How Should Network Theory Inform Population-Based Studies of HIV Transmission?

By

Scott K. Heysell

Presented to

The Department of Public Health and Preventive Medicine

Oregon Health and Science University

In partial fulfillment of the requirements of the degree of
Master of Public Health

March 31, 2004

Department of Public Health and Preventive Medicine
School of Medicine
Oregon Health and Science University

Certificate of Approval

This is to certify that the M.P.H. thesis of

Scott K. Heysell

has been approved.

[redacted]
Thesis Chair

[redacted]
Thesis Committee Member

[redacted]
Thesis Committee Member

III. TABLE OF CONTENTS:

IV. List of Figures and Tables:

V. ACKNOWLEDGEMENTS:

I sincerely thank my committee, Dr. Dongseok Choi, Dr. Mark Loveless and Dr. Todd Korthuis, for the guidance and expertise that allowed for such an enriching experience. I must also recognize the contributions of the staff of the Oregon Department of Human Services, HIV/TB/STD Section that assembled the initial dataset and offered me the opportunity for further exploration. Lastly, I had the great fortune to work with my fellow MD/MPH classmates, Tania Thomas and Jill Gelow, for five years and I consider them not only outstanding colleagues but more importantly, the dearest of friends.

VI. ABSTRACT:

The traditional approach to the epidemiology of sexually transmitted infections (STIs) has focused on an individual's behavioral traits. Likewise, models of case detection have followed the paradigm that an infected individual increases the risk of directly transmitting the disease to one's sexual partners and consequently only sexual partners should be investigated. More recently, however, a methodology that recognizes the individual as occupying a larger web-like network of social relationships has emerged. The network-based approach recognizes that the risk of infection, and similarly, the risk of transmission, depend not only upon an individual's risk factors, but perhaps more importantly, on an individual's sexual or social partners' risk factors as well. Numerous field investigations have recently demonstrated the utility of describing a population based on network parameters, however, the advantages of a network-based approach have not been studied in any larger collection of surveillance data.

Employing data that were collected from the Oregon HIV-test reporting form during the three-year period beginning 9/1/98 and ending 8/30/01 and that included behavioral patterns of MSM (men who have sex with men) and IDU (injection drug users), the following were accomplished: 1) identified the population at risk of acquiring HIV based on a traditional risk factor profile using multiple logistic regression; 2) identified the population at risk of acquiring HIV using novel network-based risk factors and compared the traditional and network-based models using best subset combinations; 3) provided a broad description of the population with respect to network-based parameters previously unexplored in surveillance data- degree of closed v. open population, concurrency, absolute and relative size of core group, core group risk

behavior, sexual/needle exchange interaction among core group and non-core or periphery individuals, and bridges among subpopulations; and 4) described the specific benefits and limitations of applying a network model as a surveillance tool.

Upon completion of the analysis, network theory showed great promise in being able to complement traditional epidemiological approaches to surveillance studies of HIV transmission. Future survey design should seek to incorporate questions that reflect network-based principles such as concurrency, duration, centrality, and social cohesion. Network-based explanations also proved useful in framing descriptive results but may best be used in populations where subjects share geographic and ethnographic proximity. In particular, the analysis suggests that network-based models may be of most use in communities with higher HIV prevalence and where HIV transmission involves a more diverse group of subpopulations with higher rates of heterosexual activity and injection drug use.

TITLE:

How should network theory inform population-based studies of HIV transmission?

RESEARCH QUESTION:

Does network-based analysis provide a more accurate and refined understanding of disease transmission in a population at risk of HIV/AIDS compared to traditional analysis based on personal risk factors alone?

SPECIFIC AIMS:

Employing data that were collected from the Oregon HIV-test reporting form during the three-year period beginning 9/1/98 and ending 8/30/01 and that included behavioral patterns of MSM (men who have sex with men) and IDU (injection drug users), the following specific aims were accomplished:

1. Identify the population at risk of acquiring HIV using a traditional analysis based on personal risk factor profile.

2. Derive a model that identifies the population at risk of acquiring HIV using novel network-based risk factors and compare the traditional and network-based models.

3. Provide a broad description of the population with respect to network-based parameters previously unexplored in surveillance data- degree of closed v. open population, general population composition, concurrency, absolute and

relative size of core group, core group risk behavior, sexual/needle exchange interaction among core group and non-core or periphery individuals, and examining bridges among subpopulations.

4. Describe the specific benefits and limitations of applying a network model as a surveillance tool.

BACKGROUND and SIGNIFICANCE:

The Origins of the Network Model

The traditional approach to the epidemiology of sexually transmitted infections (STIs) has focused on an individual's behavioral traits (1,2,3). Likewise, models of case detection have followed the paradigm that an infected individual increases the risk of directly transmitting the disease to one's sexual partners and consequently only sexual partners should be investigated. Richard Rothenberg, a leading proponent and researcher of network theory, describes the roots of the traditional, linear, approach as taking hold after the Surgeon General of the United States in 1937 published a text that called national attention to syphilis, entitled *Shadow on the Land*, and inspired the program of venereal disease control and reporting that the burgeoning public health service utilized for the next 40 years (4,5). Conceptual illustrations, such as the one offered below by Rothenberg, became commonplace and were applied to track many different STIs.

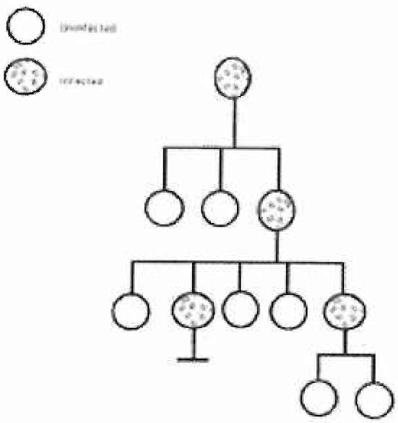


Fig. 1. Syphilis case series, with linear transmission as the underlying concept (4). Empty circles= unifected. Speckled circles= infected.

Yet, the traditional framework fails to account for the interconnections among individuals, and focuses solely on one person as the 'source' of infection.

More recently, however, a methodology that recognizes the individual as occupying a larger web-like network of social relationships has emerged (6,7,8). The network-based approach recognizes that the risk of infection, and similarly, the risk of transmission, depend not only upon an individual's risk factors, but perhaps more importantly, on an individual's sexual or social partners' risk factors as well. Infections from pathogens such as human immunodeficiency virus (HIV) and hepatitis C (HCV) that can occur through needle-sharing or other non-sexual behaviors present an added layer of complexity that a network-based model more accurately elucidates. The traditional framework has drawn criticism because of the inability to adequately predict an individual's risk of infection based on a multiplicative number of behavioral factors and a failure to account for a core group of individuals with high-risk behavior and linked by a sexual and social network that has since been described as essential for understanding the persistence of STIs in populations with otherwise ample access to treatment (1,6,9).

In contrast, a network-based approach requires an attention to the ethnographic context within which the population of interest operates and seeks to describe the weight and influence of relationships within the population. Network theory attempts to resolve the discrepancy of how the same behavior may produce opposing outcomes in different settings: the patient who seems "to get everything" and the patient who, despite possessing a multitude of risk factors, remains the image of health. A discussion of the origins of network theory will provide a useful and contextual background for explaining the specific tenets of a network-based approach to studying the transmission of HIV and other STIs.

Albert-Laszlo Barbasi has written a brilliant and comprehensive review of network science, *Linked*, that details both the historical advancements and implications of networks for future discoveries in varied fields such as economics, internet technology, and cancer research (10). To introduce the origins of network theory in *Linked*, Barbasi relates the story of the acclaimed eighteenth-century mathematician, Leonhard Euler. Among his many accomplishments, Euler had written a short paragraph that solved an entertaining problem that had been troubling the residents of Königsberg, a small town in eastern Prussia, near St. Petersburg. The Pregel River split into a Y that worked its way through the center of town, creating four distinct bodies of land, a north and a south bank along with two islands. Seven bridges connected the bodies of land and the people of Königsberg had long-wondered if it was possible for one to walk across the seven bridges and never cross the same bridge twice.

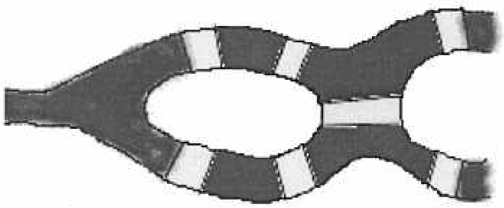


Fig. 2. The Bridges of Königsberg, pre-1875 (11). Yellow= bridges. White= land. Blue= water.

Euler created an elegant but detailed proof stating that not only did such a path *not* exist, but that a similar path cannot exist in any situation with more than two bodies of land and an odd number of bridges. In elucidating the proof Euler used the language of 'links' and 'nodes' to describe the bridges and bodies of land respectively, maintained that the path was an inherent property of the graphical relationship, and unwittingly gave birth to an entire field of graph theory that serves as the fundamental beginning for the contemporary understanding of networks. Small changes in the structure of the network, such as the addition of one more bridge in the town of Königsberg, allow the possibility of entirely different pathways to emerge.

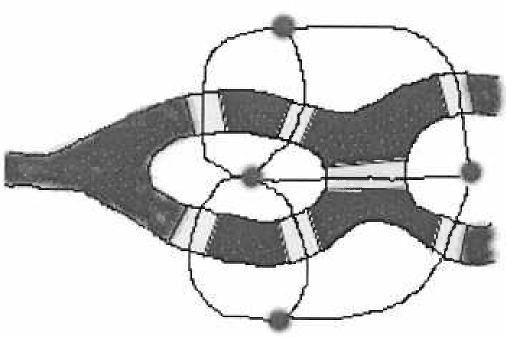


Fig. 3. Euler's graph, with nodes (red) and links (black lines) (11).

