micro-level interactions between individuals. That is, local rules beget global structures. Networks are not irreducibly complex, but can be viewed as emergent from a collection of substructures interacting with each other. Therefore, in the context of STBBI transmission, two key goals would be to: 1) Understand the generative processes leading to the formation of these relevant substructures; and 2) Understand how relevant substructures impact transmission dynamics. Having a clearer understanding of both could ultimately inform efforts to break both the chain of transmission, and the underlying structure that promotes transmission in the first place.

1.3. RATIONALE FOR CURRENT STUDY

Despite a tremendous amount of public health resources invested in their elimination, the incidence of bacterial STBBIs, such as chlamydia and gonorrhea remain unabated across the western hemisphere. Moreover, infections thought close to being eliminated at the end of the 20th century, such as syphilis, have re-emerged in the past decade. At the same time, new digital technologies, such as on-line “apps” connected to
ubiquitous smartphones, have changed the landscape of how individuals meet, socialize, and seek sexual relationships, complicating traditional public health infectious disease control tools such as case and contact investigation and follow-up.

Although SNA has emerged as a tool to help understand transmission dynamics of STBBIs, the application of SNA to real-world public health settings has been limited by its predominant use as a descriptive tool. With the advent of a rigorous statistical framework for network analyses, an opportunity exists to advance social network theory in an applied public health manner. Identification of substructures important to the transmission and acquisition of STBBIs facilitates theories on the generative processes that create these substructures. Knowledge of the generative processes underlying substructures important to the spread of STBBIs enables public health practitioners to design interventions that target network structure, which has the potential to be more effective in curbing transmission of STBBIs.

At the same time, the computationally-intensive routines used to solve the dependency issue when analyzing networks within a statistical framework allows for the simulation of networks centred around empirical observations, as well as modeling disease transmission processes over these simulated networks. When planning for public health interventions, this ultimately permits a counterfactual framework to be realized, in the form of mathematical models, which are increasingly becoming invaluable tools in the context of limited resources to address the obduracy of STBBIs.

Therefore the studies included in this dissertation will build on traditional surveillance tools and methodologies, and explore methods to understand sexual networks, and finally, help to introduce statistical network modelling as a valuable tool in
the epidemiological toolbox, providing insight into disrupting transmission networks, as well as providing a means for evidence-based program planning. With respect to implications of the generative network science paradigm, Martina Morris writes:

“\textit{There are also important practical implications if (the generative network science paradigm) is true. If it turns out that...local rules explain most of the variation in network structure that is relevant to disease spread, then we have a simple inexpensive way to rapidly assess network vulnerability for public health surveillance and some simple behavioural rules that people can be taught to recognize and change.}” 6(p. 122)

The question then, is whether overall network structure can be explained with a small number of partnership formation rules that operate at the local level, and whether or not an understanding of the role of different network substructures, as well as their interactions can help researchers understand why some networks are more successful at transmission over others.

If both can contribute to the extant SNA field, it will radically simplify data collection needs and tools, and give network analysis a central place among the tools available for STBBI research and intervention. As Adams et al. suggest, having this more nuanced understanding of social structure, and how the intersection between structure and ecological niche influences epidemic potential, can help researchers move beyond the view that networks are “…a connection of pipes that carry disease” 7(p. 328). This more nuanced perspective acknowledges that given an initial set of conditions, networks are
not chaotic and intractably complex, but that they unfold and evolve in fully predictable ways.

1.4. STUDY OBJECTIVES & DISSERTATION OVERVIEW

Taken together, the projects in this dissertation aim to combine traditional surveillance techniques and place-based analyses with recent advances in network methodologies, including molecular epidemiology of NG and statistical network modelling, in a cohesive manner, oriented towards the investigation of STBBI epidemics from the applied public health perspective.

Chapter 3 builds on pioneering work in Manitoba, and lays the foundation for understanding the context of STBBI epidemiology in Winnipeg, Canada. This chapter explores the use of traditional surveillance tools, augmented with other context-oriented analyses – specifically, place-based analyses using the Gini coefficient and population attributable fractions. Data for this first study were collected as part of routine public health surveillance between 2007 and 2016. Secular trends in four different STBBI: chlamydia (CT), gonorrhea (NG), infectious syphilis, and the human immunodeficiency virus (HIV) were examined, in addition to trends in CT/NG co-infection. The objectives of this first study were to describe the recent epidemiology of STBBI in Winnipeg, Canada; incorporate traditional surveillance tools, with place-based analyses; and compare and contrast the ecological niches of different STBBIs, in the context of their ever-evolving epidemiology.

In Chapter 4, the focus will shift to NG. Traditional case-and-contact techniques to ascertain sexual networks are compared to more recent advances in molecular epidemiology. Data for this study are from an enhanced public health investigation of NG
collected over a 10-month period in 2014 in Winnipeg, Canada. The objectives of Chapter 4 were to: 1) determine the distribution of NG subtypes in a population-based sample of NG infections reported to public health; 2) describe the subtypes by socio-demographic and sexual partnership attributes; and 3) to compare networks created through genotyping to those created by case-contact data.

The third study will exploit the ability of ERGMs to simulate both networks and epidemics over time, and will examine gaps in knowledge regarding the dynamic nature of sexual networks, and the interplay between STBBI epidemics and how they co-evolve with network structures. In addition to simulating networks and epidemics, ERGMs will be used to uncover important network substructures at different epidemic phases (e.g., growth, plateau and decline). Knowledge of prominent phase-specific substructures, as well as recognition of epidemic trajectory within each phase, has potential application in tailoring public health responses more appropriately to the epidemic at hand, making the use of public health resources more effective and efficient.
CHAPTER 2: LITERATURE REVIEW

2.1. INTRODUCTION

Despite intensive prevention and research activities, the transmission and acquisition of sexually transmitted and bloodborne infections (STBBIs) continues to persist globally, with certain populations at disproportionately higher risk for STBBIs. It has been demonstrated that interventions solely targeting individual risk behaviours have had limited success in curtailing STBBI epidemics. Given this limited success, researchers have called upon incorporating an individual’s socio-epidemiologic context to explain heterogeneity in STBBI rates, emphasizing that focusing on risk behaviors alone discounts the notion that factors at multiple levels interact to create complex risk environments, ultimately impacting the distribution of health in a population. Thus, there is a realization by epidemiologists, researchers and policymakers that in the case of STBBI epidemics, context matters, and importantly, populations are not homogenous entities, allowing STBBIs to spread uniformly within them.

Social network analysis (SNA) is one methodology which explicitly incorporates an individual’s context to explain heterogeneity. The use of SNA for understanding epidemics of STBBIs has provided a rich body of literature demonstrating the importance of networks in understanding heterogeneity observed in STBBI epidemics. Importantly, SNA has been able to explain why persons/groups can have similar risk behaviors and yet one may have a much greater risk of contracting or transmitting STBBIs.

At its simplest, social networks consist of actors, the attributes of those actors, and the ties that bind the actors together. By examining social structure, recognizing local
and global patterns and properties, detecting influential (“central”) individuals and structures, as well understanding how networks evolve over time, the SNA perspective allows researchers the ability to theorize and explain how risk is not only determined by what people do, but that risk is also intimately connected to whom they do things with.\textsuperscript{16} Sexual networks are a distinct subset of SNA, with sexual networks composed of groups of people who are connected to each other through sexual relations. How these networks are configured, such as the average number of ties individuals have with each other, the proportion of concurrent relationships, how “central” individuals are, and the places where people congregate to meet each other has been shown to influence transmission dynamics of STBBIs.\textsuperscript{5,26-34} An example illustrating the influence of sexual networks upon the presence of STBBIs was demonstrated by Laumann et al., who used a nationally-representative sample of the US population and found rates of heterosexually-spread bacterial STIs were almost five-fold higher among African-Americans, compared to Caucasians, even after adjusting for individual risk factors.\textsuperscript{23} The authors concluded that African-Americans partner choices were more homogeneous (towards other African-Americans), thus leading to segregation from other racial groups. Compounding the segregation, there was a lack of available men, due to such factors as socio-economic inequities leading to high incarceration rates, and thus non-core (e.g., less risky) African-American women were left to “choose” core (i.e., higher risk) African-American men, relative to Caucasian women.\textsuperscript{23}
2.2. HISTORY OF SOCIAL NETWORK ANALYSIS

Modern methods of SNA stem from the work of Jacob Moreno and his use of sociograms to examine the relationship between elementary school students in the 1930s. Early public health efforts by Thomas Parran and others in developing contact tracing methods were partly inspired by Moreno’s work, and was used primarily to interview cases of syphilis, identify contacts to cases and then to break the chain of transmission. This early use of Moreno’s sociograms illustrates the fundamental utility of SNA: understanding how network structure influences transmission dynamics of STBBI results not only in better insight on how STBBI propagate through subgroups in a population, but because the underlying sexual network also traces potential transmission routes, knowledge of the structure can be used to inform control of STBBI.

As Wasserman et al. note, since Moreno’s time, the use of social network analysis in its many forms had grown steadily, if not exceptionally, until the early 1990s. Prior to the advent of SNA post-1990, two fundamental insights were discovered: the connections in observed empirical networks were distributed in a nonrandom fashion, and patterns of connections could often be reduced to a set of functionally equivalent positions. That is, although the micro- and macro-structural forms of networks are almost infinitely complex, successful propagation through a network depended on a smaller set of principles, such as high connectivity, which many forms shared.

At the same time that the theoretical underpinnings of SNA were being established and developed, the practical importance of networks in the transmission of STBBI was being demonstrated by researchers from the Centers for Disease Control (Atlanta) during the early days of the HIV/AIDS epidemic. SNA has also been used
to describe the importance of “hot spots”, or places where sexual networks converge, illustrated in studies examining a variety of pathogens, including gonorrhea,\textsuperscript{33,41} HIV,\textsuperscript{42} and syphilis.\textsuperscript{43} Relatedly, the ability of SNA to find important bridges between geographically distant locations has been another prominent use.\textsuperscript{44}

\section*{2.3. SOCIAL NETWORK BASICS}

The following is a brief introduction to social network basics. It should be noted that the interdisciplinary nature of SNA has produced an abundance of network measures, with the enumeration of the universe of metrics beyond the scope of this review;\textsuperscript{37} here, the intention is only to provide a very basic overview of prominent network measures. In SNA, the basic unit of interest is the combined sets of actors, and the shared relations between them. Actors, variously referred to as nodes, vertices or points, are typically represented with points, or circles and other geometric shapes. The “ties” between actors, sometimes referred to as edges, or arcs, are represented with lines. Relationships can be binary (e.g., absent or present) or valued (e.g., “How close are you to Person X?”), directed (e.g., Person X seeks advice from Person Y, but Person Y does not necessarily seek advice from Person X) or undirected (the presence of a tie denotes a mutual relationship, e.g., friendship). The connected unit (\(ij\)) is then usually referred to as a \textit{dyad}.

Networks are typically either “one-mode” or “two-mode”. One-mode networks are classified based on direct contact between nodes, with nodes belonging to the same class, such as the familiar person-to-person relationships observed in for example, a network comprising classmates in the same school. Two-mode networks, on the other hand, are based on nodes from two separate classes, with the most common being one set of nodes composed of people, and the other set composed of places. In this example,
people are connected through shared places. For the purposes of this dissertation, the focus will only be on one-mode networks.

There are two broad classes of networks, *egocentric* and *macroscopic*, with each class having implications on how data are collected, analyzed and ultimately, interpreted. Relatively speaking, the methodology used to gather information on, and measure the influence of egocentric networks (also called an *ego-network*) is the most straightforward, although by no means is the process simple. The typical manner in which an egocentric network is constructed is through the use of cross-sectional quantitative surveys whereby a respondent (i.e., ego) is asked about other people they are connected to (i.e., alters). Thus, egocentric networks focus on individuals and their “local” connections. The actual manner in which people are connected is defined by the research question (e.g. friendship, sexual partners, etc.). Information about alters is then provided by the egos that named the alters. As Marsden notes, although studies are often distinguished by employing either a whole network or egocentric design, this delineation is somewhat artificial, as whole networks can be constructed from egocentric data if, for example egos are sampled densely enough, and their alters can be uniquely identified. In the majority of research using the egocentric network approach, however, there is an inability to connect the alters of different egos.

Macroscopic networks (sometimes referred to as sociometric networks) on the other hand, comprise both actors and the ties (or relationships) between each actor. Macroscopic networks can be further classified as either *partial* or *whole*, with partial networks being defined as a less than full account of ties among actors sampled. Partial networks include those where contact tracing has been used, either in its traditional public
health form, or more modern techniques, such as snowball or respondent-driven sampling. Whole networks (sometimes referred to as ideal networks) in contrast, contain a census of all actors and their connections in a bounded population. Information on whole networks is rare, especially in research dealing with STBBIs, due to the challenges in collecting these types of data. Therefore, empirical research on networks and STBBIs has mostly used egocentric, or partial network data. Although the use of egocentric and partial networks in research has provided invaluable insight into the dynamics of STBBIs within networks, and the influence of networks on transmission, the limitations of their use should be noted here. These include sampling biases due to the uncertainty of the comprehensiveness of the sampled network, recall biases of participants, as well as algorithms being more prone to collecting information on larger connected components, versus detecting smaller, disconnected networks.

2.4. BASIC SNA METRICS

The following section will discuss some basic measures used for the descriptive examination of networks. These measures are used to capture the concept of network topology; if indeed the underlying structure of the network does influence the epidemiology of diseases, then it is imperative to have standard ways to describe the shapes and contours of the network. As Doherty et al. note:

“Overall, these sociometric assessments characterize the network. These measures describe overall network connectedness within the population, microstructural details of network configuration, and individual position and prominence within the network and, when considered together, can provide insights into the potential of STI transmission.”

\textsuperscript{45(p. 545)}
Measure of network topology can be classified into two broad categories: *connectivity* and *centrality* (including mean partnership counts,\textsuperscript{52} core-periphery structure,\textsuperscript{53} and clustering),\textsuperscript{54} and are primarily descriptive, capturing regularities in the network structure including. The formulas used to derive each network metric are included in Appendix 1.

### 2.1.1 Connectivity

Connectivity refers to how actors in one part of the network are connected to actors in another part of the network. There are three major measures of connectivity: *reachability, distance* and *number of paths*. Reachability is simply whether or not node *i* can reach *j*; nodes *i* and *j* are said to be reachable if a direct link between the two nodes, or if there is a direct chain of contact (through other nodes) between them. The links between nodes *i* and *j* can be referred to as either *paths, walks* or *cycles*, with both paths and walks defined as a sequence of nodes and edges that connect nodes *i* and *j*. Paths are distinguished by the limitation that they cannot count the same node twice, and thus must always be moving “forward”. Walks, on the other hand, can include sequences where nodes are revisited. Cycles are paths that start and end with the same node. If there is at least one path connecting every actor in a network, the network is deemed *connected*, and all connected elements are called a *component*.

Distance refers to the number of steps along each sequence that connects nodes *i* and *j*. Finally, number of paths refers to the number of paths connecting any set of dyads in the network, with the existence of multiple paths between dyads thought to increase the efficiency of transmission between dyads.
2.1.2 Centrality

Centrality is conceptually simple, as it purports to measure which nodes are in the “centre” of the network, and is thus a measure of relative importance.\textsuperscript{55} In practice however, the concept of centrality has proven to be challenging to define, with a host of metrics proffered for this measure. A major contributor to the field of network metrics was Linton Freeman, who was one of the first researchers to “impose order” on the myriad measures of centrality, by categorizing centrality into three basic types: degree, closeness and betweenness.\textsuperscript{55,56} These three measures, along with Bonacich’s eigenvector-based measure (“Eigenvector centrality”),\textsuperscript{57} are the predominant measures used in the SNA literature to measure centrality.\textsuperscript{55} It is important to note that all measures can be used as a local measure of a specific node, or a global measure of the network itself. Additionally, it should be noted that the local form of each measure can be standardized for comparability across different networks.

Degree: Degree is the simplest measure and perhaps the most intuitive measure of centrality, as it is simply a summary of the number of ties connecting one node to other nodes in the network. In directed networks, degree can be further broken down into in-degree (the number of ties “pointing” to a node) and out-degree (the number of ties “pointing out” from the same node). For undirected networks, there is only one measure of degree. Intuitively, nodes which have the most number of ties are thought to hold a place of prominence in a network. In directed networks, this influence is somewhat more nuanced: a node with high in-degree is thought to be prominent, with many nodes directing ties towards the node with high in-degree, while a node with high out-degree is thought to be influential.
In addition to the actual crude or normalized value of degree, the degree distribution within a network (or within components of a network) provides an extremely useful way to characterize networks. Highly skewed degree distributions (such as those seen when degree distribution follows power laws, i.e., scale-free networks) suggest the existence of a high degree of heterogeneity, deviating from the assumption that degree distribution would be random, following binomial or Poisson forms; ultimately, assessing heterogeneity, and pinpointing the most influential nodes has practical importance in disease prevention and control. As well, scale-free networks tend to have highly-connected hubs, which ultimately result in shorter distances overall between nodes, impacting the efficiency of these types of networks in transmitting infections.

**Closeness**: Nodes are thought to be important if they are relatively close to all other actors, and thus *closeness* measures, on average, how close a node is from all other nodes in the network. Closeness is the inverse of distance (or “farness”) to every other actor, and is calculated by summing all distances a node is from other nodes, and then taking the inverse of the distance. “Point” closeness is the raw sum, while normalized is calculated by dividing by N-1 to get an average.

**Betweenness**: A node is thought to have high *betweenness* if multiple shortest paths pass through that node, that is, the number of times one node is on the shortest path between two other nodes. Thus, a node with high betweenness can potentially control the flow of information to the dyads it lies between. Betweenness is calculated by counting the number of shortest paths between *i* and *k* that actor *j* resides on.
2.1.3 Microstructures

The above section on network metrics can help to describe and characterize the properties of networks, both at the local level, and globally. Another property of networks, clustering, begins to ascertain the degree of cohesiveness present, either within the total network, or within the sub-networks that comprise the total network. This next section will introduce the concept of microstructures, which are smaller node groupings thought to potentially facilitate STBBI propagation, as well as define some basic terms related to microstructures. Generally speaking, network microstructures are a smaller collection of nodes (sometimes referred to as a subgraph) that can be considered a “unit”, and include the number of components, size of the largest component(s), cycles, and the number of dyads, triads and other microstructures. The dyadic and triadic censuses, that is, the counts of various types of dyads and triads, respectively, have proven in particular, to provide much insight into the potential for disease potentiation.

There are many different ways of defining a microstructure, but all start with some measure of connectedness among members. For example, a clique is a unit where every member of the subgraph is connected to every other member of the subgraph, which is a very restrictive definition, and is rarely observed empirically, except within specific settings (e.g., a family). Other examples, which are meant to loosen the restrictions used to define a clique, while at the same time retain a standardized method of defining a unit, include $k$-cores, $k$-plexes, and $n$-cliques. A unit is a $k$-core if every member of that unit has ties to at least $k$ other nodes within the unit. A node is a member of a $k$-plex if they have ties to all but $k$ other nodes, conditional on a minimum number of nodes to include in a unit. $N$-cliques capture the familiar adage, “a friend of a friend”; all
members of an \textit{n-clique} are reachable within \textit{n} steps. K-cores, k-plexes and other microstructures have been shown to be associated with STBBI prevalence; for example, being a member of a 2-core was a statistically significant risk factor for HIV infection in Bushwick, New York, while other individual-level factors, such as number of other people injecting drugs with, and race/ethnicity both lost statistical significance in multivariable models.\textsuperscript{62}

2.5. THE IMPACT OF NETWORK STRUCTURE ON TRANSMISSION

The pervasiveness of STBBIs globally poses a challenging paradox: how can transmission of pathogens reliant on exceedingly intimate forms of human contact be so widespread and sustained in a vast array of geographical and temporal spaces?\textsuperscript{63} Unlike respiratory pathogens such as influenza, which are typically randomly transmitted through highly connected networks,\textsuperscript{64-67} STBBI transmission occurs along relatively sparse and disconnected sexual contact networks.\textsuperscript{68-70} At the same time, successful transmission depends on the micro-level conditions for STBBI transmission (e.g., an infectious partner, a susceptible partner and no method of protection) aligning perfectly with this relatively sparse and disconnected network.\textsuperscript{45} In fact, one of the more important findings from early use of SNA in the HIV epidemic was the counterintuitive notion that the successful transmission of HIV in earlier phases of the global epidemic was not reliant on dense networks of sexual partners, but rather, on “core-transmitters” with high numbers of partners.\textsuperscript{39} Therefore, understanding the characteristics of networks under which successful propagation has taken place is of vital importance. Moreover, the CDC has recognized the importance of social network strategies for HIV testing, prevention and engagement in care are important.\textsuperscript{71,72}
Broadly speaking, the network characteristics influencing STBBI transmission can be categorized as either *structural* or *compositional*.\(^7\) The focus of this section of the literature review will predominantly be on the influence of network structure (i.e., the physical attributes of a network), although it may be useful to note that compositional factors tend to include things like relationship type and duration, selective mixing, and nodal characteristics, such as age structure.\(^6,7\) As well, important partnering patterns, such as concurrency and serial monogamy also fall under compositional characteristics. The concept of concurrency, as first proposed by Watts and May,\(^7\) and expounded by researchers such as Kretzschmar and Morris, is an especially powerful concept influencing the size, speed and persistence of STBBI epidemics in sparse networks,\(^15,30,79-81\) although the practical importance of concurrency in empirical settings is still highly debated.\(^8\)

One of the first clues suggesting how network structure might “facilitate or obstruct” STBBI transmission was discussed by Rothenberg et al.,\(^8\) in their study of high-risk populations in Colorado Springs.\(^28,83\) The authors noted that in contrast to injection drug users from Bushwick, New York, where a high prevalence of HIV was observed and whose members were part of a network with high centrality, the relatively low levels of HIV among their participants could be explained by HIV-positive individuals in Colorado Springs occupying positions of low centrality.\(^83,84\) Specifically, the combined dendritic and radial nature of networks in Colorado Springs were thought to result in less efficient transmission.\(^28\) Moreover, in a follow-up longitudinal study, the authors were able to demonstrate that low levels of “endogenous” transmission among their participants coincided with an increased segmentation of networks over time, as
well as an increasing deficiency of the types of network microstructures (e.g., those
displaying high connectivity) which would normally heighten the potential for STBBI propagation.\textsuperscript{29,84} Evidence for the importance of microstructures, such as cliques and
cycles has been demonstrated in the re-emergence of syphilis in the US,\textsuperscript{21} the spread of HIV among “low-risk” individuals in Malawi, Africa,\textsuperscript{85} and high prevalence of STBBI and risky injection practices among injection drug users.\textsuperscript{73,86}

In addition to microstructures, other network properties, such as density, higher
connectivity and degree centrality have been shown to be associated with STBBI
prevalence and incidence.\textsuperscript{45,76,87-89} For example, Ellen et al. observed that adolescents in
areas with high STBBI rates who had larger social networks were more likely than those
with smaller networks to acquire an infection from bacterial STBBI\textsuperscript{87}. This is consistent
with other work which has demonstrated that the majority of risk-potential linkages are
social in nature,\textsuperscript{90} as sexual ties often start off as social contacts.\textsuperscript{91} This observation is
conditional on age, as younger populations do not frequent bars and thus the social-
contact-to-sexual-partner pathway is even more pronounced.\textsuperscript{91} Additionally, more recent
digital technologies (such as “hook-up” apps like Grindr and Tinder) may be impacting
traditional ways of seeking sexual partners.\textsuperscript{92} Even in the latter case however, capturing
network information remains important, as researchers have found that sexual risk
behaviours (e.g., unprotected anal intercourse) are associated with how partners were
met.\textsuperscript{91}

On the subject of centrality, it should be noted that Ghani & Garnett, using
individual-based simulation studies have demonstrated an asymmetry in the influence of
this measure in STBBI acquisition and transmission; local centrality measures were
important for acquisition while global measures of centrality were important for transmission. Using population-based public health contact tracing data from case and contact investigations of bacterial STBBI, Wylie et al. found consistent structural forms (e.g., radial and linear types of components) that were associated with specific epidemiological properties. Radial-type components spanned fewer geographic locations than linear-type, chlamydia was found in radial-type components, whereas both chlamydia and gonorrhea were harboured in linear-type components, with linear-type components displaying more assortative mixing, in terms of numbers of partners, while radial-type networks were more dissortative. These findings have been replicated in simulation studies, as well as other empirical research.

