DEVELOPING PRACTICAL TOOLS TO INFORM THE ALLOCATION OF LIMITED HIV RESOURCES IN NORTH CAROLINA

Brooke Elizabeth Hoots

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Epidemiology.

Chapel Hill
2011

Approved by:

William C. Miller, MD, PhD, MPH

Michael E. Emch, PhD

Lisa B. Hightow-Weidman, MD, MPH

Peter A. Leone, MD

Pia D. M. MacDonald, PhD, MPH

© 2011
Brooke Elizabeth Hoots
ALL RIGHTS RESERVED

ABSTRACT

BROOKE ELIZABETH HOOTS: Developing Practical Tools to Inform the Allocation of Limited HIV Resources in North Carolina
(Under the direction of William C. Miller, MD, PhD, MPH and Peter A. Leone, MD)

In the current economy, North Carolina (NC) faces a multi-faceted HIV epidemic with limited funding and staff. As state revenue continues to decline, it is imperative that cuts to HIV program resources are based on evidence of where resources are most essential. The purpose of this dissertation was to 1) characterize the geographic distribution of HIV in order to better inform HIV resource allocation, and 2) provide practical tools to aid NC disease intervention specialists (DIS) in prioritizing their HIV partner notification caseloads. Using HIV surveillance data from 2000-2007, we identified highly localized geographic clusters, or core areas, of reported HIV cases in urban areas. These clusters were temporal in addition to spatial in nature and did not persist in the last two years of the study. The disappearance of these clusters was coincident with a dramatic increase in Internet use and distance to sexual partners among men who have sex with men (MSM). Internet-based interventions may therefore be preferable to targeting specific locations. Using DIS interview data from newly diagnosed persons (index cases), we developed a risk score algorithm to predict a sexual partnership between an index case and an undiagnosed HIV-infected partner. We identified five predictive factors—≤ four weeks between diagnosis and DIS interview, no history of crack use, no anonymous sex, fewer partners reported to DIS, and partnerships between an older case and younger partner. While the predictive power of the model was low, it is

possible to reduce the number of partners that need to be located and interviewed while maintaining high sensitivity. We developed and evaluated a second risk score algorithm to predict future violation of NC control measures (failure to disclose HIV status and/or failure to use a condom with a partner) in order to prioritize persons for case management intervention. We identified five predictive factors—identifying as a MSM, younger age, syphilis co-infection, marijuana use in the past year, history of anonymous sex, and reporting two or more sex partners to DIS. Use of this algorithm would facilitate prioritizing case management intervention for those engaging in risky behaviors that perpetuate HIV transmission.

For the people living with HIV and the Disease Intervention Specialists who serve them.

ACKNOWLEDGEMENTS

This work would not have been possible without the constant support and guidance of my committee—Peter Leone, Bill Miller, Mike Emch, Lisa Hightow-Weidman, and Pia MacDonald. In addition to being talented and well-respected researchers and clinicians, they are also incredibly approachable, kind, and generous people that always treated me as a colleague even when I am quite certain I did not merit it.

I met Peter as a student in his infectious diseases course in the fall of 2005, and within one hour of listening to him talk about his work, I knew that I wanted him to be my advisor. Peter makes public health happen. He has shown me how it is possible to be both a realist and an optimist in public health, and how to pursue my professional life and personal life passionately. Despite his insanely busy schedule, he always took the time to sit down with me to catch up and encourage me in my work and my life. He was also instrumental in forging my partnership with the NC Communicable Disease Branch for this dissertation.

I think Bill mentors more students in the Department of Epidemiology than any other professor, so I was hesitant to ask him to chair my committee because he was already dealing with the dissertation drama of at least six other women. He still agreed to take me on, however, and I could not be more grateful for having the chance to work with him. Bill is wise, humble, and calm. I would always leave our meetings with an increased understanding of an epidemiologic method and a decreased sense of panic. He has made me a better writer and a better researcher.

Mike provided invaluable feedback on the geographic analyses for this dissertation. He explained complicated methods with great patience and humor and was always available for my questions even while on sabbatical in New Zealand. I must also thank him for calling in for my defense at 6 am New Zealand time. I first met Lisa when I worked for her in 2007 abstracting a subset of the data that I ultimately ended up using for this dissertation. Lisa is genuine, honest, and always remembers the people behind the data. She was crucial in helping me highlight gaps in my discussion sections. I regret that I didn't officially meet Pia until I asked her to be on my dissertation committee. She is an amazing public health practitioner and was so helpful in guiding me through the proposal and dissertation process. She too made me a better writer by helping me summarize my thoughts more succinctly.

This dissertation would not have been possible without the generous help of the NC Communicable Disease Branch and the HIV/STD Field Services Unit. I am indebted to Evelyn Foust, Michael Hilton, and Del Williams for giving me access to the study data and to Rhonda Ashby and Todd Vanhoy for facilitating my data abstraction at the regional offices. Thank you also to Maxi Mackalo, Kristen Patterson, Albert Saunders, Tama Simms, and the DIS of Regions 3 and 4 for allowing me to invade your space for eight months and for answering all of my questions. I am also grateful to Lani Clough and Molly Fitch for patiently teaching me how to geocode, and to Dianne Enright for allowing me to use her computer software.

I am grateful to have received financial support to do this work from the Infectious Diseases Training Grant. It allowed me to have the time to abstract the study data and to complete my data analysis at the NC Communicable Disease Branch offices. I am also indebted to Nancy Colvin and Carmen Woody in the Department of Epidemiology for their constant support in spite of the fact that I am sure I asked questions that could be answered by the Student Handbook. Thank you also to Vicki Pickard in the Division of Infectious Diseases for always getting me on Peter's schedule and for making sure all of my forms were signed.

I was so lucky to meet a friend on the first day of graduate school that walked beside me through the entire program. Thank you, Anna Dow Sheahan, for always being there. I also could not have made it without the help and wisdom of many other classmates, particularly those in my dissertation support groups (DSG #1: Anna, Jessica Keys, Kim Porter, Kim Powers, Kelly Quinn, and Lizzi Torrone, and DSG#2: Nab Dasgupta, Jenn Griffin, Sarah Radke, and Cindi Snider).

Finally, thank you to my family for supporting me emotionally and often financially in this endeavor. You always told me that I could be anything I wanted and that you would believe in me. Thank you for always holding up your end of the bargain.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIDS	Acquired Immune Deficiency Syndrome
ANOVA	Analysis of Variance
ART	Antiretroviral Therapy
AUC	Area under the ROC curve
CDC	Centers for Disease Control and Prevention
CHC	Community Health Center
CI	Confidence Interval
CMV	Control Measure Violator
DHHS	Department of Health and Human Services
DIS	Disease Intervention Specialist(s)
eHARS	Enhanced HIV/AIDS Reporting System
EMM	Effect Measure Modification
ESRI©	Environmental Systems Research Insititute, Redlands, CA
FN	False Negative
FP	False Positive
GEE	Generalized Estimating Equations
HIV	Human Immunodeficiency Virus
HR	Hazard Ratio
IQR	Interquartile Range
MSM	Men who have Sex with Men
MSM/W	Men who have Sex with Men and Women
NC	North Carolina

NIM	Not in Model
OR	Odds Ratio
PCRS	Partner Counseling and Referral Services
PLWHA	People Living with HIV/AIDS
ref	Referent
RNA	Ribonucleic Acid
ROC	Receiver Operating Characteristic
RS	Risk Score
SAS©	Statistical Analysis Software, Cary, NC
SD	Standard Deviation
STD/STI	Sexually Transmitted Disease/Infection
STD*MIS	Sexually Transmitted Disease Management Information System
UK	United Kingdom
US	United States
Wt	Weight

CHAPTER ONE: SPECIFIC AIMS

In North Carolina (NC), there are approximately 35,000 people living with HIV/AIDS and 1,800 new infections reported annually. Of those living with HIV/AIDS, it is estimated that 31% are unaware of their infection.[1]

Name-based HIV reporting is mandatory in NC. When a positive HIV test result is reported to the state or local health department by a medical provider or clinical laboratory, a disease intervention specialist (DIS) is assigned to investigate and interview the HIV-infected person, or index case. The DIS is responsible for partner elicitation and notification and for providing linkage to care and HIV services for the newly diagnosed person.

In the current economic environment, public health departments are facing funding cuts and hiring freezes. With the current shortage of qualified staff at the state and local health departments, DIS are filling critical gaps in staffing.[2] At the same time, HIV testing in the state is increasing in an effort to identify undiagnosed infections. In the future, DIS will likely face larger caseloads with fewer resources and less time to devote to partner counseling and referral services. It is important to develop tools to help them perform their primary tasks more efficiently. In addition, decisions regarding where to allocate limited HIV resources will require knowledge of the spatial distribution of HIV in the community and an understanding of how sexual partnerships are affecting this spatial distribution.

This dissertation aims to characterize the geographic distribution of HIV in two regions of NC in order to better inform future allocation of HIV resources and to provide practical tools to aid DIS in prioritizing their HIV caseloads.

<u>Specific Aim 1</u>

a. To describe the geospatial distribution of newly diagnosed HIV-infected persons in two regions of NC from January 1, 2000-December 31, 2007.

<u>Hypotheses:</u>

1. Core areas, or clusters, of HIV infection will be present in urban as well as rural areas of NC.
2. Core areas of HIV infection will be less clearly defined in the 2004-2007 time period compared to the 2000-2003 time period.

b. To calculate the geographic distance between HIV-infected persons and their sexual partners and evaluate the effect of Internet use to meet sexual partners on mean distance.

<u>Hypotheses:</u>

1. Average distance between partners in the 2004-2007 time period will be greater than in the 2000-2003 time period.
2. Average distance between partners where the index case uses the Internet to find partners will be greater than that between partners where the index case does not use the Internet.

3. Average distance between male same sex partners will be greater than among heterosexual partners, and this effect will be modified by Internet use.

Overview: We used addresses and demographic information available in DIS charts of newly diagnosed HIV-positive individuals in two regions of NC. Addresses were geocoded and the spatial incidence density of HIV infection was mapped to visualize geographical core areas of infection. We examined the relationship between Internet use to meet sexual partners and mean distance between sexual partners in order to provide an indication of whether geographically-based interventions would be warranted in these areas, or whether alternative approaches, such as targeting Internet sites with HIV prevention messages, may be more effective.

Specific Aim 2

To develop and evaluate a risk score algorithm, using demographic and behavioral characteristics of the index case and sexual partners, to predict a sexual partnership between an index case and an undiagnosed HIV-infected partner.

Hypothesis: It is possible to predict undiagnosed HIV infection among named partners of index cases with reasonable accuracy using selected screening criteria.

Overview: Using demographic and risk behavior data from DIS charts, we developed and evaluated a risk score algorithm using the fewest possible variables to predict undiagnosed HIV infection in named partners. This would be a useful tool for DIS officers and would standardize the method of prioritizing follow-up of partners.

Specific Aim 3

a. To describe index cases who have violated NC HIV control measures.

Hypothesis: Violators of NC HIV control measures will be more likely to belong to racial minorities, groups that are particularly stigmatized by HIV infection.

b. To develop and evaluate a risk score algorithm, using demographic and behavioral characteristics of index cases, to predict future violation of control measures.

Hypothesis: It is possible to predict violation of HIV control measures with reasonable accuracy using selected screening criteria.

Overview: Demographic and risk behavior data from DIS charts were used to develop a second parsimonious risk score algorithm to predict future violation of HIV control measures. This would be useful to DIS officers who could spend more time ensuring that these individuals access HIV care and case management early after diagnosis.

CHAPTER TWO: BACKGROUND AND SIGNIFICANCE

Epidemiology of HIV/AIDS in the Southeastern United States

The southern region of the United States (US), as defined by the US Census Bureau, includes 16 states and the District of Columbia.[3] This region, extending from Delaware to Florida and from the East Coast to Texas, has become the epicenter of the US HIV/AIDS epidemic.[4] By the end of 2007, 40% of the 455,636 persons living with AIDS in the US resided in the southern US.[5] Furthermore, six southern states are among the 15 states with the highest AIDS death rates. While other regions of the US experienced declines in AIDS deaths, the number in the South increased from 2001 to 2005.[5]

The reasons for the HIV/AIDS burden in the South are complex. The South ranks poorly on many health indicators in addition to AIDS incidence.[4] These include overall death rate, heart disease, diabetes prevalence, stroke rate, infant mortality, and preterm birth.[6] Additionally, the southeastern US leads the nation in incidence rates of all reportable sexually transmitted infections (STIs), which increase the likelihood of HIV acquisition and transmission.[4, 7]

There are racial and ethnic disparities in the burden of HIV in the US, with higher rates among African Americans and Hispanics compared to whites.[8] African Americans account for 12% of the population in the US, but constitute approximately 18.5% (and up to 30%) of the southern states' populations.[9] Higher poverty and lack of viable employment, quality education, access to medical care, and health insurance propagate the economic inequalities among African Americans.[4, 10] These factors promote health disparities, including HIV. Higher incarceration rates of African American men also promote concurrent sexual partnerships and discordant sexual mixing patterns between lower-risk African American women and men at higher risk of HIV.[11, 12]

The HIV epidemic in the South is also unique in the fact that high HIV rates are concentrated not only in urban areas, but in rural areas as well.[13] The high proportion of the population in the South living in rural areas often experience difficulty in acquiring quality health care and greater stigma related to HIV infection.[14-16] This complicates efforts to provide HIV prevention and treatment in rural areas.

These factors demonstrate the need to improve our understanding of the epidemiology of the HIV/AIDS epidemic in the South in order to reduce new infections and ensure prevention and treatment for those living with HIV/AIDS.

North Carolina's HIV/AIDS epidemic is characteristic of that seen in the South. Like the rest of the states that constitute the South, NC has a large proportion of the population living in non-metropolitan areas (35%), a high proportion of African Americans (20%), high rates of poverty, and high rates of STIs.[6] There are approximately 35,000 people living with HIV/AIDS in NC and 1,800 new infections reported annually.[1] The number of newly reported HIV infections increased from 2004 to 2007 and has remained stable at a level that is 40% higher than the national level.[17] North Carolina ranked 13th highest among the 50 states in number of reported AIDS cases in 2005.[18] However, when looking at the number of HIV and AIDS cases in rural areas at the end of 2006, NC ranked first and second, respectively, compared to the other states.[13] Women account for a third of HIV cases in NC, and the majority of these women are African American (76%) or Hispanic (7%) and acquired

HIV through heterosexual transmission (96%).[1] Among men, 57% of transmission in NC is attributable to men having sex with men (MSM).[17]

NC faces this multi-faceted epidemic with limited resources. In the current economic environment, NC public health departments are in need of interventions that bring the biggest "bang for the buck"--in this case, interventions that produce the greatest reduction in HIV incidence with the least expenditure of limited resources. Decisions regarding where to allocate limited resources require knowledge of the spatial distribution of HIV in the community and which groups would benefit the most from targeted prevention and treatment.[19]

Core Areas and Partner Selection

An underlying notion of transmission dynamics for HIV and other STIs is that small, cohesive groups of individuals account for a disproportionate amount of transmission.[20] This is known as the "core group" hypothesis. Core groups are often defined by high numbers of sexual partnerships. Observed patterns of sexual partner acquisition reveal that the majority of people form relatively few sexual partnerships while a small minority forms many partnerships.[21] Core groups have also been defined behaviorally, such as MSM, or by occupational risk, such as commercial sex workers or long-distance truck drivers.[22, 23] Alternatively, others have defined core group members as a function of disease incidence.[24] While the attributes that define core group membership are often unspecified, the existence of the core is an important concept in the study of STI epidemiology.[21]

In urban environments, the core is often characterized by geographic compactness. In the early 1980s, Rothenberg demonstrated that gonorrhea incidence was inversely proportional to the physical distance from the core group (Figure 2.1).[25] This finding has been replicated in many inner-city locations with other STIs, including HIV.[26-31] The concept of a "core group" is therefore linked to the notion of a core area or spatial location, which acts as a reservoir for infection for other regions surrounding an urban area.[21]

The core group represents a subgroup of individuals within a sexual network whose behavior assures either the maintenance or the spread of HIV, thus making it an epidemiologic "bull's eye" for preventive approaches.[28, 32] Mathematical models have demonstrated that core group dynamics impact the reproductive rate of an STI.[33-35] The epidemic reproductive rate (R_0) model of infectious diseases has been applied to STI control as $R_0=Bcd$, where B is the STI transmission efficiency between partners, c is the rate of sexual partner change, and d is the duration of infectivity. Within core areas, R_0 is greater than or equal to one, which could allow these areas to function as reservoirs for further disease spread in surrounding communities where R_0 is less than one.[26, 36] A network analysis in Manitoba, Canada identified core and non-core areas and demonstrated bridge events from urban to rural areas of the province.[37] An intervention to reduce HIV incidence in the core area should therefore impact the community-wide disease incidence.[38] The utility of targeting interventions to core groups has been demonstrated for other STIs and suggest that similar methods may be applicable for preventing the spread of HIV.[28]

In Rothenberg's original description of the core area, he noted that contact investigation data of the gonorrhea cases showed that sexual contact tended to exhibit geographic clustering as well.[25] Two studies examining distance between high-risk sexual partners' residences in urban areas have confirmed that partner selection occurs locally (Table 2.1).[39, 40] In Baltimore, the median distance between all sexual partners was 1.7 km and even shorter among partners residing in the core. These results were closer than expected by chance, meaning that individuals were less likely to select sexual partners residing at greater distances from their own residence than those residing locally. Similarly, in Colorado Springs, median distance between sexual partners was 4.3 km, with shorter distances between partnerships where one individual was HIV-infected. Baltimore syphilis cases were also found to travel very short distances to venues in order to meet sexual partners.[41] Another study that examined a sample of sexual partnerships that were not necessarily high risk found that individuals were separated by 15.7 km on average--greater than found for high-risk partnerships, but still relatively close proximity.[42]

In the setting of urban endemic transmission, Rothenberg has proposed that local sexual partnership choices are strongly influenced by the availability of partners and personal mobility. Residents in poor urban areas are less likely to travel widely and form sexual contacts in places outside their residential area.[21] It follows that geographic considerations are important determinants of STI prevalence and infectivity.[43]

Although these studies have documented distance between sexual partners, only one has linked this distance to "neighborhoods." In the Chicago report of the distance between heterosexual partnerships, researchers used the 77 pre-defined neighborhoods, each composed of approximately 40,000 persons, finding that 24% of participants had sex partners within their neighborhood.[42] While these neighborhoods are well-defined and well-known in Chicago, they are large geographic areas and this definition of neighborhood may not transfer to other geographic locations. As many neighborhood studies use smaller geographic areas such as census tracts and block groups,[44] putting geographic distance between sexual partners in the context of smaller, census-defined neighborhoods will increase the understanding of the observed relationship between neighborhoods and STIs.

Studies demonstrating the existence of core areas, or risk spaces, have focused on urban areas; it is therefore unclear if core areas exist in rural areas. The existence of core areas in urban locations has been attributed in part to the high population density in these areas. Core group members in urban areas are at increased likelihood of forming sexual partnerships with other core group members or members of the same sexual networks.[35]

Among rural men, MSM comprise approximately 60% of AIDS cases.[13] Over the years, rural areas, which represent roughly 20% of the US population, have consistently reported 5-8% of all US HIV cases;[45] however, certain rural areas and populations are disproportionately affected—the South in particular. The South comprises 68% of all AIDS cases among rural populations, and in certain areas of the South, the rate of HIV/AIDS diagnoses is almost as high in rural areas as in urban areas.[13, 46]

Before widespread use of the Internet, rural gay and bisexual men traditionally met sex partners in physical locations, such as bars, parks, or bathhouses.[47] Many small communities do not contain gay-identified venues, particularly in conservative areas of the South where high levels of stigma and social hostility persist.[48] This suggests that rural MSM may be accessing gay-identified venues in higher prevalence urban areas. This is supported by a study that found that rural men travel long distances to participate in gay community events and to meet sex partners.[47] Isolation of gay men in rural communities can lead to difficulty finding sexual partners and might lead to riskier behaviors when sexual encounters do occur. One study found that rural men are more likely to have sex on their first date than urban men, possibly due to long travel distances and concern about limited chances for sexual encounters.[48]

The Internet, which offers anonymity and access to an identifiable sex partner pool, is an ideal venue for rural MSM.[49, 50] A study of Internet-using MSM found that they were more likely to be outside major cities and less connected with the urban gay subculture.[51]

The Role of the Internet in Partner Selection

As access has become widespread in the past decade, the Internet has become increasingly popular and successful as a means for meeting potential sex partners.[52] Between one-third and one-half of gay men said they used the Internet to look for sex in recent surveys conducted in the US and UK.[52-61] In a sample of North Carolina men between the ages of 18 and 30 newly diagnosed with HIV infection (72% of whom identified as MSM),[62] 1% of men reported meeting partners over the Internet in 2000, while 26% reported Internet partnerships in 2004.[63] Fewer data are available on the extent of Internet use among populations of predominantly heterosexual men and women. One 2003 survey in London found that 5% of heterosexual women and 10% of heterosexual men had used the Internet to find sexual partners in the previous 12 months.[64]

The Internet allows individuals to meet new sexual partners on the basis of personal selection criteria, therefore enhancing successful meetings and sexual contact.[57] Further, it is possible to select partners with similar interests, particularly similar sexual interests that might have been restricted or hidden in the past.[65] MSM frequently use the Internet to find sexual partners because it offers the "triple A" criteria: accessibility, affordability, and anonymity.[66] In addition to providing these criteria, the Internet also allows MSM to manage and reveal their sexual identities in a way that is comfortable for them.[67, 68] Studies of both rural MSM and black MSM have identified that many are drawn to the Internet out of isolation, fear of rejection, and an inability to find other men like themselves.[48, 69] Specifically, the Internet makes it possible to represent the desirable self, express desires, and manage your own identity, including HIV serostatus.[67] This self-construction and sorting of potential sex partners on the Internet is referred to as "filtering."

It is unclear if geographic distance plays a consistent role in the filtering process that occurs between potential partners on the Internet. The nature of the Internet makes it easy to contact and engage with people who are physically located a long distance from the user.[70] However, if the intention of online dating is to meet up with potential sex partners, geographic location has some impact. In a qualitative study in Australia, some online sex-seeking participants said that the people they would potentially meet needed to live within a certain geographic proximity to their own location. Others did not find distance to be an issue and arranged to meet people when they were traveling on the interstate.[70]

While bars and clubs provide spatial foci for sexual activity in urban environments, online sex-seeking is a medium for distributing sexual activity in space.[67] Because barriers of distance are reduced, the Internet has emerged as a means for linking persons who may not otherwise interact.[57] In the Australian qualitative study, the Internet allowed some people to connect across wide geographic distances and then to meet up for sex. As such, online sex seeking allowed people to extend their sexual networks and to potentially increase their rate of partner change.[70] This may alter the previous observation that core group members in urban areas are at increased likelihood of forming sexual partnerships with other core group members or members of the same sexual networks.[35]

The Internet may be creating a network of high-risk men that facilitates transmission of HIV among online sex seekers, who may spread infection to other partners met at offline venues such as bars and clubs.[52] A number of studies in urban areas of Europe and the US have found that, compared with MSM who do not seek sex on the Internet, those who do are more likely to have had an STI and are more likely to report high-risk sexual behavior such as unprotected anal intercourse.[53, 54, 56-58, 60] If rural men who date online engage in higher risk behaviors than those who date in more traditional venues, as seen in studies of urban MSM, then this behavior may predict an increase in HIV incidence in rural areas. This could lead to the existence of core areas in rural regions.

Partner Notification

North Carolina's Communicable Disease Control Law requires that sexual and needle-sharing partners of HIV-infected individuals be notified that they have potentially come into contact with HIV. In 1989, the NC Department of Health and Human Services (DHHS) began offering partner counseling and referral services (PCRS) to individuals who tested HIV-positive in anonymous testing venues.[71] HIV infections were made reportable to the state in 1990, and confidential name-based reporting replaced anonymous counseling and testing services in 1997. PCRS in North Carolina is conducted by DIS. After obtaining partner information during confidential interviews with the HIV-positive individual, or the index case, the DIS searches confidential public health records to identify partners reported previously with HIV infection and then contacts the remaining partners to inform them they might have been exposed to HIV. All notified partners receive risk-reduction counseling and are offered or referred to HIV testing services.

PCRS encompasses a range of services for HIV-positive individuals and their partners intended to reduce the spread of HIV in communities.[72, 73] Partner notification, also known as contact tracing, is the central activity in PCRS. Partners may be notified by the index case, a process known as "patient referral" or "client referral," or by a public health professional, a method known as "provider referral."[73] Some programs use a mixture of these two approaches and others use a method known as "contract referral," where the index case agrees to notify his or her partners within a certain time period or the provider will step in and complete the process. The second component of PCRS is HIV testing of the named partners. This is followed by counseling of the partners to prevent the further spread of HIV and treatment for those partners newly diagnosed as HIV-positive.

The objectives of partner notification are to 1) identify previously undiagnosed HIV-infected individuals and link them to care and 2) prevent new HIV infections through risk reduction education of notified partners.[74] When used in a population at high-risk for HIV transmission (defined by a high prevalence of HIV), partner notification can identify HIV-positive cases that might otherwise not have been tested.[73] In addition to increasing the number of people tested for HIV, those who are diagnosed with HIV can be counseled on behavior changes to reduce transmission and be referred to care for possible treatment (Figure 2.2).

Systematic reviews of partner notification referral strategies have found that provider referral, although not without problems, is a more effective method than patient referral in ensuring notification and treatment of sexual partners of HIV-positive individuals.[75, 76] The only randomized clinical trial of partner notification method found that 50% of partners in the provider-referral group were notified of their exposure to HIV compared to 7% of partners in the patient-referral group.[77] However, the effectiveness of provider-referral programs is limited by the ability of the index case to recall and willingness to provide accurate partner information.[78, 79] Some problems associated with provider referral in addition to difficulty in locating partners based on information provided by the index cases are cost and labor associated with locating partners and concerns about confidentiality.[80]

North Carolina DIS, who conduct provider referral, have a high rate of success in partner notification. PCRS data from 2001 found that NC DIS interviewed 87% of 1,603 persons newly diagnosed with HIV and elicited an average of 1.1 partners per index case (1,532 injection or sex partners identified total). Twenty-one percent of tested partners had HIV infections that were previously undiagnosed.[71]

How central an HIV-positive person is to a network deeply influences transmission rates in a community. In Colorado Springs, CO, network analysts found that HIV-positive persons had high levels of risk behavior but were located in peripheral areas of risk networks.[81] This network configuration may have explained the relatively low HIV transmission levels. In contrast, HIV-positive persons in New York City, NY occupied central positions within their needle-sharing and sexual risk networks, which helped explain the high observed levels of infection.[82]

Through network analysis, many public health departments have learned to trace “up” the chain of transmission to the HIV transmitter rather than “down” the chain to those infected. This allows transmitters to be identified for treatment and HIV/STD prevention counseling and results in the fragmentation of transmission pathways.[83]

Disease intervention specialists in North Carolina are filling critical gaps in staffing[2] and have less time to conduct provider referral. NC DIS provide PCRS for both HIV and syphilis. While index cases are given the option of notifying partners themselves, DIS are responsible for following up with all named partners. Currently North Carolina has 48 DIS to locate the approximately 1,800 newly identified HIV cases and 600 early syphilis cases in the state per year and their named partners. In 2006, the US Centers for Disease Control and Prevention (CDC) released new recommendations for routine HIV testing in non-traditional healthcare centers, including hospital emergency departments and community health centers, in order to identify the 25% of HIV-positive individuals who are unaware of their status.[84] North Carolina is one of the states receiving CDC grant money to expand HIV testing across the state in hopes of reducing the 40% of HIV-positive individuals in the state who do not know they are infected. With this increased testing, it is expected that the number of newly identified HIV-infected individuals in need of PCRS will increase.

In addition to these PCRS duties, DIS are also used for assignments outside their standard scope of work, including outbreak investigations, disease control and community awareness campaigns, and public health research. They are also incorporated into NC's bioterrorism plans and are legally available to the state epidemiologist as additional personnel should the need for increased field work arise.[2] In the current economic environment, public health departments are facing funding cuts and hiring freezes, making it unlikely that more DIS will be hired to fulfill these responsibilities.

One study calculated the number of index cases that needed to be interviewed in order to identify one newly diagnosed HIV-positive partner overall and by index case characteristics.[74] Index cases that were male, under 40 years old, Hispanic, and recently diagnosed with HIV were more likely to result in a newly diagnosed partner. By analyzing case finding by index case and named partner characteristics, it is possible to guide PCRS program improvement and target partner notification to index cases that are more likely to result in location of additional HIV positive individuals. If DIS are not able to follow-up on all named partners, it would be most beneficial from a public health standpoint to locate those most likely to be HIV-positive.

