PSC Working Paper Series

6-9-2007

The Likoma Network Study: Context, Data Collection and Initial Results

Stephane Helleringer
University of Pennsylvania, hellerin@ssc.upenn.edu

Hans-Peter Kohler
University of Pennsylvania, HPKOHLER@POP.UPENN.EDU

Agnes Chimbiri
University of Malawi, achimbiri@medcol.mw

Praise Chatonda
Market Research Center, Celtel Lilongwe, Malawi

James Mkandawire
Montfort Hospital, Ntchalo, Malawi

The Likoma Network Study: Context, Data Collection and Initial Results

Abstract
The sexual networks connecting members of a population have important consequences for the spread of sexually transmitted diseases including HIV. However, very few datasets currently exist that allow an investigation of the structure of sexual networks, particularly in sub-Saharan Africa where HIV epidemics have become generalized. In this paper, we describe the context and methods of the Likoma Network Study (LNS), a survey of complete sexual networks we conducted in Likoma island (Malawi) between October 2005 and March 2006. We start by reviewing theoretical arguments and empirical studies emphasizing the importance of network structures for the epidemiology of sexually and transmitted diseases. We describe the island setting of this study, and argue that the choice of an island as research site addresses the possible sources of bias in the collection of complete network data. We then describe in detail our empirical strategy for the identification of sexual networks, as well as for the collection of biomarker data (HIV infection). Finally, we provide initial results relating to the socioeconomic context of the island, the size and composition of sexual networks, the prevalence of HIV in the study population, the quality of the sexual network data, the determinants of successful contact tracing during the LNS, and basic measures of network connectivity.

Keywords
Africa, Age, AIDS, Biomarkers, Birth, Birth control, Births, Census, Children, Condom use, Contexts, Contraception, Data, Data Collection, Death, Demographic measures, Demography, Developing countries, Disease, Diseases, Divorce, Education, Empirical study, Employment, Epidemics, Epidemiology, Extra-marital partners, Family, Fieldwork, First sex, Gender, Geographic location, Geography, Global Positioning Systems, GPS, Health, Health Behavior, Health Surveys, HIV, HIV infection, HIV prevalence, HIV risk factors, HIV risk perception, HIV risk perceptions, HIV risks, HIV status, HIV testing, HIV tests, HIV transmission, HIV/AIDS, Household, Household informants, Household structure, Households, Interviews, Island, Life course, Likoma, Likoma Network Study (LNS) , Malawi, Marital behavior, Marital dissolution, Marital history, Marital partners, Marriage, Marriage processes, Methods, Migrants, Migration, Mortality, Network epidemiology, Non-marital partners, Out-of-wedlock childbirth, Partners, Polygamy, Population Studies, Premarital partners, Premarital sex, Public Health, Quantitative, Rapid HIV test, Relationships, Remarriage, Residential Location, Sex, Sexual activity, Sexual behavior, Sexual behaviors, Sexual initiation, Sexual network, Sexual network partners, Sexual networks, Sexual partners, Sexual partnerships, Sexually transmitted diseases, Sexually transmitted infections, Social network, Social network analysis, Social networks, Sociocentric, Socio-demographic Characteristics, Socioeconomic, Sociology, Spouses, Statistics, STD, STI, Sub-Saharan Africa, Survey Data, Surveys, Timing of marriage, Unmarried partners

Disciplines
Demography, Population, and Ecology | Social and Behavioral Sciences | Sociology

Comments
Helleringer, Stéphane, Hans-Peter Kohler, Agnes Chimbiri, Praise Chatonda, and James Mkandawire. 2007. "The Likoma Network Study: Context, Data Collection and Initial Results." *PSC Working Paper Series PSC 07-05.*

The Likoma Network Study: Context, Data Collection and Initial Results

Stéphane Helleringer Hans-Peter Kohler Agnes Chimbiri Praise Chatonda
James Mkandawire*

June 9, 2007

Abstract

The sexual networks connecting members of a population have important consequences for the spread of sexually transmitted diseases including HIV. However, very few datasets currently exist that allow an investigation of the structure of sexual networks, particularly in sub-Saharan Africa where HIV epidemics have become generalized. In this paper, we describe the context and methods of the ***Likoma Network Study*** (LNS), a survey of complete sexual networks we conducted in Likoma island (Malawi) between October 2005 and March 2006. We start by reviewing theoretical arguments and empirical studies emphasizing the importance of network structures for the epidemiology of sexually and transmitted diseases. We describe the island setting of this study, and argue that the choice of an island as research site addresses the possible sources of bias in the collection of complete network data. We then describe in detail our empirical strategy for the identification of sexual networks, as well as for the collection of biomarker data (HIV infection). Finally, we provide initial results relating to the socioeconomic context of the island, the size and composition of sexual networks, the prevalence of HIV in the study population, the quality of the sexual network data, the determinants of successful contact tracing during the LNS, and basic measures of network connectivity.

1 Introduction

Sexual networks are the primary mechanism through which HIV is spread and transformed in Sub-Saharan Africa (SSA). Theoretical network models have shown that individuals' positions within these sexual networks, and the structural characteristics of the network itself, are important determinants of HIV infection risks and disease dynamics (Ghani and Garnett 2000; Kretzschmar and Morris 1996; Newman 2002b). Several features of sexual networks that are predicted by these models to enhance the spread of HIV have been empirically documented in SSA, including *concurrency* of sexual partnerships (Morris 1997), *skewed degree distributions* of sexual networks (Anderson

*Helleringer is Ph.D. student at the University of Pennsylvania, Population Studies Center, 3718 Locust Walk, Philadelphia, PA 19104-6299, USA; *Email:* hellerin@ssc.upenn.edu. Kohler is Professor of Sociology at the University of Pennsylvania, Population Studies Center, 3718 Locust Walk, Philadelphia, PA 19104-6299, USA; *Email:* hpkohler@pop.upenn.edu. Chimbiri is Director of the Centre for Reproductive Health at the College of Medicine, University of Malawi, Private Bag 360, Chichiri, Blantyre 3, Malawi; *Email:* achimbiri@medcol.mw. Chatonda is at the market research center, Celtel Lilongwe (Malawi). Mkandawire is HIV/AIDS coordinator, Montfort Hospital in Ntchalo (Chikwawa district, Malawi). We gratefully acknowledge the support for this research through NIH grants RO1 HD044228 and RO1 HD/MH41713, and a PARC/Boettner/PSC Pilot Grant by the Population Studies Center, University of Pennsylvania. We also acknowledge the contributions of James H. Jones at Stanford University, Paul Hewett at the Population Council, and Agatha Bula and George Joaki at UNC-Lilongwe during this project.

Figure 1: Likoma Island on Lake Malawi

and May 1991; Jones and Handcock 2003b), and *large and robust connected components* (Moody et al. 2003; Moody and White 2003). Despite this evidence, empirical network studies of HIV infection risks and disease dynamics in SSA remain very limited. Available data on sexual networks are often based on small populations, frequently restricted to ego-centric rather than complete networks, and with the exception of the study described in this paper, not based on an integrated design that includes tracing of sexual networks, HIV testing, and extensive socioeconomic data for all members a population.

In this paper we describe and document the ***Likoma Network Study*** (LNS), an innovative project that conducted a *complete sexual network survey* in Likoma, a small island on Lake Malawi (Figure 1) with high HIV prevalence. The data collected as part of the LNS and described in this paper include (*i*) data on (quasi) complete sexual networks covering the young adult population (aged 18–25 years) in seven villages of Likoma, (*ii*) detailed data on the socioeconomic and demographic situation, subjective health, and HIV/AIDS related behaviors, attitudes and risk-perceptions of individuals and their sexual network partners, (*iii*) HIV status of respondents and their sexual/social network partners, (*iv*) geographic locations (GPS data) of respondents and their network partners, and (*v*) limited data on migration to and from the island. By choosing Likoma, our study takes advantage—similar to other epidemiological island studies (Cliff et al. 2000; Whittaker 1999)—of the limited range of mobility and the well-defined population boundaries of insular communities. These features imply that a high proportion of the islanders' sexual partners reside on the island, thereby increasing the probability of tracing sexual partners.

The complete network design of the LNS enables us to document the population-level structure of sexual networks in Likoma. We used these data to assess whether the structure of sexual networks in this sub-Saharan setting is compatible with broad diffusion of HIV to the general population and document the position of HIV-infected individuals within the reconstructed sexual networks (Helleringer and Kohler 2007). Studies using a design similar to that of the LNS have been conducted in different contexts (e.g, Bearman et al. 2004; Klovdahl et al. 1994), but were lacking for African populations with generalized HIV epidemics.

Our presentation of the LNS is structured as follows: First, we review empirical and theoretical studies emphasizing the role of network structure in explaining patterns of STI spread within and across populations (Section 2). Second, we present the setting of the study (Section 3). Third, we discuss practical difficulties associated with the collection of sociocentric network and describe how these difficulties might be compounded in a sub-Saharan context (Section 4.1). We then motivate the choice of an island setting to conduct a complete sociocentric network survey, and show how it addresses most of these difficulties (Section 4.2). Fourth, we describe in detail our empirical strategy and procedures for the identification of sexual networks (Section 5). Finally, we discuss initial results relating to HIV prevalence (Section 6.1), sexual behaviors and relationships (Section 6.2), and basic measures of network connectivity (Section 6.3.4).

2 Background: Network epidemiology

The classical models of mathematical epidemiology (e.g., Anderson and May 1991; Bailey 1975) rely on the assumption that sexual partners are randomly selected (i.e., the population is assumed to be well-mixed and unstructured). In this framework, two key measures to study epidemics are (1) the basic reproduction number, R_0, and (2) the final size of an epidemic s_∞. The basic reproduction number, R_0, is the expected number of secondary infections arising from a single, typical infectious individual in a completely susceptible population (Heesterbeek 2002). In a well-mixed and socially unstructured population (i.e., where individuals randomly select their partners among other members of the population), R_0 is the product of three quantities: the transmissibility of the infection τ, the duration of infectiousness δ, and the rate of contact between susceptible and infectious individuals $\bar{c}$. This latter parameter is the focus of the LNS.

Epidemics are nonlinear phenomena and R_0 is a threshold parameter. When $R_0 > 1$, an epidemic is certain in a deterministic model and has non-zero probability in a stochastic model. Strategies for disease control and eradication are aimed at bringing R_0 below the threshold of unity, i.e, when the average infection generates fewer secondary infections than necessary for replacement and the epidemic fades. In the well-mixed and unstructured case, the final size of the epidemic is given by the implicit equation $\log(s_\infty) = R_0(s_\infty - 1)$, which has exactly two roots on the interval $[0\ 1]$ when $R_0 > 1$. The smaller of these roots is the proportion of the population remaining uninfected at the end of an epidemic.

Because HIV is transmitted by intimate sexual contacts between partners, and because people employ varied and elaborate rules to choose their partners (Magruder 2006; Watkins 2004), HIV transmission dynamics in real populations are not well described by the classical epidemiological model. In other terms, $\bar{c}$ is generally a poor approximation of the patterns of contacts leading to the diffusion of an infection within a population. For instance, while African men (and to a lesser extent women) do not report having more sexual partners than men elsewhere, they tend to have more than one on-going long-term relation at any point in time. Partnerships in SSA can overlap for months, maybe years (Lagarde et al. 2001; Morris and Kretzschmar 2000). This pattern of sexual partnerships that overlap rather than follow each other sequentially (Kretzschmar and Morris 1996; Moody 2002), is one of several important characteristics of human sexual networks that violate the classical epidemiological model and importantly affect HIV infection risks and disease dynamics. *Concurrent partnerships* appear to increase the speed at which HIV spreads through a population, and have probably contributed to the rapid take-off of the HIV epidemic in SSA in the 1980s (Morris and Kretzschmar 2000). Other violations of the classical epidemiological model include: (1) *assortative mixing*, i.e., the selection of sexual partners based on their individual characteristics, can structure a network into communities within which the disease spreads rapidly, but across which the spread is slow (Laumann et al. 1994; Laumann and Youm 1999; Morris 1993); (2) *small worlds*, i.e., networks characterized by bridges joining otherwise disjoint clusters (Watts and Strogatz 1998; Watts 1999), can lead to thresholds and rapid disease diffusion to distant subpopulations; (3) *robust networks*, i.e., groups of persons tied together by more than one path in the sexual network, can decrease the ability to control the spread of HIV because redundant connections continue to transmit HIV even after some transmission paths are broken or eliminated (Moody et al. 2003; Potterat et al. 2002); (4) *skewed degree distributions*, i.e., networks containing individuals with a very high number of partners (high degree network members), can result in epidemics driven by promiscuous individuals (e.g., Liljeros et al. 2001; for a critical perspective, see Handcock and Jones 2004; Jones and Handcock 2003a).