2.6. PERSISTENCE OF FORM AND UNDERLYING GENERATIVE PROCESSES

Notably, the microstructure of networks has demonstrated remarkable resilience and stability. Wylie et al., using two different macroscopic networks created through STBBI contact tracing separated by 5 years in time, demonstrated that not only were the types of components similar across both time periods (along with their associated epidemiological properties), but also the distribution of the size of components. The researchers hypothesized that despite a change in the population at the individual level, normative patterns underlying the microstructures persist in the broader community, reproducing observed forms over time. This perspective of uncovering the underlying generative processes which ultimately result in observable, stable and persistent structural forms has recently been echoed by several researchers. That is, individuals may come and go, and their links form and dissolve, but the overall structural patterns in networks
conducive to STBBI propagation remain stable at the population level. Thus, some worthy research aims are to model the micro-to-macro link and to delineate the feedback between local rules and global structures: that is, how local rules lead to global structures, and how global outcomes then select for local rules.

Being able to predict how the contact structure emerges from simple behavioural rules, as well as testing hypotheses about the underlying generative processes of link formation and dissolution can have the potential to short-circuit methodological limitations associated with SNA, especially as it relates to data collection and inference from a sample. As Morris writes:

“If simple rules govern partner selection, then these also determine the aggregate structure in the network; what looks like an unfathomably complicated system, is, in fact, produced by a few key local organizing principles. By extension, these simple local rules are also, therefore, the key behavioural determinants of disease transmission dynamics on the network…” 6(p. 122)

2.7. ANALYTICAL FRAMEWORK

Taken together, the “generative” perspective of network emergence (i.e., behaviours at the individual level create characteristic microstructures which then are components of observable networks) and these three findings suggest a framework to explain the stability of network microstructures over time, as well as the ability of STBBIs to remain persistent within populations, despite turnover at the individual level. More explicitly, the framework proposes that individual behaviours (e.g., selective mixing, degree
distribution, etc.) create characteristic microstructures in contact networks, with both the actual substructures and their interactions having predictable influences on epidemic trajectory. Some of these micro-level processes are influenced by broader level, or macro, determinants (such as socio-economic policies or oppression/stigma), while others are influenced by individual choice, or by the composition of the network an individual happens to find themselves in.  

Because in any given population, microstructures and their interactions are multitudinous, relative to actual outcomes (i.e., the presence/absence of an epidemic), observed epidemics trace out just one instance of a universe of potential instances. Danon et al. propose a similar perspective when they write that any observed network is one “realization” of an underlying epidemic process.\textsuperscript{48} Further, they also write that only focusing on one instantiation (i.e., what is observed), without exploring other potential scenarios (i.e., the scenarios that were not observed, but had potential to happen) leaves little in the way of predictive power, impacting the ability of researchers and policymakers to plan interventions more generally.\textsuperscript{48} This framework will be used to guide the remaining chapters in the dissertation.

2.8. EXPONENTIAL RANDOM GRAPH MODELS

In order to achieve research objectives, the exponential random graph modeling (ERGM) paradigm will be used for all analyses.\textsuperscript{99,100} Briefly, the ERGM framework was chosen because it is a tie-based statistical modeling framework, comparable to logistic regression, which seeks to use micro-level information to create macro-level patterns. This framework models the probability that a tie exists (in logit form) as a linear function of a set of predictors. ERGMs can trace their origins to spatial statistics and network
At its core, ERGMs are a tool to assess whether an observed network can result from the nodal characteristics and the patterns of ties composing the network. As Goodreau et al. write, because there is “rarely a neat pattern of process and pattern…” statistical methods are valuable tools that can be used to tease apart the “micro-level foundations of structure.” The goal of modeling with ERGMs is to build a model that sufficiently reproduces features of the observed network, with the parameters of the model (i.e., the independent variables) being features of the network that are thought to be important for that network. Features may include both dyad independent terms, such as attributes of nodes and links, and dyad dependent terms, such as the distribution of microstructures, like triads. Advances in computationally intensive statistics provide practical methods [Markov Chain Monte Carlo, or (MCMC)] for solving problems related to estimation of dyad-dependent terms.

In ERGMs, $S_k$ is the vector of features (i.e., the independent variables) thought to be important to the formation of the observed network, with $k$ equal to the number of features in the model. After estimation by MCMC methods, the vector of parameter estimates, $\Theta_k$, reflect the change in log-odds for ties to actually exist by $\Theta_k$. When $\Theta_k$ is statistically significant, the interpretation is that there is a propensity in the network for the feature defined by $S_k$ to be important in generating networks with features similar to the observed network. An additional benefit of the use of MCMC methods is the ability for the ERGM framework to simulate networks, centred around parameter estimates. Moreover, disease processes, such as transmission along these simulated networks can also be simulated. By examining the structure of a set of possible networks in an aggregate manner, insights into successful propagation of STBBIs may be found.
Furthermore, intervention effects can also be investigated in simulations, which allows for the application of a counterfactual framework in order to guide policy decisions; these “in-silico” experiments can help inform the impact and progression of disease outcomes, conditional on intervention efforts.

A comprehensive application of the capabilities of ERGMs in teasing out network structure is given in Goodreau et al. Using data from the National Longitudinal Study of Adolescent Health, their goal was to “…identify the determinants of friendship formation that lead to pervasive regularities in friendship structure among adolescent students.” They considered the individuals’ overall propensity to make friends (sociality), propensity to make friends based on attributes (selective mixing), and based on their friends’ choice of friends (triad closure). These concepts encompass various levels of generative processes: individual, dyadic and triadic. They found that selective mixing and triadic closure were both at play in friendship networks, with the strength of both processes dependent on grade, ethnicity, and size and ethnic composition of the school setting.

2.9. EXPONENTIAL RANDOM GRAPH MODELS IN THE LITERATURE

The literature on the use of ERGMs for modeling of STBBIs is relatively sparse, with almost all articles related to viral STBBIs, specifically HIV and hepatitis C (HCV). No articles on the use of ERGMs to model bacterial STBBIs were found. What follows is a brief review of recent published articles using ERGMs to analyse network models of viral STBBIs transmission, illustrating the flexibility of the ERGM framework to examine substantive questions regarding the impact of networks on prevention and interventions designed to decrease the incidence viral STBBIs.
Goodreau et al. used empirical data to estimate networks and then simulated HIV transmission across these networks in an effort to examine the impact of concurrent partnerships and acute-, chronic- and late-stage contacts on the transmission of HIV in Zimbabwe.\textsuperscript{81} They found that the joint impact of concurrency and acute-stage infections helped to sustain HIV in young adults in Zimbabwe, with neither alone being sufficient to sustain the epidemic. Interaction between concurrency and acute stage infections long been suspected,\textsuperscript{6,30,111,112} as concurrency sets the stage for acute-stage infections to get transmitted efficiently (especially in populations that do not report high numbers of partners), by ensuring a partner is available for transmission during the acute stage of infection.

A study conducted with data from Peru examined male circumcision among Peruvian MSM.\textsuperscript{113} This study found that despite high HIV prevalence and willingness to undergo the procedure, male circumcision would generally be an ineffective tool in Peru. The study estimated that 5-10\% of incident cases would be averted over 25 years, even at 50\% coverage of the population,\textsuperscript{113} similar to other studies looking at male circumcision in high-resource settings.\textsuperscript{114}

ERGMs were used by Goodreau et al. to understand the determinants driving HIV incidence among men who have sex with men (MSM) in the United States (US) and Peru.\textsuperscript{115} The study found that the proportion of incidence due to contacts who were in the acute stages of infection ranged from 5\% to 29\%, much smaller than what has been reported in the literature previously. Additionally, the models illustrated the substantial contribution made by MSM who were undiagnosed, diagnosed but untreated, and
inadequately treated, highlighting the need for comprehensive prevention packages that include linkage and retention of men into care programs, once they are screened for HIV.

In a follow-up study, and using ERGMs to simulate networks and transmission, Carnegie et al. investigated optimal strategies to deploy pre-exposure prophylaxis (PrEP) as a means to mitigate HIV incidence among MSM resident in the US and Peru.\textsuperscript{116} The authors noted that although compartmental mathematical models have been used to examine optimal strategies for PrEP,\textsuperscript{117-122} the population-level effectiveness of PrEP is complicated by factors such as targeting, uptake, adherence and risk compensation.\textsuperscript{116,123-126} Modeling this complexity is a limitation of compartmental mathematical models, as the number of compartments needed to capture this complexity rises exponentially with the number of variables considered.\textsuperscript{116,127} Consistent with studies that found targeting of high-risk MSM improved efficiency of PrEP, Carnegie et al. found that targeting PrEP to the men with the highest rates of condomless anal sex with their casual partners could lead to 30\% reduction in incidence over the span of 10 years, with benefits observed outside of MSM receiving PrEP. Moreover, the models found that targeting serodiscordant couples only was an ineffective strategy at the population level. Thus, the use of ERGMs suggested some optimal targeting strategies in the deployment of PrEP for MSM in the US and Peru.

In two separate studies, Khanna et al. utilized the ERGM framework to understand HIV transmission dynamics among MSM in southern California.\textsuperscript{128,129} Both studies were parameterized by data from behavioural surveys with parameters pertaining to demographic, biological, treatment and behaviour-related processes. In the first study, Khanna et al. used a special type of ERGM, separable temporal ERGMs (or STERGMs)
to model networks dynamically over time. They investigated the impact of post-diagnosis behaviour change (PDBC) on HIV prevalence. The timing, extent, durability, heterogeneity of PDBC is not well known, and neither is the impact of variations of these factors on prevalence. The authors found that models without PDBC had HIV prevalences that were about 30% higher than those that included PDBC. Even minimally implemented PDBC, for example, screening for HIV twice every 10 years, resulted in substantial declines in HIV prevalence. The authors concluded that PDBC may be an effective means to reduce HIV among MSM. In their second study, Khanna et al. examined how individualized diagnosis interventions can aid in reducing HIV incidence, with the most effective intervention tailored to an individual’s behaviours.

Fujimoto et al. investigated 2-mode networks using ERGMs to understand the role of multiplex-nominated nodes within male sex worker networks, and how ties are mediated by venues. The framework they used was Granovetter’s theory of weak ties, whereby “weak ties” act as important bridges between clusters that may not otherwise have been connected. The authors found that weak ties were more likely to be associated with venues where MSW gathered, suggesting venues as a crucial mechanism for the diffusion of messages. Moreover, the authors stressed the utility of multiplex-nominated peers with respect to targeting, and to identify peers through multiple characteristics. In their words, “(d)eveloping more sophisticated models of peers could enhance the utility of a network-level HIV intervention by resulting in significant breadth when targeting…commercial sexual and drug networks.”

Based on previous modeling work, Hellard et al. used ERGMs to model the effectiveness of anti-viral medication to treat HCV among people who inject drugs.
(PWID), taking network structure into account. The authors found that a “Treat-your-friends” strategy outperformed other strategies, resulting in a 6.5% absolute decrease in HCV incidence when compared to randomly selecting nodes from the PWID network. This more targeted approach in focusing on PWID with HCV and their social and injection networks is consistent with other studies.

Finally, Dombrowski et al. re-analyzed two seminal network datasets (the SFHR and Project 90) using ERGMs to determine whether any further insight could be gained by using a statistical framework. The authors substantiated earlier work that demonstrated the importance of race/ethnic homophily in the formation of networks, especially among the youngest members of both networks. In their analyses, Dombrowski et al. found that Hispanics showed the strongest in-group preference in both datasets. Additionally, the authors noted the central role played by transitive closure, especially in the SFHR network, which was a network of PWID. Here, transitive closure, as a determinant of tie formation, was 10-20 times higher than other determinants. The authors surmised that the legal environment at the time, which eliminated public injection spaces, forced PWID to inject in more discreet, “hidden” locations, with the result being the increased likelihood of injecting not only with one injection partner, but that partner’s partner, due to enforced close clustering of PWID.
Chapter 3 provides the foundation for ensuing chapters by unpacking the recent epidemiology of sexually transmitted and bloodborne infections in Winnipeg, Canada. Traditional public health surveillance tools, such as rate calculations and regression models explore the epidemiology of five distinct epidemics, and orient them towards person, time and place. The concept of context is introduced here, and takes on the specific form of place-based analyses. Gini coefficients and population attributable fractions explore the distribution, and concentration of infections over time and space. The ecological niche of each pathogen is defined through the combination of traditional surveillance tools and place-based analyses.

3.1. INTRODUCTION

Sexually transmitted and bloodborne infections (STBBIs), such as chlamydia (CT), gonorrhea (NG), syphilis and the human immunodeficiency virus (HIV) infections represent a tremendous public health burden. In the United States (US) alone, the direct medical cost attributed to STBBIs has been estimated to be $16 billion USD annually. In response to increases in HIV and CT incidence in the 1990s, significant health and human resources, as well as scientific endeavour was dedicated towards prevention and control of STBBIs, and understanding their epidemiology. At the same time, in industrialized nations in the last decade of the 20th century, incidence rates for both NG and infectious syphilis were at a historical nadir, and it was not uncommon for local elimination of NG and infectious syphilis to be promoted as achievable goals for STBBI control programs.

However, the first two decades of the 21st century have demonstrated how inherently intractable STBBIs are, against their prevention and control. Despite the promise of rapid and non-invasive diagnostics in the case of CT and NG, as well as inexpensive and for the most part, effective treatment, rates of CT and NG have remained high in North America and other jurisdictions, with increases in NG infections being especially notable. In 2015, the rate for CT in the US was 479/100,000, while the rate for NG was 124/100,000, representing a 6% and 13% increase in rates for CT and
NG, respectively, since 2014. Additionally, the last decade has seen tremendous increases in the incidence of infectious syphilis, primarily among men who have sex with men (MSM), but also more recently emergent in other sub-populations. For example, after reaching a historically low rate of 2/100,000 population in 2000-2001, the infectious syphilis rate has more than tripled to 7/100,000 in 2015 in the US. In Winnipeg, Canada, Shaw et al. reported a rate of 9/100,000 in 2014-2015, also a threefold increase in rates from 2006-2013, with an especially steep increase among women in 2014-2015. Finally, prevention and control of HIV remains a challenge, despite the availability of effective treatment and relatively accessible testing. In Canada, there were an estimated 2570 new infections in 2014, for an estimated HIV incidence rate of 7.2/100,000; in comparison, HIV incidence in the US was reported to be 12/100,000.

Core group theory specifies that a small segment of the population disproportionately bears the burden of infection; given limited resources, program planning to prevent and control STBBIs should be informed by data on the distribution of STBBIs, including groups or populations who are at excess risk of infection. To this end, examination of the distribution of STBBIs by geography has contributed significant insight into theoretical and practical work. Place-based analyses are also useful from the public health perspective, as geography can serve as a proxy for a variety of contextual and socio-economic processes impacting a population’s health status, and that are difficult to disentangle and quantify empirically. Additionally, within Canada, funding and delivery of health services are allocated through the use of administrative areas defined by geographic boundaries; ultimately,
public health units are responsible for the delivery of services in their geographic area. Thus, understanding the epidemiology of STBBIs through a place-based lens lends itself naturally to planning, implementation and evaluation of services and interventions.

Previous analyses in Winnipeg have shown a higher level of geographic clustering by NG, compared to CT, most particularly within “inner-core” areas. For example, using population attributable fractions (PAF) Shaw et al. showed that 46% of incident NG infections were attributable to residents living in the inner-core areas of Winnipeg, compared to 24% of CT infections. Using Wasserheit and Aral’s phase-specific framework of STBBI epidemics as a backdrop, Elliott et al. demonstrated how the use of the Lorenz curve and Gini coefficient could inform understanding of the evolution of STBBI epidemics as they progressed through various phases, and how this understanding might inform public health practice. In a period of decreasing rates of both CT and NG, the authors showed that geographic concentration of CT was increasing, suggesting that CT was undergoing an early decline phase. Their recommendation was that further reductions in CT rates might be achieved through intensifying efforts to engage high-risk populations residing in high-incidence areas who may have not been adequately reached through other mechanisms.

Given the recent increases in NG infections, a natural question would be whether or not this increase is associated with an expansion in the geographic distribution of NG, as predicted by the phase-specific framework. Furthermore, few studies have included the examination of the relationship between epidemic phase and geographical distribution of other pathogens such as HIV and syphilis. Therefore, the purpose of this study was to
examine the long-term epidemiology of four STBBIs using traditional descriptive and inferential analytical techniques alongside techniques assessing geographical concentration of incident cases.

3.2. METHODS

Population Setting
The Winnipeg Health Region (WHR) is the largest health region in the province of Manitoba, a central Canadian province. In 2010, the population of the WHR was 700,000, or almost 70% of the population of Manitoba. The WHR is divided into 25 Neighbourhood Clusters (NCs), administrative areas used for planning purposes. Four NCs (Point Douglas North and South, and Downtown East and West) compose the WHR’s “inner-core”; historically, these NCs have had the highest rates of STBBIs, and are characterized by relatively lower socio-economic status and a higher burden of poorer health outcomes.

Data Sources

Manitoba Health STBBI Database
All laboratory-confirmed CT, NG and HIV infections are reported to the provincial ministry of health [Manitoba Health, Seniors and Active Living (MHSAL)], and records of all cases are entered into the Manitoba Health STBBI Database. Results are then referred to Regional Health Authorities based on the client’s region of residence (as defined by their postal code at time of testing) and assigned to a public health nurse (PHN) for case management. PHNs within the Winnipeg Regional Health Authority’s (WRHA) public health team ensure a case surveillance form is completed on the client.

WRHA Syphilis Database
In Manitoba, all positive syphilis results are also reportable to the provincial ministry of health, and PHNs for the WRHA are responsible for investigation of clients positive for infectious syphilis. Key variables include sociodemographic (e.g., date of birth, gender), clinical (e.g., symptoms, staging) and behavioural and social risk factor information. Data from these forms are maintained in an electronic database at the WRHA.

The Manitoba Health STBBI database was used for CT, NG and HIV data as a computer-based surveillance system is not maintained at the WRHA for these infections. A computer-based surveillance system for infectious syphilis was established at the WRHA in 2003 in response to an outbreak of infectious syphilis. Prior to December 2015, only minimal data on infectious syphilis was reported back to MHSAL, and the data kept in the WRHA infectious syphilis database was thought to more accurate and timely. Since December 2015 (when MHSAL established an enhanced infectious syphilis database in response to a province-wide outbreak of syphilis), data that reside in WRHA and MHSAL databases are now synchronized. For the purposes of this study, information from the WRHA electronic database was used as a matter of convenience.