HIV Prevention for People Living with HIV/AIDS

Advances in antiretroviral therapy have reduced rates of progression to AIDS and death, and improved the quality of life for people living with HIV/AIDS (PLWHA).[85, 86] Because PLWHA are living longer, reducing the risk of transmitting HIV to others is an important aspect of medical care for HIV-infected individuals. Most people with HIV infection want to prevent others from being infected with HIV, but may practice sexual or injection drug behaviors that put others at risk of infection. Studies in the US have found that the overall rate of continued unprotected sexual intercourse is approximately 33% among PLWHA.[87-89]

Until recently, prevention planning shied away from targeting PLWHA because of concerns about stigmatization.[90] This is a missed opportunity to avert new infections and led the Centers for Disease Control and Prevention to place a new focus on “prevention for positives” in the past few years.[91, 92] Given the potentially grave consequences of continued unprotected sexual intercourse among PLWHA, there is an urgent need for effective prevention interventions that promote disclosure of HIV status to sexual partners and increased condom use.

North Carolina Administrative Code requires that individuals diagnosed with HIV follow certain control measures to prevent the spread of HIV to others.[93] These include refraining from sexual intercourse unless condoms are used and notifying future sexual partners of HIV infection. When a DIS interviews a newly diagnosed individual, NC control measures are explained and the individual is asked to sign a document that outlines these control measures saying that he or she will adhere to them. Individuals may refuse to sign this document, but adherence is still legally required. If a previously known HIV-positive individual is named as a sexual partner of a newly diagnosed HIV index case or is reported to the state or local health department with a new STI diagnosis, he or she is considered to be in violation of North Carolina control measures. Criminal prosecution of these individuals is rare, but has occurred in North Carolina.

HIV-related stigma continues to inform perceptions and shape the behavior of PLWHA, thus making it difficult and complex to engage in safe sex. HIV meets four criteria of illnesses that invoke stigma: 1) it is widely perceived to be the infected person's responsibility, 2) it is terminal, 3) it is contagious, and 4) its effects can be outwardly visible.[94] Stigma thus complicates efforts by HIV-infected individuals to have healthy sexual relationships.[95] There is concern that criminalization of HIV may serve as a barrier to HIV prevention if it increases stigma associated with HIV infection rather than deterring behaviors that transmit HIV.[95, 96]

PLWHA in the US are often poor and members of racial or ethnic minority communities with few educational and employment opportunities and high rates of relationship violence.[97-99] Fear of rejection and the risk of violence or ostracism is a major barrier to disclosure, particularly for women living with HIV.[96, 100]

In a study to identify predictors of HIV disclosure to secondary partners among MSM, having fewer sexual partners, being out as an MSM, longer time since HIV diagnosis, knowledge of CD4 count, detectable viral load, and being white were associated with greater disclosure.[101] Disclosure to secondary partners was associated with lower serodiscordant unprotected anal intercourse.

A recent meta-analysis found that individual- and group-level interventions for PLWHA reduced unprotected sex (OR: 0.57, 95% CI: 0.40-0.82) and decreased acquisition of sexually transmitted infections (OR: 0.20, 95% CI: 0.05-0.73).[102] Successful interventions are based on behavioral theory and focus on the challenges of living with HIV, particularly on transmitting the virus to partners and managing stress related to HIV disclosure. An important component is to help PLWHA protect their partners and *themselves* by stressing the importance of decreasing risks to their own health (e.g., contracting other STIs or other strains of HIV that could confer drug resistance).[100]

Because HIV is now more like a chronic disease, the CDC has recommended prevention case management or comprehensive risk counseling and services for PLWHA.[91, 103, 104] This involves a multi-faceted approach of managing medical, mental health, and substance abuse care as well as social services on an individual level.[95] These types of interventions are intensive and are recommended specifically for complex cases in which less intense provider-based or group interventions do not seem sufficient to reduce transmission risk.[105] This approach may be particularly useful for individuals who violate NC control measures, since these individuals are known to be having unprotected sex.

Better linkage to care for PLWHA results in opportunities for prevention counseling and effective treatment. HIV pre-test counseling is no longer required in North Carolina, and post-test counseling of positives is usually completed in a short session at a time of high emotional distress and by a counselor that does not have a relationship with the tester. Alternatively, counseling by primary health care providers can help PLWHA change risky health behaviors.[106] In addition to the potential for behavior change, the reduction in plasma viral load achieved by antiretroviral therapy may decrease the transmission probability to partners.[107] In the University of North Carolina HIV outpatient clinic, 75% of patients had an indication for antiretroviral therapy at their first clinic visit.[108] If it is possible to predict individuals who are more likely to violate HIV control measures, it would be especially important to ensure that those individuals are linked to care to receive prevention counseling and possibly antiretroviral therapy (ART).

Summary

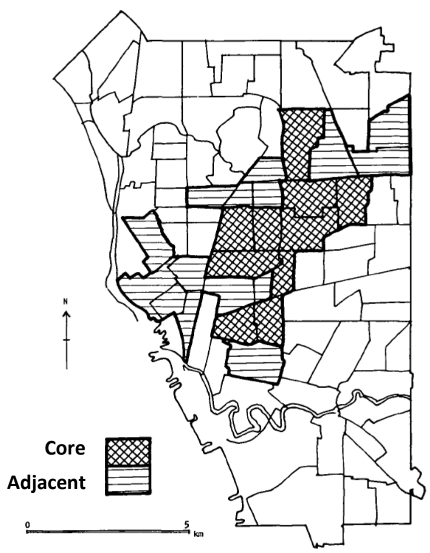
Decisions regarding where to allocate limited resources and target interventions in NC will require knowledge of the spatial distribution of HIV in both rural and urban areas and an understanding of how the Internet is affecting this spatial distribution. In addition, this dissertation aims to provide DIS with practical tools to maximize case finding in partner notification and to predict what index cases are in greatest need of secondary prevention interventions.

TABLE 2.1. Summary of literature on distance between sexual partnerships

Reference	**Location**	**Measurement**	**Distance**
Zenilman et al, 1999[39]	Baltimore, MD	Euclidean distance between partner residences recruited from two Baltimore STD clinics (TRAC study, N=296 dyads)	Median (overall)=1.7 km Median (core males)=0.5 km Median (core females)=0.3 km
Rothenberg et al, 2005[40]	Colorado Springs, CO	Euclidean distance between sexual and drug-using partner and social contact residences of persons at risk for HIV (N=3,982 dyads)	Median (overall)=3.7 km Median (HIV+ partner)=1.3 km Median (sexual partners)=4.3 km
Michaud et al, 2004[41]	Baltimore, MD	Euclidean distance between residences of early syphilis cases and their sex partner meeting venues (N=166 dyads).	Median (overall)=1.7 km
Laumann et al, 2004[42]	Chicago, IL	Euclidean distance between residences of heterosexuals and their most recent sexual partner residing in Cook County, excluding cohabitators (Chicago Health and Social Life Study, N=238 dyads)	Mean (overall)=15.7 km

FIGURE 2.1. Gonorrhea occurrence in Upstate New York, 1975-1980: A. Distribution pattern of core and adjacent census tracts in Buffalo. B. Time trends for gonorrhea case occurrence by census tract classification. Reproduced with permission from Oxford University Press.[25]

A.



B.

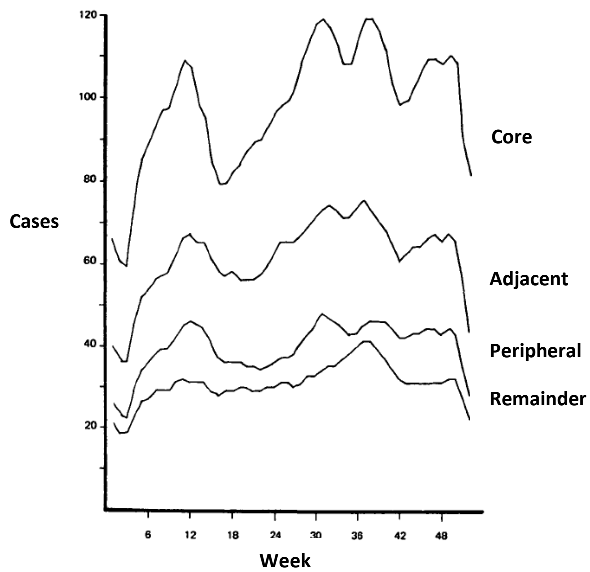
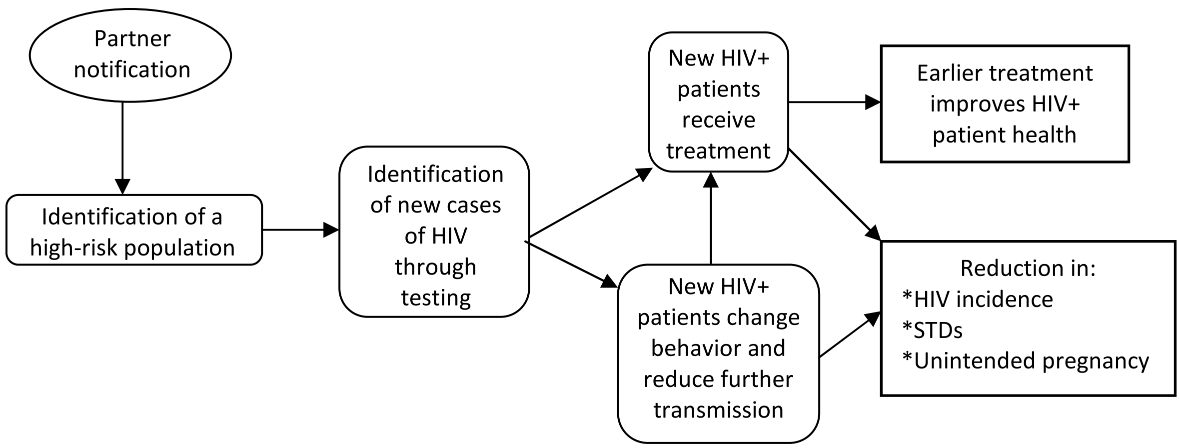


FIGURE 2.2. Analytic framework for partner notification within PCRS. Reproduced with permission from Elsevier.[73]



CHAPTER THREE: DESCRIPTION OF DATA SOURCES

Study Setting

Partner Counseling and Referral Services is completed under the guidance of seven regional offices in North Carolina. These analyses used data from North Carolina state surveillance records of newly reported cases of HIV between 2000 and 2007 in Regions 3 (Winston-Salem Regional Office) and 4 (Raleigh Regional Office). Region 3 includes the following counties: Alamance, Alleghany, Ashe, Caswell, Davidson, Davie, Forsyth, Guilford, Randolph, Rockingham, Stokes, Surry, Watauga, Wilkes, and Yadkin. Region 4 includes Chatham, Durham, Franklin, Granville, Johnston, Lee, Orange, Person, Vance, Wake, Warren, and Wilson counties (Figure 3.1). These two regions encompass approximately 40% of the state's incident HIV cases.

These regions were selected because they both contain several urban areas (Winston-Salem and Greensboro in Region 3 and Durham and Raleigh in Region 4) as well as surrounding rural areas, and because they are adjacent to each other.

Study Population

The study population consisted of individuals newly diagnosed with HIV in the state of North Carolina between January 1, 2000 and December 31, 2007 who are ten years of age or older at the time of HIV diagnosis. NC requires healthcare providers and laboratories to complete a communicable disease report card for each diagnosed case of HIV infection and send it to the local health department within 24 hours of the diagnosis (within seven days of the diagnosis before November 1, 2007) . These cases are then reported to the state health department by local health departments, where they are entered into the STD*MIS (Sexually Transmitted Disease Management Information System) database. Cases are then assigned to a DIS officer. DIS are located throughout the state and perform the initial interviews, confirmatory testing, and referrals to care. If HIV-positive individuals move to North Carolina, they should be reported to the state as new infections in North Carolina when they seek healthcare.

Selection Criteria

Selection criteria for the sub-populations to be included in each aim are described with the research designs for each aim.

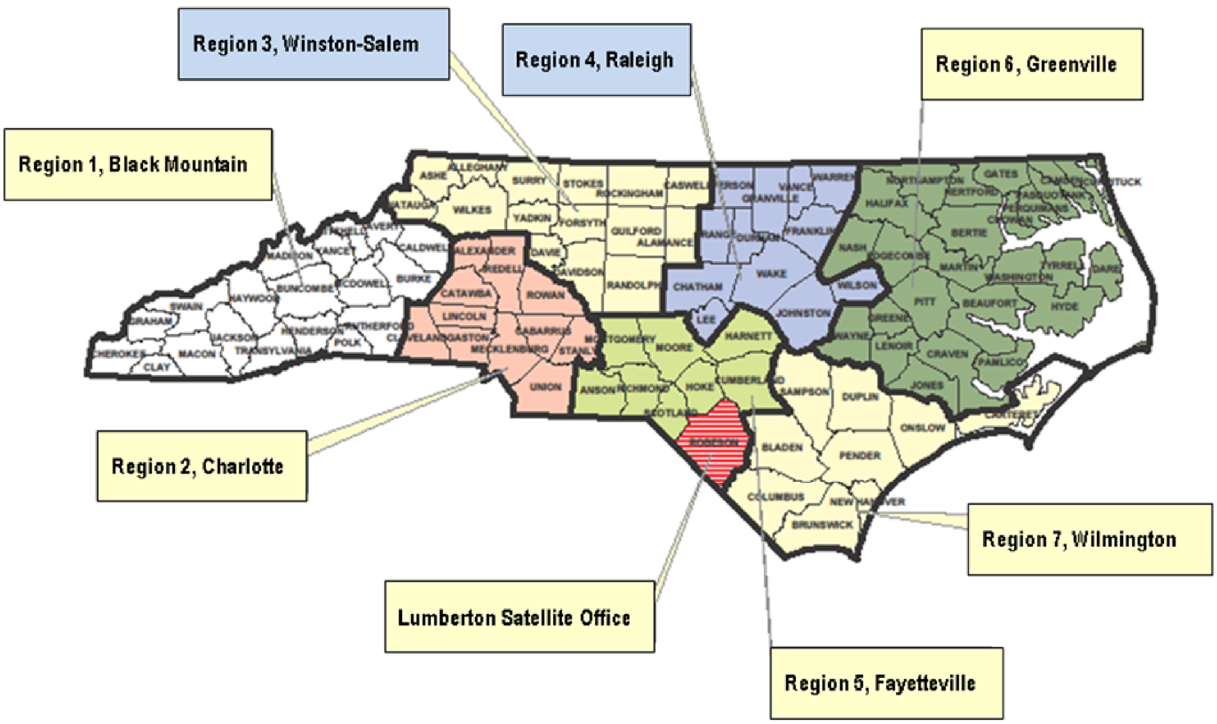
Data Collection

Demographic characteristics of the index case and their sexual partners obtained during partner notification, including date of birth, gender, race/ethnicity, and primary residential address, were available from STD*MIS.

Additional demographic and sexual behavior data, including Internet use to meet sexual partners, were abstracted from Disease Intervention Specialist (DIS) records. DIS keep a chart for each client that contains the STD*MIS entry, their notes on the interviews, and any information from the client's providers that was obtained. These charts are housed at the regional offices in Winston-Salem and Raleigh. Charts are routinely audited by regional supervisors for complete and valid information.

These data were only abstracted for the subset of persons diagnosed with HIV between January 1, 2003 and December 31, 2007 since DIS received training on asking about Internet use during an outbreak investigation of HIV among college students that began in 2003.[62] Data were abstracted using the index case abstraction form in Appendix A and entered into an Access database. Cases were not abstracted if they were unable to be located, refused the DIS interview, aged 10 years or younger, attributable to mother-to-child transmission, or reported no sexual history.

FIGURE 3.1. North Carolina HIV/STD Prevention and Care Branch regions and regional offices



CHAPTER FOUR: METHODS

SPECIFIC AIM 1

a. To describe the geospatial distribution of newly diagnosed HIV-infected persons in two regions of North Carolina from 2000-2007.

b. To calculate the geographic distance between partnerships of HIV-infected persons and evaluate the effect of Internet use to meet sexual partners on mean distance.

Study Design Overview

The goals of this aim were to examine the distribution of core areas of HIV infection in two regions of North Carolina and to examine how Internet use affected distance between partnerships of HIV-infected persons and compactness of core areas. North Carolina was an optimal location to study these research questions because of the rigorous partner notification that exists through the DIS program and because of the high numbers of HIV cases in both rural and urban areas of the state.

Specific Aim 1a: To describe the geospatial distribution of newly diagnosed HIV-infected individuals in two regions of North Carolina from 2000-2007.

Selection Criteria - Aim 1a

Inclusion criteria: All newly reported cases of HIV/AIDS ten years of age or older in North Carolina between January 1, 2000 and December 31, 2007 in Regions 3 and 4 were included in the study.

Exclusion criteria: Index cases who were homeless, had post office boxes, lived on a rural route where the rural route crossed more than one census block group, lived outside the two surveillance regions, had missing or incomplete addresses, or whose addresses failed to match during geocoding were excluded from the analysis.

Data Sources – Aim 1a

Electronic records of index cases' primary residential addresses and limited demographic information from STD*MIS (date of birth, gender, and race/ethnicity) were provided by the North Carolina Division of Public Health. All geocoding and geomasking of the address data took place at the Division's Cooper Building in downtown Raleigh in order to protect the confidentiality of the HIV-infected persons and their partners.

County boundaries, census block group boundaries, total population, percent black, percent urban, and median income were obtained from the 2000 United States Census on the Census Bureau website. Census block group population estimates for 2007 were available from ESRI (Redlands, CA).

Data Analysis - Aim 1a

Geocoding

Case and partner residential addresses were first verified using the US Postal Service address locator (Satori Software, Inc). Addresses were then geocoded in ArcGIS 9.3 (ESRI) using the

NC Integrated Statewide Road Network and county Emergency 911 street databases. College/university addresses without dormitory information were assigned to the geographic center (centroid) of the college/university, and prison/jail addresses were used for individuals currently incarcerated. Rural route addresses were examined to see if the entire rural route was in a single census block group; if so, the address was geocoded to the midpoint of the rural route.

Geomasking

Because we are mapping sensitive health data with high resolution, there was a concern about patient confidentiality. In order to mask the exact address location of the index cases so that the data could be taken outside of the Division of Public Health, we used donut geomasking.[109, 110] In this technique, each geocoded address is relocated in a random direction by at least a minimum distance, but less than a maximum distance, based on population density while retaining the address in its original census block group.

Spatial analysis

Cases were aggregated by census block group. We assigned the spatial incidence density of HIV infection (rate per census block group divided by census block group area) to the centroid of each census block group for mapping. The spatial incidence density is a measure of the number of cases per population per area. Unlike the case density (density = case count/area) and the incidence rate (rate=count/population*time), the spatial incidence density takes the population denominator and the area of case aggregation into account.

To test for the presence and location of clusters, we used Kulldorf's space-time scan statistic in the SaTScan program.[111] This is a local measure that provides information on whether or not cases are clustered and also indicates where and when clustering occurs.[112] The scan statistic is defined by a cylindrical window with a circular geographic base and the height of the cylinder corresponding to time. Each block group was considered the center of a potential cluster or high count of HIV with the radius of the cylinder varying repeatedly from zero up to a set maximum radius to include neighboring block groups, so that the maximum size of the window did not exceed 50% of the total study population. The height reflected any possible time interval of less than or equal to half the total study period, as well as the study period as a whole.[111] The window was then moved in space and time so that for each possible geographic location and size, it also visited each possible time interval.

High-rate clusters were defined as windows where the number of observed cases was greater than the number expected if cases within the window were randomly distributed in space and time, using a discrete Poisson model. The underlying population data were provided for the first and last years of the study period (2000 and 2007, respectively) and SaTScan conducted a linear interpolation to calculate population sizes for each year in between.

A likelihood ratio test statistic was calculated for each potential cluster. *P* values corresponding to the test statistic were calculated using Monte-Carlo simulation. Simulated maps of HIV cases were repeatedly generated (999 times) assuming complete spatial and temporal randomness. Clusters that had a likelihood ratio test statistic in the 95th percentile of the corresponding simulated distribution of likelihood ratio test statistics were considered significant at the 0.05 level.

Cluster detection was first conducted without adjustment for other covariates and was then repeated adjusting for the underlying race distribution (percent black) of the census block groups. Results from the SaTScan analysis were imported back into ArcGIS to generate maps identifying block groups composing significant clusters.

Specific Aim 1b: To calculate the geographic distance between HIV-infected persons and their sexual partners and evaluate the effect of Internet use to meet sexual partners on mean distance.

Selection Criteria - Aim 1b

Inclusion criteria:

- All sexual partnerships of reported cases of HIV/AIDS in North Carolina between January 1, 2000 and December 31, 2007 in Regions 3 and 4 were eligible for analysis.

Exclusion criteria:

- Partnerships where the index case or partner was homeless, had a post office box, lived on a rural route where the rural route crossed more than one census block group, had a missing or incomplete address, or an address that failed to match during geocoding were excluded from analysis.
- Partnerships where the index case lived outside the two surveillance regions were excluded from analysis.
- Partnerships where the index case or partner had a jail or prison address were excluded from analysis.
- Cohabitating partners (distance between partners is zero) were excluded from analysis.

Data Sources - Aim 1b

Primary residential addresses of sexual partners and limited demographic information from STD*MIS (date of birth, gender, and race/ethnicity) were provided by the North Carolina Division of Public Health. Addresses were geocoded as described in Aim 1a. The network, or shortest road, distance (as opposed to the Euclidean, or straight line, distance) between dyads (index case and partner address pairs) was calculated using the Network Analyst in ArcGIS. Exact addresses prior to geomasking were used to calculate distance between partners. Partners who lived together were assigned a distance of zero.

For the linear regression analysis described below, data were restricted to partnerships occurring between January 1, 2003 and December 31, 2007, since Internet use to meet sexual partners and other behavioral characteristics were only abstracted for this sub-population.

Measurements – Aim 1b

Outcome of linear regression model: Mean log distance between an index case and his/her sexual partners. Mean distance to all sexual partners was calculated for each index case. A key assumption of the linear regression model is that the outcome is normally distributed. Because the distance distribution was skewed to the right with a mean much higher than the median, distance was log-transformed to normalize the distribution for the linear regression model.

Main exposure of linear regression model: Internet use to meet sexual partners. The main exposure for this aim was Internet use to meet potential sexual partners. Internet use to meet sexual partners was defined as having used the Internet to meet at least one sexual partner reported to DIS during the partner notification interview. This was coded as "yes," "no," or "undocumented" on the abstraction form. Undocumented responses were coded as "no" for this analysis.

It is unclear when DIS were instructed to ask about Internet use routinely in their client interviews. The North Carolina Field Service Assistant Unit Manager recalled that she became aware of the need to ask about Internet use during DIS interviews from a JAMA article published in 2000 about the Internet as a risk environment for STIs.[58] DIS received training on asking about Internet use during an outbreak investigation of HIV among college students that began in 2003.[62] Therefore, Internet use may be underestimated in the earlier time period. Because of this, only DIS charts for cases diagnosed between January 1, 2003 and December 31, 2007 were abstracted to obtain information on Internet use and other behavioral and sexual HIV risk factors.

Covariates. The covariates considered as effect measure modifiers and confounders in the linear regression model examining the relationship between the main exposure and outcome are presented in Table 4.1.

Data Analysis – Aim 1b

Descriptive statistics. We first performed basic descriptive analyses on distance to sexual partners, including calculating mean distance with standard deviation and median distance with interquartile range (25^{th} and 75^{th} percentiles). We also calculated frequencies of the exposure and covariates.

Bivariable Analysis. We examined the associations of median distance in miles and log mean distance with the covariates in Table 4.1. *P* values for differences in median distances among groups were calculated with the Wilcoxon rank-sum test or Kruskal-Wallis test for covariates with more than two groups. *P* values for the differences in log mean distances among groups were calculated with t-tests or one-way ANOVA for covariates with more than two groups.

To examine whether any of the covariates were unequally distributed between the categories of Internet use, we examined *P* values from the chi-square test or the Fisher's exact test for covariates with an expected cell size less than 10.

Linear Regression. We constructed linear regression models with $\log_{10}$(mean distance between partners) as the outcome. Linear regression takes the form $E(Y_i)=\beta_0 + \beta_p X_{ip}$, where $E(Y_i)$ is the expected response at level i of predictor variable X, β_0 is the intercept (mean when X=0), and β_p is the change in Y_i for a one unit change in X_{ip}.[113]

Assessment of Effect Measure Modification. To assess effect measure modification (EMM), we examined the exposure-outcome relationship while adjusting for one covariate at a time in the linear regression model. The covariate was entered into the model individually as a main effect and as an interaction with the main exposure. If the interaction term was significant at an alpha-level of 0.05 or below, it was retained in the model.

Assessment of Confounding. Potential confounders were considered those covariates that were associated with the exposure and the outcome (among the unexposed) or identified as a confounder on the causal diagram (Figure 4.1), and those covariates that were not found to be effect measure modifiers in the bivariate models. If the potential confounder changed the unadjusted means by more than 10%, it was included for further assessment in the multivariable model.[114]

Multivariable Analysis. To generate adjusted means for the effect of Internet use on the distance between partnerships, we utilized a backwards elimination modeling strategy. The main exposure, covariates, and interaction terms (based on the assessment of EMM in the bivariable analyses) were added to the model, constituting the 'fully adjusted model.' EMM was assessed first. Confounding was examined next. The variable with the highest Wald chi-square *P* value was dropped from the full model. The dropped variable was retained in the model if the estimated mean for either of the Internet use categories changed by more than 10% from the unadjusted association; otherwise it was removed and the model was refit dropping the variable with the next highest Wald chi-square *P* value. This process was repeated with all candidate confounders until a final model was chosen.

Sample Size. Sample size places limits on the number of possible variables that can be included in a multiple linear regression model. A rule of thumb is to have at least 10 observations for predictor:

$$n \geq 10k,$$

where n is the sample size and k is the number of covariates in the model.[113] The highest number of variables that could be included in the model is presented in Table 4.2.

Limitations - Aim 1

Generalizability: These analyses used data from two adjacent regions of North Carolina and may not be generalizable to the rest of the state or other Southern states. However, these analyses were ecological in nature and are expected to provide some insight into clustering and how Internet use affects distance between a subset of partnerships to guide future, group-based interventions.

Residence may not be place of sexual activity: We geocoded home addresses because that was what was available to us. In this analysis, place of sexual activity may also have been of interest. Individuals who found partners on the Internet may have suggested a meeting location that was central to both individuals. Others may have found sexual partners while traveling. If this is true, it is likely that our distance measures overestimated the distance between sexual partners. Data on place of sexual activity would have allowed us to examine how far people were willing to travel to meet partners found on the Internet.

Definition of Internet use: Because we did not have data on which specific partnerships were formed on the Internet, our exposure was defined as an index case using the Internet to meet any sexual partners reported to DIS. Future studies could be strengthened by collecting partnership-level data on Internet use. Partnership-level data would provide more information on whether people are using the Internet to meet partners locally.

SPECIFIC AIM 2

To develop and evaluate a risk score algorithm, using demographic and behavioral characteristics of the index case and sexual partners, to predict a sexual partnership between an index case and an undiagnosed HIV-infected partner.

Study Design Overview

The goal of this aim was to develop a simple, yet effective, algorithm using the fewest possible variables to predict previously undiagnosed HIV infection among named partners. North Carolina's rigorous PCRS program made this an optimal setting to evaluate this aim, as the charts from DIS interviews provide extensive data on index cases and their partners through which to build a predictive model.

Selection Criteria

Inclusion criteria:

- All newly reported cases of HIV/AIDS in North Carolina between January 1, 2003 and December 31, 2007 in Regions 3 and 4 were eligible for analysis.

Exclusion criteria:

- Individuals with newly diagnosed HIV infection in North Carolina (index cases) who were unable to be located by a DIS for notification and referral (lost to follow-up) or those who refused DIS services were excluded from the analysis.

Measurements

Outcome: Newly diagnosed HIV infection among a located partner. The outcome for the predictive model was newly diagnosed HIV infection among a partner of an index case. This was a dichotomous (yes/no) variable where a partner was considered to be a newly diagnosed case of HIV (yes) if the HIV disposition code (Table 4.3) was a 2 (previous negative, new positive) or a 5 (previously untested, new positive). Otherwise, the partner was coded as "no" for the outcome. This category included partners that were HIV-negative or had an unknown HIV status because they could not be located or refused testing. If the partner had an HIV disposition code of K and was interviewed in another region of North Carolina (as opposed to a partner that was out-of-state), we attempted to abstract data on the partner's disposition from the regional office that located the partner.

Covariates. The potential covariates that were used to predict newly diagnosed HIV infection in the partners are presented below in Table 4.4. Potential variables included characteristics of the index case and characteristics of the partnership as reported by the index case.