While it is possible to simulate HIV disease dynamics taking these characteristics of human sexual networks into account (Hethcote et al. 1991), only detailed information on the sexual network structures in SSA allows to to properly calibrate these models. In addition, the investigation of sexual networks has important implications for disease prevention. For instance, in a simple heterogeneous epidemic model structured by degree of sexual activity, the basic reproduction number, and therefore epidemic threshold, is linearly proportional to the variance in partner numbers (e.g., Anderson and May 1991; Bailey 1975). Epidemics will therefore be more difficult to control in populations characterized by behavioral heterogeneity. The structure of sexual relations between the highest degree individuals and the general population will also determine the effectiveness of control interventions. The design of optimal interventions to control HIV therefore need to take the prevailing sexual network structure into account. However little is known about these aspects due to the lack of detailed sexual network data.

In summary, the literature reviewed above suggests a considerable potential for important and policy-relevant research based on an empirical investigation of the relationships between sexual networks, HIV infection risks, and HIV/AIDS disease dynamics. While sophisticated analytic methods have recently become available that allow these investigations (e.g., Koehly and Morris 2004), their application to high HIV-prevalence contexts in SSA has been hampered by a lack of suitable data. This is not surprising, as the empirical challenges are formidable: (*i*) network information needs to extend to (quasi) complete networks, including information on the structure of the network and the position of individuals within the network; and (*ii*) data need to be comprehensive, ranging from sexual networks to risk perceptions, sexual behaviors and health measures,

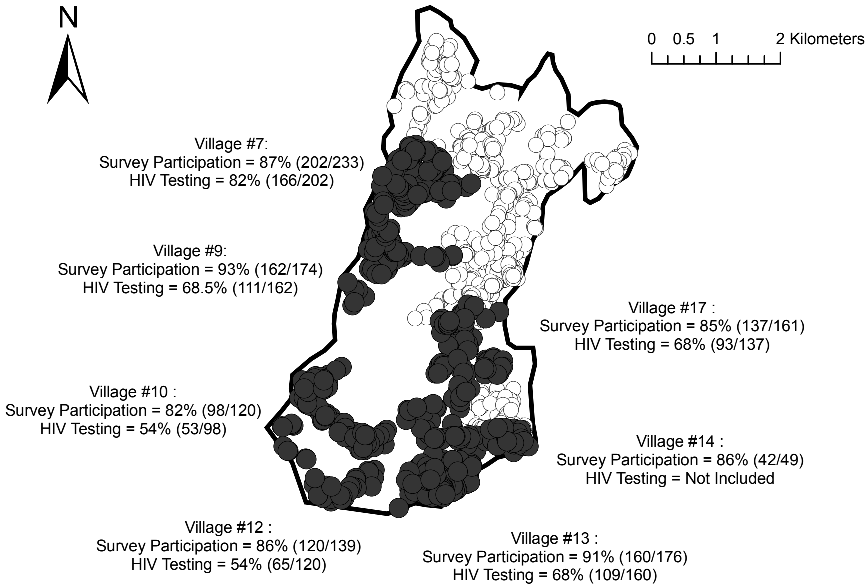

Figure 2: Geographic location of the sampled villages and village-specific participation rates. Each circle represents a dwelling unit. Dark circles represent housing units that were included in the sexual network survey. Empty circle represent housing units that were not included in this sampling frame. Denominators of the survey participation rates are the total number of eligible respondents (aged 18–35 and their spouses) in a given village, based on the initial household census. Denominators of the HIV testing participation rates are the total number of respondents who completed the sexual network survey in a given village. Island boundaries and location of dwelling units are approximate.

and be available for *both* respondents and their network partners.

3 Study Location and Context: Likoma Island

Likoma Island is located in the northern region of Lake Malawi (Figure 1). The population of the island is a little above 7,000 people living in a dozen villages, of which seven were chosen for the sexual network survey in the Likoma Network Study (Figure 2; see also Section 5). The population of Likoma is comprised primarily of the Nyanja and Tonga ethnic groups (60% of Likoma inhabitants are Nyanja by tribe, 25% are Tonga, and the remaining 15% belong to diverse ethnic groups present in mainland Malawi or Tanzania such as Tumbuka, Swahilis, Yaos, Chewas, etc.). The Tongas of Likoma are mostly found in two villages of the island (on the southern side), whereas the other (i.e., non-Nyanjas) ethnic groups reside mostly around the trading center. These differences in the ethnic composition of local populations also appear on the mainland of Malawi: if the Tumbuka ethnic group largely dominates in the northern region, the composition of the two other regions (southern and central) is more mixed.

The overall level of economic development in Likoma is quite low, and fishing is the main source of income for most households. The population of the islands is extremely young, as 50% of all inhabitants are below 15 years old, and there is a noticeable excess of women in Likoma as sex ratios at young adult ages (20–40) fluctuate around 0.85/0.9 on the island. These imbalanced sex ratios might be attributed to migration of young men to look for employment or further their schooling, as well as to differential mortality due to accidents (fishing) and HIV-related illnesses. Nevertheless, comparison with Census data (projections for 2005, NSO 2004) show that the imbalance of sex ratios is only slightly more pronounced in Likoma than in the rest of rural Malawi, where out-migration of young adult males to the larger cities of Lilongwe, Blantyre and Mzuzu, or abroad is also quite frequent.

Schooling is relatively widespread on Likoma: a large proportion of men and women between ages 15 and 49 have completed primary school. The median age at first birth in Likoma is 18 years, and among these births more than 50% take place out-of-wedlock. Indeed, marriages in Likoma happen relatively late (median age is 21 years for women and 26 for men) and are somewhat fragile (1/3 of marriages have ended after 10 years). Transportation to Likoma and Chizumulu Island is quite limited as only one boat travels weekly to the mainland of Malawi, although a few small canoes make daily trips to the Mozambican shore. Despite these constraints, inhabitants of the island travel frequently: 2/3 of males and more than 1/2 of females had gone to mainland Malawi in the year prior to the survey, while almost 1/2 of males and 1/3 of females had gone to Mozambique over that same time span. This high prevalence of travel is partially related to the fact that local waters around Likoma get agitated in March-April, and men from the islands migrate seasonally to the quieter waters of the Salima region in central Malawi (Mbenji Island). Our data does not allow estimating the duration of the various trips to the mainland, but casual observation suggests time spent outside of the island varies greatly with the purpose of the trip. Business trips to sell fish in mainland Malawi or collect firewood in Mozambique take a few days at the most. Trips to visit relatives and/or get treatment for various illnesses may last much longer.

Commercial sex workers and bar girls are virtually absent from Likoma (and neighboring Chizumulu), and inhabitants from the island primarily engage in high-risk sexual behaviors with such partners during trips to the cities of the mainland of Malawi (e.g., Nkhotakota, Nkhata Bay, Mzuzu) or Mozambique (e.g., Mtengula, Lichinga). Opportunities for sexual relationships with outsiders, however, may still arise on the island because Likoma regularly hosts a small number of visitors from the mainland of Malawi and Mozambique. For instance, small groups of inhabitants of Mozambique frequently travel to Likoma to sell firewood or grass (used as the main material for roofing on the island). Indeed, there are very few trees on Likoma, and residents of the island are completely dependent on such imports for their daily cooking needs for example. Some Mozambican traders thus spend a few days per week in Likoma, camping on the beach at the trading center. From Malawi, a small number of civil servants and NGO workers visit Likoma on a regular basis, usually residing at one of the island's guest houses at the trading centre. Various relief organizations, microfinance loan groups or NGO promoting health education have ongoing projects in Likoma. Similarly, every month a group of 20 Malawian soldiers from the mainland comes to stay at the military camp located on the southern side of Likoma (facing Mozambique). Because soldiers are generally young, well-paid by local standards and are bound to return to the mainland after their stint on Likoma, they represent very attractive partners for many local young adults. As a result, it is not uncommon to see a few girls from the island at the camp on any given night. During our time in Likoma, many local residents were quick to blame these soldiers for a large number of unwanted pregnancies and infections with STIs.

While Likoma stands out in comparison to the rest of Malawi by its secluded location and relatively high density of settlement, the island also shares many socioeconomic and demographic characteristics with the mainland population in northern Malawi. For instance, because of widespread education, a large percentage of adolescents are still enrolled in school late into their teenage years, delaying the transition to marriage. The most educated ones tend to migrate to cities in the rest of Malawi or abroad. In both the island and northern Malawi, marriage is patrilocal, i.e., women move into their husband's families after union. Polygamy is quite prevalent, but so is divorce. Usually marriages take place between people living in neighboring villages, and Helleringer and Kohler (2005) have shown the implications of such marriage patterns for social interactions and AIDS-related attitudes. Comparisons with the rural population surveyed by the Malawi Diffusion and Ideational Change Project (MDICP; see Watkins et al. 2003) confirm these overall similarities in socioeconomic context (Reniers 2003).

4 Practical challenges and sources of bias in sociocentric network studies

The collection of large-scale data on sexual networks is a challenging undertaking facing abundant practical obstacles (Doherty et al. 2005; Morris 2004). These practical difficulties are a major reason for the general scarcity of detailed sociocentric information on sexual networks, and may limit the researcher's ability to ascertain the networks transmitting HIV and other STIs.

4.1 Potential sources of incomplete-network bias

Doherty et al. (2005) identifies three difficulties, discussed in more detail below, that may lead to *incomplete-network bias*: "the incomplete ascertainment of sociometric networks is inevitable in both clinical and research settings, because (1) people may be reluctant to name all sex partners [...]; (2) they may be unable or unwilling to provide adequate contact information for locating partners; or (3) partners may be locatable but difficult to reach" (Doherty et al. 2005).

The first aspect, the misreporting of sexual partners and sexual relationships, is pervasive in all inquiries of sexual behaviors (Cleland et al. 2004), including large-scale individual centered surveys such as the DHS. This problem may also be exacerbated by the fact that misreporting of partners may differ by the type of relationships and characteristics of the respondent. The second and third aspects (insufficient information for partner tracing and failures to locate nominated parters), on the other hand, are specific to sociocentric studies of sexual networks. In developed countries, data on sexual networks are mainly collected within healthcare settings during contact tracing interviews of STI cases (e.g., Ghani et al. 1996; Klovdahl et al. 1994; Wylie et al. 2005). Such data have provided seminal insight on the role of sexual network structures in shaping disease diffusion among high-risk groups. In SSA, however, health districts generally have neither the human nor financial resources to undertake the tasks necessary for contact tracing and constructing sexual network datasets, especially in contexts where 10–20% of a population is infected with HIV (or HSV-2, Syphilis). In particular, contact tracing procedures involve collecting and managing extensive identifying information on partners of cases such as names, address, phone number, various socio-demographic characteristics etc. Even in resource-rich contexts, the collection and management of such information can be very cumbersome, and it is often the case that the information provided by respondents during contact tracing interviews is not accurate, or is not detailed enough to eventually find the nominated partner (Potterat et al. 2004). As a result, a large proportion of contacts are never traced during sociocentric studies (Ghani et al. 1996; Koumans et al. 2001; Potterat et al. 1999) and the descriptions of the networks produced during similar inquiries are somewhat partial and subject to incomplete network biases. These difficulties of contact tracing appear compounded even further in SSA, because individuals are not easily identified. For example, often there are no street names, no house number, no phone number where someone can be reached. It is also common for someone to use different names or nicknames under various circumstances or to change names after important events of the life cycle (e.g., sexual initiation among certain ethnic groups), making identification and contact tracing problematic (e.g., Potterat et al. 1999).

The problems of locating nominated parters are also akin to a *boundary specification problem* (Laumann et al. 1983). In the study of diffusion processes, it is not clear where to draw the line between members and non-members of a population. Influential individuals may well reside outside of a specific area or may not belong to a group defined by a certain criteria. In a re-analysis of the classic study of the adoption of the drug Tetracycline among a network of medical practitioners in Illinois for example (Burt 1987; Coleman et al. 1966), Van Den Bulte and Lilien (2005) shows that marketing agents were the most influential proponents of the drug. However, the network data collected by Coleman and others (Coleman et al. 1966), did not include such actors within its boundaries. In the case of HIV spread, groups of "outsiders" such as truck drivers, people from town, etc., have been identified as playing a disproportionate role within sexual networks transmitting HIV in rural areas of SSA (Caldwell et al. 1989; Hudson 1996). Such groups represent epidemiologically important *bridge populations* (Caldwell et al. 1989; Lagarde et al. 2003; Lurie et al. 2003) who continually re-introduce HIV within villages of rural areas. Because they reside outside of local communities or are highly mobile however, they are difficult to reach and may be systematically left out of complete network studies. It is thus often believed that network studies of sexual mixing in SSA would miss the key players in the network and give a distorted picture of the structural factors favoring/inhibiting HIV spread on the continent.