Yet, intuitively one notices that all links and nodes do not behave alike and clearly certain nodes are more influential to the overall network. The nodes with the most connections to other nodes have been labeled 'hubs.' Hubs exist within social networks, as many people can quickly think of a friend or acquaintance that possesses the rare gift of making numerous friends and associations and becoming truly *connected*, to borrow a colloquial descriptor that reflects scientific accuracy. Additionally, biologists have advanced that the p53 protein has the qualities of a hub in the network of cancer formation and ecologists assert that certain species act as hubs in the food web and whose existence consequently maintains an entire ecosystem's viability (10). Similarly, a single man has somewhat infamously been referred to as 'patient zero' in the AIDS epidemic, not because he was the first to be diagnosed with the disease, but because 40 of the 248 original people diagnosed with the mysterious illness by April of 1982 had either had sexual intercourse with 'patient zero' or with someone else who had (10). Functioning as a hub in the network of men having sex with men (MSM) in San Francisco, 'patient zero' estimated having an average of 250 sexual partners a year, for a decade, while also working as an international flight attendant. 'Patient zero' may not have been the one to bring the virus to the United States, but his central role in the beginnings of the epidemic are emblematic of the influential properties of hubs, or what now are commonly referred to as core groups within the literature of STIs.

Networks and Sexually Transmitted Infections

The core group is regularly defined as the individuals within a population that have a large number of sexual contacts thereby possessing the potential to infect the greatest number of people if infected themselves (3,6). By nature, the persistence of any STI requires a core group. Yet, while the core group is crucial in the maintenance of an infection within a population, the spread of an infection to distinct populations necessitates the recognition of the extent to which individuals in the core group are engaged in high-risk behavior with members of other subpopulations. Identifying such 'weak links' becomes imperative when targeted measures of prevention are discussed.

Other unique tenets of a network-based approach to STIs have been expressed in similar ways by a number of authors, but Rothenberg may arguably have been the first to outline the approach in a coherent and durable fashion (4). Rothenberg asserts that a network-based approach should ideally describe the population of interest with respect to the following properties: 1) graph connections; the frequency, distance and quality of relationships among people 2) centrality; a measure of an individual's prestige, power, or influence within the network 3) social cohesion; the type and variety of connections among subgroups within the population 4) structural equivalence; individuals with similar positional or functional relationships in the network and 5) network characteristics; overall features of the graph that defines the network.

When investigators have the ability to interview individuals on a personal level and assess who is actually sharing risk with whom, such as in an outbreak investigation, mathematical or quantitative measures of network structure can be generated (4,6,12,13). For instance, Ghani et al. describe a measure of local centrality such that at distance 1, a person's local centrality is the person's number of partners, and at distance 2, the local centrality is the sum of the person's partners' number of partners (12). *Closeness*, or global centrality, measures the shortest distance of a person to all other individuals in the network and varies from 0, when the individual is completely isolated, to 1, when the individual is directly connected to all members of the network (12). *Betweenness* refers to the number of shortest paths that an individual connects other network members, and, as a further measure of centrality, may represent the factor most influential on transmission (12). Lastly, information centrality, distinct from betweenness by considering more than the shortest paths between individuals, measures the weighted combination of the paths from an individual to all other members of the network, and has been suggested to be of importance in both the acquisition and transmission of infection (12).

Certainly, more qualitative measures of the network structure exist depending upon the question researchers wish to ask of the population, but numerous field investigations have recently demonstrated the utility of describing a population based on network parameters (14,15,16,17,18). For example, investigators at the CDC were asked to join the Mississippi State Department of Health to research a cluster of HIV in a population of young adults in a rural community (17). Traditional contact investigation identified 5 HIV-positive individuals, but a social and sexual network of 122 other individuals was established that led to a further 7 individuals (9% of the 78 tested from the social network) to be identified as HIV-positive. The dynamics of the underlying network were clearly important in the direction of treatment and prevention strategies for the community, particularly when the resources were limited.

HIV Networks at the Population Level

The advantages of a network-based approach, while evident in the model of field investigation, have not been studied in any larger collection of surveillance data. The vital question becomes whether or not surveillance data, accumulated confidentially and thus without the knowledge of who is sharing risk with whom, can be examined with network parameters, and if so, does the network-tuned analysis describe a more accurate picture of the pattern of transmission and acquisition within the population? The implications of a network-tuned analysis include a more informed provision of preventative services to an 'at-risk' network of individuals instead of groups assembled based on a stereotypical label such as prostitutes, injection drug users, or gay men, as well as the ability to identify and focus risk-reducing strategies on individuals within the network that exhibit a high degree of centrality. Identification and treatment of infected individuals remains a cornerstone of any widespread campaign to decrease the incidence of STIs. However, with pathogens such as HIV, where infections can remain largely asymptomatic for years and social factors that may increase the risk of infection may also decrease the likelihood of prompt access to medical attention, a network-based approach to disease surveillance may be particularly effective.

The HIV pandemic presents an international health crisis without contemporary parallel. The developing world has suffered a disproportionate burden of the devastating illness, witnessing the disease disrupt every aspect of emotional, social, economic, and political well-being. In Africa alone, there does not exist any greater threat to the continent's vitality, as AIDS has left 34 million people dead and has orphaned 13 million children (19). Yet in the majority of industrialized nations, extensive efforts at prevention and availability of critical anti-retroviral medication has curbed the incidence of new infection and altered the course of disease. Indeed, with the relative success of anti-retroviral medication in areas with ample access to health care, particularly after the introduction of protease inhibitors, the infection has lost the dreaded moniker of "a death sentence" (20). Yet, the stable incidence of HIV in the United States hides the changing face of the illness. For instance, the rate of AIDS cases for women in the United States has steadily increased since 1986, and as of 2000, 63% of those women were black and 18% were Hispanic (21). The demographics reveal the alarming disparities in health care that continue to exist in the country, but also that the factors associated with acquisition of HIV may be better explained within the social network and include such aspects of women's health as personal or partner's drug use and associated unprotected sexual activity, STI co-infection and lack of recognition or treatment of the infection, and lack of empowerment to influence sexual activity (21). Issues of empowerment may manifest with respect to negotiation of condom use, exchange of sex for drugs, knowledge of partner's HIV status, socioeconomic status, or abuse.

Similarly, recent outbreaks of infections such as syphilis, chlamydia and gonorrhea in populations of MSM suggest that risk behavior may be changing within such networks, and it has been postulated that the outbreaks may be a harbinger of an increase in HIV to follow, as safe sex practices are relaxed under the perception of the efficacy of anti-retroviral therapy (22,23,24,26). Indeed, the documentation of the practice of 'barebacking' within subpopulations of MSM underscores the need for a network-tuned approach to surveillance and prevention efforts. Barebacking has been defined as intentional, unprotected anal intercourse, distinct in the absence of negotiable risk, and typically engaged in regardless of seropositivity among partners (24,25). Certainly a flashpoint for political discussion and polarizing viewpoints, many have argued that the practice reflects a desire for the reclamation of identity after the fear and change of behavior that the HIV epidemic had inspired. However, Mansergh et al. did not find any differences between MSM who had engaged in the practice within the last two years compared to those who had not based on common risk factors such as age, racial/ethnic background, sexual orientation, or drug use, but did find that significantly more HIV-positive men participated in the practice (24). Thus, such high-risk sexual conduct appears most suitably described in the context of the underlying network of relationships that dictates trends, behaviors, exchange of information and ultimately, risk.

However, contrary to the national trends, the state of Oregon has not witnessed the explosion of HIV in communities of color or among women. AIDS cases in Oregon reported in 2001 revealed that 87% were white, 93% were male, and 69% were MSM (26). The proportion of MSM and IDU has remained stable over the last five years, and general population rates of reported AIDS cases and AIDS deaths have plateaued since the precipitous decline in the mid-1990's following the introduction of protease inhibitors and combination anti-retroviral therapy. Yet, despite the consistently low proportion of women, the proportion AIDS cases in Oregon by race/ethnicity from 1981 to 2001 was 10% for African American women, or five times higher than would be predicted based on the demographics of the general population, perhaps suggesting that the state will soon approximate national trends (26). Additionally, on October 1, 2001 the state began the name-to-code HIV reporting process, a confidential method of disease reporting that within the following year identified 810 cases of HIV not originally reported as AIDS cases, and 30 new AIDS cases that were previously unreported (26). Preliminary analyses from these data in Multnomah County, containing the state's largest metropolitan base, do not find any trends that are inconsistent with the natural history of infection. Ultimately, theories of Oregon's relative success with battling the epidemic point to a progressive public health infrastructure, historically early needle exchange programs, and widespread access to healthcare for the state's low-income individuals under the Oregon Health Plan.

In order to keep the state on the leading edge of HIV prevention and curb trends that have occurred in other communities across the country, Oregon has consistently searched for new methods of understanding the disease within the state's population. Before the institution of Oregon's mandatory name-to-code reporting of HIV tests, the public-sector HIV test report form was required to be sent to the state's Department of Human Services from county health departments. During the period of 9/1/1998 to 8/30/2001, however, a series of specific and personal questions regarding risk were included in the HIV test reporting form to develop additional information on behavior and help to guide prevention efforts. The initial study will be discussed in detail in the methods and preliminary results sections. Yet, in brief, 65,980 HIV test reports were received during the period and among the reports, 298 indicated a new HIV-positive finding and 312 represented a repeat or confirmatory HIV-positive test.

The surveillance data that were collected in the study period present a wealth of information currently unavailable within such a large population. Thus, to apply the teachings of network theory to the data, beyond the realm of outbreak investigation or a controlled clinical environment, would not only be a unique venture, but if successful, would carry the potential of altering the administration of disease monitoring in a systematic fashion. Identifying a more precise screening and surveillance method offers direct benefit for the future of targeted prevention such as the eventual distribution of vaccine (9). Projections for the type of vaccine against HIV include a costly injection that must be administered every six months. If that method proves to be the case then a blanketed campaign of universal vaccination will likely not be the strategy. Instead, an appropriate estimation of the truly 'at-risk' population will need to be assessed because regular outreach and follow-up will be necessary for a larger group of individuals. In the current climate of limited public health resources and with the promise of a vaccine upon the horizon, a practical application of the network paradigm is nothing short of vital.