DIS in NC have a special protocol for follow-up of index cases with acute HIV infection (HIV antibody-negative, RNA-positive cases), giving these cases the highest priority for interviewing and follow-up of partners. Therefore, partners of acute index cases were considered to be definite notifications and were removed from the model building process, but were included in assessment of algorithm performance.

Data Analysis

Bivariable Analysis. We examined the association between each of the candidate predictors in Table 4.3 and the outcome. Candidate variables were eliminated if there was a substantial proportion of missing values (> 5%). Unadjusted odds ratios and 95% confidence intervals as well as chi-square tests were calculated for categorical predictors. Continuous predictors were assessed with t-tests if their distributions were approximately normal. For highly skewed variables, the non-parametric Wilcoxon rank sum test was used. Continuous variables were categorized if the area under the receiver operating characteristic (ROC) curve for the model with only the categorical variable as a predictor did not change or was greater than the area under the ROC curve for the model with only the continuous variable.

Collinearity. We assessed the correlation of each potential predictor variable with every other potential predictor variable to avoid collinearity and potential model convergence problems. All potential predictors were categorical. We used odds ratios to assess collinearity and defined two variables as collinear if the odds ratio was 3 or greater or 0.33 or lower. Variables that were highly correlated with each other were recoded or one of the variables was selected based on the substantive meaning and relationship to other variables.

Sample Size and Number of Predictors. A predictive model's reliability is a function of the prevalence of the outcome in the study population, the total study population, the number of fitted variables in the model, and how well the variables have been measured. To estimate the number of predictors available for modeling, we used the following formula: $(3*n_1*n_2)/10N$, where n_1= the number of persons with the outcome, n_2= the number of persons without the outcome, and N= total number of observations.[115] The maximum number of predictors that could be included in the model is presented in Table 4.2.

Predictive Models.

Logistic regression. Logistic regression takes the form

$$E(Y_i) = (\exp(\beta_0 + \beta_p X_{ip}))/ (1+\exp(\beta_0 + \beta_p X_{ip}))$$

where Y_i is a binary response variable, X_{ip} is a known constant from the *i*th participant, and β_0 and β_p are parameters.[116]

Variable selection for full model. We used unconditional multiple logistic regression with generalized estimating equations to assess the relationship of the predictor variables to the outcome. Generalized estimating equations (GEE) were used to address the lack of independence between index case-partner pairs for persons with multiple partners. The statistical significance of the chi-square test from the bivariable analysis in addition to data from the collinearity analysis guided selection of variables into the full model. All variables were recoded so that they were risk factors for the outcome. Indicator variables were created for all variables with more than two categories. We used a high alpha ($P < .25$) to guide inclusion into the multivariable model to avoid exclusion of important variables that might be excluded if only bivariable analyses were used.[117] This process ensured that only variables with minimal relationship to the outcome were excluded. Interaction terms between all candidate predictors included in the final model were examined and were retained in the model if their Wald P value was $<.25$. We considered these models to be the reference models given that they have the greatest predictive power.

Variable selection for reduced model. Although the full models have greater predictive power, we examined reduced models to see if they had improved model fit. Modeling proceeded in a backward elimination process using a lower alpha ($P < .10$) to eliminate predictors with weak predictive power, starting with interaction terms. Because we used GEE, the likelihood ratio test could not be used to compare models; instead we used the

Wald *P* value to evaluate the effect of removing variables from the model. The Wald *P* value should approximate the likelihood ratio test *P* value when the sample size is large. Changes in the area under the ROC curve were examined to assess the impact of removing each variable from the model and to ensure that the overall predictive accuracy was not significantly reduced. Model fit was evaluated using the Hosmer-Lemeshow test.

Risk Score Development and Testing. We created clinical risk scores using the β-coefficients corresponding to each predictor in the final model. The β-coefficients were summed to create an overall clinical risk score for each patient. Sensitivity and specificity of the model were determined under the assumption that not all partners will be able to be interviewed by the DIS in the future. Internal validity of the resulting model and risk score sensitivity and specificity were examined using bootstrap analysis in which the partner population was resampled 1000 times with replacement.[115]

We compared different cutpoints for additive risk score totals (i.e., over a certain cutpoint, an individual would be located and interviewed). To identify an 'optimal' strategy for prioritizing DIS interviews, we examined the number of misclassification errors that would be made depending on the cutpoint used. A false positive (FP) was defined as locating and interviewing a person who turned out to be HIV-negative, whereas a false negative (FN) was defined as failing to locate and interview a person who was HIV-positive. If DIS were to locate everyone, only false positives would occur, while if DIS were to locate no one, only false negatives would occur. A FN was weighted more than a FP since it would be worse to miss an undiagnosed HIV-infected partner than to locate and test a partner that was HIV-uninfected. The following calculations were made to determine the number of errors associated with the sensitivity and specificity of the model at different cutpoints:

Number of FN = (1 – sensitivity) * HIV prevalence * N

Number of FP = (1 – specificity) * (1 – HIV prevalence) * N

Number of errors = (weight * FN) + FP,

where weight reflects the relative value of a FN compared to a FP.

Sensitivity Analyses. A sensitivity analysis was performed to ensure that unconditional logistic regression with GEE was the appropriate model to use. First, data were analyzed using unconditional logistic regression without accounting for clustering. Second, data were analyzed using logistic regression with generalized estimating equations. Finally, one partnership per case ID was randomly selected into another dataset and data were analyzed using logistic regression. There were no meaningful differences in which variables would have been chosen for inclusion in full models, so logistic regression with GEE was used since it provided slightly wider confidence intervals.

We developed a second model using only index case data to compare the performance of the primary model, which prioritizes particular partnerships, to one that prioritizes interviewing all partners of particular index cases. For this model, the unit of analysis was an index case. Model building procedures were identical to those described above except that the use of GEE was no longer necessary. To assess performance, we weighted the counts by the number of partners reported by the index cases.

DIS do not complete a standard questionnaire when interviewing index cases. Therefore the absence of a risk factor (for example, trading sex for drugs or money) is often not recorded in the case's chart, whereas it most likely would be recorded if present. The abstraction form collects data on yes/no questions as yes/no/undocumented to distinguish documented "no's" from undocumented responses. For analysis, these undocumented responses were collapsed with the "no" category in order to avoid large amounts of missing data. A sensitivity analysis was performed to address the impact of coding the "undocumented" responses as "no." The undocumented responses were instead coded as “yes” to assess the extreme alternative condition. The result of this analysis is described in Appendix B.

Limitations

Lack of data on partner notification costs. We were unfortunately unable to collect data on partner notification costs in these two regions and are therefore unable to demonstrate the cost effectiveness of using a predictive model in this capacity. We are also unable to discuss the tradeoff between resources saved in terms of DIS time and travel with the potential monetary and public health cost of failing to identify a partner with undiagnosed HIV infection. However, many health departments in the US currently provide inconsistent partner notification for HIV.[118] Predictive models used to prioritize partner interviews could ensure that the most undiagnosed HIV infections are identified given the available level of resources available for partner notification.

Generalizability. This analysis used data from two regions of North Carolina and may not be generalizable to the other field service regions in the state or to other states due to varying prevalence of risk factors in different regions. The data used to develop this model are routinely collected and available in the other regions of North Carolina. It may be worthwhile to develop similar models in the other NC regions or to test the sensitivity and specificity of this model on data from other regions. Areas outside of NC may also want to consider development of a model from routinely collected surveillance data.

Use of self-reported behaviors. Our model relies on self-reported data, and we were unable to validate any of the demographics abstracted from the DIS charts. However, the information in the charts is documented by specially trained public health professionals who work closely with patients and providers to provide and document services for patients. Multiple sources, including correctional facility databases, hospital records, and the NC testing database, are used to verify patients' self-reports. Also, the risk score algorithm would use self-reported data if applied in the field.

Missing data. Some index case-partner pairs (<5%) were missing data on time between HIV diagnosis and DIS interview and age difference between partners. These data would also likely be missing in the data if applied in the field. Therefore, we did a complete case analysis.

Validation on the same population. The performance of a predictive model is overestimated when determined on the sample of subjects that was used to construct the model. However, bootstrapping has been shown to result in stable and nearly unbiased estimates of performance when used with a large sample size.[119]

SPECIFIC AIM 3

a. To describe index cases who have violated NC HIV control measures.

b. To develop and evaluate a risk score algorithm, using demographic and behavioral characteristics of index cases, to predict future violation of control measures.

Study Design Overview

The goal of this aim was to describe a subset of index cases who violated North Carolina HIV control measures and to develop a simple, effective algorithm using the fewest possible variables to predict which index cases are likely to violate control measures for use in the future in recommending prevention case management. North Carolina's rigorous PCRS program made this an optimal setting to evaluate this aim, as the charts from DIS interviews provide extensive data on index cases through which to build a predictive model.

Selection Criteria

<u>Inclusion criteria</u>:

- All newly reported cases of HIV/AIDS in North Carolina between January 1, 2003 and December 31, 2007 in Regions 3 and 4 were eligible for analysis.

<u>Exclusion criteria</u>:

- Individuals with newly diagnosed HIV infection in North Carolina's Regions 3 and 4 who were unable to be located by a DIS for notification and referral (lost to follow-up) or those who refused DIS services were excluded from the analysis.

Data Sources

If a previously known HIV-positive individual is named as a sexual partner of a newly diagnosed HIV index case or is reported to the state or local health department with a new sexually transmitted infection diagnosis, he/she is considered to be in violation of North Carolina control measures. When a case is identified as a control measure violator (CMV), DIS create a separate chart with the case's original STD*MIS entry and a document detailing the violation and follow-up. Regions 3 and 4 keep separate files of index cases identified as control measure violators. All CMV files from January 1, 2003 through June 30, 2010 were reviewed at both regional offices. Data on date of violation, type of violation, and whether or not the CMV was reported to their local health department following DIS investigation were abstracted and entered into an Access database.

Measurements

Outcome: Violation of North Carolina HIV control measures. Violation of control measures was defined by a DIS investigation into a person's sexual behaviors following an initial DIS interview after diagnosis. The outcome for the predictive model was time to control measure violation following the initial DIS interview. A time-to-event analysis was used to account for differences in follow-up time for CMV violation after the initial DIS interview. The set of possible predictor variables included demographic characteristics and HIV risk behaviors documented in the original DIS chart.

Additional covariates. Potential predictors of control measure violation for this aim included demographic characteristics (e.g., age, race/ethnicity) and HIV risk factor information (e.g., gender/sexual orientation, drug use, Internet use, bar use). The coding of these variables is described in Table 4 under the heading "Characteristics of the index case" in Table 4.3.

Data Analysis

Descriptive Statistics. We calculated frequencies and percentages of CMVs reported to the health department, filed CMV violations per person, and violation type (e.g., STI diagnosis, partner of an HIV case, or pregnancy). We also calculated the median amount of time between the HIV diagnosis and the violation of control measures.

Bivariable Analysis. We examined the relationship between each predictor variable and the outcome using Cox proportional hazards models to estimate unadjusted hazard ratios and their associated 95% confidence intervals. Candidate variables were eliminated if more than 5% of values were missing.

Collinearity. We also assessed the association between each pair of candidate predictor variables to avoid collinearity. For dichotomous variables, we used an odds ratio to assess collinearity and determined that the two variables were collinear if the odds ratio was 3 or greater. If one variable was continuous and the other was categorical, we examined the magnitude of the difference in means in standardized units. A difference of more than 1.5 standard deviations was considered a strong association.[113] Collinear variables were recoded or one of the variables was selected based on the substantive meaning and relationship to other variables.

Sample Size and Number of Predictors. As described for Aim 2, sample size places limits on the number of possible variables that can be included in a predictive model. We will use the formula described in that section to determine the maximum number of variables that can be included in the predictive model based on the available sample sizes of control measure violators and non-control measure violators. The maximum number of predictors that could be included is presented in Table 4.2.

Predictive Models.

Cox proportional hazards model. Proportional hazards regression models the hazard rate, which is based on the number of events per interval of time. Hazard rates are comparable to incidence rates, but are conditional on survival in the immediately preceding time interval. The model takes the form

$$h_x(t) = h_0(t) * e^{\beta x},$$

where X is a vector of explanatory variables ($x_1, x_2, \ldots, x_k$), $h_0(t)$ is the baseline hazard, and $h_x(t)$ is the hazard at X=x. The interpretation of $e^{\beta x}$ in a multivariable model is the hazard ratio comparing those with x=1 to those with x=0 (referent) at all times *t* adjusted for the other variables in the model. The hazard ratio $e^{\beta x}$ is assumed to be constant across time, meaning that the ratio of the hazard function in the exposed to the hazard function in the unexposed is a fixed constant over time.[120]

Variable selection for full model. Variables for which $p<0.25$ in the bivariable analyses were selected for inclusion in the Cox proportional hazards model.[115] A time-to-event analysis was used to account for differences in follow-up time for CMV violation after the initial DIS interview. Interaction terms between all candidate predictors included in the model were examined and retained in the model if their *P* value was <.25. This model was considered the full, or "reference," model.

Variable selection for reduced model. We examined reduced models to see if they maintained model fit without loss of predictive power. Modeling proceeded in a backward elimination process using a lower alpha ($P < .10$) to eliminate predictors with weak predictive power, starting with interaction terms and then proceeding with the variable with the highest *P* value. Change in the C-statistic was used to assess variations in model performance due to collapsing across categories or removing variables. For binary outcomes, the C-statistic in time-to-event analysis is identical to the area under the ROC curve for logistic regression.[121] The modeling procedures were limited to those persons with complete data for all variables in the reference model. Model fit was evaluated using the Hosmer-Lemeshow test. The proportional hazards assumption was assessed graphically by plotting log-log survival plots. A change in C-statistic less than 0.01 was acceptable between models. The final model had the fewest covariates with minimal reduction in C-statistic and the best model fit.

<u>Risk Score Development and Testing.</u> Risk scores were created as described for Aim 2. A false positive (FP) was defined as choosing to provide additional intervention to an index case who was not going to violate control measures in the future , whereas a false negative (FN) was defined as failing to choose an index case who violates control measures in the future for additional intervention.

A FN was weighted more than a FP since it would be worse to miss a future CMV than to invest prevention resources into a person that did not go on to violate control measures. The following calculations were made to determine the number of errors associated with the sensitivity and specificity of the model at different risk score cutpoints:

Number of FN = (1 – sensitivity) * CMV prevalence * N

Number of FP = (1 – specificity) * (1 – CMV prevalence) * N

Number of errors = (weight * FN) + FP,

where weight reflects the relative value of a FN compared to a FP.

Internal validity of the resulting model and risk score sensitivity and specificity were examined using bootstrap analysis in which the partner population was resampled 1000 times with replacement.[115]

Sensitivity analysis. To examine the effect of possible misclassification of the outcome, persons who were investigated as CMVs but who were not reported to their local health department as CMVs or whose investigation outcome was unknown were recoded as not violating control measures. The final predictive model was re-run with this modified outcome and the C-statistic was examined to assess change in predictive power.

Limitations

Unmeasured covariate--Linkage to care: Linkage to care is frequently not recorded in DIS charts and, if it is, is documented after diagnosis and does not indicate whether a patient remains in care. Because we do not have data on linkage to care in this population, we are unable to discuss the level or intensity of intervention needed to reduce behaviors that lead to control measure violation. It is unknown if the CMVs were successfully linked to care following their initial DIS interview or if they were in care at the time of control measure violation. If non-CMVs were more likely to be in care compared to CMVs, linkage to care and maintenance in care for those indicated by the model may be enough to reduce the incidence of control measure violation. Counseling by primary health care providers can help PLWHA change risky health behaviors.[106] Alternatively, if linkage to care is not associated with CMV status, more intensive interventions may be required for a reduction in the number of CMVs.

Generalizability. This analysis used data from two regions of North Carolina and may not be generalizable to the other field service regions in the state or to other states due to varying prevalence of risk factors in different regions. The data used to develop this model are routinely collected and available in the other regions of North Carolina. It may be worthwhile to develop similar models in the other NC regions or to test the sensitivity and specificity of this model on data from other regions.

Use of self-reported behaviors. Our model relies on self-reported data, and we were unable to validate any of the demographics abstracted from the DIS charts. However, the information in the charts is documented by specially trained public health professionals who work closely with patients and providers to provide and document services for patients. Multiple sources, including correctional facility databases, hospital records, and the NC testing database, are used to verify patients' self-reports. Also, the risk score algorithm would use self-reported data if applied in the field.

Missing data. Some index case-partner pairs (<5%) were missing data on sexual orientation. These data would also likely be missing in the data if applied in the field. Therefore, we did a complete case analysis.

Validation on the same population. The performance of a predictive model is overestimated when determined on the sample of subjects that was used to construct the model. However, bootstrapping has been shown to result in stable and nearly unbiased estimates of performance when used with a large sample size.[119]

LIMITATIONS

The limitations of the proposed methods for the study aims are presented at the conclusion of each research design section. Here certain limitations that apply to several or all of the aims have been repeated.

Limited generalizability outside of these regions of North Carolina: This proposal focuses exclusively on patients with HIV in two regions of NC. Our findings may not be directly applicable to the rest of the state or to other southern states. However, the goals of these analyses are to develop effective intervention tools in these regions that could then be applied to other areas if successful.

Use of self-reported behaviors**:** These analyses use self-reported behaviors that we will be unable to validate. However, the information in the PCRS charts is documented by specially trained public health professionals who work closely with patients and providers to provide and document services for patients. Multiple sources, including correctional facility databases, hospital records, and the NC testing database, are used to verify patients' self-reports. Also, the goal of the predictive models is not to explain or quantify risk behaviors in this population. These data are being used to develop risk score algorithms that will use self-reported data when applied in the field.

TABLE 4.1. Potential covariates to be included in the multiple linear regression model

Covariate	**Coding**	**Notes**
Geographic characteristics		
Index case residence inside cluster	0=No 1=Yes	Significant clusters were identified from the SaTScan analysis in Aim 1a
Urbanicity of census tract where index case resides	0=≤ 25% 1=25.01-50% 2=50.01-75% 3=>75	Collected as a continuous variable, but categorized to present prevalence of Internet use; evaluated using indicator variables
Median income of census tract where index case resides	0=≤ $15,000 1= $15,001-30,000 2= $30,001-45,000 3=> $45,000	Collected as a continuous variable, but categorized to present prevalence of Internet use; evaluated using indicator variables
Characteristics of partnership		
HIV status of partner	0=Negative/unknown status 1=Positive	
Partner is same race/ethnicity as index case	0=No 1=Yes	
Age difference between index case and partner	0=Partner is > 7 years younger 1=Partner is 0-7 years younger 2=Partner is 1-7 years older 3=Partner is > 7 years older	Collected as a continuous variable, but categorized to present prevalence of Internet use; evaluated using indicator variables
Characteristics of index		
Age	0=≤ 24 years 1=25-31 years 2=32-42 years 3=≥ 43 years	Collected as a continuous variable, but categorized to present prevalence of Internet use; evaluated using indicator variables
Black race	0=Non-black 1=Black	
Hispanic ethnicity	0=Non-Hispanic 1=Hispanic	

Covariate	**Coding**	**Notes**
Internet use to meet any sexual partners reported to DIS	0=No 1=Yes	
Bar use to meet any sexual partners reported to DIS	0=No 1=Yes	Only considered as an effect measure modifier, not as a confounder, since on the causal pathway in Figure 4.1
Incarceration history	0=No 1=Yes	
History of hard drug use	0=No 1=Yes	Any drug use excluding marijuana
College or university student	Yes/ No	

TABLE 4.2. Estimated number of variables available for regression models

Aim	Type of model	Formula to determine number of variables	Parameters	Number of variables (k)
1a	Linear regression	$k \leq N/10$	N = 410	41
2	Logistic regression (Predictive model)	$k = (3*n_1*n_2)/10N$	n_1=164 n_2=1,936 N=2,100	45
3	Logistic regression (Predictive model)	$k = (3*n_1*n_2)/10N$	n_1=167 n_2=2,914 N=3,081	47

TABLE 4.3. Explanation of HIV disposition codes used by DIS

HIV Disposition Code	Meaning
1	Previous positive
2	Previous negative, New positive
3	Previous negative, Still negative
4	Previous negative, Not re-tested
5	Not previously tested, New positive
6	Not previously tested, New negative
7	Not previously tested, Not tested now
G	Insufficient information
H	Unable to locate
J	Located, Refused counseling and testing
K	Out of jurisdiction
L	Other

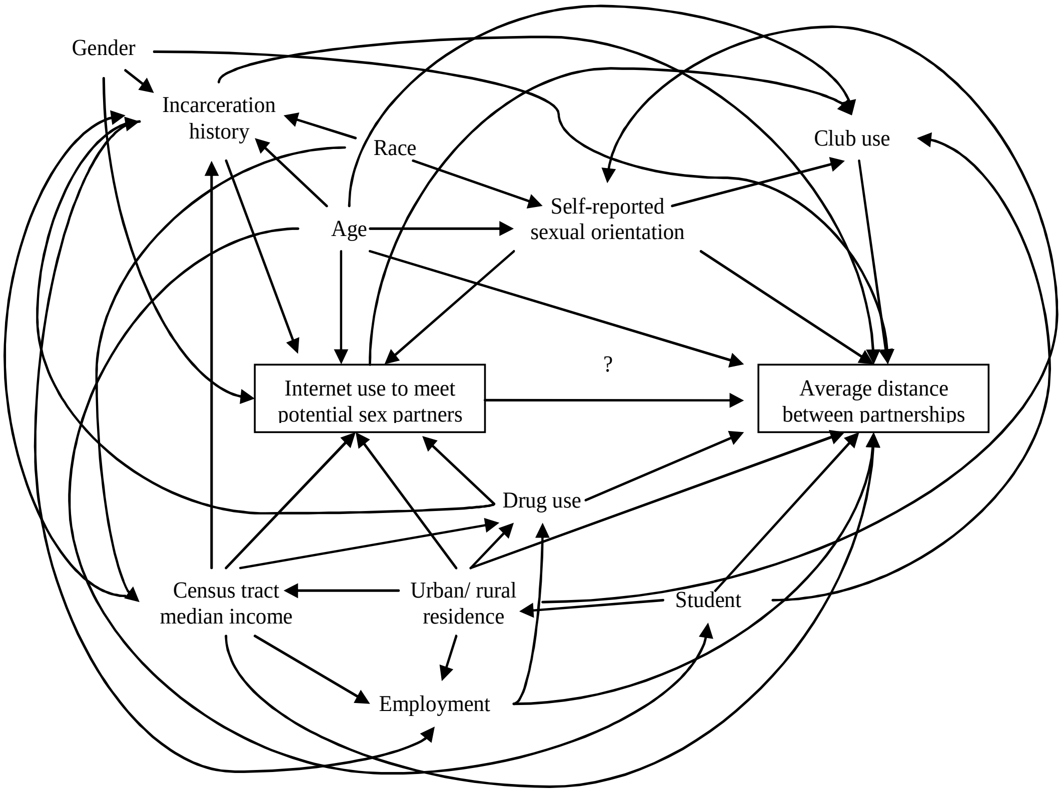
TABLE 4.4. Potential covariates to be included in model to predict undiagnosed HIV infection in named partners of HIV-infected persons

Covariate	**Coding**	**Notes**
Characteristics of the index case		
Old case (time between HIV diagnosis and DIS interview is > 1 year)	0=No 1=Yes	
Stage of infection	0=Chronic 1=AIDS	AIDS defined as CD4 count <200 or CD4 % < 14; acute infections were considered definite notifications and were not included in model building
Race/ethnicity	0=White, non-Hispanic 1=Other/unknown 2=White, Hispanic 3=Black	"Other" includes Asian American/Pacific Islander and Native American; also considered black vs. non-black and Hispanic vs. non-Hispanic
Gender/Sexual orientation	0=Female 1=MSW 2=MSM or MSM/W	This is a composite variable of gender and sexual risk group. Combining these variables allows fewer degrees of freedom to be used.
Immigrated to the US	0=No/unknown 1=Yes	
Incarceration history	0=No/unknown 1=Yes	
Current STI at diagnosis	0=No/unknown 1=Yes	
Syphilis co-infection at diagnosis	0=No/unknown 1=Yes	
History of STI	0=No/unknown 1=Yes	
Alcohol abuse	0=No/unknown 1=Yes	Looked at alcohol abuse (ever) and alcohol abuse in the year prior to diagnosis
Drug use	0=No/unknown 1=Yes	Looked at drug use (ever) and drug use in the year prior to diagnosis

Covariate	Coding	Notes
Hard drug use	0=No/unknown 1=Yes	Any drug use excluding marijuana; looked at hard drug use (ever) and hard drug use in the year prior to diagnosis
Marijuana use	0=No/unknown 1=Yes	Looked at marijuana use (ever) and marijuana use in the year prior to diagnosis
Crack use	0=No/unknown 1=Yes	Looked at crack use (ever) and crack use in the year prior to diagnosis
Injection drug use (IDU)	0=No/unknown 1=Yes	Looked at IDU (ever) and IDU in the year prior to diagnosis
Internet use to meet sexual partners	0=No/unknown 1=Yes	
Bar/club use to meet sexual partners	0=No/unknown 1=Yes	
Victim of rape/sexual assault	0=No/unknown 1=Yes	
Any anonymous sex partners	0=No/unknown 1=Yes	
Exchanged sex for drugs/money	0=No/unknown 1=Yes	
Male bisexual sex partners	0=No/unknown 1=Yes	
Known HIV-positive sex partner	0=No/unknown 1=Yes	
College student	0=No/unknown 1=Yes	
Time to DIS interview	0=≤ 28 days 1=> 29 days	Collected as continuous covariate, but was not linear in the logit
Age of index	Continuous, in years	
Number of sex partners pursued by DIS	Continuous	Also considered as a categorical variable (1 vs. 2-3 vs. ≥ 4 partners)
Characteristics of the partnership		
Same gender partnership	0=No 1=Yes	Reported by index case
Same race partnership	0=No 1=Yes	

Covariate	**Coding**	**Notes**
Age difference	0=partner >6 years older 1=partner 0-6 years older 2=partner 0-6 years younger 3=partner >6 years younger	Also considered as a dichotomous variable (partner is same age or older vs. partner is younger)
Place of meeting	Bar or club/ Internet/ College/ Other	Reported by index case

FIGURE 4.1. Causal diagram of the relationship between Internet use to meet potential sex partners and mean distance between partnerships.



CHAPTER FIVE: Geographic Core Areas, Internet Use, and Distance to Sexual Partners: A Geographic Analysis of HIV Infection in North Carolina

ABSTRACT

Background: The geographic compactness of urban core areas of HIV infection may be reduced by the availability of using the Internet to identify sexual partners. The objectives of this study were to 1) describe the geospatial distribution, including core areas, of newly diagnosed HIV-infected persons in two regions of North Carolina (NC), and 2) to examine factors associated with the geographic distance between partnerships, particularly the effect of Internet use on mean distance. Methods: We mapped the residences of HIV-infected persons and their sexual partners in two multi-county surveillance regions of North Carolina between January 1, 2000 and December 31, 2007. To test for the presence and location of clusters, we used Kulldorff's space-time scan statistic in SaTScan. We examined the association between Internet use and distance to sexual partners among men who have sex with men with a multiple linear regression model. Results: We observed highly localized geographic clustering of reported HIV cases in urban areas, supporting the existence of core areas of HIV transmission in NC. Clustering was temporal in addition to spatial in nature and did not occur after 2005. Internet use among MSM increased dramatically over our study time period, as did distance between sexual partners. The mean distances to sexual partners for MSM who used the Internet and those who did not were 16.6 miles (95% CI: 13.2, 21.4) and 11.5 miles (95% CI: 9.8, 13.5), respectively ($P < .01$). Conclusion: By connecting sexual networks outside of core areas to those of high HIV incidence core areas, Internet use may

have contributed to the lack of clustering identified in the last two years of our study period. The high prevalence of Internet use among MSM and the greater distance to sexual partners observed among Internet users suggest that Internet-based interventions aimed at fragmenting sexual networks may be preferable to interventions targeting specific locations.

INTRODUCTION

Small, cohesive groups of persons are believed to account for a disproportionate amount of transmission of HIV and other sexually transmitted infections (STIs).[20, 28] These groups of persons are referred to as "core groups." In urban environments, the core is often characterized by geographic compactness. In urban areas, gonorrhea incidence is inversely proportional to the physical distance from the core group.[25] These geographical core areas have been identified for other STIs, including HIV infection, in several inner-city locations.[26-31]

An important aspect of the linkage between core groups and geographical core areas, or high HIV incidence clusters, is that partner selection occurs locally in urban areas, maintaining STI incidence in the core area.[25] If core group members form partnerships with persons outside of the core, the core area acts as a reservoir for infection for other regions surrounding the area.[21] The core group has therefore been considered an epidemiologic "bull's eye" for prevention activities.[28, 32] An intervention to reduce HIV incidence in the geographic core area should impact the community-wide disease incidence.[38]

The geographic compactness of urban core areas may be reduced by the availability of the Internet to identify sexual partners. While bars and clubs provide spatial foci for sexual activity in urban environments, online sex-seeking is a medium for distributing sexual activity in space.[67] By allowing people to connect across wide geographic distances, the

Internet has emerged as a means for linking persons who may not otherwise interact, extending their sexual networks.[57, 70] The potential expansion of one's sexual network geographically through the Internet may alter the previous observation that core group members in urban areas are at increased likelihood of forming sexual partnerships with other core group members or members of the same sexual networks.[35] In addition, using the Internet to meet sexual partners could result in the dissolution of urban core areas.