4.2 Likoma as an "epidemiological laboratory"

In light of these challenges, some network studies have attempted to limit the time spent on contact tracing by implementing various selection schemes to enroll only a subset of the contacts named during tracing interviews. Various sampling schemes based on snowball sampling and the statistical theory of random walks, for example, have been suggested (e.g., Klovdahl 1989). The properties of network estimates derived from such data are however not clear. More recently, a new approach has been developed in which researchers ask study participants to enroll their (sexual) contacts themselves. This sampling technique is referred to as "respondent-driven sampling", and its statistical foundations have been investigated in detail (Salganik and Heckathorn 2004). However, such a method is highly vulnerable to respondents' malfeasance, especially if (financial) incentives are associated with participation in the study. The LNS we describe here does not follow this route derived from snowball sampling procedures, but instead aims at reconstructing the complete sexual network of a well-bounded population.

Other studies of complete networks have addressed the above challenges by carefully selecting study populations that have well-defined boundaries as well as a limited size that allows efficient identification of network members. For example, the classic dataset of a complete network describes friendship connections between monks of a monastery (Sampson 1969), and most of the other datasets available focus on small groups within organizations (Krackhardt 1987), small groups of families (Padgett and Ansell 1993), or other well-defined populations. Studies of sexual networks have been conducted among small numbers of members of high-risk groups, and the only available sociocentric study of sexual networks conducted among the general population similarly focused on a small, well-bounded population. Bearman et al. (2004) set their study in a high school of a rural and isolated community of the US, where the selected high school was the key *focal context* (Feld 1981) of adolescents' social life. As a result, the majority of romantic relationships identified by the study took place between members of the sampled population (i.e., between student of the selected high school). For comparison, the National Longitudinal Study of Adolescent Health (AddHealth) attempted a study similar to that of Bearman et al. (2004) but in a highly diverse urban high school. In this school, only 11% of sexual partnerships of students were with fellow students: the majority of members in the social network thus did not meet the inclusion criteria for the study, and an accurate reconstruction of the sexual network of students in this school was not possible (Bearman et al. 2004).

For the LNS, we chose an island setting—Likoma on Lake Malawi—to minimize the relevance of the boundary problem and reduce the importance of incomplete network biases. Indeed, Likoma Island (a small island of Lake Malawi) extends over only 18 square kilometers, has limited transportation to the mainland, and its population is small with just over 7,000 persons living in a dozen villages (Section 3). As a result, a limited set of identifying information allows tracing contacts nominated during this sociocentric network study. In addition, we used audio-computer assisted self-interviewing (ACASI) to minimize misreporting of sexual behaviors (Bearman et al. 2004; Bloom 1998; Mensch et al. 2003). The combination of the island setting of this study with an ACASI network survey thus enables the collection of high-quality quasi-complete sexual network data, despite the considerable challenges facing such data collection. Selected aspects of the data quality of the Likoma Network Study will be further analyzed in Section 6 below.

The choice of an island as research site also continues a long tradition of island studies in epidemiology and biology (for reviews, see Cliff et al. 2000; Whittaker 1999). Understanding of the spread of airborne diseases for example has been significantly enhanced by analyses of flu or measles epidemics in Iceland and pacific islands. One of the main advantages of an island setting for study of epidemics is that it allows identifying the mechanisms through which an infection diffuses locally through a population: whereas in mainland settings, it is often possible that a disease is constantly reintroduced within a locality through travel, on an island this occurs much less often. This aspect enables identifying internal pathways of spread that would be more difficult to observe in other settings. This is also the case in the study of HIV. Many epidemiological studies have emphasized the role of migration and mobility in diffusing HIV throughout the continent (e.g., Glynn et al. 2001), but very few have tried to assess whether there exist mechanisms that would enable the local transmission of HIV (exceptions include Coffee et al. 2005; Lurie et al. 2003). Furthermore, while movement is crucial for initial epidemic growth, it may be less relevant for incidence of HIV at latter stages of the epidemic, and especially may not be a key determinant of the generalization of infection to large proportions of the population. Proximate relationships between members of a general population are possibly much more relevant at this stage (Blanchard 2002). The island setting of our study allows observing aspects of sexual mixing that contribute significantly to disease spread, but may be confounded by movement and mobility in other settings.

5 Data Collection Procedures in the Likoma Network Study

In this section we describe the data collection procedures that were implemented as part of the LNS to reconstruct the sexual networks connecting the inhabitants of the study villages chosen for this project (Figure 2). The protocol for this study was approved by institutional review boards at the Malawi College of Medicine and the University of Pennsylvania. Community approval was obtained during meetings with local representatives (traditional chiefs, district representatives), and informed consent from the study participants was obtained prior to interviews and HIV testing.

The data collection of the LNS took place between October 2005 and March 2006 and involved three different phases. The first phase consisted of a census of all households residing in Likoma district (i.e., both Likoma and Chizumulu islands), whose main aim was to establish rosters of potential network members. The second phase consisted of an in-depth sexual network and health survey that was conducted only in a subset of the villages of Likoma island, *among all adults aged 18–35*. Finally, the third phase involved the collection of biomarkers of HIV infection among respondents of the network survey. Information on the socioeconomic status and demographic characteristics of respondents comes from both the household listing and from a short face-to-face interview conducted with all survey respondents prior to the ACASI network interviews.

5.1 Rosters of potential network partners

a) Household rosters: During the first phase of the study, we gathered extensive information about the socioeconomic characteristics of each household (e.g., housing type), as well as about the names, maiden names (for married women), potential nicknames, ages, and marital histories of all residents of a household. Ten interviewers enumerated more than 1,300 households in Likoma and 500 in Chizumulu in roughly 3 weeks. Each house of the islands was visited, and available informants were asked to answer questions about their household and its members. Eligibility criteria for informants included: being older than 18 years old and being a regular member of the household. Only 4.2% of household informants were less than 18 years old, and 85% of them were either the household head or his/her spouse. At this stage, interviewers broadly described the study goals to informants prior to gathering list of names for enrollment. Despite general enthusiasm for the study, subsequent visits to villages during the sexual network survey indicate that some of the most remote locations in Likoma (e.g., hills, village at the northern tip of the island) may have been underenumerated.

The household census also included vacated dwellings: neighbors of empty houses were asked to answer a one-page questionnaire about former residents of the house. This questionnaire included questions about family name, former head of the household, time since the house had been vacated, and reason for departure. 150 vacated dwellings were enumerated in Likoma, and roughly 30 in Chizumulu. Such information helps get better estimates of migration out of the island.

b) Migration and mortality: Because migrants and recently deceased individuals have been identified as crucial for disease spread (e.g., Coffee et al. 2005, 2007; Lurie et al. 2003; Wawer et al. 2005), household informants were asked about (*i*) temporary migrants from their households, (*ii*) household or extended family members who moved permanently out of Likoma over the last 5 years, and (*iii*) household or extended family members who died over the last 5 years. For each of these, the informant was asked to provide names, potential nicknames and sociodemographic characteristics (age, sex, education, marital status, etc.). For migrants, date of departure, reason for departure as well as destination and for deceased, date of death as well as a few questions relating to the cause of death and probable final illness were collected. The lists gathered through the household census and migration/mortality modules constitute rosters of potential social and sexual network partners that we use to link records of relationships (see below).

c) GPS data: During the household listing, we also collected the GPS coordinates of all the dwelling units we visited. We tracked the main roads and pathways people use to travel around the islands. Finally we referenced the main landmarks and public places in each village: for example, each school, church, borehole, or village center were located. This information is extensively used to identify sexual partnerships (see below).

5.2 ACASI network survey

d) Study populations: After the initial census of the island provided a list of potential network partners, we conducted a more extensive survey of sexual partnerships in some villages of the island. Seven adjacent villages were sampled for the sexual network survey (Figure 2). Sampling was purposive: we initially selected two "seed" villages in which the proportion of births out-of-wedlock reported during the household listing was significantly higher than in the other villages. We interpreted this difference as indicating either a higher prevalence of extra-marital relationships in these villages or a higher propensity to report such relations during a survey. These two villages were located on opposite sides of the island, and showed largely different village contexts: in one village, fishing is the quasi-exclusive source of income, and as a result most males spend their days (and often nights) on the lake. Women, on the other hand, often travel to the local trading center (or to the mainland) to sell their husbands' and siblings' catches. There is very little "village life" as no groceries nor bottle stores are located within the village. On the other hand, families in the other seed village have stronger ties to the mainland of Malawi, and remittances represent a significant source of income for many households. As a result, fishing is not the focus of social and economic life, and on most days large groups of people can be found gathering close to the groceries or at the village center, playing games of bawo (a local board game), chatting and, for men, drinking beer. The villages immediately bordering these seeds villages were chosen to complete the sample, so that we could capture a larger proportion of the sexual partnerships reported during the survey. Geographically, the sample thus stretches from the southern tip of the island to villages of the northern side of Likoma.

Sampled villages represent more than half of the island population. In these villages, we interviewed *all inhabitants aged 18–35 and their spouses*. We limited our sample to this age group, because it represents the age range during which most HIV infections occur (Heuveline 2003) and thus is most significant for epidemiological analyzes of the diffusion of the virus. This limitation of the sample to the younger age groups generated some confusion among the population, especially because we explained the purpose of our study as trying to identify the factors affecting the spread of HIV. Older respondents argued (rightly so) that they were also at risk of acquiring HIV and as such should have been interviewed by the survey team (and especially tested during the collection of biomarker data).

Prior to the start of the study, we conducted a pilot of the interviewing software in two separate non-sample villages of Likoma. We chose these two villages for their convenience as they were immediately bordering the trading center where the research team was staying. In the first village, we interviewed 20 respondents using a first version of the sexual network module. After getting feedback from both interviewers and respondents, we significantly revised our interviewing strategy (see below) and conducted a second, more extensive pilot with roughly 80 respondents. This second pilot proved largely successful and initial releases of the network data included relationships identified from pilot interviews (Helleringer and Kohler 2006). However, because interviews were conducted with only 60% of village inhabitants (a response rate significantly lower than in the other sample villages and one that does not allow drawing a quasi-complete picture of the village network), these data are discarded from the final analyzes. In addition, the final version of the software used during fieldwork is slightly different from the version we used during this pilot.

e) Length of the recall period: During surveys of sensitive practices (e.g., injecting drug use, commercial sex...), respondents are usually asked to recall their behaviors over short periods of time. Bell et al. (2000) for example uses a recall period of 30 days. In the case of HIV transmission however, because the period of infectivity can last for years, asking questions about such short periods is likely to omit most of the behaviors and partnerships that may have lead to infection. Instead, we chose a recall period of 3 years that likely encompasses a significant proportion of the partnerships during which HIV was transmitted in a population aged 18–35. Such a long recall period, however, is prone to forgetting of partners (Brewer and Webster 1999). To reduce this effect of partial recall in the construction of the population-level network, a sexual relationship was assumed to exist if it was reported by *at least one partner*.

f) Fixed choice design: As we argued above (Section 2), the variance of the degree distribution is an important parameter in mathematical models of STI spread. Unfortunately, the behaviors of high-degree network members is often difficult to measure in empirical studies. For example, highly sexually active survey respondents may grossly mis-estimate the number of their partnerships (e.g., Handcock and Jones 2004). Several studies have also shown that respondent's fatigue builds up quickly in network surveys (e.g., White and Watkins 2000), and respondents quickly become bored with answering the same set of questions about a (possibly large) number of different partners. As a result (and also because of software and programming constraints), we followed the common practice in network research and we imposed a cutoff on the number of partnerships to be reported. This value was set at five. Such a research design may lead to bias in estimates of network properties (Costenbader and Valente 2003; Kossinets 2006) if it is close or even below the mean of the underlying degree distribution. However, if the mean of the underlying degree distribution is well below the cutoff used during a network survey, then various structural properties (e.g., dyad or triad census) of the underlying networks are well estimated in such designs (see chap. 13, Wasserman and Faust 1994). We estimate in 6.2.1 what proportion of respondents might have reported more sexual partnerships if such a fixed choice design had not been used, and we evaluate the potential biases arising from the limitation of outdegrees at five.

g) Strategy for identifying sexual partners: Establishing complete network data by linking reported sexual partners generally involves looking up names generated during a survey into pre-existing rosters of potential network partners. While previous studies of sexual networks having used ACASI technology (e.g., Bearman et al. 2004) put this burden on respondents and asked them to directly browse through rosters of potential partners, this was found to be highly impractical in Likoma, where computer literacy is minimal. Indeed, during pilots and pre-tests, we experimented by inputting the rosters created from the household census in audio files and incorporating them in our interviewing software, so that respondents themselves could establish links. But pilot respondents were surprised and angered that a machine could know the names of actual people (raising all sorts of witchcraft accusations). Furthermore, the thousands of audio-files (*.wav files) required to enable this interviewing strategy were significantly slowing down the ACASI software.