METHODS:

Oregon's Public Sector HIV-test Report Form

Acquisition of the dataset occurred through gracious consent of Dr. Mark Loveless and the Oregon Department of Human Services, HIV/STD/TB Program. Exploratory analysis of the data collected from Oregon HIV-test reports between 9/1/1998 and 8/30/2001 found ample information on ethnographic questions suitable for a network-based approach. Thus, after preliminary research and consultation suggested the validity of the project, a formal proposal was submitted to the state's internal Institutional Review Board and the study was approved for commencement.

The data were derived from the HIV-test reports submitted by publicly funded health departments from all 36 counties during the aforementioned study period. During that period HIV-tests ordered in the public sector were required to be sent to the Oregon Department of Human Services where the information contained therein was securely entered into a statistical software package (SPSS version 11.5). A total of 65,980 initial test reports were received.

The HIV-test reports included a basic panel of demographic questions including: 1) the date that the specimen was collected, 2) the type of publicly funded test site providing the test, 3) the laboratory where the test was completed, 3) whether the test was anonymous or confidential, 4) the type of specimen (blood, saliva, plasma, or dried blood), 5) residence information including state, county and zip code, 6) race/ethnicity (Asian, Black, Hispanic, Native American, White, Other), 6) education completed, 7) reason for test, and 8) previous test history.

Upon completion of the post-test counseling, test-administrators and subjects were encouraged to fill out an additional, voluntary page of 14 questions regarding behavioral risk factors. If a given subject reported IDU or MSM behavior, then attention to the appropriate items on an additional page of the questionnaire was further advised.

The 14 questions about behavioral risk factors:

1. Client's self-assessed risk for HIV/AIDS- known to be infected/high risk/low risk/no risk/unsure of risk
2. History of male-to-male sexual contact (MSM)- yes/no
3. History of injection drug use (IDU)- yes/no
4. Client trades sex for money or drugs- yes/no
5. Sex partner of known HIV positive individual- yes/no
6. Needle sharing contact with HIV positive individual- yes/no
7. Sex partner of person known to be at risk for HIV- yes/no
8. Other risk factors- written comments
9. None of the above risks- yes/no
10. Gender of sexual partners- male/female/both/none
11. Last sex with main partner- main partner/someone else/refused to answer
12. Number of different sex partners in the past 12 months- numerical value, 0+

13a. Vaginal intercourse- never/within last 6 months/more than six months ago/refused to answer

13b. Anal receptive intercourse- never/within last 6 months/more than six months ago/refused to answer

13c. Anal insertive intercourse- never/within last 6 months/more than six months ago/refused to answer

13d. Oral sex- never/within last 6 months/more than six months ago/refused to answer

13e. Condom use (answered separately for each type of sexual encounter listed in question 13)- used condom/did not use condom/do not know whether condom used

14a. Use of alcohol or drugs with sex- always/sometimes/never

14b. Drugs used- alcohol/marijuana/cocaine/crank/heroin/crack/other

The section of the HIV-test report that was addressed to men that answered 'yes' to question 2 of the risk factor section included the following questions:

1. In the past six months have you had sexual contact with another male?- yes/no

2. In the past six months have you had sex with someone you had known for less than a few hours?- yes/no

3. If someone were to ask you whether you were gay, bisexual or heterosexual, what would you say?- gay/bisexual/heterosexual/other

4. Privately, only to yourself, do you think of yourself as gay, bisexual, or heterosexual?- gay/bisexual/heterosexual/other

5. How many of your friends are gay males?- written estimation

6. How much do you care whether other people know you had sex?- written comment

7. Which of the following activities have you participated in (went to a gay social group or activity, went to a gay bar or club, went to a gay bath house, read a gay newspaper or periodical, participated in an AIDS organization, attended a gay or religious support group)?- yes/no

The section of the HIV-test report that was addressed to subjects that answered 'yes' to question 3 of the risk factor section included the following questions:

1. Have you ever shared needles?- yes/no

2. In the past six months have you used injection drugs?- yes/no

3. If shared needles, with whom did you share?- main sex partner/other sex partners/friends/people you do not know well

4. Reasons for sharing needle- written comments

Study Population

All subjects completing an HIV-test report in the state of Oregon during the study period were initially considered eligible for participation. Only forms that were submitted by publicly-funded HIV Counseling and Testing sites have risk information recorded, thus only subjects seeking services at such public sites were included. The following table demonstrates the number of HIV-test reports received for each major subgroup:

Table 1. Number of test reports within each subgroup

All reports	65,919 (100%)
All women*	32,417 (49.2%)
All men*	31,510 (47.8%)
MSM-IDU	1,063 (1.6%)
IDU (not MSM)	9,630 (14.6%)
MSM (not IDU)	6,254 (9.5%)

*gender not recorded for 1,991 reports

Since the HIV-test report forms were free of any information regarding individual's names or unique identification numbers, then over the course of three years many subjects tested on multiple occasions and completed more than one form. Thus, only subjects that indicated that they were completing an HIV-test report form questionnaire for the first time were included in the following analysis. When including only subject's initial tests and report forms, the following table results and the loss of frequent testers from the subgroups of MSM and IDU becomes apparent:

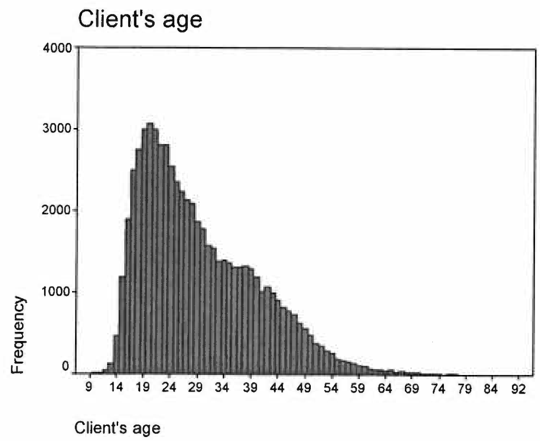
Table 2. Inclusion of only initial test report forms- number of tests within each subgroup

All reports	26,840 (100%)
All women*	13,549 (50.5%)
All men*	12,258 (45.67%)
MSM-IDU	168 (.62%)
IDU (not MSM)	2,126 (7.9%)
MSM (not IDU)	1,292 (4.8%)

*gender missing from 1,033 reports

The mean age of the subjects was 29.38 and the distribution ranged from a minimum of 0 years (neonatal) to a maximum of 92 years of age. Vertical transmission from mother to child has remained an extremely rare event in Oregon, as one such case has been reported during the 2001-2002 biennium. Similarly, the purposes of the current study seek to describe the network of transmission and acquisition of disease among a community with potentially modifiable risk factors and to whom the additional voluntary questions regarding risk factors would potentially pertain. Therefore, the distinction was made to exclude children between the ages of 0 and 9 from the analysis. 43 subjects were excluded, none of which were HIV-positive, and the 43 excluded subjects did not significantly differ from the remaining population with respect to other demographic characteristics. Figure 4 shows presents the histogram of age and demonstrates that the mean age remained relatively unchanged at 29.40:

Figure 4. Age distribution, after exclusion of subjects aged 0-9



No other identifiable characteristics were obtained other than what was listed in the description of test report questions, thus no further exclusion of subjects was made.

Study Design

Given that the data for each subject were reported at one point in time, the overall design best resembles a cross-sectional study. The primary outcome is binary, testing either HIV positive or negative. In the first section of the study, eligible subjects were compared by using multiple variables from the initial list of 14 questions regarding risk factors and pertinent demographic characteristics. A multiple logistic regression model was generated to reflect 'traditional' HIV risk factors from the list of variables and the model was tested using logistic regression. In the second section of the study a population of eligible subjects was identified by employing network-based parameters and the subsequent 'network' model was similarly tested using logistic regression and then compared to the traditional risk factors model. The third section of the study provides a general description of the population using network-based theory. The final discussion section of the study introduces potential improvements of modeling networks and further describes the limitations of such modeling techniques with surveillance data.

Laboratory Analysis

Determination of HIV status was made by specimen analysis using enzyme immunoassay (EIA) and Western Blot (WB) tests completed by the state-approved laboratory. Of the tests, 55,911 (84.8%) were performed on blood samples, 8,153 (12.4%) were performed on saliva samples, ten tests were completed with samples of plasma, and three with dried blood. Subjects with a history of IDU had oral fluid tests run slightly more often (24.0%) than the general population.