The objectives of this study were to 1) describe the geospatial distribution, including geographic core areas, of newly diagnosed HIV-infected persons in two regions of NC from 2000-2007, and 2) to examine factors associated with the geographic distance between partnerships, particularly the effect of Internet use on mean distance. Examining the distribution of HIV-infected individuals and distance between partnerships will provide an indication of whether geographically-based interventions would be warranted in these areas, or whether alternative approaches, such as targeting Internet sites with HIV prevention messages, may be more effective.

METHODS

Data Collection

NC is divided into seven HIV and STI multi-county surveillance regions. We reviewed the Sexually Transmitted Disease Management Information System (STD*MIS) database from two of these regions (Winston-Salem and Raleigh regional offices) to identify persons in whom HIV infection was diagnosed between January 1, 2000 and December 31, 2007. These two regions include 27 of NC's 100 counties and encompass approximately 40% of the state's incident HIV cases. Primary residential address and race/ethnicity of the index cases and their sexual partners were available from STD*MIS. Data from the 2000 US Census included census block group boundaries, total population, percent black, percent urban, and median income. Census block group population estimates for 2007 were available from ESRI (Redlands, CA).

Additional demographic and sexual behavior data, including Internet use to meet sexual partners, were abstracted from Disease Intervention Specialist (DIS) records. DIS maintain a chart for each index case at the regional office that contains the STD*MIS entry and their notes on the index case and partner notification interviews. These data were only abstracted for the subset of persons diagnosed with HIV between January 1, 2003 and December 31, 2007 since DIS received training on asking about Internet use during an outbreak investigation of HIV infection among college students that began in 2003.[62] Data were abstracted using a standard form and entered into an Access database. Cases were not abstracted if they were aged 10 years or younger, attributable to mother-to-child transmission, or reported no sexual history.

Case and partner residential addresses were first verified using the US Postal Service address locator (Satori Software, Inc.). Addresses were then geocoded in ArcGIS (ESRI) using the NC Integrated Statewide Road Network and county E911 street databases. College/university addresses without dormitory information were assigned to the geographic center (centroid) of the college/university, and prison/jail addresses were used for individuals currently incarcerated. Rural route addresses were examined to see if the entire rural route was in a single census block group; if so, the address was geocoded to the midpoint of the rural route. Cases with residential addresses outside of the two surveillance regions were excluded. The shortest road distance between sexual partner residences was calculated using the Network Analyst in ArcGIS. Partners who lived together were assigned a distance of zero.

Data Analysis

We assigned the spatial incidence density of HIV infection (rate per census block group divided by census block group area) to the centroid of each census block group for mapping. To test for the presence and location of clusters, we used Kulldorf's space-time scan statistic in the SaTScan program. This statistic is defined by a cylindrical window with a circular geographic base and the height of the cylinder corresponding to time. Each block group was considered the center of a potential cluster or high count of HIV with the radius of the cylinder varying repeatedly from zero up to a set maximum radius to include neighboring block groups, so that the maximum size of the window did not exceed 50% of the total study population.

High-rate clusters were defined as windows where the number of observed cases was greater than the number expected if cases within the window were randomly distributed in space and time, using a discrete Poisson model. The underlying population data were provided for the first and last years of the study period (2000 and 2007, respectively) and SaTScan conducted a linear interpolation to calculate population sizes for each year in between.

A likelihood ratio test statistic was calculated for each potential cluster. *P* values corresponding to the test statistic were calculated using Monte-Carlo simulation. Simulated maps of HIV cases were repeatedly generated (999 times) assuming complete spatial and temporal randomness. Clusters that had a likelihood ratio test statistic in the 95^{th} percentile of the corresponding simulated distribution of likelihood ratio test statistics were considered significant at the 0.05 level.

Cluster detection was first conducted without adjustment for other covariates and was then repeated adjusting for the underlying race distribution (percent black) of the census block groups. Results from the SaTScan analysis were imported back into ArcGIS to generate maps identifying block groups composing significant clusters.

We examined median distance in miles to sexual partners by geographic characteristics, partner characteristics, and index case characteristics. Cohabitators (distance=0) and partnerships where one partner resided at a jail or prison were excluded. *P* values for differences in median distances among groups were calculated with the Wilcoxon rank-sum test or Kruskal-Wallis test for covariates with more than two groups. Distance was log-transformed to normalize the distribution and mean distances were compared using t-tests or one-way ANOVA.

Internet use to meet sexual partners was defined as having used the Internet to meet at least one sexual partner reported to DIS during the partner notification interview. Associations between Internet use to meet sexual partners and index case characteristics were examined with chi-square tests. Because 90% of those reporting Internet use to meet sexual partners were men who have sex with men (MSM) or men who have sex with men and women (MSM/W), bivariate analyses were completed among MSM and MSM/W only.

The association between Internet use to meet sexual partners (exposure) and median distance to partners (outcome) was examined with a multiple linear regression model. Effect measure modification (EMM) was assessed by creating an interaction term between the exposure and potential modifier. If the interaction term had a Wald *P* value < .05, it was retained in the model. Candidate confounding variables were related to both the exposure and outcome and not found to be effect measure modifiers. Bar use to meet sexual partners was considered to be a causal intermediate between Internet use and distance since many MSM and MSM/W reported meeting partners initially on the Internet and then arranging to meet them in person at bars. Bar use was therefore not considered as a confounder. The model was reduced using a backwards elimination strategy. Variables were removed from the model if the estimated mean for either of the Internet use categories changed by less than 10% from the unadjusted association, starting with the variable with the highest Wald *P* value.

RESULTS

Of 5,940 HIV index cases recorded in STD*MIS between January 1, 2000 and December 31, 2007, 5,587 (94%) had addresses that could be geocoded within the two surveillance regions. The cases that could not be geocoded included homeless persons (n=20), post office box addresses (n=8), rural route addresses where the rural route crossed more than one census block group (n=4), and addresses that were missing or incomplete (n=321). Thirty percent (n=545) of the census block groups in the study area reported no HIV cases over the eight-year period (Figure 5.1). The mean spatial incidence density was 4.9 cases/1,000/mi^2 (range: 0-414.7 cases/1,000/mi^2), with the greatest spatial incidence densities in urban areas.

Adjusting for the underlying population at risk, six statistically significant ($P < .05$) high HIV incidence clusters (core areas) and one non-statistically significant cluster were identified using SaTScan (Figure 5.2). These core areas comprised 12% (n=221) of the 1,846 block groups analyzed and occurred between 2000 and 2005. The core census block groups combined over the eight-year period were responsible for 13% (n=709) of the HIV cases, with a mean spatial incidence density of 30.4 cases/1,000/mi^2 over the study period. All identified clusters occurred in cities with populations over 50,000. With adjustment for the underlying black population, one cluster (High Point, NC) that occurred between 2001 and 2004 disappeared and another cluster (Durham, NC) that occurred between 2000 and 2003 was reduced in size by approximately half.

Among the 3,994 partnerships where the partner's address could be geocoded (79%), 1,246 (31%) were cohabiting partners, 149 (4%) involved an index case or partner with a jail or prison address, and 24 (0.6%) had a distance that could not be calculated using Network Analyst in ArcGIS. The remaining 2,575 partnerships for 1,535 index cases were analyzed.

Most index cases resided in urban census tracts with median annual incomes of >$30,000 (Table 5.1). More than one-third of partnerships (36%) involved a partner that was also HIV-positive and 15% were discordant with respect to race/ethnicity. Most index cases were black (73%) and 42% were MSM or MSM/W. One quarter had a history of incarceration, 26% had a history of drug use other than marijuana, and 12% were college students.

The overall mean and median distance between partner residences were 19.9 and 7.5 miles, respectively (range: 0.05-276.6 mi). When cohabitators were included, the mean and median distance decreased to 12.0 and 2.8 miles, respectively. The large difference in the mean and median reflect the long tail to the right in the distribution of distances. Median distance between partners increased over the study period, from 6.5 miles in 2000 to 8.9 miles in 2007 (P =.03).

Index cases living inside core areas had significantly shorter distances to partners than cases outside of core areas (Table 5.1). The distance to partners decreased with increasing urbanicity. Cases residing in census tracts with median annual incomes less than $30,000 had the shortest median distance to partners (3.7 mi) (Table 5.1). Median distance differed significantly by age, race/ethnicity, sexual orientation, bar use to meet sexual partners, incarceration history, history of drug and alcohol abuse, and student status.

Persons who found sexual partners on the Internet traveled, on average, more than 16 miles to meet sexual partners. This distance was substantially greater than the distance for persons who did not report use of the Internet to find sexual partners ($P < .001$).

Internet use to meet sexual partners increased linearly over the study time period, from 10% of index cases reporting use in 2003 to 25% reporting use in 2007 (27% of MSM and MSM/W in 2003 to 52% in 2007). Only 8 females (2%) and 9 MSW (4%) reported Internet use to meet a sexual partner (Table 5.2). Among all Internet users, living outside a cluster, younger age, non-black race, non-Hispanic ethnicity, bar use to meet sexual partners, no history of incarceration or hard drug use, and being a college student were associated with increased Internet use to meet sexual partners. After restricting to MSM and MSM/W, only younger age, bar use to meet sexual partners, no history of incarceration, and being a college student were significantly associated with increased Internet use. Urbanicity and median income of the case's census tract were not associated with Internet use in either the entire or restricted groups.

Although MSM and MSM/W who did not use the Internet to meet sexual partners had a larger range of distances to their partners, those who used the Internet had a higher mean and median distance (Figure 5.3). The unadjusted mean distances for MSM and MSM/W who used the Internet and those who did not were 16.2 miles (95% confidence interval (CI): 12.6, 20.4) and 10.5 miles (95% CI: 9.1, 12.3), respectively (Table 5.3). Only history of incarceration remained in the final model as a confounder. After adjustment , the mean distances for MSM and MSM/W who used the Internet and those who did not were 16.6 miles (95% CI: 13.2, 21.4) and 11.5 miles (95% CI: 9.8, 13.5), respectively.

DISCUSSION

We observed highly localized geographic clustering of reported HIV cases in urban areas of two multi-county surveillance regions of NC, supporting the existence of core areas of HIV transmission. Clustering was temporal in addition to spatial in nature and did not occur after 2005. Internet use among MSM and MSM/W increased dramatically over our study time period, as did distance between sexual partners. By connecting sexual networks outside of core areas to those of the high HIV incidence core areas, Internet use may have contributed to the lack of clustering identified in the last two years of our study period.

Our estimates of distance to sexual partners were similar to those observed in a study in Colorado Springs, but are much larger than those reported in two other urban studies. In the Colorado Springs study, which included cohabitating partnerships, the median Euclidian (straight-line) distance between sexual partners was 4.3 km (2.7 miles),[40] which was almost identical to our median distance including cohabitators of 2.8 miles. The median distance including cohabitators observed in Baltimore was lower, at 1.7 km (1.1 miles).[39] In Chicago, the mean Euclidean distance between residences of heterosexuals and their most recent sexual partner excluding cohabitators was 15.7 km (9.8 miles),[42] which was much shorter than our mean distance of 19.9 miles excluding cohabitators. Median distances were not provided in the Chicago study. Our estimates may have been larger than two of the studies due to our inclusion of a much larger study area that encompasses urban and rural settings. Our estimates may be larger, but more accurate, than those found in these studies because of our use of road distance between partners rather than median Euclidian distance.

We found that distance between sexual partners was significantly shorter when the index case resided in the core. However, since all core area census block groups were classified as >75% urban, we were unable to determine if the association between core residence and distance was solely attributable to urbanicity. In the setting of urban endemic transmission, local sexual partnership choices appear to be strongly influenced by the availability of partners and personal mobility.[40] The existence of core areas in urban locations has been attributed in part to the high population density in these areas. Residents in urban areas are less likely to travel widely and form sexual contacts in places outside their residential area.[21]

In contrast, persons living in rural areas, particularly MSM and MSM/W, have fewer identifiable venues in which to meet sexual partners and have been found to travel longer distances to partners. Before widespread use of the Internet, rural gay and bisexual men traditionally met sex partners in physical locations, such as bars, parks, or bathhouses.[47] Many small communities do not contain gay-identified venues, particularly in conservative areas of the South where high levels of stigma and social hostility persist.[48] Consequently, rural MSM may be accessing gay-identified venues in higher prevalence urban areas, given that we observed that rural men travel lengthy distances for partnerships and others have suggested that rural men travel long distances to participate in gay community events and to meet sex partners.[47] Rural MSM and MSM/W were just as likely to use the Internet to meet sexual partners compared to urban MSM and MSM/W in our study.

Because we did not have data on which specific partnerships were formed on the Internet, our exposure was defined as an index case using the Internet to meet any sexual partners reported to DIS. Future studies could be strengthened by collecting partnership-level data on Internet use. Partnership-level data would provide more information on whether people are using the Internet to meet partners locally. Also, our results and interpretations are specific to two multi-county surveillance regions in NC and may not be generalizable to other areas due to varying prevalence of Internet use by location. However, the methods we used may be replicated at other sites to provide further understanding of core areas and the association between Internet use and distance to sexual partners. Finally, MSM who used the Internet in our study were more likely to report partners that could not be located compared to those that did not use the Internet and were therefore more likely to be missing in the multiple linear regression model.

Adjusting for the underlying proportion of black population in these two regions yielded slightly different results in the cluster analysis. High Point, NC was no longer a cluster site and the Durham, NC cluster was reduced in size. These are racially homogeneous areas where the high incidence of HIV may simply be a marker of the high prevalence of blacks, the racial group at highest risk for HIV infection. Adjusting for the underlying black population may have removed clusters with high prevalence of racial concordancy in partnerships. Racially concordant partnerships had a shorter median distance between partners compared to those that were racially discordant, which likely contributed to the existence of racial clusters. A study of core areas of gonorrhea transmission in Baltimore also found differences in detected clusters after adjusting for the race of the underlying population.[122] Clusters detected without adjusting for race were hypothesized to be related to broader sexual networks within the high morbidity population. Network data are needed to further investigate this hypothesis.

The nature of the Internet makes it easy to contact and engage with people who are physically located a long distance from the user.[70] Because the Internet allows persons to meet new sexual partners on the basis of personal selection criteria, successful meetings and sexual contact are enhanced.[57] Consequently, persons may be motivated to travel further in order to have a successful encounter. However, the distance a person is willing to travel is not unlimited. In a qualitative study in Australia, online sex seeking participants said that they would limit their interactions to people living within a certain geographic proximity to their own location.[70]

Online sex seeking allows people to extend their sexual networks and to potentially increase their rate of partner change.[70] Therefore, online sex seeking may alter the previous observation that core group members in urban areas are at increased likelihood of forming sexual partnerships with other core group members or members of the same sexual networks.[35] MSM and MSM/W residing in core areas that use the Internet to meet sexual partners may increase the core's influence on HIV rates in surrounding areas.[123] Alternatively, if online sex seeking increases to the point where people residing in the core are finding more partners at greater geographic distances than locally, core areas may become less persistent. Bridging events between core and non-core areas may have contributed to the lack of clustering identified in the last two years of our study period. Future surveillance data will need to be analyzed for clustering to explore this possibility.

To our knowledge, this is the first study to quantitatively examine the association of Internet use to meet sexual partners with distance between sexual partners. The high prevalence of Internet use among MSM and MSM/W in our sample and the greater distance to sexual partners observed among Internet users suggest that Internet-based interventions aimed at fragmenting sexual networks may be preferable to interventions targeting specific locations. MSM are receptive to Internet-based interventions, such as chat room or message board discussions or educational services.[124-127] Such interventions have been used to increase HIV testing, encourage disclosure to partners, and reduce high-risk behaviors that lead to HIV transmission.[128-131]

While we did identify core urban areas of high HIV prevalence, they did not persist in the later years of the study period. The disappearance of core areas was coincident with increases in Internet use and median distance to sexual partners, suggesting that Internet use may be associated with the dissolution of urban clusters. Further cluster analysis needs to be conducted in NC with more recent surveillance data in order to determine if clusters re-emerged in later years. If geographic core areas did not re-emerge, the NC Communicable Disease Branch may consider eliminating door-to-door, geographic-based HIV testing in these two surveillance regions of the state.

TABLE 5.1. Median distance to sexual partners by geographic, partnership, and index case characteristics among HIV-positive persons reported in two regions of North Carolina, 2000-2007, excluding cohabitators

Characteristic	n (%) Partnerships	n (%) Index cases	Median distance (IQR)	Log mean distance (SD)	*P* value[a]
Geographic characteristics					
Index case residence inside core area					
Yes	337 (13.1)	212 (13.8)	3.7 (1.3, 8.3)	0.59 (0.68)	<.001
No	2,238 (86.9)	1,323 (86.2)	8.0 (3.1, 21.0)	0.87 (0.64)	
Urbanicity of census tract where index resides					
≤ 25%	203 (8.0)	125 (8.2)	16.2 (8.3, 33.0)	1.17 (0.56)	<.001
25.01-50%	96 (3.8)	59 (3.9)	11.8 (5.0, 22.5)	0.96 (0.68)	
50.01-75%	161 (6.3)	89 (5.9)	12.8 (5.6, 25.6)	1.02 (0.60)	
> 75%	2,089 (82.0)	1,244 (82.0)	6.1 (2.3, 17.8)	0.79 (0.64)	
Median income of census tract where index resides					
≤ $15,000	91 (3.6)	52 (3.4)	1.7 (0.9, 3.8)	0.31 (0.58)	<.001
$15,001-30,000	811 (31.8)	511 (33.7)	3.7 (1.5, 10.8)	0.60 (0.67)	
$30,001-45,000	994 (39.0)	588 (38.8)	8.6 (3.6, 23.5)	0.93 (0.62)	
> $45,000	653 (25.6)	366 (24.1)	11.5 (6.1, 22.5)	1.06 (0.54)	
Characteristics of partnership					
HIV status of partner					
Positive	926 (36.0)	--	7.4 (2.9, 20.8)	0.85 (0.64)	.40
Negative/unknown	1,649 (64.0)		7.1 (2.6, 19.9)	0.83 (0.65)	
Partner is same race/ethnicity as index					
Yes	2,107 (85.5)	--	6.8 (2.5, 19.2)	0.81 (0.65)	<.001
No	358 (14.5)		10.0 (3.6, 22.7)	0.95 (0.60)	

Characteristic	n (%) Partnerships	n (%) Index cases	Median distance (IQR)	Log mean distance (SD)	*P* value[a]
Characteristics of index[b]					
Age in years					
≤ 24	--	239 (24.3)	10.3 (4.3, 23.6)	0.98 (0.61)	<.001
25 – 31		169 (17.2)	12.3 (4.8, 22.5)	1.00 (0.53)	
32 – 42		350 (35.5)	8.0 (3.1, 20.0)	0.89 (0.61)	
43+		227 (23.0)	6.6 (1.9, 18.1)	0.74 (0.68)	
Black race					
Yes	--	723 (73.4)	7.5 (2.9, 19.8)	0.86 (0.62)	<.001
No		262 (26.6)	12.4 (5.0, 25.3)	1.01 (0.62)	
Gender/sexual orientation					
Females	--	338 (34.6)	6.2 (2.0, 16.0)	0.75 (0.67)	<.001
MSW		229 (23.4)	5.3 (2.2, 13.0)	0.73 (0.60)	
MSM or MSM/W		410 (42.0)	14.1 (6.3, 29.2)	1.10 (0.53)	
Internet use to meet sex partners					
Yes	--	200 (20.3)	16.3 (8.3, 29.1)	1.17 (0.48)	<.001
No		785 (79.7)	7.1 (2.6, 18.7)	0.83 (0.63	
Bar use to meet sex partners					
Yes	--	257 (26.1)	12.3 (5.4, 28.6)	1.07 (0.53)	<.001
No		728 (73.9)	7.4 (2.7, 18.3)	0.84 (0.64)	
Incarceration history					
Yes	--	246 (25.0)	5.4 (2.2, 16.4)	0.75 (0.64)	<.001
No		739 (75.0)	9.6 (3.8, 22.6)	0.95 (0.61)	
History of hard drug use (excludes marijuana)					
Yes	--	253 (25.7)	5.6 (1.9, 20.6)	0.78 (0.68)	<.001
No		732 (74.3)	9.6 (3.9, 21.4)	0.94 (0.59)	

Characteristic	n (%) Partnerships	n (%) Index cases	Median distance (IQR)	Log mean distance (SD)	*P* value[a]
College student					
Yes	--	115 (11.7)	11.8 (5.6, 33.9)	1.11 (0.57)	<.001
No		870 (88.3)	8.2 (3.0, 20.0)	0.87 (0.62)	

Abbreviations: DHHS, Department of Health and Human Services; IQR, interquartile range; MSW, men who have sex with women; MSM, men who have sex with men; MSM/W, men who have sex with men and women; NC, North Carolina; SD, standard deviation

[a] *P* values were calculated using the Wilcoxon rank sum test or the Kruskal-Wallis test

[b] Restricted to HIV-positive persons reported to NC DHHS between 2003 and 2007

TABLE 5.2. Demographic and behavioral characteristics of HIV-positive persons and HIV-positive MSM and MSM/W who used the Internet to meet sexual partners in two HIV surveillance regions of North Carolina, 2003-2007

Characteristic	Internet users All n (%)	*P* value[a]	Internet users MSM and MSM/W only n (%)	*P* value[a]
Overall	200 (20.3)		180 (43.9)	
Gender/sexual orientation				
Females	8 (2.4)	<.001	--	
MSW	9 (3.9)			
MSM or MSM/W	180 (43.9)			
Index case residence inside core area				
Yes	10 (11.4)	.03	9 (36.0)	.41
No	190 (21.2)		171 (44.4)	
Urbanicity of census tract				
≤ 25%	50 (19.7)	.13	48 (43.2)	.66
25.01-50%	6 (9.7)		4 (28.6)	
50.01-75%	18 (19.2)		17 (43.6)	
> 75	126 (22.1)		111 (45.5)	
Median income of census tract				
≤ $15,000	4 (16.0)	.75	4 (44.4)	.76
$15,001-30,000	47 (21.9)		40 (46.5)	
$30,001-45,000	99 (19.3)		89 (41.4)	
> $45,000	50 (21.7)		47 (47.0)	
Age in years				
≤ 24	78 (32.6)	<.001	70 (49.3)	.04
25 – 31	47 (27.8)		41 (52.6)	
32 – 42	52 (14.9)		47 (35.9)	
43+	23 (10.1)		22 (37.3)	
Black race				
Yes	119 (16.5)	<.001	101 (40.9)	.13
No	81 (30.9)		79 (48.5)	
Bar use to meet sex partners				
Yes	110 (42.8)	<.001	104 (50.2)	<.01
No	90 (12.4)		76 (37.4)	
Incarceration history				
Yes	19 (7.7)	<.001	14 (28.6)	.02
No	181 (24.5)		166 (46.0)	
History of hard drug use (excludes marijuana)				
Yes	26 (10.3)	<.001	25 (37.9)	.28
No	174 (23.8)		155 (45.1)	

Characteristic	Internet users All n (%)		*P* value[a]	Internet users MSM and MSM/W only n (%)		*P* value[a]
College student						
Yes	50	(43.5)	<.001	46	(56.8)	<.01
No	150	(17.2)		134	(40.7)	

Abbreviations: MSW, men who have sex with women; MSM, men who have sex with men; MSM/W, men who have sex with men and women

[a] *P* values were calculated using the chi-square statistic.

TABLE 5.3. β-coefficients and mean distance to sexual partners by Internet use from linear regression model among HIV-infected MSM and MSM/W in two HIV surveillance regions in North Carolina, 2003-2007

	Unadjusted value (95% CI)	Adjusted value [a] (95% CI)
Model β-coeffecients (log_{10} miles)		
Intercept	1.02 (0.96, 1.09)	1.06 (0.99, 1.13)
Internet use to meet sex partners	0.18 (0.08, 0.28)	0.17 (0.06, 0.27)
History of incarceration	--	-0.24 (-0.39, -0.08)
Mean distance to sexual partners (miles)		
Internet use to meet sex partners		
Yes	16.2 (12.6, 20.4)	16.6 (13.2, 21.4)
No	10.5 (9.1, 12.3)	11.5 (9.8, 13.5)
P value (difference)	<.001	<.01

[a] Adjusted for history of incarceration
Abbreviations: CI, confidence interval; DHHS, Department of Health and Human Services; MSW, men who have sex with women; MSM, men who have sex with men; MSM/W, men who have sex with men and women; NC, North Carolina

FIGURE 5.1 Map of spatial incidence density of HIV (cases/1,000/square mile) for two HIV/AIDS multi-county surveillance regions of North Carolina, 2000-2007

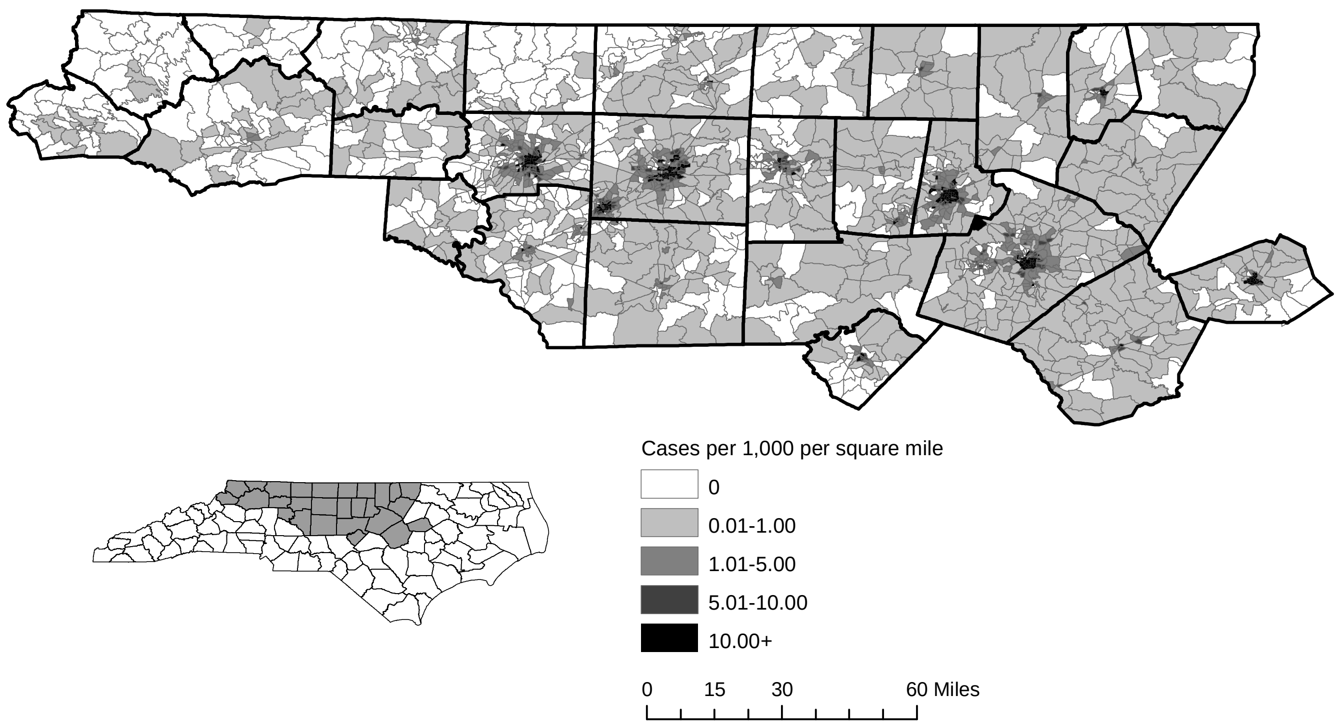


FIGURE 5.2. Census block groups comprising clusters of reported HIV cases in two HIV surveillance regions of North Carolina, 2000-2007. Most likely cluster: (1) Raleigh, 2002-2005 (*P* < .001), Secondary clusters: (2) Winston-Salem, 2001-2004 (*P* < .001), (3) Durham, 2000-2003 (*P* < .001), (4) Greensboro, 2000-2003 (*P* < .001), (5) High Point, 2001-2004 (*P* < .001), (6) Wilson, 2000-2002 (*P* < .001), (7) Sanford, 2003-2004 (*P* =.90)