Thus we developed an alternative linking strategy. Respondents were asked to mention the name of each their partner through recording headsets. For the purpose of identifying sexual partners in the rosters of potential partners, they were asked where the partner they mentioned was currently residing and where he/she was residing at the time of the relationship (if the relationship was over). If the partner was currently residing in Likoma, they were asked to provide additional details about his/her residence, i.e., in which village and where specifically in this village this person was staying. For example, from such information we know that a respondent has been involved in a relationship with "John Banda" who lives in Ulisa village close to the groceries, etc. A few questions on the socioeconomic characteristics (e.g., daily activity, age, etc.) of the partners were asked to help narrow down the list of potential matches. If the partner had never resided on Likoma, respondents were only asked to mention the first name or the initials of their partners, and the audio-files were subsequently discarded by the ACASI software. This name-generating process was repeated by the software for up to five partners. Stored audio files (i.e., including full names of partners having ever resided on Likoma) were downloaded daily by the data management staff and linkages with the village/migration rosters of potential network partners were generally conducted and checked within two days of the interview. These checks were initially conducted using phonetic name-matching routines (Blasnick 2001), and finalized through manual inspection of the village and migration rosters by the investigator present in Likoma during fieldwork. Through this process, we are able to reconstruct the network of sexual relations within which inhabitants of these villages are embedded.

It is important to note that this linkage strategy involves absolutely no "active" contact tracing, during which the researcher (or public health specialist) approaches the partner(s) of initial cases using the information gathered during interviews. Contacting a nominated parter as part of the survey would signal to other community members (including spouse and family members) that a person belongs to an extended network of (possibly extra-marital) sexual relations in a tightly knit rural/island setting. Our approach on the other hand is inclusive as *respondents are only approached by the research team on the basis of their residence and their age*. By including all inhabitants of neighboring villages, and taking advantage of the island setting, we are able to reconstruct the sexual networks without going through active and explicit contact tracing.

h) Definition and translation of "sexual partner": The goal of the sexual network survey was to elicit all sexual relationships of the respondents during the three years prior to the survey, including regular and stable relationships (e.g., with the spouse or a regular extra-marital partner) as well as relationships that were short-term and/or infrequent (e.g., a one-off relationship or a sexual encounter with a visitor), or relationships that occurred during the marriage process as part of dating and partner search. Relationships occurring on the island, but also sexual relationships occurring off the island during a respondent's travel were supposed to be reported.

To ensure the most accurate reporting of relationships, considerable effort was devoted to developing the sexual network module, and specifically the question prompting respondents to report their sexual relationships. The local language, *Chichewa*, has a specific term to designate sexual partners: "chibwenzi". This expression translates loosely as "someone who provides for one's sexual needs", and this is the term that was used throughout the survey to signify a sexual partner. Because within marriage, spouses are supposed to provide for each other far beyond sexual needs, this term does not encompass marital relationships and we collected information on behaviors within marital relations during face-to-face interviews and the household listing. There appears to be a hierarchy among "chibwenzi" relationships as some are rather infrequent and occasional, including possible one-off sexual encounters, whereas others become more established and may represent a step towards marriage: at this latter stage, they are sometimes referred to as "chitomelo" relationships, in which the man has promised to marry his partner in the near future. ACASI interviews thus asked respondents to classify their relationships according to their degree of involvement: one-night stand/infrequent partner/stable partner.

In some instances, especially among younger adolescents, the term "chibwenzi" may however not be so closely linked to sexual relations, but instead refers to "someone from the opposite sex one spends time with". Indeed, because in Likoma (and in Malawi in general) sociability is highly gendered (men spend most of their time outside of their household with other men, women with other women), spending extensive amounts of time with someone from the opposite is a special type of friendship that may eventually lead to sexual relations. But as many younger (female) respondents emphasized, this is not a necessary trajectory. Because we are interested in the *transmission* of HIV in networks, this polysemy of "chibwenzi" in Chichewa is potentially problematic: it may yield to inclusion of non-sexual relations in our analyzes of HIV diffusion, and for example lead to biases in analyzes of the relationship between sexual behavior and infection, or comparative analyzes of modes of HIV transmission. To avoid such confusion, we included a specific question about the presence of vaginal intercourse within the relationship. This question was perceived as largely redundant by many respondents, but allowed excluding a series of relationships between younger respondents that likely did not involve sexual activity prior to the survey (see below).

Finally, the term “chibwenzi” does not subsume relationships with prostitutes, as well as other forms of “traditional” relationships. Whereas such relations have garnered much attention, it is however not clear how much they contribute to the local transmission of HIV since we could not identify bar girls or commercial sex workers on the island during our fieldwork. Informal prostitution may nevertheless occur on the island, and it is said that it is mostly limited to the female owners of local breweries, who may engage in relationships with customers.

Likoma is also said to be the home of a ritual or traditional sexual practice the locals called “kutondola”. This practice seems to have its origins in the largely imbalanced sex ratios of the island: because the number of women exceeds that of men, there are spinsters who can not secure a spouse. Thus, some of these older unmarried women, after a certain age, are said to “rent” the services of a fellow villager (generally a neighbor or a relative, possibly married) to be able to bear children. Such relationships are traditionally arranged by an aunt of the woman, take place at night, and are kept in high secret, as even the woman is not supposed to know the identity of the man who visits her. The context and taboos surrounding these relations thus makes them difficult to identify during a sexual network survey. However, during our time on the island, inhabitants of Likoma indicated that “kutondola” was fading, and was mostly associated with women of older generations. In the event it still takes place, the secret surrounding the identity of the visiting male is not well kept. In fact, during the face-to-face survey we collected maternity histories, and among women having had a birth during the three years prior to the survey, only a very small proportion of them reported not knowing the identity of the father of one of their children (around 1%). Nowadays, it seems the meaning of “kutondola” has even extended to signify relationships during which a woman is abandoned by her partner after he has made her pregnant, or relations with relatives/family members (e.g., in-laws). Such relationships are adequately captured by our survey instrument.

Using the Chichewa term “chibwenzi” during the sexual network survey, therefore, gives us considerable confidence that it adequately captures all types of the sexual relationships between members of the general population. It may nevertheless leave out some relationships between members of these populations and some members of core groups (e.g., sex workers).

i) Reactions to network survey: During the study and consistent with reports from Mensch et al. (2003) for Kenya, we found that audio-CASI technology performed well with few hardware or software malfunctions, and respondents were able to complete the survey with limited training. In particular, use of the recording headsets was not problematic and audio files of partners’ names were usually of good sound quality.

Very few respondents refused to complete the computer-assisted section of the survey after having completed the face-to-face interview. Similarly, even though respondents had the possibility to skip/refuse to answer every question of the audio-survey by pressing a simple touch on the computer’s touchpad, few refused to name partners or to answer questions about partners’ residence. Such refusals usually came after 2 or 3 partners had already been named, and might indicate an inadvertent error from the respondent or fatigue. Comparatively, refusal rates for single items were significantly higher for questions about occasions of initial meeting (10% missing data), or questions relating to condom use. We suspect, however, that when a respondent wanted to keep a relationship secret, he/she used names such as "Andreya Banda" or "Esther Phiri" for their partners, which are the equivalent of "John Smith" or "Mary Jones" in the U.S. We generally had difficulties linking such common names to our village/migration rosters.

5.3 Relationship and Health Data

In addition to questions allowing the identification of sexual partners, respondents were asked a short series of relationship-specific questions during the ACASI interview. These included questions related to the initial meeting and conditions that surrounded the initiation of the relationship such as how did the two partners know each other before the start of the relationship, where did the first meeting occur, when did the first meeting occur, when did the relationship end and what were the reasons for the relationship ending; questions related to sexual activity within the relationships, including whether or not the relationship involved sexual intercourse, the frequency of sexual intercourse within the relationship, as well as use condoms during sexual intercourse; and finally a short series of questions about perceptions of HIV/STD risk at the time of the relationship.

After completion of the sexual networks part of the ACASI interview, respondents were asked to answer questions regarding their own health, including: a self-reported assessment of general health; the presence and frequency of specific symptoms such as headache, stomach ache, general weakness, joint aches, painful urination or discharge during urination, ulcers in the genital area; the use of healthcare when these symptoms occur; previous use of HIV testing services, and reasons for not being tested (e.g., distance, cost, stigma associated with testing centers); and a history of injections during healthcare, and the date, reason and location of the last injection received by the respondent. This data on injections is useful because it allows estimating the hazard rates for incidence of injections within a population (Allison 1985), and this parameter can be used in simulations of the spread of HIV through different modes of transmission.

5.4 HIV testing

A month on average after the completion of survey fieldwork, each respondent was re-visited by a member of the research team and was offered a free HIV test. The research team for this phase of data collection was composed of one nurse in charge of the overall supervision of biomarker collection, and 10 health counselors trained by the Malawian ministry of Health. This team of health counselors visited all respondents in six of the seven survey villages (see also Figure 2). The 7^{th} village—village 14 in Figure 2—could not be included due to funding and timing constraints limiting the scope and duration of the fieldwork. When approached by the health counselors, respondents were offered the opportunity to received counseling and testing in their homes using rapid HIV tests. Because individuals may be concerned about the privacy of in-home HIV tests, respondents were also given the option to be tested at another location (i.e., the team's hotel). The rapid HIV tests were conducted using a parallel testing algorithm approved by WHO (World Health Organization 2002) and the Malawian Ministry of Health. Two rapid HIV tests, Unigold (Trinity Biotech, Ireland) and Determine (Abbott, Japan) were run simultaneously at the respondent's home. Blood samples that were concordantly negative or positive were considered to be a true result. Four discordant results were obtained and were referred to local testing centers for confirmatory testing after a few weeks. Results were available to the respondents after 20 minutes, but respondents were also given the option to retrieve their test results at a latter date if they desired so.

6 Results

6.1 HIV prevalence

Overall response rate for the biomarker data collection was 74% for women and 65% for men. 21.7% of men refused to be either counseled or tested vs. only 15.4% of women. 13.5% of men and 11% of women could not be found at the time of the VCT. Overall participation rates varied a lot between villages, ranging from 54% to 82% (see Figure 2). In total, 597 respondents were tested for HIV, and among those tested, the data show an overall HIV prevalence rate of 10.6% (95% CI 7.2%-13.9%) for females and 4.7% (95% CI 2.1%-7.3%) for males.

6.2 Sexual partnerships: descriptive statistics

A total of 923 inhabitants of these villages were interviewed during the sexual network survey (421 Males and 502 Females), and the participation rate was 88% (923 participants out of 1,052 eligible respondents). The main reason for non-participation in the survey was temporary migration to the mainland of Malawi or Mozambique. Very few potential respondents refused to participate in the survey ($N = 21$, 2.5%).

6.2.1 Outdegree distributions

The outdegree of a respondent is the number of reports of partnerships made during the sexual network survey (i.e., how many partners were nominated). Males reported having been involved in an average of 2.41 relationships (Figure 3a) during the 3 years prior to the survey, whereas women reported 1.82 relationships over the same time span ($p < 0.01$). Only 2% of women and 8.4% of men reported 5 partnerships or more. The variance of the outdegree distribution was significantly lower than its mean (0.92 for women, 1.3 for men). A little over 5% of respondents reported no sexual partnerships during this survey, and among these the proportion of males was slightly higher (Figure 3a). When the period of observation is restricted to the year prior to the survey (Figure 3b), more respondents report not having been involved in a relationship (10%) and the majority of respondents report only one partnership. 28% of women and 43% of men nevertheless report more than one partnership ($p < 0.01$).

6.2.2 Characteristics of relationships: tracing rates

A total of 2,040 reports of relationships were collected from the 923 respondents during ACASI interviews, among which 1,858 were said to have involved sex (or 91%). Table 1 describes characteristics of the reports of sexual partnerships made during the network survey, and Figure 4 summarizes the steps of the linking process and the terminology we use in our analyzes.

Table 1 shows that 30% of the relationship reports by women were about marital relationships, whereas marital relationships constituted only 20% of men's reports. The majority of relationship reports collected during this survey were thus described by respondents as non-marital relations.