**Definitions**

Data on CT, NG, syphilis and HIV infections between January 1st, 2007 to December 31st, 2016 for clients residing in the WHR at the time of testing were extracted from their respective databases. Specimen collection date was used to define the year of infection; as well as to calculate age at each infection. Infections were categorized as “CT-only”, “NG-only”, “CT/NG concomitant”, “HIV” and infectious syphilis infections. For the purposes of this analysis extra-genital CT/NG infections were excluded. Also for CT/NG infections, any positive test occurring within 7 days of a previous positive test, for a
particular infection, was treated as the same infection; subsequent infections within 7 days were excluded from further analyses. Infectious syphilis was defined as all cases of primary, secondary, early latent and incubating syphilis, as per the protocol of MHSAL. Staging of infectious syphilis was based on information from the testing physician, and reviewed by two WRHA Medical Officers of Health. Population data were derived from the provincial insurance registry and provided by MHSAL. Only infections in those 9 years or older were included in these analyses.

For all infections, rates were calculated based on records with a valid postal code. The crude rate of each infection was calculated using the appropriate mid-point population as the denominator; population data were supplied by MHSAL. CT, NG and syphilis analyses are based on number of infections, and not individuals. For analyses focusing on geographical distribution, and given the variability of the epidemiology of STBBIs over time in our region, three four-year time periods were examined (2007-2009, 2010-2012, 2013-2015); additionally, the year 2016 was included in order to study the most current year of data. The geography of the WHR was defined as “Inner-Core”, which included the four Point Douglas and Downtown NCs, and “Non-Core”, which was defined as the rest of the WHR.

Analyses

Descriptive and Inferential Statistics

For basic descriptives analysis, rates were age- and sex-standardized to the Canadian population from the 2006 Canadian Census; 95% confidence intervals (95%CI) were generated using the Tiwari et al. method. For each infection, the following determinants were entered into regression models: age-group, year (2007-2009, 2010-
2012, 2013-2015 and 2016), sex and residence in the inner core. *A priori*, the sex X inner core interaction was included in models as amongst these variables, significant heterogeneity has been shown in previous analyses. Poisson regression models, with the logarithm of the population entered as an offset, were used to produce unadjusted and adjusted incidence rate ratios (UIRR/AIRR) and 95% CIs; only AIRRs are reported here.

**Geographical Distribution/Attribution**

Two measures of geographic distribution/attributions were used. First, the Gini coefficient was used to measure inequality in the geographic distribution of each bacterial infection; a Gini coefficient of 0 indicates perfect equality, in terms of geographic distribution (i.e., the infection is distributed equally among all areal units), while a coefficient of 1 indicates perfect inequality (i.e., all infections are contained in one areal unit). For these analyses, the areal unit of analysis was at the NC level (i.e., n=25). For each infection and for each time period, a Poisson regression model was used to estimate non-zero rates for each neighbourhood, adjusted for age and sex. These non-zero rates were then used in the calculation of the Gini coefficient. 95% CI were estimated using 1000 bootstrapped replications.

Second, the population attributable fraction (PAF) was used to quantify the geographical concentration of each infection. The PAF is epidemiologically interpreted as the fraction or proportion of cases in the population that can be attributed to given exposures, with the “exposure” in this case defined as residing in the inner core and is expressed here as an age- and sex-adjusted estimate. For the purposes of this study, PAFs are based on logistic regression models after the work of Greenland and Drescher.

**NG Outbreak Comparisons**
During the 10-year time period of study, two NG outbreaks were observed (2012 and 2016). In both years, the 95% CI of the age-standardized rate did not overlap the previous year’s 95% CIs (Table 3.1). For example, in 2012 the age-standardized rate for NG was 71/100,000 (95%CI: 64.3-77.3), while the age-standardized rate in 2011 was 52/100,000 (95%CI: 47.0-58.4), representing a statistically significant increase in NG rate. The distribution of age-adjusted NG rates, by NC, was examined for both outbreaks, and compared to the preceding three years of “endemic” rates (i.e., 2009-2011 and 2013-2015). For each NC, relative rates were calculated between outbreak years and the preceding endemic rate (i.e., 2012 vs. 2009-2011 and 2016 vs. 2013-2015). There were no statistically significant differences in age-standardized rates in the three-year period prior to each outbreak year, as assessed by 95%CI. Finally, Spearman’s rank correlation was used to compare the NC “profile” of NG outbreak years to the profile of NG and CT distribution during non-outbreak years.

Mapping

Dot maps for each infection were produced for each time period, based on the frequency of each infection over the entire study period. To protect the anonymity of cases, while maintaining the spatial distribution of infections, all cases were randomly allocated to one of 100 different locations within their neighbourhood (n=232) of residence. Additionally, the locations that cases were assigned to were further randomly “wobbled” up to 500 metres along the X and Y axes of each location. All analyses were conducted using Stata version 13 (College Station, Texas, USA). Gini coefficients and their 95%CIs were calculated using the ineqerr package (details in Appendix 2), while adjusted PAFs with 95% CI were estimated using the punaf package. The details of
the derivation for PAF are highly technical; Newson provides a comprehensive
explanation how PAF and 95% CIs are calculated with the punaf program.\textsuperscript{177} Dot maps
were produced using the spmap package, based on the work of Pisati.\textsuperscript{180,181} All packages
were downloaded from the Boston College Statistical Software Components (SCC)
archive. Informed consent was not obtained as this was a secondary analysis of
anonymized surveillance data.

3.3. RESULTS

\textit{Descriptive Analysis}

Over the 10-year time span of the study, a total of 37,521 infections were reported to
public health. Over this time period, CT was the most common infection (Table 3.1), with
an age-standardized rate of 457/100,000, followed by NG (64/100,000), CT/NG
(37/100,000), HIV (12/100,000) and syphilis (6/100,000). Figure 3.1 illustrates the
dynamic epidemiology of all infections; data are presented on a log-scale to more
explicitly illustrate secular trends in NG, CT/NG, syphilis and HIV. Most notably, there
has been a substantial increase in the incidence of reported syphilis infections starting in
2010; from 2010 to 2016, the age-standardized rate of syphilis increased, in an almost
monotonic manner, from 1/100,000 to 14/100,000. Also notable is the trebling of
reported NG infections from 2015 to 2016, increasing from 50/100,000 to 144/100,000,
respectively. Reflecting this increase in NG infections, CT/NG concomitant infections
more than doubled from 33/100,000 to 72/100,000 during the same time period. Over the
10-year time period, reported CT rates have ranged between 440/100,000 to 525/100,000
between 2008 and 2012. After decreasing to 407/100,000 in 2013, reported CT infections
have increased every year since 2014. Reported HIV rates did not show a clear trend over
the 10-year time period, with periods of increasing rates (2007-2010), followed by decreases (2011-2012); similar to CT, however, HIV rates have steadily increased between 2014 and 2016, from 9/100,000 to 15/100,000.
Figure 3.1: Incidence of Select Sexually Transmitted and Bloodborne Infections, Winnipeg Health Region 2007-2016 (logarithmic scale)
Table 3.1: Frequency, Crude and Age-Standardized Rates (per 100,000), Sexually Transmitted and Bloodborne Infections, Winnipeg Health Region (2007-2016)*

<table>
<thead>
<tr>
<th>Year</th>
<th>CT Rate</th>
<th>95% CI</th>
<th>NG Rate</th>
<th>95% CI</th>
<th>CT/NG Rate</th>
<th>95% CI</th>
<th>HIV Rate</th>
<th>95% CI</th>
<th>Syphilis Rate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>394.9</td>
<td>379.3 - 411.0</td>
<td>78.0</td>
<td>71.1 - 85.3</td>
<td>35.2</td>
<td>30.7 - 40.3</td>
<td>9.9</td>
<td>7.5 - 12.7</td>
<td>3.9</td>
<td>2.6 - 5.8</td>
</tr>
<tr>
<td>2008</td>
<td>525.0</td>
<td>507.1 - 543.4</td>
<td>58.4</td>
<td>52.5 - 64.8</td>
<td>40.0</td>
<td>35.2 - 45.4</td>
<td>12.3</td>
<td>9.7 - 15.4</td>
<td>1.3</td>
<td>0.6 - 2.5</td>
</tr>
<tr>
<td>2009</td>
<td>485.6</td>
<td>468.4 - 503.2</td>
<td>52.6</td>
<td>47.1 - 58.6</td>
<td>28.7</td>
<td>24.6 - 33.3</td>
<td>13.6</td>
<td>10.8 - 16.9</td>
<td>0.7</td>
<td>0.2 - 1.7</td>
</tr>
<tr>
<td>2010</td>
<td>487.3</td>
<td>470.2 - 504.8</td>
<td>42.9</td>
<td>38.0 - 48.4</td>
<td>29.9</td>
<td>25.8 - 34.5</td>
<td>15.5</td>
<td>12.5 - 18.9</td>
<td>1.3</td>
<td>0.6 - 2.5</td>
</tr>
<tr>
<td>2011</td>
<td>481.1</td>
<td>464.2 - 498.4</td>
<td>52.4</td>
<td>47.0 - 58.4</td>
<td>23.0</td>
<td>19.4 - 27.0</td>
<td>9.0</td>
<td>6.8 - 11.7</td>
<td>1.4</td>
<td>0.7 - 2.6</td>
</tr>
<tr>
<td>2012</td>
<td>443.3</td>
<td>427.3 - 459.7</td>
<td>70.6</td>
<td>64.3 - 77.3</td>
<td>45.4</td>
<td>40.4 - 50.9</td>
<td>8.1</td>
<td>6.1 - 10.7</td>
<td>2.3</td>
<td>1.4 - 3.7</td>
</tr>
<tr>
<td>2013</td>
<td>407.1</td>
<td>391.8 - 422.8</td>
<td>50.4</td>
<td>45.1 - 56.1</td>
<td>36.1</td>
<td>31.7 - 41.1</td>
<td>12.2</td>
<td>9.6 - 15.2</td>
<td>4.7</td>
<td>3.3 - 6.6</td>
</tr>
<tr>
<td>2014</td>
<td>426.9</td>
<td>411.5 - 442.9</td>
<td>43.1</td>
<td>38.2 - 48.4</td>
<td>28.2</td>
<td>24.3 - 32.6</td>
<td>9.2</td>
<td>7.0 - 11.9</td>
<td>12.1</td>
<td>9.7 - 14.9</td>
</tr>
<tr>
<td>2015</td>
<td>446.6</td>
<td>430.8 - 462.9</td>
<td>50.1</td>
<td>44.9 - 55.8</td>
<td>33.4</td>
<td>29.2 - 38.1</td>
<td>13.0</td>
<td>10.3 - 16.0</td>
<td>15.7</td>
<td>13.0 - 18.8</td>
</tr>
<tr>
<td>2016</td>
<td>476.2</td>
<td>459.9 - 493.0</td>
<td>143.5</td>
<td>134.5 - 152.8</td>
<td>71.8</td>
<td>65.5 - 78.5</td>
<td>15.4</td>
<td>12.5 - 18.7</td>
<td>14.3</td>
<td>11.7 - 17.3</td>
</tr>
<tr>
<td>Total</td>
<td>457.0</td>
<td>451.8 - 462.3</td>
<td>64.4</td>
<td>62.5 - 66.4</td>
<td>37.3</td>
<td>35.8 - 38.9</td>
<td>11.8</td>
<td>11.0 - 12.7</td>
<td>6.0</td>
<td>5.4 - 6.6</td>
</tr>
</tbody>
</table>

*CT: Chlamydia; NG: Gonorrhea; CT/NG: Chlamydia/Gonorrhea Co-infection; HIV: Human Immunodeficiency Virus
95% CI: 95% Confidence Intervals
Table 3.2: Results from Poisson Regression models, Sexually Transmitted and Bloodborne Infections, Winnipeg Health Region *

<table>
<thead>
<tr>
<th>Age group</th>
<th>CT Crude Rate</th>
<th>CT AIRR</th>
<th>CT 95% CI</th>
<th>NG Crude Rate</th>
<th>NG AIRR</th>
<th>NG 95% CI</th>
<th>CT/NG Crude Rate</th>
<th>CT/NG AIRR</th>
<th>CT/NG 95% CI</th>
<th>HIV Crude Rate</th>
<th>HIV AIRR</th>
<th>HIV 95% CI</th>
<th>Syphilis Crude Rate</th>
<th>Syphilis AIRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;20</td>
<td>960.4 Ref</td>
<td>96.8</td>
<td>96.8-128</td>
<td>189.6 Ref</td>
<td>19.3</td>
<td>19.3-198</td>
<td>110.0 Ref</td>
<td>11.0</td>
<td>11.0-11.0</td>
<td>14.9 Ref</td>
<td>1.45</td>
<td>1.29-1.61</td>
<td>2.83 Ref</td>
<td>2.83-3.07</td>
</tr>
<tr>
<td>20-29</td>
<td>1496.7</td>
<td>1.54</td>
<td>1.28-1.89</td>
<td>326.6 Ref</td>
<td>0.86</td>
<td>0.73-1.03</td>
<td>27.7 Ref</td>
<td>0.28</td>
<td>0.22-0.35</td>
<td>23.2 Ref</td>
<td>1.29</td>
<td>0.82-1.22</td>
<td>4.48 Ref</td>
<td>4.38-4.59</td>
</tr>
<tr>
<td>30-39</td>
<td>419.3</td>
<td>0.43</td>
<td>0.38-0.48</td>
<td>326.6 Ref</td>
<td>0.86</td>
<td>0.73-1.03</td>
<td>27.7 Ref</td>
<td>0.28</td>
<td>0.22-0.35</td>
<td>23.2 Ref</td>
<td>1.29</td>
<td>0.82-1.22</td>
<td>4.48 Ref</td>
<td>4.38-4.59</td>
</tr>
<tr>
<td>40-49</td>
<td>115.4</td>
<td>0.12</td>
<td>0.11-0.13</td>
<td>32.6 Ref</td>
<td>0.35</td>
<td>0.25-0.48</td>
<td>8.0 Ref</td>
<td>0.08</td>
<td>0.04-0.15</td>
<td>18.4 Ref</td>
<td>1.22</td>
<td>0.99-1.51</td>
<td>3.59 Ref</td>
<td>3.59-3.63</td>
</tr>
<tr>
<td>50-59</td>
<td>39.8</td>
<td>0.04</td>
<td>0.03-0.05</td>
<td>14.9 Ref</td>
<td>0.17</td>
<td>0.09-0.29</td>
<td>4.1 Ref</td>
<td>0.04</td>
<td>0.02-0.07</td>
<td>9.4 Ref</td>
<td>0.71</td>
<td>0.51-1.01</td>
<td>1.84 Ref</td>
<td>1.84-2.56</td>
</tr>
<tr>
<td>60+</td>
<td>6.4</td>
<td>0.007</td>
<td>0.005-0.008</td>
<td>3.8 Ref</td>
<td>0.04</td>
<td>0.01-0.07</td>
<td>0.5 Ref</td>
<td>0.01</td>
<td>0.003-0.12</td>
<td>3.2 Ref</td>
<td>0.88</td>
<td>0.63-1.24</td>
<td>0.63 Ref</td>
<td>0.63-1.24</td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007-2009</td>
<td>481.0 Ref</td>
<td>64.2</td>
<td>56.7-75.9</td>
<td>33.6 Ref</td>
<td>0.94</td>
<td>0.85-1.05</td>
<td>10.8 Ref</td>
<td>0.92</td>
<td>0.76-1.11</td>
<td>11.3 Ref</td>
<td>0.96</td>
<td>0.79-1.15</td>
<td>5.57 Ref</td>
<td>5.57-6.57</td>
</tr>
<tr>
<td>2010-2012</td>
<td>483.4</td>
<td>0.997</td>
<td>0.97-1.04</td>
<td>56.7 Ref</td>
<td>0.87</td>
<td>0.83-0.91</td>
<td>33.6 Ref</td>
<td>0.94</td>
<td>0.84-1.05</td>
<td>10.8 Ref</td>
<td>0.92</td>
<td>0.76-1.11</td>
<td>11.3 Ref</td>
<td>11.1-11.9</td>
</tr>
<tr>
<td>2013-2015</td>
<td>440.4</td>
<td>0.91</td>
<td>0.88-0.98</td>
<td>49.0 Ref</td>
<td>0.76</td>
<td>0.72-0.81</td>
<td>33.0 Ref</td>
<td>0.93</td>
<td>0.84-1.04</td>
<td>11.3 Ref</td>
<td>0.96</td>
<td>0.79-1.15</td>
<td>11.0 Ref</td>
<td>5.57-6.57</td>
</tr>
<tr>
<td>2016</td>
<td>491.5</td>
<td>1.15</td>
<td>0.97-1.30</td>
<td>147.3 Ref</td>
<td>2.27</td>
<td>2.05-2.51</td>
<td>73.3 Ref</td>
<td>2.07</td>
<td>1.84-2.33</td>
<td>15.0 Ref</td>
<td>1.27</td>
<td>1.00-1.61</td>
<td>14.8 Ref</td>
<td>14.8-15.6</td>
</tr>
</tbody>
</table>

**CT**: Chlamydia; **NG**: Gonorrhea; **CT/NG**: Chlamydia/Gonorrhea Co-infection; **HIV**: Human Immunodeficiency Virus

95% CI: 95% Confidence Intervals; AIRR: Adjusted Incidence Rate Ratios

*Inner-core includes Point Douglas and Downtown Neighbourhood Clusters
Inferential Analysis

The 20-29 year old age group had the highest rates for CT, NG and CT/NG infections (Table 3.2); adjusted for year, sex and inner-core residency, and using those under the age of 20 as reference, rates were 1.5 times higher (95%CI: 1.5-1.6) for CT infections, 1.9 (95%CI: 1.8-2.1) for NG infections, and 1.1 times (95%CI: 1.0-1.2) for CT/NG. For HIV and syphilis, the highest rates were observed in the 30-39 year old age group; rates for this age group were almost 7-fold (AIRR: 6.5; 95%CI: 4.5-9.4) and 12-fold (AIRR: 12.0; 95%CI: 7.3-19.7) higher for HIV and syphilis infections, compared to those under the age of 20. Notably, compared to those under the age of 20, rates of infection amongst those 30 or over were several-fold lower for those with CT, NG and CT/NG infections. The rate of CT amongst 30-39 year olds was approximately half of those under the age of 20, while for NG, the rate amongst 30-39 year olds was approximately 90% of under 20 year old rate. For CT, rates amongst 50-59 year olds were about 5% of those under the age of 20, compared to approximately 20% for NG. CT/NG concomitant infections were even more concentrated in the under 30 years of age group, with nominal differences in rates between those under the age of 20 and those 20-29 years of age (AIRR: 1.1; 95%CI: 1.0-1.2); rates of concomitant infections were especially low amongst those 40 years or older. In comparison, rates for HIV and syphilis were all higher amongst all age groups 30 years or older, relative to the under 20 group.

In terms of secular trends, and compared to the 2007-2009 time period, after a period of declining rates during 2010-2012 and 2013-2015, NG rates in 2016 have doubled (AIRR: 2.3; 95%CI: 2.1-2.5). Significant increases in rates for 2016 were
observed for HIV (AIRR: 1.3; 95%CI: 1.0-1.6) and syphilis (AIRR: 7.5; 95%CI: 5.2-10.7).

The excess burden of infection is illustrated in the sex X inner-core interactions. Compared to females living in non-core areas, rates for females living in the inner-core were over two-fold higher (AIRR: 2.5; 95%CI: 2.4-2.5) for CT and five-fold higher for both NG (AIRR: 5.2; 95%CI: 4.7-5.6) and CT/NG (AIRR: 5.2; 95%CI: 4.6-5.7) infections. Risk for HIV was three-fold higher for inner-core females (AIRR: 3.4; 95%CI: 2.7-4.4) and six-fold higher for syphilis (AIRR: 6.6; 95%CI: 4.1-10.6), compared to their non-core counterparts. Non-core males had lower risk than non-core females for CT (AIRR: 0.5; 95%CI: 0.5-0.6) and for CT/NG (AIRR: 0.8; 95%CI: 0.7-0.9). Males residing in the inner-core had the highest rates for HIV (AIRR: 4.3; 95%CI: 3.5-5.4) and syphilis (AIRR: 22.8; 95%CI: 15.3-34.0).