All negative EIA results were categorized as 'HIV-negative.' Positive or equivocal EIA results were confirmed with WB tests. Positive WB tests were categorized as 'HIV-positive' and negative WB tests as 'HIV-negative.' Equivocal WB tests were often followed with an additional immunofluorescent assay (IFA) test. Positive IFA results in the setting of a positive EIA and an equivocal WB were categorized as 'HIV-positive' and negative IFA results in the equivalent scenario were categorized as 'HIV-negative.' If an IFA test was not completed after an equivocal WB, then the subject was considered 'HIV-negative.' In a small subset of EIA-positive subjects, polymerase-chain reaction (PCR) was employed for certain internal and capsular proteins specific to HIV such as gp120, gp41, and p24, but the lack of widespread use of this technique precluded its use in determining serostatus for purposes of the study. In general, standard serologic assays show sensitivity and specificity rates of >99% (27,28,29,31). Likewise, saliva tests have demonstrated comparable results, and thus no identifiable reason exists to question the seropositivity outcomes reported (30). The outcome determination schema are summarized in the following table:

Table 3. Determination of HIV serostatus

EIA	WB	IFA	Serostatus
negative			'HIV-negative'
positive	negative		'HIV-negative'
positive	equivocal	negative	'HIV-negative'
positive	equivocal	positive	'HIV-positive'
positive	positive		'HIV-positive

Statistical Methods: Logistic Regression

For the variables included in the first portion of the study of traditional risk factors, it was necessary to select only individual characteristics, or those not representing dyadic or higher-order structures. The variables of 'gender,' 'race/ethnicity,' and 'age' are inherent individual traits and rarely absent from any biomedical model. The 'level of education' variable was recoded into a binary variable indicating an education completed as greater than or equal to achieving a high school diploma or GED, or less than such level. The 'reason for test' variable indicated whether or not a patient was seeking an HIV-test secondary to HIV-like symptoms. The following three variables, 'history of MSM contact,' 'history of IDU,' and 'history of exchanging sex for drugs or money,' though referring to the type of partner an individual has shared risk with, and therefore network-based in nature, the variables have long been associated with models for HIV transmission and were thus included in the first portion of the study. The variable, 'number of sexual partners in the last twelve months,' allows for the establishment of the categories of 'core' (6 or more partners), 'adjacent' (2-5 partners) and periphery (0-1 partners). Since the variables describing sexual activity and condom-use during sexual activity are often frequently included in any model of sexually-transmitted disease, univariate analysis of each covariate (anal receptive/anal insertive/vaginal/oral and condom use with activity) was completed and Pearson's correlations were analyzed to find if any variables could reasonably estimate one another and serve as proxy to reduce the number of variables in the model. The variables were collapsed together to produce, 'unprotected sex of any kind within the last six months,' and coded as binary. The model was explored for potential interaction terms from the main effects covariates using a clinically-informed approach of selection and then the change in deviance statistic was compared to the significance of a chi-square distribution with one degree of freedom at $\alpha<0.10$.

To assess the overall fit of the final regression model, the number of covariate patterns was determined. Once the number of patterns appeared to approximate the number of subjects, observed and fitted values were assessed for differences. The deviance of the non-aggregated data was used to estimate the final model's deviance. The

Hosmer and Lemeshow test was employed to assess goodness of fit. The Hosmer and Lemeshow test carries the null hypothesis that the model is a good fit, and classification tables serve as an adjunct in describing how well the model can distinguish a case from a non-case. As summary analysis, the change in Pearson's chi-square [$(res^2)/(1\text{-leverage})$], the analogue to Cook's distance, and the change in deviance [$(dev^2)/(1\text{-leverage})$] were all individually compared to the predicted probability (33). The analyses identify significant outliers that could then be examined for their overall influence on the model.

A similar logistic approach was used to analyze the network-based parameters. Variables included in this portion of the analysis referred to a higher order structure that described the subject's partner(s)' characteristics or their immediate environment. The binary variables coded as 'yes' or 'no,' included 'sex partner HIV+,' 'sex partner at risk of HIV,' 'last sex with main partner.' The 'gender of sexual partners,' is somewhat unique to a variable such as 'history of MSM contact' because it implies the gender of a subject's partner that is encountered with the most regularity. Based on pattern similarity in univariable analysis and small cells, the categories were collapsed to include 'none or only single gender sexual partner' or 'both,' that also reflects network-based hypotheses on relevant mixing. Geographic location was also included as a variable gleaned from the county that the subject resides. To provide a more qualitative scope to the subject's environment, the variable was coded as 'urban' if the subject resided in Multnomah, Washington, Clackamas, Marion, Lane or Jackson counties, and 'rural' if living in any other county in Oregon. Additionally, the subject's self-assessed risk, a subjective impression of their own partner(s)' behavior, coded as 'none,' 'low,' and 'high,' was also tested. The following three questions applied only to men answering 'yes' to a history of

MSM contact and were thought to particularly carry network importance: 'sexual contact with someone known less than 2 hours in past 6 months,' 'visiting a bath house in past 6 months,' and 'concern for external perception of sexual behavior.' The first two were coded as binary, 'yes,' and 'no.' The latter was kept with the categories, 'don't care who knows,' 'don't want a few specific people to know,' 'don't want most/everyone to know.' Lastly, 'sharing of needles in last six months,' for the IDU population was coded as 'yes' and 'no.'

The model-building strategy and model interpretation was completed in similar fashion as it was for the model of traditional risk factor variables. Furthermore, the combination of significant predictors from the traditional model and the network-tuned model were then compared to one another, in addition to comparing the full models from both approaches. Akaike's Information Criterion (AIC) was used to compare the respective models within the subset of subjects without missing data (32).

Statistical Methods: Network-Based Description

Descriptive analysis of the data in terms of network-based population determinants was completed based on the principles outlined by Aral et al. (3).

1. Absolute and relative size of core group: for the purposes of the study the core group was based purely on the number of reported sexual partners, thus core= 6+ partners in last year, adjacent= 2-5, or periphery= 0-1. The use of the terms core, adjacent and periphery are described as most reflective of a network approach, and

have been suggested to be used in further descriptions of sexual behavior (2). The absolute core group size was based on the number of respondents that reported 6+ sexual partners in the last year, and the relative size was based on the proportion of core group members compared to the total population of subjects.

2. Core group risk behavior: risk behavior described with respect to traditional risk factors as well as prevalence of HIV-positive test results.

3. Closed v. open populations: members of a high risk population having sexual connections with other sub-populations. Description was made of MSM and IDU pathways to sub-groups. For example, the proportion of MSM or self-identified 'gay' having had recent vaginal intercourse.

4. Concurrency: number of concurrent partners (different than serial monogamy), or measure of time between partners. The questionnaire did not include a direct question to assess concurrency, thus, the level of concurrency was estimated using- 'last sexual encounter with someone known less than 2 hours', and needle sharing with 'people don't know well,' for the MSM and IDU subpopulations.

5. Population composition: relevant mixing patterns and basic demographic statistics. Core, adjacent and periphery groups differences based on race, gender, and area of residence.

6. Sexual interaction among core group and adjacent or periphery individuals: the questionnaire did not specifically address core and non-core interaction, thus the determinant was estimated by the proportion of core individuals who deemed their partners 'low-risk.'

7. Bridges between subpopulations: the weak-link. By definition, the population is comprised of non-core individuals. Thus, the bridges parameter was estimated in the sub-group of the periphery individuals that estimated that their last sexual encounter was someone 'different than main partner' and that they had also had sex with a person that they deemed 'high risk' or 'HIV-positive.'

RESULTS:

Preliminary

As described earlier, a total of 65,919 HIV-test report forms were received between the dates of 9/1/98 and 8/30/01. Of the total number of forms submitted, 610 revealed HIV-positive results. Yet of the 610 positive reports, 298 were an initial HIV-positive finding, while the remaining 312 were a repeat or confirmatory test for subjects that had already been tested. The total number of subjects completing an HIV-test report form for the first time was 26,840 (40.7%), and of the remaining subjects that had been previously tested at any other time or location, and knew the result of the previous test, 35,945 (98.9%) reported having tested negative. Of the 26,840 initial-testers, 94 (.4%) tested HIV positive.

At least one in five subjects was tested anonymously, and the following table presents the distribution on choice of HIV testing by major subgroups:

Table 4. Test choice

Subgroup	Anonymous	Confidential	Not recorded	Total
MSM	38.8% (501)	59.3% (766)	1.9% (25)	1,292
IDU	23.4% (497)	72.9% (1550)	3.7% (79)	2,126
MSM-IDU	26.2% (44)	68.5% (115)	5.4% (9)	168

Information about race/ethnicity was available for 96.7% (62,087) of the reports, and the distribution for the study period is described in the following table:

Table 5. Race/Ethnicity

Race/ethnicity	Asian	Black	Hispanic	Native American	White	Other	Missing/ Unknown
% of total in study (% of general population of Oregon)*	2.6% (3.2%)	4.1% (1.6%)	21.4% (8.0%)	1.4% (1.3%)	65.3% (86.6%)	1.2% (7.3%)	3.9% ---

*based on 2001 census reports

Compared to the general population, the collection of HIV-test report forms do not appear to adequately reflect the racial/ethnic composition of the state. However, the distribution may simply represent a more accurate estimation of what race of individual was accessing public services. The Hispanic community also doubled in relative size from 1990 to 2001, 4.0% and 8.0% respectively, and the disparity may reflect a growing population during the three years of the study period.

On the HIV-test report form, the question that asked for the reason why subjects had sought to receive an HIV-test was asked in an open-ended fashion. Of the 26,840 subjects (97.2%) completed this item of the form and the vast majority, 22,388 (83.4%), reported that the test was self-initiated without presenting with any HIV-specific medical symptoms. Only 169 (0.6%) were completed for subjects that exhibited symptoms of the disease, while 2,694 (10%) were associated with prenatal care, and the remaining respondents listed a wide range of other assorted reasons.