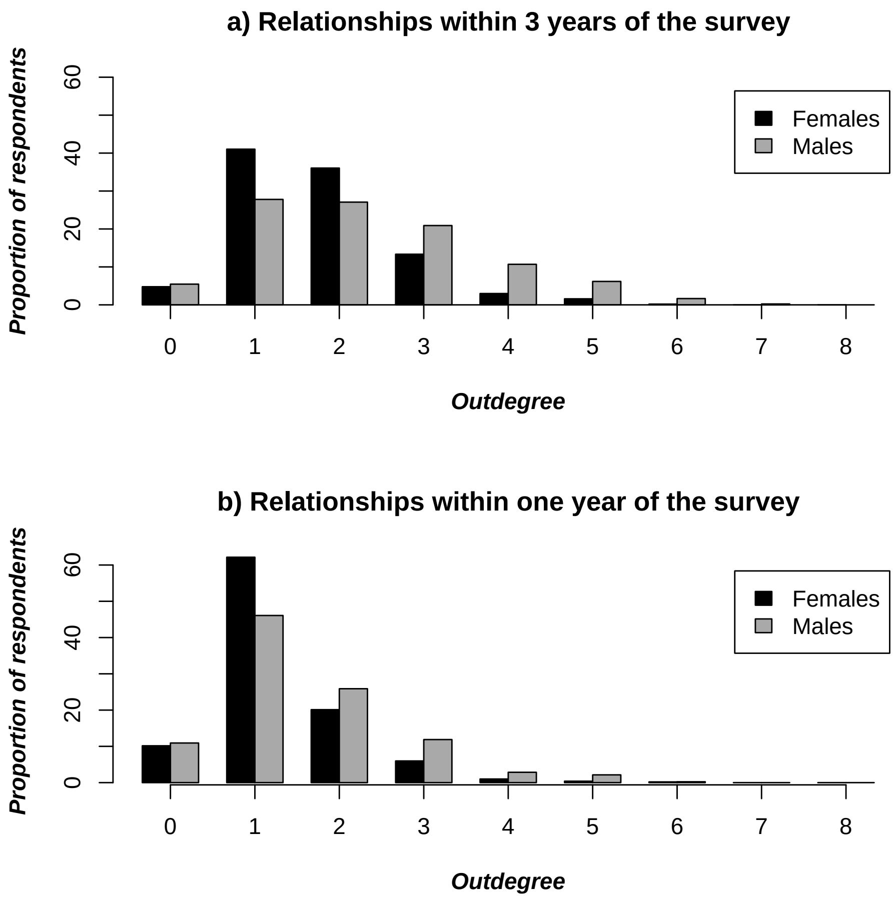

Figure 3: Outdegree distributions among survey respondents, by gender; *Source:* Likoma Network Study)

Table 1: Characteristics of relationships reported during the sexual network survey

	Relations reported by women				**Relations reported by men**			
	Marriage $N = 272$	Steady partner $N = 344$	Infrequent partner $N = 200$	One-night stand $N = 51$	Marriage $N = 200$	Steady partner $N = 401$	Infrequent partner $N = 247$	One-night stand $N = 88$
Proportion of partners who were currently residing								
in Likoma	83.5	61.3	64.6	52.1	92.5	66.9	64.2	50.6
in Chizumulu	1.1	2.7	6.2	4.2	0.0	5.0	8.2	6.0
in Mozambique	1.8	4.0	2.1	2.1	2.5	4.1	4.7	9.6
in Malawi	11.3	26.1	21.4	22.9	3.5	19.4	15.9	19.3
dead	4.1	4.0	5.2	14.6	1.5	1.5	3.0	13.2
Proportion of traced partnerships among partners who								
resided in Likoma	95.5	76.3	79.0	68.0	95.1	81.1	71.8	78.6
resided outside†	75.0	57.9	60.0	51.4	60.0	62.0	45.4	45.0
Proportion of in-sample partners among								
partners in Likoma	71.9	50.9	52.4	52.0	85.9	58.0	50.3	52.4
jointly reported ††	94.0	52.4	27.5	31.2	96.9	35.0	20.0	10.4

Notes: Reported numbers in the table are percentages (except for p-values). p-values are based on chi-square tests of associations. †Among partners who have ever resided in Likoma. ††among in-sample partnerships

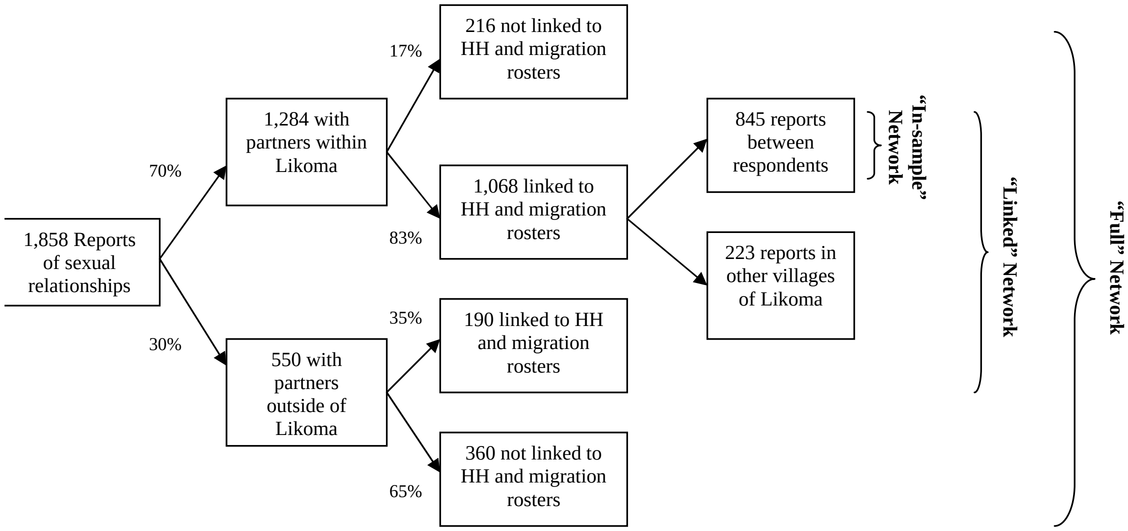

Figure 4: Flow Chart of the linking process and terminology used

One-off-encounters ("one-night-stands") represent only 7% of all reports, and infrequent partnerships account for slightly over 20% of all nominations. As a result, relationships included in the network are relatively stable. Nominations to partners residing in Likoma at the time of the survey accounted for 70% of all reports, a relatively high percentage compared to Bearman et al. (2004)'s study of a secluded US high school, where only half of the nominations were to fellow students. There were significant differences in the residence of sexual partners by type of relationship: most marital partners co-reside in Likoma, but a large proportion of non-marital sexual networking takes place *outside of the island*. Only two thirds of steady and infrequent extra-marital partnerships took place between current residents of the island, and this proportion decreases even further in the case of one-night stands (50%, see Table 1). There were also significant gender differences in patterns of geographical mixing: most non-marital partnerships of women not taking place within Likoma took place with partners residing in the mainland of Malawi. Men, on the other hand, engaged in non-marital partnerships with partners from more diverse contexts (Chizumulu or Mozambique). This pattern of sexual mixing probably reflects daily patterns of mobility. In addition, 11.3% of women but only 3.5% of men reported a marital partner outside of Likoma. Spousal separation was often due to divorce (the relations in Table 1 include some relationships that were over at the time of the survey), but may also occur because of migration of one of the spouse (generally the man). Finally, women were more likely to report having been involved during the three years prior to the survey in a partnership with someone who had since died. The proportion of deceased partners was significantly higher in short and unstable partnerships.

Of the sexual relationships involving two partners currently residing on Likoma ($N = 1,284$), we were able to trace both partners within our lists in 84.9% of the cases (80.5% of extra-marital relationships and 94% of marriages). Tracing rates of partners residing within the island did not differ systematically by gender, but less stable relationships were significantly less likely to be linked to a record of the village rosters ($p < 0.01$). Differential success in tracing partners might thus slightly bias the network we draw towards more stable/legitimate relationships. These tracing rates however compare favorably with those observed during studies of high-risk groups (such as drug users or patients of STI clinics) conducted in the US (e.g., Koumans et al. 2001; Potterat et al. 1999).

When the nominated partner is not currently residing on the island or has died (but was residing on the island at the time of the relationship, $N = 303$), on the other hand, we are able to trace him/her within our rosters of migrants from the island in 62.5% of the cases. This lower tracing rate may suggest that (*i*) migrations or deaths were underreported during the migration/mortality module, (*ii*) the timing of migration and/or death may have been misreported during the migration/mortality module.

6.2.3 Characteristics of relationships: Jointly reported relationships

While other studies often cannot evaluate the accuracy of the reporting, the network design pursued in the LNS provides several opportunities to assess the quality of the data on sexual behaviors. Because sexual relationships represent a *dyad* within a network, accurate reporting of sexual behaviors would imply that each relationship is reported by both members of the dyad. In reality, however, reports of sexual relationships are often discordant: they are reported by one, but not the other partner. For instance, the 845 relationships where both partners were respondents in the pilot study (in-sample, Figure 4), with accurate reporting, should be reported by both partners. This was the case for 57.7% of all partnership reports (Table 1). Specifically, close to 95% of marriages were jointly reported by both spouses, and 36% of reports of extra-marital relations were concordantly reported by both partners. Table 1 also shows that the proportion of reciprocated reports generally increases with the strength of a relationship: reports of stable non-marital partnerships are generally more reliable than reports of one-night-stands. The concordance of reporting increases for ongoing relationships, as 54.2% of ongoing non-marital relationships are jointly by both partners whereas this is the case in only 25% of non-marital partnerships that have finished more than a year prior to the survey. The proportion of concordantly reported non-marital relationships appears lower than proportions of concordant reports in other studies of high-risk behaviors (e.g., Adams and Moody 2006; Bell et al. 2000). However, this lower concordance rate is likely to be due to (*a*) the long recall period for sexual relationships in the Likoma study (up to 3 years prior to the survey as compared to 30 days in Bell et al. 2000), (*b*) the limit of at most five network partners may have resulted in truncation of some reports, and (*c*) reporting biases that may differ among groups and types of relationships.

6.2.4 Mixing patterns and risk factors

Table 2 provides further descriptions of the context and quality of non-marital relationships. Several characteristics of relationships differed significantly between gender and across relation type. Whereas most relationships (almost 70%) were initiated within the island, men were much more likely to engage in one-off encounters outside of Likoma. Almost half the one-night stands reported by men thus took place either in Mozambique, in Chizumulu, in Mainland Malawi or possibly elsewhere (e.g., Tanzania, South Africa). Because most relationships take place within Likoma, the large majority of partnerships are with someone the respondent was acquainted to

Table 2: Characteristics of relationships reported during the sexual network survey

	Relations reported by women				Relations reported by men			
	Steady partner	Infrequent partner	One-night stand	*p*-value	Steady partner	Infrequent partner	One-night stand	*p*-value
Context of initial meeting								
Place of first meeting								
In Likoma	75.8	77.0	62.5	0.18	74.5	74.0	51.8	< 0.01
Occasion of first meeting				0.26				< 0.01
In school	43.4	38.1	42.2	–	37.0	31.4	32.9	–
Mganda dances	23.5	31.2	22.2	–	28.8	29.5	20.2	–
During business trip	8.8	7.7	15.6	–	8.1	10.0	20.2	–
On the steamer	6.2	7.7	8.9	–	9.0	15.4	13.9	–
At a religious meeting	13.7	13.8	5.1	–	11.4	10.9	12.6	–
Type of relation				0.01				< 0.01
Did not know each other	21.3	23.4	14.6	–	19.4	25.9	37.8	–
Relatives	5.8	6.2	23.2	–	7.2	9.5	12.2	–
Friends	36.8	30.7	31.2	–	34.6	26.7	14.6	–
Acquaintances	33.7	38.5	29.1	–	36.2	36.6	32.9	–
Timing & duration of relationships				< 0.01				0.06
Started w/n last month	6.2	6.7	20.8	–	6.3	5.6	10.8	–
Still ongoing	65.2	38.4	40.0	0.2	51.7	15.4	22.2	0.04
Started w/n last year	24.4	33.9	14.6	–	27.0	29.3	37.4	–
Still ongoing	55.5	23.1	0.0	< 0.01	56.4	19.1	6.1	< 0.01
Started more than a year ago	68.3	59.4	64.6	–	65.1	65.1	51.8	–
Still ongoing	30.6	7.9	3.1	< 0.01	22.4	7.9	4.6	< 0.01
Partner's occupation				0.06				0.64
Fishing	37.9	47.2	47.5	–	–	–	–	
Business	19.3	13.9	7.5	–	26.0	26.2	25.3	
Services/government	25.6	17.2	17.5	–	6.6	4.5	8.5	
Stays at home/piece work	17.2	21.6	27.5	–	67.4	69.5	66.2	