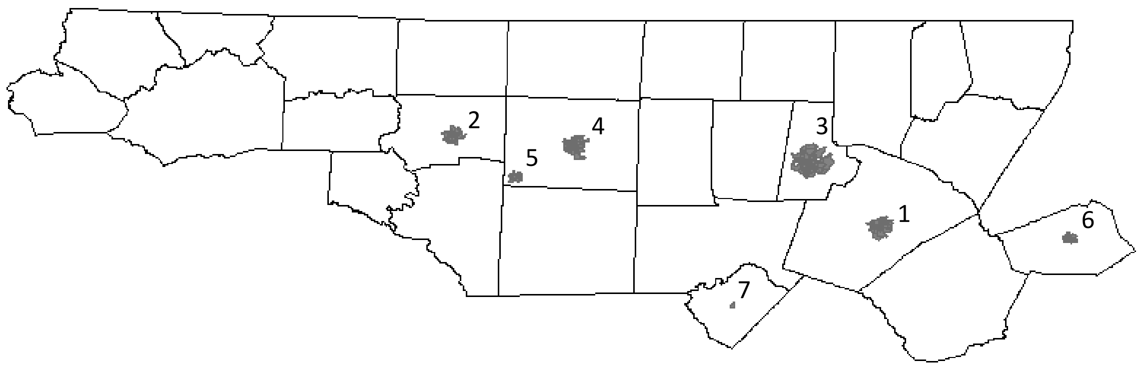
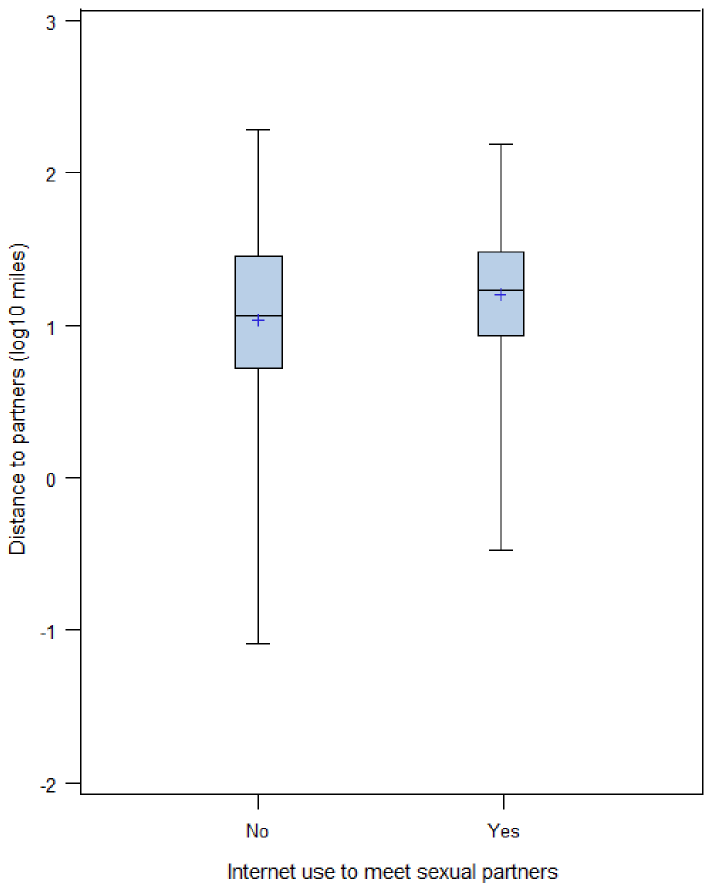


FIGURE 5.3. Boxplot of distance to sexual partners by Internet use to meet sexual partners among HIV-infected persons in two HIV surveillance regions of North Carolina, 2003-2007



CHAPTER SIX: Developing a Predictive Model to Prioritize HIV Partner Notification in North Carolina

ABSTRACT

Background: Disease Intervention Specialists (DIS) in North Carolina have less time to conduct partner notification due to competing responsibilities while simultaneously facing increased case loads due to increased HIV testing. We developed a model to predict undiagnosed HIV infection in sexual partners to aid DIS in prioritizing interviews. Methods: We abstracted demographic, behavioral, and partnership data from DIS records of HIV-positive persons reported in North Carolina between January 1, 2003 and December 31, 2007. Multiple unconditional logistic regression with generalized estimating equations was used to develop a predictive model and applicable clinical risk scores. The sensitivities and specificities of the risk scores at different cutoffs were used to examine algorithm performance. Results: We identified five factors that predict a sexual partnership between a person with newly diagnosed HIV infection and an undiagnosed partner—four weeks or fewer between HIV diagnosis and DIS interview, no history of crack use, no report of anonymous sex, fewer sexual partners reported to DIS, and sexual partnerships between an older index case and younger partner. Using this model, DIS could identify 90.2% of undiagnosed HIV infection in partners while reducing the number of partners pursued by 25%. Conclusions: While the overall predictive power of the model is low, it is possible to reduce the number of partners that need to be located and interviewed while maintaining high

sensitivity. If DIS continue to pursue all partners, the model would be useful in identifying partners in which to invest more resources for locating and testing.

INTRODUCTION

Partner notification is an established component of public health efforts to control the transmission of sexually transmitted infections (STI), including human immunodeficiency virus (HIV) infection. The objectives of partner notification are to 1) identify previously undiagnosed HIV-infected persons and link them to care, and 2) prevent new cases of HIV infection through risk reduction education of notified partners.[72] When used in a population with a high prevalence of HIV, partner notification leads to identification of HIV-infected persons that might otherwise not have been tested.[132]

Partners may be notified by the HIV-infected index case, a process known as "patient referral," or by a public health professional, a method known as "provider referral."[132] Generally, provider referral is a more effective method than patient referral in ensuring notification of sexual partners of HIV-infected individuals (3,4).[75, 76] However, the effectiveness of provider-referral programs is limited by cost and the labor associated with locating and interviewing partners (5).[80]

In North Carolina (NC), disease intervention specialists (DIS) conduct provider referral for both HIV and syphilis. Currently, 48 DIS are available to locate ~2,000 newly identified HIV cases and ~600 early syphilis cases in the state per year as well as their named sexual and drug sharing partners. Over the last few years, HIV testing efforts in the state have increased in an attempt to identify the estimated 35% of persons with undiagnosed infection.[17] With increased testing, partner notification demand has also increased.

DIS are also used for assignments outside their standard scope of work, such as community awareness campaigns and public health research, leaving less time for their traditional partner notification duties.[2] In the current economic environment, public health departments are facing budget cuts and hiring freezes, making it unlikely that more DIS will be hired to fulfill these responsibilities. If DIS are unable to trace all named partners in the future, identifying those partners most likely to be HIV-infected would be a potentially effective strategy.

In 2008, the Centers for Disease Control and Prevention (CDC) released updated partner notification guidelines for HIV and STIs that emphasize the need for setting-specific, evidence-based partner services programs.[133] A risk score algorithm based on local data to prioritize interviews and standardize partner follow-up of named partners of index cases might improve DIS efficiency. While risk scores have not been utilized by partner services programs specifically, they have been shown to successfully increase efficiency and cost-effectiveness of STD case finding activities in the past.[134-141] Using demographic and behavioral characteristics of both the index cases and named partners from DIS records, we developed and evaluated risk scores to predict undiagnosed HIV infection in named sexual partners of newly diagnosed HIV-infected persons in NC.

METHODS

Data Collection

NC is divided into seven HIV and STI surveillance regions. We reviewed the Sexually Transmitted Disease Management Information System (STD*MIS) database from two of these regions (Winston-Salem and Raleigh regional offices) to identify persons in whom HIV was diagnosed between January 1, 2003 and December 31, 2007. These two regions include 27 of NC's 100 counties and encompass approximately 40% of the state's incident HIV cases. DIS maintain a chart for each index case at the regional office that contains the STD*MIS entry and their notes on the interviews with the index and partners. Demographic, sexual behavior, and partner data were abstracted using a standard form and entered into an Access database. Cases were not abstracted if they were aged 10 years or younger, attributable to mother-to-child transmission, or reported no sexual history. Cases were excluded from analysis if they were unable to be located or refused the DIS interview. Sexual partners named by the index cases were excluded from analysis if they were previously diagnosed with HIV. The unit of analysis was an index-partner pair.

Data Analysis

The outcome was newly diagnosed HIV infection in a sexual partner. The set of possible predictor variables included demographic characteristics and risk behaviors of the index case, demographic characteristics of the named partner, and characteristics of the partnership reported by the index case. DIS in NC have a special protocol for follow-up of index cases with acute HIV infection (HIV antibody-negative, RNA-positive cases), giving these cases the highest priority for interviewing and follow-up of partners. Therefore, partners of acute index cases were considered to be definite notifications and were removed from the model building process, but were included in assessment of algorithm performance.

Generalized estimating equations were used to address the lack of independence between index case-partner pairs for persons with multiple partners. We examined the association between each predictor variable and the outcome using unadjusted prevalence odds ratios with associated 95% confidence intervals. We assessed the association between each pair of candidate predictor variables to avoid collinearity. Collinear variables were either recoded or one of the variables was selected based on the substantive meaning and relationship to other variables. Variables for which $p<0.25$ in the bivariable analyses were selected for inclusion in the multiple unconditional logistic regression model.[115] We assessed interaction terms between all candidate predictors included in the model and retained interaction terms with P values <.25. This model was considered the full, or ‘reference’, model.

We examined reduced models to see if they had adequate model fit without loss of predictive power. Modeling proceeded in a backward elimination process using a lower alpha ($P < .10$) to eliminate predictors with weak predictive power, starting with interaction terms and then proceeding with the variable with the highest P value. Changes in the area under the receiver operating characteristic (ROC) curve were used to assess variations in model performance due to collapsing across categories or removing variables. Model fit was evaluated using the Hosmer-Lemeshow test. The modeling procedures were limited to those persons with complete data for all variables in the model.

We created clinical risk scores using the β-coefficients corresponding to each predictor in the final model. The β-coefficients were summed to create an overall clinical risk score for each patient. We used 1,000 bootstrap samples with replacement to validate our model and risk score performance.

To identify an ‘optimal’ strategy for prioritizing DIS interviews, we examined the number of misclassification errors that would be made depending on the cutpoint used for additive risk score totals (i.e., over a certain cutpoint, a partner would be located and interviewed). A false positive (FP) was defined as interviewing a partner who turns out to be HIV-uninfected, whereas a false negative (FN) was defined as failing to interview a partner with undiagnosed HIV.

A FN was weighted more than a FP since it would be worse to miss an undiagnosed HIV-infected partner than to locate and test a partner that was HIV-uninfected. The following calculations were made to determine the number of errors associated with the sensitivity and specificity of the model at different risk score cutpoints:

Number of FN = (1 – sensitivity) * HIV prevalence among tested partners * N

Number of FP = (1 – specificity) * (1 – HIV prevalence among tested partners) * N

Number of errors = (weight * FN) + FP,

where weight reflects the relative value of a FN compared to a FP.

We developed a second model using only index case data to compare the performance of the above model, which prioritizes particular partnerships, to one that prioritizes interviewing all partners of particular index cases. For this model, the unit of analysis was an index case. Model building procedures were identical to those described above except that the use of GEE was no longer necessary. To assess performance, we weighted the counts by the number of partners reported by the index cases.

All analyses were conducted using SAS version 9.2 (Cary, NC).

RESULTS

A total of 3,880 index cases from the two surveillance regions were diagnosed with HIV infection and recorded in STD*MIS between January 1, 2003 and December 31, 2007 (Figure 6.1). DIS interviewed 81.3% of eligible cases. Over half of these cases (61%)

reported one or more partners to DIS for follow-up. Almost one-third of the partnerships (31.1%) involved a previously known HIV-infected partner, leaving 2,232 index-partner pairs for analysis. Approximately 42% of these pairs involved a partner that was unable to be located or refused testing.

Overall, 171 index-partner pairs (7.7%) had a partner that was newly diagnosed with HIV. DIS interviewed 18.8 index cases to identify one partner newly diagnosed with HIV. Most of the index cases in the index-partner pairs were male (68.3%) and black (66.0%) (Table 6.1). They were also in the chronic stage of HIV infection (78.3%), with only 6.1% of cases acutely infected with HIV and 15.6% identified as AIDS cases (CD4 count or percent < 200 cells/μL or 14%, respectively, or diagnosis with an AIDS-defining illness). The median age of the index cases in the pairs was 33 years (range: 15-68 years). The partner was younger than the index case in 41.0% of the index-partner pairs, and 45.1% were same gender partnerships. Thirteen percent of partnerships were between persons of different races or ethnicities.

Reporting only one partner to DIS compared to reporting four or more partners was the predictor most strongly associated with a newly diagnosed HIV-infected partner (Table 6.1). The odds of having a newly positive partner for those who reported only one partner were 2.68 times those of having a newly positive partner for those who reported four or more partners (95% confidence interval (CI): 1.64, 4.38). Index cases with acute HIV infection were less likely to have a newly diagnosed HIV-infected partner compared to those with chronic HIV infection (odds ratio (OR) 0.39, 95% CI: 0.14, 1.08). Other potentially important predictors of a newly diagnosed HIV-infected partner ($P < .05$ in bivariate analyses)were no history of crack use, no anonymous sex, exchanging sex for drugs or money, fewer than 4 weeks between time of HIV diagnosis and DIS interview, and having a younger partner. Hispanic ethnicity, having immigrated to the US, no incarceration history, HIV diagnosis at a community health center or health department, having a bisexual sex partner, heterosexual partnerships, and same race partnerships were also candidate predictors ($0.05 < p < 0.25$ in bivariate analyses) for the reference model.

Stage of infection was not a candidate predictor in the reference model since all acutely infected index cases are prioritized in the partner notification algorithm (Figure 6.2). Other candidate predictors were excluded from the reference model due to collinearity. Non-Hispanic ethnicity, being a native of the US, history of incarceration, and exchanging sex for drugs or money were highly correlated with crack use. Exchanging sex for drugs or money was also highly correlated with anonymous sex, as was same gender partnership. Therefore, these variables were excluded.

The reference model included time between HIV diagnosis and DIS interview, diagnosis location, history of crack use, history of anonymous sex, bisexual sex partner, number of partners reported to DIS, age difference between partners, and same race partnership. The relationship between crack use and undiagnosed HIV infection varied by the age difference between the index case and partner, so an interaction term between these variables was included (Table 2). The area under the ROC curve was 0.666 (95% CI: 0.619, 0.712) for this model.

After model simplification, the final model included six terms—time between HIV diagnosis and DIS interview, crack use, anonymous sex, number of sex partners pursued, age difference between partners, and the interaction between crack use and age difference (Table 6.2). The area under the ROC curve was 0.662 (95% CI: 0.619, 0.704).

The risk score for a partnership is equal to the sum of the predictors' β-coefficients. For the final model that included partnership characteristics, the risk scores ranged from zero to 3.46 for an index case that was interviewed within four weeks of diagnosis (+0.55) with no history of crack use (+1.37) or anonymous sex (+0.56) who reported one partnership to DIS (0.98) with a younger partner (+1.27 and -1.14 for the interaction term) (Table 6.2).

The overall predictive power of the model was low, as indicated by the low value for the area under the ROC curve. In order to maintain a high sensitivity, only relatively small reductions in partners pursued can be made. Using a lower risk score cutpoint (e.g., 1.00 or 1.50) entails interviewing a larger proportion of partners. Consequently, more partners who actually have undiagnosed HIV infection would be interviewed and tested, resulting in fewer false negatives. Interviewing all partners, as currently practiced, corresponds to a cutpoint of 0, with sensitivity = 100% and specificity = 0%. Using a cutpoint of 1.50 for this model, DIS would identify 95.7% of undiagnosed HIV infection in partners while reducing the number of partners pursued by 15% (Table 6.3).

If false negatives are weighted 15 times worse than false positives, the ideal cutpoint in terms of minimizing total number of errors for the model with partnership data is a risk score of 2.00 (Table 6.3). Interviewing all partners at or above 2.00 has a sensitivity of 90.2% and reduces the number of partners DIS would need to locate and interview by 26%. Increasing the tradeoff weight to 30 decreases the ideal cutpoint to 1.50. The weight for universal partner referral is infinity.

Using bootstrap techniques, validation of the model demonstrated consistent performance over 1,000 replications.

The reference model to predict undiagnosed HIV infection in any partner of an index case included five predictors: shorter time between HIV diagnosis and interview, HIV diagnosis at a community health center or health department, no history of crack use, no history of anonymous sex, and two or more partners reported to DIS (data not shown). All variables remained in the final model. The area under the ROC curve for the final model was 0.649 (95% CI: 0.602, 0.696). Comparing this model to the one predicting partnerships involving an undiagnosed HIV infection, sensitivities were lower for this model at similar reductions in number of partners pursued.

DISCUSSION

Using demographic and behavioral data collected from DIS interviews of HIV index cases, we developed a risk score algorithm to predict undiagnosed HIV infection in named sexual partners. We identified five factors that predict a partnership with an undiagnosed partner—four weeks or fewer between HIV diagnosis and DIS interview, no history of crack use, no report of anonymous sex, fewer sexual partners reported to DIS, and sexual partnerships between an older index case and younger partner. The association between history of crack use and undiagnosed HIV infection in a sexual partner varied by age difference between the index case and partner. While overall performance of the model is low with poor specificity, it is possible to reduce the number of partners that need to be located and interviewed by up to 25% while maintaining sensitivity above 90%.

In deciding to use this algorithm to reduce DIS workloads, authorities would need to decide the relative value of a false negative compared to a false positive. Currently, in pursuing all partners, a false negative is considered infinitely worse than a false positive. In order to reduce the number of partners pursued, the tradeoff between false negatives and false positives must be quantified by weighing the potential public health and monetary costs of failing to diagnose an HIV infection with the monetary costs of hiring more DIS.

Alternatively, if DIS continue to pursue all partners, the model could be a helpful tool in prioritizing partners in which to invest more time for locating and testing. Currently DIS must complete an extensive checklist of locating tactics (e.g., searching the Department of Corrections database or checking for a social networking account) before declaring that a person is unable to be located. If the algorithm indicated that a partner should not be prioritized, the locating checklist could be modified so that not all tactics are attempted on this person, particularly those that are the most time consuming (e.g., driving to the person's listed address and asking neighbors for additional locating information).

A model using only index case characteristics to predict undiagnosed HIV infection in any named partners showed reduced predictive power. The sensitivities at the ideal cutpoints for minimizing errors were lower for this model, indicating that the model with partnerships as the unit of analysis is preferred. Therefore, prioritizing particular partnerships of an index case is better than prioritizing all partners of particular index cases.

We were unfortunately unable to collect data on partner notification costs in these two regions and are therefore unable to demonstrate the cost effectiveness of using a predictive model in this capacity. However, use of this model to prioritize partner interviews could ensure that the most undiagnosed HIV infections are identified in a timely manner given the available level of resources available for partner notification. Many health departments in the US currently provide inconsistent partner notification for HIV due to limited resources[118] and may benefit from prioritizing particular cases.

Our analysis uses data from two regions of NC and may not be generalizable to the other field service regions in the state or to other states due to varying prevalence of risk factors in different regions. However, the age and racial distributions of newly diagnosed persons in these two regions are similar to those for NC as a whole.[1] It may be worthwhile to test the sensitivity and specificity of this model on data from other NC regions or develop similar models in other regions if the model proves useful in the two regions of study. Areas outside of NC may also want to consider development of a model from routinely collected surveillance data.

Several factors may contribute to the relatively poor performance of the model and the limited reduction in number of partners interviewed. The strongest predictors for having an undiagnosed HIV-infected partner, such as type of sex, were undocumented. Although the risk of transmitting HIV via saliva is very low,[142] the odds of HIV transmission during receptive anal intercourse are much higher than the odds of transmission during insertive anal sex or vaginal sex.[143, 144] Therefore, the inclusion of type of sex would likely improve the predictive power of the model. Additionally, when DIS identify a newly diagnosed positive partner, the potential transmission dynamics are difficult to determine. The partner may have infected the index case, the index case may have infected the partner, or both may have been infected through other exposures. Because the timing and directionally of infection is unknown, the partnerships reflect a mixture of transmission events. Transmission events to the index case could have different predictors that are diluting the potential predictors of transmission events to the partner, reducing the predictive capacity with the available information.

While it may seem counterintuitive that several of our model predictors are considered lower risk behaviors for HIV transmission, this may be explained by the amount of locating information those persons with lower risk profiles were able to provide DIS. Index cases that reported anonymous sex or crack use and named more sex partners were more likely to report partners that could not be located or refused testing compared to those of a lower risk profile (data not shown). Although we do not have the data to show this, persons reporting only one partner to DIS may also have been in partnerships of longer duration that resulted in more unprotected sexual acts and therefore increased transmission probability compared to persons who reported multiple partners.

Our other model predictors are consistent with predictors of HIV infection identified in other studies. Persons reporting sex with an older partner were more likely to be HIV-infected compared to persons with partners their same age or younger in previous studies.[145-147] Our finding that partnerships with index cases interviewed four weeks or fewer after their HIV diagnosis predict undiagnosed HIV infections in partners is also consistent with previous data.[148, 149] Decreased time between diagnosis and patient interview increases the number of interviews yielding locatable contacts and therefore the number of partners notified and tested.. This increases the probability of identifying a partner with undiagnosed HIV infection.

Some innovations for improving partner notification have focused on where to interview or how to interview, such as targeting provider referral to areas of high endemicity or using enhanced interviewing techniques, but few have focused on who to interview based on case characteristics.[132, 150-152] Recently the San Francisco Department of Public Health evaluated a predictive model to prioritize partner notification interviews of syphilis index cases likely to result in treated partners.[153] While we are unaware of the implementation of such a predictive model to improve the efficiency of partner notification, predictive models have been shown to successfully increase efficiency and cost-effectiveness of STD case finding activities in the past. STD clinics have implemented selective screening criteria to test patients at highest risk for acquiring hepatitis B and hepatitis C infections.[134, 135] Predictive models have also been used to develop cost-effective screening programs for chlamydia among public clinic and emergency department patients and pregnant women.[136-141]

As resources available for partner notification decrease and HIV testing and case detection increase, public health departments are in need of novel strategies to maximize the efficiency of partner notification. Using data available from DIS interviews in two surveillance regions of NC, we demonstrate that it is possible to develop a model to predict undiagnosed HIV infection in partners, albeit with less accuracy than desired. Implementation of the model would allow DIS to prioritize partner interviews when all partners cannot be pursued and would allow DIS to reduce the number of partner interviews with high sensitivity for identifying undiagnosed HIV infection. Predictive models with additional partnership data including types and number of sex acts could potentially improve performance and should be explored as evidence-based approaches to improving partner notification.

TABLE 6.1. Index case-partner pair characteristics from two HIV surveillance regions in North Carolina, 2003-2007, by partner HIV status and associated odds ratios, restricted to complete cases included in model

Characteristics	Newly HIV-infected partner n (%)	Total n (%)	Unadjusted OR (95% CI)
Overall	164 (7.8)	2,100	
Demographics of index case			
Gender/Sexual orientation			
Female	52 (7.8)	663 (31.7)	1.0 (ref)
MSW	41 (9.2)	447 (21.4)	1.19 (0.77, 1.83)
MSM and MSM/W	71 (7.3)	979 (46.9)	0.92 (0.63, 1.35)
Race/Ethnicity			
White, non-Hispanic	34 (6.1)	561 (26.7)	1.0 (ref)
Other/Unknown	4 (10.0)	40 (1.9)	1.72 (0.55, 5.41)
White, Hispanic	15 (13.2)	114 (5.4)	2.35 (1.19, 4.64)
Black	11 (8.0)	1,385 (66.0)	1.35 (0.88, 2.08)
Stage of infection			
Chronic	127 (7.7)	1,645 (78.3)	1.0 (ref)
AIDS	33 (10.1)	327 (15.6)	1.34 (0.87, 2.06)
Acute	4 (3.1)	128 (6.1)	0.39 (0.14, 1.08)
Age			
14-19	11 (7.1)	155 (7.4)	1.0 (ref)
20-29	56 (7.7)	725 (34.5)	1.10 (0.54, 2.20)
30-39	44 (7.4)	597 (28.4)	1.04 (0.51, 2.13)
40-49	32 (7.5)	429 (20.4)	1.06 (0.51, 2.20)
50-59	17 (10.4)	164 (7.8)	1.51 (0.67, 3.41)
60+	4 (13.3)	30 (1.4)	2.01 (0.58, 6.96)
Time between HIV diagnosis and interview			
≤ 4 weeks	85 (9.2)	927 (44.1)	1.40 (1.00, 1.95)
> 4 weeks	79 (6.7)	1,173 (55.9)	1.0 (ref)
College student			
No	144 (7.9)	1,832 (87.2)	1.06 (0.65, 1.73)
Yes	20 (7.5)	268 (12.8)	1.0 (ref)
Immigrated to US			
No	145 (7.4)	1,963 (93.5)	1.0 (ref)
Yes	19 (13.9)	137 (6.5)	2.02 (1.22, 3.35)
Diagnosis Location			
Other	82 (7.4)	1,111 (56.7)	1.0 (ref)
CHC or Health Department	73 (8.6)	848 (43.3)	1.18 (0.84, 1.67)

Characteristics	Newly HIV-infected partner n (%)	Total n (%)	Unadjusted OR (95% CI)
Risk behaviors of index case			
History of incarceration			
No	133 (8.2)	1,620 (77.1)	1.30 (0.86, 1.95)
Yes	31 (6.5)	480 (22.9)	1.0 (ref)
Concurrent STD at HIV diagnosis			
No	145 (8.0)	1,824 (86.9)	1.17 (0.68, 2.00)
Yes	19 (6.9)	276 (13.1)	1.0 (ref)
History of crack use			
No	153 (8.4)	1,817 (86.5)	2.27 (1.22, 4.22)
Yes	11 (3.9)	283 (13.5)	1.0 (ref)
History of anonymous sex			
No	124 (9.5)	1,308 (62.3)	1.97 (1.32, 2.93)
Yes	40 (5.1)	792 (37.7)	1.0 (ref)
Exchanged sex for drugs/money			
No	153 (8.5)	1,811 (86.2)	2.33 (1.26, 4.30)
Yes	11 (3.8)	289 (13.8)	1.0 (ref)
Bisexual sex partner			
No	156 (8.0)	1,961 (93.4)	1.42 (0.69, 2.90)
Yes	8 (5.8)	139 (6.6)	1.0 (ref)
Number of sex partners reported to DIS			
1	75 (10.5)	715 (34.1)	2.68 (1.64, 4.38)
2-3	61 (8.5)	717 (34.1)	2.13 (1.29, 3.51)
≥ 4	28 (4.2)	668 (31.8)	1.0 (ref)
Characteristics of partnership			
Age difference between index and partner			
Partner is same age or older	87 (7.0)	1,239 (59.0)	1.0 (ref)
Partner is younger	77 (8.9)	861 (41.0)	1.30 (0.95, 1.79)
Same gender partnership			
No	96 (8.3)	1,152 (54.9)	1.17 (0.83, 1.65)
Yes	68 (7.2)	947 (45.1)	1.0 (ref)
Same race partnership			
No	26 (9.8)	265 (13.2)	1.37 (0.87, 2.16)
Yes	128 (7.3)	1,743 (86.8)	1.0 (ref)

Abbreviations: CI, confidence interval; DIS, disease intervention specialist; MSW, men who have sex with women; MSM, men who have sex with men; MSM/W, men who have sex with men and women; OR, odds ratio; STD, sexually transmitted disease

TABLE 6.2. Adjusted prevalence ORs and associated β-coefficient risk scores for variables included in the reference and final models to predict undiagnosed HIV infection in a sexual partner using data from two HIV surveillance regions of North Carolina, 2003-2007

Predictor	Reference model OR (95% CI), AUC=0.665	Final model OR (95% CI), AUC=0.662	β-coefficient risk scores
Time between HIV diagnosis and interview			
≤ 4 weeks	1.76 (1.20, 2.59)	1.74 (1.22, 2.47)	0.55
> 4 weeks	1.0 (ref)	1.0 (ref)	
Diagnosis location			
Other	1.0 (ref)	--	
CHC or Health Dept	1.12 (0.76, 1.64)		
History of crack use and age difference between index/partner			
Crack use, partner is same age or older	1.0 (ref)	1.0 (ref)	
No crack use, partner is same age or older	4.72 (1.48, 15.09)	3.92 (1.41, 10.86)	1.37
Crack use, partner is younger	4.33 (1.08, 17.39)	3.56 (1.03, 12.34)	1.27
No crack use, partner is younger	5.33 (1.66, 17.08)	4.45 (1.60, 12.39)	1.49
History of anonymous sex			
No	1.66 (1.10, 2.52)	1.75 (1.17, 2.62)	0.56
Yes	1.0 (ref)	1.0 (ref)	
Bisexual sex partner			
No	1.34 (0.63, 2.86)	--	
Yes	1.0 (ref)		
Number of sex partners reported to DIS			
1	2.40 (1.37, 4.23)	2.36 (1.43, 3.89)	0.86
2-3	1.66 (0.96, 2.85)	1.69 (1.03, 2.80)	0.53
≥ 4	1.0 (ref)	1.0 (ref)	
Same race partnership			
No	1.21 (0.74, 1.98)	--	
Yes	1.0 (ref)		