Logistic Regression with Traditional Model

Analysis was completed on the 26,840 subjects testing for the first time in the state in the given study period. In the first portion of the non-network based analysis, the following variables were constructed (the reasoning for coding schema was addressed in the methodology section) and tested in univariable analysis:

1. Gender- coded as female = 0, male = 1
2. Race/Ethnicity- coded as ‘White’ (referent) = 0, ‘Black’ =1, ‘Hispanic’ = 2, ‘Other, including Asian and Native American’=3
3. Age- coded as continuous, age 10-92 years
4. Education <12 years- coded as ‘no’ = 0, ‘yes’ = 1

5. Reason for Test/Symptomatic- coded as 'no/asymptomatic' = 0, 'yes/symptomatic' = 1
6. History of MSM sexual contact- coded as 'no' = 0, 'yes' = 1
7. History of IDU- coded as 'no' = 0, 'yes' = 1
8. History of exchanging sex for drugs or money- coded as 'no' = 0, 'yes' = 1
9. Number of sex partners in the last 12 months- coded as '0-1' = 0, '2-5' = 1, '6+' = 2
10. Unprotected Intercourse of Any Kind in the Last 6 months- coded as 'yes' = 0, 'no/don't know' = 1

Initial crosstabulation of the traditional variables produced the following results (number of subjects with missing data are included):

Table 6. Initial crosstabulation of traditional variables

Variable	Number of Subjects missing data for variable (% of total)	% Negative in variable category	% Positive in variable category	Pearson Chi-Square (df)	Significance
Gender (female)	1,033 (3.8%)	53.1%	14.9%	39.06 (1)	p<.001
(male)		46.9%	85.1%		
Race/Ethnicity (white)	53 (.2%)	65.1%	58.0%	15.002 (3)	p=.002
(black)		4.1%	13.0%		
(Hispanic)		21.9%	23.2%		
(other, including Asian and Native American)		9.0%	5.8%		
Age	1,789			33.961	p<.001

(continuous, univariable regression)	(6.7%)				
Reason for Test (no symptoms)	0 (0%)	99.4%	95.7%	16.744 (1)	p<.001
(symptomatic)		.6%	4.3%		
Education (>12 yrs)	0 (0%)	68.4%	76.8%	2.274 (1)	p=.132
(<12 yrs)		31.6%	23.2%		
History of MSM activity (no)	0 (0%)	95.1%	44.9%	359.091 (1)	p<.001
(yes)		4.9%	55.1%		
History of IDU (no)	0 (0%)	94.0%	91.3%	.351 (1)	p=.310
(yes)		6.0%	8.7%		
Number of Partners in Last 12 months (periphery: 0-1)	0 (0%)	39.9%	29.8%	7.887 (2)	p=.019
(adjacent: 2-5)		40.0%	39.4%		
(core: 6+)		20.0%	30.9%		
History of Exchanging Sex for Drugs/Money (no)	0 (0%)	98.0	98.6	.110 (1)	p=.740
(yes)		2.0%	1.4%		
Unprotected Sex of Any Kind (no)	2,261 (8.4%)	10.8%	10.2%	.028 (1)	p=.865
(yes)		88.9%	89.8%		

To perform logistic regression analysis, the subjects with any missing data were deleted, reducing the number of subjects from 26,840 to 22,670 (85%). For purposes of clinical interest, all variables were kept in the model:

Table 7. Full traditional model summary

-2 Log likelihood	Cox & Snell R	Nagelkerke R Square

	Square	
791.319	.009	.208

Table 8. Hosmer and Lemeshow Test

Chi-square	df	Sig.
10.235	8	.249

Table 9. Odds ratio for full traditional model

Variable	**Odds Ratio (95% CI)**	**Significance**
Unprotected Sex of Any Kind	.857 (.383, 1.919)	P=.708
Reason for Test (Symptomatic)	2.785 (.641, 12.098)	P=.172
Education <12 years	1.482 (.855, 2.568)	P=.161
Number of Partners (periphery)	1.00 (referent)	
(adjacent)	1.323 (.739, 2.368)	P=.347
(core)	1.674 (.891, 3.145)	P=.109
Gender (male)	1.702 (.808, 3.589)	P=.162
Age	1.054 (1.036, 1.072)	**P<.001**
Race (white)	1.00 (referent)	
(black)	5.752 (2.783, 11.892)	**P<.001**
(Hispanic)	1.989 (1.075, 3.681)	**P=.029**
(other)	1.692 (.706, 4.056)	P=.239
MSM	17.50 (10.127, 30.241)	**P<.001**
IDU	1.983 (1.059, 3.713)	**P=.032**
Exchange Sex for Drugs/Money	.167 (.022, 1.270)	P=.084

Thus, the Hosmer and Lemeshow test upholds the null hypothesis and demonstrates that the model is a good fit. The variables of 'MSM,' 'IDU,' 'Age,' and 'black' and 'Hispanic' race show a significant increase in risk at α=.05. Otherwise, the variable of 'exchange sex for drugs/money,' though not strictly significant, does suggest that such a history may be somewhat protective in the study population.

Logistic Regression with Network-Tuned Parameters

Analysis was completed on the 26,840 subjects testing for the first time in the state in the given study period. In the second portion of the network-tuned parameter analysis, the following variables were constructed (the reasoning for coding schema was addressed in the methodology section) and tested in univariable analysis:

1. Geographic Location- coded as 'rural' = 0, 'urban' = 1
2. Sex or Needle Partner HIV positive- coded as 'no' = 0, 'yes' = 1
3. Sex Partner at Risk for HIV- coded as 'no' = 0, 'yes' = 1
4. Self-assessed Risk- coded as 'no risk/unsure' = 0, 'low risk' = 1, 'high risk' = 2. Individuals previously known to be HIV+ have been excluded from analysis, thus category of 'known HIV+' from this question was eliminated.
5. Gender of Sexual Partner(s)- coded as 'none or only single gender' = 0, 'both' = 1
6. Last Sexual Encounter- coded as 'main partner' = 0, 'someone else' = 1
7. MSM Contact with Someone Known <2 hours- coded as 'no' = 0, 'yes' = 1
8. Concern with External Perception- coded as 'don't know who cares' = 0, 'don't want few specific people to know' = 1, 'don't want most/everyone to know' = 2
9. Needle Sharing in the Last Six Months- coded as 'no' = 0, 'yes' =1
10. Alcohol or Drug Use with Sex- 'never' = 0, 'sometimes' = 1, 'always' = 2

Table 10. Initial crosstabulation of network variables

Variable	Number of Subjects missing data for variable (% of total)	% Negative in variable category	% Positive in variable category	Pearson Chi-Square (df)	Significance
Geographic Area of Residence (rural)	0 (0)	33.5%	13.8%	12.855 (1)	P<.001
(urban)		66.5%	86.2%		
Sex/Needle Partner HIV+ (no)	0 (0)	99.2%	94.2%	21.269 (1)	P<.001
(yes)		.8%	5.8%		
Partner at Risk of HIV (no)	0 (0)	83.0%	89.9%	2.317 (1)	P=.128
(yes)		17.0%	10.1%		
Gender of Sexual Partners (none/ single gender only)	0 (0)	95.3%	77.7%	64.656 (1)	p<.001
(both)		4.7%	22.3%		
Last Sex with Main Partner (yes)	0 (0)	78.3%	72.5%	1.355 (1)	P=.244
(no)		21.7%	27.5%		
Self Assessed Risk (none)	1,505 (5.6%)	40.7%	39.0%	3.448 (2)	P=.178
(low)		53.2%	49.2%		
(high)		6.1%	11.9%		
Concern with External Perception (do not care who knows)	816 (3%)	98.5%	71.4%	272.017 (2)	P<.001
(do not want a few specific people to know)		.6%	14.3%		
(do not want most/everyone		.9%	14.3%		

to know)					
Sex with Someone Known <2 hrs (no)	794 (3%)	99.2%	84.2%	154.938 (1)	P<.001
(yes)		.8%	15.8%		
Sharing Needles (no)	716 (2.7%)	95.8%	92.0%	3.295 (1)	p=.07
(yes)		4.2%	8.0%		
Alcohol or Drug Use with Sex (never)	3,572 (13.3%)	50.3%	53.9%	11.623 (2)	p=.003
(sometimes)		45.3%	34.2%		
(always)		4.4%	11.8%		

To perform logistic regression analysis, the subjects with any missing data were deleted, reducing the number of subjects from 26,840 to 21,221 (79.1%). For purposes of clinical interest and consistency with the analysis of the traditional model, all variables were kept in the network-based model:

Table 11. Full network model summary

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
594.444	.005	.144

Table 12. Hosmer and Lemeshow Test

Chi-square	df	Sig.
9.273	8	.320

Table 13. Odds ratio for full network-based model

Variable	**Odds Ratio (95% CI)**	**Significance**
Geographic Area of Residence (urban)	2.745 (1.219, 6.180)	**P=.015**
Sex or Needle Partner HIV+ (yes)	17.944 (6.307, 51.052)	**P<.001**
Sex Partner at Risk for HIV (yes)	1.366 (.545, 3.429)	P=.506
Self Assessment of Risk (none)	1.00 (referent)	
(low)	.699 (.363, 1.345)	P=.283
(high)	1.748 (.770, 3.968)	P=.182
Gender of Sexual Partner (both)	1.206 (.479, 3.038)	P=.690
Last Sex with Main Partner (no)	.642 (.311, 1.324)	P=.230
MSM Contact with Someone Known <2 hours	3.095 (.988, 9.70)	**P=.053**
Concern with External Perception (do not care who knows)	1.00 (referent)	
(do not want a few specific people to know)	14.282 (4.144, 49.219)	**P<.001**
(do not want most/everyone to know)	12.446 (4.057, 38.179)	**P<.001**
Alcohol or Drugs with Sex (never)	1.00 (referent)	
(sometimes)	.689 (.365, 1.303)	P=.252
(always)	1.603 (.572, 4.491)	P=.369
Needle Sharing in Last Six Months (yes)	2.654 (.954, 7.384)	P=.062

Thus, the Hosmer and Lemeshow test upholds the null hypothesis and demonstrates that the model is a good fit. The variables of 'geographic area of residence,' 'sex or needle partner HIV+,' 'MSM contact with someone known <2 hours,' and 'concern with external perception' show a significant increase in risk at α=.05. Otherwise, variables such as 'needle sharing in the last six months,' and a 'high self-assessed risk' are suggestive of a significant increase in risk of testing HIV+.

Comparison of the Full Traditional and Network-Based Models

As demonstrated above, both are 'good fitting models' by the Hosmer and Lemeshow test, but it is important to compare the models head-to-head using Akailke's Information Criterion (AIC) = -2 log likelihood + [(2)*(# of parameters)]. The model with the lower AIC is generally considered a 'better' model.