**Geographic Distribution/Attribution**

**PAF**

Using PAF to measure the proportion of infections attributable to inner-core residents, a substantial amount of heterogeneity was found both between type of infection and across time periods (Table 3.3). Inner-core residents were responsible for less than 25% of infections for both CT and HIV across all time periods, while PAF was highest for NG (range: 36.4%-45.9%) and CT/NG (range: 39.2%-46.6%) infections. For all time periods, the PAF for NG and CT/NG infections was significantly higher (p<.05) than the PAF for CT infections, while the PAF for NG and CT/NG infections was significantly higher than HIV for all time periods except 2016. In the 2007-2009 time period, 46% (95%CI: 41.8-49.7) of NG infections were attributable to inner-core residents; in 2013-2015 and 2016,
Table 3.3: Indices (and 95% Confidence Intervals) of Geographic Concentration, Sexually Transmitted and Bloodborne Infections in the Winnipeg Health Region, by Time Period

<table>
<thead>
<tr>
<th>Index</th>
<th>Year</th>
<th>CT</th>
<th>NG</th>
<th>CT/NG</th>
<th>HIV</th>
<th>Syphilis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>2007-2009</td>
<td>20.8 (18.5-23.0)</td>
<td>45.9 (41.8-49.7)</td>
<td>46.6 (41.0-51.8)</td>
<td>22.1 (13.5-29.8)</td>
<td>19.8 (0.02-0.37)</td>
</tr>
<tr>
<td></td>
<td>2010-2012</td>
<td>20.3 (18.0-22.5)</td>
<td>42.8 (38.7-46.6)</td>
<td>40.8 (35.0-46.1)</td>
<td>21.7 (13.1-29.4)</td>
<td>23.0 (0.02-0.42)</td>
</tr>
<tr>
<td></td>
<td>2013-2015</td>
<td>21.4 (19.2-23.4)</td>
<td>36.6 (32.2-40.7)</td>
<td>42.2 (36.7-47.2)</td>
<td>20.3 (12.2-27.8)</td>
<td>27.2 (18.2-35.6)</td>
</tr>
<tr>
<td></td>
<td>2016</td>
<td>17.1 (14.0-20.2)</td>
<td>36.4 (31.6-40.9)</td>
<td>39.2 (32.1-45.5)</td>
<td>22.5 (9.8-33.4)</td>
<td>48.4 (33.5-59.9)</td>
</tr>
<tr>
<td>Gini*,**</td>
<td>2007-2009</td>
<td>0.315 (0.224-0.405)</td>
<td>0.544 (0.423-0.666)</td>
<td>0.613 (0.499-0.727)</td>
<td>0.659 (0.533-0.785)</td>
<td>0.504 (0.375-0.633)</td>
</tr>
<tr>
<td></td>
<td>2010-2012</td>
<td>0.336 (0.239-0.427)</td>
<td>0.511 (0.373-0.649)</td>
<td>0.554 (0.425-0.684)</td>
<td>0.649 (0.523-0.775)</td>
<td>0.708 (0.590-0.825)</td>
</tr>
<tr>
<td></td>
<td>2013-2015</td>
<td>0.327 (0.172-0.373)</td>
<td>0.484 (0.355-0.613)</td>
<td>0.565 (0.431-0.699)</td>
<td>0.646 (0.515-0.778)</td>
<td>0.376 (0.264-0.489)</td>
</tr>
<tr>
<td></td>
<td>2016</td>
<td>0.273 (0.183-0.363)</td>
<td>0.483 (0.358-0.608)</td>
<td>0.554 (0.414-0.694)</td>
<td>0.594 (0.458-0.731)</td>
<td>0.589 (0.462-0.716)</td>
</tr>
</tbody>
</table>

CT: Chlamydia; NG: Gonorrhea; CT/NG: Chlamydia/Gonorrhea Co-infection; HIV: Human Immunodeficiency Virus

*Age- and sex-adjusted; **Gini coefficient closer to “0” indicates equal geographical dispersion while a coefficient closer to “1” indicates unequal dispersion
the PAF had decreased to 36.6% (95%CI: 32.2-40.7) and 36.4% (95%CI: 31.6-40.9), respectively. PAF for syphilis increased every time period, from 19.8% (95%CI: 0.02-0.37) during 2007-2009 to 48.4% (95%CI: 33.5-59.9) in 2016.

Similar to PAF, the lowest values for the Gini coefficient (indicating a more even geographic distribution) was seen amongst CT infections (range: 0.27-0.34). Unlike PAF, however, the highest Gini coefficient was observed amongst HIV infections (range: 0.59-0.66). With the exception of syphilis, Gini coefficients for all infections have been decreasing since the 2007-2009 time period. Figures 3.4-3.7 contain maps of all STBBI cases in the WHR, by time period. The emergence of syphilis in this time period can clearly be seen.

Figure 3.2 shows the distribution of age-adjusted and relative rates of NG by NCs, comparing outbreak years vs. preceding years. From 2009-2011 (54/100,000) to 2012 (71/100,000), the average relative rate by NC was 1.7 (95%CI: 0.7-3.1) with 21 out of 25 NCs experiencing an increase. In comparison, from 2013-2015 (49/100,000) to 2016 (144/100,000), the average increase in relative rate by NC was 3.4 (95%CI: 1.8-6.3), with all NCs experiencing an increase in rates. Generally speaking, the largest increases in relative rates from 2013-2015 to 2016 were observed among NCs with lower rates in 2013-2015. However, it should be noted that in absolute terms, NCs with the highest NG rates in 2013-2015 continued to contribute the largest number of NG cases in 2016.
Figure 3.2: Age-Adjusted and Relative Rates of Gonorrhea by Neighbourhood Clusters, Outbreak Years (2012 & 2016) vs. Endemic Rate (2009-2011 & 2013-2015)

2012 vs 2009-2011 denotes the relative rate between the two time periods (right axis)

2016 vs 2013-2015 denotes the relative rate between the two time periods (right axis)
Figure 3.3 compares the distribution, by NC, of NG rates in outbreak years to 2009-11 CT and NG rates, and 2013-2015 CT and NG rates. It can be seen that the “profile” of NG rates, by NC, in the 2016 outbreak closely matches that of CT in 2009-2011 and 2013-2015. In order to quantify this relationship, Spearman’s rank correlation was used to correlate the distribution of NG rates in outbreak years to those of CT and NG in 2009-2011 and 2013-2015 (Table 3.4). For the 2016 NG outbreak, correlation was highest with the distribution of 2013-2015 CT rates (\( \rho \): 0.91, p<.0001). For the 2012 NG outbreak, correlation was highest with 2009-2011 CT rates (\( \rho \): 0.73, p<.0001).
Figure 3.3: Age-Adjusted Rates of Gonorrhea (NG) & Chlamydia (CT) by Neighbourhood Clusters, Gonorrhea Outbreak Years (2012 & 2016) vs. Endemic Rate (2009-2011 & 2013-2015)

Table 3.4: Spearman’s Rank Correlation, Comparing Age-Adjusted Rates in Gonorrhea Outbreak Years, by Neighbourhood Cluster

<table>
<thead>
<tr>
<th>Endemic Rate of Infection, by Time Period</th>
<th>Gonorrhea Outbreak Years</th>
<th>2016</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009-11</td>
<td>0.834*</td>
<td>0.681*</td>
<td></td>
</tr>
<tr>
<td>2013-15</td>
<td>0.762*</td>
<td>0.726*</td>
<td></td>
</tr>
<tr>
<td>Chlamydia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009-11</td>
<td>0.835*</td>
<td>0.731*</td>
<td></td>
</tr>
<tr>
<td>2013-15</td>
<td><strong>0.910</strong></td>
<td>0.708*</td>
<td></td>
</tr>
</tbody>
</table>

*p < 0.0001
3.4. DISCUSSION

Using population-based public health surveillance data from a large urban setting, this study demonstrated the dynamic epidemiology of STBBIs over a 10-year time span, and reiterated the high rates of infection observed in Winnipeg in previous studies. The 2016 rate for CT (476/100,000) in Winnipeg is similar to the 2015 CT rate (479/100,000) in the US, while much higher than the 2012 Canadian CT rate (299/100,000).\textsuperscript{146,182} The 2016 NG rate in Winnipeg (128/100,000) was higher than both the 2015 US (123/100,000) and 2012 Canadian (36/100,000) rates.\textsuperscript{146,182} The 2016 rate for infectious syphilis in Winnipeg (14/100,000) was almost double that of the 2015 US rate (7.5/100,000) and almost threefold the 2012 reported rate in Canada (6/100,000).\textsuperscript{146,182}

**NG: Early Growth Phase**

Although all pathogens showed differing epidemiological patterns within the study period, results suggest the reported incidences of CT, NG, CT/NG, HIV and syphilis have all been increasing since 2013. Wasserheit and Aral proposed a phase-specific framework to understand the epidemiology of STBBIs.\textsuperscript{139} According to this framework, STBBIs undergo growth, hyperendemic, decline and endemic phases. Importantly, Wasserheit and Aral also proposed specific strategies appropriate for each phase,\textsuperscript{139} although there has been a lack of empirical evidence as to which interventions are most effective for each phase.\textsuperscript{183} Increasing NG incidence, and a higher degree of geographical dispersion (as measured by PAF and the Gini coefficient) suggests that NG may be undergoing an early growth phase in the city of Winnipeg. The findings from analyses using PAF and Gini coefficients align with our findings that increases in NG rates have been observed across all NCs in 2016, and that in terms of geographic
distribution, 2016 NG rates resemble more recent years of CT rates the most. The fact that NG seems to be moving away from its historical niche in the inner-core of Winnipeg has the potential to complicate control and prevention efforts, as the typical response to dampen transmission is to intensify interventions amongst targeted populations. A more broadly dispersed geographical base from which NG can be acquired and transmitted will be a challenge for Winnipeg’s public health units, especially in the context of an established and heterogeneous outbreak of infectious syphilis.

**Ecological Niche**

Our results also demonstrate that although each pathogen is transmitted similarly through sexual contact (with the exception of HIV, which can be transmitted through contaminated syringes or other drug paraphernalia), each pathogen has also managed to inhabit and define its own ecological niche. CT and NG infections are most prevalent amongst those under the age of 30, with a rapid decrease in rates after the age of 30. There was marked heterogeneity in this decrease between CT and NG, with CT declining at a much faster rate by age group, compared to NG. CT/NG concomitant infections occurred almost exclusively amongst those under the age of 30. In contrast, incidence rates for HIV and syphilis are highest amongst 30-39 year olds, and virtually non-existent among those under the age of 20.

This study also demonstrated the utility of using concentration indices in order to understand each pathogen’s ecological niche. Similar to previous studies, CT was shown to be more geographically dispersed, when compared to NG and NG/CT infections. This was illustrated both by the Gini coefficient, as well as PAF. Using PAF, a greater proportion of incident NG infections (compared to CT infections) could be
averted if prevention and control efforts were focused on inner-core residents. For example, in 2013-2015, 21% of CT infections were attributable to inner-core residents, compared to 37% of NG infections. However, with respect to NG, what is also notable is that in the context of increasing incidence, a trend towards less geographic concentration was detected, as measured by PAF. Incidence of NG in 2016 was over two-fold higher than in 2007-2009; at the same time, PAF decreased from 46% to 36% (p<.05), suggesting that infection had spread from inner-core to non-core residents. This was supported by a decreasing trend in the Gini coefficient, from 0.54 in 2007-2009 to 0.48 in 2016 for NG, although it should be noted that the decrease in the Gini coefficient was not statistically significant at the p<.05 level.

Individuals with CT/NG concomitant infections seem to show an even greater degree of geographic clustering, compared to those with NG-only infections, with this clustering based in the inner-core of Winnipeg. A Dutch study using routinely collected surveillance data found concomitant infections in 7% of incident bacterial diagnoses. This same study found concomitant infections were more likely found in specific risk groups, such as MSM, and those of Surinamese or Antillean origin; similarly, a UK study found an association between concomitant infections and black ethnicity. Other studies have found concomitant infections associated with younger age, repeat infections, HIV, higher social deprivation, and lower income. Thus, evidence suggests concomitant infections are associated with highly-clustered, and more “tightly-defined” populations at higher risk of infections from STBBIs. Clearly, concomitant infections occur through some complex interaction between age, ethnicity, socio-economic status and social marginalization and vulnerability, suggesting populations
vulnerable to concomitant CT/NG infections should be prioritized. In contrast, although HIV showed a high degree of geographic clustering, the PAF of HIV suggests that cases were not concentrated in the inner-core, but were concentrated in areas outside the core. The mixed results from syphilis cases (low PAF up to 2016) may be due to syphilis evolving from primarily impacting MSM to also including heterosexual/bisexual populations.

**Outbreaks and Overlap**

Comparison of the two NG outbreaks, and the relative changes by geographic distribution during non-outbreak and outbreak years suggests two inter-related interpretations. First, not all outbreaks are created equal. The 2012 NG outbreak saw a 44% increase in rates (from 2009-2011), with modest increases in rates across some, but not all, NCs. In contrast, the 2016 NG outbreak saw a three-fold increase in rates, with all NCs experiencing an increase in rates; in fact, all but 2 NCs saw a doubling in rates from 2013-2015 to 2016. Furthermore, the magnitude, velocity (i.e., the steep increase from 2015 to 2016) and uncharacteristically across-the-board increase geographically suggest factors beyond behaviour change may also be responsible for the 2016 outbreak. This is especially true as a contemporaneous and dramatic rise in CT rates was not observed (Table 3.1). Changes in pathogen-side factors, such as a higher prevalence of asymptomatic and/or antibiotic resistant subtypes; population-level screening patterns, and even testing technology are potential contributing factors. To the best of our knowledge, no evidence of a pervasive antibiotic resistant NG subtype has been demonstrated in Canada, although data from 2016 has not yet been produced nationally. Currently, no systematic mechanism exists in the WRHA to assess NG symptoms at the
time of infection, leaving comparisons over time challenging. Discussions with public health staff and physicians at key testing clinics have not revealed any unusual patterns in NG symptoms, nor treatment failures. In fact, PHNs have noted an increase in purulent NG cases (personal communication, A. Lapple); however, since this evidence is anecdotal, the possibility of an asymptomatic (or at least a less acute) strain of NG being responsible for the increased infections detected still exists. It should be noted that without knowledge of testing/screening patterns, it is not known whether an increase in asymptomatic NG would result in an increase, or a decrease, in case finding. Increased screening has been shown to increase case finding; however, heterogeneity in cases detected (e.g., more females over males, and specific geographic areas being over-represented) would be expected. \(^{192}\) That increases were observed across both sexes and all geographic areas suggests that screening is not the only contributor. As well, since nearly all NG testing involves CT testing in Manitoba, an increase in CT infections would have been expected. Finally, CT/NG testing technology in Manitoba has not changed since 2008 (personal communication, J.L. Wylie) making it unlikely that the increase in NG infections was due to a testing factor, unless there were wholesale increases in testing sensitivity to NG infections only, which would be an improbable scenario. Likely, subtle changes in behavioural, pathogen, and partnering/partnership characteristics combined synergistically to create conditions ripe for NG rates to substantially increase in Winnipeg in 2016. Because there are no systematic mechanisms to evaluate NG symptoms, the possibility exists that less-symptomatic strains of NG may have contributed to the increased infections observed. A delay in treatment seeking (due to less severe symptoms), mixed with any combination of an increase in number of
partners, an increase in concurrent partnerships, or decreasing condom use in certain types of partnerships, may be a potential explanation.

Second, and especially more so for the 2016 NG outbreak, a “spillover” effect was observed whereby geographic distribution of NG more closely resembled that of CT. NCs with elevated rates of NG in 2016 overlapped with areas of elevated CT rates in 2013-2015. Differences in epidemiology between NG and CT are thought to be driven by the higher likelihood of NG infections being symptomatic; given symptomatic cases are more likely to seek care, NG is thus more likely to be transmitted in “core” populations that are marked by high connectivity.\textsuperscript{193-195} That is, the “reachable path” of infection is determined by the pathogen’s properties.\textsuperscript{6} Therefore shared geographic space between NG and CT could be circumstantial evidence of a change in the properties of the pathogen itself, vis-à-vis a reduction in symptoms, or an increase in antimicrobial resistance in the strains that were circulating in 2016. Genotyping of strains would help to determine whether either of these shifts have occurred. From a network perspective, the geographic overlap between 2013-2015 CT and 2016 NG cases highlights the importance of bridging populations that may connect networks where CT is dominant, to those where NG is more likely to appear, that is, from “higher-risk” to “lower-risk” networks.\textsuperscript{196} Alternatively, instead of “bridging” between distinct sexual networks, which implies separate networks that are occasionally bridged by select individuals who are uniquely positioned to link different networks, the possibility exists of a larger, continuous sexual network, marked by a denser core where sexual contacts and opportunities for spread are more frequent, and decreasing density as the periphery of the network is reached. Being less symptomatic, CT is able to take advantage of all parts of the continuous network,
while NG would be restricted to the core. A less symptomatic strain of NG would thus simply be allowed to spread from the core, to other parts of the network. This scenario is distinct from bridging by at least two aspects: sexual networks are not distinct, and spread to other parts of the sexual network is not contingent on a select few individuals, but is more a permeation, through multiple paths that are already inherent in the network. From a broader, systems-level perspective, and echoing the modeling work of Robinson et al., both scenarios suggest that within any locality, the “universe” of existing sexual and social networks form the pool of potential STBBI cases; differences in the epidemiology of NG and CT (and the populations that are then seen as vulnerable to these pathogens) are an expression of the interaction between pathogenic properties and the transmission networks which most successfully transmit each specific pathogen. From the prevention aspect, the focus should then not solely be on CT vs. NG *per se*, but the syndemic nature of STBBIs.

**Strengths and Limitations**

Our analyses benefit from a number of strengths, including the use of population-based data sources, instead of using self-reported data on infections, or data from select clinical populations. As Manitobans are covered by universal health insurance, ability to pay is not a barrier to health care access. An additional strength was that five distinct epidemics were examined, thus enabling comparisons between epidemics. Postal code data allowed for place-based analyses. There are also several limitations to our study. First and foremost, behavioural and partner-side information was limited. Second, data likely underestimate the actual incidence of STBBIs, at the individual level, as databases are based on laboratory confirmation of infections. Relatedly, data may overestimate the
degree to which incidence is increasing, as the analysis was based on the number of infections, and not cases. Specifically, increased testing/screening may account for the increases in observed incidence. Moreover, if increases in testing/screening are limited to certain geographic areas (e.g., those with lower socio-economic conditions), which might occur in more targeted prevention/intervention efforts, the potential to artificially bias patterns in the Gini coefficient and/or PAF exists. However, it should be noted that increased testing/screening would likely result in increasing inequality, which is the opposite of what was observed in our NG data. Increases in testing/screening in geographic areas, while testing rates remaining the same in areas historically at higher risk of NG infections would account for a lessening of the inequality in the distribution of NG infections. Fourth, the analyses examining geographic spread in outbreak years was limited to NG only; CT/NG infections and infectious syphilis were not examined. Future work should examine whether outbreaks of infectious syphilis and CT/NG also follow the geographic distribution of CT in preceding years. Finally, CT/NG infections were treated separately from CT and NG infections; further research should ascertain whether grouping CT/NG co-infections into their respective individual infections would result in the same findings. For the purposes of this dissertation, given the abundance of evidence in the literature demonstrating the uniqueness of individuals co-infected with CT/NG, it was decided *a priori* to group individuals with concomitant infections separately.

**Conclusion**

Our findings demonstrate NG rates being drastically on the rise in Winnipeg, Canada. This increase is occurring within the context of more heterosexual transmission of syphilis, continuing syphilis transmission among MSM, and slight increases in HIV
and CT rates. As predicted by the phase-specific framework of Wasserheit and Aral,\textsuperscript{139} this growth phase is occurring through a wider geographic and demographic base, with areas impacted being more similar to the geographic distribution of CT, suggesting an overlap in risk groups. Given limited public health resources for following up cases and their contacts and non-declining rates in other STBBIs, focused prevention initiatives, and efforts to understand the determinants making sub-populations vulnerable to multiple STBBIs should be prioritized.
Figure 3.4: Map of Sexually Transmitted and Bloodborne Infection Cases in the Winnipeg Health Region, 2007-2009
Figure 3.5: Map of Sexually Transmitted and Bloodborne Infection Cases in the Winnipeg Health Region, 2010-2012

Includes only those greater than 15 years of age (2010-2012).
Figure 3.6: Map of Sexually Transmitted and Bloodborne Infection Cases in the Winnipeg Health Region, 2013-2015
Figure 3.7: Map of Sexually Transmitted and Bloodborne Infection Cases in the Winnipeg Health Region, 2016

Includes only those greater than 15 years of age (2016).
Chapter 3 demonstrated the utility of combining place-based analyses with traditional surveillance tools. We demonstrated that gonorrhea was experiencing a growth phase in Winnipeg, and confirmed theoretical predictions of geographic dispersion during the growth phase of an epidemic. We also found that within the context of a growth phase, gonorrhea occupies similar spaces as chlamydia. In Chapter 4, our focus turns to gonorrhea and network analyses. In addition to networks traced out through case-and-contact investigations, another tool in the armamentarium of network analysis, molecular epidemiology, is introduced. The “view” of the sexual networks afforded by case-and-contact investigations and molecular epidemiology is compared and contrasted. Some simple network metrics are introduced alongside bread-and-butter surveillance techniques. As well, a simple technique for testing hypotheses with network data, conditional uniform graph testing, is applied to sexual networks constructed through case-and-contact investigations.
CHAPTER 4: SEXUAL NETWORK AND GENOTYPIC ANALYSIS
OF AN OUTBREAK OF GONORRHEA IN WINNIPEG, CANADA

4.1. INTRODUCTION

Gonorrhea (NG), the infection caused by Neisseria gonorrhoeae, is the second most commonly reported sexually transmitted infection (STI) worldwide. Globally, an estimated 106 million NG infections occur annually. Along with chlamydial infections, NG is one of the most significant drivers of public health resources in developed countries, and is a major contributor to STI-related morbidity. Despite being highly preventable, and the widespread availability of non-invasive diagnostic tests, as well as inexpensive treatment options, STIs such as NG have shown remarkable resilience to treatment and control efforts, and have experienced a resurgence over the last decade in many parts of the world. For example, in the United States, NG rates have increased from 98.1/100,000 in 2009 to 123.9/100,000 in 2015. During this same time period, rates of NG have more than doubled from 30.9/100,000 to 75.8/100,000 in England. Since 1997, the Canadian rate of NG has also more than doubled from 14.9 per 100,000 to 36.2 per 100,000 in 2012.