Abbreviations: AUC, area under the ROC curve; CHC, community health center; DIS, disease intervention specialists; OR, odds ratio; ref; referent

TABLE 6.3. Algorithm performance characteristics across selected risk scores, given the prevalence of undiagnosed HIV infection among partners in two HIV surveillance regions of North Carolina, 2003-2007

Risk scores	Sensitivity (95% CI)	Specificity (95% CI)	Percent pursued (95% CI)	Number of FNs/FPs[a]	Total errors[b]	Total errors, weight=15	Total errors, weight=30
RS ≥ 0	100	0	100	0/992	922	922	922
RS ≥ 0.50	99.4 (98.0, 100.0)	0.9 (0.5, 1.0)	99.1 (98.7, 99.5)	1/914	915	929	944
RS ≥ 1.00	98.8 (96.9, 100.0)	4.4 (3.4, 5.3)	95.9 (95.0, 96.8)	1/882	883	897	912
RS ≥ 1.50	95.7 (92.8, 98.7)	16.9 (15.3, 18.6)	84.1 (82.5, 85.6)	4/767	771	827	887
RS ≥ 2.00	90.2 (85.5, 94.6)	28.1 (26.1, 30.2)	73.4 (71.5, 75.2)	8/663	671	783	903
RS ≥ 2.50	66.5 (59.3, 74.3)	53.3 (50.8, 55.5)	48.3 (46.3, 50.7)	27/431	458	836	1,241
RS ≥ 3.00	33.5 (26.7, 40.7)	79.6 (77.7, 81.3)	21.4 (19.8, 23.3)	52/188	240	968	1,488

[a]FN = (1-Sensitivity)*Prevalence*1000, FP=(1-Specificity)*(1-Prevalence)*1000
[b]FNs and FPs are equally weighted
Abbreviations: CI, confidence interval; FN, false negative; FP, false positive

FIGURE 6.1. Flow chart of study selection criteria using data from two HIV surveillance regions in North Carolina, 2003-2007

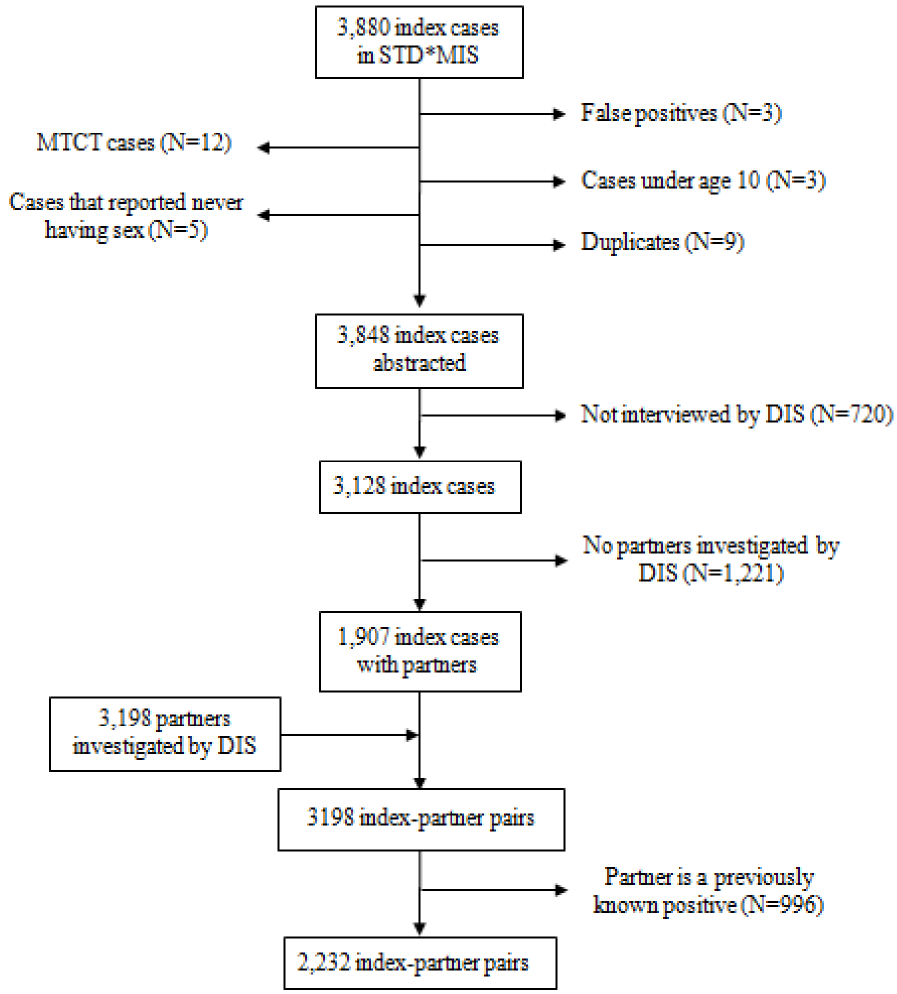
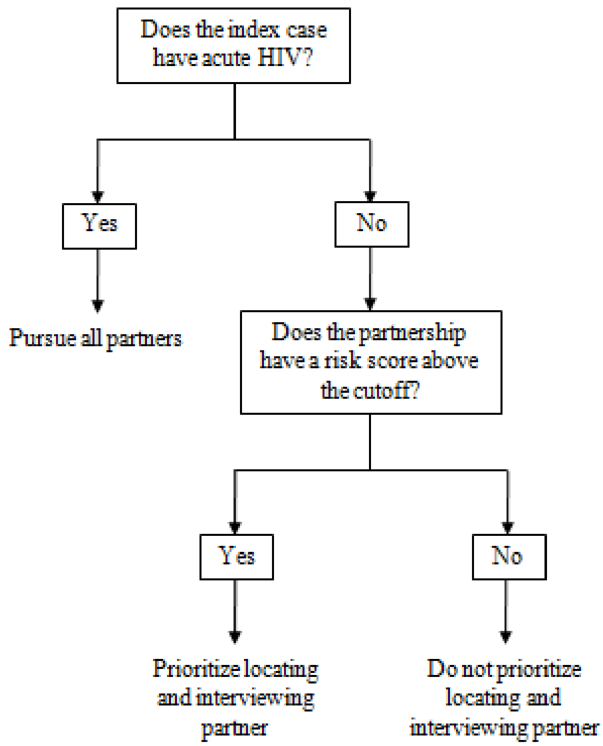


FIGURE 6.2. Algorithm for prioritizing partner interviews using data from two HIV surveillance regions of North Carolina, 2003-2007



CHAPTER SEVEN: A Predictive Model to Prioritize Prevention Interventions for People Living with HIV/AIDS in North Carolina

ABSTRACT

Background: North Carolina (NC) has control measures for people living with HIV/AIDS (PLWHA) to reduce unprotected sex and acquisition of sexually transmitted infections that could lead to further HIV transmission. Identifying persons likely to violate control measures and linking them with case management services soon after diagnosis is a potentially efficient and cost-effective prevention strategy. We developed and evaluated risk scores to predict future control measure violation in order to prioritize persons for case management intervention. Methods: We abstracted demographic, behavioral, and partnership data from disease intervention specialists' records of HIV-positive persons reported in NC between January 1, 2003 and December 31, 2007. Risk scores were developed using a Cox proportional hazards model. The sensitivities and specificities of the risk scores at different cutoffs were used to examine algorithm performance. Results: We identified five factors that predict violation of NC control measures—identifying as a man who has sex with men, younger age, syphilis co-infection at the time of HIV diagnosis, marijuana use in the past year, history of anonymous sex, and reporting two or more sex partners to DIS during partner notification. Using this algorithm, referring 23% of the population to case management intervention would capture over half of control measure violators. Conclusions: While the overall predictive power of the model is moderate, it is possible to prioritize case management intervention for those engaging in risky behaviors that perpetuate HIV

transmission. Predictive models should be explored as evidence-based approaches to implementing limited interventions for PLWHA.

INTRODUCTION

People living with HIV and AIDS (PLWHA) are living longer and more sexually active lives due to advances in antiretroviral therapy.[154, 155] Because PLWHA are living longer, reducing the risk of transmitting HIV to others is an important aspect of medical care for HIV-infected persons. Most PLWHA respond to their HIV diagnosis by adopting lower risk sexual behaviors,[156, 157] but approximately 33% of PLWHA continue to have unprotected sexual intercourse that may put others at risk for HIV infection.[87-89]

North Carolina (NC) has control measures in place for PLWHA to minimize the spread of HIV to others.[93] The control measures include refraining from sexual intercourse unless condoms are used and notifying future sexual partners of HIV infection. If a previously known HIV-infected person is named as a sexual partner of a newly diagnosed HIV index case or is reported to the state or local health department with a new sexually transmitted infection (STI) diagnosis, he or she is considered to have violated NC control measures. Criminal prosecution of these persons is rare, but has occurred in NC.

The Centers for Disease Control and Prevention (CDC) currently recommends prevention case management or comprehensive risk counseling and services for PLWHA.[91, 103, 104] These interventions are intensive, but are effective in reducing unprotected sex and sexually transmitted infection (STI) acquisition. Case management is recommended specifically for complex cases in which provider-based or group interventions are unlikely to reduce transmission risk.[105] Case management may be particularly useful for persons who violate control measures.

In the current economic environment, public health departments' budgets are being cut, necessitating judicious use of case management. Identifying persons likely to violate control measures and linking them with case management services soon after diagnosis is a potentially efficient and cost-effective prevention strategy. Using demographic and behavioral characteristics of persons newly diagnosed with HIV in NC, we developed and evaluated risk scores to predict future control measure violation in order to prioritize persons for case management intervention.

METHODS

Data Collection

North Carolina is divided into seven HIV/STD surveillance regions that have their own disease intervention specialists (DIS) that provide diagnosis and partner notification. We reviewed the Sexually Transmitted Disease Management Information System (STD*MIS) database from two of these regions (Winston-Salem and Raleigh regional offices) to identify persons aged ≥ 10 years in whom HIV was diagnosed between January 1, 2003 and December 31, 2007. These two regions include 27 of North Carolina's 100 counties and encompass approximately 40% of the state's incident HIV cases. Cases were excluded if they were unable to be located, refused DIS interview, were attributable to mother-to-child transmission, or reported no sexual history. The University of North Carolina Institutional Review Board approved all study procedures.

DIS maintain a chart for each index case at the regional office that contains the STD*MIS entry, interview notes with the index case and his/her partners, and information from the client's providers. Data on demographics, HIV risk factors, and sexual behaviors were abstracted from these charts onto a standard form and entered into an Access database.

When a case is identified as a control measure violator (CMV), DIS create a separate chart with the case's original STD*MIS entry and a document detailing the violation and follow-up. All CMV files from January 1, 2003 through June 30, 2010 were reviewed at both regional offices. Data on date of violation, type of violation, and actions taken were abstracted and entered into the database.

Data Analysis

The outcome of our predictive model was future violation of HIV control measures. Violation of control measures was defined by a DIS investigation into a person's sexual behaviors following an initial DIS interview after diagnosis with HIV infection. The set of possible predictor variables included demographic characteristics and HIV risk behaviors documented in the original DIS chart following HIV diagnosis in NC.

We examined the relationship between each predictor variable and the outcome using Cox proportional hazards models to estimate unadjusted hazard ratios and their associated 95% confidence intervals. We also assessed the association between each pair of candidate predictor variables to avoid collinearity. For dichotomous variables, we used an odds ratio to assess collinearity and determined that the two variables were collinear if the odds ratio was 3 or greater. If one variable was continuous and the other was categorical, we examined the magnitude of the difference in means in standardized units. A difference of more than 1.5

standard deviations was considered a strong association.[113] Collinear variables were recoded or one of the variables was selected based on the substantive meaning and relationship to other variables.

To describe gender and male sexual risk groups, we created a composite variable with three categories: men who reported having sex with men (MSM) and men who reported having sex with men and women (MSM/W), men who did not report MSM (i.e. heterosexual men), and women. The referent category was set to "women" because this group had the lowest risk for control measure violation.

Variables for which $P < .25$ in the bivariable analyses were selected for inclusion in the Cox proportional hazards model.[115] A time-to-event analysis was used to account for differences in follow-up time for CMV violation after the initial DIS interview. Interaction terms between all candidate predictors included in the model were examined and retained in the model if their $P < .25$. This model was considered the full, or "reference," model.

We examined reduced models to see if they maintained model fit without loss of predictive power. Modeling proceeded in a backward elimination process using a lower alpha ($P < .10$) to eliminate predictors with weak predictive power, starting with interaction terms and then proceeding with the variable with the highest *P* value. Change in the C-statistic was used to assess variations in model performance due to collapsing across categories or removing variables. For binary outcomes, the C-statistic in time-to-event analysis is identical to the area under the receiver operating characteristic curve for logistic regression.[121] The modeling procedures were limited to those persons with complete data for all variables in the reference model. Model fit was evaluated using the Hosmer-Lemeshow test. The proportional hazards assumption was assessed graphically by plotting log-log survival plots.

A change in C-statistic less than 0.01 was acceptable between models. The final model had the fewest covariates with minimal reduction in C-statistic and the best model fit. The modeling procedures were limited to those persons with complete data for all variables in the model.

We created clinical risk scores using the β-coefficients corresponding to each predictor in the final model. The β-coefficients were summed to create an overall clinical risk score for each patient. Internal validity of the resulting model and risk score sensitivity and specificity were examined using bootstrap analysis in which the partner population was resampled 1,000 times with replacement.

To develop an optimal strategy for predicting future violation of control measures and intervening with these individuals, we examined the trade-off between the number of misclassification errors that would be made depending on the cutpoint used for additive risk score totals (i.e., over a certain cutpoint, an index case would be considered a potential future CMV that could benefit from additional intervention). A false positive (FP) was defined as choosing to provide additional intervention to an index case who was not going to violate control measures in the future , whereas a false negative (FN) was defined as failing to choose an index case who violates control measures in the future for additional intervention.

A FN was weighted more than a FP since it would be worse to miss a future CMV than to invest prevention resources into a person that did not go on to violate control measures. The following calculations were made to determine the number of errors associated with the sensitivity and specificity of the model at different risk score cutpoints:

Number of FN = (1 – sensitivity) * CMV prevalence * N

Number of FP = (1 – specificity) * (1 – CMV prevalence) * N

Number of errors = (weight * FN) + FP,

where weight reflects the relative value of a FN compared to a FP. All analyses were conducted using SAS software, Version 9.2 of the SAS System for Windows (Cary, NC).

Sensitivity Analysis

To examine the effect of possible misclassification of the outcome, persons who were investigated as CMVs but who were not reported to their local health department as CMVs or whose investigation outcome was unknown were recoded as not violating control measures. The final predictive model was re-run with this modified outcome and the C-statistic was examined to assess change in predictive power.

RESULTS

Among 3,880 index cases from January 1, 2003 to December 31, 2007, 3,848 cases were eligible and DIS interviewed 3,128 (81.3%). Of these, 169 (5.4%) persons were interviewed for control measure violations.

Most of the index cases were male (70.1%) and black (65.8%) (Table 7.1). Most were also in the chronic stage of HIV infection, with only 2.8% of cases diagnosed with acute HIV (HIV antibody-negative, RNA-positive cases)and 24.5% identified as AIDS cases (CD4 count or percent < 200 cells/µL or 14%, respectively, or diagnosis with an AIDS-defining illness). The median age of the index cases was 38 years (range: 14-83 years). Almost one-third (30.6%) had engaged in sex with an anonymous partner and 22.8% reported two or more sex partners to DIS.

Of the 169 persons interviewed for violation of control measures (and therefore considered CMVs in this analysis), most (71.6%) were reported to their local health department as CMVs (Table 7.2). Forty-eight of those investigated may not have been considered true CMVs following DIS investigation, including 22 people not reported to the health department and 26 persons with an unknown outcome following DIS investigation.

DIS conducted 194 investigations of 169 persons; 148 CMVs had one filed investigation, while 21 had two or more (Table 7.1). Over half of the 194 investigations (54.1%) resulted from an index case being reported to the health department with another STI (failure to use condoms). Fifty-four cases (27.9%) were partners of newly infected HIV cases that reported being unaware of the CMV's HIV infection (failure to disclose status). Other reasons for initiating investigations included report by a partner of unprotected sex (failure to disclose status and failure to use condoms) and pregnancy or a pregnant partner (failure to use condoms). Five violations involved multiple reasons for investigation. The median time between HIV diagnosis and first violation was 2.8 years (range: 0.05-20.7 years).

Internet use to meet sexual partners at the time of HIV diagnosis was strongly associated with violating control measures (HR 3.0, 95% confidence interval (CI): 2.2, 4.2) (Table 7.1). Other potentially important predictors of violating control measures recorded at the initial DIS interview were identifying as an MSM or an MSM/W, younger age, being a college student, marijuana use in the past year, meeting sexual partners in bars or clubs, history of anonymous sex, and reporting two or more sexual partners to DIS at the initial interview. Acute HIV infection at diagnosis, black race, non-Hispanic ethnicity, US birth, syphilis co-infection at time of HIV diagnosis, no history of alcohol abuse, and history of sexual assault were also candidate predictors for the reference model.

The reference model included stage of HIV infection at diagnosis, gender/sexual orientation, age, syphilis co-infection, marijuana use in the past year, history of anonymous sex, and number of sex partners reported to DIS (Table 7.3). Age was included as a continuous variable. The relationship between gender/sexual orientation and violation of control measures differed by history of anonymous sex, so an interaction term between these variables was included. The C-statistic was 0.736 (95% CI: 0.698, 0.774) for this model.

Only stage of HIV infection at diagnosis was removed during model simplification. The C-statistic for the final model was 0.737 (95% CI: 0.698, 0.774).

Currently no HIV-infected persons are linked to case management intervention in North Carolina. We are interested in linking a small proportion of the HIV-infected population of these two regions to intervention with a high sensitivity for identifying control measure violators. Linking no one to intervention corresponds to a cutpoint of 1.50, with sensitivity = 0% and specificity = 100% (Table 7.4).

By not intervening with anyone, false negatives currently have a weight of zero. If false negatives are instead weighted ten times worse than false positives, the ideal cutpoint in terms of minimizing errors is -0.25, which corresponds to a sensitivity of 53% and specificity of 79%. Intervening with all persons at or above 0.70 means that 23% of the population would be linked to intervention and 53% of future CMVs would receive intervention. As the tradeoff weight increases, the cutpoint decreases (Figure 7.1). This results in intervening with a larger proportion of the population and providing intervention for more CMVs.

Using bootstrap techniques, validation of the model demonstrated consistent performance over 1,000 replications.

Sensitivity Analysis

When persons who were not reported to their local health department were recoded as non-CMVs, the predictive power of the final model increased (C-statistic=0.755). Similarly, when those who were not reported to their local health department and those persons with an unknown outcome were recoded as non-CMVs, the predictive power of the model was higher, although slightly lower than the model where only those not reported were recoded as not having the outcome (C-statistic=0.752).

DISCUSSION

Using demographic and behavioral data collected from initial DIS interviews of HIV-infected persons following diagnosis in NC, we developed a risk score algorithm to predict future violation of NC control measures. We identified six factors that predict future control measure violation—identifying as a MSM or MSM/W, younger age, syphilis co-infection at time of HIV diagnosis, marijuana use in the past year, history of anonymous sex, and reporting two or more sex partners to DIS during partner notification. The association between gender/sexual orientation and control measure violation varied by history of anonymous sex status.

The final model had moderate predictive performance. While sensitivity is low when intervening with a small proportion of the HIV-infected population, using the risk score algorithm to recommend intervention for 23% of the population (risk score $\geq$ -0.25) would capture over half of control measure violators. Specificity is high at this cutpoint, indicating that few false positives would occur and intervention resources would be spent primarily on those at highest risk for violation.

In considering use of this algorithm, public health authorities would need to decide the relative value of a false negative compared to a false positive. Currently, by intervening with no one beyond providing linkage to care recommendations and some case management to those acutely diagnosed with HIV at diagnosis, a false negative has a weight of zero (i.e., a false positive is considered infinitely worse than a false negative). The tradeoff must be quantified by weighing the potential public health cost of potential continued HIV transmission by HIV-infected persons aware of their status against the cost of providing limited case management. Determination of the tradeoff could be done formally with a cost-utility analysis, or health departments could take a more intuitive approach and consider current cost constraints and resource limitations.

Because we do not have data on linkage to care in this population, we are unable to discuss the level or intensity of intervention needed to reduce behaviors that lead to control measure violation. It is unknown if the CMVs were successfully linked to care following their initial DIS interview or if they were in care at the time of the control measure violation. If non-CMVs were more likely to be in care compared to CMVs, linkage to care and maintenance in care for those indicated by the model may be enough to reduce the incidence of control measure violation. Counseling by primary health care providers can help PLWHA change risky health behaviors.[106] Alternatively, if linkage to care is not associated with CMV status, more intensive interventions, such as psychosocial support or group-level education and support classes, may be required for a reduction in the number of CMVs.

Our analysis uses data from two regions of NC and may not be generalizable to the other field service regions in the state due to varying prevalence of risk factors in different regions. However, the age and racial distributions of newly diagnosed persons in these two regions are similar to those for NC as a whole.[1] It would be worthwhile to validate this model in other regions. Areas outside of NC with available resources for bridging case management may also want to consider assessing this model performance or developing a comparable local model to prioritize intervening with HIV-infected individuals at highest risk for non-disclosure or continued unprotected sex.

Our outcome is dependent on report of a violation to the regional office and follow-up based on DIS discretion and therefore may not represent the actual incidence of control measure violation. If a person does not contract an additional STI or is not reported by a sexual partner as failing to disclose their status or use a condom, they would not be considered a CMV despite engaging in risky sexual behaviors. Between 21 and 50% of HIV-positive MSM reported unprotected anal intercourse with a serodiscordant or serostatus unknown partner in US studies of MSM.[158-161] It is therefore likely that the number of actual CMVs is significantly higher. In addition, some of the cases classified as CMVs may not actually have violated control measures, particularly those where there was not enough evidence or the DIS decided it was unnecessary to report the HIV-infected person to the health department. However, the sensitivity analysis showed that reclassifying the outcome by considering those not reported to the health department and those with unknown investigation outcome non-CMVs, the predictive power of the model increased.

Data on the HIV status of the partners involved in CMV investigations were also unavailable for this analysis. Therefore, we are unable to discuss the prevalence of serosorting among the CMVs who were investigated due to STI acquisition. While unprotected sex among seroconcordant HIV-infected persons does not lead to further HIV transmission, NC control measures require condom usage regardless of the partners' infection statuses. Also, whether serosorting or not, these individuals acquired STIs that may further negatively impact their health and increase the likelihood of HIV transmission to HIV-uninfected partners.

Control measure violators in NC were more likely to be black and identify as MSM or MSM/W—populations disproportionately affected by HIV-related stigma. This stigma continues to inform perceptions and shape the behavior of PLWHA, making disclosure of HIV status difficult. Stigma thus complicates efforts by HIV-infected individuals to engage in safe sexual relationships.[95] There is concern that criminalization of HIV may serve as a barrier to HIV prevention if it increases stigma associated with HIV infection rather than deterring behaviors that transmit HIV (Shriver 2000, Burris 2008).[95, 96] Providing limited case management to reduce unsafe sexual behaviors may be more effective in reducing control measure violation compared to the current system of investigation and potential prosecution.

Traditional prevention case management involves a multi-faceted approach of managing medical, mental health, and substance abuse care as well as social services on an individual level, and is effective in promoting behavior change.[95] While it can be expensive, a meta-analysis of clinical trials showed that case management for PLWHA was successful in reducing unprotected sex by 43% and decreasing acquisition of sexually transmitted infections by 80%.[89] Successful interventions are based on behavioral theory and focus on the challenges of living with HIV, particularly on transmitting the virus to partners and managing stress related to HIV disclosure. An important component is to help PLWHA protect their partners and *themselves* by stressing the importance of decreasing risks to their own health (e.g., contracting other STIs or other strains of HIV that could confer drug resistance).[100]

While service providers in various disciplines generally agree on the basic functions of case management, there is not a consensus on the scope of services offered. Case management can take on a broad array of service definitions, from a few phone calls to an HIV-infected person following diagnosis to encourage linkage to care to home visits that occur over an extended period of time. NC DHHS may therefore consider monitoring CMV incidence while providing less-intensive case management services initially to limit costs before implementing more rigorous interventions.

As resources for HIV prevention decrease, we are in need of novel strategies to maximize the efficiency of targeted prevention interventions. Targeting HIV-infected persons likely to violate control measures in the future for case management intervention could be useful in reducing risky behaviors that lead to further HIV transmission. Using data available from DIS interviews in two surveillance regions of North Carolina, we developed a model to predict control measure violation among HIV-infected cases reported to the state health department. Implementation of the model would allow authorities to implement case management intervention for a small proportion of the HIV-infected population that are known to be engaging in activities that perpetuate HIV transmission. Predictive models with additional data on linkage to care to potentially improve performance should be explored as evidence-based approaches to implementing limited interventions for PLWHA.