Traditional Model: 791.391 + (2*12) = 815.391

Network-Based Model: 594.44 + (2*14) = 622.44

However, to be most accurate, it is necessary to compare within the same subset of subjects without any missing data. When all variables are applied and subjects missing any data are removed, the total number of subjects is reduced to 18,286 (68.1%):

Table 14. Model summary of traditional model in reduced subset (n=18, 286)

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
450.867	.006	.211

Table 15. Hosmer and Lemeshow Test of traditional model in reduced subset

Chi-square	df	Sig.
5.203	8	.736

Table 16. Odds ratio for traditional model in reduced subset (n=18,286)

Variable	**Odds Ratio (95% CI)**	**Significance**
Unprotected Sex of Any Kind	.880 (.262, 2.959)	P=.837
*Reason for Test (Symptomatic)	-	P~1.00
Education <12 years	1.348 (.635, 2.863)	P=.437
Number of Partners (periphery)	1.00 (referent)	
(adjacent)	1.819 (.780, 4.241)	P=.166
(core)	2.469 (.969, 6.290)	P=.058
Gender (male)	2.046 (.774, 5.406)	P=.149
Age	1.062 (1.039, 1.086)	**P<.001**
Race (white)	1.00 (referent)	
(black)	3.428 (1.137, 10.330)	**P=.029**
(Hispanic)	1.477 (.588, 3.708)	P=.406
(other)	1.305 (.386, 4.416)	P=.669
MSM	20.082 (9.772, 41.270)	**P<.001**
IDU	2.006 (.809, 4.976)	P=.133
Exchange Sex for Drugs/Money	.418 (.054, 3.229)	P=.403

*When reduced to 18,286, all patients that answered 'symptomatic' for the Reason for Test variable were subsequently dropped therefore the variable becomes irrelevant. Odds ratios above reflect model run with the Reason for Test variable removed.

Now examining the network model in the same subset of 18,286 subjects without any missing data:

Table 17. Model summary of network model in reduced subset (n=18,286)

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
476.261	.005	.166

Table 18. Hosmer and Lemeshow Test of network model in reduced subset

Chi-square	df	Sig.
11.842	8	.158

Table 19. Odds ratio for network model in reduced subset (n=18,286)

Variable	**Odds Ratio (95% CI)**	**Significance**
Geographic Area of Residence (urban)	3.803 (1.337, 10.817)	**P=.012**
Sex or Needle Partner HIV+ (yes)	22.841 (7.783, 67.031)	**P<.001**
Sex Partner at Risk for HIV (yes)	1.163 (.545, 3.429)	P=.790
Self Assessment of Risk (none)		
(low)	.705 (.340, 1.460)	P=.347
(high)	1.654 (.676, 4.046)	P=.271
Gender of Sexual Partner (both)	.712 (.278, 1.825)	P=.479
Last Sexual Main Partner (no)	.626 (.283, 1.383)	P=.246
MSM Contact with Someone Known <2 hours	3.048 (.953, 9.749)	**P=.060**
Concern with External Perception (do not care who knows)	1.00 (referent)	
(do not want a few specific people to know)	13.277 (3.562, 49.491)	**P<.001**
(do not want most/everyone to know)	13.834 (4.359, 43.906)	**P<.001**
Alcohol or Drugs with Sex (never)	1.00 (referent)	
(sometimes)	.775 (.387, 1.552)	P=.472
(always)	1.712 (.550, 5.328)	P=.353
Needle Sharing in Last Six Months (yes)	2.510 (.816, 7.720)	P=.108

Utilizing the AIC comparison:

Traditional Model= 450.867 + (2*12) = 474.867

Network-Based Model= 476.261 + (2*13) = 502.261

Though in direct head-to-head comparison the traditional model outperforms the network-based model, the results are not overwhelmingly in favor of the traditional model. The next logical step then, is to develop a 'best' combination of the two models that incorporates the strongest of variables from the two approaches.

A forward procedure using likelihood ratios reveals the same four variables as a conditional combination: 'Age,' 'MSM,' 'Sex or Needle Partner HIV+,' and 'Geographic Area of Residence.' Interestingly, a backward conditional procedure that completed 16 iterations settled on a model that included the traditional variables of 'Age,' 'Gender' and 'MSM,' and the network-based variables of 'Sex or Needle Partner HIV+,' 'Self-Assessed Risk,' and 'Geographic Area of Residence.'

Thus, employing the 'best' combination model from the backward conditional procedure for comparison (with n=18,286):

Table 20. Model summary of the 'best' combination of network and traditional variables

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
426.044	.008	.255

Table 21. Hosmer and Lemeshow Test for 'best' combination (n=18,286)

Chi-square	df	Sig.
5.120	8	.745

Table 22. Odds Ratios for 'Best' Combination Model (n=18,286)

	Sig.	Odds Ratio	95.0% C.I. for Odds Ratio	
			Lower	Upper
MSM	p<.001	21.165	10.104	44.338
Sex or Needle Partner HIV+	p<.001	23.885	8.116	70.288
Geographic Area of Residence	p=.020	3.470	1.220	9.871
Gender	p=.049	2.637	1.005	6.921
Age	p<.001	1.056	1.033	1.079
Self Assessed Risk (none)	p=.054			
(low)	p=.139	.574	.276	1.197
(high)	p=.283	1.600	.679	3.775

And in the full dataset again, we pick up more subjects with the reduced or 'best' combination model. Now n=23,404 (and a total of 66 positive cases):

Table 23. Model summary of the 'best' combination (n=23,404)

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
703.659	.009	.227

Table 24. Hosmer and Lemeshow Test for 'best' combination (n=23,404)

Chi-square	df	Sig.
11.032	8	.200

Table 25. Odds ratios for 'best' combination model (n=23,404)

	Sig.	Odds Ratio	95.0% C.I. for Odds Ratio	
			Lower	Upper
MSM	p<.001	15.910	8.810	28.730
Sex or Needle Partner HIV+	p<.001	19.456	7.528	50.284
Geographic Area of Residence	p=.002	3.881	1.664	9.049
Gender	p=.133	1.786	.838	3.804
Age	p<.001	1.050	1.032	1.068
Self Assessed Risk (none)	p=.005			
(low)	p=.490	.816	.458	1.453
(high)	p=.016	2.286	1.166	4.481

And adding the variables Age^2 and Age^3 to model more precisely the continuous variable of Age demonstrates an even greater lowering of the log likelihood:

Table 26. Model summary of 'best' combination (n=23,404) including exponential modeling of the continuous variable, Age

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
681.761	.010	.252

Network-Based Descriptive Analysis

1. The Absolute and Relative Size of the Core Group

The core group was defined by the number of sexual partners that the subject reported as having in the last 12 months. The designation of 'core' was given to a subject having 6 or more partners. Of the 26,840 initial test reports, 5,389 (20.1% of the total population) were in the core group. Approximately 17% of the core group reported that the number of partners in the last year exceeded 10, while the range extended, for a small percentage of respondents, well into the hundreds. Table 27 summarizes the absolute and relative size of the core group with respect to the overall study population:

Table 27. Categories of number of sexual partners- defining the core group

	Frequency	Percent
periphery, 0-1 partners in last 12 months	10709	39.9
adjacent, 2-5 partners in last 12 months	10742	40.0
core, >6 partners in last 12 months	**5389**	**20.1**
Total	26840	100.0

2. Core Group Risk Behavior

A summary of basic risk behavior within the core group is displayed in the following table and demonstrates that the core group's basic risk behavior does not differ greatly from the total population:

Table 28. Core group risk behavior

Risk Group	Percentage of Core Group (n=5,389)	Percentage of Total Population (n=26,840)
MSM-IDU	.9%	.6%
MSM (not IDU)	6.7%	4.8%
IDU (not MSM)	8.2%	7.9%
Sex or Needle Partner HIV+	.8%	.8%
Exchange Sex or Drugs for Money	2.1%	1.1%

.

3. Concurrency

Since a formal question of concurrency was not addressed in the HIV-test report form, the level of concurrency in the population was estimated with the question, 'last sex with main partner,' for the MSM subgroup, the additional question of 'sex with someone known less than 2 hours,' and for the IDU subgroup, 'needle sharing frequency.' Ideally, concurrency would be measured as a continuous variable, but in this study population an estimated descriptor served as proxy. Tables 28-30 describe the estimators of concurrency for the core, adjacent, and periphery groups separately.

Table 29. Core group respondent's last sexual partner

Risk Behavior	% last having sex with 'someone else'	Number in Core group	% of Core Group
MSM-IDU	61.7	27	.5
MSM	50.7	183	3.4
IDU	40.6	177	3.3
Sex/Needle Partner HIV+	51.2	22	.4
Exchange Sex/Drugs for Money	56.5	65	1.2

Table 30. Adjacent group respondents' last sexual partner

Risk Behavior	% last having sex with 'someone else'	Number in Adjacent Group	% of Adjacent Group
MSM-IDU	46.4	32	.3
MSM	40.8	247	2.3
IDU	30.8	322	3.0
Sex/Needle Partner HIV+	30.4	21	.2
Exchange Sex/Drugs for Money	51.3	54	.5

Table 31. Periphery group respondents' last sexual partner

Risk Behavior	% last having sex with 'someone else'	Number in Periphery Group	% of Periphery Group
MSM-IDU	30.8	11	.1
MSM	28.1	86	.8
IDU	14.8	129	1.2
Sex/Needle Partner HIV+	16.3	11	.1
Exchange Sex/Drugs for Money	21.0	11	.1

For the subpopulation of MSM within the core group, at least 50% of respondents reported having sex with someone known for less than 2 hours in the past six months, regardless of IDU status (50.7% for MSM only, and 61.7% for MSM-IDU). The proportions begin to drop in the adjacent (46.4% MSM only, and 40.8% of MSM-IDU) and periphery groups (30.8% MSM only, 28.1% of MSM-IDU).

For the subpopulation IDU, the relative proportions of needle-sharing frequency do not differ greatly among the core, adjacent, or periphery groups. However for those respondents answering the question, the proportion remains high: core group needle sharing frequency (sometimes 24.2%, often 11.3%, almost always 7.7%), adjacent group (sometimes 18.5%, often 11.6%, almost always 3.8%), and periphery group (sometimes 20.4%, often 7.9%, almost always 5.7%).