An expansion of core groups has thought to underlie the resurgence of NG rates. Of note, increases in the rates of sexual partner acquisition, driven by social media technology, and among those vulnerable to HIV infection, the availability of highly active antiviral therapy, pre-exposure prophylaxis, and associated increases in risk behaviour (such as condomless penetrative sex), have all thought to contribute to increased STI incidence. Increasing incidence of NG is particularly worrisome in the context of N. gonorrhoeae’s ability to evolve and develop resistance to many anti-
microbial treatments.\textsuperscript{212-214} Recently, the American Centers for Disease Control (CDC) has declared resistance in NG a serious public health issue, and requiring urgent action, as an increasing incidence of NG, coupled with increased rates of resistance may lead to the possibility of NG infections being untreatable in the future. This potential of NG being untreatable has underscored the urgency to develop effective prevention programs.

At the core of effective prevention programming lies epidemiological intelligence related to the distribution of NG (and other STIs) in a population, the excess burden of risk in particular subpopulations, and secular trends in incidence.\textsuperscript{153,171,215-218} Though not without important limitations, surveillance data have played a critical role in this regard.\textsuperscript{183,219-224} Observing differences in STI rates may lead to the formation of theories used to inform intervention development.\textsuperscript{171} Alongside traditional epidemiological tools characterizing those at risk by person, place and time, significant contributions have been made by network-type analyses\textsuperscript{,27,33,225-229} and molecular epidemiology\textsuperscript{230-232} in understanding and contextualizing transmission patterns of STIs, with some authors suggesting the combination of network-type analyses and molecular epidemiology may offer even deeper insights into the distribution and structure of STI epidemics.\textsuperscript{233-235}

In terms of molecular epidemiology, the literature has shown that certain strains of NG have been linked to geographic, or demographic clusters (such as sexual orientation).\textsuperscript{226,229,232,236} For example, eight of 21 major strains identified among London residents were clustered geographically, persisting for a shorter duration compared to those that were not clustered.\textsuperscript{237} Additionally, several strains observed in the United Kingdom (UK) were predominant in men who have sex with men (MSM),\textsuperscript{229} while the most frequently observed strains differed between heterosexuals and MSM in Wales.\textsuperscript{232}
Furthermore, among MSM, NG strain distribution has shown to be differentiated by sex act (i.e., those reporting only recent receptive oral sex), as well as number of sex partners and race or ethnicity.238,239

As an evolving organism, NG presents challenges for strain discrimination; generally speaking, genotype-based methods have replaced phenotypic methodologies. Of these, *Neisseria gonorrhoea* Multi Antigen Sequence Typing (NG-MAST) offers the best combination of discriminatory power, cost-effectiveness, robustness, rapidness and reproducibility in discriminating between NG strains, and has thus been the method of choice for molecular epidemiology of NG.240 Genetic relatedness of subtypes (STs) can also be ascertained through NG-MAST;240 finally, NG-MAST is supported by an online, open-access database (http://www.ng-mast.net/), enabling comparability across countries. Within Canada, in 2015, the three most common STs were ST-2400 (12.1%), ST-9663 (7.4%) and ST-9514 (6.1%).241 Both ST-2400 and ST-5985 were reported as being previously present in Canada, while ST-9663 first appeared in 2013 in the provinces of Ontario, Quebec, Alberta and British Columbia.200 Knowing ST also confers knowledge of pathogen properties; both ST-2400 and ST-9663 are predominantly comprised of multi-drug resistant profiles, including probable chromosomal mediated resistance, ciprofloxacin resistance, and decreased susceptibility to ceftriaxone and cefixime.200 Knowledge of resistance and susceptibility profiles, in turn, can inform clinical and public health treatment responses as well as strategies to decrease incidence of future infections.

This study sought to compare biological information from NG-positive individuals and their contacts with individual level data. Therefore, the objectives of this
study were to: 1) determine the distribution of NG subtypes in a population-based sample of NG infections reported to public health; 2) describe the subtypes by socio-demographic and sexual partnership attributes; and 3) to compare networks created through genotyping to those created by case-contact data. To the best of our knowledge, this is the first study incorporating the use of NG-MAST with public health data in Winnipeg.

4.2. METHODS

Population Setting
The Winnipeg Health Region (WHR) is the largest health region in the province of Manitoba, a central Canadian province. In 2010, the population of the WHR was 700,000, or almost 70% of the population of Manitoba. The WHR is divided into 25 Neighbourhood Clusters (NCs), administrative areas used for planning purposes, with each NC containing approximately the same number of individuals. There is marked heterogeneity in STBBI rates between NCs, with higher rates associated with the “inner-core” – four NCs areas disproportionately impacted by low socio-economic status and other vulnerabilities. Informed consent was not obtained as this was a secondary analysis of anonymized surveillance data.

Data Sources

Enhanced NG Database
All laboratory-confirmed CT and NG infections are reported to Manitoba Health, Seniors and Active Living (MHSAL), who then refer all cases to the regional health authority responsible for follow-up (usually the regional health authority in which the case resided at the time of testing). In response to an NG outbreak in the WHR in 2012, public health
staff of the Population and Public Health Program of the Winnipeg Regional Health Authority (WRHA) implemented enhanced surveillance of NG cases from January 1, 2014 to October 31, 2014. Enhanced surveillance investigation involved augmenting routine investigation of cases and contacts. In Manitoba, public health staff are mandated by the Public Health Act to initiate case and contact investigations of NG cases.

Enhanced surveillance involved the following four main areas: 1) asking a more robust set of questions related to the case; 2) interrogating sexual and behavioural interactions with their contacts; 3) dyad-specific questions, including location of where they met their contacts, and duration and type of relationship; and 4) genotyping of positive urine specimens collected as part of screening.

Upon referral by MHSAL, the WRHA’s public health staff contact each NG case; from there, a contact list is created containing rudimentary information on each case’s contact. Public health staff then follow-up and notify contacts that they have been in contact with someone who has tested positive for NG. Information collected from case and contact investigations was entered into an Epi-Data database (Christiansen & Lauritsen, 2010; Denmark) by a single research assistant.

Definitions
Specimen collection date was used to define the date of infection, and to calculate age at each infection. For the purposes of this analysis extra-genital infections were excluded. Any positive test occurring within 7 days of a previous positive test, for a particular infection, was treated as the same infection; subsequent infections within 7 days were excluded from further analyses. Cases who resided in the Point Douglas North and South
and Downtown West and East NCs were categorized as residing in the “inner-core”; all others were considered “non-core”.

**Laboratory Testing & Genotyping**

All specimens were tested at Cadham Provincial Laboratory (CPL), the sole public health laboratory in Manitoba. CPL is responsible for performing the majority (95%) of CT and NG diagnostics in the province of Manitoba.\(^{242}\) Samples are dual-tested for CT and NG. Since 2007, the Hologics/GenProbe Aptima NAAT assay has been used for detection of CT and NG from urethral/cervical swab and urine specimens.

NG-MAST was used to determine the molecular genotypes of gonorrhea-positive specimens as previously described.\(^{243}\) DNA was extracted from residual Aptima specimens using the QIAamp viral RNA minikit according to manufacturer’s instructions. DNA sequencing was performed at the National Microbiology Laboratory (Winnipeg, MB, Canada). Sequence types were assigned by concatenating porin gene (*por*) and transferrin-binding protein gene (*tbpB*) sequences and submitting to the NG-MAST website (http://www.ng-mast.net/).

**Statistical Analyses**

For objectives 1 and 2, and in addition to visualization of the distribution of all STs, the five most frequently reported STs were described by socio-demographic and clinical characteristics. Specifically, the distribution of the five most-frequently reported STs was compared by sex, age group, residency in the inner-core, chlamydia (CT) co-infection, and whether the case was symptomatic at the time of reporting (yes/no/unknown or don’t know). Pearson’s chi-square test of association was used to assess statistical significance. The most frequently reported ST was described by socio-demographic and clinical
characteristics, and compared to remaining STs in bivariate analyses. Multivariable logistic regression models were used to assess the association between the above characteristics and being infected with the most frequently reported ST. For multivariable analyses, age, sex and residency in the inner-core were kept as confounding variables. Variables significant at the p<.20 level in bivariate analyses were included in multivariable analyses. All two-way interactions were assessed and included if significant at the p<.05 level. Adjusted odds ratios (AORs) and their 95% confidence intervals (95%CIs) are reported. In total, 200 NG cases were investigated in the study period; only cases who were subtyped were kept for statistical and network analyses (N=126/200; 63%). Table 1 includes behavioural data collected as part of follow-up public health investigations; however, only a fraction of those subtyped had follow-up information (N: 62/126; 50%). Therefore, behavioural data are presented only as descriptive information and were not included as part of multivariable analyses.

**Network Analyses**

Sexual networks constructed from case-contact investigations were visualized using the Ucinet (Kentucky, Analytic Technologies, 2008) suite of programs (Version 6). The distribution of component sizes was calculated for the constructed sexual network, and compared to the distribution of component sizes when genotyping data were used to group cases. Sexual networks were characterized with the use of univariate statistics, including density, dyad- and triad-census, degree distribution (in- and out-), and in- and out-degree centralization. Dyadic census classifications are based on Holland and Leinhardt,\textsuperscript{244} while triadic census classifications are based on Davis and Leinhardt.\textsuperscript{245} Degree centralization was calculated based on the work of Freeman.\textsuperscript{246} This measure is
meant to convey the amount of inequality in degree distribution, with a score=0 denoting all nodes have the same degree distribution, and a score=1 denoting one node has the maximum score for degree distribution, while all other nodes have the minimum score. Measures of centrality, such as betweenness and closeness were not calculated due to preliminary analyses demonstrating a lack of connectedness among network members, beyond dyads and triads.

In order to assess the significance of the degree centralization observed, conditional uniform graph (CUG) tests were performed. Similar to the method used by Grund and Densley, 1000 random graphs were created with the same density and number of nodes as the empirical data. Degree centralization was calculated from each of the 1000 random graphs, and two-sided $z$-scores were used to assess whether or not the empirical observation was drawn from the simulated distribution. Component sizes, univariate statistics and the generation of random graphs for CUG testing were performed using the *nwcommand* suite of programs (Grund, 2015) in Stata (V13, College Station, TX).

**4.3. RESULTS**

In total, from the 126 cases who were subtyped, 41 STs were found, with 21 STs appearing more than once. Figure 4.1 shows the distribution of STs; with $n=22$ (or 17% of all STs), ST-3671 was the most predominant ST, followed by ST-11531 ($n=14$), ST-2992 ($n=14$), ST-3307 ($n=7$) and ST-9663 ($n=8$). These five STs were the only STs with more than 10 isolates, and together composed 51% ($n=62$) of all STs genotyped. Table 4.1 displays the demographic and clinical characteristics associated with each of the five most common circulating STs. CT co-infection was significantly associated at the $p<.05$
level, with a high degree of co-infection (i.e., >65%) observed amongst those infected with ST-3671, ST-11531 and ST-3307.

Characteristics of those infected with ST-3671 (the most common) were compared to all other subtypes in bivariate analyses (Table 4.2). Those infected with ST-3671 differed significantly (at the p<.05 level) by age group (p=.002) and by CT co-infection (p=.012). In multivariable logistic regression analyses (Table 4.3), age group remained statistically significant, while the CT co-infection and inner-core residency interaction was the only two-way interaction significant at the p<.05 level. In this interaction, CT co-infection was over six-fold higher (AOR: 6.5; 95%CI: 1.6-27.3) in non-core areas among those infected with ST-3671, compared to non-core cases not infected with ST-3671. For non-core cases, CT co-infection was 73% in ST-3671 cases, compared to 36% of non-core cases not infected with ST-3671. For inner-core cases, CT co-infection was 50% in ST-3671 cases, while co-infection was 38% for cases not infected with ST-3671. It should be noted that although in absolute numbers, only 30% of ST-3671 infections were among inner-core residents, in terms of rates, the prevalence of ST-3671 was twice as high among inner-core residents (5/100,000), compared to non-core residents (2.4/100,000).
Figure 4.1: Distribution of Gonorrhea Subtypes, Winnipeg Health Region (2014, N=126)
Table 4.1: Gonorrhea Subtypes by Socio-Demographic and Testing Characteristics, Winnipeg Health Region (2014)

<table>
<thead>
<tr>
<th></th>
<th>ST-3671</th>
<th>ST-11531</th>
<th>ST-2992</th>
<th>ST-3307</th>
<th>ST-9663</th>
<th>All other Subtypes</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>13</td>
<td>64.1%</td>
<td>7</td>
<td>50.0%</td>
<td>6</td>
<td>85.7%</td>
<td>25.0%</td>
</tr>
<tr>
<td>Male</td>
<td>8</td>
<td>38.1%</td>
<td>8</td>
<td>57.1%</td>
<td>7</td>
<td>50.0%</td>
<td>14.3%</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>100.0%</td>
<td>14</td>
<td>100.0%</td>
<td>14</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td><strong>Age Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;15</td>
<td>0</td>
<td>0.0%</td>
<td>1</td>
<td>7.1%</td>
<td>1</td>
<td>7.1%</td>
<td>0.0%</td>
</tr>
<tr>
<td>15-19</td>
<td>13</td>
<td>61.9%</td>
<td>2</td>
<td>55.0%</td>
<td>5</td>
<td>57.1%</td>
<td>28.6%</td>
</tr>
<tr>
<td>20-24</td>
<td>5</td>
<td>23.8%</td>
<td>7</td>
<td>50.0%</td>
<td>5</td>
<td>55.0%</td>
<td>57.1%</td>
</tr>
<tr>
<td>25+</td>
<td>3</td>
<td>14.3%</td>
<td>4</td>
<td>28.6%</td>
<td>3</td>
<td>21.4%</td>
<td>38.6%</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>100.0%</td>
<td>14</td>
<td>100.0%</td>
<td>14</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td><strong>Residence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Core</td>
<td>15</td>
<td>71.4%</td>
<td>50</td>
<td>50.0%</td>
<td>10</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Inner-Core</td>
<td>6</td>
<td>28.6%</td>
<td>50</td>
<td>50.0%</td>
<td>4</td>
<td>28.6%</td>
<td>28.6%</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>100.0%</td>
<td>14</td>
<td>100.0%</td>
<td>14</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td><strong>Chlamydia Positive</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>7</td>
<td>33.3%</td>
<td>11</td>
<td>78.6%</td>
<td>5</td>
<td>57.1%</td>
<td>78.6%</td>
</tr>
<tr>
<td>Yes</td>
<td>14</td>
<td>66.7%</td>
<td>3</td>
<td>21.4%</td>
<td>2</td>
<td>14.3%</td>
<td>21.4%</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>100.0%</td>
<td>14</td>
<td>100.0%</td>
<td>14</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td><strong>Symptomatic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1</td>
<td>4.8%</td>
<td>4</td>
<td>28.6%</td>
<td>2</td>
<td>14.3%</td>
<td>14.3%</td>
</tr>
<tr>
<td>Yes</td>
<td>14</td>
<td>66.7%</td>
<td>7</td>
<td>50.0%</td>
<td>3</td>
<td>38.6%</td>
<td>38.6%</td>
</tr>
<tr>
<td>UK/ Don't know</td>
<td>6</td>
<td>28.6%</td>
<td>3</td>
<td>21.4%</td>
<td>5</td>
<td>35.7%</td>
<td>35.7%</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>100.0%</td>
<td>14</td>
<td>100.0%</td>
<td>14</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td><strong>Number of sex partners</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>35.7%</td>
<td>5</td>
<td>71.4%</td>
<td>1</td>
<td>50.0%</td>
<td>50.0%</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>14.3%</td>
<td>2</td>
<td>25.0%</td>
<td>2</td>
<td>28.6%</td>
<td>28.6%</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>21.4%</td>
<td>1</td>
<td>12.5%</td>
<td>0</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>4+</td>
<td>2</td>
<td>14.3%</td>
<td>1</td>
<td>12.5%</td>
<td>0</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>UK/ Don't know</td>
<td>2</td>
<td>14.3%</td>
<td>0</td>
<td>0.0%</td>
<td>0</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>100.0%</td>
<td>8</td>
<td>100.0%</td>
<td>7</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td><strong>Case used condoms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>1</td>
<td>7.1%</td>
<td>0</td>
<td>0.0%</td>
<td>1</td>
<td>14.3%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Some of the time</td>
<td>5</td>
<td>35.7%</td>
<td>3</td>
<td>37.5%</td>
<td>4</td>
<td>37.1%</td>
<td>37.1%</td>
</tr>
<tr>
<td>Most of the time</td>
<td>3</td>
<td>21.4%</td>
<td>4</td>
<td>50.0%</td>
<td>0</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

P*: 0.136 0.059 0.235 0.002 0.052 0.706 0.578
<table>
<thead>
<tr>
<th></th>
<th>Always</th>
<th>UK/ Don't know</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Do you use condoms with:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some partners</td>
<td>42.9</td>
<td>25.0</td>
<td>36.8</td>
</tr>
<tr>
<td>All partners</td>
<td>14.3</td>
<td>25.0</td>
<td>18.8</td>
</tr>
<tr>
<td>UK/ Don't know</td>
<td>42.9</td>
<td>12.5</td>
<td>37.4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>14</td>
<td>100.0</td>
<td>83.3</td>
</tr>
</tbody>
</table>

| **Alcohol Use**  |        |                |       |
| No               | 21.4   | 25.0           | 23.6  |
| Yes              | 57.1   | 75.0           | 69.4  |
| UK/ Don't know   | 21.4   | 12.5           | 20.0  |
| **Total**        | 14     | 100.0          | 83.3  |

| **Drug Use**     |        |                |       |
| No               | 42.9   | 37.5           | 43.7  |
| Yes              | 35.7   | 21.4           | 40.7  |
| UK/ Don't know   | 21.4   | 12.5           | 20.0  |
| **Total**        | 14     | 100.0          | 83.3  |

*based on Pearson’s Chi-square test; ** Inner-core includes Point Douglas and Downtown Neighbourhood
Table 4.2: Subtype-3671 vs. all Other Gonorrhea Subtypes, Winnipeg Health Region (2014)

<table>
<thead>
<tr>
<th></th>
<th>ST-3671</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>13</td>
<td>61.9</td>
<td>47</td>
<td>44.8</td>
</tr>
<tr>
<td>Male</td>
<td>8</td>
<td>38.1</td>
<td>58</td>
<td>55.2</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>100.0</td>
<td>105</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Age Group</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.002</td>
</tr>
<tr>
<td>&lt;15</td>
<td>0</td>
<td>0.0</td>
<td>4</td>
<td>3.8</td>
</tr>
<tr>
<td>15-19</td>
<td>13</td>
<td>61.9</td>
<td>22</td>
<td>21.0</td>
</tr>
<tr>
<td>20-24</td>
<td>5</td>
<td>23.8</td>
<td>38</td>
<td>36.2</td>
</tr>
<tr>
<td>25+</td>
<td>3</td>
<td>14.3</td>
<td>41</td>
<td>39.0</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>100.0</td>
<td>105</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Residence</strong></td>
<td>0.147</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Core</td>
<td>15</td>
<td>71.4</td>
<td>57</td>
<td>54.3</td>
</tr>
<tr>
<td>Inner-Core</td>
<td>6</td>
<td>28.6</td>
<td>48</td>
<td>45.7</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>100.0</td>
<td>105</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Chlamydia Positive</strong></td>
<td>0.012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>7</td>
<td>33.3</td>
<td>65</td>
<td>63.1</td>
</tr>
<tr>
<td>Yes</td>
<td>14</td>
<td>66.7</td>
<td>38</td>
<td>36.9</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>100.0</td>
<td>105</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Symptomatic</strong></td>
<td>0.223</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1</td>
<td>4.8</td>
<td>13</td>
<td>12.4</td>
</tr>
<tr>
<td>Yes</td>
<td>14</td>
<td>66.7</td>
<td>49</td>
<td>46.7</td>
</tr>
<tr>
<td>UK/Don’t Know</td>
<td>6</td>
<td>28.6</td>
<td>33</td>
<td>41.0</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>100.0</td>
<td>105</td>
<td>100.0</td>
</tr>
</tbody>
</table>

*based on Pearson’s Chi-square test; **Inner-core includes Point Douglas and Downtown Neighbourhood
Table 4.3: Adjusted Odds Ratios (OR) and 95% Confidence Intervals (95%CI) from Logistic Regression Models, Socio-demographic and Testing Characteristics Associated with Subtype-3671 (>15 years of age)

<table>
<thead>
<tr>
<th></th>
<th>AOR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td><em>Ref</em></td>
</tr>
<tr>
<td>Male</td>
<td>0.42 (0.12-1.47)</td>
</tr>
<tr>
<td><strong>Age Group</strong></td>
<td></td>
</tr>
<tr>
<td>15-19</td>
<td>5.95 (1.12-31.75)</td>
</tr>
<tr>
<td>20-24</td>
<td>2.49 (0.47-13.16)</td>
</tr>
<tr>
<td>25+</td>
<td><em>Ref</em></td>
</tr>
<tr>
<td><em><em>Residence</em> X Chlamydia Infection Interaction</em>*</td>
<td></td>
</tr>
<tr>
<td>Non-Core, No Chlamydia Co-infection</td>
<td><em>Ref</em></td>
</tr>
<tr>
<td>Non-Core, Chlamydia Co-infection</td>
<td>6.49 (1.55-27.29)</td>
</tr>
<tr>
<td>Inner-Core, No Chlamydia Co-infection</td>
<td>2.38 (0.37-15.34)</td>
</tr>
<tr>
<td>Inner-Core, Chlamydia Co-infection</td>
<td>1.27 (0.23-7.13)</td>
</tr>
</tbody>
</table>

N=126

*Inner-core includes Point Douglas and Downtown Neighbourhood*
Network Analyses

Figure 4.2 shows a visualization of the sexual network constructed from case and contact investigations, grouped by the five most frequent subtypes. As can be seen, the sexual network was highly-fragmented, consisting mainly of dyads, triads and tetrads, and lacking long chains and/or closed loops. This visualization is confirmed by Table 4.4, which shows the distribution of component sizes; of the 85 components found in the data, 61% (52/85) were simply dyads, where a case would name one contact, and the contact would not be named by any other case. The size of the largest component was 6.