Continued on next page

Table 2 *continued*

	Relations reported by women				Relations reported by men			
	Steady partner	Infrequent partner	One-night stand	p-value	Steady partner	Infrequent partner	One-night stand	p-value
Frequency of intercourse				0.93				0.02
Weekly or more frequent	52.7	51.9	50.0	–	67.0	57.5	55.4	–
Less than weekly	47.3	48.1	50.0	–	33.0	42.5	44.6	–
Condom use				0.57				0.65
Never	34.4	36.2	45.8	–	32.8	33.2	33.7	–
Sometimes	41.2	38.8	37.5	–	36.0	41.0	37.3	–
Always	24.3	25.0	16.7	–	31.2	25.7	28.9	–
Reason for condom use				0.09				0.38
Prevent STIs	62.8	60.0	65.4	–	57.2	66.0	63.0	–
Prevent pregnancy	32.1	35.8	19.2	–	37.7	29.4	29.6	–
Partner insisted	5.1	4.2	15.4	–	5.1	4.6	7.4	–
HIV risk perception				0.07				0.06
Worried a lot	12.8	23.3	22.9	–	13.7	14.4	22.9	–
Worried a little	30.1	22.7	20.8	–	23.4	28.4	31.3	–
Not worried at all	54.1	52.9	54.1	–	60.5	55.9	42.2	–

Reported numbers in the table are percentages (except for p-values). p-values are based on chi-square tests of associations.

prior to starting the relationship. Possibly related to these differences in meeting places, the occasions and specific contexts during which men and women meet different types of partners also varied quite largely. School and traditional events (Mganda) represent the two settings during which most partnerships are initiated. Men were more likely to meet their short-term partners during business trips or while on the steamer. Religious gatherings contributed to the formation of more than 10% of partnerships, but women in particular report that very few unstable partnerships were formed during such events. Further differences were found with respect to the type of relationship that existed between partners prior to the initiation of the sexual partnership. A large proportion (25%) of non-marital sexual partnerships are initiated between partners who did not know each other or had just met. Men were slightly more likely to engage in partnerships with women they did not know, and these partnerships often led to "one-night stands" or unstable relationships. Less than 10% of extra-marital relationships took place with a relative, but close to 23% of the one-night stands reported by women were with someone they were related to (either by blood or by marriage, i.e., in-laws).

Reported starting times of relationships differed significantly across types of relationships: over 60% of all extra-marital relationships reported during the sexual network survey had started more than a year prior to data collection, but this proportion was significantly higher among the steady relationships. In particular, more than 20% of the one-night-stands reported by women had started during the month immediately preceding the survey, but this was the case in only 6% of their more stable relationships. The same pattern held for males, as close to half of the one-night stands they report took place during the year prior to the survey. Not surprisingly, relationships classified as "steady partnerships" were reported to last longer than other types of extra-marital relationships by both men and women. A small proportion of one-night stands having started more than a year prior to the survey were still ongoing at the time of the survey, suggesting possible misclassification of partnerships.

With respect to relationship-specific risk factors for HIV infection, male respondents reported significantly more frequent sexual activity within all types of relationships than women. This pattern was especially apparent in steady relationships. Consistent condom use was reported in 25–30% of all extra-marital relationships, in general to prevent infection with STDs. Finally, levels of worry about HIV did not differ largely across relationship types, even though respondents engaged in one-night stands were more likely to be worried (a little or a lot) than respondents in more stable relationships.

6.3 Characteristics of the reconstructed network

6.3.1 Indegree distributions

The reports of partnerships described thus far define the underlying connectivity of the network. The insertion of individuals within this network is further defined by the distribution of indegrees, i.e., the number of times a respondent is nominated by someone else during the survey. The distributions of indegrees over 3 years and over one year prior to the survey are depicted in Figure 5. These distributions differ significantly from the outdegree distributions described in Section 6.2.1: over the full reporting period, the mode of the indegree distribution is 1 and a large number of respondents have indegree 0. Only slightly less than 20% of all respondents were

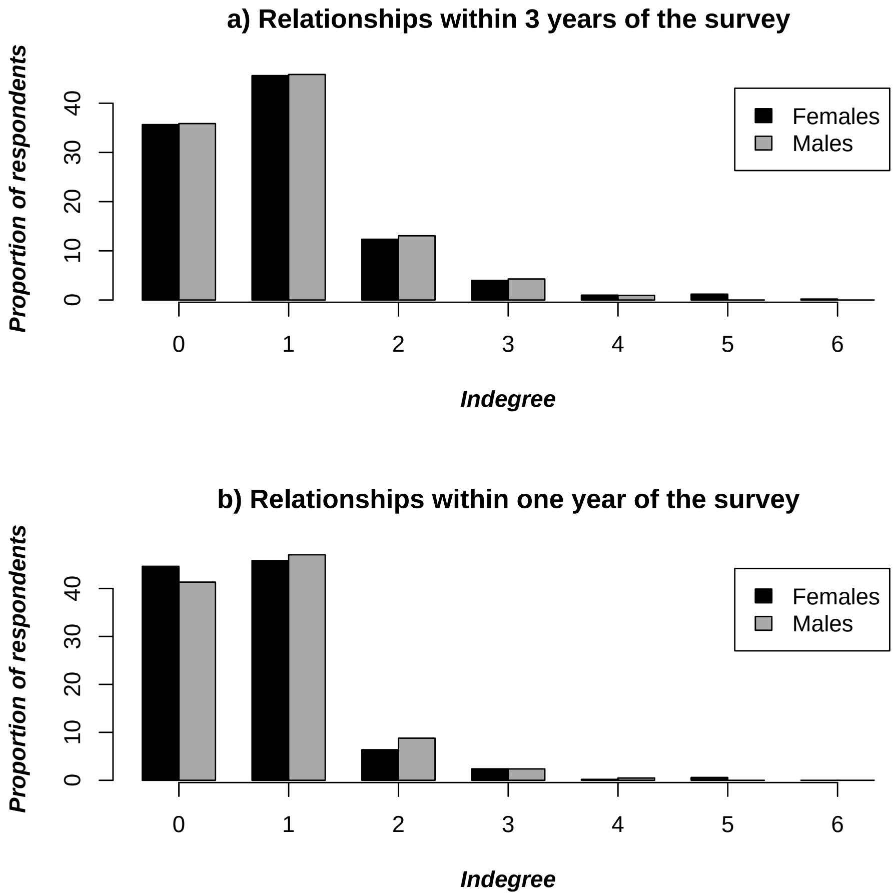

Figure 5: Indegree distributions among survey respondents, by gender; *Source:* Likoma Network Study)

nominated by more than 2 other respondents during the survey. Among respondents who themselves nominated more than two partners, this proportion is only raised to 24%. There were no significant differences in indegrees between men and women. Outdegree and indegree distributions differ because (*i*) some network members have had only partners who lived outside of the sample (and thus are not interviewed), (*ii*) women reported significantly fewer partnerships than men (see Nnko et al. 2004) and (*iii*) the limit of at most five partnerships to be reported may have resulted in some respondents not being nominated by their partners during the survey.

Figure 6 displays the indegrees of contacts, i.e., the number of times each network member *who was not a survey respondent* was nominated during the survey. The mode of this distribution is 1, and only a handful of network members who were not interviewed were reported more than once. In this distribution, nobody has degree 0 because being nominated by at least one respondent is the criteria for inclusion in the network.

6.3.2 Total degree distributions:

Whereas most studies of sexual behaviors in SSA are based on individual reports of partnerships (Cleland et al. 2004) (i.e., outdegrees), several analyses derived from the LNS (e.g., Helleringer

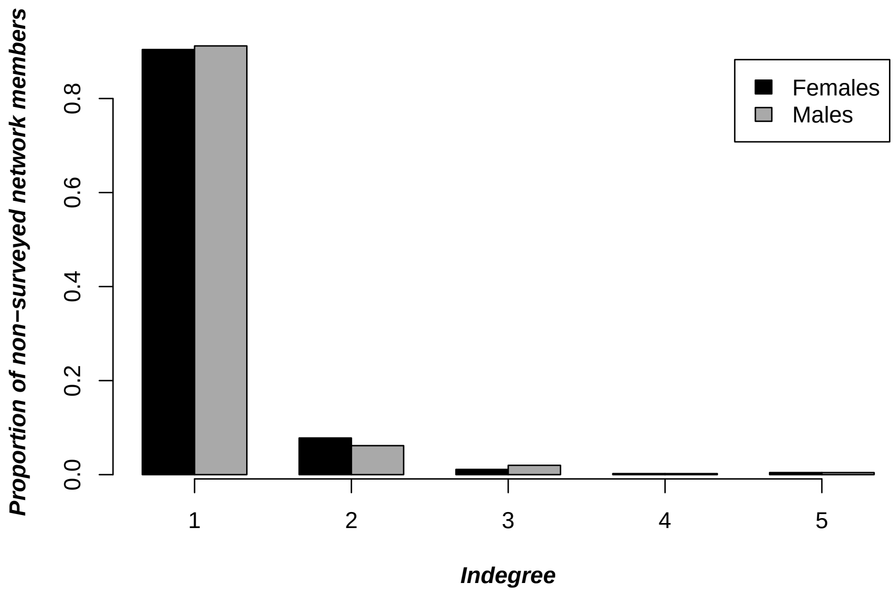

Figure 6: Indegree distribution among non-surveyed network members, by gender; *Source:* Likoma Network Study)

and Kohler 2007; Jones et al. 2007), on the other hand, build on *total degree distributions*. These distributions combine the partnerships reported by a respondent with the partnerships others report about a respondent. Figure 7 thus ignores the difference between reporting and being reported, and considers that a *relationship exists between two individuals as long as it is reported by at least one of the two parties* (symmetric sociomatrix). All sexual relationships that are reported by at least one partner are thus included, thereby possibly reducing the potential impact of underreporting of sexual relationships. The properties and structural features of the network defined by these distributions are analyzed in (Helleringer and Kohler 2007).

The average total degree of women over the years prior to the survey was 2.2, vs. 2.6 for men ($p < 0.01$). This was the case even though a slightly higher proportion of males had no partnerships over this time span (Figure 7a). During the year prior to the survey, the average total degree of women was 1.55 vs. 1.81 for men ($p < 0.01$). 6.7% of women and 12.6% of men had 5 or more partnerships. Only very few respondents ($N = 27$) were not sexually active over the entire recall period.