4. Closed v. open populations

To reflect a degree of interaction between closed and open populations and describe members of a high-risk population having sexual connections with other sub-populations, the MSM population was examined in greater detail. In men who had reported a history of MSM contact and completed the question regarding self-description ('gay,' 'bisexual,' 'heterosexual'), the proportion engaging in incongruous sexual activity was analyzed as an estimator of closed and open population interaction.

The following tables separate core, adjacent, and peripheral groups:

Table 32. Core group of MSM subpopulation

Self Description	Vaginal Sex % in last 6 mo,	% ever	Anal Receptive Sex % in last 6 mo,	% ever
Gay (n=105)	**5.6**	**37.1**	74.2	90.7
Heterosexual (n=27)	12.5	79.2	**4.5**	**22.7**

Table 33. Adjacent group of MSM subpopulation

Self Description	Vaginal Sex % in last 6 mo,	% ever	Anal Receptive Sex % in last 6 mo,	% ever
Gay (n=180)	**6.9**	**36.3**	64.2	82.1
Heterosexual (n=62)	77.6	91.0	**17.2**	**37.5**

Table 34. Periphery group of MSM subpopulation

Self Description	Vaginal Sex % in last 6 mo,	% ever	Anal Receptive Sex % in last 6 mo,	% ever
Gay (n=75)	**1.6**	**25.8**	35.8	62.7
Heterosexual (n=39)	50.0	88.9	**6.1**	**39.4**

Perhaps even more revealing, examination of MSM that reported the greatest degree of concern about external perception of their sexuality, or alternatively the most *closeted*, also reported much greater degrees of recent vaginal intercourse:

Table 35. MSM that reported 'do not want anyone to know' about their history of male-to-male sexual behavior

Subgroup	Vaginal Sex % in last 6 mo	% ever
Core (n=60)	54.5	77.2
Adjacent (n=121)	63.6	77.3
Peripheral (n=66)	34.6	69.2

5. Population Composition

The potential for relevant mixing patterns were explored for each major subgroup based on gender, race, and geographic location and separated by core, adjacent, and peripheral categories:

Table 36. Core group- population composition

Risk Subgroup (number in the subgroup)	Geographic Location: %urban / % rural	Race/Ethnicity: % of subgroup white = W, black = B, Hispanic = H, Asian = A, Native American = N, Other = O	Gender: % female / % male
MSM-IDU (49)	78.7 / 21.3	W = 73.9, B = 6.5, H = 4.4, A = 3.2, N = 4.3, O = 2.2	0 / 100
MSM (361)	79.5 / 20.5	W = 70.8, B = 3.2, H = 16.2, A = 4.2, N = 1.4, O = 2.2	0 / 100
IDU (442)	68.0 / 32.0	W = 74.3, B = 2.7, H = 9.3, A = .2, N = 5.0, O = 1.8	40.0 / 60.0
Sex/Needle Partner HIV+ (43)	58.1 / 41.9	W = 65.1, B = 2.3, H = 16.3, A = 2.3, N = 2.3, O = 2.3	56.1 / 43.8
Exchange Sex/Drugs for Money (113)	66.1 / 32.9	W = 62.3, B = 9.6, H = 21.1, A = .9, N = 1.8, O = 1.8	61.4 / 38.6

Table 37. Adacent group- population composition

Risk Subgroup (number in the subgroup)	Geographic Location: %urban / % rural	Race/Ethnicity: % of subgroup white = W, black = B, Hispanic = H, Asian = A, Native American = N, Other = O	Gender: % female / % male
MSM-IDU (64)	68.1 / 31.9	W = 85.5, B = 5.8, H = 8.7, A = 0.0, N = 0.0, O = 0.0	0 / 100
MSM (623)	75.2 / 24.8	W = 73.6, B = 2.4,	0 / 100

		H = 17.0, A = 2.6, N = 1.1, O = 1.6	
IDU (838)	64.2 / 335.8	W = 83.0, B = 2.9, H = 5.2, A = .6, N = 4.5, O = .6	40.2 / 58.8
Sex/Needle Partner HIV+ (75)	54.4/ 45.6	W = 78.5, B = 1.3, H= 13.9, A = 0.0, N = 2.5, O = 1.3	56.6 / 43.4
Exchange Sex/Drugs for Money (118)	69.6/ 30.4	W = 60.9, B = 7.0, H = 27, A = 2.6, N = 0.0, O = .9	63.7 / 36.3

Table 38. Periphery group- population composition

Risk Subgroup (number in the subgroup)	Geographic Location: %urban / % rural	Race/Ethnicity: % of subgroup white = W, black = B, Hispanic = H, Asian = A, Native American = N, Other = O	Gender: % female / % male
MSM-IDU (54)	73.1.1 / 26.9	W = 80.8, B = 0.0, H = 11.5, A = 0.0, N = 0.0, O = 1.9	0 / 100
MSM (311)	71.9 / 28.1	W = 66.1, B = 1.9, H = 23.0, A = 2.9, N = 1.0, O = 1.9	0 / 100
IDU (846)	64.3 / 35.7	W = 83.4, B = 2.2, H = 5.2, A = .8, N = 2.6, O = .5	35.0 / 65.0
Sex/Needle Partner HIV+ (96)	70.7/ 29.3	W = 78.3, B = 4.8, H = 9.3, A = 3.3, N = 2.2, O = 0.0	54.9 / 45.1
Exchange Sex/Drugs for Money (64)	54.8/ 45.2	W = 61.3, B = 12.9, H = 17.7, A = 3.2, N = 1.6, O = 0.0	33.3 / 66.7

6. Sexual Interaction Among Core Group and Adjacent or Periphery Individuals

Since the questionnaire did not address the concept specifically, the degree of interaction among core group and adjacent or periphery individuals was estimated by the proportion of the core group that deemed their partners 'low risk.' 2,263 individuals

(40.2%) in the core group thought that they were at little or no risk of HIV infection. Of the 40.2% in the core group, the gender and geographic distributions do not change significantly from what was previously reported, 76.9% of women and 73.1% of men live in urban areas. However, a disproportionately higher number of the core group that deemed their partners 'low risk' were Hispanic (34.4%), and a high proportion of that subgroup of Hispanics (85.8%) live in urban areas.

7. Bridges Among Subpopulations

The degree of bridging among subpopulations can be estimated by examining those subjects in the adjacent and periphery groups that that answered yes to having last had sex with someone other than their main partner (31.0% of the adjacent group and 10.6% of the periphery group):

Table 39. Women in adjacent/periphery groups reporting sex with someone other than main partner in last six months

Gender of Sexual Partner(s)	% Vaginal Sex in Last 6 months	% Anal Receptive Sex in Last 6 Months	% Oral Sex in Last Six Months
Men	89.9	6.9	56.7
Women	2.7	0.0	60.5
Both	7.2	9.7	74.8

Table 40. Men in adjacent/periphery groups reporting sex with someone other than main partner in last six months

Gender of Sexual Partner(s)	% Vaginal Sex in Last 6 months	% Anal Receptive Sex in Last 6 Months	% Oral Sex in Last Six Months
Men	22.7	42.3	73.7
Women	92.8	.4	53.4
Both	56.2	27.2	69.9

DISCUSSION:

The Traditional and Network-Based Models

After comparison of the traditional and the network-based models one can conclude that within the confines of the study design, the network-based variables warrant further study and certainly necessitate the incorporation of network-based theory in large population-based surveillance projects. Clearly, the greatest benefits of a network-based approach are lost when one is unable to map directly who is sharing risk with whom. With the increased demand for protection of privacy, such studies, for better or worse, will be increasingly difficult to accomplish. Therefore, designing epidemiological questions that are rooted in network theory and refer to the environment of the subject's sexual, needle-sharing, or friendship partners without compromising that partner's personal identity or privacy is of greater importance. A combination of risk factors inspired both from traditional and network-based sources will likely prove to be of most use to future population-based surveys.

The variables that were assembled for the network-based model may also perform better against a traditional model in a setting other than the public sector in the state of Oregon. As discussed earlier, contrary to the national trends, Oregon has not observed a surge of HIV in the heterosexual population. Therefore, the variable of 'history of MSM,' perhaps even more so in the study period from 1998-2001 than now, will likely be the strongest predictor in any model of a similar setting. However, in a community where

HIV transmission may resemble other STIs and involve a larger number of diverse sub-populations, a network-based model may be of greater predictive value.

Furthermore, framing of descriptive results into network-based explanations provides a unique account of the study population that may better inform public health interventions. For instance, one question focused on the sexual interaction of core group members with other 'low-risk' individuals that may be in a non-core group. Approximately 2,263 (40.2%) of the core group thought that their sexual partners were low-risk. One may wonder whether efforts at education as to what constitutes risk may be of benefit for individuals within the core group, or an alternative hypothesis may be that a certain portion of the core group is engaging in numerous 'low-risk' encounters over the course of the previous year. Such encounters, particularly if engaged in concurrently with other relationships, may be more likely to occur without the protection of condoms or other safe-sex practices contrary to 'high-risk' encounters at a bath-house or with a commercial-sex worker where one may be more likely to use a barrier device. Thus, prevention efforts must also consider other venues for intervention beside historically 'high-risk' areas.

With respect to employing network-based explanations to descriptive results, future studies may wish to focus on smaller populations than the cachement area for an entire state. It may be more reasonable to divide subjects into core, adjacent, and peripheral groups for comparison when the population shares a realistic geographic and ethnographic proximity. County-level analyses may be most feasible when attempting to balance issues such as geographic specificity with ethical considerations of personal privacy.