Genotyping data allowed cases and contacts to be clustered together, as shown in Figure 4.2 and Table 4.4. When genotyping data were used to cluster cases and their contacts together, the potential size of each component increased. Of the 33 potential components now found when genotyping data were incorporated, only 32% (10/33) were dyadic in nature, with a component comprised of 45 cases and their contacts now the largest component (ST-3671). Although some subtypes (e.g., ST-9663) were clustered around the inner-core, many subtypes were distributed across the city of Winnipeg. For example, the two most common subtypes, ST-3671 and ST-11531 were found in 7 out of 12, and 6 out of 12 Community Areas (CAs), respectively (data table not shown). In contrast, with the exception of one case, ST-9663 was found in only 2 out of 12 CAs. It should be emphasized here that having the same ST is only suggestive of actual linkage among cases; validation of an actual linkage between cases would need to be made with epidemiological data.
Figure 4.2: Sexual Contact Network from Case and Contact Investigations and Gonorrhea Subtypes, Winnipeg Health Region (2014)
Univariate descriptive statistics of the sexual network constructed from public health investigations are shown in Table 4.5. A total of 221 nodes were included, with 142 ties between them, for a density of less than 0.003. Dyadic census indicated that the majority of dyads were asymmetric (i.e., only the case would name the contact, and not the reverse), likely due to contacts not being named or unsuccessful attempts to make contact with named contacts. The dyadic nature of the case-contact data is also illustrated in the results from the triadic census, with the majority of the structures being unidirectional (012; N=28,336) and bidirectional (102; N=1,308) dyads. In discussing the triad census, reference is made to Appendix 3, which contains illustrations of the triadic forms. Of structures that were triads, the vast majority were where a case would name two contacts, but would not be named by their contact (021D; N=56). Of note, there was only once instance where a contact was named by two different individuals (021U). With respect to in-degree distribution, 63% of nodes were only nominated by one other person, while over 85% of nodes nominated one or fewer individuals. Results from CUG testing indicated that observed in-degree centralization value of 0.0062 was significantly less than expected (Fig 4.3; z-score:-2.56; p<.05) when compared to the distribution of in-degree centralization scores from 1000 random graphs with the same number of nodes and density. Thus, the inequality of being nominated by another individual was less in the observed data, compared to the expected; this was likely due to almost all nodes being nominated exactly once (i.e., an in-degree of 1). Finally, for ST-9663, it should be noted that public health investigations revealed an epidemiological connection to Ontario, Canada, as well as the presence of bisexual partnerships in this network.
Table 4.4: Component Sizes, Case and Contact Investigations, and from Gonorrhea Genotyping, Winnipeg Health Region (2014)

<table>
<thead>
<tr>
<th>Size</th>
<th>Case &amp; Contact only</th>
<th>Case &amp; contact with genotyping (potential component sizes if cases all epidemiologically-linked)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>52</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>--</td>
<td>2</td>
</tr>
<tr>
<td>28</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>45</td>
<td>--</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 4.5: Univariate Statistics, Sexual Network Constructed from Case & Contact Investigations, Winnipeg Health Region (2014)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Total Network</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nodes 221</td>
</tr>
<tr>
<td></td>
<td>Ties 142</td>
</tr>
<tr>
<td></td>
<td>Density 0.003</td>
</tr>
<tr>
<td>Dyad Census</td>
<td></td>
</tr>
<tr>
<td>Mutual</td>
<td>6</td>
</tr>
<tr>
<td>Asymmetric</td>
<td>130</td>
</tr>
<tr>
<td>Null</td>
<td>24,174</td>
</tr>
<tr>
<td>Triad Census*</td>
<td></td>
</tr>
<tr>
<td>003</td>
<td>1,744,916</td>
</tr>
<tr>
<td>012</td>
<td>28,336</td>
</tr>
<tr>
<td>021D</td>
<td>56</td>
</tr>
<tr>
<td>021U</td>
<td>1</td>
</tr>
<tr>
<td>021C</td>
<td>7</td>
</tr>
<tr>
<td>030T</td>
<td>0</td>
</tr>
<tr>
<td>030C</td>
<td>0</td>
</tr>
<tr>
<td>102</td>
<td>1,308</td>
</tr>
<tr>
<td>120D</td>
<td>0</td>
</tr>
<tr>
<td>120U</td>
<td>0</td>
</tr>
<tr>
<td>120C</td>
<td>0</td>
</tr>
<tr>
<td>111D</td>
<td>6</td>
</tr>
<tr>
<td>111U</td>
<td>0</td>
</tr>
<tr>
<td>201</td>
<td>0</td>
</tr>
<tr>
<td>210</td>
<td>0</td>
</tr>
<tr>
<td>300</td>
<td>0</td>
</tr>
<tr>
<td>In Degree</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>80 (36.2)</td>
</tr>
<tr>
<td>1</td>
<td>140 (63.4)</td>
</tr>
<tr>
<td>2</td>
<td>1 (0.45)</td>
</tr>
<tr>
<td>Out Degree</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>126 (57.0)</td>
</tr>
<tr>
<td>1</td>
<td>62 (28.1)</td>
</tr>
<tr>
<td>2</td>
<td>20 (9.1)</td>
</tr>
<tr>
<td>3</td>
<td>12 (5.4)</td>
</tr>
<tr>
<td>4</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Centralization Measures</td>
<td></td>
</tr>
<tr>
<td>In-degree centralization</td>
<td>0.0062</td>
</tr>
<tr>
<td>Out-degree centralization</td>
<td>0.0153</td>
</tr>
</tbody>
</table>

*Inner-core includes Point Douglas and Downtown Neighbourhood
Figure 4.3: Results from Conditional Uniform Graph Simulations of In-degree and Out-degree Centralization Measures*

*Based on 1000 conditional uniform graphs

4.4. DISCUSSION

Canada’s National Microbiology Laboratory (NML) reported the first appearance of ST-9663 in four Canadian provinces in 2013, with subsequent spread to Manitoba and New Brunswick provinces by the following year. Our surveillance data support these findings, as cases in 2014 were linked, epidemiologically, to case(s) in Ontario, Canada. Thus, from the best available evidence, ST-9663 managed to spread to 6 (out of 10) provinces in Canada within the span of two years; this despite NG being more concentrated in the most marginalized and vulnerable of populations. Of note, analysis of the ST-9663 subtype revealed that both women and bisexual men were infected with this subtype. Not only did ST-9663 manage to gain a foothold in other
provinces with surprising rapidity, it emerged in Manitoba within epidemiologically important populations.

**Presence in Core Populations and Emergent Network Features**

The transmission success of an infectious organism is contingent on its ability to propagate to at least one susceptible host, either before the infected host naturally clears the infection, or prior to the successful completion of treatment. That is, the duration of infectivity has to be of sufficient length for the infected host to have unprotected sex with at least one other susceptible sexual partner, and for the infection to still be present prior to this interaction.²⁵⁰ Because the duration of infectivity for NG is thought to be shorter than for CT, NG is most effectively transmitted through networks with very high connectivity, through any combination of a high number of sexual partners, rapid turnover of partners, and/or a high prevalence of concurrency.²⁵¹,²⁵² From the “point of view” of NG, only highly-connected portions of the overall sexual network are most conducive to ongoing transmission.⁹⁶

Viewed from this perspective, that ST-9663 rapidly crossed provincial borders soon after its introduction, and was present in epidemiologically-relevant core groups, aligns well with network theory, as both pathways are examples of assortative mixing amongst “high-risk” groups.²⁵¹-²⁵⁴ Likewise, the higher prevalence of CT co-infection amongst those infected with ST-3671, as well as the association with younger age is congruent with this perspective, as CT and NG co-infection has been shown to be associated with higher-risk, and more marginalized sub-populations.¹⁸⁵-¹⁹¹,²⁵⁵,²⁵⁶ Less than 30% of those infected with ST-3671 resided in the inner-core, although this proportion may have been higher if all ST-3671 cases were subtyped. Rate-wise however, the
prevalence of ST-3671 was higher among inner-core residents (5/100,000), compared to inner-core residents (2.3/100,000). Regardless, other subtypes had higher prevalences than ST-3671 in inner-core populations, which may be an indication of ST-3671 “escaping” inner-core populations. Further research is needed to verify this had indeed occurred.

Mathematical modeling studies have shown that compared to disassortative or neutrally-mixed networks, assortative networks can more rapidly form highly-connected structures, such as ‘large components’; these structures, in turn, facilitate percolation of pathogens to other parts of the connected component. Moreover, Newman et al. have shown that although large components form more rapidly when mixing is assortative, the size of these large components tend to be smaller, when compared to those created through other types of mixing. Thus, the authors illustrate how relatively small, but highly-connected structures can act to sustain transmission of pathogens, despite low average density in the overall network. At the same time, connectivity in these highly-connected structures is robust to node removal. Therefore, the topological properties of highly-connected structures offer pathogens the best of both worlds: an efficient landscape for transmission, while also ensuring connectivity is maintained in the face of any individual nodes’ removal. For our purposes, it then follows that assortative networks, comprised of “higher-risk” individuals would then offer the best chance of transmission success for NG, at least during the initial growth phase. Thus, network-level features, which emerge out of, but yet are independent of individual behaviour, could help explain the rapid spread of ST-9663, as well as the association between ST-3671 and CT co-infection.
That ST-3671 and ST-9663 were both found in higher-risk groups aligns with network theory; what is not apparent from our data is how one feature (assortative mixing) can result in (at least) two different transmission patterns. That is, if a network-level feature such as high connectivity is fundamental to the spread of certain pathogens, what accounts for the differential spread between subtypes of the same pathogen, and what accounts for variations in secular trends over time? Although the answer to these questions is beyond the scope of our analyses and data, the literature suggests that once a pathogen is introduced into a susceptible population, non-linear, complex dynamics ultimately impact the course of an epidemic.\textsuperscript{157,257,258} Although high connectivity ensures a higher probability of ongoing transmission, and given the initial introduction itself is a random process, interactions between various determinants, post-introduction, such as host factors (e.g., the individual composition of the nodes in the network and host response to the infection); variations in pathogen characteristics (e.g., resistance, infectivity and virulence); behavioural factors; and the impact of control programs, make the ultimate course of the initial infection inherently stochastic, and highly unpredictable.\textsuperscript{157,259-262} In terms of secular trends, mathematical tools, such as the logistical map and bifurcation diagrams, have shown that system behaviours are sensitive to parameter values,\textsuperscript{257} while Yorke et al. have shown that changes in the pathogen, or in “socio-sexual” factors related to NG can rapidly impact both the “moving equilibrium” state, and the incidence of NG,\textsuperscript{157} a finding which has been echoed by other researchers.\textsuperscript{262} Thus, rapid increases in the incidence of NG are not uncommon, nor are they unexpected.
Phase-Specific Trajectories of Subtypes

Given the availability and timeliness of new genotypic technologies, one future area of empirical and theoretical work could be in applying Wasserheit and Aral’s phase-specific theory of epidemics to the trajectories of individual subtypes, or families of subtypes. The typical usage of the phase-specific theory has been in describing how epidemics evolve at an aggregate or population level. However, Blanchard has suggested changing the perspective towards the host population, as well as the pathogen itself. Shifting perspective to that of the pathogen, and understanding how phase-specific trajectories of different subtypes combine, either additively, or multiplicatively, to produce population-level incidence/prevalence is a worthy area of study. For example, during outbreak periods of NG, does one subtype (or family of subtypes), dominate? If so, focusing on identification and control of dominant subtypes through more intensive case and contact follow-up of specific transmission chains, could shape the planning and on-the-ground implementation of public health resources. Furthermore, knowledge of subtype characteristics (such as likelihood of being symptomatic and/or antimicrobial resistance) can help inform screening and treatment strategies. Conversely, if one subtype does not dominate, an intriguing avenue of study would be to determine how different subtypes competing in the same ecological niche act independently, yet synergistically, to impact population-level incidence.

That 51% of all NG infections in our sample were from just five subtypes suggests that certain subtypes would dominate in an outbreak situation. Given this reality, an important research step would be to systematically assess pathogen-side characteristics that create a competitive advantage for some subtypes over others. As alluded to in
Chapter 3, there exists very little systematically collected data on symptoms at the time of NG infection in the WRHA, and no systematized mechanisms to link these types of data with molecular data. If such types of linking were made available, a more standardized process of follow-up of cases to ascertain and document the presence (and absence) of symptoms by PHNs would help fill in some of the gaps in knowledge regarding symptoms, and whether or not a lack of symptoms indeed confers a competitive advantage. Alternatively, as more concentrated follow-up for every NG infection would be resource intensive, sentinel surveillance sites could be set up at key clinical locations to assess trends in symptom presentation over time. Here, documentation of treatment failure and successes could also be a complementary objective, in order to assess antimicrobial resistance. Knowledge of contacts, characteristics of contacts, and meeting venues would also make speculative links identified by molecular means more definitive. Having identified cases infected with dominant subtypes, perhaps a combination of more intensive follow-up of these cases and their contacts, alongside delivering innovative intervention strategies such as partner expedited therapy specifically to these transmission chains may result an alteration of epidemic trajectory. In this way, molecular data could complement surveillance data to identify and prioritize cases for more focused interventions.

Another application of this perspective is in understanding the role of bridging between one community and another, and how bridging sustains the transmission success of a given sub-type. Bridging between high prevalence populations has been shown to be associated with NG infections,\textsuperscript{33,217,263} as it has for other STBBIs, such as HIV.\textsuperscript{264,265} For example, in Alberta, Canada, De et al. demonstrated how a local venue acted as a hub for
NG transmission between communities.\textsuperscript{33} Using network analyses, Wylie et al. hypothesized that the occasional transmission and acquisition of NG to, and from, different communities/subpopulations, and thus avoidance of “burning out” locally, was integral for maintaining NG endemicity within the province of Manitoba.\textsuperscript{217} From the pathogen’s perspective, a highly-connected large component may offer the best chance for transmission success during its initial growth phase, while (uni- and bi-directional) bridging populations to other communities, networks, or parts of networks may be essential for its endemic phase.

**Multiple Streams of Evidence**

We found that networks constructed from case-contact data very sparse, and not well-connected, and unlike previous studies conducted in Manitoba, lacking long transmission chains, and closed loops.\textsuperscript{95,217} This sparseness in networks has also been noted in a recent Canadian study, with the authors suggesting that molecular data may help reveal some of the hidden links between cases.\textsuperscript{233} The fragmented networks may be a result of fewer public health resources to follow-up cases and their contacts (as compared to previous study periods),\textsuperscript{95} a reticence by cases to divulge information, or it could be a reflection of changes in partnering behaviour, with NG cases having more contact with casual partners. The sparseness of the case-contact data, as measured by density, alongside the preponderance of asymmetric (i.e., uni-directional) dyads, and the lower than expected levels of in-degree centralization would suggest an underreporting of ties between individuals.

Irrespective of the underlying reasons, our results demonstrate that combining traditional case-contact interviews and molecular data has the potential uncovering of
some of the hidden connections between NG cases. When molecular data were included with case-contact data, much larger components were suggested, and cases were potentially linked to a broader geographic space in Winnipeg, than what would have been known through public health investigations alone. These results align with a study of routinely collected HIV surveillance data from New York City that combined contact tracing and molecular epidemiology to examine HIV transmission networks. The authors found that molecular data was a more reliable method of constructing transmission chains. Echoing earlier calls in the late 20th century for “second generation surveillance” with the goal of combining behavioural data with traditional surveillance data, perhaps the timing is suitable for incorporating molecular technologies with routinely collected behavioural and surveillance data.

Strengths and Limitations

A major strength of the present study was the combination of NG-MAST data with surveillance data, likely the first time this study has been reported in Winnipeg, Canada. Second, our data were population-based in that all reported NG infections were collected in a centralized location. This study also had several limitations, most notably the fact that not all cases had samples available for NG-MAST, thus biasing our results only to those who had an NG-MAST performed. Second, behavioural and partner data were available only for those cases who were contacted for follow-up. Third, without a definitive epidemiological connection, individuals belonging in the same subtype group could only be potentially linked to each other, as the same subtype could have been introduced at different times through individuals disconnected to each other. That is, molecular data by itself can only suggest linkages, and any potential linkage should be
validated through other epidemiological means. However, our strong assertion is that molecular data needs to be combined with case-contact investigations to maximize its utility.

**Conclusion**

In conclusion, our results demonstrate that incorporation of molecular data reveals connections between cases that were not apparent from traditional case-contact investigations alone, leading to many more cases being potentially linked together, and over a wider area of Winnipeg than previously demonstrated. Only a handful of subtypes were responsible for the majority of infections in our sample, which suggests certain subtypes have a competitive advantage over other subtypes, either at the pathogen level, or in their ability to access highly transmissible networks. Host-level characteristics may also confer a competitive advantage to some subtypes over others; for example, some subtypes may take advantage of a lack of herd immunity in certain populations/networks. Identification of the mechanisms behind why some subtypes proliferate over others, coupled with early identification of the most successful subtypes may help to alter the epidemic trajectory of NG.
Chapter 4 compared and contrasted two methods of constructing sexual networks, and applied basic social network metrics in order to describe observed networks. Through the use of conditional uniform graphs, Chapter 4 also introduced a simple method of hypothesis testing for network data. We were able to link the association that ST-3671 had with chlamydia and gonorrhea coinfection with our results from Chapter 3 – that those dually-infected with chlamydia and gonorrhea were likely to be younger, with a disproportionate burden borne among inner-core residents. Chapter 5 builds on the work of the previous chapters, and advances network analysis by exploring the use of an integrated statistical framework that uses separable temporal exponential random graph models to estimate dynamic networks over time, and which also allows the modeling of infectious diseases within networks. Network formation (and dissolution), and the transmission dynamics of different pathogens across networks can all be explored within this framework. We test how local formation rules impact network structure, and how network structure can then alter the course of an epidemic. The interface between pathogen biology and network topology is explored.
CHAPTER 5: THE ROLE OF ASSORTATIVE MIXING AND DURATION OF INFECTION ON EPIDEMIC POTENTIAL OF SEXUALLY TRANSMITED INFECTIONS: A SIMULATION STUDY OF NETWORK DYNAMICS

5.1. INTRODUCTION

Sexually transmitted infection (STI) epidemics arise as a result of complex interactions between pathogens, individuals, the socio-economic environment within which individuals and their partners are embedded, and the social and sexual networks that are a byproduct of their interactions in social and sexual space. At the pathogen level, however, the transmission success of an STI is fundamentally and uniquely contingent on the pathogen’s duration of infectivity being of sufficient length for the infected host to have unprotected sexual contact with at least one other susceptible sexual partner. From this foundational dyad-level interaction, the infection is transmitted along sexual partnership networks, with the complex configurations and structures of sexual networks implicated in the dynamics of STI epidemiology. Therefore, the duration of infectivity is an underlying factor in a pathogen’s reachable path, and it is the pathogen’s reachable path that ultimately is responsible for determining which portions of the overall network (and the populations comprising these sub-networks) the pathogen is exposed to.

Gonorrhea (NG) and chlamydia (CT) infections are the most widespread STIs globally, and are responsible for an enormous burden on public health resources. Although both are transmitted sexually, their epidemiologic profiles differ
Rates of CT are substantially higher than NG, and CT is typically more widespread, both geographically and socio-demographically.\textsuperscript{141,146,171,183} The differentials between the epidemiology of NG and CT are driven partially by differences at the pathogen level. Because the duration of infectivity for NG is shorter than for CT, due to biological characteristics and the interaction of cases with the medical system,\textsuperscript{275-277} it is thought that NG has evolved to exploit, and proliferate within sexual networks marked by high connectivity – characterized by a high number and rapid turnover of sexual partners (i.e., “super-spreaders”), a high prevalence of concurrent partnerships, or both.\textsuperscript{78,193,251-253,278-281} Moreover, studies have also shown how a relatively small “core group” of individuals can effectively sustain endemicity of NG infections within a local population.\textsuperscript{251,252}

In practice, because social partnerships are not random, but are defined and constrained by societal mechanisms, and because socio-economically marginalized populations have less social mobility,\textsuperscript{20,22,41,63,282-289} the populations at highest risk for NG infections are those who are socio-economically marginalized, relative to those at risk for CT infections.\textsuperscript{41,141,216,249} Therefore, patterns in the epidemiology of CT and NG infections are a consequence of specific interactions between these two pathogens and the ecological niches they have evolved to exploit; the ecological niches then manifest themselves as population characteristics that are then attached or associated with each pathogen.\textsuperscript{195,260} From this perspective, the epidemiologic profiles of CT and NG are significant not only for the purposes of risk quantification, but they can also offer clues for understanding underlying network structure, and the dynamics of transmission.
Mixing Patterns and Network Generation

In terms of understanding underlying network structure, mathematical modeling studies have demonstrated how the mixing patterns of individuals generate observed network structure.\textsuperscript{6,17,31,290} For example, compared to disassortative or neutrally-mixed networks, assortative mixing (i.e., “like with like”) can more rapidly form highly-connected structures, such as “large components”.\textsuperscript{252,272,276,291} Three features of large components formed by assortative mixing are crucial: first, and by definition, nodes within these structures are highly homogeneous; second, connectivity within these highly-connected structures is particularly robust to individual node removal; and third, the relative size of large components formed by assortative mixing tends to be smaller.\textsuperscript{252,253} Therefore, small, highly-connected sub-networks that retain connectivity even with the removal of individual nodes can act as a suitable ecological niche for the shorter duration of infectivity for NG. From this example, one can see how a generative view of network formation allows insight into how scientists can conceptualize social interactions, and how different pathogens then exploit the products of these interactions for evolutionary success.\textsuperscript{269} Moreover, certain features of sexual networks at the population level have been demonstrated to persist over time, despite a turnover in population at the individual level.\textsuperscript{95} Thus, the need to uncover the underlying generative processes which ultimately result in observable, stable and persistent structural forms has recently been echoed by several researchers\textsuperscript{5,6,17,98}. Knowledge of how partnership dynamics lead to global network structures, and how these global structures are then exploited by different pathogens could then inform strategic STI prevention and intervention initiatives.
**Assortativity and Heterogeneity**

Although assortative mixing is thought to be a key generative feature in networks where NG proliferates, bridging between high prevalence populations has also been shown to be crucial to the transmission success of NG infections. Using network analyses, Wylie et al. hypothesized that the occasional transmission and acquisition of NG to, and from, different communities/subpopulations was integral to NG endemicity within the province of Manitoba. This finding aligns with research demonstrating that individual variation and heterogeneity in number of sexual partners, and in disassortative mixing between partners are all associated with persistence of STIs. Thus, at the same time that assortative mixing creates the foundation for the proliferation of pathogens, a certain amount of heterogeneity also seems to drive STI epidemics.