6.3.3 Comparison of total degree distributions and outdegree distributions:

Graphically, it appears that the gender differences in total degree largely parallel differences in outdegree described above (section 6.2.1). Tables 3 and 4 compare these distributions more systematically and evaluate how they differ. The correlation between total degree and outdegree is high as it reaches 0.85 among all respondents. There are significant differences between male and female respondents: the correlation coefficient for males is 0.92, whereas it is only 0.76 for females. This indicates that in this study, as has been noted elsewhere (e.g., Nnko et al. 2004), women tend to report fewer relationships than men do. And as a result of our study design, the total degree of

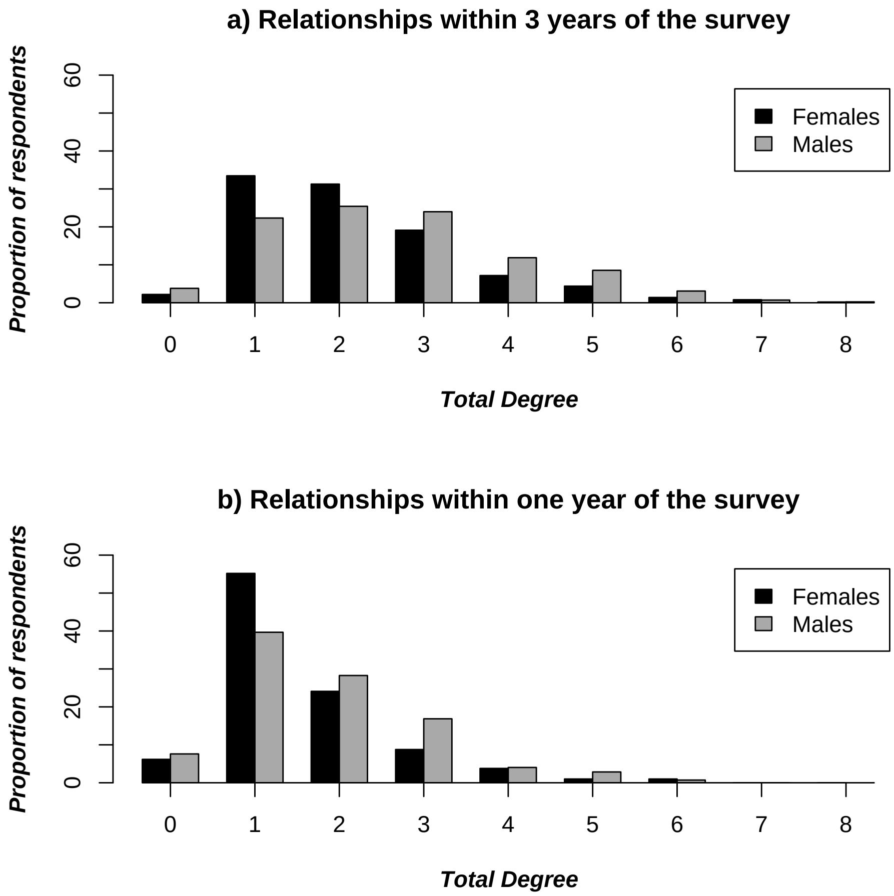

Figure 7: Total degree distributions among survey respondents, by gender; *Source:* Likoma Network Study)

Table 3: Average total degree and outdegree by gender and age among never-married respondents

	Female respondents			Male respondents		
	Total degree	Outdegree	*n*	Total degree	Outdegree	*n*
Age groups						
Less than 20	2.59(1.61)	1.62(1.08)	93	2.21(1.44)	1.92(1.32)	75
20–24	2.38(1.53)	1.89(1.16)	66	2.69(1.49)	2.35(1.41)	109
25–29	2.44(1.2)	2.27(1.22)	18	3.91(1.93)	3.29(1.62)	24
30–34	1.8 (1.3)	1.4 (1.67)	5	1.12 (0.99)	1.0(1.06)	8
35 and older	1.5 (0.57)	1.25 (0.5)	4	3.25 (1.7)[a]	3.0(1.63)[a]	4

A test of a linear trend in degree across age groups was significant at the .1 level. Standard deviations are in parentheses

Table 4: Average total degree and outdegree by gender and age among ever-married respondents

	Female respondents			Male respondents		
	Total degree	Outdegree	n	Total degree	Outdegree	n
All relations						
Less than 20	2.2(1.05)	2.05(0.94)	20	–	–	0
20–24	2.57(1.35)	2.0(0.95)	93	2.58(1.43)	2.25(1.26)	31
25–29	1.96(1.08)	1.64(0.81)	109	2.86(1.46)	2.57(1.38)	65
30–34	1.53 (0.81)	1.48(0.72)	60	2.55(1.31)	2.28(1.27)	49
35 and older	1.62 (0.79)[a]	1.5 (0.75)[a]	32	2.34(1.25)	2.12(1.22)	56
Non-marital relations						
Less than 20	1.65(1.22)	1.35(0.98)	20	–	–	0
20–24	1.79(1.47)	1.19(1.06)	93	2.0(1.63)	1.42(1.26)	31
25–29	1.12(1.13)	0.78(0.84)	109	2.0(1.57)	1.61(1.38)	65
30–34	0.66(0.81)	0.56(0.72)	60	1.51(1.17)	1.26(0.99)	49
35 and older	0.87(0.94)[a]	0.78(0.83)[a]	32	1.37(1.28)[a]	1.19(1.28)	56

[a], A test of a linear trend in degree across age groups was significant at the .01 level. Standard deviations are in parentheses

women is often higher than their outdegree. This is the case for 30.3% of female respondents vs. only 20% of male respondents ($p < 0.01$).

Furthermore, the patterns of differences between total degree and outdegree vary not only by gender, but also by age and marital status (Tables 3 and 4). In particular, the largest differences are observed among never-married women under age 25: such respondents spontaneously report only 70% of all relationships they appear to have engaged in. The gap between total and outdegree is much narrower for male and ever-married female respondents. In particular, males generally report between 85 and 90% of the relations they appear to have engaged in.

There are several trends in the reporting of sexual partnerships by age that emerge within this study. First of all, we were not able to detect any significant differences in reporting of partnerships among never-married women over age. This might be due to the fact that there are few women in their 30's who have never married, and the statistical test of trends might lack power. Among never-married males, on the other hand, both outdegree and total degree appear to increase significantly with age. Among ever-married respondents, total degree and outdegree appear to decline with age for women. This is the case both for all types of relations and for non-marital relations. For men, on the other hand, only the total number of non-marital relationships (total degree) appears to decline with age. This not the case for the number of non-marital relationships reported by men (outdegree). This finding indicates that *women* may be reluctant to report partnerships with older men, or alternatively that older men may be more likely to exaggerate the extent of their sexual networking.

6.3.4 Component size distribution:

Combined together in a large *sociomatrix* (Wasserman and Faust 1994), the reports of sexual relationships described above define a sexual network connecting residents of Likoma and residents of the mainland. We thus identified a network of order 1,803 connected by 1,614 unique relationships. The properties and structural features of this network are analyzed in detail in Helleringer

Table 5: Distribution of component sizes among respondents reporting at least one partner over the 3 years prior to the survey

Component size	Number of components	Proportion of all network members (in %)	Cumulative distribution (in %)
2	112	12.4	12.4
3	60	10.0	22.4
4	35	7.8	30.2
5	14	3.9	34.1
6	14	4.6	38.7
7	5	1.9	40.6
8	1	0.4	41.0
9	6	3.0	44.0
10	4	2.2	46.2
11	1	0.6	46.8
16	1	0.9	47.9
24	1	1.3	49.2
34	1	1.9	51.1
883	1	48.9	100.0

and Kohler (2007). In this section, we provide a few measures of network connectivity.

A fundamental measure of the connectivity of such a structure is the distribution of component sizes. A component is a set of individuals members of the network in which each pair is connected by at least one path (Wasserman and Faust 1994). These units are important because a person **A** may transmit infection to **B** only if there is a path of sexual relations (indirectly) connecting them. Over a 3 year time span, the sexual network identified by this study contains a total of 256 separate components (Table 5). The distribution of component sizes is highly skewed: more than 86% of the identified components are of size five or smaller, but include only 34% of all sexually active respondents. On the other hand, two thirds of network members are embedded in 35 components of size six or larger. Moreover, 883 network members constitute a single"giant" component of individuals connected through sexual partnerships having taken place during the three years prior to the survey. A substantial fraction of members of the giant component (40%) had at most 2 partners during the three years prior to the survey. The connectivity of the sexual network therefore occurs not because of high rates of partner change across all network members, but as a result of a generally moderate number of relationships with partners who have (or have had) other partners, who in turn may have had other partners and so on.

When we restrict the period of observation to the year prior to the survey (Table 6), a larger proportion of network members are included in dyads (28.2%). Nevertheless, large connected components still emerge within the network: close to half of all network members belong to components of size 4 and above, and close to one quarter of all network members are included in structures of size 25 and larger. In particular, 239 network members are connected into a single component and Helleringer and Kohler (2007) shows that this component is robust as more than a third of its members are linked together through multiple chains of sexual relations.

Table 6: Distribution of component sizes among respondents reporting at least one partner over the year prior to the survey

Component size	Number of components	Proportion of all network members (in %)	Cumulative distribution (in %)
2	204	28.2	28.2
3	75	15.6	43.8
4	39	10.8	54.6
5	11	3.8	58.4
6	8	3.3	61.7
7	5	2.4	64.1
8	3	1.6	65.7
9	3	1.9	67.6
10	1	0.7	68.3
11	1	0.8	69.1
12	3	2.5	71.6
13	1	0.9	72.5
18	1	1.2	73.7
25	1	1.7	75.4
46	1	3.2	78.6
70	1	4.8	83.4
239	1	16.6	100.0

7 Conclusion

The ***Likoma Network Study*** (LNS) constitutes—to our best knowledge—the first sociocentric study of sexual networks among a general population of SSA. Using these data, Helleringer and Kohler (2007) for instance document the existence of a large and robust sexual network. Half of all sexually active respondents were linked together in a *giant network component*, and more than one quarter were connected together through multiple independent chains of sexual relations. Such structural features of sexual networks have been associated with epidemic spread of STIs in high-risk groups (Moody et al. 2003; Newman 2002a; Potterat et al. 2002; Rothenberg et al. 1998), but prior to this study had never been documented among the general population. In addition, analyzes of HIV prevalence revealed important differences in the structural position of HIV-positive individuals, and counter-intuitively, HIV prevalence was higher in the sparser regions of the network than the densely connected components. Greater exposure to external infections through older partners or partners from the mainland was the main factor explaining this ecological variation in HIV prevalence across network locations.

In the present paper, we describe and evaluate the data collection procedures implemented during the LNS. We provide initial results relating to the socioeconomic context of the island, the prevalence of HIV in the study population, the quality of the sexual network data and the size, composition and (some) structural features of sexual networks. Our analyses show that the LNS has been able to identify the large majority of sexual relationships reported by survey respondents and provide evidence that the collection of accurate data on quasi-complete sexual networks in the general population is feasible in a sub-Saharan context. Nevertheless, close to 20% of nominated contacts residing in Likoma were not linked during the survey. Because errors in link tracing affect network density, the consequences for structural measurements may be important and such errors may generate *incomplete network bias* (Doherty et al. 2005). The network we reconstruct may be systematically sparser than the actual sexual networks of young adults on Likoma Island. Analyses based on the LNS are likely to underestimate the overall network connectivity.

Despite the potential biases affecting the collection of network data during the LNS, the data collected as part of the Likoma Network Study thus promise to substantially increase our knowledge about significant—but only poorly understood—hypotheses related to questions in several important areas relevant for understanding the spread of HIV: (*a*) *Selection of sexual partners:* What criteria of partner selection do individuals use to choose their partners? And how do these choices aggregate into larger sexual networks that provide potential routes for the transmission of HIV within a population? (*b*) *Network characteristics:* How do networks and the structural position of individuals within networks affect AIDS-related behaviors and HIV infection risks? (*c*) *Networks and life-course transitions:* How do important events during adolescence and young adulthood (e.g., marriage, migration, HIV-infection, death of partner) affect the structure of an individual's sexual network, and what are the implications of these changes for HIV risks? (*d*) *Disease dynamics and prevention:* Do the specific structures of sexual networks accelerate or slow the spread of HIV? Do these networks shield some groups from HIV while increasing exposure for others? Are there behaviors that appear particularly risky within the networks? Do these behaviors take place within the local populations and are there "bridges" between the islands' sexual networks and the larger cities of the mainland?

References

Adams, J. and J. Moody (2006). To tell the truth: Measuring concordance in multiply reported network data. *Social Networks*, forthcoming.

Allison, P. (1985). Survival analysis of backward recurrence times. *Journal of the American Statistical Association 80*, 315–322.

Anderson, R. M. and R. M. May (1991). *Infectious diseases of humans: Dynamics and control*. Oxford: Oxford University Press.

Bailey, N. (1975). *The mathematical theory of infectious disease*. New York: Hafner Press.

Bearman, P. S., J. Moody, and K. Stovel (2004). Chains of affection: The structure of adolescent romantic and sexual networks. *American Journal of Sociology 110*(1), 44–91.

Bell, D., I. Montoya, and J. Atkinson (2000). Partner concordance in reports of joint risk behaviors. *Journal of the Acquired Immune Deficiency Syndromes 25*, 173–181.

Blanchard, J. F. (2002). Populations, pathogens, and epidemic phases: Closing the gap between theory and practice in the prevention of sexually transmitted diseases. *Sexually Transmitted Infections 78*(Supplement 1), 183–188.

Blasnick, M. (2001). Gsoundex: Stata module to implement soundex algorithm. Unpublished manuscript, Boston College Department of Economic. Available online at http://ideas.repec.org/c/boc/bocode/s420901.html.

Bloom, D. E. (1998). Technology, experimentation, and the quality of survey data. *Science 280*(5365), 847–848.

Brewer, D. D. and C. M. Webster (1999). Forgetting of friends and its effects on measuring friendship networks. *Social Networks 21*(4), 361–373.

Burt, R. S. (1987). Social contagion and innovation: Cohesion versus structural equivalence. *American Journal of Sociology 92*(6), 1287–1335.

Caldwell, J. C., P. Caldwell, and P. Quiggin (1989). The social context of AIDS in sub-Saharan Africa. *Population and Development Review 15*(2), 185–234.