TABLE 7.1. Population characteristics of index cases reported in two HIV surveillance regions in North Carolina between 2003 and 2007 and unadjusted hazard ratios by CMV status, restricted to complete cases

Characteristic	Total n (%)	Control Measure Violators n (%)	Unadjusted HR (95% CI)
Overall	3,081	167 (5.4)	
Stage of infection			
Chronic	2,238 (72.6)	133 (5.9)	1.0 (ref)
AIDS[a]	756 (24.5)	27 (3.6)	0.5 (0.3, 0.8)
Acute[b]	87 (2.8)	7 (8.1)	1.6 (0.7, 3.3)
Black race			
No	1,054 (34.2)	46 (4.4)	1.0 (ref)
Yes	2,027 (65.8)	121 (6.0)	1.5 (1.1, 2.1)
Hispanic ethnicity			
No	2,831 (91.9)	159 (5.6)	1.5 (0.7, 3.1)
Yes	250 (8.1)	8 (3.2)	1.0 (ref)
Gender/sexual orientation			
Female	921 (29.9)	27 (2.9)	1.0 (ref)
MSW	882 (28.6)	32 (3.6)	1.1 (0.7, 1.9)
MSM and MSM/W	1,278 (41.5)	108 (8.5)	2.8 (1.8, 4.2)
Age (years)			
14-19	105 (3.4)	16 (9.6)	13.3 (5.5, 32.5)
20-29	713 (23.1)	60 (35.9)	6.7 (3.1, 14.7)
30-39	916 (29.7)	50 (29.9)	3.3 (1.5, 7.3)
40-49	858 (27.8)	34 (20.4)	2.1 (0.9, 4.7)
50+	489 (15.9)	7 (4.2)	1.0 (ref)
Immigrated to US			
No	2,750 (89.3)	160 (5.8)	2.4 (1.1, 5.1)
Yes	331 (10.7)	7 (2.1)	1.0 (ref)
College student			
No	2,868 (93.1)	145 (5.1)	1.0 (ref)
Yes	213 (6.9)	22 (10.3)	2.5 (1.6, 3.9)
Syphilis co-infection			
No	2,849 (92.5)	140 (4.9)	1.0 (ref)
Yes	232 (7.5)	27 (11.6)	2.2 (1.4, 3.3)
History of alcohol abuse			
No	2,645 (85.9)	151 (5.7)	1.4 (0.8, 2.3)
Yes	436 (14.2)	16 (3.7)	1.0 (ref)
Marijuana use in past year			
No	2,598 (84.3)	121 (4.7)	1.0 (ref)
Yes	483 (15.7)	46 (9.5)	1.9 (1.3, 2.6)
Internet use to meet sexual partners			
No	2,642 (85.8)	115 (4.4)	1.0 (ref)
Yes	439 (14.3)	52 (11.9)	3.1 (2.2, 4.2)

Characteristic	Total n (%)		Control Measure Violators n (%)		Unadjusted HR (95% CI)	
Bar use to meet sexual partners						
No	2,388	(77.4)	101	(4.2)	1.0	(ref)
Yes	698	(22.6)	66	(9.5)	2.4	(1.7, 3.2)
History of sexual assault						
No	2,981	(96.8)	158	(5.3)	1.0	(ref)
Yes	100	(3.3)	9	(9.0)	2.1	(1.0, 4.0)
History of anonymous sex						
No	2,138	(69.4)	96	(4.5)	1.0	(ref)
Yes	943	(30.6)	71	(7.5)	1.9	(1.4, 2.6)
Number of sexual partners reported to DIS						
0 or 1	2,380	(77.3)	98	(4.1)	1.0	(ref)
≥ 2	701	(22.8)	69	(9.8)	2.2	(1.6, 3.0)

[a]CD4 count or percent less than 200 or 14%, respectively, or diagnosis with an AIDS-defining illness
[b]HIV antibody-negative, RNA-positive cases
Abbreviations: CI, confidence interval; CMV, control measure violator; DIS, disease intervention specialist; HR, hazard ratio; MSM, men who have sex with men; MSM/W, men who have sex with men and women; MSW, men who have sex with women

TABLE 7.2. Characteristics of 169 index cases reported between 2003 and 2007 in two HIV surveillance regions of North Carolina who violated HIV control measures

Characteristic	n (%)
Individual reported to Health Department	
No	22 (13.0)
Yes	121 (71.6)
Unknown	26 (15.4)
Number of filed violations	
1	148 (87.6)
2	18 (10.7)
3	2 (1.2)
4	1 (0.6)
Violation type (all violations)	
STI diagnosis[a]	105 (54.1)
Partner to HIV case	54 (27.9)
Partner reported unprotected sex	26 (13.4)
Pregnant or pregnant partner	4 (2.0)
Multiple reasons	5 (2.6)

[a]chlamydia (n=2) , gonorrhea (n=13), syphilis (n=90), trichomoniasis (n=6)
Abbreviations: STI, sexually transmitted infection

TABLE 7.3. Adjusted prevalence ORs and associated 95% confidence intervals for variables included in the final model using data from two HIV surveillance regions of North Carolina, 2003-2007

Predictor	Reference model HR (95% CI)[a]	Final model HR (95% CI)[b]	β-coefficient risk score
Stage of infection			
Chronic	1.0 (ref)	NIM	
AIDS	0.7 (0.5, 1.2)		
Acute	1.1 (0.5, 2.4)		
Gender/sexual orientation and history of anonymous sex			
Female, no anonymous sex	1.0 (ref)	1.0 (ref)	
Female, history of anonymous sex	3.3 (1.5, 7.1)	3.2 (1.5, 7.0)	1.17
MSW, no anonymous sex	1.3 (0.6, 2.7)	1.3 (0.6, 2.6)	0.23
MSW, history of anonymous sex	3.7 (1.8, 7.5)	3.6 (1.8, 7.2)	1.28
MSM or MSM/W, no anonymous sex	3.2 (1.8, 5.6)	3.2 (1.8, 5.5)	1.15
MSM or MSM/W, history of anonymous sex	3.1 (1.7, 5.5)	3.0 (1.7, 5.5)	1.11
Age in years at DIS interview (continuous)			-0.05*Age
Age 30 v. age 40	1.6 (1.4, 2.0)	1.7 (1.4, 2.0)	
Syphilis co-infection			
No	1.0 (ref)	1.0 (ref)	
Yes	1.8 (1.2, 2.8)	1.8 (1.2, 2.8)	0.59
Marijuana use in past year			
No	1.0 (ref)	1.0 (ref)	
Yes	1.4 (0.9, 1.9)	1.4 (1.0, 2.0)	0.33
Number sex partners pursued by DIS			
0 or 1	1.0 (ref)	1.0 (ref)	
≥ 2	1.7 (1.2, 2.3)	1.7 (1.3, 2.4)	0.55

[a]C-statistic= 0.736

[b]C-statistic = 0.737

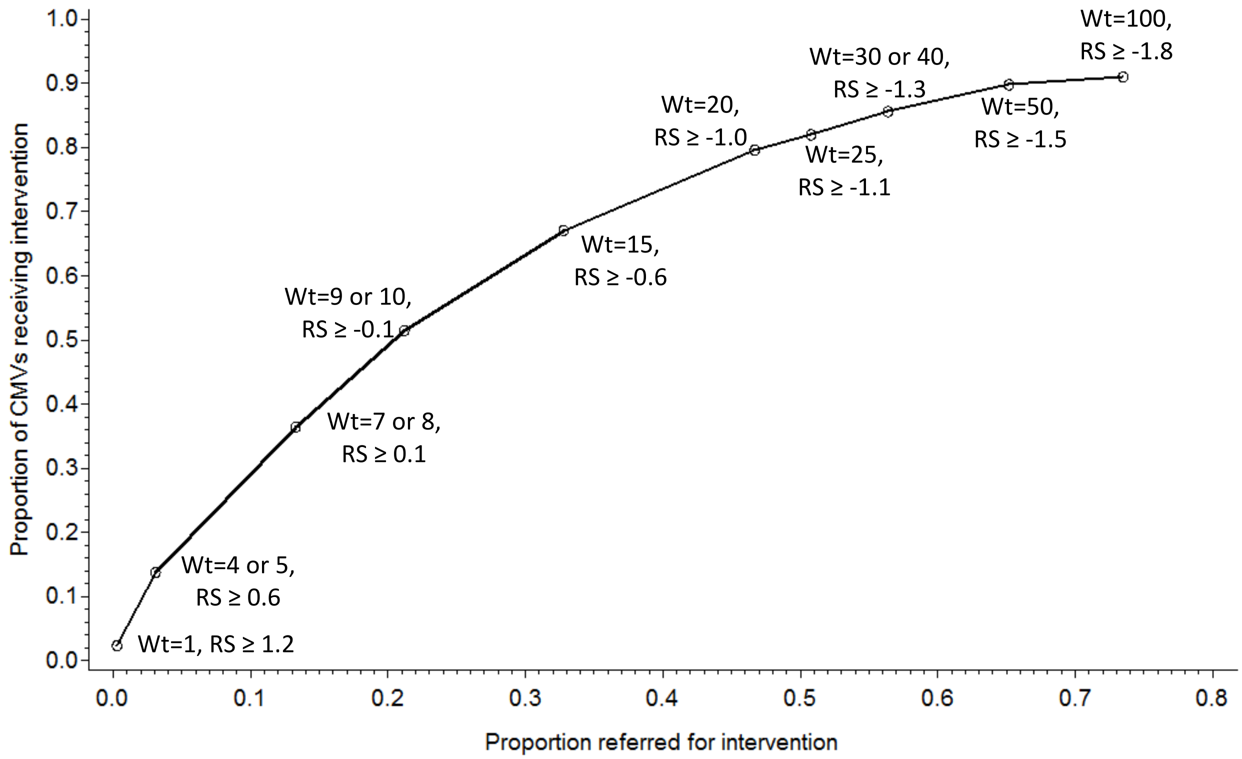
Abbreviations: CI, confidence interval; DIS, disease intervention specialist; HR, hazard ratio; MSM, men who have sex with men; MSM/W, men who have sex with men and women; MSW, men who have sex with women; NIM, not in model

TABLE 7.4. Performance characteristics of the algorithm across selected risk scores, given the prevalence of CMVs in the current study population of index cases from two HIV surveillance regions in North Carolina, 2003-2007

Risk Score Cutoff	Sensitivity	Specificity	Percent referred for case management	Number of FNs/FPs[a]	Total errors, weight=1	Total errors, weight=10	Total errors, weight=20
RS ≥ -1.75	91.0 (86.4, 95.0)	27.5 (25.9, 29.0)	73.5 (72.0, 74.9)	5/686	691	736	786
RS ≥ -1.50	89.8 (85.1, 94.0)	36.2 (34.5, 37.9)	65.2 (63.5, 66.8)	6/604	610	664	724
RS ≥ -1.25	85.6 (80.4, 90.8)	45.3 (43.5, 47.1)	56.4 (54.5, 58.0)	8/518	526	598	678
RS ≥ -1.00	79.6 (73.7, 85.8)	55.2 (53.3, 56.9)	46.7 (44.9, 48.4)	12/424	436	544	664
RS ≥ -0.75	70.7 (63.8, 78.0)	63.9 (62.1, 65.6)	38.0 (36.2, 39.6)	16/342	358	502	662
RS ≥ -0.50	61.1 (53.6, 68.8)	71.6 (70.1, 73.2)	30.2 (28.5, 31.7)	22/270	292	490	709
RS ≥ -0.25	53.3 (45.8, 60.7)	79.2 (77.9, 80.7)	22.5 (21.0, 23.9)	26/197	223	457	717
RS ≥ 0	40.1 (32.7, 47.5)	85.9 (84.7, 87.2)	15.5 (14.2, 16.7)	33/134	167	464	794
RS ≥ 0.25	29.3 (22.9, 36.7)	90.9 (89.9, 92.0)	10.2 (9.1, 11.2)	39/86	125	476	866
RS ≥ 0.50	21.6 (15.8, 27.9)	94.5 (93.7, 95.3)	6.3 (5.5, 7.2)	43/52	95	482	912
RS ≥ 0.75	12.0 (7.1, 17.4)	97.8 (97.3, 98.4)	2.7 (2.1, 3.3)	48/21	69	501	981
RS ≥ 1.00	5.4 (2.3, 9.1)	99.3 (98.9, 99.6)	1.0 (0.7, 1.4)	52/8	60	528	1,048
RS ≥ 1.50[b]	0	100	0	55/0	55	550	1,100

[a]FN = (1-Sensitivity)*Prevalence*1000, FP=(1-Specificity)*(1-Prevalence)*1000
[b]Equivalent to no case management
Abbreviations: CMV, control measure violation; FN, false negative; FP, false positive; RS, risk score

FIGURE 7.1. Proportion of population referred for intervention and proportion of CMVs receiving intervention by different tradeoff weights using data from two HIV surveillance regions of North Carolina, 2003-2007



CHAPTER EIGHT: DISCUSSION

Thirty years into the epidemic, there are approximately 35,000 people living with HIV/AIDS in NC and 1,800 new infections reported annually.[1] The number of newly reported HIV infections has remained stable over the last few years at a level that is 40% higher than the national level.[17] This divergence between the incidence in NC and the national incidence is characteristic of the Southern US, where the HIV epidemic is multi-faceted. Like the rest of the South, the NC epidemic spans urban and rural areas, and heterosexuals and MSM.

NC faces this complex epidemic with limited resources. In the current economic environment, public health departments are facing funding cuts and hiring freezes. DIS, whose primary responsibility is partner elicitation and notification for HIV and syphilis in NC, are filling critical gaps in staffing. In addition, HIV testing efforts in the state have increased in an attempt to identify the estimated 35% of persons with undiagnosed infection. With increased testing, partner notification demand has also increased. It is therefore important to develop practical tools to help DIS perform their primary tasks more efficiently.

Public health departments in NC are also in need of strategies that bring the biggest "bang for the buck"--in this case, interventions that produce the greatest reduction in HIV incidence with the least expenditure of limited resources. Decisions regarding where to allocate resources require knowledge of the spatial distribution of HIV in the community and which groups would benefit the most from targeted prevention and treatment.[19] This dissertation aimed to characterize the geographic distribution of HIV in two regions of NC in order to better inform future allocation of HIV resources and to provide practical tools to aid DIS in prioritizing their HIV caseloads.

Summary of Findings

In this dissertation, we described several findings that may inform the utility of future interventions as well as two risk score algorithms to prioritize partner notification and prevention case management for HIV-infected persons. In our first specific aim, we identified highly localized geographic clusters of reported HIV cases in urban areas, supporting the existence of core areas of HIV transmission in NC. However, these clusters were temporal in addition to spatial in nature and did not persist in the last two years of our study period. The disappearance of these clusters was coincident with a dramatic increase in Internet use to meet sexual partners among MSM and MSM/W. Internet use was associated with a greater mean distance to sexual partners, suggesting that online sex-seeking may be changing the phenomenon of local partner selection by linking sexual networks that otherwise may not have come into contact with each other.

In our second specific aim, we developed a risk score algorithm to predict undiagnosed HIV infection in sexual partners to aid DIS in prioritizing interviews. We identified five factors that predict a sexual partnership between a person with newly diagnosed HIV infection and an undiagnosed partner—four weeks or fewer between HIV diagnosis and DIS interview, no history of crack use, no report of anonymous sex, fewer sexual partners reported to DIS, and sexual partnerships between an older index case and younger partner. While the overall predictive power of the model was low, it is possible to reduce the number of partners that need to be located and interviewed while maintaining high sensitivity. Using this model, DIS could identify 90.2% of undiagnosed HIV infection in partners while reducing the number of partners pursued by 25%.

We developed and evaluated risk scores to predict future control measure violation in order to prioritize persons for case management intervention in our third specific aim. We identified five factors that predicted violation of NC control measures—identifying as a man who has sex with men, younger age, syphilis co-infection at the time of HIV diagnosis, marijuana use in the past year, history of anonymous sex, and reporting two or more sex partners to DIS during partner notification. As in the second aim, the overall predictive power of the model was moderate. However, use of this algorithm would facilitate prioritizing case management intervention for those engaging in risky behaviors that perpetuate HIV transmission. Referring 23% of the population to case management intervention using this algorithm would capture over half of control measure violators.

Public Health Significance and Future Directions

Aim 1: Our findings in Aim 1 suggest that HIV testing events targeting specific locations may not be effective in identifying high HIV transmission networks since core areas did not persist after 2005. This may explain why recent geographic-based HIV testing events have failed to identify many new HIV infections. Between 2006 and 2009, the Get Real, Get Tested campaign in NC tested over 4,500 people through door-to-door community testing events in high HIV morbidity areas and identified 38 new HIV infections (0.8%).[162] This positivity rate was much lower than screening positivity rates observed for non-traditional testing sites (1.5%) and community health centers (1.5%).[1] Based on the low yield from Get

Real, Get Tested, NC DHHS decided in 2010 that door-to-door testing should be minimal and focused. The disappearance of core areas in our study may lend weight to eliminating door-to-door, geographic-based testing altogether.

The high prevalence of Internet use among MSM and the greater distance to sexual partners observed among Internet users suggest that Internet-based interventions may be preferable to interventions targeting specific locations. MSM are receptive to Internet-based interventions, such as chat room or message board discussions or educational services.[124-127] Such interventions have been used to increase HIV testing, encourage disclosure to partners, and reduce high-risk behaviors that lead to HIV transmission.[128-131]

HIV surveillance data after 2007 should be analyzed for clusters of infection in order to determine if clusters really dissolved in NC following 2005 or if they re-emerged in later years. Data from the other surveillance regions in NC could also be examined to see if a similar reduction in clustering occurred concurrently with increased online sex-seeking. Network analyses measuring Internet use and spatial and geodesic distance to partners could also provide evidence as to whether or not Internet use is linking sexual networks and leading to increased HIV transmission outside of a core group. Mathematical models may provide the best indication of whether or not Internet use could feasibly lead to the dissolution of core HIV areas and the level of Internet use that would need to be reached in a population to produce such an outcome.

Aim 2: Several of our model predictors in the second aim risk score algorithm were lower risk behaviors for HIV transmission. While this may seem counterintuitive, it is likely explained by the fact that persons with lower risk profiles were able to provide DIS with more locating information for their sexual partners. Index cases that reported anonymous sex or crack use and named more sex partners were more likely to report partners that could not be located or refused testing compared to those of a lower risk profile. Some DIS may perceive index cases with higher risk behaviors as more worthwhile pursuits with respect to identifying newly infected partners because they are engaging in risk behaviors that facilitate HIV transmission. Our finding that lower risk index cases named more newly infected partners may be an important result to share with NC DIS in order to shape prioritization of partner notification.

Due to the relatively poor performance of the risk score algorithm in predicting new HIV infection in a sexual partner, it is unlikely that NC DHHS will implement the algorithm for partner notification. However, as noted above, we still believe it is useful in shaping perceptions of which index cases are more likely to yield newly positive partners and may be a useful tool in determining the amount of time DIS should spend attempting to locate a partner. Currently DIS must complete a 17-item checklist of locating tactics before declaring that a partner is unable to be located for PCRS. If the algorithm predicts that a particular partnership is less likely to result in the identification of a newly infected partner, DIS could complete a reduced checklist for this partner before declaring that he is unable to be located.

For the foreseeable future, DIS will continue to notify all named partners of HIV-infected index cases since it is mandated in the NC Administrative Code. However, if the current economic environment persists and DIS continue to be overloaded with cases to pursue without improvements in work conditions (e.g., better remuneration and prospects for career advancement), NC may need to modify the Administrative Code to specify a prioritization scheme for PCRS. A risk score algorithm would be particularly useful in this instance. Models with increased predictive performance could be pursued by including variables that we were unable to include. The strongest predictors for having an undiagnosed HIV-infected partner, such as type of sex, were undocumented in our study. Although the risk of transmitting HIV via saliva is very low,[142] the odds of HIV transmission during receptive anal intercourse are much higher than the odds of transmission during insertive anal sex or vaginal sex.[143, 144] Therefore, the inclusion of type of sex would likely improve the predictive power of the model. DIS occasionally note type of sex in the frequency of sex field in STD*MIS (e.g., 2x vaginal sex). In order for type of sex to be included in a predictive model, a specific field for this variable would need to be included on the field report form completed for each partner. Other variables that we were unable to include were length of the partnership and frequency of sex, which are also important predictors of HIV transmission. These data were only abstracted for located partners in our study because we initially proposed to include only partners with known HIV status in our model. These variables were predictive in a model that included only partners with known HIV status, and would therefore likely be predictors in a model that included all named partners regardless of HIV status at the time of partner notification.

Public health authorities are currently engaged in a reappraisal of PCRS that is long overdue. This includes a proposal for a randomized control trial to evaluate the effectiveness of partner notification in identifying new cases of HIV. Since most people with newly diagnosed HIV in other states are not currently interviewed for PCRS, a randomized trial that compares early versus delayed PCRS interviews should not present an ethical problem.[163] If most partners are tested for HIV prior to delayed provider referral, the cost of partner notification may outweigh the benefit. This result could shift states like NC to reconsider notification of all partners.

Aim 3: Our third aim showed that control measure violators in NC were more likely to be black and identify as MSM or MSM/W—populations disproportionately affected by HIV-related stigma. This stigma continues to inform perceptions and shape the behavior of PLWHA, making disclosure of HIV status difficult. Criminalization of HIV may serve as a barrier to HIV prevention if it increases stigma associated with HIV infection rather than deterring behaviors that transmit HIV.[95, 96] Providing limited case management to reduce unsafe sexual behaviors may be more effective in reducing control measure violation compared to the current system of investigation and potential prosecution.

While this risk score algorithm also showed only moderate predictive power, we feel that this algorithm is implementable in its current form. Because case management can take on a broad array of service definitions, NC DHHS may consider providing only less-intensive, low cost case management services such as a few phone calls to an HIV-infected person following diagnosis to encourage linkage to care. Currently, most index cases are passively referred to care by their diagnosing physician, post-test counselor, or DIS. Active referral via follow-up phone calls by a bridging case manager would be beneficial for all HIV-infected persons, but could be prioritized first for those identified in the algorithm as engaging in high-risk sexual behaviors and struggling with HIV disclosure to sexual partners. NC DHHS ultimately needs to weigh the potential public health cost of potential continued HIV transmission by HIV-infected persons aware of their status against the cost of providing limited case management in order to determine the proportion of the population receiving intervention. Determination of the tradeoff could be done formally with a cost-utility analysis, or health departments could take a more intuitive approach and consider current cost constraints and resource limitations. Ultimately, costs could be constrained by implementing only minimal case management services initially.

An important follow-up study of Aim 3 would be to examine the prevalence of serosorting among CMVs in order to examine whether STIs are acquired from HIV seroconcordant or serodiscordant partners. Such a study would provide evidence on whether serosorting reduces a person’s risk of acquiring other STIs or only HIV and could inform future recommendations on serosorting as a risk reduction practice for PLWHA. Because DIS interview an index case’s named partners following an alleged CMV violation, the data on the index case’s partnerships in the year prior to violation and the HIV status of the partners should be available in STD*MIS. However, the risk of prosecution for control measure violation may dissuade an individual from naming any known serodiscordant partners, making it difficult to obtain an unbiased estimate of the association between serosorting and STD acquisition.

Data on linkage to care in NC would be helpful in determining the level or intensity of intervention needed to reduce behaviors that lead to control measure violation. It was unknown in our study if the CMVs were successfully linked to care following their initial DIS interview or if they were in care at the time of the control measure violation. If non-CMVs were more likely to be in care compared to CMVs, linkage to care and maintenance in care for those indicated by the model may be enough to reduce the incidence of control measure violation. The CDC is currently sponsoring a national patient survey called the Medical Monitoring Project designed to answer questions about healthcare utilization after

HIV diagnosis. North Carolina is one of the study sites, and it may be possible through the Enhanced HIV/AIDS Reporting System (eHARS) to link study participation with NC DIS records. This would allow us to examine the effect of linkage to care on future control measure violation.

Conclusion

In 2009 the National Coalition of STD Directors reported on a study of the effects of the economic crisis on STI programs and public health infrastructure. The lead author of the study noted that funding cuts to these programs "threaten our national ability to control both sexually transmitted diseases and our entire public health infrastructure." As state revenues continue to decline, it is imperative that cuts to HIV program resources are based on evidence of where resources are most essential. Together, these three aims have added to our understanding of where and how limited resources could be allocated most efficiently to reduce HIV transmission in NC.

APPENDIX A. Index Case Abstraction Form

Eligibility

A1: Abstractor: ___ ___ ___ **A2:** Date of Abstraction ___ ___ / ___ ___ / ___ ___

A3a: Case ID: ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

A3b: Case ID: ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

A4: _Lot #: ___ ___ ___ **--** ___ ___ **A5:** Region: __R1 __R2 ___R3 ___R4 __R5 __R6

A6: Was the case interviewed directly by DIS? ___1 YES → **A7**: Date: __ __ / __ __ / __ __

___2 NO → **DO NOT ABSTRACT**

A8: Is the case above the age of 10? ___1 YES ___2 NO → **DO NOT ABSTRACT**

Demographics

B1: **HIV/AIDS code(s)** ___900 ___901 ___950

B2: **Syphilis code(s)** ___710 ___720 ___730 ___740 ___745 ___750 ___745

B3: Age ___ ___ **B4:** Date of Birth ___ ___ / ___ ___ / ___ ___

B5: Race ___1 W ___2 B ___3 A/PI ___4 AI/AN ___5 O/U **B6**: Ethnicity ___1 NON-His ___2 His

B7: Gender ___1 Male ___2 Female ___3 Transgender (Circle: Male to Female OR Female to Male)

B8: Pregnant ___1 YES **a)** ___ ___ weeks ___2 NO ___3 Unk/not doc

B9: Children ___1 YES ___2 NO ___3 Unk/not doc **B9a)** Number of children ___ ___

B10: Immigrated to the US? ___1 YES ___2 NO ___3 Unk/not doc **B11**: Year: _________

B12: Immigrated from: ___1 Central Am. ___2 South Am. ___3 Africa ___4 Other

a) specify country_________________

B13: Employment status: ___1 Employed ___2 Unemployed ___3 Unk/not doc

B14: Occupation / place of work:_______________________________________

B15: Comments

Institutions

C1: Incarcerated currently or previously? ___1 YES ___2 NO ___3 Unk/not doc

C2: IF YES: ___1 Jail ___2 Prison ___3 Both ___4 Unk/not doc

C3: Current College or University Student ___1 YES ___2 NO ___3 Unk/not doc

C4: Recent college graduate (within past 12 months) ___1 YES ___2 NO ___3 Unk/not doc

C5: College(s) attendance: a)__ __ ________________ b)__ __ ________________

code name code name

C6: Comments:

HIV History

D1: Date of first positive test: ___ ___ / ___ ___ / ___ ___

D2: Previous HIV test(s): ___1 YES ___2 NO ___3 Unk/not doc

D3: *IF YES:* Date of last negative test: ___ ___ / ___ ___ / ___ ___

D4: Diagnosis location

___1 Student Health
___2 Private MD
___3 Emergency Dept
___4 Health Department
___5 Prenatal
___6 Delivery
___7 Community screening
___8 Institutional screening
___9 Hospital

D4a: Specify location name ________________

D5a: CD4 count __________________ **D5b**: Test date ___ ___ / ___ ___ / ___ ___

D6a: Viral load __________________ **D6b**: Test date ___ ___ / ___ ___ / ___ ___

D7: Is index case a Control Measure Violator?: ___1 YES → **D7a**: Date: __ __ / __ __ / __ __

___2 NO

D8: Were HIV control measures signed by index case?: ___1 YES ___2 NO ___3 Unk

D9: Comments

STD History

F1: CURRENT STD @ time of HIV dx

___1 YES
___2 NO
___3 unk/not doc

F2: Specify CURRENT STD(s)

___1 Chlamydia
___2 Gonorrhea
___3 Genital Herpes
___4 Warts/HPV
___5 Chancroid
___6 Genital Herpes
___7 Trichomonas/Trich
___8 Syphilis
___9 Other
___10 unk/not doc

F3: PAST HX STD

___1 YES
___2 NO
___3 unk/not doc

F4: Specify PAST STD(s)

___1 Chlamydia
___2 Gonorrhea
___3 Genital Herpes
___4 Warts/HPV
___5 Chancroid
___6 Genital Herpes
___7 Trichomonas or Trich
___8 Syphilis
___9 Other
___10 unk/not doc

F5: Comments

Drug, Internet, and Club Use

G1: Any recreational drug use? ___$_1$ YES ___$_2$ NO → **SKIP** ___$_3$ Unk/not doc → **SKIP**

	a) Any use?	**b) Last 12 months?**	**c) Comments**
G2: Marijuana			
G3: Crack			
G4: Cocaine			
G5: Heroin			
G6: Methamphetamine			
G7: Ketamine			
G8: GHB			
G9: Viagra/Cialis/Levitra			
G10: Poppers			
G11: Club drugs			
G12: Ecstasy			
G13: Injection drug use			

H0: Any Internet/chat line use? ___$_1$ YES ___$_2$ NO → **SKIP** ___$_3$ Unk/not doc → **SKIP**

Web site code (for any reason)	**Web site name or chat line name/PH #**	**Found sex partners**	**Anonymous sex partners**	**H6: Comments**
H1: ___ ___				
H2: ___ ___				
H3: ___ ___				
H4: ___ ___				
H5: ___ ___				

I0: Any Bar or Club use? ___$_1$ YES ___$_2$ NO → **SKIP**___$_3$ Unk/not doc → **SKIP**

Bar / Club code for any reason	**Bar / Club name**	**Found sex partners**	**Anonymous sex partners**	**I11: Comments**
I1: ___ ___				
I2: ___ ___				
I3: ___ ___				
I4: ___ ___				
I5: ___ ___				

Sexual Risk

J1: Ever been raped or sexually assaulted? ___$_1$ YES ___$_2$ NO ___$_3$ Unk/Not doc

J2: Any anonymous sex partners? ___$_1$ YES ___$_2$ NO ___$_3$ Unk/Not doc

J3: Exchanged sex for drugs or money? ___$_1$ YES ___$_2$ NO ___$_3$ Unk/Not doc

J4: Male bisexual sex partners? ___$_1$ YES ___$_2$ NO ___$_3$ Unk/Not doc

J4: Knew partner was HIV+ before having sex with them? ___$_1$ YES ___$_2$ NO ___$_3$ Unk/Not doc

J5: Any sex partners from out of state? ___$_1$ YES ___$_2$ NO ___$_3$ Unk/Not doc

J6: Any sex partners from another country? ___$_1$ YES ___$_2$ NO ___$_3$ Unk/Not doc

J7: Gender of sex partners (ever): ___$_1$ Male ___$_2$ Female ___$_3$ Unk/Not doc

J8: Gender of sex partners (last year): ___$_1$ Male ___$_2$ Female ___$_3$ Unk/Not doc

J9: Number of sex partners ever ___ ___ ___ OR ___ unk/not doc

J10: Number of sex partners in last year ___ ___ ___ OR ___ unk/not doc

J11: Number of sex partners listed in STD*MIS: ___ ___

J12: NUMBER OF SEX PARTNERS ABSTRACTED ___ ___ ___

J13: Any sex partners/contacts who are known positives (and not in lot)? ___$_1$ YES ___$_2$ NO

J14: Which ones?

a) CONTACT # ___ ___ Field record ___________________

b) CONTACT # ___ ___ Field record ___________________

J15: Comment

APPENDIX B. Results of Aim 2 Sensitivity Analysis

Of the variables included in the final predictive model in Aim 2, history of crack use and history of anonymous sex were collected from the DIS narratives as ‘yes,’ ‘no,’ or ‘undocumented.’ When all undocumented responses for crack use and anonymous sex were recoded as ‘yes’ rather than ‘no,’ the odds ratios for undiagnosed HIV-positive partners were no longer significant and were on the other side of the null such that crack use and anonymous sex were risk factors for the outcome. They were no longer included in the predictive model, which reduced the area under the ROC for the model to 0.617. The sensitivities and specificities of this model at different interview coverage levels were slightly lower than those of the final model with partnership data above.

REFERENCES

1. NC Department of Health and Human Services. *Epidemiologic Profile for HIV/STD Prevention and Care Planning, December 2010*. Available at: http://www.epi.state.nc.us/epi/hiv/epiprofile1210.html. Accessed on January 10, 2011.

2. MacDonald PD, Nelson AL, Hightow-Weidman L, et al. Disease intervention specialists as a resource in a public health emergency. *Biosecur Bioterror* 2007;5(3):239-48.