Using an individual-based stochastic simulation model, the goal of the present study was to investigate the synergy between assortativity and heterogeneity. The hypothesis is that for STI epidemics that reach the endemic phase, a delicate balance exists between assortative and heterogeneous mixing. Finally, the impact that duration of infectivity has on the relative contributions of the two types of mixing will be investigated.

**5.2. METHODS**

**Simulation Design**

We used a stochastic simulation model to evaluate the impact of assortative mixing on network formation, and the interplay between duration of infectivity and levels of assortative mixing. The intent of our simulation was to create illustrative models, with the
goal being to understand key qualitative differences in epidemic potential as model parameters are varied, while others are kept constant, in order to isolate the effect of assortative mixing. Our main question was to understand the impact of varying assortative mixing on the epidemic potential of STIs. Additionally, in an attempt to investigate the observed differences between NG and CT, we also investigated how duration of infectivity interacted with varying levels of assortative mixing.

We used an integrated statistical framework based on separable temporal exponential random graph models (STERGMs), a type of exponential random graph model (ERGM) that can be utilized in modeling network dynamics over time and implemented in Statnet, a software package available for the R programming language. The Statnet software package allows the estimation of network models and the simulation of epidemics over these estimated networks. Detailed description of the theory and methods underlying Statnet are found elsewhere. Briefly, two formulas are required: a formation and a dissolution formula, in order to mimic the natural formation and dissolution processes seen in partnership behaviour. Models are solved with Markov Chain Monte Carlo methods. Features such as average degree (e.g., number of sexual partners), triangles, and mixing matrices can all be entered as parameters in the formation formula, while dissolution processes are more limited. For the purposes of the study, average momentary degree, the proportion of mixing between high- and low-risk groups, and the average momentary degree for high-risk groups are used as model parameters. Estimates from the statistical model are then used to generate networks, and infectious disease epidemics can then be simulated over estimated networks.

A heterosexual population of 1000 individuals over 10 years was simulated, and a
number of simplifying assumptions were made, in terms of mixing and pathogen parameters. The model did not include terms for births, deaths, or aging. For simplicity, two activity groups were created to denote high and low levels of sexual partnerships; assortativity was based on the level of partnering between these two activity groups. The high activity group comprised 10% of the population for both males and females. Formation parameters included average momentary degree for both high and low activity groups, and the level of assortativity between the two groups. Average momentary degree for low activity groups was set at 0.8, based on Morris et al., while average momentary degree for the high risk group was arbitrarily set to 50% higher, or 1.2. Note that a momentary mean degree greater than 1 includes the potential for concurrent relationships. Thus, “high activity” was defined as the potential for concurrent relationships, while “low activity” precluded concurrency, as concurrency has been shown to promote connectivity in sexual networks. Momentary mean degree was used due to the availability of published data; degree distribution could have also been used to parameterize models – however, there is a lack of published data on degree distributions and defining “high risk” vs. “low risk” groups would have been highly arbitrary. Assortativity was varied between 50% to 99%; this parameter controlled the degree of mixing between high and low activity groups. In terms of dissolution, average relationship duration length of low activity groups was set to 18 months, while average relationship duration length for high activity groups was set to 6 months, based on Robinson et al. An SIS (susceptible-infectious-susceptible) framework was used for simulating epidemics, with probability of transmission based on Kretzschmar. Probability of transmission was made the same for both males and females. In order to investigate the impact of duration of infectivity and to
simulate the differences between NG and CT, two duration lengths were used: 12 months for “CT” (or the longer-duration pathogen),\(^{296}\) while duration was arbitrarily set to 50% less for “NG” (or the shorter-duration pathogen) at 6 months. Each step of the model was measured in weeks; thus, for a 10-year period, 520 weeks were modeled. Ten infections were randomly distributed (5 in females and 5 in males) at Time 1. All parameters detailed in Table 5.1. It is important to note that since Statnet relies on multiple stochastic simulations to arrive at model solutions, parameter values are randomly varied at each model run; for example, although the average duration length of partnerships among members of the high-risk was set at 18 months, this value randomly varied, centred on 18 months, in each model run.

For each pathogen (i.e., the longer duration and shorter duration pathogens) assortativity was varied between 50% to 99% in 10% increments; therefore, for each pathogen, 6 assortativity scenarios were ran, for a total of 12 scenarios (2 pathogens X 6 assortativity scenarios). For each assortativity and pathogen scenario, results reported are averaged over a total of 100 runs; thus, 1,200 simulations were run in total. Cumulative incidence was calculated for each scenario, while secular trends in incidence and prevalence for each pathogen, under each scenario, were visualized. Networks were visualized and described, with a focus on growth and endemic phases. The following formulas were used in our models:

**Formation formula**

\[ Y = \text{edges} + \text{nodematch} + \text{nodefactor} \]
**Dissolution formula**

\[ Y = \text{offset}(edges) + \text{nodematch}(duration) \]

In the formation formula, “edges” denotes the target average number of ties at a specific time point (derived from the average momentary degree), “nodematch” is the term used in R to specify the amount of assortativity between risk groups, and “nodefactor” is the term used to specify how much higher the expected number of partnerships in high-risk groups ought to be (in our case, this was set to 50% higher than in low-risk groups – see Table 1). In the dissolution formula, “offset(edges)” is a term to control for the amount of partnerships in the model, while “nodematch(duration)” specifies different relationship lengths by risk groups.

**Table 5.1: Model Parameters for Network Models**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of nodes</td>
<td>1000</td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>500</td>
<td>n/a</td>
</tr>
<tr>
<td>Females</td>
<td>500</td>
<td>n/a</td>
</tr>
<tr>
<td>Years simulated</td>
<td>10 (520 weeks)</td>
<td>n/a</td>
</tr>
<tr>
<td>Proportion in high activity group</td>
<td>10%</td>
<td>n/a</td>
</tr>
<tr>
<td>Mean degree</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High activity group</td>
<td>1.2</td>
<td>50% higher than low activity group Morris, M. et al. 2009</td>
</tr>
<tr>
<td>Low activity group</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>Relationship duration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High activity group</td>
<td>6 months</td>
<td>Robinson, K. et al. 2012</td>
</tr>
<tr>
<td>Low activity group</td>
<td>18 months</td>
<td>Robinson, K. et al. 2012</td>
</tr>
<tr>
<td>Assortativity</td>
<td>50%-99%</td>
<td>n/a</td>
</tr>
<tr>
<td>Duration of infection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long duration</td>
<td>12 months</td>
<td>Geisler WM. 2010</td>
</tr>
<tr>
<td>Short duration</td>
<td>6 months</td>
<td>50% shorter than the long duration pathogen</td>
</tr>
</tbody>
</table>
5.3. RESULTS

Figures 5.1 and 5.2 show the prevalence of the longer and shorter duration pathogens, respectively, over the 10-year time period, as assortativity is increased. For the longer duration pathogen, prevalence remains at 50% of the population once equilibrium is reached for almost all levels of assortativity. Prevalence remains below 50% for the scenario where 99% of partnerships are within risk groups.

Figure 5.1: Prevalence of Long Duration Pathogen From Estimated Networks as Assortativity Between High- and Low-Activity Groups is Modified*  

*Averaged over 100 model runs per assortativity scenario; assortativity is measured as percentage of partnerships within each risk group. Prevalence is measured as number of infections.
For the shorter duration pathogen, increasing assortativity had a much more substantial impact on prevalence (Fig. 5.2). In the scenario where assortativity was at 50%, prevalence of the shorter duration pathogen was just under 30% once equilibrium was achieved. Prevalence declined linearly as a function of increased assortativity; that is, as assortativity increased, prevalence decreased. At the highest level of assortative mixing, prevalence was under 10%.

**Figure 5.2: Prevalence of Short Duration Pathogen From Estimated Networks as Assortativity Between High- and Low-Activity Groups is Modified**

![Graph showing prevalence over time with different levels of assortativity](image)

*Averaged over 100 model runs per assortativity scenario; assortativity is measured as percentage of partnerships within each risk group. Prevalence is measured as number of infections.*
Average cumulative incidence for both types of pathogens is shown in Figure 5.3. For the longer duration pathogen, cumulative incidence decreased from 10,748.8 cases at 50% assortativity to 9,500.9 cases at 99% assortativity, or a decrease of approximately 11% in cumulative number of cases over the 10-year period. For the shorter duration pathogen, the decrease was much steeper. At 50% assortativity, the cumulative incidence was 10,183, while at 99% assortativity, the cumulative incidence was 3823.1, for a decrease of 63% in cumulative number of cases.

**Figure 5.3: Average Cumulative Incidence of Simulated Epidemics of Long- and Short-Duration Pathogens Within Networks Estimated from Separable Temporal Exponential Random Graph Models as Assortativity Between High- and Low-Activity Groups is Modified**

*Averaged over 100 model runs per assortativity scenario; assortativity is measured as percentage of partnerships within each risk group. Incidence is measured as number of infections.*
Table 5.2 shows the overall prevalence rate (over the 10-year time period) as a function of assortativity, stratified by type of pathogen, and by risk group. For longer duration pathogens, and for both risk groups, there is a slight decrease in overall prevalence rate as assortativity increases. However, the rate ratio between high- and low-risk groups remains stable as assortativity increases, with the prevalence rate in the high-risk group approximately double that of the low risk group. For the shorter duration pathogen, increasing assortativity had a substantial impact on the low-risk group, with prevalence rates decreasing from 19.6 per 100 population to 3.2 per 100 as assortativity increased from 50% to 99%. The impact of increasing assortativity on the prevalence rates in the high-risk group was less linear, with higher prevalence rates observed when assortativity was at 50% and at 99%. However, the rate ratio between high- and low-risk groups widened as a function of increasing assortativity, from a three-fold difference, to just under a 20-fold difference in prevalence rates.

Table 5.2: Prevalence Rate (per 100 population) and Rate Ratio, by Risk Group and Assortativity Level

<table>
<thead>
<tr>
<th>Assortativity Level</th>
<th>Prevalence Rate, High Risk Group</th>
<th>Prevalence Rate, Low Risk Group</th>
<th>Rate Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Longer Duration Pathogen</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50%</td>
<td>83.6</td>
<td>44.7</td>
<td>1.87</td>
</tr>
<tr>
<td>80%</td>
<td>76.8</td>
<td>39.7</td>
<td>1.93</td>
</tr>
<tr>
<td>99%</td>
<td>74.1</td>
<td>39.6</td>
<td>1.87</td>
</tr>
<tr>
<td><strong>Shorter Duration Pathogen</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50%</td>
<td>58.5</td>
<td>19.6</td>
<td>2.9</td>
</tr>
<tr>
<td>80%</td>
<td>47.8</td>
<td>9.0</td>
<td>5.3</td>
</tr>
<tr>
<td>99%</td>
<td>56.3</td>
<td>3.2</td>
<td>18.1</td>
</tr>
</tbody>
</table>

* Averaged over 100 model runs per assortativity scenario; assortativity is measured as percentage of partnerships within each risk group
Figures 5.4 and 5.5 show the differential in prevalence rates between the two risk groups, at different levels of assortativity over time. Of note is the rapid growth phase for the shorter duration pathogen as assortativity reached 99% (Figure 5.4); a plateau in prevalence was reached at 39 weeks. In contrast, prevalence did not plateau until approximately week 80 when assortativity levels were at 50% and at 80%. In contrast, and for the longer duration pathogen, equilibrium levels were reached most rapidly at the 50% assortativity level for both high- and low-risk groups (Figure 5.5). The “tighter” clustering of the large component, and the presence of more closed loops, compared to networks formed by low assortative mixing (Figure 5.8), may have produced a network that was more efficient for transmission.

Figures 5.6 and 5.7 are visualizations of the sexual networks and the course of infection over time in the first 100 weeks of the simulation, for the longer and shorter duration pathogen, respectively. The large, well-connected component produced when mixing is at 50% assortativity can clearly be seen in both figures; of interest, with respect to the longer duration pathogen (Figure 5.6) is how well disseminated the infection spreads by time 100, even when assortativity is at its highest. In comparison, for the shorter duration pathogen, infection is mostly concentrated in the single large component by time 100 (Figure 5.7). Finally, Figure 5.8 shows the largest giant component at 50% assortativity and at 99% assortativity; the giant component for the latter is much smaller, has more closed loops, and is almost entirely composed of members of the high-risk group.
Figure 5.4: Prevalence Rate of Simulated Epidemics of Short-Duration Pathogens Within Networks Estimated from Separable Temporal Exponential Random Graph Models, by Risk Group and Assortativity Level*

*Averaged over 100 model runs per assortativity scenario; assortativity is measured as percentage of partnerships within each risk group.
Figure 5.5: Prevalence Rate of Simulated Epidemics of Long-Duration Pathogens Within Networks Estimated from Separable Temporal Exponential Random Graph Models, by Risk Group and Assortativity Level*

*Averaged over 100 model runs per assortativity scenario; assortativity is measured as percentage of partnerships within each risk group
Figure 5.6: Network Visualization Showing Infected and Non-Infected Nodes, Simulated Epidemic of Short Duration Pathogen Within Network Estimated from Separable Temporal Exponential Random Graph Models, at 50% Assortative Mixing.

*Time in months
Figure 5.7: Network Visualization Showing Infected and Non-Infected Nodes, Simulated Epidemic of Short Duration Pathogen Within Network Estimated from Separable Temporal Exponential Random Graph Models, at 99% Assortative Mixing*  

*Time in months
Figure 5.8: Largest Component in Network Estimated from Separable Temporal Exponential Random Graph Models at 2 Assortative Mixing Levels
5.4. DISCUSSION

We used stochastic simulation modeling based on an integrated statistical framework to assess the impact of assortative mixing on the epidemic potential of STIs. Our results demonstrated that the degree to which assortative mixing impacts epidemic potential is dependent on the duration of infectivity of the pathogen. Generally speaking, our models showed that higher levels of assortative mixing tended to mute STI epidemic potential (as measured by cumulative incidence), with the impact being more profound for pathogens with shorter infectivity periods. In effect, for pathogens with shorter infectivity periods, increased assortative mixing altered the epidemic potential from a generalizing epidemic to one that was more likely to be, using the terminology of Moses et al., “concentrated” or “truncated”.

Figure 5.9 illustrates the framework for describing epidemics, as envisaged by Moses et al; broadly speaking, as epidemics go through the incipient, growth, plateau and decline phases, they can assume three main types of trajectories: generalizing, concentrated and truncated. A generalizing epidemic is one where the reach of the pathogen is universal, and irrespective of risk group; a concentrated epidemic is one that is concentrated only within localized high risk groups; and a truncated epidemic is a mixture of the generalizing and concentrated epidemics, where the epidemic is concentrated in non-localized higher-risk groups, as well as their local partners.
Our findings align generally with STI theory, as Anderson and May observed early on that heterogeneity drives epidemics; they demonstrated that variability in number of partners had a direct correlation with epidemic size – the larger the variability, the larger the epidemic. Aside from mathematical modeling studies, numerous examples exist in the literature whereby bridging populations are thought to maintain and/or sustain STI epidemics, including men who have sex with men and women (MSMW), migrant workers, urban-to-rural travellers, tourists, and other types of multiplex relationships. The fact that heterogeneity fuels STI epidemics makes intuitive sense. Without a new susceptible pool, infections will die out; heterogeneity in partnership types introduces both a new population that can acquire STIs, as well as provide an avenue for infectious cases to “escape” for ongoing transmission. As Adams and Morris suggest, whether directly or indirectly, the levels of connection that an infected case has with other parts of a network ultimately determines the course of an epidemic.
Our main finding, that an increase in assortative mixing had a larger impact on shorter duration pathogens is also supported by STI theory. In order for NG to proliferate, a relatively well-connected network needs to exist, due to its shorter duration of infectivity.\(^\text{195}\) In our models, the high rate ratio between high- and low-risk groups existed not because the shorter duration pathogen had more success in the now highly assortative high-risk group; in fact, prevalence remained at almost the same level in the high-risk group, irrespective of the level of assortativity. Rather, the shorter duration pathogen “died out” in the low-risk group. Without constant re-introduction of NG into the low-risk group from members of the high-risk group, the prevalence of NG never has a chance of reaching the growth phase, as evidenced by Figure 5.4.

Figure 5.10 provides an example of the potential impact assortative mixing may have on epidemic potential. Two risk groups are represented in the figure – a “high-risk” group (in black) and a “low-risk” group (in blue). The first row illustrates the transmission pathway in a sexual network generated through low assortative mixing (i.e., a higher probability of mixing between the two risk groups), while the second row illustrates the transmission pathway in a sexual network created through high assortative mixing; each column represents a different time period, moving forward in time from left to right. In the low assortative mixing scenario (where high- and low-risk groups are free to form partnerships), the prevalence of the pathogen increases in the low risk from 1 to 3 to 5 cases – that is, the low risk group experiences a growth phase in the pathogen over this time period. Conversely, in the high assortative mixing scenario, where mixing is generally confined to within-group mixing (although not completely so), after initial introduction into the low-risk group from Person 2 -> Person A, the infection gradually
dies out, due to: 1) Person A not partnering with any new partners; and 2) no new infections are introduced from the high risk-group to the low-risk group.

An important point to note is that in the low assortative mixing scenario, since the pathogen “spreads” from the low-risk group to the high-risk group, this could be considered a “generalizing” epidemic. Conversely, in the high assortative mixing scenario, the infection/pathogen remains in the high risk group, and thus this trajectory could be considered a “concentrated” epidemic, as it is sustained in the high risk-network only. Also note that the prevalence of the pathogen does not change in the high-risk group in either of the assortative mixing scenarios, as was observed in our results.

**Implications for Public Health**

With this simple model we have illustrated how assortativity might impact network structure, and how that structure might then alter the course of an epidemic. Using only two parameters, our models reproduced network features and epidemiologic patterns that are consistent with empirical research. This includes smaller large component sizes, and the tighter clustering in the high-risk group under the high assortative mixing scenario, as well as the differentials in the prevalence of the longer and shorter duration pathogens, akin to empirical observations of the differences between CT and NG.

However, it should be emphasized that the intent of this study was for illustrative, and not predictive, purposes. To this end, more realistic models, calibrated to local empirical observations are needed. Nevertheless, our study demonstrates the potential for network models to understand transmission dynamics, and the key features of transmission that may impact the course of an epidemic. Because our models can be
parameterized with egocentrically-collected data, there is a clear opportunity for improved collection and provision of network-type data. In North America, since public health is tasked with the follow-up and management of STI cases and their contacts, the most logical step would be to engage with public health providers in order to envision how to collect network-oriented data. More robust and timely data on degree distribution of cases and their contacts, and including formal questions on concurrent relationships, as well as how populations mix would facilitate the development of more realistic models.

In terms of implications for interventions, our results suggest that a certain amount of heterogeneous mixing is needed for an infection to evolve from a localized to a generalized epidemic. This finding highlights the importance of identifying the links and bridges between different sub-populations, as they act to enervate transmission to other newly susceptible populations. At the same time, although finding the key individuals who act as links between different parts of a sexual network can be vitally important work, especially in the context of smaller outbreaks, or for those whose sequelae may cause serious morbidity, and be highly preventable, such as in the case of syphilis,\textsuperscript{153,303} a balance needs to be struck between chasing an “old” infection, and prevention of further ensuing infections. It may be just as vital to understand why linkages exist, and under what conditions and contexts are they maintained, as this may lead to more strategic use of resources to identify and target key populations, vs. chasing of individuals. Our findings are consistent with Morris et al.’s \textit{generative} perspective on sexual networks,\textsuperscript{6} in that local rules beget global structures. Recognition of global structures can lend insight into the local rules that created them, while understanding of
why local rules exist may then, in turn, provide a meaningful waypoint for more distal prevention schemes.