Limitations and Improvements

The greatest limitations to the study arise secondary to the survey design. Subjects were asked a series of personal questions in a medical setting that is already fraught with heightened sensitivity and discomfort. For many in such a population, it may have been the first time in numerous years to visit a medical provider. Despite the assurance of anonymity or confidentiality the subjects may have been reluctant to share such detailed information about their sexual history. In such instances, the recall bias would favor the null hypothesis for certain variables where subjects might perceive others to judge their answers as deviant or risky. Therefore, subjects may have answered questions such as 'self-assessed risk,' 'last sex with main partner,' or 'alcohol or drug use with sex' more conservatively or declined to answer altogether. It is likely that the recall bias would disproportionately affect the variables of the network-based model. It also remains unclear to what degree the process of administering the questionnaire was standardized. A range of medical providers in a variety of different public clinic settings carried out the survey over the course of three years. It is conceivable that certain providers may have prompted or delved deeper for information while others may have tired with the process or viewed the questionnaire as superfluous or a mere formality.

Needless to say, the most obvious consequence of the study design was the degree of missing data. With the study limited to first-time testers it was thought that the number of subjects with missing data would be minimized. To compare the traditional and network-based models head-to-head it was necessary to eliminate all subjects with any missing data, dropping the total n= 26,840 to n= 18,286 (68.1%). However, the total number of subjects testing HIV+ fell from 94 to 40 (42.5%). The subjects with missing data tested positive in 54/8554 (0.63%) of cases, while those without missing data tested positive in 40/18,286 (0.22%) of cases. The discrepancy was not explained by any one variable in particular. For instance, analysis was completed on the variable with the largest number of subjects having missing data, 'alcohol or drug use with sex,' 3,572 (13.3%). Of the 3,572 subjects that were missing data for that variable, 18 (0.50%) tested HIV+, leaving 76 to test HIV+ out of the remaining 23,268 that had data for the question (0.33%). A chi-square test (2.789, p=.095) found no statistical difference between the two groups. Unfortunately, it appears the subjects that tested HIV+ were simply more likely to leave one or more questions blank but without any predominant pattern.

With only 40/18,286 (0.22%) of subjects testing HIV+, it is important to restate that the false positive rate for the HIV test does not approach or exceed the prevalence in the study population and indeed remains negligible. Causes for a false-positive EIA test include such conditions as concurrent hematologic malignancy, autoimmune disorders, or chronic renal failure (17). However, all subjects that were EIA positive had a WB test performed and in the rare instance when the WB test was read as equivocal a final IFA test was also completed. The recent generation of EIA tests has reported a sensitivity and specificity of greater than 99.9% in serum from subjects with known HIV-1 infection (17). When combined with a confirmatory WB test the specificity improves further. Thus, the false positive rate does not present any statistical concern, however the overall prevalence of only 0.22% HIV+ subjects may have limited the study's ability to test the novel network-based variables where such variables may be more appropriately employed in communities with a higher prevalence of HIV.

The second issue involving the study design concerns the manner in which the questions were phrased. Since the questionnaire was not conceived with the original intent of testing a network-based hypothesis, certain questions were better suited while other aspects of network theory were not specifically addressed. Questions that can be included in future epidemiological studies that better incorporate key principles of network theory are assembled in the following table:

Table 41. Complementary epidemiological questions that better address network theory

Network-based principle	Example of question*
Concurrency	"How many sexual partners do you maintain at one time?," "Do you have one or more long-term partners other than your main sexual partner?-if so, tell me about that other long-term partner(s)?"
Duration	"How long is your typical sexual relationship?"
Partner's Behavior/Centrality	"How many partners does your partner(s) maintain?," "How much control does your partner exert over your sex life?," "Are all your sexual partners from your own social group?"
Friendship/Social Cohesion	"Where do you acquire information about sex or HIV?," "With whom do you talk about sex?"

*Needle-sharing or other behaviors may be substituted for sexual practices

The complementary questions that address network-based principles can elucidate further whether a subject may be acting as a central hub, as a link to separate subpopulations, or as a distant outlier without a great degree of influence in the network. In addition, the complementary questions provide a better understanding of where to target future public health intervention.

Conclusion

Network theory shows great promise in being able to complement traditional epidemiological approaches to HIV transmission. Future survey design should seek to incorporate questions that reflect network-based principles such as concurrency, duration, centrality, and social cohesion. Network-based explanations should also be considered in framing descriptive results from populations where subjects share geographic and ethnographic proximity. In particular, network-based models may be of most use in communities with higher HIV prevalence and where HIV transmission involves a more diverse group of subpopulations with higher rates of heterosexual activity and injection drug use.

REFERENCES:

1. Jolly AM., Muth SQ., et al. Sexual networks and sexually transmitted infections: a tale of two cities. *Journal of Urban Health.* 78(3):433-45, 2001 Sep.

2. Youm Y., Laumann EO. Social network effects on the transmission of sexually transmitted diseases. *Sexually Transmitted Diseases*. 29(11):689-97, 2002 Nov.

3. Aral SO. Sexual network patterns as determinants of STD rates: paradigm shift in the behavioral epidemiology of STDs made visible. *Sexually Transmitted Diseases.* 1999; 26: 262–264.

4. Rothenberg R., Narramore J. The relevance of social network concepts to sexually transmitted disease control. *Sexually Transmitted Diseases*. 23(1):24-9, 1996 Jan-Feb.

5. Parran T. Shadow on the Land. New York: Reynal and Hitchcock, 1937.

6. Ghani AC., Garnett GP. Risks of acquiring and transmitting sexually transmitted diseases in sexual partner networks. *Sexually Transmitted Diseases.* 27(10):579-87, 2000 Nov.

7. Potterat JJ., Rothenberg RB., et al. Network structural dynamics and infectious disease propagation. *Int J STD AIDS* 1999; 10: 182–185.

8. Korenromp EL., de Vlass SJ., et al. Estimating the magnitude of STD cofactor effects on HIV transmission: how well can it be done? *Sexually Transmitted Diseases.* 28(11):613-21, 2001 Nov.

9. Rothenberg R. How a net works: implications of network structure for the persistence and control of sexually transmitted diseases and HIV. *Sexually Transmitted Diseases.* 28(2):63-8, 2001 Feb.

10. Barbasi, A. Linked. Cambridge: Perseus Publishing, 2002.

11. Math Forum. http://mathforum.org/isaac/problems/bridges2.html. Retrieved June, 2003.

12. Kretzschmar M., Sexual network structure and sexually transmitted disease prevention: a modeling perspective. *Sexually Transmitted Diseases.* 27(10):627-35, 2000 Nov.

13. Kretzschmar M,. van Duynhoven YT., et al. Modeling prevention strategies for gonorrhea and Chlamydia using stochastic network simulations. *American Journal of Epidemiology.* 144(3):306-17, 1996 Aug 1.

14. Miller M., Neaigus A. Networks, resources and risk among women who use drugs. *Science & Medicine.* 52(6):967-78, 2001 Mar.

15. Rosenberg D., Moseley K., et al. Networks of persons with syphilis and at risk for syphilis in Louisiana: evidence of core transmitters. *Sexually Transmitted Diseases.* 26(2):108-14, 1999 Feb.

16. Rothenberg R., Kimbrough L., et al. Social network methods for endemic foci of syphilis: a pilot project. *Sexually Transmitted Diseases.* 27(1):12-8, 2000 Jan.

17. Cluster of HIV-infected adolescents and young adults--Mississippi, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 49(38):861-4, 2000 Sep 29.

18. Flom PL., Friedman SR., et al. Stigmatized drug use, sexual partner concurrency, and other sex risk network and behavior characteristics of 18- to 24-year-old youth in a high-risk neighborhood. *Sexually Transmitted Diseases.* 28(10):598-607, 2001 Oct.

19. UNAIDS. Sub-Saraharan Africa. Fact Sheet 2002. Available at http://www.unaids.org. Accessed March 10, 2003.

20. National Institute of Allergy and Infectious Diseases. Fact Sheet: Treatment of HIV. January 2002. Available at http://www.niaid.nih.gov/factsheets/treat-hiv.htm. Accessed March 15, 2003.

21. Dr. Mark Loveless, Oregon Department of Human Services- HIV/STD/TB Program, and Oregon Health and Sciences University- Department of Public Health and Preventive Medicine. Personal Correspondence.

22. US Centers for Disease Control and Prevention. Increases in unsafe sex and rectal gonorrhea among men who have sex with men – San Francisco, California, 1994–1997. *MMWR* 1999, 48: 45–58.

23. US Centers for Disease Control and Prevention. Resurgent bacterial sexually transmitted disease among men who have sex with men – King County, Washington, 1997–1999. *MMWR* 1999, 48: 773–777.

24. Mansergh G., Marks G. et al. 'Barebacking' in a diverse sample of men who have sex with men. *AIDS* (03.08.02) 2002;16:653-659.

25. Sheon N, Plant A. Protease dis-inhibitors? The gay bareback phenomenon. *HIV InSite* website newsletter, 11 November 1997. Available at http://hivinsite.uscf.edu. Accessed June 20, 2003.

26. HIV/AIDS Epidemiological Profile. Oregon and Multnomah County. Oregon Department of Human Services. Prepared for Multnomah HIV Prevention Planning Group. November, 2002.

27. US Centers for Disease Control and Prevention. Update: HIV Counseling and Testing Using Rapid Tests. *MMWR.* March 27, 1998

28. Sloand EM., Pitt E., et al.. HIV testing: State of the art. *JAMA.* 266(20):2861-6, 1991 Nov 27.

29. Mylonakis E., Paliou M., et al. Laboratory testing for infection with the human immunodeficiency virus: established and novel approaches *American Journal of Medicine*.109: 568-576, 2000 Nov.

30. Gallo D., George J., et al. Evaluation of a system using oral mucosal transudate for HIV-1 antibody screening and confirmatory testing. *JAMA*. 277(3):254-258, 1997 Jan.

31. Kosko DA., Becker KL. HIV counseling, testing, and prevention in primary care. *Lippincott's Primary Care Practice.* 4(1):29-39; quiz 40-2, 2000 Jan-Feb.

32. Akaike, H. A new look at the statistical model identification, *IEEE Trans., Auto. Control.* 19: 716-723, 1974.

33. Hosmer, D. and Lemeshow, S. Applied Logisitic Regression, 2nd Edition. New York: John Wiley and Sons, 2000.