Cleland, J., J. Boerma, M. Carael, and S. Weir (2004). Monitoring sexual behaviour in general populations: A synthesis of lessons of the past decade. *Sexually Transmitted Infections 80*(Supplement 2), 1–7.

Cliff, A. D., P. Haggett, and M. R. Smallman-Raynor (2000). *Island Epidemics*. Oxford: Oxford University Press.

Coffee, M., G. Garnett, M. Mlilo, H. Voeten, S. Chandiwana, and S. Gregson (2005). Patterns of movement and risk of HIV infection in rural Zimbabwe. *Journal of Infectious Diseases 191*(Supplement 1), S159–167.

Coffee, M., M. N. Lurie, and G. P. Garnett (2007). Modelling the impact of migration on the HIV epidemic in South Africa. *AIDS 21*(3), 343–350.

Coleman, J. D., E. Katz, and H. Menzel (1966). *Medical Innovation: A Diffusion Story*. Indianapolis: Bobbs-Merrill.

Costenbader, E. and T. W. Valente (2003). The stability of centrality measures when networks are sampled. *Social Networks 25*(3), 283–307.

Doherty, I. A., N. S. Padian, C. Marlow, and S. O. Aral (2005). Determinants and consequences of sexual networks as they affect the spread of sexually transmitted infections. *Journal of Infectious Diseases 191*, S42–S54.

Feld, S. L. (1981). The focused organization of social ties. *American Journal of Sociology 86*(5), 1015–1035.

Ghani, A., C. Ison, H. Ward, G. Garnett, G. Bell, G. Kinghorn, J. Weber, and S. Day (1996). Sexual partner networks in the transmission of sexually transmitted diseases. an analysis of gonorrhea cases in Sheffield, UK. *Sexually Transmitted Disease 23*, 498–503.

Ghani, A. C. and G. P. Garnett (2000). Risks of acquiring and transmitting sexually transmitted diseases in sexual partner networks. *Sexually Transmitted Diseases 27*, 579–587.

Glynn, J. R., J. Ponnighaus, A. C. Crampin, F. Sibande, L. Sichali, P. Nkhosa, P. Broadbent, and P. E. Fine (2001). The development of the HIV epidemic in Karonga district, Malawi. *AIDS 15*(15), 2025–2029.

Handcock, M. S. and J. Jones (2004). Likelihood-based inference for stochastic models of sexual network evolution. *Theoretical Population Biology 65*, 413–422.

Heesterbeek, J. A. P. (2002). A brief history of r_0 and a recipe for its calculation. *Acta Biotheoretica 50*(3), 189–204.

Helleringer, S. and H.-P. Kohler (2005). Social networks, risk perceptions and changing attitudes towards HIV/AIDS: New evidence from a longitudinal study using fixed-effect estimation. *Population Studies 59*(3), 265–282.

Helleringer, S. and H.-P. Kohler (2006). The structure of sexual networks and the spread of HIV in sub-Saharan Africa: Evidence from Likoma Island, Malawi. Paper presented at the annual meeting of the Population Association of America, Los Angeles, CA, March 30–April 2, 2006.

Helleringer, S. and H.-P. Kohler (2007). Sexual network structure and the spread of HIV in Africa: Evidence from Likoma Island, Malawi. *AIDS*, accepted (pending minor revisions).

Hethcote, H. W., J. W. van Ark, and J. Longini, I. M. (1991). A simulation model of AIDS in San Francisco: I. Model formulation and parameters. *Mathematical Biosciences 106*(2), 203–222.

Heuveline, P. (2003). HIV and population dynamics: A general model and maximum-likelihood standards for East Africa. *Demography 40*(2), 217–245.

Hudson, C. P. (1996). AIDS in rural Africa: A paradigm for HIV-1 prevention. *International Journal of STD & AIDS 7*(4), 236–243.

Jones, H. H., S. Helleringer, and H.-P. Kohler (2007). Statistical models for sexual networks on-Likoma Island, Malawi: Implications for sexual behavior and HIV control. Paper to be presented at the annual meeting of the Population Association of America, New York, NY, March 29–31, 2007 Online available at http://paa2006.princeton.edu.

Jones, J. and M. S. Handcock (2003a). Sexual contacts and epidemic thresholds. *Nature 425*, 605–606.

Jones, J. H. and M. S. Handcock (2003b). An assessment of preferential attachment as a mechanism for human sexual network formation. *Proceedings of the Royal Society of London Series B-Biological Sciences 270*, 1123–1128.

Klovdahl, A., J. Potterat, D. Woodhouse, J. Muth, S. Muth, and W. Darrow (1994). Social networks and infectious disease: The Colorado Springs Study. *Social Science and Medicine 38*, 79–88.

Klovdahl, A. S. (1989). *Urban Social Network: Some Methodological Problems and Possibilities*. Ablex Publishing.

Koehly, L. Goodreau, S. and M. Morris (2004). Exponential family models for census and sampled network data. *Sociological Methodology 34*, 241–270.

Kossinets, G. (2006). Effects of missing data in social networks. *Social Networks 28*(3), 247–268.

Koumans, E., T. Farley, J. Gibson, C. Langley, M. Ross, M. McFarlane, J. Braxton, and M. St Louis (2001). Characteristics of persons with syphilis in areas of persisting syphilis in the United States: Sustained transmission associated with concurrent partnerships. *Sexually Transmitted Diseases 28*, 497–503.

Krackhardt, D. (1987). Cognitive social structures. *Social Networks 9*, 109–134.

Kretzschmar, M. and M. Morris (1996). Measures of concurrency in networks and the spread of infectious disease. *Mathematical Biosciences 133*(2), 165–195.

Lagarde, E., B. Auvert, M. Caraël, M. Laourou, B. Ferry, E. Akam, T. Sukwa, L. Morison, B. Maury, J. Chege, I. N'Doye, and A. Buvé (2001). Concurrent sexual partnerships and HIV prevalence in five urban communities of sub-Saharan Africa. *AIDS 15*, 877–884.

Lagarde, E., M. Schim van der Loeff, C. Enel, B. Holmgren, R. Dray-Spira, G. Pison, J. Piau, V. Delaunay, S. M'Boup, I. Ndoye, M. Coeuret-Pellicer, H. Whittle, and P. Aaby (2003). Mobility and the spread of human immunodeficiency virus into rural areas of West Africa. *International Journal of Epidemiology 32*, 744–752.

Laumann, E., J. Gagnon, T. Michael, and S. Michaels (1994). *The social organization of sexuality:*

Sexual practices in the United States. Chicago: University of Chicago Press.

Laumann, E., P. Marsden, and D. Prensky (1983). The boundary specification problem in network analysis. In R. S. Burt and M. J. Minor (Eds.), *Applied Network Analysis: A Methodological Introduction*, pp. 18–34. London: Sage Publications.

Laumann, E. and Y. Youm (1999). Racial/ethnic group differences in the prevalence of sexually transmitted diseases in the United States: A network explanation. *Sexually Transmitted Diseases 26*(5), 250–261.

Liljeros, F., C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Aberg (2001). The web of human sexual contacts. *Nature 411*(6840), 907–908.

Lurie, M., B. Williams, K. Zuma, D. Mkaya-Mwamburi, G. Garnett, J. Sweat, M.D. Gittelsohn, and S. Karim (2003). Who infects whom? HIV-1 concordance and discordance among migrant and non-migrant couples in South Africa. *AIDS 17*, 2245–2252.

Magruder, J. (2006). Marital shopping and epidemic AIDS. Unpublished working paper, Department of Economics, Yale University.

Mensch, B. S., P. C. Hewett, and A. Erulkar (2003). The reporting of sensitive behavior among adolescents: A methodological experiment in Kenya. *Demography 40*(2), 247–268.

Moody, J. (2002). The importance of relationship timing for diffusion. *Social Forces 81*, 25–46.

Moody, J., M. Morris, J. Adams, and M. Handcock (2003). Epidemic potential in human sexual networks. Unpublished working paper, Ohio State University.

Moody, J. and D. R. White (2003). Structural cohesion and embeddedness: A hierarchical concept of social groups. *American Sociological Review 68*, 103–127.

Morris, M. (1993). Epidemiology and social networks: Modeling structured diffusion. *Sociological Methods & Research 22*(1), 99–126.

Morris, M. (1997). Sexual networks and HIV. *AIDS 11*, S209–S216.

Morris, M. (2004). Overview of network survey designs. In M. Morris (Ed.), *Network Epidemiology*. Oxford: Oxford University Press.

Morris, M. and M. Kretzschmar (2000). A microsimulation study of the effect of concurrent partnerships on the spread of HIV in Uganda. *Mathematical Population Studies 8*(2), 109–133.

Newman, M. (2002a). Spread of epidemic disease on networks. *Physical Review E 66*(1), 016128–11.

Newman, M. E. J. (2002b). Spread of epidemic disease on networks. *Physical Review E 66*(1), art. no.–016128.

Nnko, S., J. T. Boerma, M. Urassa, G. Mwaluko, and B. Zaba (2004). Secretive females or swaggering males? an assessment of the quality of sexual partnership reporting in rural Tanzania. *Social Science and Medicine 59*(2), 299–310.

Padgett, J. F. and C. K. Ansell (1993). Robust action and the rise of the Medici, 1400–1434. *American Journal of Sociology 98*(6), 1259–1319.

Potterat, J. J., Z.-R. H., S. Q. Muth, R. B. Rothenberg, D. L. Green, J. E. Taylor, M. S. Bonney, and H. A. White (1999). Chlamydia transmission: Concurrency, reproduction number, and the epidemic trajectory. *American Journal of Epidemiology 150*(12), 1331–1339.

Potterat, J. J., L. Phillips-Plummer, S. Q. Muth, R. B. Rothenberg, D. E. Woodhouse, T. S. Maldonado-Long, H. P. Zimmerman, and J. B. Muth (2002). Risk network structure in the early epidemic phase of HIV transmission in Colorado Springs. *Sexually Transmitted Infections 78*(Supplement 1), 159–163.

Potterat, J. J., D. E. Woodhouse, S. Q. Muth, R. Rothenberg, W. W. Darrow, A. S. Klovdahl, and J. B. Muth (2004). Network dynamism: History and lessons of the Colorado Springs study. In M. Morris (Ed.), *Network Epidemiology*. Oxford: Oxford University Press.

Reniers, G. (2003). Divorce and remarriage in rural Malawi. *Demographic Research Special Collection 1*(6), 175–206. Available online at http://www.demographic-research.org.

Rothenberg, R. B., C. Sterk, K. E. Toomey, J. J. Potterat, D. Johnson, M. Schrader, and S. Hatch (1998). Using social network and ethnographic tools to evaluate syphilis transmission. *Sexually Transmitted Diseases 25*, 154–160.

Salganik, M. and D. Heckathorn (2004). Sampling and estimation in hidden populations using respondent-driven sampling. *Sociological Methodology 34*, 193–239.

Sampson, S. F. (1969). A novitiate during a period of change: An experimental and case study of relationships. Unpublished Ph.D. dissertation, Cornell University.

Van Den Bulte, C. and G. Lilien (2005). Medical innovation revisited: Social contagion versus marketing effort. *American Journal of Sociology 106*(5), 1409–1435.

Wasserman, S. and K. Faust (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

Watkins, S., J. R. Behrman, H.-P. Kohler, and E. M. Zulu (2003). Introduction to "Research on demographic aspects of HIV/AIDS in rural Africa". *Demographic Research Special Collection 1*(1), 1–30. Available online at http://www.demographic-research.org.

Watkins, S. C. (2004). Navigating the AIDS epidemic in rural Malawi. *Population and Development Review 30*(4), 673–705.

Watts, D. and S. Strogatz (1998). Collective dynamics of 'small-world' networks. *Nature 393*, 440–443.

Watts, D. J. (1999). *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton: Princeton Studies in Complexity.

Wawer, M., R. Gray, N. Sewankambo, D. Serwadda, X. Li, O. Laeyendecker, N. Kiwanuka, G. Kigozi, M. Kiddugavu, T. Lutalo, et al. (2005). Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *Journal of infectious diseases 191*, 1403–1409.

White, K. and S. C. Watkins (2000). Accuracy, stability and reciprocity in informal conversational networks in rural Kenya. *Social Networks 22*(4), 337–355.

Whittaker, R. (1999). *Island Biogeography*. Oxford: Oxford University Press.

World Health Organization (2002). *HIV Assays: Operational Characteristics. Report 12: Simple/Rapid Tests, Whole Blood Specimens*. Geneva, Switzerland: World Health Organization. Online availabe at http://www.who.int/diagnostics_laboratory/publications/hiv_assays_rep_12.pdf.

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