3. US Census Bureau. *Census Regions and Divisions of the United States*. Available at: http://www.census.gov/geo/www/us_regdiv.pdf. Accessed on May 8, 2009.

4. Reif S, Geonnotti KL, Whetten K. HIV Infection and AIDS in the Deep South. *Am J Public Health* 2006;96(6):970-3.

5. Centers for Disease Control and Prevention, *Cases of HIV Infection and AIDS in the United States and Dependent Areas, 2005*, in *HIV AIDS Surveillance Reports*. 2006.

6. The Kaiser Family Foundation. *State Health Facts*. Available at: http://www.statehealthfacts.org. Accessed on May 8, 2009.

7. Centers for Disease Control and Prevention. Sexually Transmitted Disease Surveillance, 2003. *Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention; 2004*.

8. Hall I, Geduld J, Boulos D, et al. Epidemiology of HIV in the United States and Canada: Current Status and Ongoing Challenges. *J Acquir Immune Defic Syndr* 2009;51(S1).

9. US Census Bureau. *Census 2000. Demographic Profiles: 100-percent and Sample Data*. Available at: http://www.census.gov/prod/2001pubs/c2kbr01-5.pdf. Accessed on May 8, 2009.

10. Doherty IA, Leone PA, Aral SO. Social determinants of HIV infection in the Deep South. *Am J Public Health* 2007;97(3):391; author reply 391-2.

11. Adimora AA, Schoenbach VJ. Contextual factors and the black-white disparity in heterosexual HIV transmission. *Epidemiology* 2002;13(6):707-12.

12. Adimora AA, Schoenbach VJ, Martinson F, et al. Concurrent sexual partnerships among African Americans in the rural south. *Ann Epidemiol* 2004;14(3):155-60.

13. Centers for Disease Control and Prevention. Cases of HIV Infection and AIDS in Urban and Rural Areas of the United States, 2006. *HIV AIDS Surveillance Reports* 2006;13(2).

14. Heckman TG, Somlai AM, Peters J, et al. Barriers to care among persons living with HIV/AIDS in urban and rural areas. *AIDS Care* 1998;10(3):365-75.

15. Mondragon D, Kirkman-Liff B, Schneller ES. Hostility to people with AIDS: risk perception and demographic factors. *Soc Sci Med* 1991;32(10):1137-42.

16. Reif S, Golin CE, Smith SR. Barriers to accessing HIV/AIDS care in North Carolina: rural and urban differences. *AIDS Care* 2005;17(5):558-65.

17. NC Department of Health and Human Services. *New estimates profile new HIV infections in North Carolina* Available at: http://www.ncdhhs.gov/pressrel/2008/2008-9-12-New-estimates-profile-new-HIV.htm. Accessed on May 8, 2009.

18. Centers for Disease Control and Prevention. *National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention: North Carolina--2007 Profile*. Available at: http://www.cdc.gov/nchhstp/stateprofiles/North_Carolina/North_Carolina_Profiles.htm. Accessed on May 9, 2009.

19. Law DC, Serre ML, Christakos G, et al. Spatial analysis and mapping of sexually transmitted diseases to optimise intervention and prevention strategies. *Sex Transm Infect* 2004;80(4):294-9.

20. Yorke JA, Hethcote HW, Nold A. Dynamics and control of the transmission of gonorrhea. *Sex Transm Dis* 1978;5(2):51-6.

21. Garnett GP, Anderson RM. Contact tracing and the estimation of sexual mixing patterns: the epidemiology of gonococcal infections. *Sex Transm Dis* 1993;20(4):181-91.

22. Bwayo J, Plummer F, Omari M, et al. Human immunodeficiency virus infection in long-distance truck drivers in east Africa. *Arch Intern Med* 1994;154(12):1391-6.

23. Moses S, Plummer FA, Ngugi EN, et al. Controlling HIV in Africa: effectiveness and cost of an intervention in a high-frequency STD transmitter core group. *Aids* 1991;5(4):407-11.

24. Thomas JC, Tucker MJ. The development and use of the concept of a sexually transmitted disease core. *J Infect Dis* 1996;174 Suppl 2:S134-43.

25. Rothenberg RB. The geography of gonorrhea. Empirical demonstration of core group transmission. *Am J Epidemiol* 1983;117(6):688-94.

26. Ellen JM, Hessol NA, Kohn RP, et al. An investigation of geographic clustering of repeat cases of gonorrhea and chlamydial infection in San Francisco, 1989-1993: evidence for core groups. *J Infect Dis* 1997;175(6):1519-22.

27. Zenilman JM, Glass G, Shields T, et al. Geographic epidemiology of gonorrhoea and chlamydia on a large military installation: application of a GIS system. *Sex Transm Infect* 2002;78(1):40-4.

28. Scribner RA, Johnson SA, Cohen DA, et al. Geospatial methods for identification of core groups for HIV/AIDS. *Subst Use Misuse* 2008;43(2):203-21.

29. Alvarez-Dardet C, Marquez S, Perea EJ. Urban clusters of sexually transmitted diseases in the city of Seville, Spain. *Sex Transm Dis* 1985;12(3):166-8.

30. Arya OP, Rees E, Turner GC, et al. Epidemiology of penicillinase-producing Neisseria gonorrhoeae in Liverpool from 1977 to 1982. *J Infect* 1984;8(1):70-83.

31. Zenilman JM, Whittington WL, Frazier D, et al. Penicillinase-producing Neisseria gonorrhoeae in Dade County, Florida: phenotypic characterization of isolates from 1983, 1984, and 1986. *Sex Transm Dis* 1988;15(3):158-63.

32. Wasserheit JN, Aral SO. The dynamic topology of sexually transmitted disease epidemics: implications for prevention strategies. *J Infect Dis* 1996;174 Suppl 2:S201-13.

33. Blanchard JF. Populations, pathogens, and epidemic phases: closing the gap between theory and practice in the prevention of sexually transmitted diseases. *Sex Transm Infect* 2002;78 Suppl 1:i183-8.

34. Boily MC, Lowndes C, Alary M. The impact of HIV epidemic phases on the effectiveness of core group interventions: insights from mathematical models. *Sex Transm Infect* 2002;78 Suppl 1:i78-90.

35. Brunham RC. Core group theory: a central concept in STD epidemiology. *Venereology* 1997;10:28-31.

36. Woodhouse MEJ, Dye C, Etard J-F, et al. Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. *Proc Natl Acad Sci USA* 1997;94:338-342.

37. Jolly AM, Wylie JL. Gonorrhoea and chlamydia core groups and sexual networks in Manitoba. *Sex Transm Infect* 2002;78 Suppl 1:i145-51.

38. Becker KM, Glass GE, Brathwaite W, et al. Geographic epidemiology of gonorrhea in Baltimore, Maryland, using a geographic information system. *Am J Epidemiol* 1998;147(7):709-16.

39. Zenilman JM, Ellish N, Fresia A, et al. The geography of sexual partnerships in Baltimore: applications of core theory dynamics using a geographic information system. *Sex Transm Dis* 1999;26(2):75-81.

40. Rothenberg R, Muth SQ, Malone S, et al. Social and geographic distance in HIV risk. *Sex Transm Dis* 2005;32(8):506-12.

41. Michaud JM, Johnson SM, Ellen J. Comparison of sex partner meeting venues and residences of syphilis cases in Baltimore. *Sex Transm Dis* 2004;31(4):239-42.

42. Laumann EO, Ellingson S, Mahay J, et al. *The Sexual Organization of the City*. Chicago: University of Chicago Press, 2004.

43. Rothenberg R, Muth SQ. Large-network concepts and small-network characteristics: fixed and variable factors. *Sex Transm Dis* 2007;34(8):604-12.

44. Krieger N, Chen JT, Waterman PD, et al. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter?: the Public Health Disparities Geocoding Project. *Am J Epidemiol* 2002;156(5):471-82.

45. Steinberg S, Fleming P. The geographic distribution of AIDS in the United States: is there a rural epidemic? *J Rural Health* 2000;16(1):11-9.

46. Hall HI, Li J, McKenna MT. HIV in predominantly rural areas of the United States. *J Rural Health* 2005;21(3):245-53.

47. Williams ML, Bowen AM, Horvath KJ. The social/sexual environment of gay men residing in a rural frontier state: implications for the development of HIV prevention programs. *J Rural Health* 2005;21(1):48-55.

48. Horvath KJ, Bowen AM, Williams ML. Virtual and physical venues as contexts for HIV risk among rural men who have sex with men. *Health Psychol* 2006;25(2):237-42.

49. Bowen A, Williams M, Horvath K. Using the internet to recruit rural MSM for HIV risk assessment: sampling issues. *AIDS Behav* 2004;8(3):311-9.

50. Bull SS, McFarlane M, Rietmeijer C. HIV and sexually transmitted infection risk behaviors among men seeking sex with men on-line. *Am J Public Health* 2001;91(6):988-9.

51. Ross MW, Tikkanen R, Mansson SA. Differences between Internet samples and conventional samples of men who have sex with men: implications for research and HIV interventions. *Soc Sci Med* 2000;51(5):749-58.

52. Liau A, Millett G, Marks G. Meta-analytic examination of online sex-seeking and sexual risk behavior among men who have sex with men. *Sex Transm Dis* 2006;33(9):576-84.

53. Benotsch EG, Kalichman S, Cage M. Men who have met sex partners via the Internet: prevalence, predictors, and implications for HIV prevention. *Arch Sex Behav* 2002;31(2):177-83.

54. Bolding G, Davis M, Hart G, et al. Gay men who look for sex on the Internet: is there more HIV/STI risk with online partners? *Aids* 2005;19(9):961-8.

55. Elford J, Bolding G, Davis M, et al. The Internet and HIV study: design and methods. *BMC Public Health* 2004;4:39.

56. Elford J, Bolding G, Sherr L. Seeking sex on the Internet and sexual risk behaviour among gay men using London gyms. *Aids* 2001;15(11):1409-15.

57. Kim AA, Kent C, McFarland W, et al. Cruising on the Internet highway. *J Acquir Immune Defic Syndr* 2001;28(1):89-93.

58. McFarlane M, Bull SS, Rietmeijer CA. The Internet as a newly emerging risk environment for sexually transmitted diseases. *Jama* 2000;284(4):443-6.

59. Mettey A, Crosby R, DiClemente RJ, et al. Associations between internet sex seeking and STI associated risk behaviours among men who have sex with men. *Sex Transm Infect* 2003;79(6):466-8.

60. Rietmeijer CA, Bull SS, McFarlane M, et al. Risks and benefits of the internet for populations at risk for sexually transmitted infections (STIs): results of an STI clinic survey. *Sex Transm Dis* 2003;30(1):15-9.

61. Taylor M, Aynalem G, Smith L, et al. Correlates of Internet use to meet sex partners among men who have sex with men diagnosed with early syphilis in Los Angeles County. *Sex Transm Dis* 2004;31(9):552-6.

62. Hightow LB, Leone PA, Macdonald PD, et al. Men who have sex with men and women: a unique risk group for HIV transmission on North Carolina College campuses. *Sex Transm Dis* 2006;33(10):585-93.

63. Torrone E. Personal communication. December 2008.

64. Bolding G, Davis M, Hart G, et al. Heterosexual men and women who seek sex through the Internet. *Int J STD AIDS* 2006;17(8):530-4.

65. Ross MW, Mansson SA, Daneback K, et al. Characteristics of men who have sex with men on the internet but identify as heterosexual, compared with heterosexually identified men who have sex with women. *Cyberpsychol Behav* 2005;8(2):131-9.

66. Cooper A. *Sex and the Internet: A guidebook for clinicians*. New York: Brunner-Routledge, 2002.

67. Davis M, Hart G, Bolding G, et al. E-dating, identity and HIV prevention: theorising sexualities, risk and network society. *Sociol Health Illn* 2006;28(4):457-78.

68. Davis M, Hart G, Bolding G, et al. Sex and the Internet: gay men, risk reduction and serostatus. *Cult Health Sex* 2006;8(2):161-74.

69. Miller M, Serner M, Wagner M. Sexual diversity among black men who have sex with men in an inner-city community. *J Urban Health* 2005;82(1 Suppl 1):i26-34.

70. Couch D, Liamputtong P. Online dating and mating: the use of the internet to meet sexual partners. *Qual Health Res* 2008;18(2):268-79.

71. Centers for Disease Control and Prevention. Partner Counseling and Referral Services to Identify Persons with Undiagnosed HIV --- North Carolina, 2001. *MMWR* 2003;52(48):1181-1184.

72. Centers for Disease Control and Prevention. *HIV partner counseling and referral services*. Available at: http://www.cdc.gov/hiv/PUBS/pcrs.htm. Accessed on May 2, 2009.

73. Hogben M, McNally T, McPheeters M, et al. The effectiveness of HIV partner counseling and referral services in increasing identification of HIV-positive individuals a systematic review. *Am J Prev Med* 2007;33(2 Suppl):S89-100.

74. Bocour A, Udeagu C-C, Renaud T, et al. *HIV partner notification: which interviews yield the most new positive partners in NYC?*National Summit on HIV Diagnosis, Prevention, and Access to Care. November 19-21, 2008. Arlington, VA.

75. Macke BA, Maher JE. Partner notification in the United States: an evidence-based review. *Am J Prev Med* 1999;17(3):230-42.

76. Mathews C, Coetzee N, Zwarenstein M, et al. A systematic review of strategies for partner notification for sexually transmitted diseases, including HIV/AIDS. *Int J STD AIDS* 2002;13(5):285-300.

77. Landis SE, Schoenbach VJ, Weber DJ, et al. Results of a randomized trial of partner notification in cases of HIV infection in North Carolina. *N Engl J Med* 1992;326(2):101-6.

78. Andrus JK, Fleming DW, Harger DR, et al. Partner notification: can it control epidemic syphilis? *Ann Intern Med* 1990;112(7):539-43.

79. Gunn RA, Montes JM, Toomey KE, et al. Syphilis in San Diego County 1983-1992: crack cocaine, prostitution, and the limitations of partner notification. *Sex Transm Dis* 1995;22(1):60-6.

80. Cates W, Jr., Handsfield HH. HIV counseling and testing: does it work? *Am J Public Health* 1988;78(12):1533-4.

81. Rothenberg RB, Potterat JJ, Woodhouse DE, et al. Social network dynamics and HIV transmission. *Aids* 1998;12(12):1529-36.

82. Friedman SR, Neaigus A, Jose B, et al. Sociometric risk networks and risk for HIV infection. *Am J Public Health* 1997;87(8):1289-96.

83. Ghani AC, Ison CA, Ward H, et al. Sexual partner networks in the transmission of sexually transmitted diseases. An analysis of gonorrhea cases in Sheffield, UK. *Sex Transm Dis* 1996;23(6):498-503.

84. Branson BM, Handsfield HH, Lampe MA, et al. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. *MMWR Recomm Rep* 2006;55(RR-14):1-17; quiz CE1-4.

85. Centers for Disease Control and Prevention. Guidelines for the use of antiretroviral agents in HIV-infected adults and adolescents. *MMWR: Recommend Rep* 1998;47(RR-5):1-5.

86. National Institutes of Health. *Report of the NIH Panel to Define the Principles of Therapy of HIV Infection*. Bethesda, MD: National Institutes of Health, 1997.

87. Kalichman SC. HIV transmission risk behaviors of men and women living with HIV-AIDS: prevalence, predictors, and emerging clinical interventions. *Clin Psychol Sci Pract* 2000;7(1):32-47.

88. Bova C, Durante A. Sexual functioning among HIV-infected women. *AIDS Patient Care STDS* 2003;17(2):75-83.

89. Crepaz N, Marks G. Serostatus disclosure, sexual communication and safer sex in HIV-positive men. *AIDS Care* 2003;15(3):379-87.

90. Collins C, Morin S, Shriver Mea. *Designing Primary Prevention for People Living with HIV*. San Francisco, CA: AIDS Research Institute, University of California, San Francisco, 2000.

91. Centers for Disease Control and Prevention. *HIV Prevention Strategic Plan: Extended Through 2010*. Available at: http://www.cdc.gov/hiv/resources/reports/psp/. Accessed on May 8, 2009.

92. Centers for Disease Control and Prevention. *Advancing HIV Prevention: Interim Technical Guidance for Selected Interventions*. Available at: http://www.cdc.gov/hiv/topics/prev_prog/AHP/resources/guidelines/Interim-Guidance.htm. Accessed on May 8, 2009.

93. North Carolina Administrative Code. Communicable Disease Control. *Control Measures--HIV*. 10 NCAC 41A.0202. 1988.

94. Herek G. AIDS and stigma. *Am Behav Sci* 1999;42:1102-1112.

95. Shriver MD, Everett C, Morin SF. Structural interventions to encourage primary HIV prevention among people living with HIV. *Aids* 2000;14 Suppl 1:S57-62.

96. Burris S, Cameron E. The case against criminalization of HIV transmission. *Jama* 2008;300(5):578-81.

97. Barkan SE, Melnick SL, Preston-Martin S, et al. The Women's Interagency HIV Study. WIHS Collaborative Study Group. *Epidemiology* 1998;9(2):117-25.

98. Centers for Disease Control and Prevention. *HIV/AIDS among women*. Available at: http://www.cdc.gov/hiv/topics/women/resources/factsheets/women.htm. Accessed on May 5, 2009.

99. Gielen AC, Ghandour RM, Burke JG, et al. HIV/AIDS and intimate partner violence: intersecting women's health issues in the United States. *Trauma Violence Abuse* 2007;8(2):178-98.

100. Teti M, Bowleg L, Cole R, et al. A Mixed Methods Evaluation of the Effect of the Protect and Respect Intervention on the Condom Use and Disclosure Practices of Women Living with HIV/AIDS. *AIDS Behav* 2009.

101. Simon Rosser BR, Horvath KJ, Hatfield LA, et al. Predictors of HIV disclosure to secondary partners and sexual risk behavior among a high-risk sample of HIV-positive MSM: results from six epicenters in the US. *AIDS Care* 2008;20(8):925-30.

102. Crepaz N, Lyles CM, Wolitski RJ, et al. Do prevention interventions reduce HIV risk behaviours among people living with HIV? A meta-analytic review of controlled trials. *Aids* 2006;20(2):143-57.

103. Centers for Disease Control and Prevention. *Guidance: HIV Prevention Case Management*. Available at: http://www.cdc.gov/hiv/topics/prev_prog/CRCS/resources/PCMG/. Accessed on May 7, 2009.

104. Centers for Disease Control and Prevention, *Comprehensive Risk Counseling and Services (CRCS) Implementation Manual*. 2006: Atlanta, GA.

105. Effects of a behavioral intervention to reduce risk of transmission among people living with HIV: the healthy living project randomized controlled study. *J Acquir Immune Defic Syndr* 2007;44(2):213-21.

106. Richardson JL, Milam J, McCutchan A, et al. Effect of brief safer-sex counseling by medical providers to HIV-1 seropositive patients: a multi-clinic assessment. *AIDS* 2004;18(8):1179-86.

107. Quinn TC, Wawer MJ, Sewankambo N, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. *N Engl J Med* 2000;342(13):921-9.

108. Gay CL, Napravnik S, Eron JJ, Jr. Advanced immunosuppression at entry to HIV care in the southeastern United States and associated risk factors. *Aids* 2006;20(5):775-8.

109. Hampton KH, Fitch MK, Allshouse WB, et al. Mapping health data: improved privacy protection with donut method geomasking. *Am J Epidemiol*;172(9):1062-9.

110. Allshouse WB, Fitch MK, Hampton KH, et al. Geomasking sensitive health data and privacy protection: an evaluation using an E911 database. *Geocarto Int*;25(6):443-452.

111. Kulldorff M, Athas WF, Feurer EJ, et al. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. *Am J Public Health* 1998;88(9):1377-80.

112. Pfeiffer D. *Spatial analysis in epidemiology*. Oxford ; New York: Oxford University Press, 2008.

113. Kleinbaum DG. *Applied regression analysis and other multivariable methods*. 4th ed. Belmont, CA: Thomason/Brooks/Cole, 2008.

114. Maldonado G, Greenland S. Simulation study of confounder-selection strategies. *Am J Epidemiol* 1993;138(11):923-36.

115. Harrell FE. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. New York: Springer, 2001.

116. Kleinbaum DG, Klein M. *Logistic regression: a self-learning text*. 2nd ed. New York: Springer, 2002.

117. Sun GW, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol* 1996;49(8):907-16.

118. Katz DA, Hogben M, Dooley SW, Jr., et al. Increasing public health partner services for human immunodeficiency virus: results of a second national survey. *Sex Transm Dis*;37(8):469-75.

119. Steyerberg EW, Harrell FE, Jr., Borsboom GJ, et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 2001;54(8):774-81.

120. Cox DR, Oakes D. *Analysis of survival data*. New York: Chapman and Hall, 1984.

121. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143(1):29-36.

122. Jennings JM, Curriero FC, Celentano D, et al. Geographic identification of high gonorrhea transmission areas in Baltimore, Maryland. *Am J Epidemiol* 2005;161(1):73-80.

123. Wohlfeiler D, Potterat JJ. Using gay men's sexual networks to reduce sexually transmitted disease (STD)/human immunodeficiency virus (HIV) transmission. *Sex Transm Dis* 2005;32(10 Suppl):S48-52.

124. Bolding G, Davis M, Sherr L, et al. Use of gay Internet sites and views about online health promotion among men who have sex with men. *AIDS Care* 2004;16(8):993-1001.

125. Bull SS, McFarlane M, King D. Barriers to STD/HIV prevention on the Internet. *Health Educ Res* 2001;16(6):661-70.

126. Rhodes SD. Hookups or health promotion? An exploratory study of a chat room-based HIV prevention intervention for men who have sex with men. *AIDS Educ Prev* 2004;16(4):315-27.

127. Klausner JD, Levine DK, Kent CK. Internet-based site-specific interventions for syphilis prevention among gay and bisexual men. *AIDS Care* 2004;16(8):964-70.

128. Chiasson MA, Shaw FS, Humberstone M, et al. Increased HIV disclosure three months after an online video intervention for men who have sex with men (MSM). *AIDS Care* 2009;21(9):1081-9.

129. Blas MM, Alva IE, Carcamo CP, et al. Effect of an online video-based intervention to increase HIV testing in men who have sex with men in Peru. *PLoS One*;5(5):e10448.

130. Rhodes SD, Hergenrather KC, Duncan J, et al. A pilot intervention utilizing Internet chat rooms to prevent HIV risk behaviors among men who have sex with men. *Public Health Rep*;125 Suppl 1:29-37.

131. Rosser BR, Oakes JM, Konstan J, et al. Reducing HIV risk behavior of men who have sex with men through persuasive computing: results of the Men's INTernet Study-II. *AIDS*;24(13):2099-107.

132. Hogben M. Partner notification for sexually transmitted diseases. *Clin Infect Dis* 2007;44 Suppl 3:S160-74.

133. Recommendations for partner services programs for HIV infection, syphilis, gonorrhea, and chlamydial infection. *MMWR Recomm Rep* 2008;57(RR-9):1-83; quiz CE1-4.

134. Sellors J, Zimic-Vincetic M, Howard M, et al. Predictors of positivity for hepatitis B and the derivation of a selective screening rule in a Canadian sexually transmitted disease clinic. *J Clin Virol* 1998;11(1):85-91.

135. Gunn RA, Murray PJ, Brennan CH, et al. Evaluation of screening criteria to identify persons with hepatitis C virus infection among sexually transmitted disease clinic clients: results from the San Diego Viral Hepatitis Integration Project. *Sex Transm Dis* 2003;30(4):340-4.

136. Sellors JW, Pickard L, Gafni A, et al. Effectiveness and efficiency of selective vs universal screening for chlamydial infection in sexually active young women. *Arch Intern Med* 1992;152(9):1837-44.

137. Stergachis A, Scholes D, Heidrich FE, et al. Selective screening for Chlamydia trachomatis infection in a primary care population of women. *Am J Epidemiol* 1993;138(3):143-53.

138. Finelli L, Nakashima AK, Hillis S, et al. Selective screening versus presumptive treatment criteria for identification of women with chlamydial infection in public clinics: New Jersey. *Am J Obstet Gynecol* 1996;174(5):1527-33.

139. Marrazzo JM, Fine D, Celum CL, et al. Selective screening for chlamydial infection in women: a comparison of three sets of criteria. *Fam Plann Perspect* 1997;29(4):158-62.

140. Al-Tayyib AA, Miller WC, Rogers SM, et al. Evaluation of risk score algorithms for detection of chlamydial and gonococcal infections in an emergency department setting. *Acad Emerg Med* 2008;15(2):126-35.

141. Chen MY, Fairley CK, De Guingand D, et al. Screening pregnant women for chlamydia: what are the predictors of infection? *Sex Transm Infect* 2009;85(1):31-5.

142. Cohen MS, Shugars DC, Fiscus SA. Limits on oral transmission of HIV-1. *Lancet* 2000;356(9226):272.

143. Baggaley RF, White RG, Boily MC. HIV transmission risk through anal intercourse: systematic review, meta-analysis and implications for HIV prevention. *Int J Epidemiol*;39(4):1048-63.

144. Seidlin M, Vogler M, Lee E, et al. Heterosexual transmission of HIV in a cohort of couples in New York City. *AIDS* 1993;7(9):1247-54.

145. Hurt CB, Matthews DD, Calabria MS, et al. Sex with older partners is associated with primary HIV infection among men who have sex with men in North Carolina. *J Acquir Immune Defic Syndr*;54(2):185-90.

146. Bingham TA, Harawa NT, Johnson DF, et al. The effect of partner characteristics on HIV infection among African American men who have sex with men in the Young Men's Survey, Los Angeles, 1999-2000. *AIDS Educ Prev* 2003;15(1 Suppl A):39-52.

147. Knolle H. Age preference in sexual choice and HIV transmission. *AIDS* 1990;4(7):698.

148. Golden MR, Stekler J, Kent JB, et al. An evaluation of HIV partner counseling and referral services using new disposition codes. *Sex Transm Dis* 2009;36(2):95-101.

149. Taylor MM, Mickey T, Winscott M, et al. Improving partner services by embedding disease intervention specialists in HIV-clinics. *Sex Transm Dis*;37(12):767-70.

150. Han Y, Coles FB, Muse A, et al. Assessment of a geographically targeted field intervention on gonorrhea incidence in two New York State counties. *Sex Transm Dis* 1999;26(5):296-302.

151. Brewer DD, Garrett SB. Evaluation of interviewing techniques to enhance recall of sexual and drug injection partners. *Sex Transm Dis* 2001;28(11):666-77.

152. Brewer DD, Garrett SB, Kulasingam S. Forgetting as a cause of incomplete reporting of sexual and drug injection partners. *Sex Transm Dis* 1999;26(3):166-76.

153. Marcus JL, Katz MH, Katz KA, et al. Prediction model to maximize impact of syphilis partner notification--San Francisco, 2004-2008. *Sex Transm Dis*;37(2):109-14.

154. Lima VD, Harrigan R, Bangsberg DR, et al. The combined effect of modern highly active antiretroviral therapy regimens and adherence on mortality over time. *J Acquir Immune Defic Syndr* 2009;50(5):529-36.

155. Baker JV, Peng G, Rapkin J, et al. CD4+ count and risk of non-AIDS diseases following initial treatment for HIV infection. *AIDS* 2008;22(7):841-8.

156. Crepaz N, Marks G, Liau A, et al. Prevalence of unprotected anal intercourse among HIV-diagnosed MSM in the United States: a meta-analysis. *AIDS* 2009;23(13):1617-29.

157. Weinhardt LS, Kelly JA, Brondino MJ, et al. HIV transmission risk behavior among men and women living with HIV in 4 cities in the United States. *J Acquir Immune Defic Syndr* 2004;36(5):1057-66.

158. Chen SY, Gibson S, Weide D, et al. Unprotected anal intercourse between potentially HIV-serodiscordant men who have sex with men, San Francisco. *J Acquir Immune Defic Syndr* 2003;33(2):166-70.

159. Berg RC, Grimes R. Do traditional risk factors predict whether men who have sex with men engage in unprotected anal intercourse? The need for locally based research to guide interventions. *Health (London)*.

160. Snowden JM, Raymond HF, McFarland W. Seroadaptive behaviours among men who have sex with men in San Francisco: the situation in 2008. *Sex Transm Infect*;87(2):162-4.

161. Gorbach PM, Weiss RE, Jeffries R, et al. Behaviors of recently HIV-infected men who have sex with men in the year postdiagnosis: effects of drug use and partner types. *J Acquir Immune Defic Syndr*;56(2):176-82.

162. Davis AB. *Get Real. Get Tested.* 2010 STD Prevention Conference. Atlanta, GA: March 8-11, 2010.

163. Golden MR. HIV partner counseling and referral services: finally getting beyond the name. *Am J Prev Med* 2007;33(2 Suppl):S84-5.