**Strengths and Limitations**

The major strength of our study was the use of an integrated statistical framework to both estimate complex networks and simulate sexually transmitted epidemics over time. The major limitation was that this was a simple model built for illustrative purposes. Given the illustrative nature of the study, no attempt was made to test for statistical significance. The size of the population modelled, as well as the number of experiments planned was limited by the processing power of a personal desktop computer. The prevalence found in our models was extremely high, and likely a function of the small size of networks. Larger and more complex network models would require a substantive upgrade in processing power, but would be important in the investigation of more realistic networks. Finally, aside from assortativity and a crude measure of risk (e.g., a low-risk group where average momentary degree was < 1, and a high-risk group where average momentary degree was >1), and infection duration, no other parameters were varied; future research should include sensitivity tests for all parameter values, as well as methods like Latin hypercube sampling to more rigorously test out potential values in the parameter space in order to have more confidence in the models’ findings. At the very least, parameters used to define average momentary degree, high vs. low risk groups, relationship duration and duration of infection should be systematically varied and assessed. Duration of infection was set to a minimum of 6 months, which might be considered too long in the face of established STBBI control programs in North America. However, durations of less than 6 months resulted in a substantial number of epidemics
never reaching equilibrium, or epidemics never being established. This was likely a result of the small network size modelled. It should be noted that a recent study of NG from the United Kingdom found a median time of transmission between cases of 3.4 months (with an upper bound of 8 months), so the parameter used in our models for duration of infection is not entirely unreasonable. Future work should endeavor to create larger networks, with variable infection durations, which would likely also result in prevalence/incidence that is more reflective of reality.

An important limitation to note is the possibility of achieving similar results had other parameters been varied; thus, our results should not be interpreted as being conclusive. Modeling merely has suggested a possible relationship between assortativity and duration of infection. Rigorous testing of all parameters and the cataloging of the impact of their variation, on epidemic trajectories would lend more weight to our conclusions. However, it should be noted that few examples exist of known sexual network structures, especially those that are longitudinal in time. Even if many examples existed, the generalizability of networks from one location to another is unknown. Toy models such as the one constructed in this study, by necessarily simplifying relationships, parameters and assumptions, may further more fundamental theoretical knowledge development. Average momentary degree was used in network construction in the present analyses – this decision was made (vs. using degree distribution) as there is a dearth of published data on degree distribution, and no data at the local (i.e., WRHA) level. Moreover, as the goal was to assess the impact that varying the assortativity of high- and low-risk groups has on epidemic potential, use of the degree distribution became highly problematic when it came to actually defining what constituted a “high-
risk” vs. a “low-risk” group. For example, settling on the proportion of members of the high-risk group who had (e.g.) more than 3 partners, compared to the low-risk became a highly arbitrary exercise, without empirical evidence to guide the definition. For explicative purposes, then, average momentary degree was used, and although also arbitrary, the definition of high-risk groups having an average momentary degree 50% higher than low-risk groups had, at least, an intuitive feel.

A future area of work could include collection of important network parameters, such as degree distribution and relationship duration. It should be noted that the collection of such data through case-and-contact investigations would result in a bias towards positive cases; thus, average momentary degree captured through case-and-contact investigations may be artificially high, creating networks that are likely more dense than what would occur at the population level. Serial cross-sectional surveys of some representative sample of the sexually-active population would be a desirable research design, and would result in less bias.

It should also be noted that generally speaking, and controlling for contact rate, the probability of transmission for bloodborne infections is much higher, relative to sexual transmission. As well, other factors, such as the prevalence of injection equipment sharing, and the rate of partner change need also be considered in choosing parameter values for the transmission probability of bloodborne infections, and may impact local network structure; equipment sharing and high rates of partner exchange common with certain intravenous drugs (i.e., cocaine) can quickly amplify an epidemic amongst persons who inject drugs.
Conclusion

In conclusion, our results demonstrate that local rules privileging certain types of mixing over other types of mixing can impact global network topology. Network structure can then dictate the course of an epidemic. Understanding the socio-environmental context and pressures that individuals and populations face, and which then manifest themselves as partnership choices, could lead to well-tailored and more distally-oriented interventions that could more systematically disrupt the chain of transmission.
Figure 5.10: An Example of the Impact of Assortative Mixing on Epidemic Potential

Low risk individuals are those who have, on average, 1 or fewer partners at any one time.

High risk individuals are those who can average 1 or more partners at any one time.
CHAPTER 6: DISCUSSION, IMPLICATIONS, & CONCLUSIONS

6.1. SUMMARY OF RESEARCH FINDINGS

Taken together, the included studies affirm the important insights that are made available when context is incorporated into the analyses of public health data. The production of social diseases, of which STBBIs are a very exclusive form, is complex, inter-dependent and non-linear. However, persistent differentials between sub-populations suggest that the potential trajectories of STBBIs, and the populations they impact, are limited, and that the system that acts to sustain STBBIs, while maintaining distinct differentials in STBBI burden, is not irreducibly complex. The analyses presented in this dissertation offer methods and a perspective that can be used to unpack complex systems, and identify crucial features that may reveal key processes behind their production.

The analyses presented in Chapter 3 described the epidemiology of five distinct STBBI epidemics in Winnipeg, Canada, using a combination of traditional surveillance tools and place-based analyses. The analyses described how each epidemic had its own ecological niche – CT was widespread, both geographically and socio-demographically, while NG was more tightly clustered in space, especially in the inner-core. CT/NG co-infections had, overall, the narrowest distribution, in terms of space and age range, while HIV showed geographic clustering, but in areas outside of the inner-core. CT, NG, and CT/NG co-infections impacted mostly those under 30, while HIV was nearly non-existent in those under the age of 25. The distribution of syphilis cases did not show clear patterns across the study period, which was likely due to syphilis evolving from primarily impacting MSM to also including heterosexual/bisexual populations. NG was shown to
be undergoing an unprecedented growth phase, with clear signs of geographic dispersion. The expansion of NG infections seemed to be closely linked to CT over time and space. The results suggest that investigating the ecological niche of each epidemic reveals their underlying transmission dynamics; that NG in an outbreak situation seemed to follow CT closely in time and geographical space could lead to better targeting of resources for sub-populations vulnerable to NG infection.

The analyses in Chapter 4 compared and contrasted sexual networks constructed from case and contact investigations and those constructed from molecular genotyping. An analysis of networks constructed from case and contact investigations found very few linked components, with components mostly comprised of dyads and triads; incorporation of molecular genotyping revealed many potential links between cases and their contacts, demonstrating the benefits of having both types of tools available for public health investigations. Analyses revealed that the bulk of reported NG infections could be attributed to a few subtypes, suggesting that successful subtypes “crowd out” the field. The most common subtype was associated with CT/NG co-infections, which helped to partially explain its transmission success, as CT/NG co-infections have been shown to be associated with the youngest age groups, as demonstrated in Chapter 3. The interaction between pathogen characteristics and host populations, and the relative contributions of both, should be studied to understand why some subtypes proliferate over others. Early identification of the most successful subtype(s) may facilitate muting of the epidemic trajectory of NG.

Using mathematical modeling techniques, specifically, stochastic simulation models, the analysis presented in Chapter 5 demonstrated that local partnership rules
impacted global network structure, which then impacted the epidemic potential of a simulated pathogen. The impact of altering local partnering rules on the epidemiology of pathogens was dependent on the biological characteristics of the pathogen itself, demonstrating how intricately intertwined the biology of the pathogen and network topology are. Pathogens with shorter durations of infectivity were more sensitive to changes in network structure that were brought about by different levels of assortative mixing. The results suggest that understanding the socio-environmental pressures that shape partnership choices could lead to upstream structural interventions that could more systematically disrupt transmission chains. A limitation of this work was that host-pathogen characteristics were not modelled.

6.2. STUDY STRENGTHS AND UNIQUE CONTRIBUTIONS TO THE LITERATURE

The analyses contained in this dissertation help to advance the notion that context matters, when it comes to understanding the determinants of STBBI epidemics. The theoretical orientation of all analyses is explicitly generative – population-level patterns in network structure and the distribution of disease burden are ultimately determined by interactions at the local level. This theoretical orientation implies that empirical epidemiological observations are merely proxies for the underlying complex and inter-relational processes between society, individuals, and pathogens. A unique strength of the analyses is the explicit framing of applied public health practice within this generative framework, which has the potential to provide unique insights in the construction of novel prevention and intervention programs.
For example, secular trends of five distinct epidemics were examined, and placed in the context of their dynamic epidemiology and geographical foothold. Conventional epidemiological analyses typically only examine one or two pathogens, and typically do not address inequality in their distributions. Furthermore, the analysis in Chapter 3 is one of the first to confirm, through the use of inequality measures, the lessening inequality in NG as it experiences a growth phase, as predicted by STBBI theory. The shared geographical space that NG shared with CT suggests that within the most current outbreak, there is an overlap between the determinants of CT and NG. Therefore, knowledge of previous high incidence areas of CT may facilitate the judicious use of resources in the control of NG, which is unique in the sense that typically CT and NG are thought of as separate and distinct entities.

The analysis in Chapter 4 fills a gap in the literature of studies combining modern molecular genotyping techniques with traditional case and contact investigations in the investigation of NG infections. Our results demonstrated how molecular genotyping could reveal the potentially hidden connections between cases and their contacts that were not captured through public health investigations. Given the tremendous amount of time and resources allocated for case and contact investigation, future research, in the form of pilot projects would be a wise investment in terms of assessing utility and cost-effectiveness of the timely collection, analysis and dissemination of molecular data for NG. However, it would likely be some time before these technologies become widely available for routine public health usage, improvements in case-contact investigations can be simultaneously undertaken. Additionally, case-and-contact investigations are currently the only way partnering behaviours are captured, and provide a definitive validation of
links suggested by molecular data. Thus, the analysis in Chapter 4 could significantly contribute to the dialogue regarding what features should be kept from public health systems that rely on case-and-contact investigations, and how these features should evolve once molecular-typing technologies are routinely used, since Manitoba is one of the very few localities that offer centralized laboratory services that are linked to agencies that can perform NG-MAST.

The methods used in Chapter 5 demonstrated the potential of the powerful analytical paradigm offered by the integrated statistical framework used by Statnet for network modelling. Practical applications have so far been limited to research-based or academic institutions. The use of very simple models, and the insight they provided in Chapter 5 could potentially lay the foundation for a more rigorous approach to data collection, modelling and knowledge translation that could be applied to public health practice. Public health needs to nimble and responsive to constantly emerging threats; the ongoing challenge for public health is to make “right” decisions in the absence, or lack of empirical data. The counterfactual paradigm offered through mathematical modeling is an attractive way forward; and the insights offered in Chapter 5 from simple models may motivate more collaborations between public health and research/academic institutions.

6.3. IMPLICATIONS

The specific implications for each set of analyses are explicated in Chapters 3 through to 5. Presented here is a general overarching discussion that offers implications from the projects as a whole. The insights offered here can potentially help to re-frame and advance public health surveillance, in light of the lessons learned from incorporating context in the analysis of STBBI data.
Context-Based Surveillance Framework

Figure 6.1 is a schematic of traditional case-based surveillance data. Here, the burden of disease can be described in the context of person and time. Place, or geography, typically is incorporated to describe where cases reside, and thus can be seen as ancillary to the individual.

Figure 6.1: Schematic of Traditional Case-Based Surveillance

Figure 6.2 illustrates the use of traditional case-based surveillance in the comparison of different pathogens. Here, Pathogens A, B and C are meant to represent NG, infectious syphilis, and CT, respectively. The fundamentals of basic descriptive epidemiology can be seen in this schematic – describing the burden of disease by population characteristics, over time, and comparing and contrasting the disease burden across pathogens according to the characteristics of the population.
In Figure 6.3, a more holistic and context-based surveillance framework is proposed. Taking the lessons learned from the projects of this dissertation as a whole, the below schematic incorporates multiple levels of data. Included are data from the molecular, case-based, case-contact, place-based and socio-behavioural levels.

Figure 6.4 illustrates the generalizability of the context-based surveillance data when incorporating case-contact data. Without the case-contact layer, all cases would be described independent of each other. Incorporation of the case-contact layer reveals the linkages between Cases 1, 2, 3 & 4 and between Cases A, B & C; thus, two clusters containing Cases 1-4 in one, and Cases A-C in the other, are observed. In terms of the
implications for public health, since linkages are now known, treatment and follow-up of cases can be more focused, and case-finding efforts can start to be implemented.

**Figure 6.4: Schematic of Context-Based Surveillance: Incorporating Case-Based and Case-Contact Data**

![Diagram](image)

Figure 6.5 illustrates the incorporation of the place-based layer to the same data from Figure 6.4. Here, it can be shown that not only are there direct linkages between Cases 1-4 and Cases A-C, but each cluster is also located in a specific geographic location. Again, from the public health perspective, case-contact and place-based information can form the basis of evidence-informed allocation of resources, as well as reveal potential important key populations and linkages. In sum, with each additional layer, more insight is gained, and as more insight is gained, prevention and control efforts have more intelligence to inform program decision-making.
Figure 6.5: Schematic of Context-Based Surveillance: Incorporating Case-Based, Case-Contact, and Place-Based Data

Figure 6.6 demonstrates the incorporation of the molecular level of data to case-based, case-contact and place-based data. Here, although no members of Cluster 1 have named contacts in Cluster 2, and vice-versa, molecular epidemiology reveals the linkages between the two clusters. Thus, it can be surmised that the two clusters are, in actuality, part of the same outbreak. Note that this scenario is similar to what was observed in Chapter 4, in terms of potential linkages. Figure 6.7 illustrates the incorporation of all proposed layers of the context-based surveillance framework. Here, the addition of the socio-behavioural layer adds even more context to the outbreak. For example, the underlying socio-behavioural layer could be attendance at the same music festival, a drug-sharing circle, or gang affiliation.
Cases X and Y were added to Figure 6.7 to elucidate how the synthesis of knowledge from all layers may potentially be used to identify at-risk or vulnerable cases before they become infected, or to identify infected cases who do not know their infection status or have not yet sought care. Since X and Y form part of the same socio-behavioural demographic of the infectious cases, it may be practical to contact Cases X and Y, either to pre-emptively offer protective technologies, like condoms, PrEP, partner expedited therapy, and/or to offer screening services.
Application of the Context-Based Surveillance Framework

This section contains some practical applications of the context-based surveillance framework to the analyses contained in Chapters 3-5. The intent here is to demonstrate how having a more holistic surveillance system which incorporates multiple layers of data can help fill in the gaps in knowledge about local STBBI epidemics, and how these multiple lines of evidence can be integrated into a single framework.

Figure 6.8 illustrates how molecular epidemiology allowed potential linkages to be made between isolates, dyads and triads who were not linked together through case and contact data. This figure is meant to reflect the data captured in Chapter 4; when molecular data were added, multiple geographies and several more individuals and their contacts could be potentially linked together as part of a transmission chain.
Figures 6.9 and 6.10 illustrate the co-infection scenario from Chapters 3 and 4, and shows one potential explanation for the association between CT/NG co-infection and specific sub-populations (as observed in Chapter 3), and a specific NG subtype and CT/NG co-infection (as observed in Chapter 4). Figure 6.9 shows a possible scenario where NG co-infection is successfully established within a sexual network. Here, two hallmarks exist: all cases are linked socio-behaviourally; and there are not long gaps between cases being infected. NG (blue dashed line) is transmitted quickly as there is a rapid turnover in partnerships; concurrency is explicitly captured here in the trio of cases that are stacked on top of one another. Note that CT does not require as many partner exchanges to persist in the illustration. In this scenario, CT/NG co-infection is a product of CT travelling along the same networks as NG, with the specific network configuration being amenable for both the propagation of CT and NG. A biological interaction may exist (where infection with one pathogen increases the infectivity of the other at the biological level), but this scenario is not explicitly captured here.
Figure 6.9: Schematic of Context-Based Surveillance: Co-infection Success

Figure 6.10 demonstrates a scenario where NG fails to be established, and thus, CT/NG dual infection does not persist. The concurrent partnership has been eliminated, and there is now a large gap in time between the 3rd and 4th case. As can be seen, NG is successfully propagated up to the 3rd case, at which point a large gap appears before any further partner exchange, at which point the duration of infectivity for NG has passed. Note that CT can still be transmitted within this transmission chain. This scenario could explain the association that ST-3671 had with CT/NG co-infection; ST-3671 was the most common subtype, and therefore likely accessed sexual networks that were most efficient at percolation of NG. In networks that were inefficient at transmission, ST-3671 died out. So therefore, the success of ST-3671 is somewhat tautological – ST-3671 was the most common subtype because it was the most successfully transmitted. Likely, ST-3671 made inroads into both highly-connected networks, and those marked by lower connectivity. In highly-connected networks (which by definition are also efficient at transmitting CT), NG flourishes alongside CT. In low-connectivity networks, NG is less likely to persist.
Figures 6.11 and 6.12 illustrate the transmission dynamics of a pathogen in networks that are formed at low- and high-levels of assortativity, reflecting observations from the simulation models in Chapter 5. The scenarios illustrated here are similar to that of the co-infection scenarios. In Figures 6.11 and 6.12, the black lines denote membership in the high-risk group (i.e., where concurrency is a possibility), while the solid blue lines denote membership in the low-risk group. Figure 6.10 presents a scenario where there is heterogeneous mixing - members of the high-risk group are as likely to form partnerships with their own group members as they are with low-risk group members. Here, the presence of links to high-risk group members maintain circulation of the pathogen within the low-risk group.
Figure 6.11: Schematic of Context-Based Surveillance: Low Level of Assortative Mixing

Figure 6.12 shows how transmission in the low-risk group may be muted in the presence of highly assortative mixing. Here, after the initial infection from a high-risk group member, the infection dies out as there is no secondary infection. The low-partnering rate shields other members of the low-risk group from acquiring infections, while homogeneous mixing ensures that re-seeding of the pathogen in the low-risk group by high-risk group members is unlikely.

Figure 6.12: Schematic of Context-Based Surveillance: High Level of Assortative Mixing
6.4. CONCLUSION

The results from the studies included in this dissertation demonstrate the critical insights gained when context is used as a lens to interrogate public health data on STBBIs. Pathogens have their own ecological niche; the interactions between pathogens, individuals, and their risk environment can be uncovered through examination of epidemiological trends and patterns. Knowledge of pathogen characteristics, mixing patterns of individuals and populations, and the network topology that arises out of mixing patterns can be integrated to form the basis of a more strategic public health approach. A potential way in which knowledge of STBBI epidemics from multiple levels can be integrated into a single cohesive framework – from the socio-behavioural, through to the geographical, and relationship levels – is presented.

The projects in this dissertation were motivated by a central paradox: why patterns in the epidemiology of STBBIs seem to persist and sustain themselves, despite being reliant on the most intimate form of social contact, which itself is driven by locally-unique rules and customs. Our results suggest that to understand the trajectory of infections is to understand the interaction between specific pathogens, the socio-spatial space where they may or may not proliferate, and the pressures that shape and govern rules of interaction.
BIBLIOGRAPHY


145


129. Khanna A, Goodreau SM, Wohlfeiler D, Daar E, Little S, Gorbach PM. Individualized diagnosis interventions can add significant effectiveness in reducing


179. Jolliffe D, Krushelnytskyy B. INEQERR: Stata module to calculate bootstrap standard errors for indices of inequality. Center for Economic Research and Graduate Education.

180. Pisati M. spmap: Stata module to visualize spatial data. 2007.


196. Doherty IA, Shiboski S, Ellen JM, Adimora AA, Padian NS. Sexual bridging socially and over time: a simulation model exploring the relative effects of mixing and concurrency on viral sexually transmitted infection transmission. Sexually transmitted diseases 2006;33:368-73.


214. Xia M, Whittington WL, Shafer WM, Holmes KK. Gonorrhea among men who have sex with men: outbreak caused by a single genotype of erythromycin-resistant


280. Leung KY, Kretzschmar M. Concurrency can drive an HIV epidemic by moving R0 across the epidemic threshold. AIDS 2015;29:1097-103.


## APPENDIX 1: SOCIAL NETWORK BASIC METRICS CALCULATION

<table>
<thead>
<tr>
<th></th>
<th>Individual-Level</th>
<th>Graph/Sub-Graph-Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Crude</td>
<td>Normalized</td>
</tr>
<tr>
<td>Degree Centrality</td>
<td>$\sum \frac{d_i}{N-1}$</td>
<td>$\sum \frac{d_i}{N-1}$</td>
</tr>
<tr>
<td></td>
<td>$\sum d_i / (N-1)$</td>
<td>$\sum d_i / (N-1)$</td>
</tr>
<tr>
<td>Closeness Centrality</td>
<td>$\frac{1}{\sum D_{ij}}$</td>
<td>$\frac{N-1}{\sum D_{ij}}$</td>
</tr>
<tr>
<td></td>
<td>$\frac{1}{\sum D_{ij}}$</td>
<td>$\frac{N-1}{\sum D_{ij}}$</td>
</tr>
<tr>
<td>Betweenness Centrality</td>
<td>$\frac{g_{ij}P_k}{g_{ij}}$</td>
<td>$\frac{g_{ij}P_k}{g_{ij}}$</td>
</tr>
<tr>
<td>Maximum number of ties</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Density</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Where \( N \) = number of nodes and \( L \) = Number of ties

APPENDIX 2: DERIVATION OF GINI INDEX

Derivation of Gini Coefficient:

Point estimates and 95% confidence intervals were calculated using the *ineqerr* program in Stata. The *ineqerr* program uses the following formula to calculate the Gini coefficient, based on Jenkins\textsuperscript{305}

\[ G = 1 + \left( \frac{1}{N} \right) - \left( \frac{2}{mN^2} \right) \sum_{i=1}^{n} (N = -i + 1)y_i \]

where individuals are ranked in ascending order of \( y_i \), and \( m \) is the mean value of interest. The *ineqerr* program uses Stata’s native *bootstrap* program to estimate 95% confidence intervals.
APPENDIX 3: TRIAD CENSUS


Accessed: April 27, 2017